Université de Montréal

Analytical Strategies for the Comprehensive Profiling of Histone Post Translational Modifications by Mass Spectrometry and Implications for Functional Analyses

Par
Paul Drogaris

Département de Chimie
Faculté des Arts et des Sciences

Thèse présentée à la Faculté des Études Supérieures
en vue de l'obtention du grade de
Philosophiæ Doctor (Ph.D.)
en chimie

Novembre 2010

Université de Montréal
Faculté des Études Supérieures

Cette thèse intitulée :

Analytical Strategies for the Comprehensive Profiling of Histone Post Translational Modifications by Mass Spectrometry and Implications for Functional Analyses

Présentée par
Paul Drogaris

A été évaluée par un jury composé des personnes suivantes:

Pierre Thibault ..... Directeur de recherche
Alain Verreault..... Codirecteur de recherche
Joëlle Pelletier.....Président-rapporteur
Jean-François Masson..... Membre du jury
Daniel Figeys ..... Examinateur externe
Pascal Chartrand .....Représentant du Doyen de la Faculté

Resumé

Le long bio-polymère d'ADN est condensé à l'intérieur du noyau des cellules eukaryotes à l'aide de petites protéines appelées histones. En plus de leurs fonctions condensatrices, ces histones sont également la cible de nombreuses modifications post-traductionnelles (MPT), particulièrement au niveau de leur section N-terminale. Ces modifications réversibles font partie d'un code d'histones épi-génétique transmissible qui orchestre et module dynamiquement certains événements impliquant la chromatine, tels l'activation et la désactivation de gènes ainsi que la duplication et la réparation d'ADN. Ces modifications sont impliquées subséquemment dans la signalisation et la progression de cancers, tels que la leucémie. En conséquence, l'élucidation des modifications d'histones est importante pour comprendre leurs fonctions biologiques.

Une méthodologie analytique a été mise au point en laboratoire pour isoler, détecter, et quantifier les MPT d'histones en utilisant une approche rapide à deux volets à l'aide d'outils bioinformatiques spécialisés. La méthodologie développée en laboratoire a été validée en utilisant des histones de souche sauvage ainsi que deux types d'histones mutants déficients en enzymes acétyltransferase. Des trois sources d'histones utilisées, la seule MPT qui a démontré un changement significatif est l'acétylation de l'histone H3 à lysine 56 (H3K56ac). L'expression et la stœchiométrie de cette MPT, issue de cellules de souche sauvage et de cellules mutantes, ont été déterminées avec précision et comparées.

Les fonctions de balayage polyvalentes d'un instrument à trappe ionique quadrupôle linéaire hybride ont été utilisées pour améliorer la détection de protéines intactes. Le mode de balayage « enhanced multiply charged » (EMC) a été modifié pour contenir et détecter les ions de protéines intactes situées dans la trappe ionique linéaire. Ce mode de balayage nommé « targeted EMC » (tEMC) a permis de quadrupler le niveau de sensibilité (signal/interférence), et quintupler la résolution du mode de balayage conventionnel. De plus, la capacité de séparation des charges du tEMC a réduit de façon significative les effets de « space charge » dans la trappe ionique linéaire. La résolution

supérieure du mode tEMC a permis de différencier plusieurs isoformes modifiées, particulièrement pour l'histone H3. L'analyse des peptides d'histones trypsiques à l'aide du mode de balayage « MRM » a permis le séquençage et la quantification de MPT avec un haut degré de précision. La seule MPT qui était sous-exprimée entre l'histone de souche sauvage et le mutant DOT1L fut la méthylation de l'histone H3 lysine 79 (H3K79me1).

Les effets de deux inhibiteurs d'enzymes HDAC (HDACi) sur l'expression de MPT d'histone ont été évalués en utilisant la méthodologie analytique mentionnée. Les histones extraites de cellules normales et cancéreuses ont été exposées à du Vorinostat (SAHA) ou du Entinostat (MS-275) pour une période de 24 à 72 heures. Deux histones furent principalement affectées, soit H3 et H4. Étonnamment, les mêmes effets n'ont pas été détectés lorsque les cellules normales ont été traitées avec le HDACi pour une période de 48 à 72 heures. Une méthode absolue de quantification avec une courbe d'étalonnage a été développée pour le peptide H3K56ac. Contrairement à certaines publications, nos résultats démontrent que cette MPT est présente dans les cellules mammifères avec une stœchiométrie très basse (< 0,1%) et n'est pas surexprimée de façon significative après le traitement au HDACi.

Mots Clés: Histones, modifications post-traductionnelles, spectrométrie de masse, nanoLC-MS/MS, trappe ionique linéaire, protéomique quantitative.

Summary

In eukaryotic cells, the lengthy DNA biopolymer is condensed into the cell nucleus with the aid of small packaging proteins called histones. In addition to their packing functions, histones are also targets of numerous post translational modifications (PTMs), especially on their N-terminus. These reversible modifications are believed to be constituents of a heritable epigenetic "histone code" that dynamically orchestrate and modulate chromatin-based events such as gene activation and silencing, DNA replication and repair, and are also involved in the downstream signaling and progression of cancers, such as leukemia. Thus, the elucidation of histone PTMs is important in understanding their biological function.

An analytical workflow was designed and set-up in the laboratory to isolate, detect, and quantitate histone PTM, using a two-pronged, unbiased, and rapid approach with specialized bioinformatic tools. The workflow was validated using histones from wild-type, and 2 mutants deficient in acetyltransferase activity. Between the three histone sources, the only PTM that demonstrated any change was acetylation at histone H3 lysine 56 (H3K56ac). The down-regulation and stoichiometry of this PTM was accurately assessed between wild-type and mutant cells.

The versatile scan functions of a hybrid quadrupole-linear ion trap instrument were exploited to enhance the detection of intact histone proteins. The enhanced multiply charged (EMC) scan was modified in order to contain and detect intact protein ions within the linear ion trap. This targeted EMC (or tEMC) resulted in not only a 4-fold increase in signal-to-noise, but also a 5-fold increase in resolution. Furthermore, the charge separation capability of the tEMC dramatically reduced space charge effects within the linear ion trap. The superior resolution of the tEMC mode allowed for the discimination of many modified histone isoforms, especially for histone H3. Using the bottom-up strategy with multiple reaction monitoring (MRM), histone peptides were quantified and sequenced with a high degree of precision. The only PTM that was down-

regulated between wild-type and DOT1L mutant histones was methylation at histone H3 lysine 79 (H3K79me1).

The effects of two clinically relevant small molecule HDAC inhibitors (HDACi) on histone PTMs patterns were assessed using the analytical workflow developed. Histones derived from both normal and cancer cells were exposed to either Vorinostat (SAHA) or Entinostat (MS-275) over a 24- to 72 hour period. The two core histones primarily affected were H3 and H4. Surprisingly, the same effects were not observed when normal cells were treated with three doses of SAHA at 24-hour intervals over a 72-hour period. An absolute quantitation method using a calibration curve was developed for H3K56ac. In opposition to other published litterature, our findings demonstrate that this PTM is present in very low stoichiometry (< 0.1%) in mammalian cells, and exhibits no significant up-regulation in different cell lines treated with several types of HDACi.

Keywords: Histones, post-translational modifications, mass spectrometry, nanoLC-MS/MS, linear ion trap, quantitative proteomics.

Table of contents

List of tables

List of figures

List of abreviations

1-D	One dimensional
2-D	Two dimensional
A	Adenine
ac	Acetylation
ADP	Adenosine diphosphate
ACN	Acetonitrile
AGC	Automatic gain control
amu	Atomic mass unit
API	Atmospheric pressure ionization
APL	Acute promyelocytic leukemia
BCA	Bicinchoninic acid
bp	base pair(s)
BSA	Bovine serum albumin
C	Cytosine
C. albicans	*Candida albicans*
CAD	collision activated dissociation
cal. conc.	calculated concentration
CDK	Cyclin-dependent kinase
CE	capillary electrophoresis
CES	Collision energy spread
CID	collision induced dissociation
CML	chronic myelogenous leukemia
CRM	Charged residue model
CV	Coefficient of variation
Cyt C	Cytochrome C
Da	Dalton
DC	Direct current
DFT	Dynamic fill time

DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferase
E. coli	*Escherichia coli*
ECD	Electron capture dissociation
EMC	Enhanced multiply charged scan
EMS	Enhanced MS scan
EPI	Enhanced product ion scan
ESI	Electrospray ionization
ETD	Electron transfer dissociation
FAIMS	High field asymmetric waveform ion mobility spectrometry
fg	femtogram
FT-ICR	Fourier transform ion cyclotron resonance
FT-MS	Fourier transform mass spectrometer
fwhm	Full width at half maximum
G	Guanine
GC	Gas chromatography
h	hour
HAT	Histone acetyltransferase enzyme
HDAC	Histone deacetylase enzyme
HDACi	Histone deacetylase enzyme inhibitor
HPLC	High performance liquid chromatography
IDA	Information dependant acquisition
IEM	ion evaporation model
IT	ion trap
kbp	kilo base pairs
LC-MS	Liquid chromatography interfaced with electrospray ionization mass spectrometry
LC-MS/MS	Liquid chromatography interfaced with electrospray ionization and tandem mass spectrometry
LC	Liquid chromatography
LC-UV	Liquid chromatography with ultraviolet detection

LE	Leucine enkephalin
LIT	Linear ion trap
M	Concentration expressed in molar units
MALDI	Matrix-assisted laser desorption ionization
MCS	Multiply charged separation
MDS	myelodysplastic syndrome
me	methylation
me1	monomethylation
me2	dimethylation
me3	trimethylation
μL	volume in microliters
MIDAS	MRM initiated data acquisition and sequencing
MRM	Multiple reaction monitoring
MS	Mass spectrometry
ms	millisecond
MSD	Mass selective detector
MS/MS	Mass spectrometry/Mass spectrometry (tandem mass spectrometry)
MW	Molecular weight
m/z	mass-to-charge ratio
NAD^+	Nicotinamide adenine dinucleotide
NCP	Nucleosome core particle
ng	nanogram
PA	propionic anhydride
pg	picogram
PITC	phenyl isothiocyanate
PTH	phenylthiohydantoin
PTMs	Post translational modifications
Q0	Non-resolving quadrupole
Q1	First resolving quadrupole
Q3	Third resolving quadrupole

QqLIT	Quadrupole-linear ion trap mass spectrometer
QqQ	Triple quadrupole mass spectrometer
Q-TOF	Quadrupole-time of flight mass spectrometer
Q-TRAP	Quadrupole-linear ion trap mass spectrometer
S. cerevisiae	*Saccharomyces cerevisiae*
s	second
SAHA	Suberoylanilide hydroxamic acid
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SIDT	single ion in droplet theory
SILAC	stable isotope labeling by amino acids in cell culture
stoichio.	Stoichiometry
RF	Radio frequency
rH3	Recombinant yeast histone H3
RIC	Reconstructed ion chromatogram
ROS	Reactive oxygen species
RP	Reversed phase
RSD	Relative standard deviation
SCX	Strong cation exchange
T	Thymine
tEMC	Targeted enhanced multiply charged scan mode
TFA	Trifluoroacetic acid
Th	Thompson
TIC	Total ion chromatogram
TOF	Time-of-flight mass analyzer
UV	Ultraviolet detection
WT	Wild-type
v/v	Concentration expressed as volume per volume

Acknowledgements

A successful Ph.D. project is never the work of solely one person. There is a great number of other players involved, both actively and on the sidelines. I will take this opportunity to thank the many people who made this dream a reality.

My journey into the realm of mass spectrometry began in 1995, after registering for an undergraduate level course given by Dr. Robin Rye at Concordia University. It was during this semester when I realized how powerful mass spectrometry truly was. A few months later, I would land my first job as a bench chemist at Phoenix life sciences, and eagerly waited for my opportunity to get my hands on the various Sciex API III instruments present in the facility. I would have to wait patiently for another two years, and another company, before my chance would materialize. My first teacher and mentor was Roger Demers at Maxxam Analytics, who took me under his wings and taught me not only the basics on a Sciex API III instrument, but also method development for the bioanalysis of drugs and metabolites from biological matrices. For this opportunity and invaluable experience, I will be forever grateful.

My next break came in 2001 during the ASMS conference in Chicago. This is where I had the opportunity to meet Dr. Pierre Thibault. After keeping in contact for several years, we had met again one night in early 2005 at a local conference in Montreal. After learning that he had become a professor at the Université de Montréal and started his own lab, I jumped at the opportunity to join his group. A few months later, I was accepted in the chemistry program, and started my new academic journey. I would like to thank Pierre for giving me the chance to realize my research potential. Again, for the opportunity to pursue graduate studies in mass spectrometry, I will be forever grateful.

With Pierre opening the door to pursue graduate studies, my co-director Alain Verreault provided a constant supply of biological samples, which became a gold mine of information waiting to be discovered. I am thankful for not only his expertise, patience, and willingness to share his knowledge, but also for motivating me to probe and go

further. A special thanks to all the members in his lab, including Hugo Wurtele, Jason Tanny, Benoit Guillemette, Valérie Villeneuve, Neda Delgoshaie, and Eun-Hye Lee for not only being a great group of scientists, but wonderful people to work with. I can only wish that the people I get to work with in my future endeavors would be the same as in the Verreault lab.

I would like to thank all the past and present members of the Thibault lab. They have not only been great colleagues and friends, but also provided much needed support in times of discouragement. I will miss them all greatly.

I would like to thank my parents George and Georgia, my brother Costa, and my wife Véronic for their relentless support, love, and encouragement. I would have not been able to accomplish this journey without them.

A special thanks goes out to all my past and present teachers and training partners in various martial art disciplines, including coaches Tony Blauer and Phil Hughes, sensei Tetsuji Ishizuka, Joel Boucher, Marcel Lefebvre, Patrick Dubé, Marc-André Côté, senseis Claude Gagnon and Josette Lanteigne, and Guru Randall Goodwin, for helping to bring out and forge the warrior spirit of my Spartan heritage.

Finally, I would like to dedicate this Ph.D. thesis to the late Eleni Alevizatos (1917-2000), beloved mother and grandmother, who succumbed to the metastatic effects of liver cancer. If any of this research could one day contribute to the some type of cancer therapy, then it would have been worth the difficult Ph.D. journey I embarked upon.

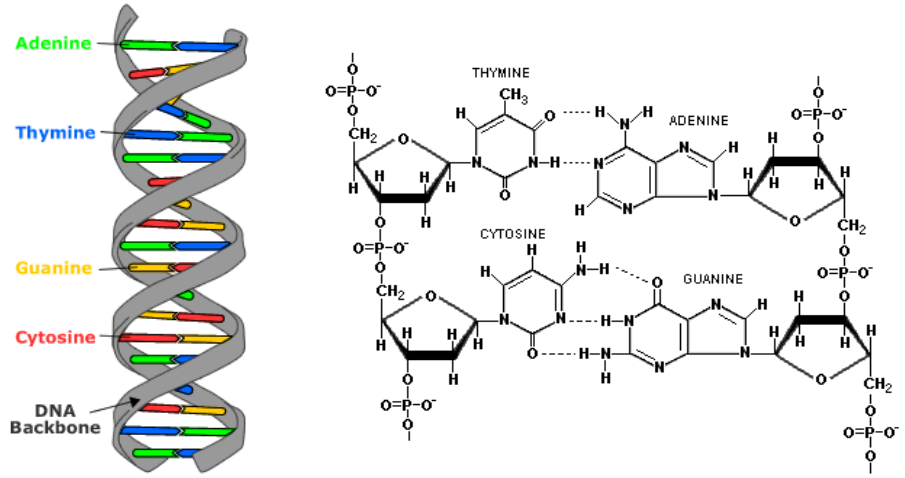
"*Shikin Haramitsu Daikomyo*"

In every experience, there is the potential for self-discovery, clarity, and improvement…

1. Introduction

1.1 DNA structure and properties

The organizational and biological blueprint of an organism's genetic information is contained in its DNA sequence. At the most basic structural level, DNA consists of a constant deoxyribose sugar moiety, and four variable and alternating nucleotide bases [1]: adenine (A), thymine (T), cytosine (C), and guanine (G), (see Figure 1.1). Antiparallel polynucleotide chains can form hydrogen bonds with each other, joining A-T and C-G base pairs (bp). X-ray diffraction studies using DNA fibers by Watson and Crick [2] enabled the elucidation of the DNA structure, and the formulation of the classical "B-form" model. According to their pioneering work, DNA is a repeating right-handed double helix, requiring 10 base residues to complete a full turn, with a precise 36 degree rotation angle between subsequent bp. The majority of DNA molecules adopt this B-form *in vivo*, while being stabilized by "high water activity" and along a spine of hydration [3]. Biophysical studies have shed a great deal of insight about the hydration dynamics [4] and conformational stabilization [5] of B-DNA in solution using short synthetic oligomers. DNA can also adopt two other well-characterized conformations, notably the A- and Z-forms [6]; these depend on the chemical environment and DNA sequence. The A-form arises when DNA becomes dehydrated, resulting in a more compact double helix, with 11 bp per turn. The Z-confirmation can form in solutions of high ionic strength, with alternating CG-CG and CA-TG sequences [7], resulting in a left-handed helix. The classical model was amended primarily by the work of Dickerson and co-workers [8]. Using x-ray diffraction analysis of DNA dodecamer crystals, they concluded that the bp bond rotation had a slightly different range (28 to 42 degrees). Furthermore, the bp were not co-planar, but rather could rotate in opposite directions; this "propeller twisting" [9] enabled a high degree of bp stacking. One of the most important topological properties of DNA is its ability to be supercoiled [10, 11] and twisted periodically into multiple superhelices without physical damage, enabling the macromolecule to be highly compacted.



(a) (b)

Figure 1.1: The helical structure of DNA, with its four nucleotide base pairs (a). Chemical structure of DNA, with negatively charged phosphate groups (b). From references [12] and [13].

1.2 Histones: the packaging proteins of DNA

All essential proteins necessary for the biological functioning of an organism are encoded in its DNA sequence, forming a very long biopolymer that twists onto itself to generate a thermodynamically stable helical structure. Almost 3 billion bp make up the human genome [14] (6 billion bp per diploid cell), and the entire DNA sequence from a cell would easily span a linear distance of roughly two meters when annealed and unwrapped. This very lengthy sequence of nucleotides cannot occupy the limited size of the cell nucleus (about 10^{-5} m in diameter) without some form of packaging. As a result, nature has evolved an elegant solution to accommodate and package DNA in a highly ordered manner. The key proteins involved in this packaging process in eukaryotic organisms are called *histones* [15]. These proteins are relatively small (10 to 20 kDa), and highly basic; their entire amino acid sequence contains over 20% lysine and arginine residues. Together, DNA and histones form chromatin fibers (see Figure 1.2), which are the physiologically relevant substrate for important cellular activities such as DNA transcription, replication, and repair. At the first level of organization, 147 bp of DNA are wrapped in roughly 1 ¾ turn around an octamer of histone proteins, forming a

nucleosome, or nucleosome core particle (NCP). At first, DNA was believed to reside in "uniform supercoils", however, the existence of NCPs was demonstrated by several groups including Clark [16], Hewish [17], and Noll [18] in chromatin digestion studies using micrococcal nuclease and DNase I. Unlike DNA strands, chromatin digestion was incomplete and did not form small oligonucleotides; this result led to the conclusion that the protein component was not evenly distributed along the DNA strands, suggesting the presence of a "subunit". Although these experiments were simplistic in design, the same results and observations by different research groups were convincing enough and led to further investigation by emerging instrumental techniques such as electron microscopy and high resolution x-ray crystallography. Chromatin extracted from eukaryotic cells was found to be present in a linear array of repeating spheres ("beads on a string") by Olins [19]; this further corroborated the coiling of DNA to form the NCP. Finally, Luger [20] and Davey [21] provided concrete evidence at the atomic level using crystallized human α-satellite DNA and recombinant histone proteins. Their studies showed that 147 bp of DNA underwent excessive curvature to coil around a histone protein core forming a superhelix, whose folding was dependant on electrostatic interactions between the histones and DNA.

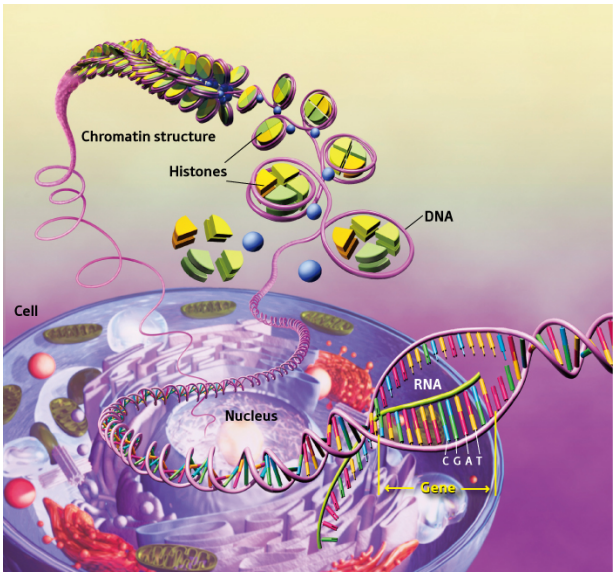


Figure 1.2: Histone proteins aid in the packaging of DNA into the cell nucleus.
From reference [22].

A closer inspection of the NCP reveals an octameric histone complex. The protein complex is made up of four different types of core histones, namely, H2A, H2B, H3, and H4 (see Figure 1.3). The association of the four core histones is also ordered and specific: two pairs of H2A and H2B molecules assemble into two H2A-H2B dimers, forming a tetramer, while H3 and H4 form an $(H3\text{-}H4)_2$ tetramer [23, 24]. All four core histones share common structural motifs, including a hydrophobic core, referred to as the histone-fold or globular domain, comprising three α-helices, and hydrophilic extremities, i.e., conformationally flexible N-termini or N-terminal "tails", as well as short C-terminal ends. The histone-fold domain is primarily involved in maintaining structural integrity of the NCP, while the more basic histone tails, rich in lysine and arginine residues, protrude outside of the NCP structure. The core of the histone-fold domains from two different histones associate interfaces together in a head-to-tail fashion that results in inter-molecular, six α-helix bundles, forming a "handshake motif" by the two forms of heterodimers: H2A-H2B and H3-H4 [25]. Both the α-helices and the "paired-element motifs" of the helix/loop-and-strand/helix (HSH) structures contribute to form grooves in the histone-fold octamer [26], which serve as "docking pads" for DNA. Moreover, specific arginine residues [27] in the globular domain provide DNA anchoring points for negatively charged phosphate groups by electrostatic and hydrogen bonding.

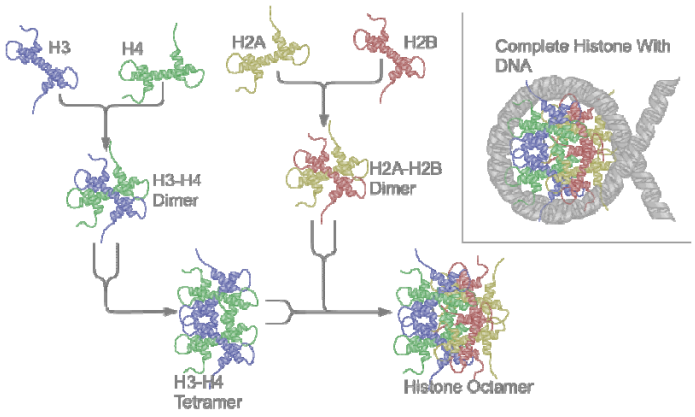


Figure 1.3: Formation of the octameric complex by the four core histones. From reference [28]

As Arents points out [26], an important and well-conserved lysine-arginine pair for DNA interaction exists in the HSH2 loop between helix II and strand B. The core histones bind to the DNA backbone at precise locations [20], namely the L1L2 loops, and four $\alpha 1 \alpha 1$ helices in the globular domain.

1.2.1 Higher order chromatin architecture

Repeating units of nucleosomes are compacted further into a helical coil of nucleosomes. This is achieved by additional structural components and chromatin associated proteins to form the "30 nm fiber" [29]. This constitutes the secondary level of DNA-histone organization. Chromatin fibers are stabilized and maintained in a higher order structure by another type of histone, referred to as linker histone H1 [30]. In mammals, there are six different types of histone H1 variants (H1.0 to H1.5), each possessing different biological properties, and named according to their chromatographic behavior [31]. Histone H1 is structurally and functionally different from the four core histones comprising the nucleosome. Compared to the histones present in the NCP, histone H1 is longer and heavier (about 23 kDa), with three distinct structural domains [32]. These are characterized by an 80 residue long globular domain, attached to a short N-terminal and a long C-terminal extension, both of which are highly basic [33].

The globular domain of H1 also contains two surfaces lined with positively charged amino acids that provide two additional DNA binding regions in NCPs. A short sequence of DNA resides outside of the nucleosome, and varies in length depending on species and tissue. This linker DNA provides contact points for the N- and C-terminal extensions of histone H1. Linker histone H1 helps to stabilize and maintain the higher ordered chromatin fiber by neutralizing the negative electrostatic charges on the DNA molecule, and also by interacting with NCPs. Biophysical studies have shown that although linker histone H1 helps maintain chromatin stability, it is not totally essential. For example, Carruthers and colleagues [34] have shown that nucleosomal arrays from histone octamers depleted of histone H1 and twelve 208 bp tandem repeats of *Lytechinus* 5S rDNA were able to fold in salt solutions of varying concentrations. Although unstable

in solution, the arrays showed a high degree of folding. The extensive basic N- and C-terminal extensions of linker histone H1 provide the driving force to stabilize and maintain the higher order chromatin structure. Tertiary structure formation involves further packaging of chromatin into thick and condensed fibers. Two models (shown in Figure 1.4) have been proposed to describe this higher order architecture: the solenoid and "zig-zag" models [35]. Due to its highly compacted nature, the spatial arrangement and localization of individual nucleosomes within the 30 nm fiber is not resolvable. Although it is very difficult to determine with absolute certainty which model holds true *in vivo*, experimental evidence points to the "zig-zag" model as being the most plausible one [36-38]. In this model, nucleosomes become "interacting partners", imparting more stabilization.

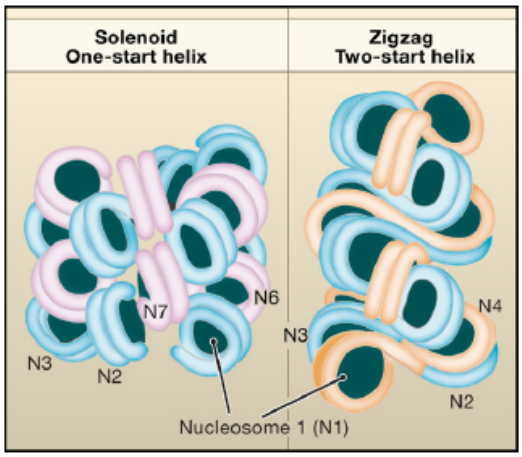


Figure 1.4: Two models of chromatin structures, i.e., the solenoid (left) and "zig-zag" models. From reference [39].

Further folding of the chromatin fiber into a metaphase chromosome necessitates the formation of radial and helical loops of 50 to 100 kbp. Construction of this highly organized chromosomal configuration is aided by the support of a nuclear "matrix" or "scaffold" [6]. One of the major components of this scaffolding is topoisomerase II (topo II). A study conducted by Takeyasu demonstrated that topo II was an integral component necessary for the condensation of mitotic chromosomes and had the ability to join two DNA strands together. Moreover, this process required histone H1-bound chromatin for

the formation of higher order structures in an ATP-independent manner [40]. Other chromatin-associated proteins are involved in this higher order packing [41]; these include MeCP2, which binds to methylated DNA and stabilizes loops, MENT, a serpin family protein that aids in oligomerization, and Polycomb, a large protein complex involved in gene silencing and can also attach to nucleosomes. In addition, inner nuclear membrane proteins such as the lamins [42] provide binding sites for chromatin within the cell nucleus. All of these structural components work together in concert to ultimately achieve the final packaged state of the chromosome.

1.2.2 Histone modifications

The preceding sections have described histone proteins solely as packaging agents of chromatin, resulting in nucleosomes with a rather rigid globular domain, and flexible N-terminal tails. Histones have a number of amino acids, including basic lysine and arginine residues, spread over the length of both the globular domain and the flexible N-terminal tails that can be covalently modified. These post-translational modifications (PTMs) of histones have been known for many years; for example, acetylation was first reported in 1963 by Phillips in histones isolated from calf thymus extracts [43]. To date, there are over 60 PTM sites on histones that have been identified by techniques such as Edman sequencing and mass spectrometry [44]; this list continues to grow as research intensifies. The most common and well-characterized histone PTMs include acetylation, methylation, and phosphorylation. Some of these PTMs are shown in Figure 1.5. A great number of lysine residues can be modified by acetylation and methylation (in mono-, di-, and trimethyl forms). Arginine residues can be methylated (mono- and dimethyl forms), while serine and threonine can be phosphorylated. Other less commonly observed PTMs include conjugation of histone lysine residues to small proteins such as Ubiquitin and Sumo, as well as ADP-ribosylation of histone arginine, aspartic and glutamic acid side chains. The majority of these PTMs occur on the N- and C-terminal tails of histones, yet a small number of important modifications have also been detected in the globular domain, e.g. histone H3 lysine 79 methylation. The multitude of these PTMs leads to an ever increasing complexity and combinatorial “explosion” of possibilities.

The fact that histone proteins show a high degree of sequence conservation between different species, and their modifications often have similar functions in various biological systems suggests that these PTMs are important and possess evolutionarily conserved roles in cell biology. The modification of histones is dictated by both internal and external cellular stimuli, and ultimately leads to changes in the structure and function of chromatin, mediating important biological processes such as gene transcription, DNA replication, and damage repair [45]. These alterations in chromatin can be achieved by three known mechanisms: (a) disrupting DNA-histone electrostatic interactions ("*cis*" effects), and (b) recruitment of enzymes and chromatin remodeling complexes ("*trans*" effects), and (c) replacement of canonical histones by histone variants, e.g., H3.1 by replacement H3.3, and H2A by H2A.Z.

Figure 1.5: A partial list of known histone PTMs that occur on the N-terminal tails and globular domains of the core histones. From reference [46].

One of the best known and most studied *cis*-effects of chromatin remodeling is the acetylation of lysine residues on core histones. The incorporation of an acetyl group removes the positive charge on the lysine side chain, thus negating the electrostatic contact with the DNA backbone. This PTM favours the unraveling of a specific region of chromatin, making DNA more accessible, and allowing gene transcription to proceed. In the yeast *S. cerevisiae*, lysine 14 on the H3 N-terminal tail can be acetylated by the

enzyme Gcn5, which activates gene transcription [47]. As demonstrated by Cairns [48], H3K14ac is an active mark for transcription recognized by the bromo domain of the RSC complex, an ATP-dependent nucleosome remodelling enzyme that promotes gene activation by disrupting histone-DNA interactions. Thus, histone modifications can also serve as a "docking site" for larger enzyme complexes and remodeling machinery. These *trans*-effect chromatin marks direct the engagement of other binding proteins, which can signal downstream cellular processes. Another example of such a *trans* (but opposing) effect is the methylation of lysine 9 on histone H3 [49]. This modification can be generated by the Suv39h methyltransferase, providing an anchoring point for the chromo domain heterochromatin binding protein 1 (HP1). Once bound to chromatin that contains K9-methylated H3 molecules, HP1 can mediate the formation of heterochromatic regions, forcing DNA to be inaccessible, and promoting gene silencing. The third mechanism involves the substitution of core histones by sequence variants in nucleosomes. These histone variants are utilized to establish chromatin states with specialized functions. In mammalian cells, one such histone variant is H3.3, which resembles the canonical H3 histone, but differs only in four amino acid residues.

In the fruit fly *D. melanogaster* and human cells, this H3 variant can be deposited into chromatin in a replication-independent manner, and directed into sites of transcriptionally active genes [50]. These small differences in amino acids are crucial for the deposition of H3.3 into specific regions, and its function after incorporation into chromatin [51]. One of these differences is at position 31; in H3.1, this residue is an alanine, while in H3.3, the alanine is replaced by a serine or threonine. Unlike the alanine residue in the canonical histone H3, serine and threonine can be phosphorylated, triggering important downstream cellular events. Allis and colleagues [52] have demonstrated that Ser-31 phosphorylation occurs only in metaphase, and is localized in specific areas adjoining centromeres. Suggested functions for this residue difference and PTM include protection of euchromatin, and substitution of H3.3 by the canonical H3.

1.3 Epigenetics: Going beyond the DNA sequence

A recent study performed by Fraga [53] examined phenotypic differences in monozygous (MZ) twins. Although genetically identical at birth, these MZ twins developed significant differences over time, e.g., risk of disease progression. This begs the following question: what mechanism(s) is responsible for these alterations over time? One major finding in the study was that older twins showed drastic changes in their DNA methylation and histone acetylation profiles. Hence, the answer could lie beyond the genetic information encoded in the DNA sequence. The regulation of phenotypes by non-DNA related mechanisms brings us to the concept of *epigenetics*. The idea that phenotypic changes can be caused by non-genotypic changes is not entirely new. Originally coined by Conrad Waddington in 1942 (see Figure 1.6), epigenetics was the term used to "describe events that could not be explained by genetic principles" [54]. A more recent definition by Allis would describe epigenetics as the collection of modifications to chromatin producing different states of gene expression from the same genome [46]. In multicellular organisms, most cells have the same genotype, yet become differentiated and organized to fulfill many distinct roles.

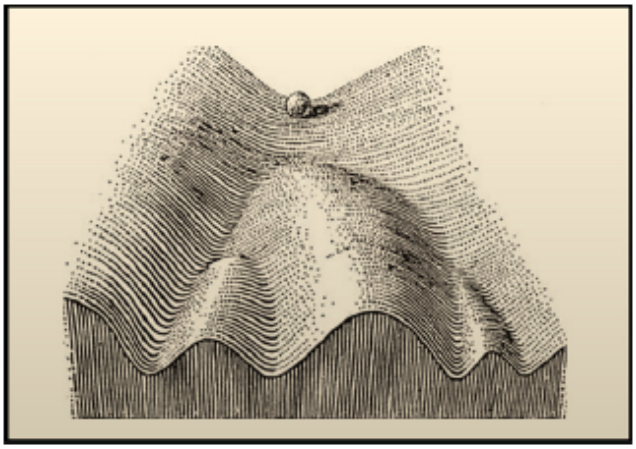


Figure 1.6: The classical epigenetic "landscape" by Waddington. In this figure, a cell is represented by a ball at the top of the hill. A number of slopes are present, representing a limited number of pathways that the cell can take, leading to different cellular fates. From reference [54]

There are two main epigenetic mechanisms in mammalian cells: (1) DNA methylation and (2) histone modifications [55]. The vast majority of DNA methylation occurs on cytosine residues in 5'-CpG-3' dinucleotides. In regions where CpG dinucleotides occur frequently (known as CpG "islands"), methylation of cytosine is a mark of gene silencing [56].

The second mechanism is the modification of specific amino acid residues of histone proteins. As described the previous section, the diverse types of histone PTM can modulate when and how genes are expressed, using chromatin as a template for synchronizing cellular events. These modifications usually do not occur as single events on chromatin, but rather in combination with each other at precise residues and defined times within the cell cycle, causing spatio-temporal changes in chromatin structure that lead to gene transcription or repression. An example of histone PTMs working in concert to modulate chromatin remodelling is shown in Figure 1.7.

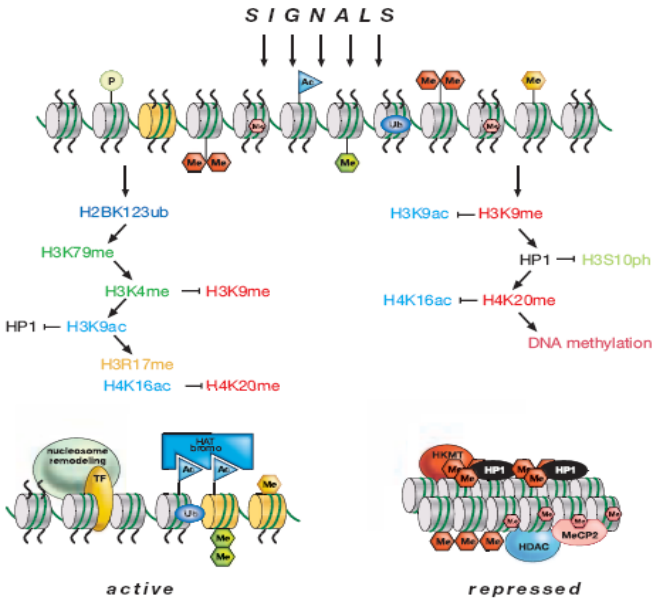


Figure 1.7: Histone PTMs are a series of synchronized events working synergistically or antagonistically to generate active (left) or repressed (right) chromatin states. From reference [46].

The combinatorial nature and phenotypic changes of these modifications lead to the formulation of the "histone code" hypothesis by Allis, Strahl and Jenuwein [57, 58]. According to them, these PTMs form an "epigenetic language" that is reversible and heritable, without necessitating changes in the DNA sequence. The multiple sites of modification present on the small segments of the N-terminal tails makes them "hotspots" for enzymatic activity. Addition of a modification on a specific residue of an N-terminal tail can affect the binding and the catalytic rate of another enzyme, either inductively or antagonistically. Some research groups have provided evidence that supports the existence of this histone code by modification "switches". For example, Lo and colleagues have shown that H3 serine-10 phosphorylation up-regulates the acetylation of H3 lys-14 by several HATs, including Gcn5, PCAF, and p300 [59]. This modification sequence is linked to transcriptional activation.

Epigenetics can be viewed as a double edged sword. On one hand, epigenetic mechanisms are involved in regulating important DNA metabolic processes such as gene transcription and DNA repair, which help maintain a high degree of genomic integrity and stability [60]. This regulation is maintained by a multitude of "feedback mechanisms", and interplay between different enzymes, such as DNA methyltransferases and chromatin modifying complexes. However, epigenetic mechanisms are often disrupted in diseases such as cancer. Essentially, cancer is genetic disease in which gene expression has become abnormal and cellular homeostasis is unbalanced. Some common attributes of different types of cancers include unregulated cell proliferation, point mutations, loss or silencing of genes, tumor growth, invasion, and metastasis [61]. However, the molecular complexity of this disease is further expanded by dysfunctional epigenetic regulation, such as defects in DNA cytosine methylation and histone modifications.

Both genetic and epigenetic factors can work in tandem to interfere with normal cellular activity, producing two types of aberrations in gene expression that contribute to cancer: (1) induction of oncogenes, and (2) silencing of tumor suppressor genes [46]. One example is the oncogenic activation of the *ras* gene [62]. This occurs by a genetic point

mutations that cause single amino acid substitutions in the Ras protein, and high spontaneous mutation rates, especially in pancreatic cancers [63]. The presence of the positive K-*ras* mutation in patients with lung adenocarcinomas is associated with poor prognosis and low survival rates. The second mechanism by which cancer may be induced is by epigenetic repression of genes that are important to regulate and maintain healthy cellular functions. One key epigenetic mechanism involved in tumorigenesis is CpG island promoter hypermethylation [64]. In bladder cancers, the CDK inhibitor and tumor suppressor $p16^{INK4a}$ becomes silenced by extended hypermethylation of the DNA region encoding the gene [65]. This allows malignant cells to bypass senescence and proliferate without constraint. Histone modifications can also contribute to abnormal gene silencing during carcinogenesis. Baylin and colleagues demonstrated that, in cancer cells, the silenced promoter of tumor suppressor gene hMLH1 resides in an area whereby histones are de-acetylated and subsequently methylated at H3K9 [66]. In mixed lineage leukemia (MLL), a chromosome translocation aberrantly fuses the *AF10* and *MLL* genes (see Figure 1.8). The AF10-MLL oncogenic fusion protein aberrantly recruits the histone methyltransferase hDOT1L, resulting in abrogation of the H3K4me mark characteristic of active chromatin, and induces H3K79 hypermethylation, leading to the overexpression of the *HoxA9* gene and leukemogenesis [67].

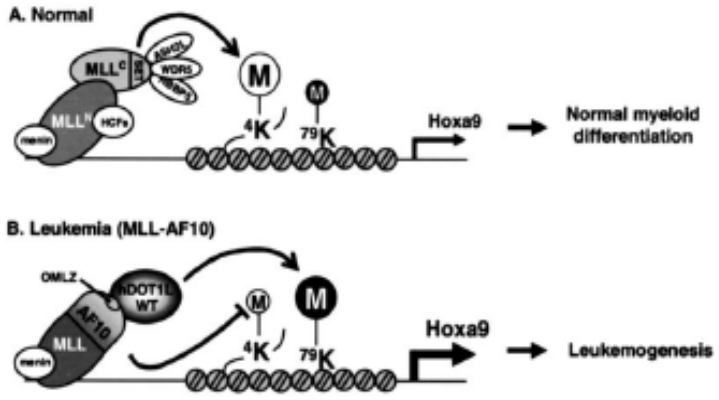


Figure 1.8: Normal myeloid differentiation mediated by histone modifications characteristic of healthy cells (top). Fusion proteins (MLL-AF10) cause aberrant H3K79 methylation and transcription of the *Hoxa9* gene, leading to leukemia (bottom). From reference [67].

1.3.1 Epigenetic drug therapy: a promising avenue for cancer treatment

Biochemical processes often require strict regulation by opposite sets of enzymes. In diseases such as cancer, these processes become dysregulated, leading to the silencing of genes that are involved in tumor suppression. The realization that abnormal epigenetic gene inactivation can be reversed led to an interest in studying the enzymes that propagate epigenetic marks as potential therapeutic drug targets. As discussed in the previous section, an important epigenetic process that is often dysregulated in cancer cells is DNA methylation. When CpG islands become hypermethylated by DNA methyltransferases (DNMTs), this can lead to inappropriate inactivation of tumor suppressor genes. Studies of DNA methylation during the 1960s led to the discovery of methylation inhibitors, such as 5-azacytidine [68]. This small molecule has anti-proliferative properties and inhibits tumor growth. Further optimization resulted in the development of structural analogues such as 5-aza-2'-deoxycytidine [69] that have less toxicity and side effects. This research also prompted studies of other enzyme inhibitors as potential epigenetic therapies. These inhibitors exert their pharmacological effects [70] by administration as prodrugs, metabolic activation via phosphorylation, and incorporation into DNA as a modified nucleosides that subsequently interfere with DNA methylation. Continued research into DNMT inhibitors has led to their use as epigenetic chemotherapies that reverse the aberrant gene silencing that occurs in cancer cells. For example, in myelodysplastic syndrome (MDS), the genes encoding the CDK inhibitors p15 and p16 become hypermethylated. Upon treatment with the DNMT inhibitor 5-aza-2-deoxycytidine (decitabine), the hypermethylated state of these genes was reversed in responsive patients, restoring gene function and controlling the proliferation of malignant cells [71]. Clinical studies led to the approval by the US food and drug administration of 5-azacytidine for the treatment of all subtypes of MDS [72], while several other structural analogues are in different phases of drug development.

Histone modifications such as acetylation are also a highly regulated. This regulation is mediated by the reversible actions of both histone acetyltransferases (HATs) and deacetylases (HDACs). In cancers such as acute promyelocytic leukemia (APL) [73],

HDAC enzymes contribute to epigenetic gene silencing through aberrant recruitment by fusion proteins to specific loci, forming a complex with other enzymes, and remodeling chromatin into a condensed state. This deacetylation effect is amplified in cancer, since HDACs can undergo overexpression in certain malignancies such as colon cancer [74]. Wilson and colleagues demonstrated that HDAC4 can promote proliferation of cancerous colon cells by down regulating the *p21* gene promoter. Hence, HDACs have also become viable therapeutic drug targets. In humans, there are 18 different types of HDAC enzymes, which are divided into 4 classes [75]: I, IIa, IIb, III, and IV. Every type with the exception of class III enzymes can be inhibited by "classical" small molecule HDAC inhibitors (HDACi). Class I and II HDACi possess a bidentate pharmacophore that can coordinate to and chelate an essential zinc ion within the active site of class I and II HDACs [76]. The third class of HDACs is structurally related to Sir2 yeast proteins, uses NAD^+ as a co-substrate, rather than simply hydrolyzing the acetyl group from lysine residues in the target protein. The class III HDACs are not inhibited by "classical" HDACi that block class I and class II enzymes. Once HDACi are bound to an enzyme's catalytic core, deacetylase activity is blocked, histones become hyperacetylated, and target genes activated through chromatin remodeling (see Figure 1.9). Several research groups, including Marks and colleagues, have shown that HDACi can block proliferation or kill tumor cells by triggering G1/S cell cycle arrest, differentiation, increased levels of reactive oxygen species (ROS), and apoptosis [77]. Furthermore, HDACi can trigger hyperacetylation of non-histone proteins such as transcription factors (E2F1, p53, STAT1). Transcription factor and histone hyperacetylation can act synergistically to shift the equilibrium to reactivate tumor suppressor genes. To date, only two HDACi have been approved by the US food and drug administration for treatment of cutaneous T-cell lymphoma. The most successful and extensively studied HDACi is the Merck chemical Vorinostat (*Zolinza*), a hydroxamic acid based molecule that inhibits a wide spectrum of class I and II HDAC enzymes. However, the ability of pan-HDACi to alter the acetylation patterns of other proteins has raised concern about toxicity and potential side effects; this has led to research aimed at developing more selective HDACi that target a smaller group of enzymes.

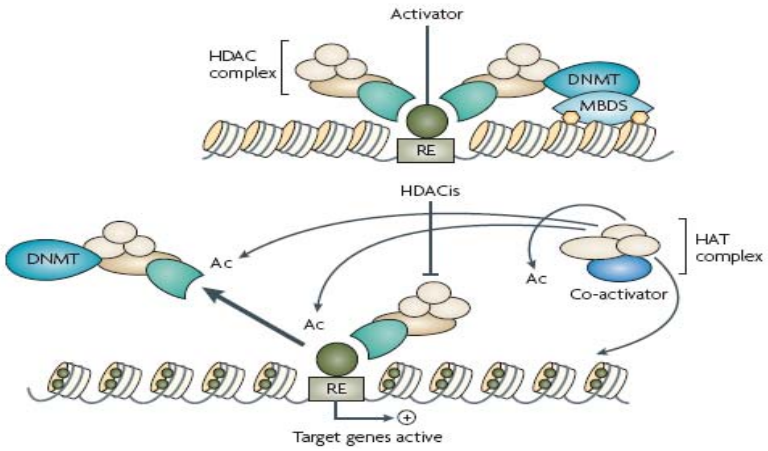


Figure 1.9: Epigenetic drug therapy. In acute promyelocytic leukemia (APL), chromosomal translocations result in fusion proteins that recruit DNMTs and HDAC enzymes to silence tumor suppressor genes. Treatment with HDAC inhibitors can disrupt these complexes, hyperacetylate histones, and reactivate the required genes which trigger cell cycle arrest, differentiation or apoptosis of cancer cells. From reference [78].

1.4 Protein analysis: the early years

The word protein is derived from the Greek word "*proteios*", meaning "of first importance or rank" [1]. The term was coined by Jons Jacob Berzelius [79] in 1838 to highlight the biological significance of these molecules. Proteins are present in virtually all cells and are involved in almost all cellular processes, from mitosis to catalysis of xenobiotics. The structural building blocks of proteins are amino acids, which occur naturally in 20 different forms [80]. These structural components are responsible for synthesizing the vast amounts of different proteins in all organisms, from single cell amoeba to complex mammals. The amino acid sequence determines the protein's structure, chemical properties, folding, and ultimately, its function.

A great interest in the structural biology of proteins started in 1950's, with the pioneering work of Corey and Pauling. Their research focused on bond lengths, atomic arrangements, and peptide configurations, which eventually led to the elucidation of an

important structural motif: the α-helix [81]. This caused a great interest in protein and peptide research in the scientific community, and fuelled the start of the research explosion into protein structure and function. During the same period, there was another important contribution made in analyzing protein structure: the Edman degradation [82]. This reaction enables the N-terminal sequencing of proteins by sequentially removing one amino acid at a time from the polypeptide backbone. The reaction is a three-step process [80], whereby the polypeptide is reacted with phenyl isothiocyanate (PITC) to form a stable phenylthiohydantoin (PTH) derivative. The PTH-amino acid can be easily analyzed using chromatographic methods, such as HPLC. Repeated cycles of Edman degradation will liberate the amino acids, thus allowing the determination of the protein sequence. Edman degradation has serious limitations [83], including a practical limit of 40 to 60 residues before side reactions make sequencing unreliable, the necessity of a substantial amount of purified protein, and long reaction times. Furthermore, if the N-terminus is modified, the amino acid will not be derivatized, and the method would fail completely [84]. Despite these limitations, it was the principle method that researchers would use for the next thirty years.

1.5 Mass spectrometry: the driving force of proteomics

A sample containing multiple components from an *in vivo* source can provide a serious challenge from an analytical perspective. "Classical" methods of analysis consist of gas or liquid chromatography with various types of spectroscopic detection (i.e. ultra-violet, fluorescence). Although rugged and reliable, these techniques can sometimes result in ambiguity or uncertainty in the identity of the peaks detected. Mass spectrometry (MS) provides another analytical methodology to detect various analytes, including peptides and proteins. Instead of using a spectroscopic property to detect and measure the component of interest, MS discriminates analytes based on their mass-to-charge (*m/z*) ratio [85]. Analytes are converted to gaseous ions, and their *m/z* ratio is recorded onto a mass spectrum. Since the mass is used to identify chemical species, this results in much better specificity in analytical determinations compared to spectroscopic methods. Hence, using MS to investigate a proteome (the protein complement to the genome) resulted in a new paradigm of protein analysis, i.e., proteomics. During the

1990's, this new approach in protein analysis would take over as the method of choice for protein and peptide analysis due to its speed, sensitivity, and specificity.

1.5.1 The emergence of electrospray ionization

A key breakthrough that would change the landscape of protein analysis by mass spectrometry was the optimization of the electrospray ionization (ESI) process for large biomolecules by John Fenn [86] and colleagues at Yale University. Using data and observations reported by both Rayleigh and Dole [87, 88], Fenn's group capitalized on this preliminary research and successfully applied it to larger charged species, i.e., proteins [89]. This ionization method uses high voltage (kilo volt range) applied to a narrow diameter capillary with the aid of a counter current flow of drying gas (e.g. nitrogen). Species that are ionizable in solution are placed in a solvent. This solution is passed through the capillary in order to produce charged species in the gaseous phase. During this process (depicted in Figure 1.10), liquid emerging from the tip of the capillary generates the Taylor cone [90] and experiences opposing electrostatic and surface tension forces [91], forming a mist of droplets. Solvent molecules gradually evaporate until the Rayleigh instability limit [92] is reached. Beyond this point, repulsive coulombic forces between species of the same charge overcome the surface tension of the sprayed droplet, releasing ions into the gaseous state. Additives present in the liquid effluent (i.e. acid) undergo charge transfer reactions with analytes in the gaseous phase, producing protonated adducts and pseudomolecular ions [93]. An important feature of ESI is the generation of intact multiply charged ions without fragmentation or decomposition. There are currently two models used to rationalize the generation of ions from charged droplets in the gaseous phase by ESI. The first is the charged residue model (CRM) [94], or single ion in droplet (SIDT) [95] theory, formulated by Dole. In this description, it is assumed that total solvent evaporation results in one analyte molecule being contained per charged droplet at the end of a series of coulombic explosions; this analyte molecule then accepts a charge from the last droplet formed. In the ion evaporation model (IEM) developed by Iribarne and Thomson [96], ions are desolvated

and ejected by a sufficient amount of electrostatic force from the surface of "highly electrified cloud droplets" containing many charges.

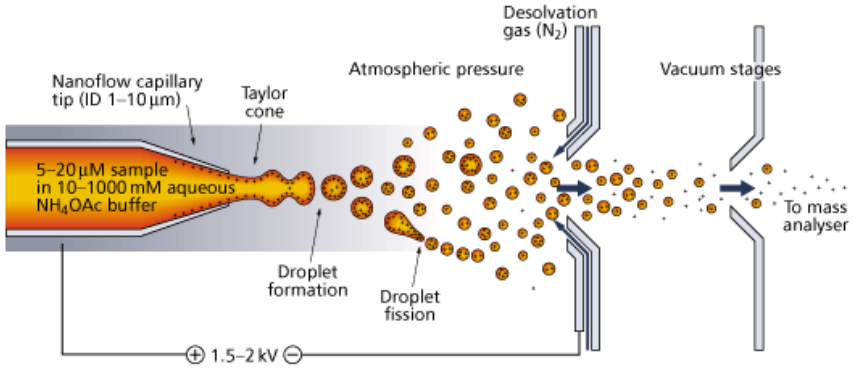


Figure 1.10: A close-up view of the formation of gas phase ions during electrospray ionization. From reference [97].

1.5.2 Analytical LC-MS instrumentation for proteomics

There are five main components that make up an MS-based analytical system for proteomics: (a) a sample introduction device, (b) an ion source, (c) a mass analyzer, (d) a detection module, and (e) data processing station. Each component will be discussed in more detail below.

1.5.2.1 Sample introduction for mass spectrometry

The two most common sample introduction devices for MS include liquid and gas chromatographs. Capillary electrophoresis (CE) has also been used to introduce samples into the mass spectrometer [98]. For protein and peptide analysis, capillary HPLC and CE are the two most widely used separation methods [99]. Prior to mass analysis, complex protein extracts or tryptic digests require some form of column separation to decrease the sample complexity and resolve the components individually. This can be accomplished by chromatographic or electrophoretic separation methods. LC column separation by partition chromatography on reversed phase (RP) sorbents [100] is the conventional method for both protein and peptide applications. Two dimensional chromatography with

strong cation exchange (SCX) [101] sorbents installed before the RP column enable orthogonal separations and higher amounts of starting material to be loaded. Alternatively, CE methods in free solution using an applied voltage to separate components based on electrophoretic mobilities have been used to separate complex samples with high efficiency [102].

1.5.2.2 The ionization source

The ion source is present at the exit of the separation column, taking in the liquid effluent and converting it to gaseous phase ions. The ESI emitter, transfer capillary, and electronic control components are contained within the source housing. Since the report by Fenn's group, ESI has been researched and improved extensively; over the past twenty years, these improvements include atmospheric pressure ionization (API) sources [103], pneumatically-assisted ESI [104], heat-assisted ESI [105], and nano-ESI for higher sensitivity [106, 107]. For proteomics applications, ion sources have been designed to accommodate lower and stable flow rates in the nL/min range, enabling the detection of very low abundance species from biological extracts. Narrow bore emitters also aid in the generation of smaller diameter microdroplets in the submicron range [108], thus resulting in more efficient desolvation and transfer of ions into the gas phase. Modular microfluidic devices (chip-LC) [109] comprising of chromatographic media, transfer lines, and nano ESI emitters have also been developed to integrate all components and simplify routine nano-ESI operation. Using nano-ESI sources, detection limits in the attomole range can be achieved.

1.5.2.3 The mass analyzer

The principle hardware that allows for discriminating ions based on their *m/z* ratio is the mass analyzer. The ability of a mass analyzer to differentiate neighboring ions is defined as the mass resolution [94]. This can be described by the following equation:

$$R = m/\Delta m \quad (1)$$

where *R* is the mass resolution, *m* is the *m/z* ratio under investigation, and *Δm* is the spacing between two neighboring mass spectral peaks (resolving power). Hence, the higher the value of R, the better the ability of the mass analyzer to separate two closely spaced *m/z* values. There are several types of mass analyzers (see Figure 1.11) for proteomic applications [110], each with their inherent advantages and disadvantages. These include the following: ion trap (IT), triple quadrupole (QqQ), time-of-flight (TOF), Orbitrap, and Fourier transform ion cyclotron resonance (FT-ICR). These analyzers can also be used in tandem in a hybrid type instrument. Storage based mass analyzers like IT use a three dimensional RF-field to capture and release ions [111]. The IT analyzer can been used to detect peptides and proteins, enabling the determination of *m/z* ratio, while taking advantage of their rapid duty cycle. However, because of low mass resolution (about 1000) and space charge effects, stand-alone ion trap instruments are not totally suited for the analysis of complex proteomic samples.

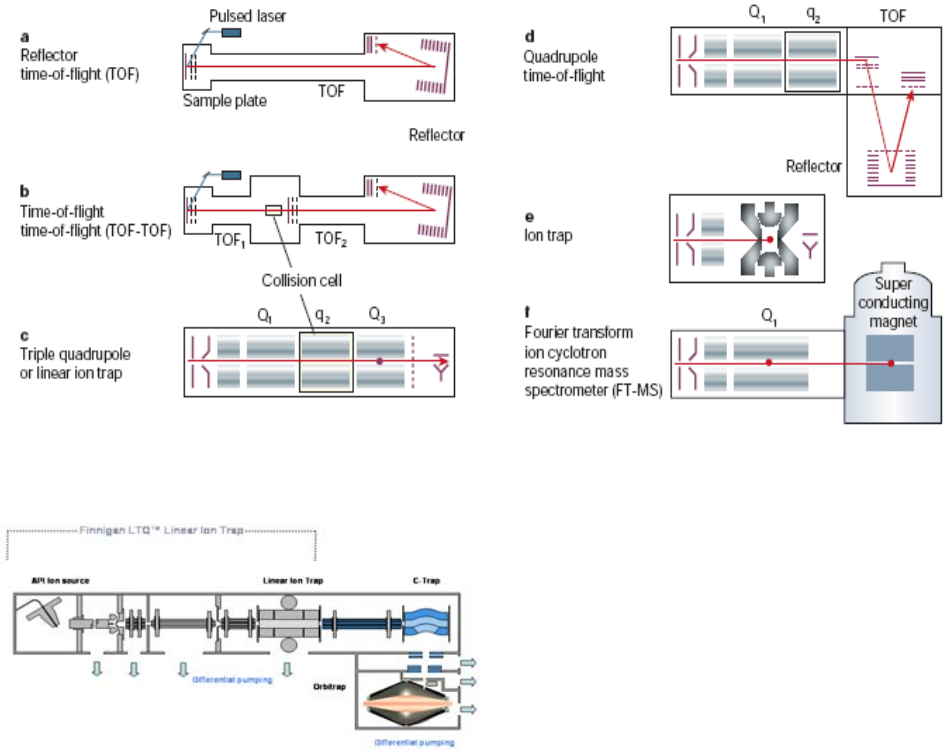


Figure 1.11: Examples of different mass analyzers used for proteomics (a-g). From references [110] and [112].

Both TOF and quadrupole-TOF (Q-TOF) instruments rely on the principle of separating ions in time during their travel through a fixed flight tube of a certain length [113]. Accurate calculation of this flight time results in the determination of the *m/z* ratio. Development of the reflectron as an "ion mirror" by Mamyrin [114] allowed TOF analyzers to better focus ions and elongate the flight path, resulting in higher resolution. The TOF analyzers can achieve mass resolution between 10 000 to 20 000, while more recent instrumentation can reach up to 40 000. The Orbitrap mass analyzer recently developed by Makarov [115] consists of 2 electrodes that form an electrostatic field. Ions are injected into the device and oscillate back and forth along the inner electrode. The frequency of these oscillations is used to calculate the *m/z* ratio with very high precision; mass resolutions of up to 100 000 are attainable with this type of mass analyzer. Lastly, FT-ICR (or FT-MS) is another type of mass analyzer used for both intact protein and peptide analysis. In FT-MS, ions are introduced into a cell (Penning trap) in the presence of a magnetic field [116]. Once the ions are trapped and in orbit, they can be excited by the application of a rotating RF electric field oscillating at the frequency of the ions of a given *m/z* value. Like the Orbitrap analyzer, cyclotronic frequencies for each ion can be converted to *m/z* ratios. FT-MS analyzers allow the highest attainable mass resolution, i.e., over 100 000.

1.5.2.4 Tandem mass spectrometry

The accurate determination of the ion *m/z* ratio is only one of the powerful features of a mass analyzer within the LC-MS analytical system. The second attribute that further extends the mass analyzers' capability is fragmentation of precursor ions and detection of fragment ions. This process is called tandem MS or MS/MS, and is the most widely used technique for fragmenting ions generated through "soft ionization" methods (ESI) [117]. The analysis of fragment ions in a single MS experiment allows for higher selectivity, structural elucidation, and sequencing of both proteins and peptides. A good example is the triple quadrupole mass spectrometer (see Figure 1.12). Triple quadrupole and hybrid triple quadrupole instruments with two mass analyzers and an RF-only collision cell in series [118] (QqQ and QqLIT) enable tandem MS capability, and provide a versatile mass analyzer with multiple types of scan modes, including precursor ion

scan, neutral loss, and multiple reaction monitoring (MRM). These scanning modes allow a researcher to not only obtain further structural information, but also to quantitate analytes with a high degree of precision while increasing signal-to-noise (S/N) ratios for detected species within complex mixtures. The MRM scan mode has been applied extensively by Aebersold's group to detect and quantify peptides with high precision and sensitivity in a targeted approach [119]. Other instrumental platforms use a TOF, Orbitrap, and FT-ICR to detect fragment ions with higher resolution.

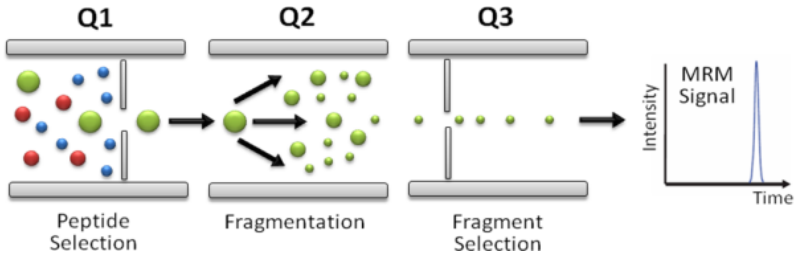


Figure 1.12: Principle of operation of the MRM mode in a triple quadrupole mass analyzer for peptide quantitation. From reference [120].

There are two principles modes of fragmentation used in conventional bottom-up proteomic analyses to elucidate the structure and sequence of peptides. The first is collision-induced dissociation (CID), and the second is electron transfer dissociation (ETD). These mechanisms require a collision or activation cell to fragment ions prior to mass analysis in a second stage.

1.5.2.5 Peptide fragmentation by collision-induced dissociation

In this mechanism, ions detected in the first mass analyzer (i.e., quadrupole or Orbitrap) are transferred to a collision cell along an ion path of increasing electric potential. Depending on the type of instrument, this cell can be an RF only quadrupole, or even an IT analyzer. Neutral molecules (i.e., helium, nitrogen) within the collision cell serve as targets for inelastic collisions with peptide ions. Upon collision, the precursor ions are activated to a higher energy state, whereby a portion of the translation energy imparted to peptide ions is converted to internal energy [121]. This energy is then

redirected and randomized throughout the peptide molecule, leading to the most thermodynamically stable bond cleavage (i.e., the path of least resistance) and fragment generation. The end result of this activation process is a charged fragment ion and a neutral species. This can be summarized in the following two-step reaction:

$$M_p^+ + N \rightarrow M_p^+ (act) \rightarrow m_f^+ + n^0 \quad (2)$$

Where M_p^+ is the precursor peptide ion, N is the neutral collision gas molecule, M_p^+ (act) is the activated precursor peptide ion, m_f^+ is the charged peptide fragment, and n^0 is the neutral peptide fragment.

Basic sites at the N-terminus and amino acid side chains provide anchor points for protonation. These protons are not totally static, but can move about, creating a heterogeneous population of peptide ions with the same sequence but protonated at different amide linkages [122]. According to the mobile proton model [123], these protons can migrate along the peptide backbone by internal solvation, and upon activation by CID, contribute in the fragmentation of the peptide in a "charge-directed" manner. In low-energy (up to 50 eV) collision instrumentation (i.e. quadrupole, Q-TOF,) this fragmentation scheme generates a predictable series of peptide fragment ion pairs by cleavage along the amide backbone [124], along with loss of water and ammonia. The fragmentation and generation of peptide ions is depicted in Figure 1.13.

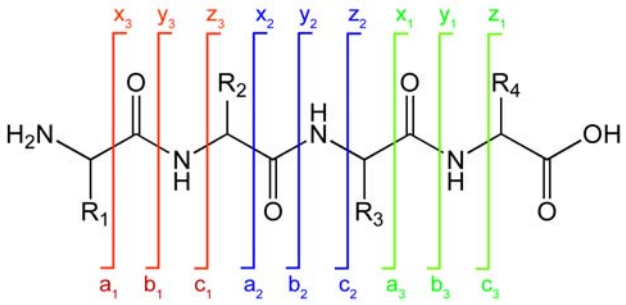


Figure 1.13: Peptide fragment ion generation. The types of ions produced depend upon where the cleavage site occurs. From references [125] and [126].

Fragments ions generated from peptide amide bond cleavages are classified as either *b*- or *y*-type ions, depending on which terminal end the fragmentation is initiated (N- or C-terminus) [127]. Cleavages of bonds adjacent to the amide bond result in either *a,x* or *c,z* ion pairs (see figure 1.13). Under CID conditions, the predominant ion pairs formed are *b* and *y* ions. Each amino acid will produce its own characteristic *b* and *y* fragment ions (see Figure 1.14); thus, the sequence of a given peptide can be deduced based on the type of fragment ions generated by tandem MS.

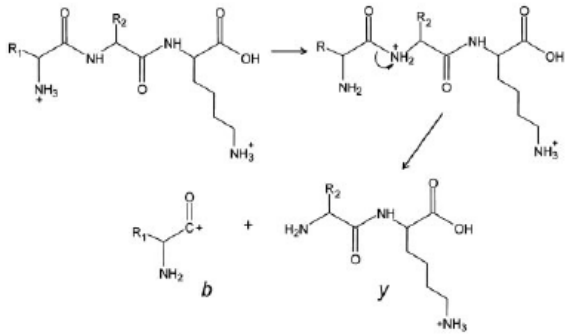


Figure 1.14: The CID fragmentation mechanism demonstrating the formation of *b*- and *y*-type ions. From reference [128].

1.5.2.6 Peptide fragmentation by electron transfer dissociation

Although a very powerful activation technique, CID does have inherent limitations. These include lower fragmentation efficiency of higher charge state ions (i.e., +4 and greater), as well as inducing loss of labile post translational modifications (PTMs), i.e, phosphorylation, sulfonation, and glycosylation (e.g. O-GlcNAcylation). The application of CID has been used successfully by many groups in the sequencing and localization of phosphorylation sites; for example, Gygi demonstrated the sequencing of over 2000 phosphopeptides from the nucleus of HeLa cells in a large scale study using CID [129]. Even though very functional and reproducible, tandem MS by CID is only providing a limited view of the modified peptides present in a given proteome. To further enable and extend the capabilities of tandem MS to sequence larger peptides, proteins, and labile PTM, new activation methods were required. A key breakthrough was reported in 1998 by McLafferty and Zubarev in the development of a new activation process called electron capture dissociation (ECD) [130]. In this technique, low energy, near-

thermal electrons are reacted with multiply charged protein/peptide cations [131], generating an odd-electron, hypervalent species. The dissociation of this radical species and subsequent rearrangement leads to fragmentation of the peptide. The bonds that are preferentially ruptured during ECD fragmentation are disulphide bridges and N-Cα by a *non-ergodic* mechanism [132]; these bonds are broken faster than the rate at which the vibrational energy is redirected throughout the molecule. Unlike CID, the most prominent ions formed in ECD are *c*- and *z*-type (see Figure 1.15). Furthermore, the fragmentation mechanism is gentle enough to preserve labile bonds and PTMs, as shown by Kelleher in the localization of carboxy and sulfated peptide residues [133]. The challenging aspect of ECD is the requirement of using FT-ICR instrumentation capable of producing the electrons, and the physical space for the capture with cations to occur. Also, FT-ICR instruments are the most expensive type of mass spectrometers to both purchase and maintain. This led researchers to the development of ECD-like fragmentation mechanism, but using more common and available instrumentation, such as quadrupoles and ion traps. The pioneer of electron transfer dissociation (ETD) was Donald Hunt at the University of Virginia. Instead of using FT-ICR with an electron generating source, his group used anionic molecules with a low affinity for electrons (i.e. anthracene) to deliver the electron, and a linear quadrupole ion trap as the reaction cell to perform ETD [128]. The anionic reagent molecules are produced in a chemical ionization source, and injected into a linear ion trap, where they react with peptide cations for a brief period of time (~100 milliseconds). The excess reagent is then purged out of the reaction cell, peptides fragmented, then scanned out. Since its introduction in 2004, ETD has been well characterized and improved. In a large scale study using both CID and ETD for the analysis of yeast extracts, Coon and colleagues demonstrated that ETD outperformed CID in the detection and sequencing of higher charge state species (+4 and higher), while at the same time sequencing peptides unique to ETD [134]. At present, ETD has been introduced in hybrid instruments, such as LTQ-Orbitrap, enabling both CID and ETD on the same platform on a chromatographic time scale [135]. Phosphopeptides can be easily sequenced and localized, while preserving the phosphate group on the peptide backbone. This allows for increased peptide sequencing rates, greater protein sequence coverage, and discovery of novel PTMs.

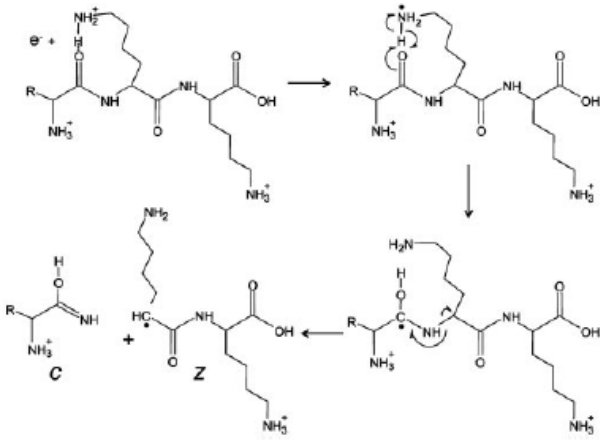


Figure 1.15: The ETD fragmentation mechanism demonstrating the formation of c- and z-type ions. From reference [128].

1.5.2.7 The detection module

After mass analysis, ions must make contact with a detector, and generate a digital signal. Ultimately, the sensitivity of the instrument relies upon how well the detector can sense the ions emerging from the mass analyzer. The two most common types of detectors used in mass spectrometers are the electron multiplier, and the microchannel plate (MCP) [94]. Electron multipliers can be either discreet or continuous [85], the latter form is usually referred to as a channel electron multiplier (CEM). The discreet form has a series of "cup" shaped dynodes coated with a surface that will release secondary electrons upon impact by an incident ion. Each dynode is held at higher potential, and will attract the electrons emitted from the previous dynode, causing an increasing cascade of electrons. The output gain from such a device can be from 10^6 up to 10^8. The final current generated can be transformed to a measurable signal using an analog-to-digital converter. The CEM is the continuous form of this detector. Rather than using discreet set of dynodes, the CEM is present in a curved open tube coated with a semi-conductive material. A beam of ions will generate the secondary ion cascade effect, and produce gains up to 10^8. A linear electron multiplier can be downsized to a few μm in diameter; the coupling of a high number (10^6) of these microscaled components linked

together forms a microchannel plate (MCP). The gain of such a detector is lower than a CEM (i.e., $10^3 - 10^4$). To improve the sensitivity, MCP's are usually stacked together.

1.5.2.8 Database searching and processing of proteomics data

A typical LC-MS run of a complex tryptic digest can result in the acquisition of thousands of MS/MS spectra. Manual interrogation of every spectrum in order to assign the peptide sequence and identify the source protein would be a very time consuming and laborious process. To expedite this task, internet database searching was designed and is now available. Mathematical algorithms for data processing were developed to assign a statistical score between an acquired and an "ideal" MS/MS spectrum. Peptide MS/MS spectra can be uploaded to a server (e.g. Mascot, see Figure 1.16) [136] and searched against a library of proteins. Raw data files are first converted to a mascot generic file (.mgf) which can be read by the search algorithm. A spectral library (e.g. NCBInr) can be selected, and the modifications of interest specified. After the search is initiated, the results are relayed back and provide a researcher with the most statistically relevant and plausible peptide sequences and PTMs present. A list of identifications includes the source protein, peptide sequence, modifications, and the Mascot score. In principle, a higher Mascot score for a given peptide is indicative of a more confident sequence assignment.

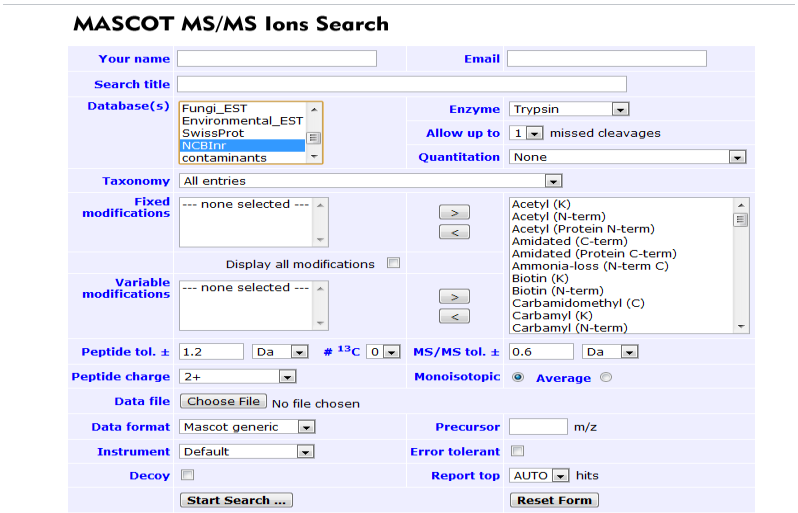


Figure 1.16: The Mascot server page for peptide MS/MS database searching (www.matrixscience.com).

1.6 The analysis of histones by mass spectrometry

Histone proteins are a challenging class of biomolecules to prepare, analyze, and sequence. The high content of basic amino acids renders them very polar with a high isoelectric point (pI ~ 11). The many, near isobaric forms, and combinatorial nature of PTMs on the N-terminus further complicates analysis and identification. Furthermore, biological relevant epigenetic markers may be present at very low abundance (< 1% of the unmodified precursor protein). The extent of histone PTM can be quite large; multiple modifications can occur on the same protein on different residues at different stages of the cell cycle. These hurdles require highly sensitive, robust, and versatile analytical instrumentation in order to quantitate and sequence histones and their PTMs. Hence, mass spectrometry is an obvious choice to characterize histones due to high sensitivity, speed of analysis, quantitative capability, and precise sequencing. The first reported usage of MS to identify methylated lysine residues in histones was by Murray in 1964 [137]. However, this was a rather isolated and rare publication; MS would not be used or documented routinely in histone research until the turn of the 21st century. Many research groups have attempted to tackle the challenge of analyzing histones by MS over the past 10 years, using both top-down and bottom-up sequencing strategies. Nevertheless, the stoichiometry, multiplicity, and half-life of these modifications still present a considerable challenge to any researcher. The use of top-down MS for analyzing intact histones [138],[139] was initiated by Neil Kelleher at the University of Illinois. This method is capable of providing complete sequence coverage and separation of many different isoforms within a group of core histone molecules. However, top down MS usually requires expensive FT-ICR instrumentation, hardware alteration, and sophisticated data processing software. Furthermore, the sample preparation for top down MS necessitates extensive HPLC fractionation, and a high amount of starting material. Only a few specialized laboratories can invest in such instrumentation. The alternative strategy, i.e., bottom-up sequencing, has been the most widely used approach for analyzing histones. Using enzymatically derived peptides (i.e., tryptic digestion) instead of intact histones simplifies the analysis and data interpretation, and more commonly available instrumentation (ion trap, Q-TOF) can be used. An early report by Burlingame

[140] was the first attempt at comprehensively mapping out histone PTMs; this consisted of tryptic digestion and MALDI-TOF analysis to identify histone H3 modifications from chicken erythrocytes. A more streamlined and robust workflow was developed by Hunt's group [141], utilizing chemical derivatization and labeling, HPLC fractionation, different enzymes, and high resolution LC-MS. This strategy was successfully used to map out histone H3 PTMs from HeLa cells. Mapping out and characterizing histone PTMs is an important first task, yet quantifying the relative abundance of histone PTMs under different experimental conditions is also vital in understanding their biological role and impact. Many reports to date utilized some form of chemical addition to perform differential PTM quantitation. These include stable isotopic labeling [142] to determine the *in-vitro* stoichiometry of acetylation at each lysine residue in the N-terminal tail of histone H4. Other *in-vivo* based methods consist of metabolic labeling with SILAC [143], and iTRAQ reagents [144] to compare the abundance of proteins or PTMs in two different samples (e.g. histones derived from control cells versus histones obtained from cells treated with DNA damaging agents). Although the labeling chemistry is well understood and functional, these procedures do suffer in reproducibility and dynamic range. Label-free methods represent the second approach for quantifying histone PTMs. This technique uses no incorporation of a chemical label at any stage of the sample preparation; peptides are analyzed in their natural state after proteolysis. However, this label-free approach requires rugged sample introduction (i.e., stable chromatography), and some form of bioinformatic tools to perform data mining. Without any software programs to aid in the data processing and interpretation, label-free methods result in a great deal of laborious and time consuming manual inspection of raw LC-MS data. Very few software tools are available commercially to aid researchers in applying label-free proteomic analyses. Some groups have addressed this need by designing in-house software to perform label-free analysis of their raw data. For example, VEMS software [145] and differential mass spectrometry (dMS) [146] have been documented for analyzing raw data; recently, Garcia's group also reported a new algorithm (PILOT PTM) [147] to rapidly search raw data in order to identify histone modifications. Regardless of which strategy is utilized, total comprehensive analysis of histone PTMs remains extremely challenging.

1.7 Research Objectives

Firstly, this Ph.D. research project focused on developing analytical strategies to analyze and characterize intact histone proteins and tryptic peptides. An analytically robust workflow had to be developed by combining all the necessities of performing such analyses, namely a rugged sample preparation procedure, modern high resolution MS instrumentation, and in-house developed bioinformatic data mining tools. These components were used together in order to comprehensively analyze and quantitate known and novel histone PTMs in a variety of biological matrices. The extent of histone PTMs can vary over a large range (some are very low in abundance, while others are very abundant). This range can extend up to three orders of magnitude, placing a heavy burden on the detection capabilities of current MS instrumentation. Modifications of interest can be present at very low stoichiometry, and result in high degrees of analytical error and uncertainty in measurement. The determination of analytical error is rarely assessed or reported in the literature; this became a second objective in this section of the project. Even in a rather simple organism such as yeast (*S. cerevisiae*), a great deal of histone PTMs are present in a large distribution of abundances. One key modified histone H3 peptide, i.e., acetylated lysine 56 (H3K56ac) is known to play a vital role in the DNA damage response in *S. cerevisiae* [148]. Current methods used to characterize and quantitate this PTM rely on either immunoblots, or generic LC-MS experiments to detect this peptide. The last objective of this section was to reproducibly quantitate the relative abundance of this important PTM in WT and mutant yeast strains using our in-house developed workflow.

Secondly, novel uses of existing MS instrumentation were investigated and developed. Not all laboratories have the luxury of possessing an instrumental platform from each major MS vendor. In some instances, researchers must use and adapt the technology at hand to answer biological questions. Ion trapping instruments are both versatile and readily available. However, these instruments are rather limited in their proteomic applications due to lower resolution, and space-charge effects. Alterations in the instrument software and scanning functions of a hybrid quadrupole-linear ion trap

instrument were tested and optimized in order to expand the analytical capabilities of the mass spectrometer, and apply these improvements to perform both intact histone profiling and bottom-up sequencing on the same MS platform. The objective of this work was to determine if trapping instruments could be used to analyze intact proteins from complex mixtures, and use this newly developed scan mode to profile changes in histone PTMs in WT and DOT1L-deficient DT40 cells.

Thirdly, using the analytical workflow developed and validated, the effects of HDAC inhibitors on histones derived from both normal and cancer cells were examined. Most published literature rely on immunoblots to assess which modification sites are affected by exposure to HDAC inhibitors. These methods are limited by the potential of cross-reactivity, poor quantitative capability, and may lead to erroneous results and conclusions. In mammalian cells, histone PTMs can vary in a large range of abundance and isobaric species, thus making detection and quantitation challenging. The primary objective of this study was to determine with a high degree of accuracy which core histones were affected, which PTMs were up-regulated, and the extent of the fold-change in histone PTMs induced by treating cells with HDAC inhibitors. Furthermore, we wanted to shed some light on important biological questions, such as: (1) do HDAC inhibitors affect normal cells in the same manner as cancer cells? (2) Does H3K56ac have the same function in mammalian cells as in yeast cells? (3) Can this information be used to develop more specific HDACi as therapeutic agents with fewer side effects?

1.8 Contents of the Thesis

The second chapter of the thesis presents an overview of the experimental aspects used in the generation and interpretation of proteomics data.

The third chapter of the thesis presents the development and validation of an analytical workflow to analyze and quantitate histone PTMs. A dual approach is described, whereby both intact mass profiling and peptide sequencing are utilized. The reproducibility of the workflow was assessed in order to generate analytical figures of merit. A tailored sample preparation for histones was adapted and developed, while applying in-house developed bioinformatic tools (Mass Sense peptide miner and clusterer) to probe the data. Modern high resolution LC-MS instrumentation (Agilent TOF and Waters Q-TOF premier) was used to analyze and sequence histone peptides. The methodology was validated using histones extracted from yeast wild-type and mutant strains to establish structure-function relationships.

The fourth chapter of the thesis presents the development of a new scanning mode on the AB Sciex 4000 Q-trap hybrid quadrupole-linear ion trap instrument. The modified scanning mode (named targeted EMC) was designed by altering the scanning parameters and voltages of the instrument. This resulted in both higher sensitivity and resolution, while minimizing space charge effects. Figures of merit and practical instrumental limits were established. Histones from chicken DT40 cells were analyzed to demonstrate the applicability of this new scanning mode to profile both intact histones, while using conventional MRM mode to localize and quantitate differentially abundant histone PTMs on the same instrument.

The fifth chapter of the thesis presents the results and findings of an HDAC inhibitor study. Two clinically relevant HDAC inhibitors, namely Vorinostat (SAHA, marketed as *Zolinza*) and Entinostat (MS-275), were selected for investigation. Both normal and leukemic K562 cells were used in this study. The methodology developed and presented in the third chapter of the thesis was applied to determine the global effects

on histone PTMs by these small molecule inhibitors. Several known acetylation sites on histones H3 and H4 exhibited a significant increase in acetylation of histones derived from K562 cells treated with HDAC inhibitors. Surprisingly, the HDAC inhibitors did not increase histone acetylation in non-transformed cells. Every type of MS platform in the laboratory was exploited to gain a better understanding of how histone PTMs were affected by the various HDACi. Targeted MRM was also used to detect and quantitate low abundance PTMs.

Lastly, the sixth chapter presents a conclusion that links all the work and findings presented in the thesis, as well as potential future perspectives for the analysis of histones by mass spectrometry.

1.9 References

1. Stryer, L., *Biochemistry*. 1988, New York: W.H. Freeman and Company.

2. Watson, J.D. and F.H. Crick, *Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid.* Nature, 1953. **171**(4356): p. 737-8.

3. McConnell, K.J. and D.L. Beveridge, *DNA structure: what's in charge?* J Mol Biol, 2000. **304**(5): p. 803-20.

4. Schneider, B., et al., *A systematic method for studying the spatial distribution of water molecules around nucleic acid bases.* Biophys J, 1993. **65**(6): p. 2291-303.

5. Drew, H.R., et al., *Structure of a B-DNA dodecamer: conformation and dynamics.* Proc Natl Acad Sci U S A, 1981. **78**(4): p. 2179-83.

6. Wolffe, A., *Chromatin: Structure and Function*. Second ed. 1995, New York: Academic Press.

7. Bacolla, A. and R.D. Wells, *Non-B DNA conformations, genomic rearrangements, and human disease.* J Biol Chem, 2004. **279**(46): p. 47411-4.

8. Dickerson, R.E., et al., *The anatomy of A-, B-, and Z-DNA.* Science, 1982. **216**(4545): p. 475-85.

9. Fratini, A.V., et al., *Reversible bending and helix geometry in a B-DNA dodecamer: CGCGAATTBrCGCG.* J Biol Chem, 1982. **257**(24): p. 14686-707.

10. Kouzine, F. and D. Levens, *Supercoil-driven DNA structures regulate genetic transactions.* Front Biosci, 2007. **12**: p. 4409-23.

11. Yang, Y., et al., *Effects of localized bending on DNA supercoiling.* Trends Biochem Sci, 1995. **20**(8): p. 313-9.

12. *Understanding Evolution*. 02-November-2010; DNA strand. Available from: http://evolution.berkeley.edu/evolibrary/article/side_o_0/reviewdna_01.

13. *Chemical structure of DNA*. 30-Jan-2011; Available from: http://www.scientificpsychic.com/fitness/dna0.gif.

14. Venter, J.C., et al., *The sequence of the human genome.* Science, 2001. **291**(5507): p. 1304-51.

15. Holde, K.E.v., *Chromatin*. 1989: Springer-Verlag.

16. Clark, R.J. and G. Felsenfeld, *Structure of chromatin.* Nat New Biol, 1971. **229**(4): p. 101-6.

17. Hewish, D.R. and L.A. Burgoyne, *Chromatin sub-structure. The digestion of chromatin DNA at regularly spaced sites by a nuclear deoxyribonuclease.* Biochem Biophys Res Commun, 1973. **52**(2): p. 504-10.

18. Noll, M., *Internal structure of the chromatin subunit.* Nucleic Acids Res, 1974. **1**(11): p. 1573-8.

19. Olins, A.L. and D.E. Olins, *Spheroid chromatin units (v bodies).* Science, 1974. **183**(4122): p. 330-2.

20. Luger, K., et al., *Crystal structure of the nucleosome core particle at 2.8 A resolution.* Nature, 1997. **389**(6648): p. 251-60.

21. Richmond, T.J. and C.A. Davey, *The structure of DNA in the nucleosome core.* Nature, 2003. **423**(6936): p. 145-50.

22. *The role of chromatin structure in the regulation of gene switching*. 02-November 2010; Histone proteins packaging DNA into the cell nucleus. Available from: http://www.rikenresearch.riken.jp/eng/frontline/6134.

23. Kornberg, R.D., *Chromatin structure: a repeating unit of histones and DNA.* Science, 1974. **184**(139): p. 868-71.

24. Davey, C.A., et al., *Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution.* J Mol Biol, 2002. **319**(5): p. 1097-113.

25. Placek, B.J. and L.M. Gloss, *The N-terminal tails of the H2A-H2B histones affect dimer structure and stability.* Biochemistry, 2002. **41**(50): p. 14960-8.

26. Arents, G. and E.N. Moudrianakis, *The histone fold: a ubiquitous architectural motif utilized in DNA compaction and protein dimerization.* Proc Natl Acad Sci U S A, 1995. **92**(24): p. 11170-4.

27. Garcia-Perez, M., M. Pinto, and J.A. Subirana, *Nonsequence-specific arginine interactions in the nucleosome core particle.* Biopolymers, 2003. **69**(4): p. 432-9.

28. *Histone Octamer*. 02-November 2010; Assembly of core histones into the octomeric complex. Available from: http://en.wikipedia.org/wiki/Histones.

29. Adkins, N.L., M. Watts, and P.T. Georgel, *To the 30-nm chromatin fiber and beyond.* Biochim Biophys Acta, 2004. **1677**(1-3): p. 12-23.

30. Woodcock, C.L., A.I. Skoultchi, and Y. Fan, *Role of linker histone in chromatin structure and function: H1 stoichiometry and nucleosome repeat length.* Chromosome Res, 2006. **14**(1): p. 17-25.

31. Raghuram, N., et al., *Molecular dynamics of histone H1.* Biochem Cell Biol, 2009. **87**(1): p. 189-206.

32. Hartman, P.G., et al., *Studies on the role and mode of operation of the very-lysine-rich histone H1 in eukaryote chromatin. The three structural regions of the histone H1 molecule.* Eur J Biochem, 1977. **77**(1): p. 45-51.

33. Widom, J., *Chromatin structure: linking structure to function with histone H1.* Curr Biol, 1998. **8**(22): p. R788-91.

34. Carruthers, L.M., et al., *Linker histones stabilize the intrinsic salt-dependent folding of nucleosomal arrays: mechanistic ramifications for higher-order chromatin folding.* Biochemistry, 1998. **37**(42): p. 14776-87.

35. Khorasanizadeh, S., *The nucleosome: from genomic organization to genomic regulation.* Cell, 2004. **116**(2): p. 259-72.

36. 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.* Proc Natl Acad Sci U S A, 1998. **95**(24): p. 14173-8.

37. Staynov, D.Z., *DNase I digestion reveals alternating asymmetrical protection of the nucleosome by the higher order chromatin structure.* Nucleic Acids Res, 2000. **28**(16): p. 3092-9.

38. Schalch, T., et al., *X-ray structure of a tetranucleosome and its implications for the chromatin fibre.* Nature, 2005. **436**(7047): p. 138-41.

39. Tremethick, D.J., *Higher-order structures of chromatin: the elusive 30 nm fiber.* Cell, 2007. **128**(4): p. 651-4.

40. Hizume, K., et al., *Topoisomerase II, scaffold component, promotes chromatin compaction in vitro in a linker-histone H1-dependent manner.* Nucleic Acids Res, 2007. **35**(8): p. 2787-99.

41. Woodcock, C.L., *Chromatin architecture.* Curr Opin Struct Biol, 2006. **16**(2): p. 213-20.

42. D'Angelo, M.A. and M.W. Hetzer, *The role of the nuclear envelope in cellular organization.* Cell Mol Life Sci, 2006. **63**(3): p. 316-32.

43. Phillips, D.M., *The presence of acetyl groups of histones.* Biochem J, 1963. **87**: p. 258-63.

44. Kouzarides, T., *Chromatin modifications and their function.* Cell, 2007. **128**(4): p. 693-705.

45. Ehrenhofer-Murray, A.E., *Chromatin dynamics at DNA replication, transcription and repair.* Eur J Biochem, 2004. **271**(12): p. 2335-49.

46. Allis, C.D., T. Jenuwein, and D. Reinberg, *Epigenetics*. 2007, New York: Cold Spring Harbor Laboratory Press.

47. Marmorstein, R., *Structure of histone acetyltransferases.* J Mol Biol, 2001. **311**(3): p. 433-44.

48. Kasten, M., et al., *Tandem bromodomains in the chromatin remodeler RSC recognize acetylated histone H3 Lys14.* EMBO J, 2004. **23**(6): p. 1348-59.

49. Lachner, M., et al., *Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins.* Nature, 2001. **410**(6824): p. 116-20.

50. Ahmad, K. and S. Henikoff, *The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly.* Mol Cell, 2002. **9**(6): p. 1191-200.

51. Henikoff, S. and K. Ahmad, *Assembly of variant histones into chromatin.* Annu Rev Cell Dev Biol, 2005. **21**: p. 133-53.

52. Hake, S.B., et al., *Serine 31 phosphorylation of histone variant H3.3 is specific to regions bordering centromeres in metaphase chromosomes.* Proc Natl Acad Sci U S A, 2005. **102**(18): p. 6344-9.

53. Fraga, M.F., et al., *Epigenetic differences arise during the lifetime of monozygotic twins.* Proc Natl Acad Sci U S A, 2005. **102**(30): p. 10604-9.

54. Goldberg, A.D., C.D. Allis, and E. Bernstein, *Epigenetics: a landscape takes shape.* Cell, 2007. **128**(4): p. 635-8.

55. Bernstein, B.E., A. Meissner, and E.S. Lander, *The mammalian epigenome.* Cell, 2007. **128**(4): p. 669-81.

56. Goll, M.G. and T.H. Bestor, *Eukaryotic cytosine methyltransferases.* Annu Rev Biochem, 2005. **74**: p. 481-514.

57. Jenuwein, T. and C.D. Allis, *Translating the histone code.* Science, 2001. **293**(5532): p. 1074-80.

58. Strahl, B.D. and C.D. Allis, *The language of covalent histone modifications.* Nature, 2000. **403**(6765): p. 41-5.

59. 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.* Mol Cell, 2000. **5**(6): p. 917-26.

60. Lund, A.H. and M. van Lohuizen, *Epigenetics and cancer.* Genes Dev, 2004. **18**(19): p. 2315-35.

61. William B. Coleman, G.J.T., *The Molecular Basis of Human Cancer*. 2002, New Jersey: Humana Press.

62. Rodenhuis, S. and R.J. Slebos, *Clinical significance of ras oncogene activation in human lung cancer.* Cancer Res, 1992. **52**(9 Suppl): p. 2665s-2669s.

63. Smit, V.T., et al., *KRAS codon 12 mutations occur very frequently in pancreatic adenocarcinomas.* Nucleic Acids Res, 1988. **16**(16): p. 7773-82.

64. Hanahan, D. and R.A. Weinberg, *The hallmarks of cancer.* Cell, 2000. **100**(1): p. 57-70.

65. Gonzalez-Zulueta, M., et al., *Methylation of the 5' CpG island of the p16/CDKN2 tumor suppressor gene in normal and transformed human tissues correlates with gene silencing.* Cancer Res, 1995. **55**(20): p. 4531-5.

66. Fahrner, J.A., et al., *Dependence of histone modifications and gene expression on DNA hypermethylation in cancer.* Cancer Res, 2002. **62**(24): p. 7213-8.

67. Okada, Y., et al., *hDOT1L links histone methylation to leukemogenesis.* Cell, 2005. **121**(2): p. 167-78.

68. Sorm, F., et al., *5-Azacytidine, a new, highly effective cancerostatic.* Experientia, 1964. **20**(4): p. 202-3.

69. Graham, J.S., S.B. Kaye, and R. Brown, *The promises and pitfalls of epigenetic therapies in solid tumours.* Eur J Cancer, 2009. **45**(7): p. 1129-36.

70. Stresemann, C. and F. Lyko, *Modes of action of the DNA methyltransferase inhibitors azacytidine and decitabine.* Int J Cancer, 2008. **123**(1): p. 8-13.

71. Daskalakis, M., et al., *Demethylation of a hypermethylated P15/INK4B gene in patients with myelodysplastic syndrome by 5-Aza-2'-deoxycytidine (decitabine) treatment.* Blood, 2002. **100**(8): p. 2957-64.

72. Kaminskas, E., et al., *FDA drug approval summary: azacitidine (5-azacytidine, Vidaza) for injectable suspension.* Oncologist, 2005. **10**(3): p. 176-82.

73. Lin, R.J., et al., *Transcriptional regulation in acute promyelocytic leukemia.* Oncogene, 2001. **20**(49): p. 7204-15.

74. Wilson, A.J., et al., *HDAC4 promotes growth of colon cancer cells via repression of p21.* Mol Biol Cell, 2008. **19**(10): p. 4062-75.

75. de Ruijter, A.J., et al., *Histone deacetylases (HDACs): characterization of the classical HDAC family.* Biochem J, 2003. **370**(Pt 3): p. 737-49.

76. Marks, P.A. and R. Breslow, *Dimethyl sulfoxide to vorinostat: development of this histone deacetylase inhibitor as an anticancer drug.* Nat Biotechnol, 2007. **25**(1): p. 84-90.

77. Richon, V.M., et al., *Histone deacetylase inhibitor selectively induces p21WAF1 expression and gene-associated histone acetylation.* Proc Natl Acad Sci U S A, 2000. **97**(18): p. 10014-9.

78. Altucci, L., et al., *RAR and RXR modulation in cancer and metabolic disease.* Nat Rev Drug Discov, 2007. **6**(10): p. 793-810.

79. Wisniak, J., *Jöns Jacob Berzelius: A Guide to the Perplexed Chemist.* Chem. Educator, 2000. **5**: p. 343-350.

80. Creighton, T.E., *Proteins: Structures and Molecular Properties*. 1984, New York: W.H. Freeman and Company.

81. Pauling, L. and R.B. Corey, *Atomic coordinates and structure factors for two helical configurations of polypeptide chains.* Proc Natl Acad Sci U S A, 1951. **37**(5): p. 235-40.

82. Edman, P., *On the mechanism of the phenyl isothiocyanate degradation of peptides.* Acta Chem. Scand, 1956. **10**: p. 761-768.

83. Voet, D., *Biochemistry*. 2nd ed. 1995, New York: John Wiley and Sons.

84. Steen, H. and M. Mann, *The ABC's (and XYZ's) of peptide sequencing.* Nat Rev Mol Cell Biol, 2004. **5**(9): p. 699-711.

85. Skoog, D.A. and J.J. Leary, *Principles of Instrumental Analysis*. 4th ed. 1992, New York: Harcourt Brace College.

86. Fenn, J.B., et al., *Electrospray ionization for mass spectrometry of large biomolecules.* Science, 1989. **246**(4926): p. 64-71.

87. Fenn, J.B., *Electrospray ionization mass spectrometry: How it all began.* J Biomol Tech, 2002. **13**(3): p. 101-18.

88. Clegg, G.A. and M. Dole, *Molecular beams of macroions. 3. Zein and polyvinylpyrrolidone.* Biopolymers, 1971. **10**(5): p. 821-6.

89. Whitehouse, C.M., et al., *Electrospray interface for liquid chromatographs and mass spectrometers.* Anal Chem, 1985. **57**(3): p. 675-9.

90. Higuera, F.J., *Ion evaporation from the surface of a Taylor cone.* Phys Rev E Stat Nonlin Soft Matter Phys, 2003. **68**(1 Pt 2): p. 016304.

91. van Honschoten, J.W., N. Brunets, and N.R. Tas, *Capillarity at the nanoscale.* Chem Soc Rev, 2010. **39**(3): p. 1096-114.

92. Tamashiro, M.N. and H. Schiessel, *Rayleigh instability of charged aggregates: Role of the dimensionality, ionic strength, and dielectric contrast.* Phys Rev E Stat Nonlin Soft Matter Phys, 2006. **74**(2 Pt 1): p. 021412.

93. Rossi, D.T. and M.W. Sinz, *Mass Spectrometry in Drug Discovery*. 1st ed. 2002, New York: Marcel Dekker.

94. Gross, J.H., *Mass Spectrometry*. 2004, Berlin: Springer-Verlag.

95. Kerbarle, P. and L. Tang, *From ions in solution to ions in the gas phase: the mechanism of electrospray mass spectrometry.* Anal. Chem., 1993. **65**(22): p. 972A-986A.

96. Iribarne, J.V. and B.A. Thomson, *On the evaporation of small ions from charged droplets.* Journal of Chemical Physics, 1976. **64**(6): p. 2287-2294.

97. *Electrospray Ionization*. 02-November-2010; The electrospray ionization process. Available from: http://www.rsc.org/chemistryworld/Issues/2003/February/index.asp.

98. Schoenherr, R.M., et al., *CE-microreactor-CE-MS/MS for protein analysis.* Anal Chem, 2007. **79**(6): p. 2230-8.

99. Wang, H. and S. Hanash, *Multi-dimensional liquid phase based separations in proteomics.* J Chromatogr B Analyt Technol Biomed Life Sci, 2003. **787**(1): p. 11-8.

100. Snyder, L.R., J.J. Kirkland, and J.L. Glajch, *Practical HPLC method development*. 2nd ed. 1997, New York: John Wiley and sons.

101. Wolters, D.A., M.P. Washburn, and J.R. Yates, 3rd, *An automated multidimensional protein identification technology for shotgun proteomics.* Anal Chem, 2001. **73**(23): p. 5683-90.

102. Moini, M., *Capillary electrophoresis mass spectrometry and its application to the analysis of biological mixtures.* Anal Bioanal Chem, 2002. **373**(6): p. 466-80.

103. Bruins, A.P., *Liquid chromatography-mass spectrometry with ionspray and electrospray interfaces in pharmaceutical and biomedical research.* J Chromatogr, 1991. **554**(1-2): p. 39-46.

104. Bruins, A.P., T.R. Covey, and J.D. Henion, *Ion spray interface for combined liquid chromatography/atmospheric pressure ionization mass spectrometry.* Anal. Chem., 1987. **59**(22): p. 2642-2646.

105. Ikonomou, M.G. and P. Kerbarle, *A Heated ESI Source In Mass Spectrometry for Analytes in Aqueous Solutions.* J. Am. Soc. Mass Spectrom., 1994. **5**: p. 791-799.

106. Wilm, M. and M. Mann, *Analytical properties of the nanoelectrospray ion source.* Anal Chem, 1996. **68**(1): p. 1-8.

107. Xie, J., et al., *Microfluidic platform for liquid chromatography-tandem mass spectrometry analyses of complex peptide mixtures.* Anal Chem, 2005. **77**(21): p. 6947-53.

108. Ku, K.K., et al., *Direct measurement of electrospray droplets in submicron diameter using a freezing method and a TEM image processing technique.* Journal of Aerosol Science, 2001. **32**: p. 1459-1477.

109. Yin, H., et al., *Microfluidic chip for peptide analysis with an integrated HPLC column, sample enrichment column, and nanoelectrospray tip.* Anal Chem, 2005. **77**(2): p. 527-33.

110. Aebersold, R. and M. Mann, *Mass spectrometry-based proteomics.* Nature, 2003. **422**(6928): p. 198-207.

111. Jonscher, K.R. and J.R. Yates, 3rd, *The quadrupole ion trap mass spectrometer--a small solution to a big challenge.* Anal Biochem, 1997. **244**(1): p. 1-15.

112. *Orbitrap Mass Analyzer*. 02-November-2010; Available from: http://www.google.ca/imgres?imgurl=http://cbsu.tc.cornell.edu/vanwijk/images/massspec/massspec3_large.png&imgrefurl=http://cbsu.tc.cornell.edu/vanwijk/mass_spec.htm&usg=__TNOXtecjwy8XVeV91DI6axMgeoQ=&h=540&w=720&sz=30&hl=en&start=0&sig2=p_4QOZ4jivAMMA_TseLVFw&zoom=1&tbnid=Ije65JnlALkAjM:&tbnh=103&tbnw=135&ei=DFnQTLP1GYSdlgee9tXfBQ&prev=/images%3Fq%3DOrbitrap%26um%3D1%26hl%3Den%26sa%3DN%26biw%3D

1280%26bih%3D479%26tbs%3Disch:1&um=1&itbs=1&iact=hc&vpx=364&vpy=62&dur=4042&hovh=194&hovw=259&tx=151&ty=117&oei=DFnQTLP1GYSdlgee9tXfBQ&esq=1&page=1&ndsp=24&ved=1t:429,r:2,s:0.

113. Chernushevich, I.V., A.V. Loboda, and B.A. Thomson, *An introduction to quadrupole-time-of-flight mass spectrometry.* J Mass Spectrom, 2001. **36**(8): p. 849-65.

114. Mamyrin, B.A., *Laser Assisted Reflectron Time-of-Flight Mass Spectrometry.* Int. J. Mass Spectrom. Ion Proc., 1994. **131**: p. 1-19.

115. Hu, Q., et al., *The Orbitrap: a new mass spectrometer.* J Mass Spectrom, 2005. **40**(4): p. 430-43.

116. Barrow, M.P., W.I. Burkitt, and P.J. Derrick, *Principles of Fourier transform ion cyclotron resonance mass spectrometry and its application in structural biology.* Analyst, 2005. **130**(1): p. 18-28.

117. Niessen, W.M.A. and J. van der Greef, *Liquid Chromatography-Mass Spectrometry*. Chromatographic Science Series. Vol. 58. 1992, New York: Marcel Dekker.

118. Hopfgartner, G., et al., *Triple quadrupole linear ion trap mass spectrometer for the analysis of small molecules and macromolecules.* J Mass Spectrom, 2004. **39**(8): p. 845-55.

119. Picotti, P., et al., *Full dynamic range proteome analysis of S. cerevisiae by targeted proteomics.* Cell, 2009. **138**(4): p. 795-806.

120. *SRM/MRM assays for targeted proteomic analysis*. 02-November-2010; Available from: http://www.mrmatlas.org/mrmassays.php.

121. Sleno, L. and D.A. Volmer, *Ion activation methods for tandem mass spectrometry.* J Mass Spectrom, 2004. **39**(10): p. 1091-112.

122. Kinter, M. and N.E. Sherman, *Protein Sequencing and Identification Using Tandem Mass Spectrometry*. 2000, New York: Wiley-Interscience.

123. Wysocki, V.H., et al., *Mobile and localized protons: a framework for understanding peptide dissociation.* J Mass Spectrom, 2000. **35**(12): p. 1399-406.

124. Hunt, D.F., et al., *Protein sequencing by tandem mass spectrometry.* Proc Natl Acad Sci U S A, 1986. **83**(17): p. 6233-7.

125. Roepstorff, P. and J. Fohlman, *Proposal for a common nomenclature for sequence ions in mass spectra of peptides.* Biomed Mass Spectrom, 1984. **11**(11): p. 601.

126. *Peptide Fragmentation*. 01-Feb-2011; Available from: http://www.chempep.com/ChemPep-Generic-Term_clip_image006.jpg.

127. Biemann, K., *Mass spectrometry of peptides and proteins.* Annu Rev Biochem, 1992. **61**: p. 977-1010.

128. Syka, J.E., et al., *Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry.* Proc Natl Acad Sci U S A, 2004. **101**(26): p. 9528-33.

129. Beausoleil, S.A., et al., *Large-scale characterization of HeLa cell nuclear phosphoproteins.* Proc Natl Acad Sci U S A, 2004. **101**(33): p. 12130-5.

130. Zubarev, R.A., N.L. Kelleher, and F.W. McLafferty, *Electron capture dissociation of multiply charged protein cations. A nonergoic process.* J. Am. Chem.Soc., 1998. **120**: p. 3265-3266.

131. McLafferty, F.W., et al., *Electron capture dissociation of gaseous multiply charged ions by Fourier-transform ion cyclotron resonance.* J Am Soc Mass Spectrom, 2001. **12**(3): p. 245-9.

132. Haselmann, K.F., et al., *Electron capture dissociation of weakly bound polypeptide polycationic complexes.* Rapid Commun Mass Spectrom, 2002. **16**(24): p. 2260-5.

133. Kelleher, N.L., et al., *Localization of labile posttranslational modifications by electron capture dissociation: the case of gamma-carboxyglutamic acid.* Anal Chem, 1999. **71**(19): p. 4250-3.

134. Good, D.M., et al., *Performance characteristics of electron transfer dissociation mass spectrometry.* Mol Cell Proteomics, 2007. **6**(11): p. 1942-51.

135. McAlister, G.C., et al., *A proteomics grade electron transfer dissociation-enabled hybrid linear ion trap-orbitrap mass spectrometer.* J Proteome Res, 2008. **7**(8): p. 3127-36.

136. Perkins, D.N., et al., *Probability-based protein identification by searching sequence databases using mass spectrometry data.* Electrophoresis, 1999. **20**(18): p. 3551-67.

137. Murray, K., *The Occurrence of Epsilon-N-Methyl Lysine in Histones.* Biochemistry, 1964. **3**: p. 10-5.

138. Boyne, M.T., 2nd, et al., *Precise characterization of human histones in the H2A gene family by top down mass spectrometry.* J Proteome Res, 2006. **5**(2): p. 248-53.

139. Pesavento, J.J., C.A. Mizzen, and N.L. Kelleher, *Quantitative analysis of modified proteins and their positional isomers by tandem mass spectrometry: human histone H4.* Anal Chem, 2006. **78**(13): p. 4271-80.

140. Zhang, K., et al., *Identification of acetylation and methylation sites of histone H3 from chicken erythrocytes by high-accuracy matrix-assisted laser desorption ionization-time-of-flight, matrix-assisted laser desorption ionization-postsource decay, and nanoelectrospray ionization tandem mass spectrometry.* Anal Biochem, 2002. **306**(2): p. 259-69.

141. Syka, J.E., et al., *Novel linear quadrupole ion trap/FT mass spectrometer: performance characterization and use in the comparative analysis of histone H3 post-translational modifications.* J Proteome Res, 2004. **3**(3): p. 621-6.

142. Smith, C.M., et al., *Mass spectrometric quantification of acetylation at specific lysines within the amino-terminal tail of histone H4.* Anal Biochem, 2003. **316**(1): p. 23-33.

143. Su, X., et al., *Histone H4 acetylation dynamics determined by stable isotope labeling with amino acids in cell culture and mass spectrometry.* Anal Biochem, 2007. **363**(1): p. 22-34.

144. Ross, P.L., et al., *Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents.* Mol Cell Proteomics, 2004. **3**(12): p. 1154-69.

145. Matthiesen, R., et al., *VEMS 3.0: algorithms and computational tools for tandem mass spectrometry based identification of post-translational modifications in proteins.* J Proteome Res, 2005. **4**(6): p. 2338-47.

146. Wiener, M.C., et al., *Differential mass spectrometry: a label-free LC-MS method for finding significant differences in complex peptide and protein mixtures.* Anal Chem, 2004. **76**(20): p. 6085-96.

147. Baliban, R.C., et al., *A novel approach for untargeted post-translational modification identification using integer linear optimization and tandem mass spectrometry.* Mol Cell Proteomics, 2010. **9**(5): p. 764-79.

148. Ozdemir, A., et al., *Histone H3 lysine 56 acetylation: a new twist in the chromosome cycle.* Cell Cycle, 2006. **5**(22): p. 2602-8.

2. Experimental Section

When this Ph.D. project was initiated, there was no analytical workflow developed in-house in order to prepare and analyze histone PTMs by mass spectrometry. The four key elements responsible for the success of this project were the following: (1) a sample preparation and digestion procedure designed for histones, (2) nano scale HPLC, (3) state of the art MS detection for quantitation and sequencing, and (4) the use of specially designed bioinformatic tools and database search engines to rapidly analyze proteomics raw data. The following sections provide greater detail on these important aspects.

2.1 Sample Preparation for Histones

Histones derived from cellular extracts are present in many different modified forms. To simplify the sample complexity and allow comparison between different experimental conditions, an HPLC separation method using a reversed-phase C_{18} column was developed. This resulted in core histones being almost baseline resolved from each other. A solvent system using acetonitrile and water both containing trifluoroacetic acid (0.1% v/v) was used with gradient elution (e.g. 5 to 90% organic component) to separate the intact histones. Histones contain a great deal of basic residues, especially on their N-terminus. Direct tryptic digestion would result in the generation of many small proteolytic fragments, many of which would not retain on a conventional C_{18} column. Although sufficient for mapping studies, direct tryptic digestion is not totally suitable for quantitative studies, where the generation of many digestion products can result in a PTMs being present on multiple peptides. To circumvent this challenge, a different approach was taken. Using a chemical derivatization method initially developed and reported by Hunt and colleagues [1], we tested and fine-tuned the reaction for histone preparation in our laboratory. HPLC purified histones were placed in basic media (0.1 M ammonium bicarbonate buffer, pH 8), and reacted with a propionic anhydride reagent. This reaction modifies all free lysines and converts them to a propionyl lysine. Digestion was performed with trypsin in order to reproducibly and efficiently produce a limited set of derived peptides. Since all lysines become occupied, trypsin cannot cleave at these residues. After the chemical reaction with proprionic anhydride, only arginine cleavage

on the C-terminal end can occur. The chemical reaction between propionic anhydride and lysine is shown below in Figure 2.1:

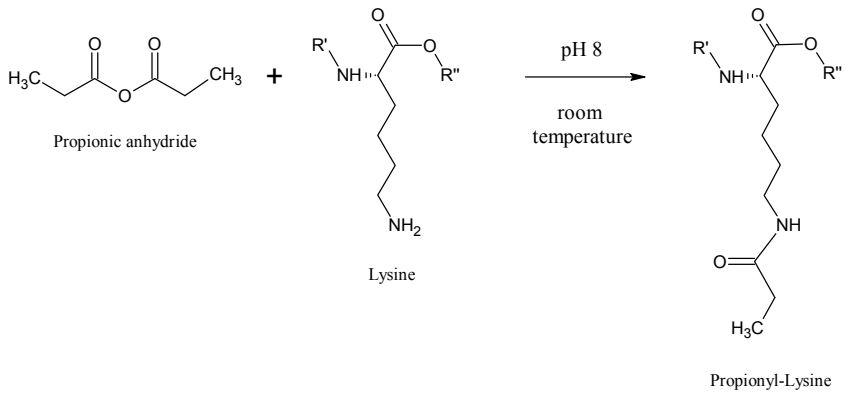


Figure 2.1: Chemical derivatization of lysine residues using propionic anhydride, resulting in the generation of a propionyl lysine.

2.2 Biological Models

For each individual chapter reported in this thesis, a different cell line was used. The primary reasons for the usage of different types of cells were availability and scope of the research. In the development of our histone analysis workflow, we validated our two-pronged strategy using histones extracted from yeast. The use of this organism had several advantages, including a relatively small genome, easy manipulation of its cell cycle, generation of mutant strains, and capability of growing large volumes of cell culture [2]. Moreover, we used mutant strains deficient in acetyltransferase activity that are known to modify a specific residue (i.e., H3K56). Hence, we could easily predict and validate our findings on structure-function relationships without too many biological changes. A more complicated biological model comprising of many differentially expressed PTMs could have impeded the validation process, and led to confounding and complicated results. In chapter four of this thesis, the histones that we investigated by targeted EMC and MRM were derived from DT40 cells. This cell line was derived from an avian leukosis virus (AVL)-induced bursal lymphoma [3]; features of this cell line include facile genetic manipulation and production of gene knockouts, and rapid generation times. This enabled the preparation of a single methyltransferase mutant

(DOT1L) in DT40 cells. Since this enzyme is known to methylate primarily one residue (H3K79), only a limited set of predictable peptides are differentially expressed between WT and mutant strains; this facilitated the demonstration of the novel targeted EMC scan mode, as well as MRM for relative peptide quantitation. In chapter five of this thesis, we used histones derived from K562 cells. These cells originated from human erythroid cells from a patient afflicted with chronic myelogenous leukemia (CML). These cells not only proliferate rapidly (12 hour doubling time) [4], but can be induced for erythroid differentiation when treated with nucleotide containing compounds [5]. As a model of transformed human cells, we selected the erythroleukemic cell line K562 for the following reasons. Several research groups at the Institute for Research in Immunology and Cancer (IRIC) and the "Centre de Recherche de l'Hôpital Maisonneuve-Rosemont" (HMR, a research institute with a close affiliation with IRIC) focus their studies on hematological cancers, with a particular emphasis on leukemia. We therefore hoped that our studies of the effects of HDAC inhibitors on K562 cells might foster collaborations with other research teams at IRIC and HMR. K562 cells are well defined genetically. As is the case in most patients suffering from CML, the K562 cell line carries a specific chromosome translocation. The translocation generates an abnormal chromosome, known as the Philadelphia chromosome [6], which encodes the key oncogenic protein BCR-ABL that drives the development of CML. The BCR-ABL fusion protein is a dysregulated (hyperactive) tyrosine kinase that is essential for the progression of CML; no other known genetic alteration appears to be involved in triggering CML. Because of this, patients suffering from leukemias driven by the BCR-ABL fusion protein (not only CML patients, but also a fraction of patients with acute lymphoblastic leukemia and acute myelogenous leukemia) are often treated with imatinib [7], a pharmacological agent that inhibits the kinase activity of BCR-ABL. However, about 30 to 40% of CML patients express an isoform of the BCR-ABL kinase that is poorly inhibited by imatinib because of a mutation in the kinase domain [8]. Hence, other chemotherapeutic agents, possibly in combination with imatinib derivatives, are needed to treat these patients. Because SAHA and other HDAC inhibitors seem to be most efficacious against hematological cancers, it seemed appropriate to study the extent to which HDACi affect histone acetylation in K562 cells. These cells are malignant and represent an ideal *in vitro* system for studying

the effects of small molecule HDAC inhibitors on the expression of histone modifications in cancerous cells. Histones from normal cells isolated from human lung fibroblasts [9] (e.g. WI-38, IMR90) were used as a control, and represent a healthy cellular state. Unlike K562 cells, IMR90 and WI-38 fibroblasts are not transformed or immortalized. Therefore, they have a short finite lifespan, which is empirically defined as the maximal number of passages for which the cells can be propagated in culture before they cease to proliferate (typically between 20 and 50 passages depending on culture conditions). These diploid cells proliferate more slowly than K562 cells, are closely indicative of cell biology *in vivo* [10], and can enter a replicative senescent state after a limited number of growth cycles [11]. IMR90 and WI-38 fibroblasts have been extensively studied to define the molecular mechanisms that trigger oncogene- and DNA damage-induced senescence. Since one of the mechanisms by which HDAC inhibitors have been reported to exert their anti-proliferative effect on cancer cells is the induction of DNA damage leading to either apoptosis or senescence, we felt that IMR90 and WI-38 fibroblasts might be particularly sensitive to increases in histone acetylation triggered by HDACi.

2.3 Nano scale HPLC

To maximize sensitivity using MS detection, nano scale separations were utilized. One dimensional reversed phase chromatography was applied to perform either intact histone or tryptic peptide separation prior to MS detection. The separation module used was either an Agilent 1100 series, a Waters Acquity, or an Eksigent 2D-nano LC system, all capable of generating low and stable flow rates in the nL/min range. Such low flow rates allow peptide sensitivity to reach low femto- and even attomole levels. Both a loading and an analytical column were connected to the liquid chromatograph. Column dimensions were 100 mm x 150 μm I.D. for the analytical column, and 4 mm x 360 μm I.D. for the trapping column. Both columns were packed in-house using fused silica and Teflon tubing; the columns were packed with C_{18} stationary phase. The use of a loading column enabled the loading of more material for MS analysis, and also provided an on-line cleanup of the sample prior to the analytical separation. Both intact histones and tryptic peptides were analyzed using the same HPLC method. Intact histones were

diluted with LC-MS mobile phase, while tryptic peptides were directly injected onto the HPLC system. The solvent system consisted of water and acetonitrile, both containing dilute formic acid (0.2% v/v). After a three minute loading step at 10 μL/min, the loading column was brought in-line with the analytical column, and gradient elution at 600 nL/min initiated to elute and separate histones.

2.4 MS detection and sequencing

In this Ph.D. project, a variety of MS instrumentation was used to analyze and sequence histones and their PTMs. The theory behind these instruments was covered in the previous chapter (section 1.5.2). More details about the experimental aspects used in each chapter are described below. All instrumental parameters are fully described in the following chapters.

2.4.1 Agilent TOF, Waters Q-TOF, and AB Sciex 4000 Q-trap for differential expression profiling and peptide sequencing (chapter three)

In our dual, two-pronged approach, both intact histones and tryptic peptides are analyzed. Intact histone LC-MS analysis was performed on a Waters Q-TOF Premier mass spectrometer. The instrument was operated in positive ion mode, while scanning a mass range between *m/z* 400 to 1600, with a mass resolution of 10 000. Application of a Q-TOF mass analyzer allowed the many different core histone isoforms to be resolved. Deconvolution of mass spectra by mathematical algorithms (MaxEnt) generated a mass profile for direct comparison between wild-type and mutant yeast strains, and selection of primary histone protein candidates for peptide analysis. Tryptic peptides for global expression profiling were analyzed using an Agilent TOF instrument, using the same parameters as above. The TOF analyzer provided a good linear range (three orders of magnitude), with accurate peptide mass determination. To sequence peptides and locate PTMs, the same extracts were re-injected onto the Waters Q-TOF premier using LC-MS/MS in data-dependant acquisition mode. Tandem MS was performed on the three most intense ions from the MS survey scan. Precursor ion selection was performed in the

quadrupole mass analyzer, while scanning product ions in the TOF analyzer. This allowed for high resolution fragment ion analysis as well. The stoichiometry of one peptide detected (H3K56ac) was further assessed using the AB Sciex 4000 Q-trap. Using data acquired in the global LC-MS/MS analysis, an MRM method was built using precursor and product ions generated on the Q-TOF instrument. A combination of three MRM peptide ion pairs was monitored, using a dwell time of 50 milliseconds. When the precursor and product ions are known, this allows for the creation of a highly sensitive and precise quantitation method in MRM mode using a triple quadrupole instrument. The stoichiometry of H3K56ac in wild type and mutant yeast strains was measured with high accuracy using MRM.

2.4.2 AB Sciex 4000 Q-trap for targeted EMC and differential PTM profiling (chapter four)

The 4000 Q-trap mass spectrometer was used to detect and profile intact proteins. The enhanced multiply charged (or EMC) scan performed in the linear ion trap (LIT) allows for selectively evacuating singly charged ions while trapping multiply charged ions. This was initially designed by AB Sciex to enhance the sensitivity and detection of peptides. The MCS voltage is the trapping voltage applied to the exit lens of the LIT to contain multiply charged species while allowing singly charged ions to escape. The MCS voltage was optimized using infusion at 0.8 V. Two model analytes were used for this optimization: apomyoglobin and leucine enkephalin. However, the EMC scan failed to perform adequately in LC-MS mode due to space-charging effects. With the aid of AB Sciex, we modified the conventional EMC mode by changing the RF applied to the protein ions, gating all the ions through Q0, and selectively transmitting only a limited *m/z* range into the LIT in a sequential fashion. This way, the LIT was not overloaded with too many ions. After trapping the ions, the charge separation step was applied to remove excess singly charged ions from the LIT. Using BioAnalyst software tools, the mass profiles could be generated and compared. This modified scan mode (targeted EMC) allowed for not only higher sensitivity and resolution, but also minimized space charge effects. Histones derived from DT40 wild type and mutant cells were examined. As

expected, differences in the histone H3 intact profiles were detected between wild type and DOT1L mutant. The abundance and localization of differentially expressed histone PTMs were monitored by MRM on the same instrument. Precursor and product ion pairs were generated by importing histone H3 sequences into AB Sciex MRM pilot software, and generating theoretical MS/MS spectra. MRM ion pairs were designed based on the most intense fragments to maximize the MS response. Three to four ion transitions were monitored for each peptide using a rapid dwell time (10 milliseconds) without compromising sensitivity. To confirm the sequence of the peptides detected by MRM, MS/MS in the LIT was performed in a data-dependant fashion using enhanced product ion (EPI) scan when a peptide peak was detected.

2.4.3 Agilent Q-TOF, Thermo LTQ-Orbitrap XL, and AB Sciex 4000 Q-trap for differential PTM expression and targeted MRM (chapter five)

In the HDACi exposure study, we used three different instrument platforms to profile intact histones and tryptic peptides. The Agilent Q-TOF instrument was used to profile intact histones from control and HDACi treated cells using the same parameters as in section 2.3.1. For global LC-MS/MS analysis of histone peptides, the Thermo LTQ-Orbitrap XL was used for both expression profiling and MS/MS sequencing. This way, both types of information can be acquired from the same injection. Precursor ions were detected in the Orbitrap mass analyzer. The instrument was operated in positive ion mode, while scanning a mass range between *m/z* 300 to 2000, with an Orbitrap mass resolution of 60 000. The higher mass resolution allowed for better discrimination of near isobaric species. The three most intense ions detected in the Orbitrap survey scan were selected for tandem MS and transferred to the LTQ ion trap. A decision tree was designed to select ions for either CID or ETD fragmentation, based on *m/z* and charge state. Lower charge state peptide ions were fragmented by CID, while higher charge state by ETD. MS/MS sequencing was performed in the LTQ ion trap by either fragmentation mechanism. Low abundance peptides (i.e. H3K56ac) were detected and sequenced by both the AB Sciex 4000 Q-trap and the Thermo LTQ-Orbitrap XL. An MRM method based on precursor-product ion pairs was built to screen for H3K56ac in different cell

types. A high resolution absolute quantitation method was designed on the Thermo LTQ-Orbitrap XL to detect and sequence H3K56ac using only CID. A calibration curve for H3K56ac was prepared and each individual standard was analyzed on the Thermo LTQ-Orbitrap XL to determine the absolute concentration of this peptide in histones derived from various biological sources.

2.5 Bioinformatic tools and database searching

2.5.1 Peptide map generation

Raw data acquired from complex proteomic analyses contain many peptides, contaminants, singly charged species, etc. Manual inspection of the raw data for differential expression profiling would be too laborious and time consuming. Software tools were developed in-house to expedite the analysis of raw data. The software tools utilized were Mass Sense peptide miner and clusterer. The peptide miner detection software converts LC-MS/MS raw data from high resolution acquisition into a two-dimensional peptide maps (see Figure 2.2). Peptides are detected and highlighted graphically based on *m/z* value, LC retention time, and charge state.

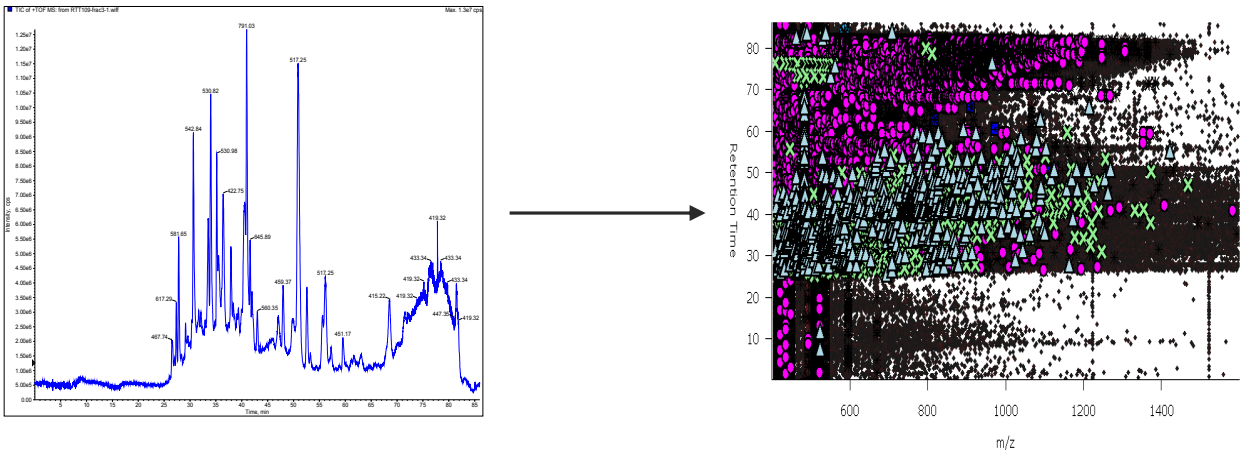


Figure 2.2: Raw LC-MS total ion chromatogram (left) is converted into a two-dimensional peptide map by Mass Sense peptide miner detection software (right).

2.5.2 Clustering analysis

Peptide maps generated from replicate injections of different experimental conditions (e.g. control vs. mutant) were imported into the clusterer software. Peptide maps across multiple LC-MS runs were aligned with each other, and detected ions clustered together based on pre-defined *m/z* and LC retention time tolerances (see Figure 2.3). For example, using high resolution LTQ-Orbitrap data, the *m/z* tolerance is usually set to 0.02 Th, and 1 minute LC retention time variation. This segmentation analysis of all ions detected allows for the rapid comparison of abundance values across different experimental conditions, and greatly reduces the time required to examine the raw data. The results of this clustering analysis can be exported into a Microsoft Excel spreadsheet for further analysis and generation of graphs (e.g. scatter and volcano plots).

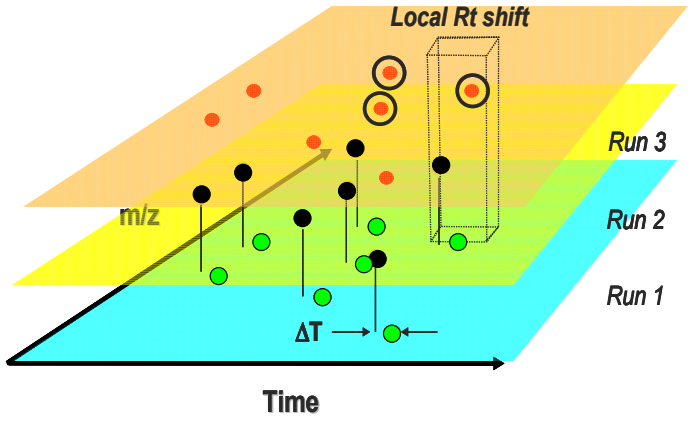


Figure 2.3: Two-dimensional peptide maps generated by Mass Sense peptide miner overlaid together from triplicate LC-MS runs. Alignment of peptide clusters across triplicate runs using Mass Sense peptide clusterer. Peptide ions of different *m/z* values are represented by different colors (i.e., red, green, black circles). Peptide ions are grouped together based on *m/z*, LC retention time, and charge state, with pre-defined tolerances.

2.5.3 Database searching

Raw data files from global LC-MS/MS were uploaded to the Mascot server (www.matrixscience.com). This is a web based search engine that allows rapid assignment of ion spectra and peptide sequences from vast amounts of acquired MS/MS data. Search parameters such as type of enzyme, number of missed cleavages, peptide

charge state, modifications and database were pre-defined prior to launching a Mascot search. For histone peptide MS/MS spectra analysis, the enzyme used was trypsin, and due to the chemical reaction with propionic anhydride, the maximum allowed missed cleavages was selected, i.e., nine. This high number was used in order to not discard any peptides bearing a long amino acid sequence. Modifications included acetylation, propionylation and methylation of lysine, methylation of arginine, phosphorylation of serine, threonine and tyrosine, and oxidation of methionine. Histone sequences from different species were obtained from the Swiss-Prot database (http://ca.expasy.org/sprot). Once the search is completed, the search engine reports all peptides found in a compiled list, along with their probability scores. Furthermore, the peptides identified by the Mascot search were linked together with the clustered data generated by Mass Sense. At the end of this process, all vital information including peptide *m/z*, LC retention time, abundance, and Mascot MS/MS identification are all grouped together in one Excel file. The processing of LC-MS/MS raw data is summarized in Figure 2.4.

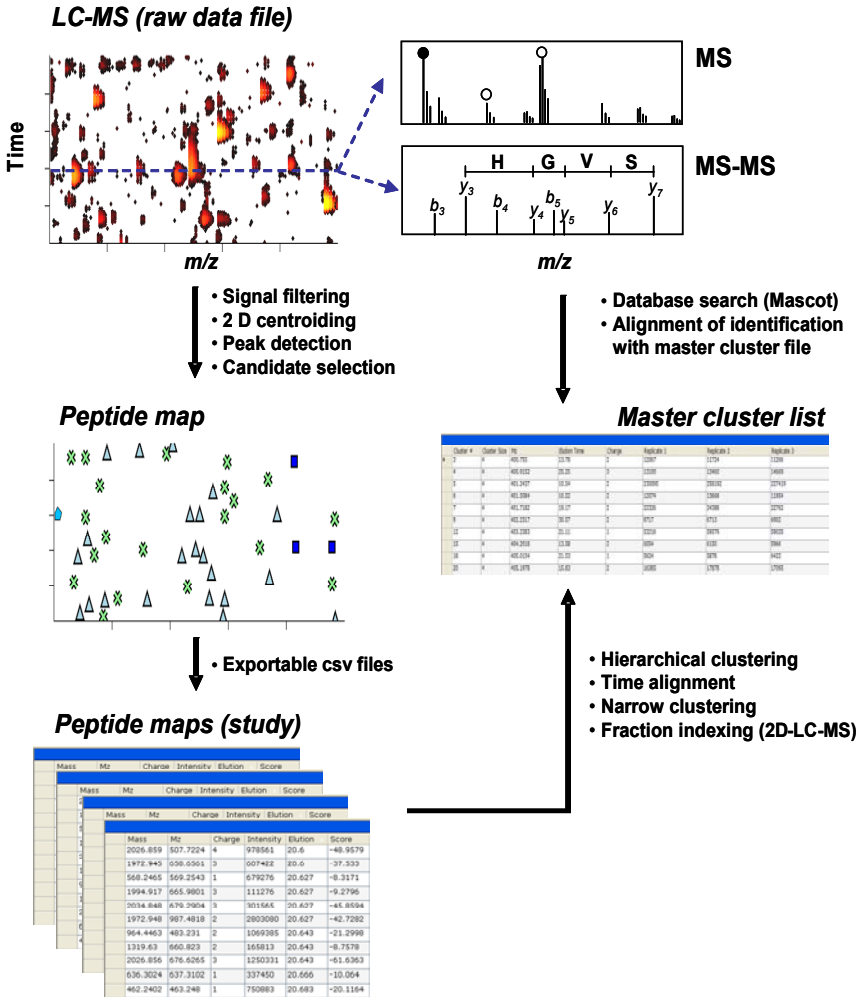


Figure 2.4: A visual flowchart of the treatment of raw LC-MS/MS data using bioinformatic tools and Mascot database searching. From reference [12].

2.6 References

1. Garcia, B.A., et al., *Chemical derivatization of histones for facilitated analysis by mass spectrometry.* Nat Protoc, 2007. **2**(4): p. 933-8.

2. Brandt, W.F., K. Patterson, and C. von Holt, *The histones of yeast. The isolation and partial structure of the core histones.* Eur J Biochem, 1980. **110**(1): p. 67-76.

3. Winding, P. and M.W. Berchtold, *The chicken B cell line DT40: a novel tool for gene disruption experiments.* J Immunol Methods, 2001. **249**(1-2): p. 1-16.

4. Koeffler, H.P. and D.W. Golde, *Human myeloid leukemia cell lines: a review.* Blood, 1980. **56**(3): p. 344-50.

5. Osti, F., et al., *Human leukemia K562 cells: induction to erythroid differentiation by guanine, guanosine and guanine nucleotides.* Haematologica, 1997. **82**(4): p. 395-401.

6. Cilloni, D. and G. Saglio, *CML: a model for targeted therapy.* Best Pract Res Clin Haematol, 2009. **22**(3): p. 285-94.

7. Breccia, M., F. Efficace, and G. Alimena, *Imatinib treatment in chronic myelogenous leukemia: What have we learned so far?* Cancer Lett, 2011. **300**(2): p. 115-21.

8. Stein, B. and B.D. Smith, *Treatment options for patients with chronic myeloid leukemia who are resistant to or unable to tolerate imatinib.* Clin Ther, 2010. **32**(5): p. 804-20.

9. Norwood, T.H. and W.R. Pendergrass, *The cultured diploid fibroblast as a model for the study of cellular aging.* Crit Rev Oral Biol Med, 1992. **3**(4): p. 353-70.

10. Hayflick, L. and P.S. Moorhead, *The serial cultivation of human diploid cell strains.* Exp Cell Res, 1961. **25**: p. 585-621.

11. Place, R.F., E.J. Noonan, and C. Giardina, *HDACs and the senescent phenotype of WI-38 cells.* BMC Cell Biol, 2005. **6**: p. 37.

12. Jaitly, G., et al., *Comprehensive profiling of peptides and proteins from large-scale proteomics experiments using one and two dimensional nanoLC-MS/MS.* (manuscript in preparation for Journal of Proteome Research), 2011.

3. Comprehensive profiling of histone modifications using a label-free approach and its applications in determining structure-function relationships

Paul Drogaris, Hugo Wurtele, Hiroshi Masumoto, Alain Verreault, and Pierre Thibault

Analytical Chemistry, 2008, Sep 1; 80 (17):6698-6707

3.1 Abstract

A two-pronged approach using specialized peptide detection and clustering tools was developed to profile changes in histone post-translational modifications (PTMs). The extent and nature of modification was inferred by comparing the mass profiles of intact core histones from nano LC-MS experiments. Histones displaying changes in their intact mass profiles were fractionated, derivatized with propionic anhydride, and digested with trypsin prior to nano LC-MS analyses. Our methodology was validated by comparing the abundance of histone PTMs in wild type (WT) and mutant strains of *Saccharomyces cerevisiae* lacking the histone acetyltransferase Rtt109 and a nucleosome assembly factor known as Asf1. Both Rtt109 and Asf1 were previously found to be essential for acetylation of histone H3 lysine 56 (H3K56ac), a modification that plays an important role in the response to genotoxic agents that interfere with DNA replication. The generation of ion abundance distribution plots enabled a rapid and comprehensive profiling of histone peptides. Our analytical methodology and data mining approach showed that most common histone PTMs were unaffected in mutant yeast cells lacking Rtt109 and Asf1. However, a subpopulation representing 17% of all H3 histones in wild type cells showed an acetylated K56 residue that was significantly reduced in both mutant strains. Our generic strategy for histone PTM profiling can be applied to assess global changes in histone PTMs across sample sets and to establish structure-function relationships.

3.2 Introduction

In eukaryotic cells, long DNA chains are compacted within the cell nucleus through their association with octamers of highly basic histone proteins. This generates a complex and dynamic nucleoprotein structure known as chromatin. The basic repeating subunit of chromatin is the nucleosome core particle (NCP) [1], which consists of 147 base pairs (bp) of DNA wrapped nearly twice around a histone octamer. This octamer is composed of two molecules each of four types of core histones (H2A, H2B, H3, and H4). Adjacent NCPs are connected by variable lengths of linker DNA (between 18 and 65 bp), and histone H1 family members bind to the linker DNA to promote folding of chromatin into a higher-order structure known as the 30-nm fibre [2]. The globular domains of the four core histones, known as histone-fold domains, are structurally related to each other and responsible for the majority of histone-histone and histone-DNA contacts in NCPs [1]. In contrast, the N-terminal "tails" are solvent accessible and protrude beyond the DNA gyres of NCPs. Several residues within the histone N-terminal tails are sites of post-translational modifications (PTMs), and recent advances in mass spectrometry (MS) have also uncovered a bewildering diversity of PTMs within the globular domains of core histones [3, 4]. Histone PTMs can affect chromatin structure directly by modifying the affinity of histones for specific segments of nucleosomal DNA. In addition, histone PTMs can act indirectly, by serving as docking platforms for either nonhistone chromosomal proteins, histone modifying enzymes or ATP-dependent chromatin remodeling machines [5]. Histone PTMs participate in many DNA metabolic processes, such as nucleosome assembly, transcription, DNA repair and chromosome segregation. In addition, some histone PTMs function in a heritable epigenetic "histone code" that mediates stable transcriptional silencing in proliferating cells [6]. Histone PTMs often silence the expression of tumour suppressor genes [7]. Thus, many inhibitors of histone modifying enzymes are under close scrutiny as potentially novel cancer chemotherapeutic agents [8]. Hence, the identification and accurate quantitation of histone PTMs is crucial in understanding their biological functions and assessing the effects of novel pharmacological agents that target histone-modifying enzymes.

MS has been successfully applied to study histone PTMs using both bottom-up [9, 10] and top-down sequencing strategies [11, 12]. MS is ideal for PTM characterization because of its high sensitivity, specificity, and ability to pinpoint exact modification sites. Even complex patterns of PTMs can be determined in either proteolytic digests or intact proteins. However, the diversity, low stoichiometry, and short *in vivo* half-life of some histone PTMs still represent a formidable challenge from an analytical MS perspective. At least eight different types of PTMs have been reported in histones, including acetylation, methylation of lysine (mono-, di- and trimethylation) and arginine residues, phosphorylation, as well as larger PTMs such as sumoylation, ubiquitylation and ADP-ribosylation [5]. Over 60 modification sites have been detected using either antibodies or MS [5]. Further complexity is imparted by the existence core histone sequence variants. These variants perform unique biological functions and often do not exhibit the same PTMs as canonical core histones [13].

The ability to detect and quantify the differential distribution of histone PTMs makes MS an important tool not only in chromatin research, but also in the discovery of novel pharmacological agents that inhibit histone modifying enzymes such as histone deacetylases (HDACs) [14, 15]. A few research groups have addressed the challenge of quantifying histone PTMs using MS. One approach for rapid profiling of global changes in PTMs of intact histones has been reported in acute myeloid leukemia cell lines treated with HDAC inhibitors (HDACi) [16]. However, the identification of modification sites cannot be achieved using histone mass profiles alone due to the natural sample heterogeneity, the presence of confounding peaks, and overlapping PTMs. Recent advances in high resolution Fourier Transform mass spectrometry using emergent "top-down" approaches enabled the direct sequencing of intact histones and their variants by tandem MS, albeit larger amounts of fractionated nuclear extracts are typically required than with other MS approaches [11, 12]. Enzymatic digestion of histones using endoprotease Glu-C to generate relatively large *N*-terminus peptides was also reported to identify sites of modifications using electron capture dissociation [17] or electron-transfer dissociation [18]. Alternative strategies for "bottom-up" sequencing, using either label-free approaches or stable isotope labeling [19-23] have been investigated for comparing

differential distribution of histone PTMs. A recent publication, exploiting LC-MS/MS and data mining software, demonstrated the use of label-free ion profiling for monitoring changes in histone PTMs in response to HDACi treatment [24]. Hunt and colleagues reported the development of a "bottom-up" MS sequencing platform [25] utilizing chemical derivatization and stable isotope labeling to profile the PTMs of histone H3 in mammalian cells [26]. The widespread use and applicability of the aforementioned approaches is limited by the requirement of either growing cells in stable isotope media to promote metabolic labeling of histones, or using multiple chemical derivatization and labelling steps with stable isotopes, extensive sample preparation, and manual inspection of raw data from LC-MS/MS runs. Furthermore, very few data mining approaches exist for in-depth raw data analysis, and most published methods do not demonstrate any analytical figures of merit for reproducibility or their ability to assess the extent of modifications.

In this report, we present an approach that addresses the inherent difficulties in profiling the relative abundance of PTMs across different histone sample sets. We describe a two-pronged strategy using both intact histones and tryptic peptides. First, intact histones are screened for changes in their mass profiles that could be attributed to PTMs. Second, peptide maps generated from LC-MS analysis of the corresponding tryptic digests are compared to identify peptides showing statistically meaningful changes in abundance across conditions and replicate runs. This is facilitated using scatter and volcano plots that readily reveal outlier histone peptides differentially abundant across sample sets. Both data-dependant and targeted nano LC-MS/MS analysis are used to sequence these outlier peptides and locate their PTMs. Using this approach, less time is spent on manually inspecting all the raw data, and the differential abundance of peptides is quantified in a reproducible and unbiased manner. The validity of this approach will be illustrated by expression profiling of PTMs in histones derived from wild type (WT) cells and mutants lacking Rtt109 (*rtt109*Δ cells), and Asf1 (*asf1*Δ cells). Rtt109 is required for the acetylation of newly synthesized H3 histones during the S-phase of the cell cycle, while Asf1 is a histone chaperone that presents new histones as substrates for Rtt109-catalyzed acetylation in *Saccharomyces cerevisiae* [27].

3.3 Experimental section

3.3.1 Reagents

Nonidet P40 (NP-40) and zymolyase were obtained from Fluka, and hydrochloric acid for histone extraction from Sigma. Total protein was measured using the micro-BCA assay (Pierce). Microbore HPLC C_{18} columns for intact histone fractionation were purchased from Phenomenex; capillary HPLC columns for LC-MS were packed in-house using Jupiter C_{18} (3 μm) particles from Phenomenex, and fused-silica tubing (Polymicro Technologies). Trapping columns were packed in-house using the same bulk material in Teflon tubing (Supelco). Solvents for chromatographic analysis were all HPLC grade (Fisher Scientific and in-house Milli-Q water). Trifluoroacetic acid (TFA), anhydrous methanol, ammonium bicarbonate, ammonium hydroxide, D_6-acetic anhydride and propionic anhydride were all purchased from Sigma. Formic acid was obtained from EMD Science. Porcine modified trypsin (sequencing grade) was obtained from Promega.

3.3.2 Yeast Biological Model

Yeast strains were constructed and grown in-house. The haploid wild-type *S. cerevisiae* strain (MAT**a** *his3Δ1 lys2Δ0 leu2Δ0 ura3Δ0*), and the *asf1*Δ::KanMX and *rtt109*Δ:KanMX mutants were derived from the BY4743 heterozygous diploid collection [28]. The *hst*3Δ *hst*4Δ double mutant strain was previously described [29].

Recombinant yeast histone H3 was expressed and purified from *Escherichia coli* as previously described [30], except that the buffers contained three HDAC inhibitors of different specificity: 100 mM sodium butyrate, 1 μM trichostatin A and 10 mM nicotinamide. The wild-type and mutant strains were treated with zymolyase to digest the cell wall and generate spheroplasts, which were lysed in a buffer containing 0.5 % NP-40, and histones were extracted with 0.25 M hydrochloric acid.

3.3.3 Sample Preparation and Protein Digestion

After acid extraction, the total protein content in the yeast histone-enriched samples was measured with the micro-BCA assay, using BSA as the reference standard. Intact core histones (~15 μg) were separated using an Agilent 1100 HPLC system equipped with a microfraction collector. Separations were performed on a microbore Jupiter C_{18} column (3 μm, 300 Å), 150 x 1 mm i.d., with a solvent system consisting of 0.1 % TFA in water (v/v) (solvent A), and 0.1% TFA in acetonitrile (v/v) (solvent B). Gradient elution was performed from 5 to 90% B in 60 min at 35 μL/min. Fractions were collected in a 96-well plate at 1-min intervals. Individual histone peaks that eluted over multiple fractions were pooled together. Histone fractions were dried in a Speed-Vac evaporator and resuspended in 0.1 M ammonium bicarbonate buffer (without pH adjustment). This solution of intact histones was derivatized in 1:1 volume ratio with a propionic anhydride (PA) reagent solution (deionized water/propionic anhydride, 3:1 (v/v)). The reaction mixture was maintained at a basic pH by the addition of concentrated ammonium hydroxide to each tube. After a 30 minute derivatization period at room temperature, samples were evaporated a second time, resuspended in 0.1 M ammonium bicarbonate, and digested overnight at 37 °C using 1 μg of trypsin. Samples were acidified with 5% TFA in water (v/v) prior to LC-MS analysis. For targeted analysis of H3 tryptic peptides, propionic anhydride was substituted with D_6-acetic anhydride; derivatization and digestion were performed using the same protocol as outlined above.

3.3.4 Nano LC-MS of Analysis of Intact Histones

Intact histone analysis was performed on a Waters Q-TOF Premier mass spectrometer equipped with a nanoelectrospray interface. The instrument was operated in positive ion mode and scanned from *m/z* 400 to 1600. Whole protein samples were diluted to 200 ng/μL using the initial LC-MS mobile phase (see below). An aliquot was loaded onto a capillary C_{18} trapping column (4 mm x 360 μm I.D.) for 3 minutes at 10 μL/min using a 90:10 water/acetonitrile solution containing 0.2 % formic acid (v/v). Trapped histones were eluted onto the analytical column (100 mm x 150 μm i.d.) using a

solvent system consisting of 0.2% formic acid in water (v/v) (solvent A), and 0.2% formic acid in acetonitrile (v/v) (solvent B). Gradient elution was performed from 10 to 90% B in 60 min at 600 nL/min. Intact histone mass profiles were generated from acquired mass spectra using the MaxEnt1 deconvolution algorithm in the MassLynx software.

3.3.5 Nano LC-MS Analysis of Tryptic Peptides

Tryptic digests from histone peptide fractions were injected in triplicate for ion profiling on either an Agilent MSD TOF or a Waters Q-TOF Premier mass spectrometer equipped with a nanoelectrospray interface. The instrument was operated in positive ion mode and scanned from *m/z* 400 to 1600, using the same HPLC columns and solvents as for the intact histone analysis. Gradient elution was performed from 0 to 50% B in 60 min at 600 nL/min. Targeted nano LC-MS/MS analyses of H3 tryptic peptides using multiple reaction monitoring (MRM) were performed on an AB/Sciex API 4000 Q-Trap mass spectrometer equipped with a Nanospray II interface. A 50 ms dwell time was applied to each unique histone peptide ion MRM transition; tryptic peptides were eluted from the HPLC column using the same conditions specified above.

3.3.6 Ion Abundance Profiling and Peptide Sequencing

All raw data files (.wiff) from triplicate injections on the MSD TOF were converted to 2-D peptide maps using the Mass Sense [31] peptide detection software (version 2.5) to identify peptides based on *m/z* values, retention time, abundance, and charge state. An intensity threshold of 100 counts (or 5 counts for Q-TOF analyses) was set as a cutoff for peak detection. Segmentation analysis was performed by clustering peptide ions across each condition based on their respective *m/z* ratio, charge, and LC retention time with Peptide Clusterer software (version 2.5) using *m/z* tolerance of ± 0.05 Th and a time difference of 1 min. Normalization of retention time is then performed on the initial peptide cluster list using a dynamic and nonlinear correction. A moving-average, time window interpolation scheme is used to compute the time shifts for each

peptide across the different data sets. This alignment confines the retention time distribution to less than ±0.1 min (< 0.3% RSD) on average. Clustered data sets were exported to Microsoft Excel to generate either scatter or volcano ion distribution plots. Sequencing of histone peptides was performed using the Waters Q-TOF Premier mass spectrometer equipped with a nanoelectrospray interface. The instrument was operated in positive ion mode and scanned from *m/z* 400 to 1600 using MS/MS in data dependant mode. The same HPLC conditions were used as for the ion abundance profiling. Database searches were performed against the SWISSPROT database using Mascot. Histone peptides with Mascot scores greater than 30 were selected and manually verified to confirm their sequence assignment. Expression data and Mascot searches were aligned together using the Peptide Clusterer software.

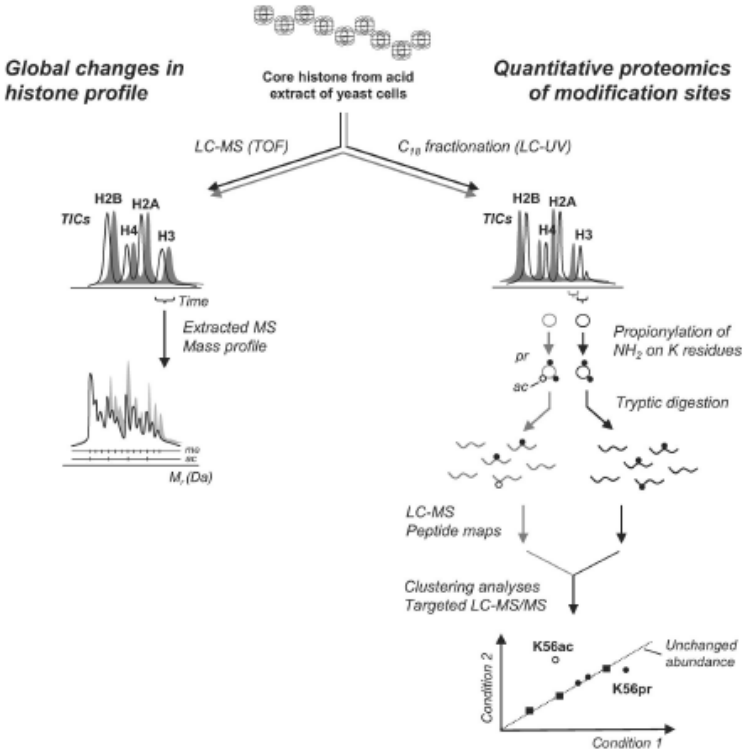


Figure 3.1: Overview of the two-pronged analytical strategy combining protein identification and data mining approaches

3.4 Results and Discussion

3.4.1 Nano LC-MS of Intact Histones

Our two-pronged strategy is summarized in figure 3.1. As a first step, intact histones were screened to monitor any changes in their mass reconstruction profile. The total ion chromatograms from the LC-MS analysis of core histones is shown in figures 3.2a and b for acid extracts from WT and mutant strain *rtt109*Δ. Histones eluted between 25 and 45 min in the following order: H4, H2B, H2A, and H3 (figure 3.2a and b). The peak labelled 1 contains both H4 and H2B, while H2A and H3 eluted in peaks 2 and 3, respectively. For histones H2A and H4, the mass profiles of intact histones derived from WT, *asf1*Δ, and *rtt109*Δ cells showed no significant changes and were virtually superimposable (figure 3.2c). However, histones H2B and H3 showed significant changes. In *S. cerevisiae*, there are two forms of histone H2B that exhibit few differences at the amino acid sequence level, and are encoded by distinct genes: H2B.1 (encoded by *HTB1*) and H2B.2 (encoded by *HTB2*). In the H2B profiles, the unmodified H2B.1 protein is observed (14121 Da), along with two extra acetylations (14163 and 14205 Da, respectively). The 14121 Da component corresponds exactly to the mass predicted for removal of the N-terminal methionine without N-terminal amino group acetylation. Consistent with this, previous studies had suggested that a fraction of H2B molecules are not acetylated on their N-terminal amino group, unlike H2A and H4 [32]. When comparing intact H2B.1 mass profiles derived from WT and *asf1*Δ cells, there is a distinct change in the peak abundance for all three peaks (figure 3.2c). However, this was not the case when WT histones were compared to those derived from *rtt109*Δ cells; the intact H2B.1 mass profiles were superimposable (figure 3.2c). Based only on the intact mass profile, this observation suggests that a higher proportion of H2B.1 is observed in WT cells compared to that of *asf1*Δ mutant cells. A second set of peaks with +42 Da differences were also observed, with a +16 Da offset from the first set of H2B.1 peaks. This second set of mirror peaks was rationalized as being chemical artifacts from the oxidation of Met 62 during the sample preparation process.

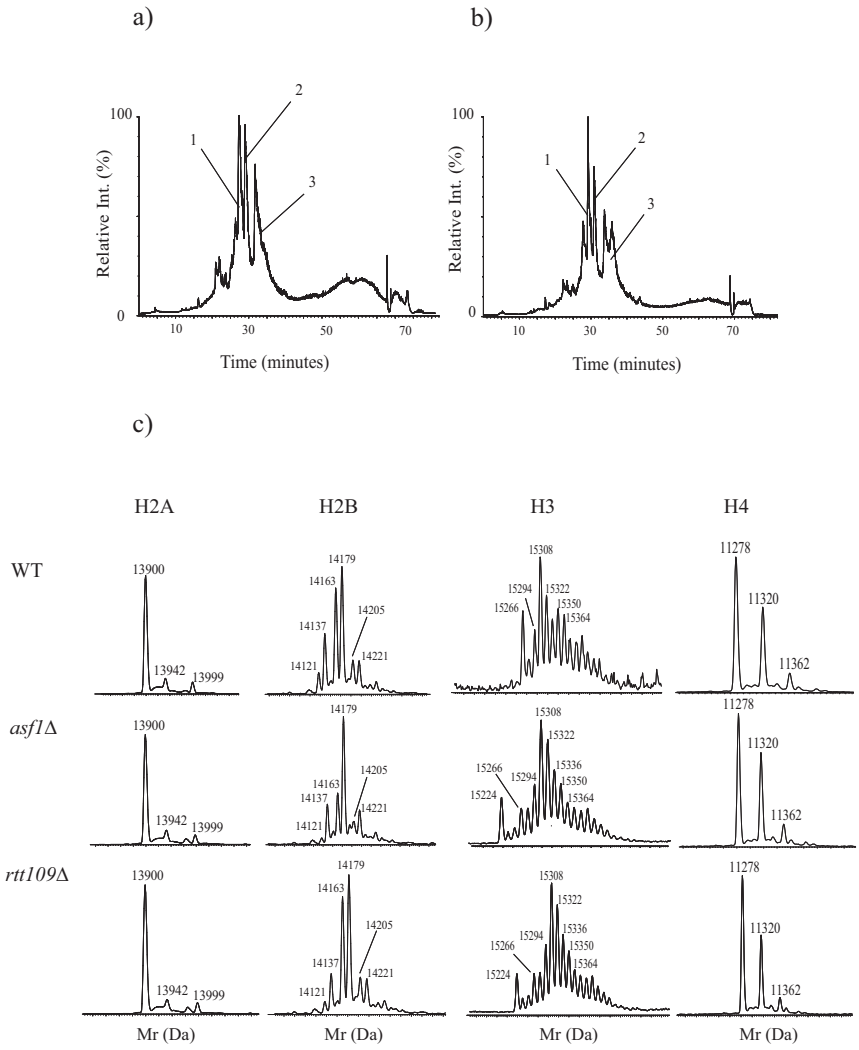


Figure 3.2: Nano LC-MS total ion chromatograms of intact core histones from (a) WT and (b) mutant yeast strain *rtt109*Δ. Histones eluted within the peaks labelled 1-3. (c) Intact histone mass profiles were generated for the four core histones (H2A, H2B, H3, and H4) derived from each yeast strain.

Another significant change was observed when comparing the chromatographic profiles of H3 derived from WT cells with those obtained from the two yeast mutant strains. A tailing shoulder was apparent in the LC-MS chromatogram of H3 from WT cells (figure 3.2a), while in the *rtt109*Δ mutant this late-eluting peak was partly resolved (figure 3.2b). Mass reconstructions were performed for both the main peak and shoulder (figure 3.2a); histone H3 was only found in the main peak, while the later component was a nonhistone protein contaminant. Comparison of intact profiles of histone H3 derived

from WT cells and each mutant revealed a more complex distribution of variants spaced by 14 and 42 Da (figure 3.2c). Closer examination of the H3 profiles indicated a subpopulation with a peak at 15 266 Da in WT cells that was shifted to 15 224 Da in both *asf1*Δ and *rtt109*Δ mutant strains (figure 3.2c). The addition of 42 Da over the intact mass predicted for unmodified *S. cerevisiae* H3 (15 224 Da for H3 with the initiator methionine removed and the N-terminal amino group nonacetylated) suggested the presence of abundant acetylation in WT cells.

3.4.2 Sample Preparation and Reproducibility of Abundance Measurements

To determine the heterogeneity and site occupancy of modified histones, yeast cell extracts were fractionated into individual core histones by HPLC prior to chemical derivatization of free lysine residues using PA based essentially on previously reported protocols with a few noted changes [33, 34]. The propionylation reaction is required due to the high number of lysine and arginine residues present in histones, resulting in a large number of small tryptic peptides that are hydrophilic and therefore poorly retained by reversed-phase HPLC. Derivatization appends a propionyl group to both free and mono-methylated lysines, which cannot be cleaved by trypsin. Di- and trimethylated lysine residues are not propionylated, but are not cleaved by trypsin either [34]. Thus, *in vitro* propionylation restricts trypsin digestion to the C-terminal side of arginine residues. This results in larger tryptic fragments that are better retained on a reversed-phase HPLC column. Consequently, these hydrophobic peptides elute at higher acetonitrile concentrations, thereby increasing sensitivity under nanoelectrospray conditions.

The addition of a propionyl group provides a traceable mass shift (addition of +56 Da) that can be easily distinguished from lysine residues modified *in vivo*. In preliminary experiments, we compared a single round of propionylation to the method of Garcia et al. using multiple rounds of propionic anhydride in methanol [33]. These experiments revealed that a significant decrease in ion abundance was observed for tryptic peptides that had undergone multiple derivatization steps compared to the same peptides using only one round of propionylation. For example, using two rounds of

propionylation, the intensity of the H3K56ac peptide decreased by 48% with a CV of 26%; this value was slightly higher than that observed for the single propionylation (see below). Also, the propionylation reagent was prepared and delivered in water rather than methanol to minimize the methyl esterification of Glu and Asp residues [34] to less than 1%, in addition to a significant decrease in amidation products. The single propionylation procedure provided a more rapid and streamlined sample preparation by reducing the number of evaporation steps, while minimizing variability in the peptide abundance measurements. Furthermore, during our data analysis and manual validation of peptides, there was no evidence of underivatized peptide present in the samples prepared using a single round of propionylation. These comparative analyses prompted us to use a single round of propionylation to maintain good reproducibility while avoiding the generation of side reaction products and chemical artifacts. In order to verify that the propionylation and tryptic digestion produced reproducible and representative histone peptides, the variability of the entire sample preparation process, including the nano LC-MS analyses, was ascertained by collecting three replicate injections of yeast histones. Chemical derivatization and tryptic digestion were performed on the fractionated core histones. After digestion, each histone fraction was analyzed by nano LC-MS in triplicate, and peptide maps were generated for cross sample comparison. The results of the reproducibility tests for all detected ions contained within the histone H3 fraction are shown in figure 3.3. A total of 273 ion clusters with intensity values ranging over 2 orders of magnitude were identified in all replicates. The variability across three fractionated samples and triplicate analyses by nano LC-MS produced a CV $\leq$ 15% for all peptides detected. A direct linear relationship was observed when comparing individual and average intensities of all ions detected on the scatter plot, highlighting the reproducibility of the approach. Accordingly, changes in abundance measurements representing more than 25 % can be identified reliably with a degree of confidence necessary for expression profiling and MS/MS sequencing.

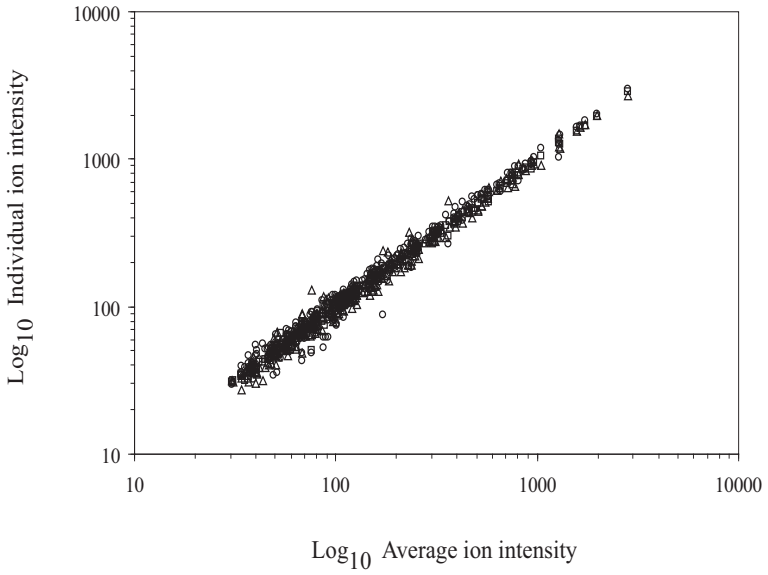


Figure 3.3: Scatter plot of reproducibility tests from sample preparation and nano LC-MS analysis using histone H3 fractions from three separate LC fractionation replicates. Circles, squares, and triangles correspond to the intensity of all ions detected in the first, second, and third fractionation replicates (same sample, fractionated three times), respectively.

3.4.3 Linearity and Dynamic Range of Quantitative Proteomics Approach for Histone Modifications

Recombinant histone H3 purified from *E. coli* is devoid of PTMs [35]. Hst3 and Hst4 are the deacetylases for H3K56ac in *S. cerevisiae* [29, 36, 37]. In *hst3*Δ *hst4*Δ yeast mutants, virtually all of histone H3 (98%) is K56-acetylated [29]. In order to evaluate system performance and dynamic range of peptide detection, histone H3 samples purified from either *E. coli* or *hst3*Δ *hst4*Δ mutant yeast cells were fractionated by HPLC to isolate comparable amounts of each form of histone H3 (~ 2 µg). After propionylation and tryptic digestion, the two digest samples were mixed in varying proportions ranging from 1 to 100 % (1:99, 5:95, 20:80, 50:50, 0:100) of H3 from *hst3*Δ *hst4*Δ mutant strains spiked with increasing amounts of unmodified histone H3 purified from *E. coli.* Each sample was prepared separately rather than by aliquoting a stock mixture to mimic variability in sample preparation. Triplicate injections were performed for each sample. A

total of 31 identified peptide clusters detected in at least two replicate injections were related to histone H3.

For the acetylated K56 peptide from histone H3, we next compared the calculated and observed abundance changes for all 10 combinations of spike amount ratio (fold-change). Results from these analyses are shown in figure 3.4, along with error bars defining maximum variations observed for individual ratios. Also reported on this plot is the error bar observed for peptide clusters unrelated to H3K56ac, which are present in constant levels among the different spike samples (e.g., VTIQKKDIKLAR, KSTGGKAPR, YKPGTVALR, KQLASKAAR). In figure 3.4, a linear relationship (r^2 = 0.995) was obtained between the observed and calculated abundance ratio measurements, with an average CV value of 9.8%. It is noteworthy that the observed abundance ratios were within 2.7% of their corresponding calculated values with larger variations noted for low spike levels showing higher variability of lower abundance ion signals. These experiments indicated that reliable change in abundance measurements can be monitored over two orders of magnitude in intensity ratio.

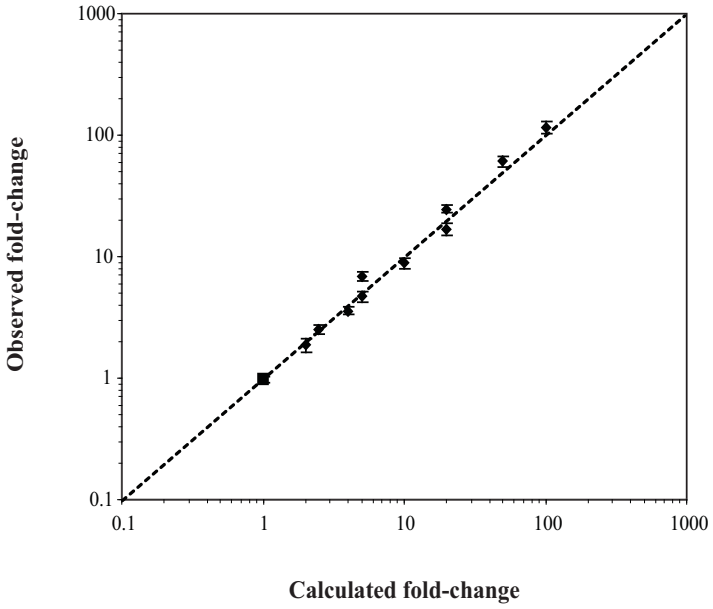


Figure 3.4: Ratio of observed and calculated intensity changes for histone H3 tryptic peptides from *hst3*Δ *hst4*Δ double mutant yeast cells spiked into a tryptic digest of recombinant histone H3 purified from *E. coli*. Spiking ratios were as follows: 1:99, 5:95, 20:80, and 50:50 of H3 from *hst3*Δ *hst4*Δ double mutant strain versus recombinant

histone H3. Error bars calculated for all pairwise ratios of H3K56ac tryptic peptide ions. The graph shows a nearly linear relationship between the experimentally observed fold-change increase in K56ac (solid diamonds) as a function of the calculated fold-change. In contrast, the abundance of H3 peptides (VTIQKKDIKLAR, KSTGGKAPR, YKPGTVALR, KQLASKAAR, denoted by the superimposed solid squares) that are not modified in either recombinant histones or H3 purified from *hst3*Δ *hst4*Δ double mutant cells was constant over the same range of molar ratio of the two sources of histones.

3.4.4 Differential Expression Profiling and Peptide Sequencing

Off-line LC-UV fractionation of histones from WT cells and a representative yeast mutant strain (*rtt109*Δ) gave almost superimposable chromatograms with the sequential elution of H2B, H4, H2A, followed by H3 (figure 3.5). It is noteworthy that off-line fractionation was achieved using TFA as an ion pairing reagent to improve separation efficiency and resolution of histone variants in contrast to nano LC-MS analyses where formic acid was selected to minimize ion suppression effects (figure 3.2) The only difference noted in the LC-UV chromatogram of histones from the yeast mutant strain (*rtt109*Δ) was the presence of a peak closely eluting with that of histone H3, which was not observed in the WT histone extract (figure 3.5b). This peak was at first suspected to be isoforms of H3, and, hence, was collected and pooled in the same tube as H3. Histones H2B and H3 were well resolved by the LC column and eluted in discrete fractions, while H4 and H2A eluted together in the same fraction.

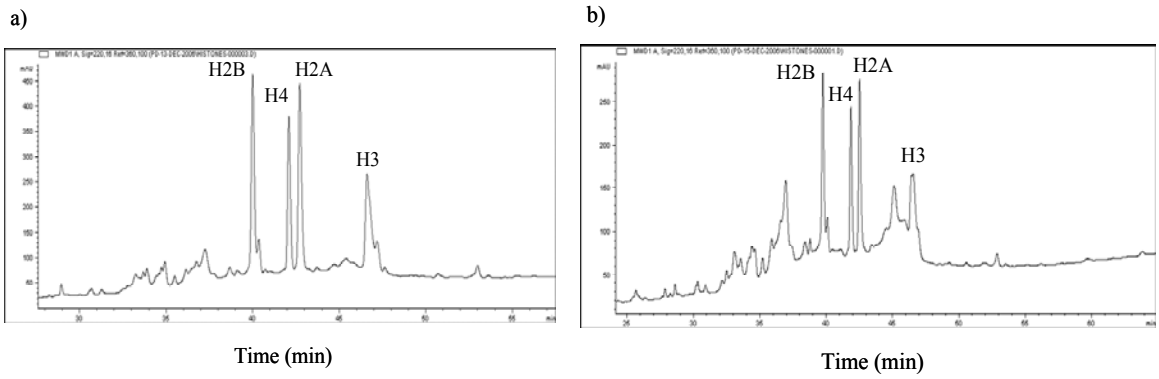


Figure 3.5: LC-UV fractionation of core histones from (a) WT and (b) *rtt109*Δ mutant. The fractionation was performed by reversed-phase HPLC on a Phenomenex Jupiter C_{18} column, 150 x 1 mm i.d. Flow rate 35 µL/min. Fractions were collected every minute in a 96-well plate. Gradient elution was performed from 5 to 90% acetonitrile in 60 min.

Since only histones H2B and H3 showed significant changes in their intact mass profiles, these two histones were the most promising candidates for the second step of our two-pronged strategy. However, as previously pointed out, minor modifications can be missed when generating intact mass profiles, due to overlaps in the mass reconstructions when many isoforms and isobaric shifts are present [16]. To avoid missing small changes in PTMs and demonstrate proof of principle, histones H2A and H4 were also subjected to the next step of our analytical strategy. After propionylation and tryptic digestion, each histone fraction was injected in triplicate for abundance profiling. The nano LC-MS raw data files were then converted into peptide maps for clustering analysis. All peptide clusters from each condition were exported to an Excel spreadsheet for direct comparison. Scatter plots of peptide intensity enable the rapid identification of histone peptide ions showing differential abundance between two sample sets. For histones H2A, H2B, and H4, no PTMs were found differentially expressed among WT and mutant cells. Only histone H3 showed differentially expressed histone peptides. Scatter plots of H3 peptides derived from WT and either *rtt109*Δ or *asf1*Δ mutants are shown in figure 3.6. The scatter plot for the yeast mutant strain *rtt109*Δ indicated that almost all identified peptide ions had no statistically significant change in ion intensities compared to those of the WT strain (figure 3.6a). The only notable exception (solid diamond) was observed for a doubly-protonated peptide ion at *m/z* 638.9^{+2} showing a 79-fold decrease in ion abundance in the yeast mutant strain *rtt109*Δ. This ion was subsequently identified as tryptic peptide FQKSTELLIR bearing an acetylated K56 residue (see below). The reconstructed ion chromatograms (RIC) demonstrated that this peptide ion was highly abundant in the H3 sample from WT cells, but was almost completely absent in histones derived from the two mutant yeast strains (inset figure 3.6a). The scatter plot obtained for the comparison of histone H3 from *asf1*Δ mutant and WT strains (figure 3.6b) showed an ion abundance profile similar to that observed for the yeast mutant strain *rtt109*Δ. Again, the only statistically significant change in ion abundance observed was for *m/z* 638.9^{+2}, where we observed a 153-fold decrease in acetylation, while all other ions showed no change in abundance. The RIC of the corresponding ions showed similar profiles to those of the mutant strain *rtt109*Δ (figure 3.6b inset).

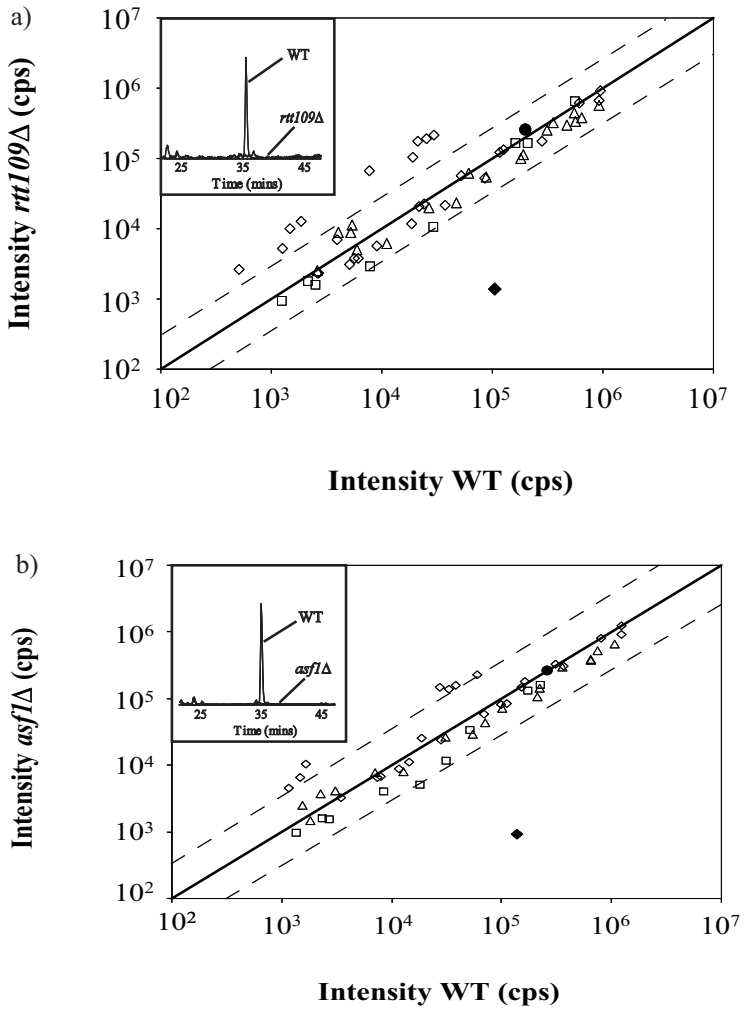


Figure 3.6: Scatter plots of ion distribution from nano LC-MS analyses of histone peptides from WT, *rtt109*Δ, and *asf1*Δ mutant cells. All data points plotted from triplicate nano LC-MS analyses with p-values of < 0.05. Peptides with an assigned MS/MS spectrum are denoted by diamonds for the H3 fraction, triangles for the H4 and H2A fractions, and squares for the H2B fraction. (a) Comparison of histone peptides from WT and *rtt109*Δ mutant cells. (b) Comparison of histone peptides from WT and *asf1*Δ mutant cells. The data point denoted by a solid diamond is a differentially abundant histone peptide containing a K56ac residue derived from histone H3. Conversely, the data point denoted by a solid circle corresponds to a propionylated K56 residue that exhibited no difference in abundance. Insets correspond to the RIC for *m/z* 638.9^{+2} corresponding to the H3K56ac tryptic peptide.

Nano LC-MS/MS analysis of significant outliers was performed to obtain peptide sequences and determine their PTMs. Tandem MS sequencing of the ion at *m/z* 638.9^{+2} revealed this to be H3K56ac peptide, a histone H3 peptide containing an acetylated lysine

56 (K56) residue; the MS/MS fragmentation produced an almost complete y ion series (figure 3.7a). In contrast, the ion at *m/z* 703.4^{+2} was determined to be equally abundant in histones from WT and mutant cells (figure 3.6a and b), and corresponds to a histone H3 peptide containing a monomethylated and propionylated K79 residue (figure 3.7b). Mascot searches revealed that the rest of the outliers in the scatter plots were in fact non-histone proteins (for example, yeast peptides derived from transposon Ty1-DR1) that were coextracted with histones during sample preparation.

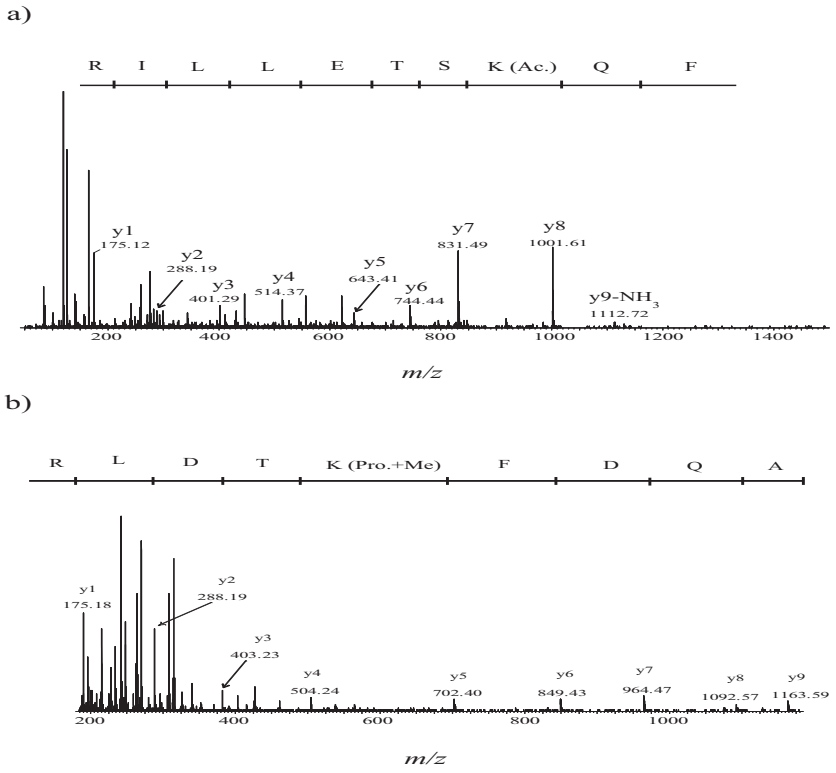


Figure 3.7: Histone peptide sequencing by tandem mass spectrometry. (a) MS/MS spectrum of *m/z* 638.9^{+2} identified as an H3K56ac peptide. (b) MS/MS spectrum of *m/z* 703.4^{+2} revealed a monomethylated (and propionylated) K79 residue in histone H3.

Interestingly, the peptide FQKSTELLIR comprising a propionylated K56 residue, hence a free ε-amino side chain in the native H3 histone, was also observed in high abundance in the tryptic digests of all yeast strains. No other modified form of the K56 side chain (e.g. methylation, ubiquitylation, etc.) was detected in our histone extracts. The scatter plots of ion abundance (figures 3.6a and b) indicated that the peptide with a free

K56 ε-amino group showed a variation of less than 27% between the mutant and WT strains. The variability in the type of modification observed for K56 suggests that only a minor subpopulation of histone H3 bear an N-acetylated side chain residue. To determine with higher precision the stoichiometry of H3 K56 acetylation in both WT and *rtt109*Δ strains, free NH_2 groups on a separate extract of histone H3 were derivatized using D_6-acetic anhydride, and the corresponding tryptic digests were analyzed by nano LC-MS/MS using positive ion MRM mode on an API 4000 Q-Trap mass spectrometer. The abundance ratio obtained from these targeted LC-MS/MS experiments revealed that *in vivo* acetylation of K56 accounts for 17% of all histone H3 molecules in the WT strain whereas this proportion decreased to only 0.4% in the *rtt109*Δ strain (data not shown). Accordingly, the present ion profiling approach enabled the identification of stoichiometric changes in modifications representing only 17% of all H3 histones. The abundance changes and modification sites determined for all core histones are summarized in figure 3.8 for yeast histones derived from WT, *rtt109*Δ and *asf1*Δ mutant strains.

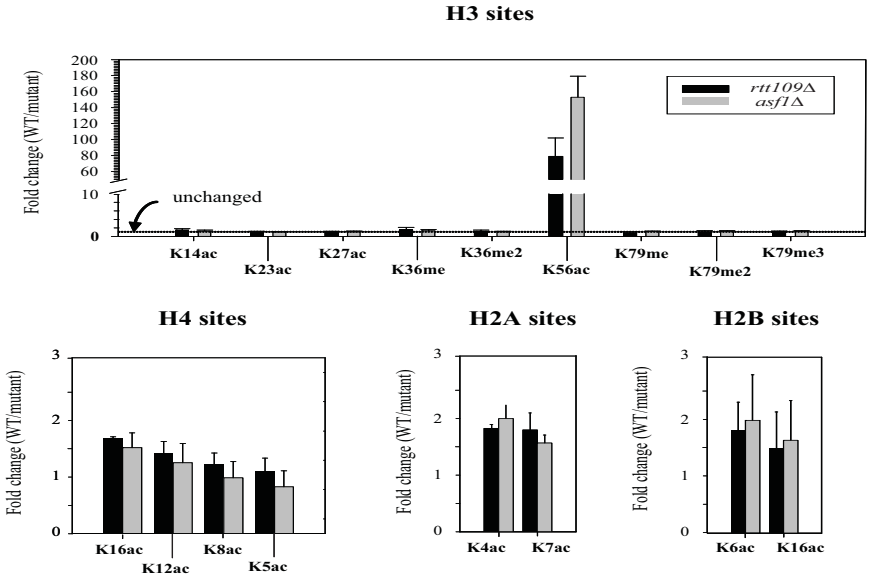


Figure 3.8: Modification maps and corresponding fold-changes in all the core histone PTMs detected in our analysis. The only modification that was differentially expressed was H3K56ac. The other modifications detected showed equal expression. Gray bars correspond to the relative abundance of histone PTMs in WT versus *asf1*Δ mutant cells. Black bars correspond to the relative abundance of histone PTMs in WT versus *rtt109*Δ mutant cells. ac: acetyl-lysine. me, me2, me3: mono-, di- and trimethyl-lysine.

3.5 Conclusion

The use of ion profiling and data mining approaches enabled the identification of K56ac as the only histone H3 modification site targeted by the histone acetyltransferase Rtt109 and the histone chaperone Asf1. Mutant strains lacking the acetyltransferase Rtt109 or the histone-binding protein Asf1 showed minimal K56ac, representing typically 0.4% of all H3 histones. None of the other histone PTMs examined substantially changed in abundance between WT and mutant cells (figure 3.8). H3K56ac showed a 79- and 153-fold increase in WT cells compared with histone H3 from *rtt109*Δ and *asf1*Δ cells, respectively. This difference may not be significant because the residual levels of H3K56ac in both mutant strains were very low and near the limit of detection of MS instrumentation, making it difficult to assess their abundance with high accuracy.

Identifying and profiling changes in histone modifications can be a daunting and challenging task. The methodology presented here is a rapid, comprehensive, and reliable procedure to profile and quantitate changes in the abundance of many histone modifications using scatter plots of ion distribution. The use of the MassSense peptide miner software developed in-house allows the generation of peptide maps for convenient clustering of ion abundance across sample replicates and conditions, thereby reducing the manual interpretation of large LC-MS data sets. This data mining approach can be used to identify differentially abundant modifications that can be subsequently validated using targeted analysis with isotope labeling and MRM. Among the plethora of ions generated by tryptic digestion, histone peptides could be identified and their relative abundance quantified in a reproducible fashion. Without the need for extensive sample preparation or peptide labeling, our generic profiling approach can be exploited in any scenario requiring a comparison of the relative abundance of histone PTMs in different samples. First, as illustrated here, our method is suitable to pinpoint which of the many known histone PTM sites are significantly affected by mutation of histone-modifying enzymes or other proteins. This approach will be very helpful to define the *in vivo* substrates of histone modifying enzymes. These findings should obviously be verified with *in vitro* enzyme assays to confirm that the changes in PTM abundance identified by MS are not

indirect effects of the mutation of a given histone-modifying enzyme. Second, our method makes it relatively straightforward to compare and contrast histone PTMs under two physiological conditions (e.g. control cells and cells treated with cancer chemotherapeutic agents). Third, our method may prove useful to profile global changes in histone PTMs during cell cycle progression or other changes in cell physiology. Finally, using tandem affinity purification strategies, it is now relatively simple to purify complexes of chromatin-modifying enzymes bound to their histone substrates. Our profiling method could be helpful to identify histone PTMs that are specific to certain chromatin-modifying enzymes.

3.6 Acknowledgment

IRIC receives infrastructure support funds from the Fonds de la Recherche en Santé du Québec (FRSQ) and from a Canadian Institutes for Health Research (CIHR) multi-resource grant. This work was carried out with the financial support of operating grants from the National Science and Engineering Research Council (NSERC) Strategic Grant Initiative (PT), the Canada Research Chair (AV, PT), and the CIHR (AV).

3.7 References

1. Davey, C.A., et al., *Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution.* J Mol Biol, 2002. **319**(5): p. 1097-113.

2. Robinson, P.J. and D. Rhodes, *Structure of the '30 nm' chromatin fibre: a key role for the linker histone.* Curr Opin Struct Biol, 2006. **16**(3): p. 336-43.

3. Mersfelder, E.L. and M.R. Parthun, *The tale beyond the tail: histone core domain modifications and the regulation of chromatin structure.* Nucleic Acids Res, 2006. **34**(9): p. 2653-62.

4. Zhang, L., et al., *Identification of novel histone post-translational modifications by peptide mass fingerprinting.* Chromosoma, 2003. **112**(2): p. 77-86.

5. Kouzarides, T., *Chromatin modifications and their function.* Cell, 2007. **128**(4): p. 693-705.

6. Jenuwein, T. and C.D. Allis, *Translating the histone code.* Science, 2001. **293**(5532): p. 1074-80.

7. Jones, P.A. and S.B. Baylin, *The epigenomics of cancer.* Cell, 2007. **128**(4): p. 683-92.

8. Altucci, L., et al., *Acute myeloid leukemia: therapeutic impact of epigenetic drugs.* Int J Biochem Cell Biol, 2005. **37**(9): p. 1752-62.

9. Johnson, L., et al., *Mass spectrometry analysis of Arabidopsis histone H3 reveals distinct combinations of post-translational modifications.* Nucleic Acids Res, 2004. **32**(22): p. 6511-8.

10. Zhang, K., et al., *Histone acetylation and deacetylation: identification of acetylation and methylation sites of HeLa histone H4 by mass spectrometry.* Mol Cell Proteomics, 2002. **1**(7): p. 500-8.

11. Pesavento, J.J., et al., *Shotgun annotation of histone modifications: a new approach for streamlined characterization of proteins by top down mass spectrometry.* J Am Chem Soc, 2004. **126**(11): p. 3386-7.

12. Thomas, C.E., N.L. Kelleher, and C.A. Mizzen, *Mass spectrometric characterization of human histone H3: a bird's eye view.* J Proteome Res, 2006. **5**(2): p. 240-7.

13. Kamakaka, R.T. and S. Biggins, *Histone variants: deviants?* Genes Dev, 2005. **19**(3): p. 295-310.

14. Dokmanovic, M. and P.A. Marks, *Prospects: histone deacetylase inhibitors.* J Cell Biochem, 2005. **96**(2): p. 293-304.

15. Kelly, W.K., et al., *Phase I study of an oral histone deacetylase inhibitor, suberoylanilide hydroxamic acid, in patients with advanced cancer.* J Clin Oncol, 2005. **23**(17): p. 3923-31.

16. Zhang, L., et al., *Differential expression of histone post-translational modifications in acute myeloid and chronic lymphocytic leukemia determined by high-pressure liquid chromatography and mass spectrometry.* J Am Soc Mass Spectrom, 2004. **15**(1): p. 77-86.

17. Garcia, B.A., et al., *Pervasive combinatorial modification of histone H3 in human cells.* Nat Methods, 2007. **4**(6): p. 487-9.

18. Mikesh, L.M., et al., *The utility of ETD mass spectrometry in proteomic analysis.* Biochim Biophys Acta, 2006. **1764**(12): p. 1811-22.

19. Knapp, A.R., et al., *Quantitative profiling of histone post-translational modifications by stable isotope labeling.* Methods, 2007. **41**(3): p. 312-9.

20. Ong, S.E., G. Mittler, and M. Mann, *Identifying and quantifying in vivo methylation sites by heavy methyl SILAC.* Nat Methods, 2004. **1**(2): p. 119-26.

21. Smith, C.M., et al., *Mass spectrometric quantification of acetylation at specific lysines within the amino-terminal tail of histone H4.* Anal Biochem, 2003. **316**(1): p. 23-33.

22. Su, X., et al., *Histone H4 acetylation dynamics determined by stable isotope labeling with amino acids in cell culture and mass spectrometry.* Anal Biochem, 2007. **363**(1): p. 22-34.

23. Zhu, H., et al., *Residue-specific mass signatures for the efficient detection of protein modifications by mass spectrometry.* Anal Chem, 2002. **74**(7): p. 1687-94.

24. Beck, H.C., et al., *Quantitative proteomic analysis of post-translational modifications of human histones.* Mol Cell Proteomics, 2006. **5**(7): p. 1314-25.

25. Syka, J.E., et al., *Novel linear quadrupole ion trap/FT mass spectrometer: performance characterization and use in the comparative analysis of histone H3 post-translational modifications.* J Proteome Res, 2004. **3**(3): p. 621-6.

26. Hake, S.B., et al., *Expression patterns and post-translational modifications associated with mammalian histone H3 variants.* J Biol Chem, 2006. **281**(1): p. 559-68.

27. Tsubota, T., et al., *Histone H3-K56 acetylation is catalyzed by histone chaperone-dependent complexes.* Mol Cell, 2007. **25**(5): p. 703-12.

28. Pan, X., et al., *A robust toolkit for functional profiling of the yeast genome.* Mol Cell, 2004. **16**(3): p. 487-96.

29. Celic, I., et al., *The sirtuins Hst3 and Hst4p preserve genome integrity by controlling histone H3 lysine 56 deacetylation.* Curr Biol, 2006. **16**(13): p. 1280-9.

30. Poveda, A., et al., *Hif1 is a component of yeast histone acetyltransferase B, a complex mainly localized in the nucleus.* J Biol Chem, 2004. **279**(16): p. 16033-43.

31. Bonneil, E., et al., *Proceedings of the 55th ASMS conference, Indianapolis, IN, June 3-7, 2007.* 2007.

32. Song, O.K., et al., *An Nalpha-acetyltransferase responsible for acetylation of the N-terminal residues of histones H4 and H2A.* J Biol Chem, 2003. **278**(40): p. 38109-12.

33. Garcia, B.A., et al., *Characterization of phosphorylation sites on histone H1 isoforms by tandem mass spectrometry.* J Proteome Res, 2004. **3**(6): p. 1219-27.

34. Garcia, B.A., et al., *Chemical derivatization of histones for facilitated analysis by mass spectrometry.* Nat Protoc, 2007. **2**(4): p. 933-8.

35. Luger, K., et al., *Characterization of nucleosome core particles containing histone proteins made in bacteria.* J Mol Biol, 1997. **272**(3): p. 301-11.

36. Maas, N.L., et al., *Cell cycle and checkpoint regulation of histone H3 K56 acetylation by Hst3 and Hst4.* Mol Cell, 2006. **23**(1): p. 109-19.

37. Thaminy, S., et al., *Hst3 is regulated by Mec1-dependent proteolysis and controls the S phase checkpoint and sister chromatid cohesion by deacetylating histone H3 at lysine 56.* J Biol Chem, 2007. **282**(52): p. 37805-14.

4. Enhanced Protein Detection Using a Trapping Mode on a Hybrid Quadrupole Linear Ion Trap (Q-Trap)

Paul Drogaris, J.C. Yves Le Blanc, Jennifer E. Fitzgerald, Noel F. Lowndes, Alain Verreault, and Pierre Thibault

Analytical Chemistry, 2009, Aug 1; 81 (15): 6300-6309

4.1 Abstract

A novel method to improve the detection of protein ions using a linear ion trap mass spectrometer is presented. A scan function combining charge separation with segmented transmission of multiply charged ions was developed to enhance the sensitivity and resolution of the linear ion trap for the nanoLC-MS analysis of intact proteins. The analytical benefits of the present method are particularly apparent in protein analyses, where the increased proportion of multiply charged ions can exacerbate space-charge effects and compromise the dynamic range of the linear ion trap instrument. The enhanced ion storage and charge separation capabilities of our targeted and enhanced multiply charged scan mode provided a 4-fold increase in signal-to-noise and 5-fold increase in resolution, thus enabling the detection of closely related protein isoforms. The application of this method is demonstrated for low femtomole detection of protein standards and nuclear extracts enriched in histone proteins. The enhanced resolution of this scan mode also enabled us to monitor subtle changes in the methylation of a subpopulation of histone H3 that occurs in chicken DT40 cells lacking specific methyltransferase activity. The extent of the fold change and PTM site localization was performed using predictive software tools and targeted multiple reaction monitoring analysis of histone peptides. Monomethylation of Lys 79 in histone H3 (H3K79me1) was down-regulated by 240-fold in methyltransferase deficient cells.

4.2 Introduction

The proteome of eukaryotic cells is relatively complex and further diversified by over 200 possible post translational modifications (PTM) [1]. Proteomics consists in applying global experimental approaches to determine the biological function of proteins [2]. Mass spectrometry (MS) is now the method of choice to characterize the proteins present in cells and tissues due to its speed, sensitivity, and selectivity. Two MS strategies can be utilized to characterize proteins and their PTM, namely the "bottom-up" [3, 4] and the "top-down" approaches [5, 6]. The bottom-up strategy consists of performing enzymatic and chemical digestion of protein slices from SDS-polyacrylamide gels or HPLC fractions using enzymes or chemicals that cleave the polypeptide chain at specific amino acid residues. This is followed by MS analysis to establish a peptide mass fingerprint, and tandem MS to determine sequence information. In the "top-down" strategy, the intact protein is introduced directly into the gaseous phase prior to mass analysis; fragmentation of the whole protein ions gives rise to sequence information.

When studying the vast array of PTM, initial MS profiling of intact proteins from cellular extracts can provide a rapid overview of the types and extent of post-translational modifications. Covalent protein modifications such as methylation, acetylation, and phosphorylation can be inferred by the observations of common mass shifts (+14, +42, and +80 Da, respectively) in the deconvoluted mass spectrum. There are numerous types of instrumentation available for MS-based proteomic profiling. These include: time-of-flight (TOF), Q-TOF, ion trap, Orbitrap, and Fourier transform ion cyclotron resonance (FT-ICR) [7, 8]. The use of FT-ICR is the gold standard for performing intact protein analysis due to its high mass resolution and accuracy. More recently, the profiling of intact proteins and the identification of partial sequences using collision activated dissociation (CAD) and multistage activation have been reported using hybrid LTQ-Orbitrap MS [9]. Q-TOF instruments with medium range resolution have also been used successfully [10-12] to study intact proteins and generate sequence tags from MS/MS data. Ion traps modified to perform ion-ion reactions in the gas phase for charge state reduction of intact proteins [13-15] have shown to be a robust and low-cost alternative to

FT-ICR and Orbitrap mass analyzers for whole protein analysis. However, high-end resolution instruments are not always available in most analytical laboratories, leaving researchers to multitask the usage of instrumentation at hand.

Ion trap instruments that have a small footprint, rapid acquisition rate, and provide good sensitivity are ideal instruments for performing quick, preliminary PTM screening on intact proteins. An earlier report by Claverol and colleagues [16] demonstrated that intact proteins eluted from 1- and 2-D gels could be stored and detected using an LCQ deca ion trap instrument. However, the detection levels were significantly high, and the samples analyzed were model standard proteins of low complexity. Although the mass spectrometer could generate enough resolution to identify single isoforms of unmodified proteins, a sample of higher complexity containing many PTM would have been too challenging to characterize and distinguish all the components present. Historically, the use of ion traps for intact protein analysis has been limited by two factors: low mass resolution and space-charge effects. Protein charge state determination is not possible due to the low resolution of the instrument. Trapping capacity is hampered by introducing a high number of multiply charged protein ions in a limited physical area, leading to space-charge. This last phenomenon has been extensively studied [17-19] to understand its detrimental effects on mass analysis, which include apparent shifts in *m/z* ratio, peak broadening, and loss of ion signal.

The introduction of two dimensional (2-D) linear ion traps (LITs) [20, 21] has offered a new technology for investigating intact proteins due to their increased ion storage capacity and faster scan rates for resonance ejection of ions. For example, Reid and colleagues [22] performed top-down analysis and sequencing of *Staphylococcus aureas* aldolase site-directed mutagenesis products using an LTQ ion trap instrument. While this study was impressive in showing the capabilities of the instrument, a relatively simple mixture was interrogated, and no analysis was performed on an LC time scale. Another study by Hunt and colleagues [23] was performed on intact ubiquitin, bovine serum albumin (BSA), and large histone peptides from Glu-C digestion using an LIT with electron transfer dissociation (ETD) capability. This demonstrated the ability of

these instruments to contain and detect simple protein or peptide mixtures with on-line LC-MS. A recent report by Kelleher and colleagues [24] was one of the few studies that established the ability of LITs to analyze intact yeast proteins fractionated from whole cell lysates by on-column chromatography.

The 4000 Q-Trap, a hybrid quadrupole linear ion trap, is another mass spectrometer that can be used to profile intact proteins for rapid PTM screening. The third quadrupole on this instrument serves as both a conventional RF/DC resolving mass filter and a 2-D LIT using axial ion ejection [25]. The scanning and trapping capabilities present on the same instrument result in versatile scan modes that are very useful in proteomics applications [26], such as precursor ion, constant neutral loss, and enhanced multiply charged (EMC) scan. The EMC scan mode enables elimination of singly charged ions from the LIT prior to mass analysis. Originally developed on conventional triple quadrupole and Q-TOF instruments [27, 28], the charge state separation can be performed in the LIT once analyte ions are trapped and thermalized. Adjustment of the DC voltage on the LIT exit lens or IQ3 results in the preferential release of singly charged ions, while multiply charged species are retained in the linear trap prior to mass analysis. The removal of singly charged species can result in higher signal-to-noise ratios and mass spectra with less chemical noise, while simultaneously reducing space-charge effects. In essence, the LIT is filled to its maximum analytical capacity with ions of interest, in this particular case, multiply charged ions.

In this report, we describe a novel trapping method on the 4000 Q-Trap based on a targeted EMC (tEMC) to capture and detect intact protein ions in the LIT with increased mass resolution and dynamic range compared to conventional scanning and trapping functions. This was accomplished by modifying a conventional scan function, and selecting narrow *m/z* regions sequentially into the LIT followed by charge separation to minimize the presence of singly charged ions prior to mass analysis and detection. Model compounds and cellular extracts were used to optimize MS parameters, benchmark performance, and establish figures of merit. The analysis of histone proteins from complex cellular extracts by LC-MS is used to illustrate the utility of the tEMC scan

mode to generate molecular mass profiles for rapid screening of changes in PTM abundance in intact proteins from different sample sets. Finally, targeted peptide detection and sequencing using predictive software tools on the same instrumental platform was performed to localize and quantitate differentially expressed histone PTM.

4.3 Experimental Section

4.3.1 Reagents and Separation Media

Protein concentrations in cell extracts and fractionated samples were measured using the micro-BCA assay kit (Pierce, Rockford, IL). Microbore HPLC C_{18} columns (150 mm x 1 mm i.d.) for histone fractionation were purchased from Phenomenex (Torrance, CA). Capillary LC-MS columns were packed in-house using Jupiter C_{18} particles, (300 Å, 3 μm, Phenomenex) and fused-silica tubing (Polymicro Technologies, Phoenix, AZ). Trapping columns were packed in-house using Jupiter C_{18} material in Teflon tubing (Supelco, Bellefonte, PA). Solvents for instrumental analysis were all HPLC grade (Thermo Fisher Scientific, Whitby, ON, Canada and Milli-Q water, in-house). Trifluoroacetic acid (TFA), ammonium bicarbonate, apomyoglobin, cytochrome C, leucine enkephalin, and propionic anhydride were purchased from Sigma (Oakville, ON, Canada). Formic acid (FA) was obtained from EM Science (Mississauga, ON, Canada). Porcine modified trypsin (sequencing grade) was obtained from Promega (Madison, WI).

4.3.2 Histone Purification

Recombinant yeast histone H3 (rH3) was expressed and purified under denaturing conditions as previously described [29] . Histones were isolated by acid extraction from wild-type and *DOT1L*-deficient chicken DT40 cells using the protocol described by Hake and colleagues [30]. Using reversed phase HPLC, rH3 in 8M urea was fractionated several times to concentrate and purify the histone. The protein content was determined using the micro-BCA assay, with BSA as the calibration standard. Using this assay, the

concentration of rH3 was 142 ng/µL. The rH3, as well as histones partially purified from wild-type (WT) and *DOT1L*-deficient cells, were diluted with the initial mobile phase and injected directly onto the LC-MS for analysis. The chicken histone samples were fractionated and prepared as outlined in a previous report [31]. Briefly, core chicken histones from DT40 cells were purified in the same manner by reversed phase HPLC (~10 µg total protein). Individual histone fractions of DT40 cells were dried in a SpeedVac evaporator, derivatized with propionic anhydride, and digested overnight at 37° C using 1 µg of trypsin. Enzymatic digests were acidified with 5% TFA in water (v/v) prior to LC-MS analysis.

4.3.3 Mass Spectrometry

All infusion and LC-MS experiments were carried out on an AB/MDS Analytical Technologies 4000 Q-Trap hybrid quadrupole-LIT instrument, equipped with the Nanospray II ion source, and controlled through Analyst 1.4 software. For the conventional quadrupole experiment, the Q3 mode was used and scanned from *m/z* 400 – 1600. The conventional Enhanced MS (EMS) and Enhanced Multiply Charged (EMC) modes were performed in a looped experiment while monitoring *m/z* 400 – 1600 at 1000 amu/sec, with an LIT fill time of 20 msec. The targeted EMC (tEMC) scan is essentially EMC controlled through the MS^3 method editor. This scan mode is a modified EMC scan function that provides control of the Q1 resolution as well as the center of mass of the transmission window. All tEMC scans were performed with Q1 set to open resolution using an 8 volt offset, thus providing uniform transmission windows of ~ 150-200 *m/z* units through Q1. The original MS^3 method table (.mtb) was modified to allow for ion cooling as RF levels are lowered prior to charge separation in the LIT and finer adjustment of the multiple charge separation (MCS) voltage using the collision energy spread (CES parameter) within the MS^3 mode. All tEMC parameters, scan ranges, and LC-MS data acquisition was controlled using the MS^3 method editor. Mass ranges from *m/z* 400 to 1600 in 200 *m/z* increments were entered once in the tEMC method editor, and used for infusion and LC-MS experiments. No other manual set-up was required. A total of five narrow *m/z* windows were introduced and mass analyzed sequentially to

monitor from *m/z* 400 to 1600 in a looped fashion to cover the multiply charged state distribution of proteins ions. The MCS voltage and Q3 empty time (both controlled by CES parameter and excitation time in MS^3) were set to 0.8 volts and 200 msec, respectively.

Peptide detection and sequencing were performed on the same instrumental platform. Chicken histone H3 amino acid sequences from the Swiss-Prot database were imported into AB/MDS Analytical Technologies MRM pilot software. The four most intense *b* and *y* fragment ions were chosen to generate theoretical MRM transitions in Analyst software. The masses of the precursor ions containing chemically modified (propionylated) H3 peptides were calculated using molecular weight calculator 6.4 (http://ncrr.pnl.gov/software/). A 10 ms dwell time was used to monitor the multiple reaction monitoring (MRM) transitions. The MRM experiment was followed by an enhanced product ion (EPI) scan in information dependant acquisition (IDA) mode to confirm the sequence of the detected peptides using the MIDAS strategy. EPI scan was monitored at 1000 amu/sec, with a fixed LIT fill time of 25 ms and Q0 trapping switched on.

4.3.4 NanoLC-MS of Intact Histones

Yeast rH3 and crude histone extracts from WT and *DOT1L*-deficient chicken DT40 cells were serially diluted from 50 ng/μL down to 1 ng/μL using the initial LC-MS mobile phase [95:5 water:acetonitrile, 0.2% FA (v/v)]. The Agilent 1100 series HPLC system with dynamic flow splitting was used as the separation module. An aliquot was injected onto a C_{18} trapping column (4 mm x 360 μm i.d.) for 3 min at 10 μL/min. Isolated proteins were eluted onto a C_{18} analytical column (100 mm x 150 μm i.d.) at 600 nL/min using a solvent system of 0.2% FA in water (A), and 0.2% FA in acetonitrile (B). Gradient elution was performed from 5 to 80% B in 80 min. Intact histone mass profiles were generated from acquired mass spectra using Bayesian protein reconstruction in BioAnalyst tools 1.4.

4.3.5 NanoLC-MS of Histone Tryptic Peptides

Histone tryptic peptides from WT and *DOT1L* mutant were analyzed using the same chromatographic system, columns, and mobile phases as outline above. Gradient elution was performed from 0 to 50% B in 60 min. Samples were injected a single time for the MIDAS strategy, and in triplicate using MRM only for quantitation, after the peptide ion transitions and sequences were confirmed.

4.4 Results and Discussion

4.4.1 Trapping Mode Conceptualization

Quadrupole ion traps have the ability to trap and scan analytes sharing the same physical space. However, there is a practical limit on the number of ions that can be stored in ion traps before distortion of the electric fields and decreased performance is observed due to space-charge effects. To circumvent this, features such as automatic gain control (AGC) and dynamic fill time (DFT) were implemented to provide optimum ion storage capacity for trapping instruments. AGC and DFT can be applied efficiently for small molecules and peptides, whose charge density is relatively small, but are far more problematic to implement for multiply charged protein ions, where space-charge effects readily override the benefits of AGC and DFT. Thus, introducing too many ions of different charge states into the trapping device is a major hurdle in the analysis of intact protein ions. This prompted us to develop the current capture and detection strategy on the 4000 Q-Trap, the concept of which is outlined in figure 4.1. The voltage ramping and duty cycle time allocation for each part of the tEMC scanning sequence are shown in figure 4.1a. After introducing and cooling the ions, the exit lens voltage is lowered to evacuate singly charged species (charge separation). This voltage is controlled via the MCS parameter, which is expressed as the potential difference between the exit lens and the Q3 rod offset. After the charge separation step, the auxiliary AC and RF drive are ramped simultaneously for mass selective axial ejection of ions (mass analysis). Finally,

any residual ions trapped in the system are emptied from the LIT prior to the next fill step. The functional aspects of the tEMC scanning mode are depicted in figure 4.1b. First, a fraction of the ions generated at the source are introduced into the LIT by gating them with the first quadrupole, where the resolution is set to open, and maximum voltage offset (8 V). This allows for a narrow yet uniform *m/z* window to be transmitted (~ 150-200 *m/z* units) into the LIT. This *m/z* region is the result of using an 8 volt offset on Q1, and is the maximum *m/z* window that can be transmitted. This process is repeated five times within the MS^3 method editor in a looped experiment to obtain enough narrow regions to cover the *m/z* range of interest (*m/z* 400 – 1600) and generate a complete mass spectrum. Once thermalized, singly charged species are evacuated from the LIT by applying an MCS of 0.8 V at the exit lens of the LIT. The MCS acts an energy "barrier" that is proportional to the charge state of the ion (i.e., energy 'barrier' = *z* x MCS). Hence, singly charged species experience a lower energy barrier and are evacuated first, while higher charge states will remain in the LIT at the same applied MCS potential difference. Once charge separation is completed, the remaining multiply charged species are scanned out of the LIT. The acquired mass spectral segments are then summed together in Analyst software to generate a full mass spectrum, and Bioanalyst tools are used to generate the molecular mass profile. The overall ion population during the scanning sequence is shown in figure 4.1c, illustrating that the number of singly charged ions is significantly decreased at each stage of the scanning sequence, thereby minimizing the overfilling of the ion trap and space-charge effects. The total capture and release of all ion species, both singly and multiply charged, is not possible; there is a loss of both types of species. However, singly charged ions are preferentially released, and their abundance is significantly reduced compared to multiply charged ions.

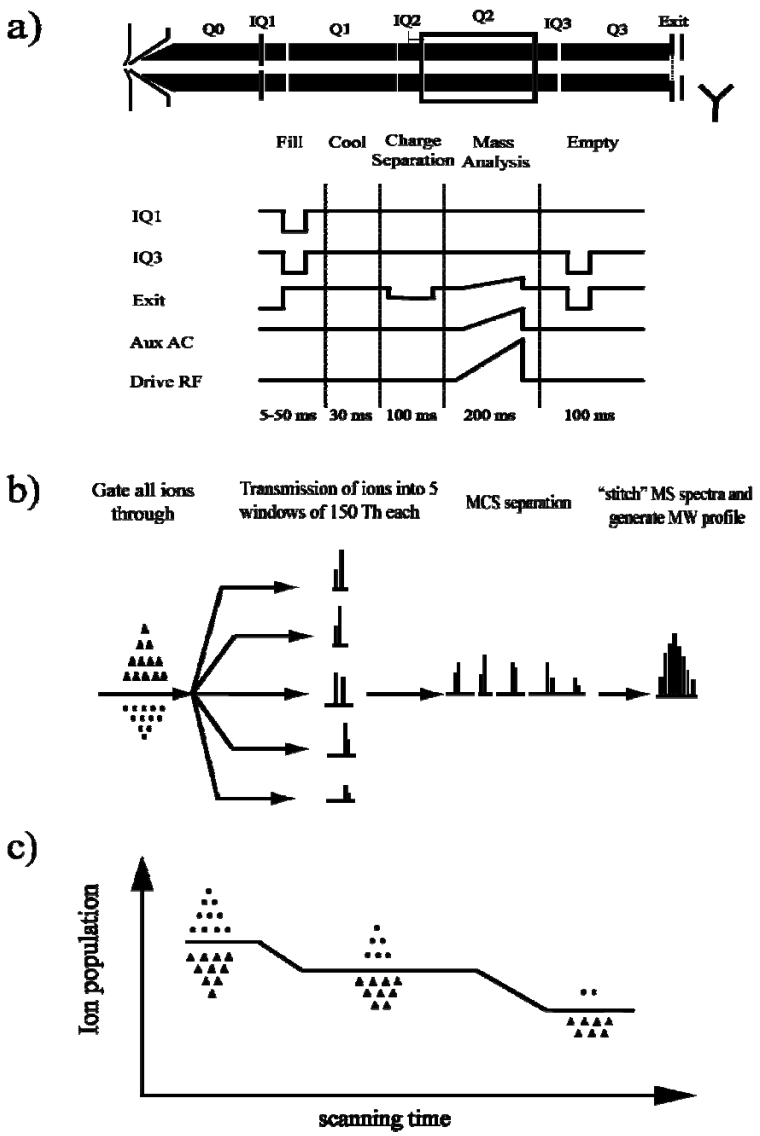


Figure 4.1: Overview of the ion trapping strategy for the targeted enhanced multiply charged (tEMC) scanning mode. (a) Schematic diagram of the Q-trap instrument and potential energy diagram for different lenses across the path of ions undergoing charge separation. (b) Schematic representation of segmented tEMC scan illustrating the transmission of narrow *m/z* regions, MCS separation, molecular mass reconstruction. (c) Distribution of singly (circles) and multiply charged ions (triangles) over the scanning time.

4.4.2 MCS Voltage Optimization and Space-Charge Effects

During the initial phase of development, we used conventional scan modes to determine if charge separation could be applied to intact proteins showing a distribution of multiply charged ions. These scan modes performed efficiently during infusion experiments and small scale optimization, but were significantly compromised during nanoLC-MS experiments. Figure 4.2a depicts the MCS voltage optimization using apomyoglobin (0.5 pmol/μL) as a model protein, spiked with 1 ng/μL of the singly charged leucine enkephalin (LE). The MCS was varied in 0.1 V increments, and 10 MCA scans were acquired for each value. The intensity of the $[M+25H]^{25+}$ ion from apomyoglobin ions (*m/z* 679) is plotted along with the $[M+H]^{+}$ (*m/z* 556) for LE. At low MCS values, the abundance of the *m/z* 556 for LE was significantly reduced during the infusion. As observed in Figure 4.2a, there is a critical voltage range between 0.7 and 0.8 V, where the apomyoglobin ions dominated the mass spectrum. The abundance of the LE ions increased rapidly beyond 0.9 V, resulting in progressively reduced charge separation. We used an optimized MCS voltage of 0.8 V in subsequent nanoLC-MS experiments.

The performance of the EMC detection mode for the nanoLC-MS analysis of intact proteins was evaluated for an acid extract of chicken histones and compared to that of a conventional quadrupole scan. Figure 4.2b shows the nanoLC-MS analysis of a 100 ng injection of intact chicken histones using a conventional Q3 scan. This scan mode was used for comparison purposes to provide an unaltered representation of the ion population without space-charge effects, while keeping transmission conditions through the ion optics (including collision cell) constant. Core histone peaks eluted between 42 to 50 minutes, with H4 and H2A eluting at 46 min, H2B at 47.5 min, and H3 at 50 min. The mass spectrum for the H3-containing peak, denoted by *, is shown as an inset to figure 4.2b. When the same sample was analyzed using the conventional EMC scan mode, there was a significant loss in sensitivity, and no histone protein ions were detected (figure 4.2c). The inset of Figure 4.2c shows the mass spectrum taken at the same elution time as in Figure 4.2b and illustrates that no envelope of histone H3 protein ions could be clearly

discerned under conventional LIT trapping. Similar effects were observed using enhanced MS (EMS). Because no ion gain control is implemented for the EMC scan, nanoLC-MS introduces a large number of ions of all charge states simultaneously into the LIT, leading to a loss of the multiply charged ion current during the elution of histones. This was evident from an overall decrease in ion current at 40 min with the elution of more abundant proteins (figure 4.2c). Using complex protein extracts, the amount of other components present, as well as the flux of singly charged species entering the LIT, is unknown. An extensively purified protein sample of known concentration was required to evaluate the onset of space-charge and understand the ion capacity and instrumental limits of the EMC mode. For these reasons, we compared the ion chromatograms and mass spectra of purified recombinant yeast histone H3 (rH3) and equine cytochrome C (Cyt C), whose protein concentrations were accurately determined.

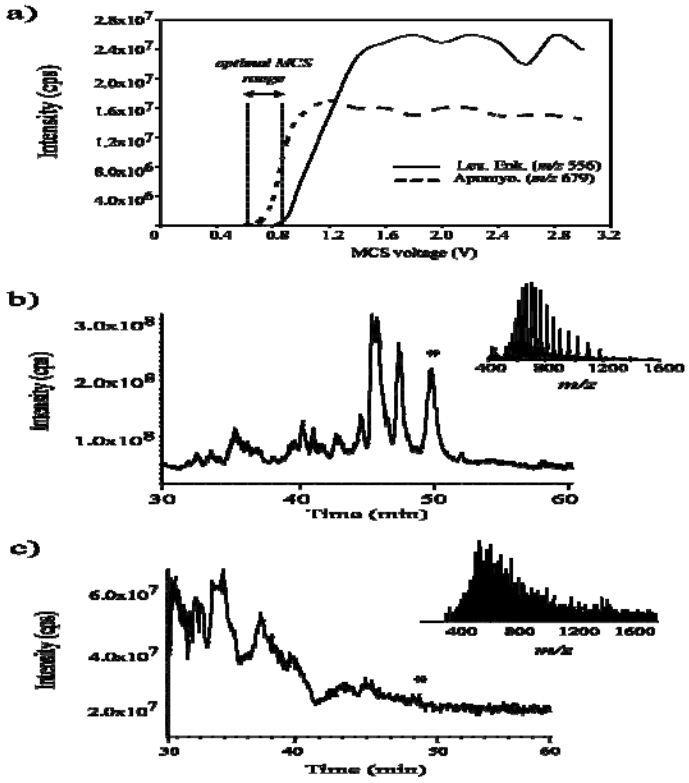


Figure 4.2: MCS voltage optimization and nanoLC-MS analysis using the conventional EMC scanning mode. (a) Changes in ion intensity for the $[M+25H]^{25+}$ ion from apomyoglobin (*m/z* 679) and the $[M+H]^{+}$ (*m/z* 556) of leucine enkephalin in response to incremental changes in MCS voltage. Total ion chromatograms of 100 ng of intact

chicken core histones from nanoLC-MS analysis using (b) Q3 and (c) conventional EMC scanning modes. Insets show mass spectra obtained from the peak denoted by *.

4.4.3 Characterization of Protein Ion Profiles Using Q3 and tEMC Modes

As indicated in the previous section, the relatively limited analytical performance of conventional EMC mode when charge state and ion flux cannot be determined accurately results in deleterious charge space-effects and reduced ion transmission. To overcome these limitations, we designed a modified EMC mode by changing the MS^3 method table to perform ion cooling and RF ramping before charge separation, and enable finer control of the MCS voltage. Instead of scanning the entire mass range all at once, we partitioned the effective scanning range into five regions of ~ 150-200 *m/z* units that were scanned in a sequential looped fashion. We prepared serial dilutions of Cyt C ranging from 0.01-60.0 ng/uL and analyzed them by nanoLC-MS using Q3 and tEMC modes. Panels a and b of figure 4.3 show a comparison of extracted mass spectra obtained from the nanoLC-MS analysis of 0.2 ng/μL (100 fmol injection) Cyt C using Q3 and tEMC modes, respectively. Two salient features are immediately noticeable when comparing the two corresponding mass spectra. First, a higher noise level was observed from *m/z* 600 to 1100 in the Cyt C spectrum obtained with the Q3 scan (figure 4.3a). This contrasts with the tEMC mass spectrum where multiply charged ions were detected with higher signal-to-noise, while the absolute intensity of the analyte ions remained the same (figure 4.3b). Second, the full width at half-maximum (fwhm) is significantly larger for the Q3 scan compared to that of the tEMC acquisition (insets of Figure 4.3). Typically, we observed an overall reduction of 50% in fwhm in the tEMC scan compared with that of a conventional Q3 scan. The increased resolution observed for the tEMC scan is attributed in part to the segmented *m/z* range selected, although the scanning cycle period remained comparable in both modes (i.e., 3 s/cycle for the Q3 and tEMC modes). In Q3 scan, the instrument is scanning a larger *m/z* range and thus spending less effective time on detected ions, whereas the tEMC mode benefits significantly from the longer collection time of individual ions for each segment, resulting in more resolved mass spectral peaks.

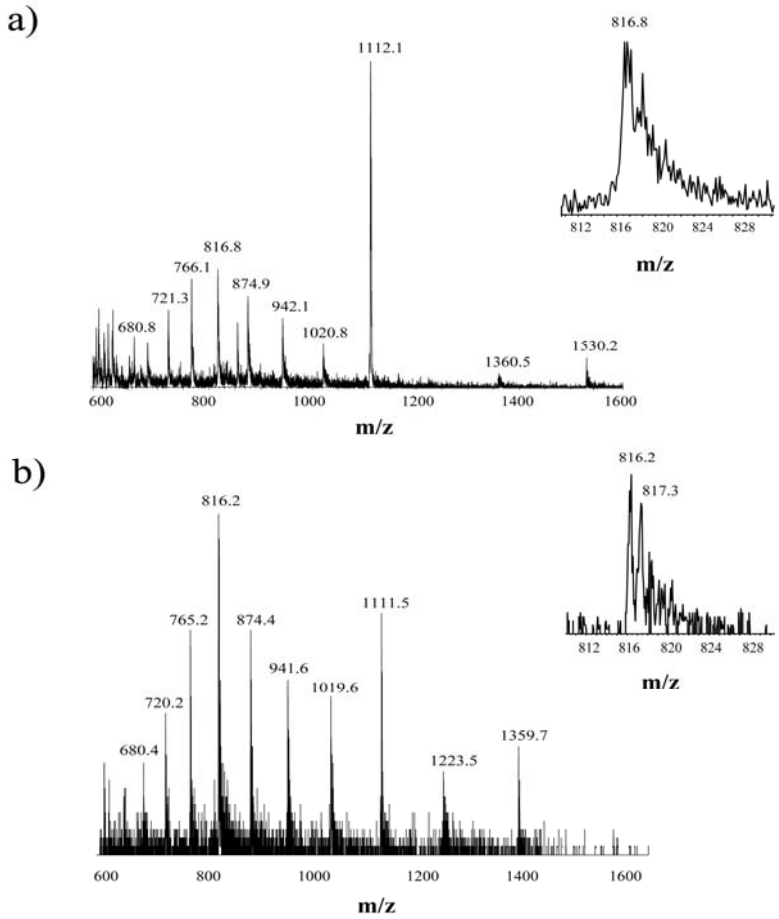


Figure 4.3: Enhancement of resolution in nanoLC-MS analysis of cytochrome C using tEMC. Comparison of extracted mass spectra obtained from the nanoLC-MS analysis of 0.2 ng/mL (100 fmol) of cytochrome C using (a) Q3 and (b) tEMC modes. Insets show the expanded view of the $[M+15H]^{15+}$ region (*m/z* 816).

The analytical performances of the Q3 and tEMC modes were further evaluated using rH3 to define practical limits of operation. The purified stock solution of rH3 was diluted serially to obtain protein concentrations of 0.2-20 ng/µL that were subsequently injected on the nanoLC-MS system. Extracted mass spectra obtained from the injection of 0.2 and 2 ng/µL are presented in figure 4.4. Mass spectra comparing the data collected in the Q3 and tEMC modes are shown in panels a and b of figure 4.4, respectively. An on-column sample loading of 1 ng of rH3 approached the detection limit of conventional Q3 scan, and the charge state envelope was interspaced with considerable background

and chemical noise (figure 4.4a). The molecular mass profile derived from the Q3 scan showed two prominent components at 15225 and 15268 Da. The rH3 protein was purified from *Escherichia coli* and does not comprise any PTM [32]; hence, explaining the simple molecular mass profiles observed here compared to the extensive acetylation and methylation usually taking place on the wild-type histone H3 from other organisms. The expected average mass of yeast histone H3 without the initiator Met residue is 15225 Da, in close agreement with that observed in figure 4.4a. The additional peak at 15268 Da is most likely a carbamylation artifact, because rH3 was isolated under denaturing conditions with 8 M urea prior to HPLC fractionation.

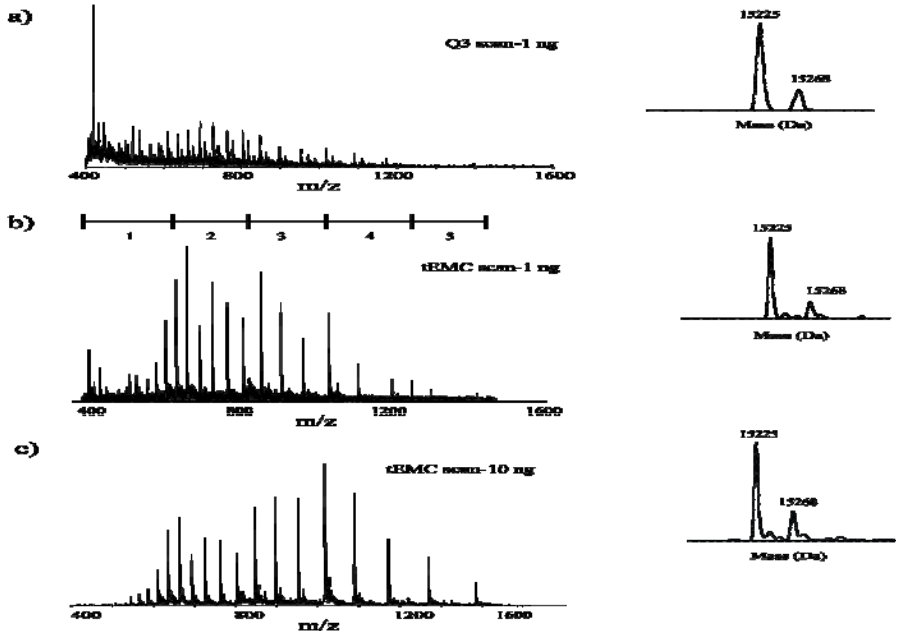


Figure 4.4: Increased sensitivity in nanoLC-MS analysis of intact proteins using tEMC. Extracted mass spectra obtained from the on-column injection of 1 ng of recombinant yeast histone H3 (rH3) using (a) Q3 and (b) tEMC modes. Extracted mass spectrum for the on-column injection of 10 ng of rH3 using the (c) tEMC scanning mode. Insets show the reconstructed molecular mass profile for the corresponding mass spectrum.

The analysis of the same sample using tEMC is shown in figure 4.4b and provided enhanced resolution and sensitivity compared to that of the Q3 scan (figure 4.4a). This was clearly evidenced by a 4-fold increase in signal-to-noise and 5-fold

increase in resolution (inset figure 4.4b). Interestingly, when 10 times more rH3 was injected, not only was the tEMC scan mode able to contain and detect the intact protein ions, but a bias in ion distribution favoring lower charge state protein ions was also observed (figure 4.4c).

To evaluate ion storage capacity of the LIT with respect to analyte concentration and space-charge limitations, we injected 10 ng of rH3 under nanoLC-MS conditions using fill times ranging from 2 to 500 ms. Figure 4.5a shows the changes in intensity for different multiply charged rH3 protein ions for the LIT fill times examined. Higher charge state ions ($[M+22H]^{22+}$ at *m/z* 693 and $[M+21H]^{21+}$ at *m/z* 726) showed a steady increase in ion abundance up to a fill time of 20 ms, after which a decline in intensity was observed. Conversely, lower charge state ions ($[M+12H]^{12+}$ at *m/z* 1269 and $[M+11H]^{11+}$ at *m/z* 1385) showed an extended linear progression in ion intensity up to 100 ms, followed with an abrupt decrease beyond this fill time period. It is noteworthy that the change in ion intensity was also accompanied by peak broadening and a shift in *m/z* values, which was more exacerbated for lower charge state ions (figure 4.5b). The proportional changes in *m/z* values resulted in a progressive increase of up to 4 Da in the apparent molecular mass of rH3 over fill times of 2 to 500 ms.

The bias in the distribution of multiply charged ions according to fill time was unexpected given the uniform storage of ions in the LIT. Indeed, the progressive increase in fill time is expected to provide a higher number of ions with a consistent distribution in multiply charged ions. However, when long fill periods increase the ion population beyond the capacity of the LIT (estimated to 10^5 ions for the 4000 Q-Trap), space-charge phenomena will have a more profound effect not only on ions of higher charge state, but also on those affected by the timing of the ejection and analysis sequence. For example, when Q1 is operated in RF-only mode, the LIT will typically store all ions from *m/z* 300-3000, irrespective of the scan region selected. Ions below the lower boundary of the scanning region will be ejected prior to the acquisition of the selected *m/z* range. Because the analysis sequence acquires segments from low to high *m/z* values, ions of a higher

charge state will experience greater ion population and space-charge effects than those of lower charge state.

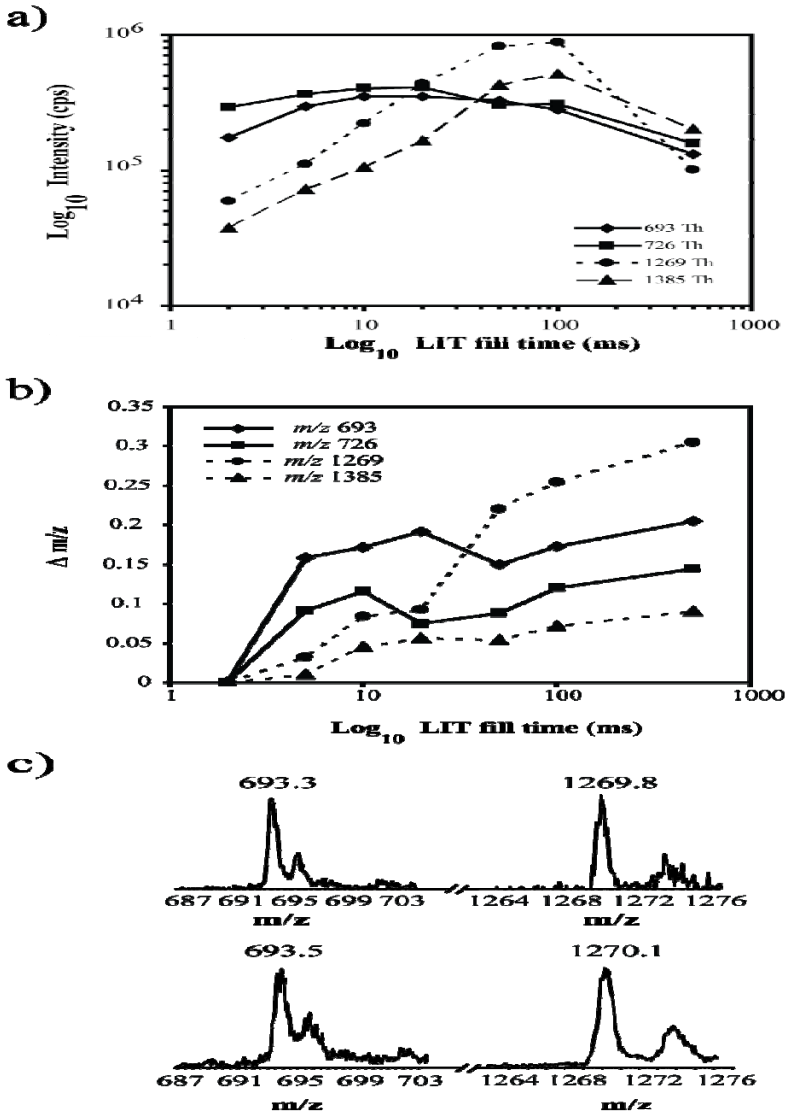


Figure 4.5: Effect of space-charge on ion intensity and shift in *m/z* values using the tEMC scanning mode. Variation in (a) ion intensity and (b) *m/z* values with LIT fill time for different multiply charged protein ions of rH3 (*n* =1). (c) LIT fill times of 2 ms (top) and 500 ms (bottom) differentially affect peak broadening and *m/z* shifts for rH3 ions with a high number (+22, *m/z* 693, left) and lower number (+12, *m/z* 1270, right) of positive charges.

This is illustrated in figure 4.5c for two scanning segments representing multiply protonated ions of rH3 with either 12 (*m/z* 1270) or 22 charges (*m/z* 693). As indicated, peak broadening and *m/z* shift are more significant for higher charge state ions as the fill

time is increased. Unlike EMS scan, where DFT can dynamically monitor the ion flux to avoid LIT overfilling, EMC does not possess this feature at the present time. The alternative taken in this work was to provide further control on the population of ions entering the LIT section by using Q1 as a high pass/low pass filter therefore transmitting a subset of the ion population and limiting deleterious space-charge effects. On the basis of these preliminary experiments and in order to maximize sensitivity while minimizing space-charge effects in tEMC mode, we used a LIT fill time of 20 ms during LC-MS acquisition.

4.4.4 LC-MS of Complex Histone Extracts

We evaluated the analytical potential of the tEMC scanning mode using acid-extracted proteins from wild-type chicken DT40 cells. These extracts contain not only histones, but also many other non-histone proteins [30]. Serial dilutions ranging from 1 to 20 ng/μL were prepared from the original protein extracts prior to nanoLC-MS analyses. A Q3 scan was used as a benchmark for instrument performance and was compared with conventional and modified EMC acquisition. We monitored the molecular mass profiles of histone H3 because of its extent in microheterogeneity and occurrence of different modifications. Panels a and b of figure 4.6 show the total ion chromatograms (TIC) and molecular profiles (inset) corresponding to the on-column injection of 5 and 50 ng of histone acid extracts derived from WT DT40 cells using a Q3 scan mode. The last eluting peak at 50.5 min (denoted by *) corresponds to histone H3, and the molecular mass profiles showed a relatively broad peak centered at 15329 Da (figure 4.6b). The complexity and heterogeneity of histone H3 isoforms could not be resolved by this scan mode, nor could any preliminary PTM screening be achieved. In contrast, panels c and d of figure 4.6 present the nanoLC-MS analyses of the same samples using a conventional EMC scan. The on-column injection of 5 ng revealed clearly detectable histone peaks with significantly higher resolution enabling the separation of histone isoforms. The 14 Da mass shifts observed in figure 4.6c suggest extensive methylation of the protein at different sites. When sample loading was increased up to 50 ng, space-charge effects led to an almost complete loss of ion current. This is exemplified in figure 4.6d for the on-

column injection of 50 ng where no clear peak was observed. The TIC and molecular mass profiles obtained using our tEMC acquisition mode are shown in panels e and f of figure 4.6.

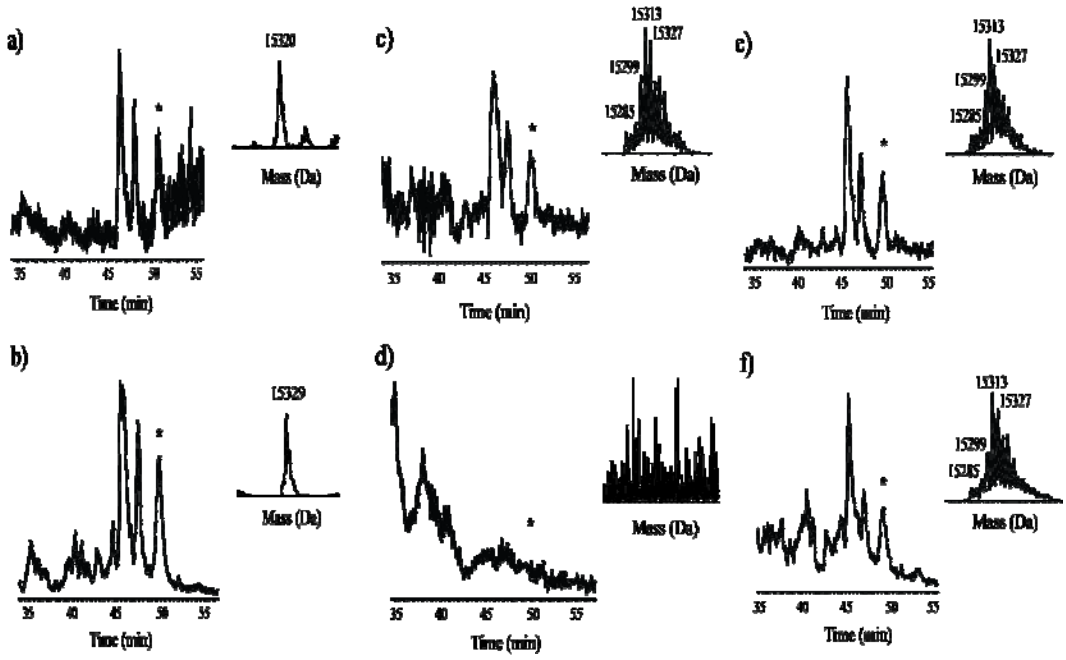


Figure 4.6: NanoLC-MS analysis of histones derived from acid extracts of WT chicken DT40 cells using different scanning modes. Total ion chromatograms (TIC) corresponding to the on-column injection of (a) 5 ng and (b) 50 ng of acid-soluble proteins using the Q3 scan. TIC for the on-column injection of (c) 5 ng and (d) 50 ng of acid-soluble proteins using the conventional EMC scan. TIC for the on-column injection of (e) 5 ng and (f) 50 ng of acid-soluble proteins using the modified (targeted) EMC scan. Insets show the reconstructed molecular mass profile for the histone H3 peak denoted by *.

Initial inspection of the TIC revealed better peak shapes and higher sensitivity for the different sample sizes that were examined. The difference in analytical performance between the conventional and modified EMC acquisition modes for 50 ng injection can be explained by deleterious space-charge effects that arise from unbalanced ion distribution and result in an overfilled LIT. The segmentation of *m/z* scanning regions enabled a partitioning of the ion population to minimize space-charge effects when numerous protein ions are present simultaneously. The histones were detectable and

eluted at the expected retention time; other acid-soluble components were also present and eluted before the proteins of interest. Furthermore, the molecular mass profiles of histone H3 were comparable at low and high sample loading (figure 4.6e and f). These results demonstrate that the tEMC scan mode can be an effective approach to reduce space charge-effects, extend the dynamic range of conventional EMC, and enhance resolution and mass measurements with performances approaching those of TOF instrumentation.

4.4.5 Profiling Changes in Histone Modifications Using tEMC

We next evaluated the potential of tEMC to profile changes in modifications of intact histone proteins caused by specific histone-modifying enzyme activities. *DOT1L* is a non-SET domain enzyme known to methylate histone H3 predominantly at Lys79 [33, 34]. Mutant cells lacking this enzyme should therefore exhibit a different histone H3 methylation pattern compared with that of wild-type cells. Protein mixtures that were acid-extracted from wild-type and *DOT1L*-deficient cells were analyzed and compared to correlate the extent of modification for the different core histones. The molecular mass profiles for all four core histones obtained from the corresponding nanoLC-MS analyses (50 ng on-column injection) are shown in figure 4.7. H2A profiles from wild type and *DOT1L*-deficient cells showed a series of peaks consistent with the removal of the initiating Met residue, followed by N-terminal acetylation (13851 Da). Mass shifts of +42 and +80 Da were also observed, suggesting possible acetylation and phosphorylation (13893 and 13931 Da) on the major H2A peak. This histone is known to be acetylated on the N-terminal amino group and lysine residues within the N-terminal tail. It is also known to be phosphorylated at Ser-1 and Thr-120 [35]. A relatively limited number of abundant modifications were also observed for H2B, as evidenced from the corresponding molecular mass profiles (figure 4.7). The most prominent component of H2B was consistent with the unmodified H2B-I variant lacking the initiating Met (13791 Da). The two other prominent forms that were observed (13819 and 13833 Da) could be attributed to other sequence variants of the protein (i.e., H2B V and H2B VII,

respectively). The profile for intact histone H4 was more heterogeneous than those of H2A and H2B. The unmodified form of the protein, with initiating Met removed and N-terminal acetylation, is observed at 11278 Da. The molecular mass profiles of H4 derived from both wild-type and *DOT1L*-deficient cells present a distribution of isoforms spaced by 14 and 42 Da, corresponding to Lys methylation and acetylation, respectively. The molecular mass profiles for H2A, H2B, and H4 from WT and *DOT1L*-deficient cells are essentially superimposable, indicating that the absence of the DOT1L methyltransferase did not lead to any significant change in the extent and heterogeneity of modifications for these core histones. However, this was not the case for histone H3, where subtle but detectable changes in modification patterns were observed (figure 4.7). The unmodified form of histone H3 lacking the initiator Met is present at 15257 Da, along with several distinct 14 Da mass shifts, suggesting a high degree of methylation. For the most part, the histone H3 profiles from WT and mutant cells share the same masses. However, the proportion of a number of isoform peaks is different. For example, histone H3 from cells lacking *DOT1L* showed lower abundance of the peaks at 15285 and 15 313 Da compared to H3 prepared from WT cells. An extra peak at 15383 Da was also present in the sample from *DOT1L* mutant cells, but was significantly less abundant in H3 from WT cells. As exemplified in figure 4.7, the possibility of probing subtle changes in histone modification patterns using the tEMC approach demonstrates the utility of this novel trapping mode for rapid profiling of PTM on intact proteins. Obviously, this would be impossible to achieve using regular quadrupole or ion trap scan modes because of their limited resolution.

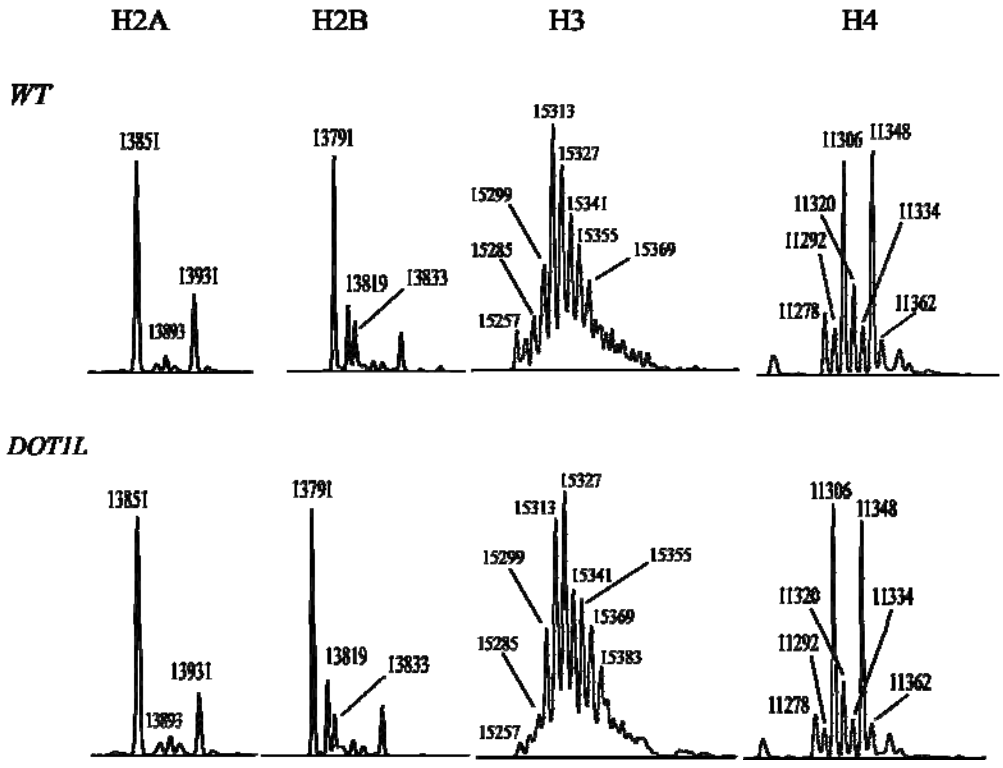


Figure 4.7: Changes in intact histone modifications between wild-type (WT) and *DOT1L*-deficient chicken DT40 cells detected by the tEMC scan mode. Molecular mass profiles of individual core histones were obtained from the mass spectra of peaks identified in the corresponding nanoLC-MS analysis (50 ng on-column injection).

4.4.6 Peptide Quantitation and Sequencing Using Targeted MRM

We subjected the DT40 cellular extracts to peptide analysis and sequencing in order to localize the histone H3 modification sites and quantify their relative abundance. After HPLC fractionation, propionylation, and sample preparation, the histone H3 tryptic digests from WT and mutant cells were injected onto the 4000 Q-trap using MRM for quantitation and the MIDAS [36, 37] strategy for peptide sequence confirmation. Theoretical MRM transitions were designed by in silico digestion of imported chicken histone H3 sequences from the Swiss-Prot database using MRM pilot software. Peptide transitions with known histone H3 methylation sites (i.e., K9, K27, K36, and K79) were incorporated into the MRM list. Two types of normalization were applied to the data. First, the protein amount was normalized on the basis of LC-UV peak areas collected for each histone fraction. Second, during the LC-MS analysis, the unmodified histone H3

tryptic peptide YRPGTVALR was monitored to normalize the amount of protein loaded, and to compensate for fluctuations in MS response. Peak area ratios for each modified histone H3 peptide were calculated using the YRPGTVALR peptide. The TIC from the MRM analysis of the H3 tryptic fraction is shown in figure 4.8a. The TICs from WT and mutant digests looked similar upon initial inspection, except for the peak at approximately 40 min (denoted by *). This peak was significantly more intense in the WT than in the mutant. An expanded view of this peak is shown as an inset in 4.8a, which revealed a doubly charged ion at *m/z* 703.4, containing four overlaid MRM transitions. This ion corresponded to the in silico generated H3K79me1 peptide with the following sequence: EIAQDFKTDLR. Figure 4.8b compares the abundance in WT and *DOT1L*-deficient cells of *y5* fragment ion originating from the *m/z* 703.4^{+2} precursor. This MRM transition belonged to a chemically modified (propionylated) K79 bearing a single methylation for both WT and mutant cells. The MRM peak for this transition was very intense in the WT (about 6.0×10^4 cps), while it was reduced to about 500 cps (shown 20 times expanded) in the mutant. The MS/MS spectrum of the *m/z* 703.4^{+2} precursor ion is shown in figure 4.8c. Tandem MS sequencing confirmed the identity of this peptide as a monomethylated (and propionylated) K79 residue, and provided an almost complete *y* ion series. The dimethylated form of the same peptide was detected in the WT at very low abundance, but not in the mutant sample. No H3K79me3 was detected in either WT or mutant. The peak area ratio generated from MRM analysis enabled the calculation of the fold change difference of H3K79me1 at 240-fold. The peak area ratio of H3K79me1 for the WT analysis had a CV of 3.6%, while for the mutant, the CV was 15% (n=3). The higher variation in the mutant sample analysis could be attributed to the low abundance of the peptide being close to the limit of detection. However, this result was still within acceptable confidence limits [38]. Other histone H3 tryptic methylated peptides were also detected at K9, K27, and K36, but these did not exhibit any significant changes in abundance between WT and mutant analyses (data not shown).

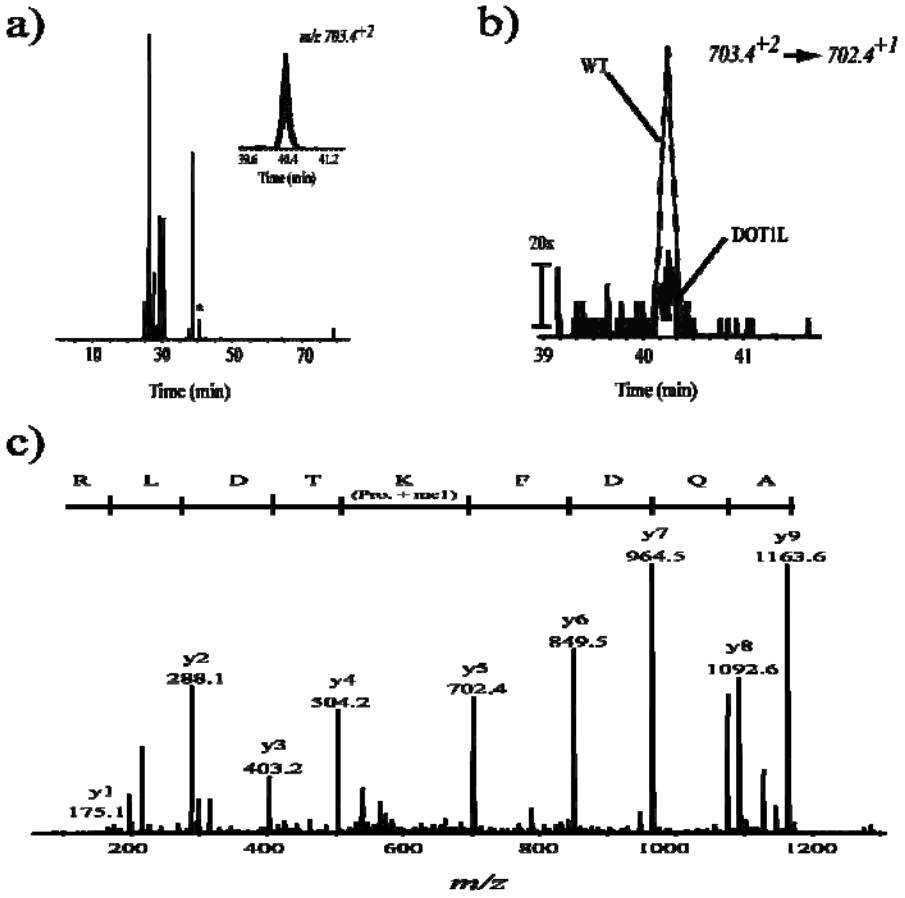


Figure 4.8: Peptide quantitation and sequencing using MRM. (a) TIC of the WT tryptic digest, with the peak showing significant change denoted *. Inset of panel (a) shows an expanded view of this denoted peak with a precursor *m/z* ratio of 703.4^{+2}. (b) MRM transitions comparing the *y5* fragment ion between WT and *DOT1L* mutant analyses. (c) MS/MS sequencing of *m/z* 703.4^{+2} confirming the identity of H3K79me1.

4.5 Conclusion

This study demonstrated the successful development and implementation of a novel tEMC trapping mode for profiling intact proteins and their modifications on the 4000 Q-Trap mass spectrometer. A combination of a segmented ion gating strategy and mass selective charge separation were successfully applied to decrease space-charge effects in the LIT. This approach also resulted in higher mass resolution enabling the separation of different protein isoforms. This was demonstrated for protein standards and complex mixtures of acid-extracted proteins that contained, among other proteins, core histones with a wide diversity of sequence variants and PTMs. Our method provided

enhanced resolution and sensitivity compared to conventional quadrupole or EMC scans. The improved resolution of the tEMC mode facilitated the profiling of intact histones in wild- type and *DOT1L* mutant cells to determine global changes in modifications across sample sets and to establish structure-function relationships. This was evidenced in the present report for the profiling of histone modifications in wild-type and *DOT1L* mutant chicken DT40 cells, where changes in the methylation pattern could be identified in only 10% of the histone H3 population. This methodology can be of practical use to enhance the performance of existing trapping instruments in laboratories where an increase in sensitivity and resolution are required. The scan mode modifications enabled automated acquisition of LC-MS data for enhanced protein detection and did not require any hardware additions or adjustments. Finally, using conventional bottom-up sequencing and predictive software tools (MRM pilot), we were able to localize differentially expressed histone H3 PTM and quantitate their relative abundance

4.6 Acknowledgment

N.L. and J.F. gratefully acknowledge support from Science Foundation Ireland (Grant 07/In.1/B958), Cancer Research Ireland (Grant CR105GRE), and the Irish Research Council for Science, Engineering, and Technology (Grant RS/2003/33). The Institute for Research in Immunology and Cancer (IRIC) receives infrastructure support funds from the Fonds de la Recherche en Santé du Québec (FRSQ) and from Canadian Institutes for Health Research (CIHR) multiresource grant. This work was carried out with the financial support of operating grants from the National Science and Engineering Research Council (NSERC), Strategic Grant Initiative (P.T.), the Canada Research Chair program (A.V., P.T.), and the CIHR (A.V.).

4.7 References

1. Walsh, C.T., *Posttranslational Modification of Proteins: Expanding Nature's Inventory*, ed. J. Murdzek. 2006: Roberts and Company. 8-10.

2. Yates, J.R., 3rd, *Mass spectrometry. From genomics to proteomics.* Trends Genet, 2000. **16**(1): p. 5-8.

3. Wysocki, V.H., et al., *Mass spectrometry of peptides and proteins.* Methods, 2005. **35**(3): p. 211-22.

4. Wolters, D.A., M.P. Washburn, and J.R. Yates, 3rd, *An automated multidimensional protein identification technology for shotgun proteomics.* Anal Chem, 2001. **73**(23): p. 5683-90.

5. McLafferty, F.W., et al., *Top-down MS, a powerful complement to the high capabilities of proteolysis proteomics.* Febs J, 2007. **274**(24): p. 6256-68.

6. Kelleher, N.L., *Top-down proteomics.* Anal Chem, 2004. **76**(11): p. 197A-203A.

7. Aebersold, R. and M. Mann, *Mass spectrometry-based proteomics.* Nature, 2003. **422**(6928): p. 198-207.

8. Erickson, B., *Linear ion trap/Orbitrap mass spectrometer.* Anal Chem, 2006. **78**(7): p. 2089.

9. Macek, B., et al., *Top-down protein sequencing and MS3 on a hybrid linear quadrupole ion trap-orbitrap mass spectrometer.* Mol Cell Proteomics, 2006. **5**(5): p. 949-58.

10. Nemeth-Cawley, J.F. and J.C. Rouse, *Identification and sequencing analysis of intact proteins via collision-induced dissociation and quadrupole time-of-flight mass spectrometry.* J Mass Spectrom, 2002. **37**(3): p. 270-82.

11. Nemeth-Cawley, J.F., B.S. Tangarone, and J.C. Rouse, *"Top Down" characterization is a complementary technique to peptide sequencing for identifying protein species in complex mixtures.* J Proteome Res, 2003. **2**(5): p. 495-505.

12. Schirm, M., et al., *Identification of unusual bacterial glycosylation by tandem mass spectrometry analyses of intact proteins.* Anal Chem, 2005. **77**(23): p. 7774-82.

13. Stephenson, J.L., et al., *Ion/ion chemistry as a top-down approach for protein analysis.* Curr Opin Biotechnol, 2002. **13**(1): p. 57-64.

14. Reid, G.E., et al., *Charge-state-dependent sequence analysis of protonated ubiquitin ions via ion trap tandem mass spectrometry.* Anal Chem, 2001. **73**(14): p. 3274-81.

15. Reid, G.E., et al., *Gas-phase concentration, purification, and identification of whole proteins from complex mixtures.* J Am Chem Soc, 2002. **124**(25): p. 7353-62.

16. Claverol, S., et al., *Characterization of protein variants and post-translational modifications: ESI-MSn analyses of intact proteins eluted from polyacrylamide gels.* Mol Cell Proteomics, 2003. **2**(8): p. 483-93.

17. Walz, J., et al., *Ion storage in the rf octupole trap.* Phys Rev A, 1994. **50**(5): p. 4122-4132.

18. Cooks, R.G., C.D. Cleven, and K.A. Cox, *Mass shifts and local space charge effects observed in the quadrupole ion trap at higher resolution.* International Journal of Mass Spectrometry and Ion Processes, 1995. **144**(1-2): p. 47-65.

19. Dobson, G., et al., *Investigation into factors affecting precision in ion trap mass spectrometry using different scan directions and axial modulation potential amplitudes.* J Mass Spectrom, 2004. **39**(11): p. 1295-304.

20. Schwartz, J.C., M.W. Senko, and J.E. Syka, *A two-dimensional quadrupole ion trap mass spectrometer.* J Am Soc Mass Spectrom, 2002. **13**(6): p. 659-69.

21. Hager, J.W., *A new linear ion trap mass spectrometer.* Rapid Communications in Mass Spectrometry, 2002. **16**: p. 512-526.

22. Scherperel, G., et al., *'Top-down' characterization of site-directed mutagenesis products of Staphylococcus aureus dihydroneopterin aldolase by multistage tandem mass spectrometry in a linear quadrupole ion trap.* Analyst, 2006. **131**(2): p. 291-302.

23. Coon, J.J., et al., *Protein identification using sequential ion/ion reactions and tandem mass spectrometry.* Proc Natl Acad Sci U S A, 2005. **102**(27): p. 9463-8.

24. Parks, B.A., et al., *Top-down proteomics on a chromatographic time scale using linear ion trap fourier transform hybrid mass spectrometers.* Anal Chem, 2007. **79**(21): p. 7984-91.

25. Londry, F.A. and J.W. Hager, *Mass selective axial ion ejection from a linear quadrupole ion trap.* J Am Soc Mass Spectrom, 2003. **14**(10): p. 1130-47.

26. Le Blanc, J.C., et al., *Unique scanning capabilities of a new hybrid linear ion trap mass spectrometer (Q TRAP) used for high sensitivity proteomics applications.* Proteomics, 2003. **3**(6): p. 859-69.

27. Thompson, B.A. and I.V. Chernushevich, *A New Scan Mode for the Identification of Multiply Charged Ions.* Rapid Commun Mass Spectrom, 1998. **12**: p. 1323-1329.

28. Chernushevich, I.V., et al., *Charge state separation for protein applications using a quadrupole time-of-flight mass spectrometer.* Rapid Commun Mass Spectrom, 2003. **17**(13): p. 1416-24.

29. Dyer, P.N., et al., *Reconstitution of nucleosome core particles from recombinant histones and DNA.* Methods Enzymol, 2004. **375**: p. 23-44.

30. Shechter, D., et al., *Extraction, purification and analysis of histones.* Nat Protoc, 2007. **2**(6): p. 1445-57.

31. Drogaris, P., et al., *Comprehensive profiling of histone modifications using a label-free approach and its applications in determining structure-function relationships.* Anal Chem, 2008. **80**(17): p. 6698-707.

32. Luger, K., et al., *Characterization of nucleosome core particles containing histone proteins made in bacteria.* J Mol Biol, 1997. **272**(3): p. 301-11.

33. Huyen, Y., et al., *Methylated lysine 79 of histone H3 targets 53BP1 to DNA double-strand breaks.* Nature, 2004. **432**(7015): p. 406-11.

34. Okada, Y., et al., *hDOT1L links histone methylation to leukemogenesis.* Cell, 2005. **121**(2): p. 167-78.

35. Aihara, H., et al., *Nucleosomal histone kinase-1 phosphorylates H2A Thr 119 during mitosis in the early Drosophila embryo.* Genes Dev, 2004. **18**(8): p. 877-88.

36. Mollah, S., et al., *Targeted mass spectrometric strategy for global mapping of ubiquitination on proteins.* Rapid Commun Mass Spectrom, 2007. **21**(20): p. 3357-64.

37. Unwin, R.D., et al., *Multiple reaction monitoring to identify sites of protein phosphorylation with high sensitivity.* Mol Cell Proteomics, 2005. **4**(8): p. 1134-44.

38. *US FDA Guidelines for Bioanalytical Method Validation*. 08-Feb-2011; http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm070107.pdf.

5. Clinically Relevant Histone Deacetylase Inhibitors Enhance Histone H3/H4 Acetylation More Readily in Transformed Cells than in Normal Cells

Paul Drogaris, Valérie Villeneuve, Christelle Pomiès, Eun-Hye Lee, Véronique Bourdeau, Eric Bonneil, Gerardo Ferbeyre, Alain Verreault, and Pierre Thibault

Journal of Proteome Research, 2010, submitted

5.1 Abstract

Histone deacetylase inhibitors (HDACi) represent a promising avenue for cancer therapy. However, in order to improve their clinical efficacy, it is important to obtain a more detailed representation of how HDACi affect histone acetylation in normal and cancer cells. Towards this goal, we applied a two-pronged mass spectrometric strategy to determine the impact of clinically relevant HDACi on histone acetylation in human cells. Intact histone profiling revealed that H3 and H4 acetylation strongly increased after treatment of several transformed cell lines with a number of HDACi, whereas H2A and H2B only showed small changes in acetylation. Our LC-MS/MS analyses indicated that HDACi dramatically increased the acetylation of several residues in the terminal tails of H3 and H4. Absolute quantification with a calibration curve spiked with the synthetic peptide standard and targeted mass spectrometry approaches demonstrated that H3K56 acetylation (H3K56ac) is rare in transformed human cells and, unlike other sites of acetylation, the abundance of H3K56ac did not increase upon prolonged incubation with inhibitors of each class of HDAC. In addition, we found that SAHA, a potent HDACi currently used for treatment of cutaneous T-cell lymphoma, globally increased H3 and H4 N-terminal tail acetylation in transformed cells. Remarkably, SAHA had little effect on histone acetylation in normal human fibroblasts. These results demonstrate that, at least under certain conditions, SAHA (and possibly other HDACi) globally affects histone acetylation in transformed cells while having minimal effects in normal cells.

Keywords

Deacetylase, HDAC inhibitor, HDACi, SAHA, MS-275, post-translational modifications, histone, nano-LC/MS/MS

5.2 Introduction

Eukaryotic DNA is packaged into a nucleoprotein structure known as chromatin. The primary function of this structure is to compact DNA within the cell nucleus. The fundamental repeating unit of chromatin is the nucleosome core particle (NCP), which is composed of 147 bp of DNA wrapped nearly twice around the surface of an octamer of small basic proteins known as histones [1]. This octamer is formed from two molecules each of four types of core histones: H2A, H2B, H3, and H4. Although the structure of chromatin is inherently dynamic, it also considerably restricts access to genetic information [2]. Cells have evolved mechanisms to either control access to DNA packaged into chromatin. These include ATP-driven nucleosome remodelling machines, histone variants and histone post-translational modifications (PTMs) [3]. Each core histone consists of a conformationally flexible N-terminal extension, commonly referred to as "histone tail", and a globular domain that mediates protein-protein interactions and DNA binding within NCPs [1]. The N-terminal tails of core histones protrude beyond the DNA gyres of NCPs, and are therefore readily accessible to histone-modifying enzymes [1]. These tail domains contain multiple residues that are covalently modified by a wide diversity of PTMs [3]. Advances in mass spectrometry (MS) have also uncovered a number of PTMs in the globular domains of core histones [4-6]. Histone PTMs occur either alone or in combination, leading to the notion of a "histone code", a generic term used to describe how specific patterns of histone PTMs control the binding of effector proteins to NCPs and, as a result, control a variety of processes that require access to DNA [7].

The most intensively studied histone PTMs include acetylation, methylation, and phosphorylation [3]. Modifications such as ADP ribosylation, ubiquitylation and sumoylation also occur on histones, but the modified residues are poorly defined. The addition and removal of an acetyl group on lysine residues, respectively mediated by histone acetyltransferases (HATs) and deacetylases (HDACs), is a highly dynamic and regulated process that contributes to transcriptional activation and silencing [8]. In humans, there are 18 known HDAC enzymes that are subdivided into four classes [9].

The seminal discovery that the histone deacetylase inhibitor (HDACi) trichostatin A (TSA) triggered cell cycle arrest and differentiation of Friend erythroleukemic cells promoted numerous subsequent studies on various types of cancer cells [10]. In diseases such as cancer, HDACs contribute to oncogenesis via at least two distinct mechanisms: overexpression of individual HDACs [11] and aberrant recruitment of HDACs to specific chromosomal loci by oncogenic fusion proteins (*e.g.* PML-RARα) [12] These perturbations of HDAC function often lead to silencing of tumor suppressor genes. Changes in the abundance of specific histone PTMs also occur in cancer cells. For example, a study of numerous cancer cell lines, normal tissues and primary tumors revealed that a global decrease in H4K16 acetylation and H4K20 methylation is a recurring feature of many types of cancers [13].

The involvement of HDACs in cancer has raised hope that these enzymes may represent valuable targets in drug discovery programs. Recent clinical trials demonstrated that, at least for hematological cancers [14-16], small molecules that inhibit HDACs can be effective pharmacological agents, either when administered alone or in combination with other drugs [17]. Through their ability to influence gene expression, HDACi exhibit a number of anti-proliferative effects, such as cell cycle arrest, differentiation, angiogenesis inhibition and apoptosis [9, 18]. In hematological malignancies such as leukemia, the oncogenic effects of fusion proteins encoded by chromosomal translocations implicating retinoid receptors are inhibited by HDACi [19]. A significant number of HDACi including suberoylanilide hydroxamic acid (SAHA, also known as vorinostat or *Zolinza*), entinostat (MS-275), romidepsin (*Istodax*), and belinostat (PXD-101) are at various stages of drug development. At present, only vorinostat [20] and romidepsin [21] have been approved by the US food and drug administration for treatment of cutaneous T-cell lymphoma. HDACi such as SAHA are non-selective [22]. SAHA and many other clinically promising HDACi that contain hydroxamic acid moieties inhibit multiple HDACs because they chelate a zinc divalent cation that is essential for activity of class I and class II HDACs [18]. There is a growing debate over the use of pan-HDACi that cause non-specific inhibition of several distinct HDACs versus development of HDACi that would target specific classes of enzymes or even

perhaps single enzymes. In addition to histones, HDACs deacetylate many important protein substrates, such as transcription regulators (*e.g.* p53, Rb, E2F1, nuclear hormone receptors) [18]; hence, inhibiting multiple HDACs could be cytotoxic to normal cells and lead to undesirable side effects. On the other hand, it can be argued that pan-HDACi are effective at killing cancer cells precisely because they interfere with the deacetylation of multiple substrates. Currently, it is not known whether inhibition of histone deacetylation, as opposed to other protein substrates, plays a major role in the anti-neoplastic effects of pan-HDACi. As a first step to address these difficult questions, it is important to determine to what extent HDACi impair histone deacetylation in normal and cancer cells.

The ability to identify and profile stoichiometric changes in histone modifications is of paramount importance to ascertain the pharmacodynamic effects of HDACi and to establish structure-function relationships between potential drug candidates and their enzyme targets. Moreover, a better understanding of HDACi-induced changes in the abundance of specific histone acetylation sites could lead to biomarker discovery for therapeutic drug monitoring [23]. Immunoblotting has been used extensively to probe changes in histone modifications under a wide variety of physiological conditions [3]. However, in our experience, many antibodies are not absolutely modification site-specific and often are prone to cross-react with sites that lie within similar primary sequence contexts in the same type of core histone molecules. In addition, even with the use of fluorescent dye-labelled antibodies, immunoblotting has a very limited linear range compared with mass spectrometry-based (MS) approaches. Thus, MS has become the method of choice to detect and profile changes in histone PTMs due to its speed, sensitivity, and selectivity. There are three main requirements needed for successful characterization of histone PTMs in a reproducible and comprehensive fashion: a tailored sample preparation strategy, modern LC-MS instrumental platform(s), and bioinformatic tools for rapid data mining. Several manuscripts that analyzed and quantified histone PTMs were deficient in one of these three aspects. An early report by Smith and colleagues described the application of a low resolution LCQ ion trap to profile global H4 acetylation levels in yeast cells [24]. Hunt and colleagues pioneered the identification and quantitation of histone PTMs in various species using a combination of chemical

modification, and high-resolution tandem MS [25-27]. Although this workflow is rugged and lead to many pioneering insights into histone biology, it suffers from many sample preparation steps: multiple rounds of propionylation and chemical labelling that can lead to significant sample loss and side reactions. Another drawback of this workflow was that it relied solely on manual investigation of raw data and lacked bioinformatic tools. This renders any high throughput data analysis impossible. The groups of Jensen and Hendrickson utilized a high-resolution tandem MS instrument combined with database search engines and ion profiling software to address the combinatorial complexity of histone modifications by LC-MS/MS [28-30]. However, their sample preparation strategy utilized direct trypsin digestion, which, unlike derivatization of free lysines, precludes accurate determination of acetylation stoichiometry.

In this report, we conducted a comprehensive label-free quantitative proteomics study to profile histone modifications in normal and cancer cells treated with different HDACi treatment. Our experimental strategy included a two pronged-approach that we previously developed and validated to identify histone PTMs whose abundance changes between wild-type and yeast HAT mutant strains [31]. Global changes in histone modifications are first determined from the mass profiles of intact histones. In a second step, the modification sites and their respective stoichiometric changes are established by a combination of propionylation, trypsin digestion and LC-MS/MS analyses. Bioinformatic tools are exploited to rapidly sift through vast amounts of raw MS data. In this study, we also designed a targeted approach using multiple reaction monitoring (MRM), and high-resolution MS combined with absolute quantitation with a synthetic peptide and a [$^{13}C_6{}^{15}N_1$]-labelled internal standard peptide, to detect and accurately quantify the acetylation of specific amino acid residues whose stoichiometry of acetylation is very low.

5.3 Experimental Section

5.3.1 Chemicals and Other Reagents

Trichostatin A (TSA, Sigma), valproic acid (Sigma), nicotinamide (Sigma), nicotinic acid (Sigma), phenyl butyrate (Calbiochem) were solubilized in water. MS-275 (Calbiochem) and Vorinostat (SAHA, Cayman chemical) were solubilized in DMSO. Human H3K56ac peptide (YQKacSTELLIR) and its isotopically-labeled analogue (YQKacSTELL[$^{13}C_6{}^{15}N_1$]-IR) were synthesized by Midwest Biotech (Fishers, IN). HPLC columns for histone fractionation were purchased from Phenomenex (Torrance, CA) and Canadian Life Sciences. Pre-columns (4 mm x 360 μm I.D.) and nanoLC-MS/MS columns (100 mm x 150 μm i.d.) were packed in-house using Jupiter C_{18}, 3 μm particles (Phenomenex) and fused silica tubing (Polymicro Technologies). Milli-Q water and Optima grade acetonitrile (Fisher Scientific) were used as HPLC solvents. Trifluoroacetic acid (TFA), ammonium bicarbonate, and propionic anhydride were obtained from Sigma. Formic acid was obtained from EMD Science. Porcine modified trypsin (sequencing grade) was obtained from Promega. The protein content in histone extracts was measured using the Micro-BCA assay (Thermo Fisher Scientific). Cell viability was assessed using the alamar blue assay kit (Invitrogen).

5.3.2 Cell Culture and Histone Preparation

K562 cells were cultured in RPMI-1640, IMR90 and WI-38 cells in DMEM with 10% FBS and 1% Pen-Strep in a 5% CO_2 incubator at 37°C. For experiments with HDACi, cells were seeded in 150 cm^2 dishes at 0.5 million cells/mL in 25 mL of culture medium. Cells treated with HDACi were collected and lysed in hypotonic lysis buffer (10 mM Tris-HCl pH 7.5, 3mM $CaCl_2$, 2mM $MgCl_2$, 1 mM DTT, 1X EDTA-free protease and phosphatase inhibitor cocktail, 1 mM PMSF, 10 mM sodium butyrate, 10 μM trichostatin A and 10 mM nicotinamide). Cells were lysed in a metallic dounce homogenizer and nuclei were collected by centrifugation (13,500 rpm, 15 min, 4 ºC). The

nuclei were then resuspended in 0.4M H_2SO_4 to extract histones as previously described [32]. Histones were dissolved in 0.2% formic acid and stored frozen at -80° C.

5.3.3 Immunoblotting

Histone proteins were resolved on SDS-polyacrylamide gels and transferred onto PVDF membranes. After transfer, the presence of equal amounts of core histones was assessed using either Ponceau S staining or by immunoblotting using an antibody against non-modified histone H4. The membranes were then probed with the primary antibodies against H3K56ac (Upstate/Millipore 07-677 or Epitomics 2134-1), H4K12ac (Cell Signaling 2591S) and H4K16ac (Upstate/Millipore 07-329) overnight at 4°C. The blots were then washed extensively in 1X Tris-buffered saline (25mM Tris-HCl pH 8.0, 137mM NaCl, 3mM KCl). After incubation with secondary antibodies (anti-rabbit IgG coupled to horseradish peroxidase) for one hour at room temperature, the blots were washed extensively in 1X Tris-buffered saline (25mM Tris-HCl pH 8.0, 137mM NaCl, 3mM KCl) containing 0.1% (w/v) Tween-20, and antibody binding was detected by enhanced chemiluminescence (GE Healthcare).

5.3.4 Flow Cytometry and Cell Viability

After HDACi treatments, 1 million cells were washed with cold PBS and fixed in a 70% cold ethanol-PBS solution overnight at -20ºC. Cells were washed three times with PBS and DNA stained with a propidium iodide (PI) solution (50 µg⁄ml PI, 0.1% sodium citrate, and 0.2 mg⁄ml ribonuclease A) for 30 min at 4°C. DNA content was measured by flow cytometry (BD Biosciences). Viability of cells after HDACi treatment was assessed using the Alamar Blue cell viability assay (Invitrogen). Cells were seeded in a 96-well plate at 5000 cells ⁄ well and incubated for three hours at 37ºC with 10% Alamar Blue. Absorbance was read at 570 and 600 nm to determine cell viability by calculating the percentage of Alamar Blue reduction [33].

5.3.5 Sample Preparation and Protein Digestion

The total protein content of enriched histone samples was measured with the Micro-BCA assay. Approximately 15 μg of starting material was separated using an Agilent 1100 HPLC system equipped with a micro-fraction collector. Separations were performed on a Jupiter microbore C_{18} column (3 μm, 300 Å; 150 x 1 mm i.d.) with a solvent system consisting of 0.1% TFA in water (v/v) (solvent A), and 0.1% TFA in acetonitrile (v/v) (solvent B). Gradient elution was performed from 5 to 70% B in 60 min at 15 μL/min. Fractions were collected in a 96-well plate at 20 second intervals and pooled together to combine individual histone peaks. For the targeted analysis of H3K56ac, approximately 50 μg of starting material was separated using an Agilent 1200 HPLC system equipped with a fraction collector. Separations were performed on an ACE standard bore C_8 column (5 μm, 300 Å; 150 x 4.6 mm i.d.) at 0.7 mL/min. Histone H3 fractions were collected in 2 mL Eppendorf tubes at 1 minute intervals. Multiple fractions were pooled together as described above. Histones were prepared and digested as previously described [31]. LC fractions were dried in a Speed-Vac evaporator and resuspended in 0.1 M ammonium bicarbonate buffer (without pH adjustment). Intact histones were derivatized in 1:1 volume ratio with a propionic anhydride reagent. After a 30 minute derivatization period at room temperature, samples were evaporated a second time, resuspended in 0.1 M ammonium bicarbonate, and digested overnight at 37° C using 1 μg of trypsin. Samples were acidified with 5% TFA in water (v/v) prior to LC-MS analysis. Under our conditions, propionylation proceeded nearly to completion, as judged by negligible trypsin cleavage at non-propionylated lysine residues, which we verified by MS. Near complete propionylation of non-modified lysine residues is important to determine the accurate stoichiometry of histone acetylation because trypsin would otherwise cleave after non-modified lysine. This would obviously lead to an under-estimation of the fraction of histone molecules that are not acetylated at a given site.

5.3.6 Intact Histone Profiling

Intact histone analysis was performed on an Agilent 6520 Q-TOF mass spectrometer equipped with a nanoelectrospray interface. The instrument was operated in positive ion mode (2800 V at interface) and scanned from *m/z* 400 to 1600. Equivalent amounts of histone acid extracts were diluted down to a concentration of 100 ng/µL with initial LC-MS mobile phase (see below). An aliquot was injected onto an in-house packed capillary C_{18} trapping column (4 mm x 360 µm I.D.) for 3 min at 10 µL/min using a loading solvent (95:5 water:acetonitrile containing 0.2% formic acid (v/v)). Histones were eluted onto the C_{18} analytical column (100 mm x 150 µm i.d.) using 0.2% formic acid in water (v/v) (solvent A), and 0.2% formic acid in acetonitrile (v/v) (solvent B). Gradient elution was performed from 5 to 90% B in 60 min at 600 nL/min. Intact histone mass profiles were generated from acquired mass spectra using the MaxEnt1 deconvolution algorithm in Mass Hunter software.

5.3.7 Ion Profiling from LC-MS/MS Analysis and Peptide Sequencing

LC-MS/MS analyses were performed using a Thermo Electron LTQ-Orbitrap XL mass spectrometer equipped with a nanoelectrospray interface. Tryptic digests were injected in triplicate using an Eksigent nano 2DLC system. The same columns and solvent systems were used as described above in the intact profiling section. The instrument was operated in positive ion mode and scanned from *m/z* 300 to 1600, resolution set to 60 000, with a target value of 1.0×10^6. For tandem MS/MS sequencing, the LTQ ion trap was scanned from *m/z* 50 to 2000, with a target value of 1.0×10^4. One full survey scan in the Orbitrap was followed by data dependant MS/MS acquisition on the three most intense ions. A decision tree based on peptide *m/z* ratio and charge state was used to trigger either CID or ETD fragmentation. A normalized collision energy of 35 was used to generate fragment ions in the LTQ by CID, while a 100 ms activation time was used for ETD ion-ion reactions.

All raw data files (.raw) from triplicate injections on the LTQ-Orbitrap XL mass spectrometer were converted into peptide maps using Mass Sense (version 2.5), an in-house peptide detection software used to identify peptides based on *m/z* values, retention time, abundance, and charge state. An intensity threshold of 10 000 counts was set as a cut-off for precursor ion peak detection. Segmentation analysis was performed by clustering identified peptide ions across each treatment condition based on their respective *m/z* ratio, charge, and LC retention time using an *m/z* tolerance of ± 0.02 Th and a time difference of 1 min. Normalization of LC retention time is then performed on the initial peptide cluster list using a dynamic and nonlinear correction. Once clustered, all peptide intensities were averaged and a mean ratio was calculated to generate a correction factor used to normalize for changes in abundance and differences in amount of peptide loaded. Normalized intensities are then used to generate a volcano plot for differential histone PTM expression profiling. Raw data files were converted to Mascot generic files (.mgf) using *Mascot Distiller* (version 2.2.1) and *Mascot Daemon* (version 2.2.2). A mass tolerance of 0.02 Th was used for the precursor ions, and 0.5 Th for the fragment ions. A maximum of 9 missed cleavages was permitted during the Mascot searches. The following variable modifications were included in the search parameters: acetylation, methylation+propionylation, dimethylation, and trimethylation, and propionylation on lysine residues, methylation on arginine residues, phosphorylation on serine, threonine and tyrosine residues, and oxidation on methionine residues. Manual MS/MS spectra verification was performed on each modified peptide with Mascot scores greater than 20 to confirm the sequence assignment. Mascot searches were downloaded from the Mascot server into an Excel spreadsheet, from which redundant and low scoring peptides were removed. All MS/MS spectra corresponding to modified histone tryptic peptides are provided as supplementary information (http://pubs.acs.org/journal/jprobs).

5.3.8 Targeted Analysis of H3K56ac by Absolute Quantitation and Multiple Reaction Monitoring

Stock solutions at 1 mg/mL in deionized water were prepared for both K56-acetylated synthetic peptides: a peptide containing only ^{12}C and a peptide where a leucine

was fully [^{13}C]-labelled (see below). Unmodified yeast histone H3 was expressed in *E. coli* and purified as described [34], and then resuspended in 0.1 M ammonium bicarbonate buffer. Approximately 3 µg of recombinant yeast histone H3 was derivatized, digested, and spiked with increasing amounts of synthetic human H3K56ac peptide (YQK(ac)STELLIR) ranging from 5 to 1000 pg/mL. To each solution, a constant amount of the [^{13}C-leucine] labelled internal standard (IS) peptide (YQK(ac)STELLI*R, wherein I* is a [$^{13}C_6{}^{15}N_1$]-labelled isoleucine), was added to obtain a final concentration of 100 pg/mL IS in each vial. A 20-µL aliquot of each mixture was injected onto the LTQ-Orbitrap XL mass spectrometer, and analyzed using the same parameters as described in the LC-MS/MS analysis section.

A multiple reaction monitoring (MRM) method was designed to detect and quantify human H3K56ac tryptic peptide using an AB Sciex 4000 Q-trap hybrid triple quadrupole-linear ion trap mass spectrometer equipped with a Nanospray II interface. Working solutions of human K56ac peptide and its labelled analogue were prepared at a concentration of 10 µg/mL in 50:50 water:methanol with 0.2% formic acid (v/v/v) for nanoelectrospray infusion. Source parameters and ion optic voltages were optimized by constant infusion of the working solutions at 600 nL/min. The electrospray voltage was set to 3800V, with a declustering potential of 120V, and an optimized collision energy of 32V. The four most intense ions were chosen to build MRM transitions within Analyst software. Both H3K56ac and its non-modified counterpart (H3K56pr) were monitored in parallel; another H3 tryptic peptide (YRPGTVALR) was also monitored to normalize for the amount of peptide loaded and fluctuations in the electrospray response.

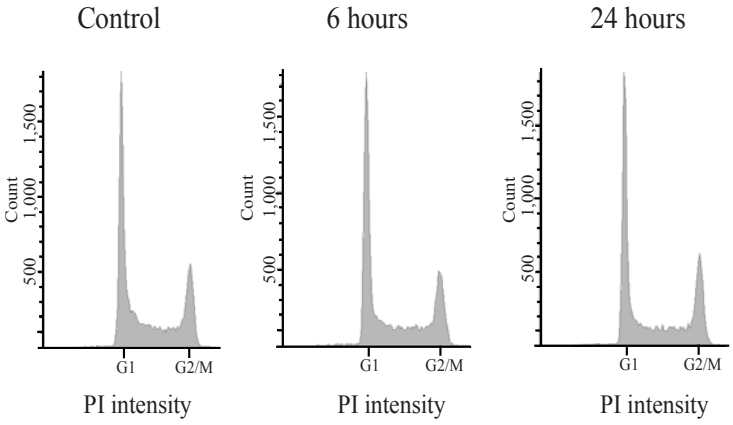
5.4 Results and Discussion

5.4.1 Cell Viability and Cell Cycle Progression in Response to HDACi

HDACi can trigger apoptosis or cell cycle arrest in transformed cell lines [9, 18]. However, we wanted to analyze changes in histone acetylation under conditions where HDACi-treated cells did not lose viability or cease to proliferate. This was important

because newly synthesized histones H3 and H4 are acetylated at several lysine residues and incorporated into chromatin throughout the genome during S-phase [35-38]. Because new H3/H4 molecules are very abundant (new histones represent 50% of total histones in G2 cells), they likely make a significant contribution to global levels of acetylation in proliferating cells. We first examined whether HDACi treatment caused cell cycle arrest using propidium iodide staining of DNA and flow cytometry. Asynchronously growing control K562 (human erythroleukemic) cells show populations of cells distributed in three distinct regions of fluorescence intensity, corresponding to G1, S, and G2/M phases (figure 5.1a). From the relative proportion of the peaks, we estimated that no more than 20% of the cells were in S-phase. No significant change in the FACS profile was noted for K562 cells exposed to 1 μM MS-275 or SAHA (data not shown) for a period of up to 24 h (figure 5.1a), showing that these two HDACi did not induce any obvious cell cycle arrest. Consistent with this, cell numbers doubled during the 24h period when they were exposed to either 1μM MS-275 or 1μM SAHA. Cell viability was monitored using Alamar Blue, a non-fluorescent chemical that is converted into fluorescent Resorufin upon metabolic reduction [39]. Control cells and cells treated with either 1μM MS-275 or 1μM SAHA for periods up to 24 h showed no obvious difference in viability (figure 5.1b). Thus, exposure to 1μM MS-275 or 1μM SAHA for 24h did not cause appreciable cell cycle arrest or lethality in K562. Our results are consistent with the fact that the same concentrations of MS-275 or SAHA were not cytotoxic to several other human cancer cell types [22].

a)



b)

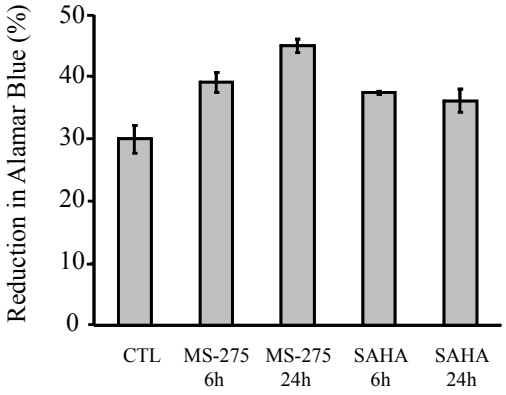


Figure 5.1: Cell cycle distribution and viability of K562 cells are not affected by treatment of MS-275. (a) Propidium iodide (PI) staining of DNA detected by flow cytometry analyses of K562 cells treated for 6 hour- and 24 hour-treatment with 1 μM of the class I HDAC inhibitor MS-275. (b) Cell viability of K562 cells after 6- or 24-hour continuous exposure to either 1μM SAHA or 1μM MS-275. Cell viability was measured by staining with Alamar Blue as described in the experimental section.

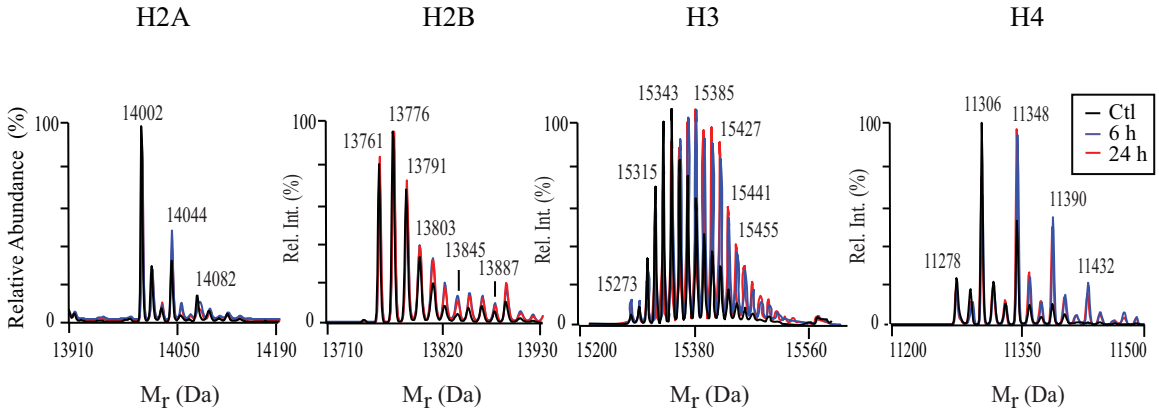
5.4.2 Global Histone Acetylation Probed by Intact Mass Profiling

We first examined the changes in molecular mass distribution of intact histones upon HDACi treatment using LC-MS on the Agilent 6520 Q-TOF mass spectrometer. In each case, a total of 500 ng of acid extracts enriched in histones were injected onto the

LC-MS system. The total ion chromatogram (TIC) of histones from K562 cells that were not treated with HDACi shows that core histones H4, H2B, H2A and H3 eluted successively between 50 and 65 min under our chromatographic conditions (see figure A.1 in appendix 1). The mass spectra of core histones eluted at different time points were deconvoluted and overlaid for comparison. Figures 5.2a and 5.2b show the intact mass profiles of the four core histones from control K562 cells and those exposed to either SAHA or MS-275 for incubation periods of 6 and 24 h. In most eukaryotes, the initiator methionine is removed from the four major types of core histones: H2A, H2B, H3 and H4 [40]. The most prominent peak for H2A is at 14002 Da (figures 5.2a and 5.2b). This is consistent with the calculated molecular mass for the type 1 form of histone H2A (Uniprot P0C0S8) without the initiator methionine, but including acetylation of the N-terminal amino group (Mcalc.: 14002.3 Da). Consistent with our data, N-terminal acetylation occurs in the vast majority of H2A molecules [40]. The most significant change observed for H2A upon treatment with SAHA or MS-275 is a small increase in the abundance of the peak at 14044 Da corresponding to a diacetylated form H2A. However, no significant increase in the abundance of more highly acetylated forms of H2A was observed upon HDACi treatment, and the intact H2A profiles were comparable during a time course from 0h to 24h with both SAHA and MS-275 (figures 5.2a and 5.2b). Many minor sequence variants of histone H2B are encoded in the human genome [41]. In both normal and transformed cells (figures 5.2 and 5.3), the first prominent peak for histone H2B (13761 Da) very closely matched the predicted mass for the type 1H (Mcalc.: 13760.9 Da ; Uniprot Q93079) H2B variant devoid of N-terminal methionine and amino group acetylation. We examined the abundance distribution of H2B isoforms upon treatment with HDACi (figures 5.2a and 5.2b). Another prominent H2B peak at 13791 Da (figures 5.2 and 5.3), which likely corresponds to lysine or arginine mono-methylation of the type 3B H2B variant (Mcalc.: 13790.95 Da). Importantly, the H2B profiles were similar for all HDACi conditions examined, except for an increase in acetylation inferred by in the abundance of peaks at 13803 Da (+1 ac, 5% increase) and 13845 Da (+2 ac, 15% increase). These peaks may correspond to acetylated forms of the type 1H variant, but we cannot rule out the possibility of lysine tri-methylation.

The most significant changes in the intact histone mass profiles were observed for histones H3 and H4 (figures 5.2a and 5.2b). Unlike H2A, H2B and H3, which exist as multiple sequence variants, there is a single H4 protein encoded by several genes [41]. Essentially all H4 molecules are acetylated on their N-terminal amino group [40]. This N-terminally acetylated form of H4 is present at 11278 Da (Mcalc.: 11278.2 Da), along with peaks at 11306 Da (+28 Da, dimethylation) and up to three additional acetylated forms at 11348, 11390, and 11432 Da. All these H4 peaks were detected in both normal and transformed cells (figures 5.2 and 5.3). A rapid increase in mono-, di- and tri-acetylation was observed as early as 6h after addition of SAHA. Similar changes in acetylation patterns were observed in cells treated with MS-275, although the more highly acetylated forms of H4 were apparent only after 24h exposure. Unlike histone H2A and H4, virtually all the H3 molecules lack acetylation of the N-terminal amino group [37]. In rapidly proliferating cells, the replication-coupled histone H3 variants (H3.1 and H3.2) are far more abundant than the replication-independent H3.3 protein [42]. In K562 and normal cells, the unmodified form of H3.2 (Mcalc.: 15256.8 Da) was not detected, but that of H3.1 (Mcalc.: 15272.9 Da) was observed in low abundance at 15273 Da (figures 5.2 and 5.3). The intact mass profile of H3.1 revealed a heterogeneous distribution of molecular species with incremental shifts of +14 and +42 Da, corresponding to multiple forms of methylation and acetylation. Treatment of K562 cells with SAHA or MS-275 caused an overall shift to more highly acetylated forms. However, as observed for H4, the most highly acetylated forms of H3 were accumulated more rapidly with SAHA (6h) than with MS-275 (24h). Based on intact mass profiling, both SAHA and MS-275 predominantly increased the acetylation of H3 and H4, whereas acetylation of H2A and H2B was affected to a much lesser extent.

a) SAHA



b) MS-275

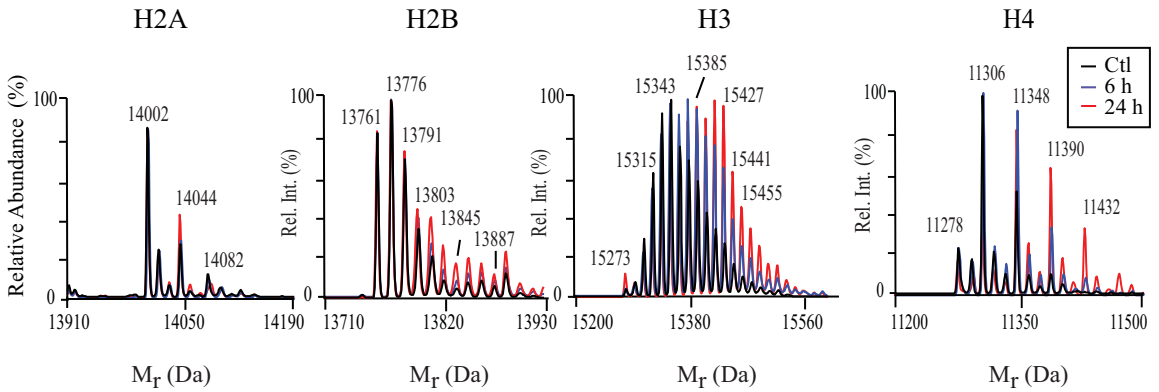


Figure 5.2: LC-MS analyses of intact histones derived from transformed cells. Intact histone profiles from K562 cells treated with SAHA (a) and MS-275 (b). Color lines indicate histones purified from control (black), 6 h (blue) and 24 h (red) treatment with 1μM SAHA or 1μM MS-275.

To determine if the HDACi had the same effect on normal cells, we treated normal human diploid fibroblasts (IMR-90 and WI-38 cells) with the same concentration of SAHA or MS-275 that was used for K562 cells. A longer treatment time was performed for IMR-90 (48h) and WI-38 (72h) because these cells proliferate more slowly than K562 cells. As observed with the K562 cell line, the proliferation and viability of WI-38 cells were not affected by these HDACi at a concentration of 1 μM (data not shown). Interestingly, the intact mass profiles of histones derived from either control or

HDACi-treated IMR-90 and WI-38 cells showed essentially no difference (figure 5.3). The profiles were almost superimposable for each core histone with no significant shift to higher acetylated forms. These results indicate that, at least at the concentration that we used, SAHA and MS-275 can trigger a global increase in H3 and H4 acetylation in cancer cells without significantly affecting global acetylation in normal fibroblasts.

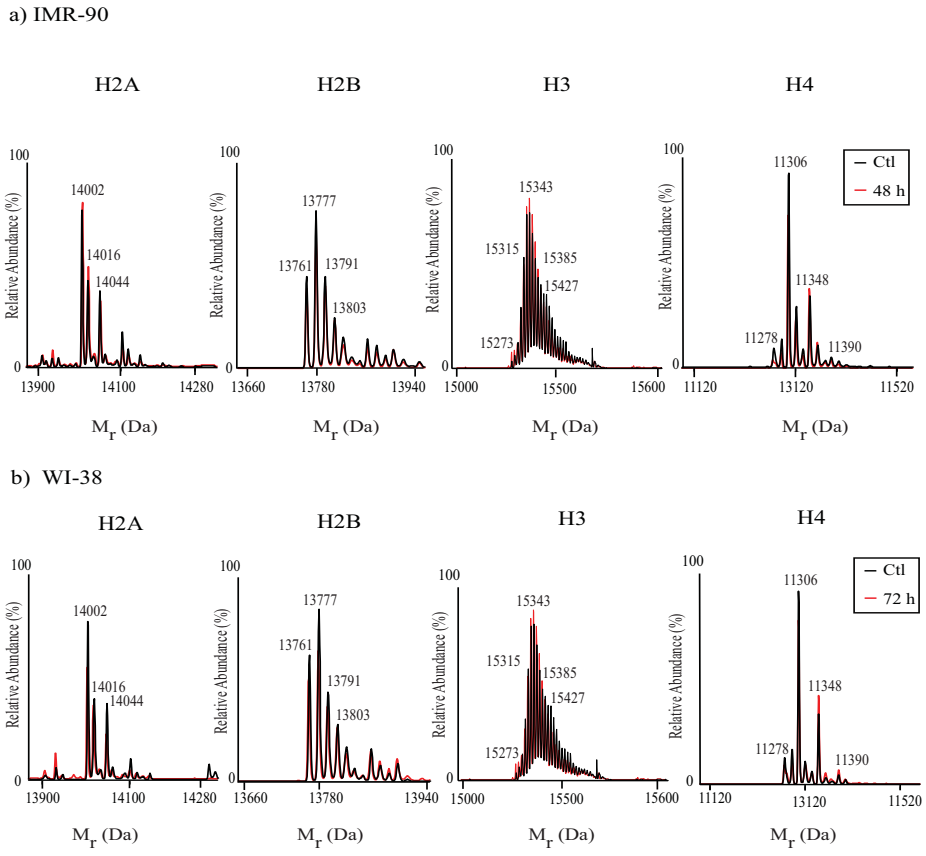


Figure 5.3: LC-MS analyses of intact histones derived from normal diploid fibroblasts. Intact histone profiles for SAHA-treated IMR-90 cells (a) or WI-38 cells (b). Color lines indicate histones purified from control cells (black), and cells with 1μM SAHA for the indicated times (red).

5.4.3 Mapping modification sites by propionylation, tryptic digestion and LC-MS/MS analyses

To determine the location and the extent by which each site is modified upon HDACi treatment, we purified individual histones by reversed-phase HPLC and propionylated free lysine residues prior to tryptic digestion and LC-MS/MS analyses [31]. The amount of protein in acid extracts containing histones was quantitated using the Micro-BCA assay, and a total of 15 μg was separated by HPLC to isolate individual histones. A typical chromatographic separation of histone extracts from K562 control cells resulted in near baseline separation of the four major core histones (see figure A.2 in appendix 1). We focused primarily on the identification of modification sites on H3 and H4, since these core histones showed the most significant global changes upon treatment with either SAHA or MS-275. H3 and H4 histones from all cells and HDACi conditions (n=3) were subsequently derivatized with propionic anhydride, digested with trypsin, and analyzed by LC-MS/MS on the LTQ-Orbitrap XL. Internal calibration using the polysiloxane ion at *m/z* 445.0012 provided *m/z* measurements within 10 ppm of the theoretical mass and facilitated the unambiguous identification of modifications with closely related masses (*e.g.* acetylation versus trimethylation).

The injection of individual histones rather than a complex acid extract digest was preferred to maximize the number of peptides detected and sequenced. This approach resulted in the detection of a multitude of tryptic peptides including those originally containing free and acetylated lysine residues from histones H3 and H4. Peptide maps representing the coordinates of all peptide ions (e.g. *m/z*, charge, LC retention time, peak intensity) were obtained from each raw LC-MS data file and aligned across replicates and conditions to correlate changes in abundance of identified peptide ions. A list of all identified peptides and their abundance is presented in appendix 1 (tables A.1 to A.10). The relative standard deviation observed for peptide ions detected across all three replicate LC-MS/MS analyses varied between 17-24%, demonstrating the reproducibility of the present approach. For all identified H3 and H4 tryptic peptides, the significance and fold change in abundance caused by SAHA and MS-275 treatment is depicted in

volcano plots (figures 5.4a and b). Histone H3 peptides detected after 6 and 24 hours of exposure to HDACi are denoted by triangles and squares, respectively. Histone H4 peptides detected after 6 and 24 hours of exposure to HDACi are denoted by circles and diamonds, respectively. The fold change was determined by calculating the abundance ratio of the same peptide in histones from control and HDACi-treated cells. A large proportion, representing ~90% of all identified peptides, were unaffected by treatment with either SAHA or MS-275. These are clustered near the origin (p-values > 0.05, fold change < 2). Two quadrants are highlighted by dashed boxes in the volcano plots for peptides showing a statistically significant fold change in either control or HDACi treated cells for triplicate measurements (p-values < 0.05, fold change > 2). MS/MS sequencing with either CID or ETD fragmentation revealed that all H3 sites acetylated upon HDACi treatment were located within the first 36 residues of histone H3. Most notably, K_9, K_{14}, K_{18} and K_{23} were primary sites of hyperacetylation upon treatment with either SAHA or MS-275. For example, figure 5.4c shows the CID MS/MS spectrum of precursor peptide *m/z* 493.27^{2+}, corresponding to the tryptic peptide K_9(ac)STGGK_{14}(ac)APR, highlighted by arrows in figure 5.4a. This peptide showed a 5- to 17-fold increase in abundance upon SAHA or MS-275 treatment. Unexpectedly, the dimethylated tryptic peptide K_{27}(me2)SAPATGGVK_{36}(me2)K_{37}(pr)PHR identified by ETD fragmentation increased approximately 2- to 4-fold in abundance under the same conditions (figure 5.4a). The combined used of CID and ETD fragmentation facilitated the localization of modification sites and enabled the identification of 41 differentially modified histone H3 peptides following addition of SAHA or MS-275 to asynchronously growing K562 cells. The use of decision tree driven MS/MS sequencing was used to maximize the protein sequence coverage and identify any peptides that might have been missed if only one fragmentation mode had been used. The majority of ions (70%) were identified using CID as doubly charged species, while ETD identified 30% of ions as triply or more highly charged species. Our LC-MS/MS analysis of histone H3 tryptic peptides resulted in a sequence coverage of 63%, representing 43% of known PTM sites detected.

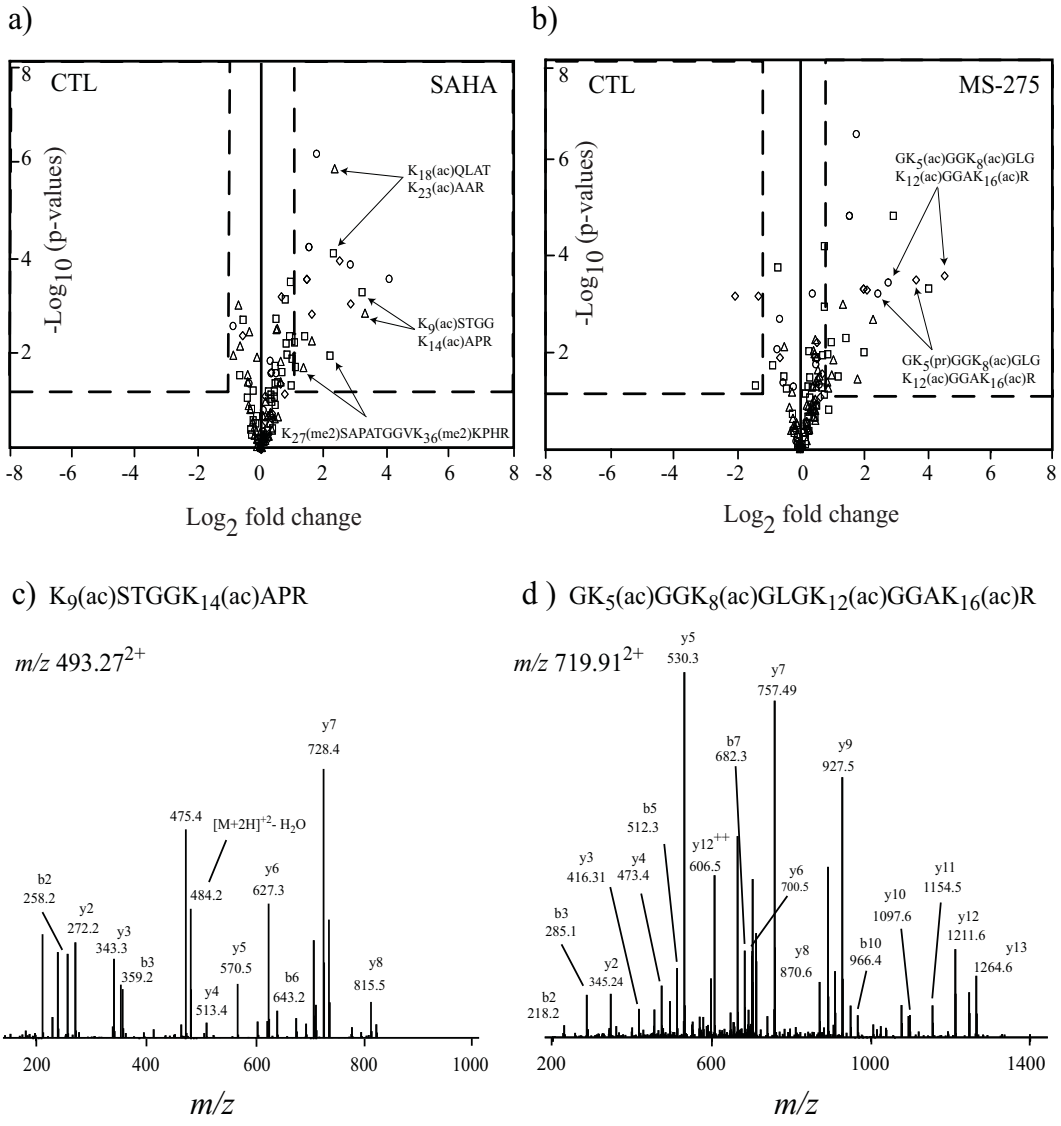


Figure 5.4: Histone H3 and H4 peptide ion intensity profiling using volcano plots. K562 cells were either left untreated (CTL) or treated with either 1µM SAHA (a) or 1µM MS-275 (b). Histone H3 peptides after a 6-hour HDACi treatment are denoted by triangles and those after a 24-hour treatment by squares. Histone H4 peptides after a 6-hour HDACi treatment are denoted by circles and those after a 24-hour treatment by diamonds. Arrows show examples of modified histone peptides whose abundance is significantly affected by HDACi. Tandem MS sequencing of an N-terminal histone H3 peptide (c), and an N-terminal histone H4 peptide (d), revealed acetylation sites that increased in abundance in response to HDACi treatment. See Appendix 1 for intensity measurements used for the generation of volcano plots.

Volcano plots were also generated for histone H4 tryptic peptides to identify residues showing the most significant changes in acetylation after SAHA (figure 5.4a) or MS-275 treatment (figure 5.4b). Following propionylation, tryptic digestion of histone

H4 yielded an N-terminal 14-amino acid peptide (from residues 4 to 17) with four lysine residues (K_5, K_8, K_{12}, and K_{16}). This 14-residue peptide was detected in multiple forms corresponding to different modification patterns. As observed for H3, amino acids located within the N-terminal tail of H4 were the most prominent targets of acetylation. Upon HDACi treatment, we observed a gradual increase in the di-, tri- and tetra-acetylated forms and the fully acetylated peptide clearly dominated after exposure to MS-275 (7- and 24-fold increase at 6h and 24h, respectively). The mono-acetylated form did not show a significant change in abundance after exposure to any HDACi (see Appendix 1). Figure 5.4d shows the MS/MS spectrum for the tetra-acetylated tryptic peptide. Careful examination of the fold changes in acetylation triggered by SAHA and MS-275 revealed important differences between these two inhibitors. Indeed, treatment with 1 μM SAHA yielded a 17-fold increase in acetylation at 6h with a subsequent reduction to 7-fold at 24h, suggesting a limited long-term inhibitory activity of SAHA compared with MS-275. The fact that MS-275 caused a more robust long-term hyper-acetylation than SAHA was also obvious from the intact mass profiles of H4 (figures 5.2a and b). Some of the sites of acetylation whose abundance clearly increased in response to HDACi were also analyzed by immunoblotting (figure 5.5). For example, we found that histone H4K12 and H4K16 acetylation strongly increased in acetylation in transformed cells treated with 1μM SAHA. Consistent with our intact mass analysis, this increase was not observed when normal fibroblasts were treated with SAHA for up to 72h. Finally, Our LC-MS/MS analysis of histone H4 tryptic peptides resulted in a sequence coverage of 75%, representing 46% of known PTM sites detected.

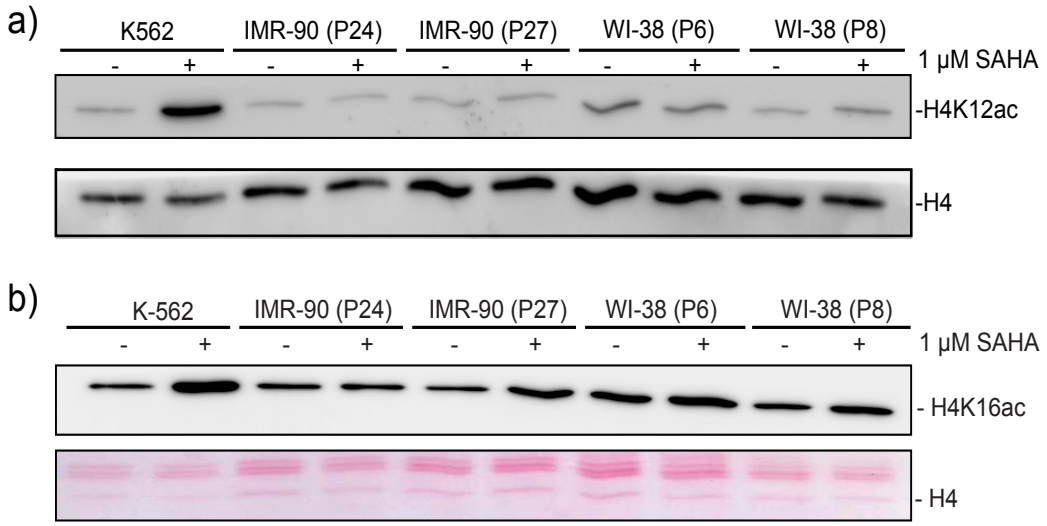


Figure 5.5: 1μM SAHA increases H4K12ac and H4K16ac in transformed cells (K562), but not in normal diploid fibroblasts (IMR-90 and WI-38). For normal fibroblasts, the passage number is indicated. Immunoblots were probed with antibodies against H4K12ac (a) and H4K16ac (b). Similar loadings of histone samples from control (minus signs) and SAHA-treated cells (plus signs) were verified by stripping the upper blot and reprobing it with an antibody that detects unmodified histone H4 (a) or by Ponceau S staining prior to immunoblotting (b).

5.4.4 Abundance of H3K56ac Measured by Targeted MRM and Absolute Quantification

Our profiling of tryptic peptides by LC-MS/MS confirmed that lysine residues located within the N-terminal tails of H3 and H4 were the most prominent acetylation sites following HDACi treatment. Indeed, we observed very few acetylated tryptic peptides within the globular domains of the corresponding histones. This prompted us to examine the stoichiometry of acetylation on these sites using more sensitive and quantitative proteomics approaches. We were particularly interested to look for H3K56ac because several recent publications have implicated this modification in replication-coupled nucleosome assembly and the DNA damage response [43-48]. In addition, based on immunohistochemistry, it has been argued that H3K56ac is a diagnostic and prognostic marker of many different types of tumours [43]. To detect H3K56ac and determine its stoichiometry, we used two approaches: MRM on an AB Sciex 4000 Q-Trap hybrid triple quadrupole-linear ion trap mass spectrometer, and a high-resolution

quantitative approach on a Thermo Orbitrap-LTQ XL instrument. An absolute quantitation method was developed using both an H3K56ac synthetic peptide and an isotopically labelled standard of the target tryptic peptide (YQK_{56}(ac)STELLI*R), where I* is a [$^{13}C_6{}^{15}N_1$]-labelled isoleucine residue. The unmodified (K56 propionylated) ion was detected in high abundance in the non-targeted LC-MS/MS analyses described earlier. An MRM method was built based on the MS/MS spectrum of both synthetic peptides (figures 5.6a and b). Four precursor/fragment ion pairs were selected to build a specific MRM method for the detection of the corresponding tryptic peptides (see table A.11 in the appendix). We also monitored the unmodified form (YQK_{56}(pr)STELLIR), together with another H3 peptide (YRPGTVALR) to normalize for the amount of H3 present in different samples and for changes in the response of the electrospray signal. Following HPLC fractionation of histone extracts (50 μg of starting material), our MRM assay was used to screen for the presence of H3K56ac in K562 and HeLa cells treated with various HDACi, including sodium butyrate, SAHA, MS-275, and nicotinamide. The modified K56ac residue was detected at low abundance in all cell types and conditions tested consistent with the LC-MS/MS analyses described in the previous section. Figure 5.6c shows LC-MS/MS results of H3K56ac in K562 cells treated with SAHA for 6h. Four MRM transition channels corresponding to the formation of diagnostic *y*- and *b*-type fragment ions were monitored. This strategy unambiguously identified a K562 cell-derived peptide that contained H3K56ac and eluted at 26 min (figure 5.6c). First, the K56-acetylated peptide co-eluted with the [$^{13}C_6{}^{15}N_1$-isoleucine]-labelled internal standard (figures 5.6c and d). Second, as expected, fragment ions from the K562 cell-derived H3K56ac peptide were 7 Th lighter than the corresponding fragment ions from the internal standard containing [$^{13}C_6{}^{15}N_1$]-isoleucine (figures 5.6a and b). The stoichiometry of acetylation was calculated using the peak areas of the ion signal corresponding to the K56ac peptide divided by the sum of the areas of K56ac and K56pr peptides. The stoichiometry of H3K56ac in transformed human cells (K562 or HeLa) is shown in table V-1. The stoichiometry of K56ac determined by MRM was 0.03% in all conditions tested, and did not change upon prolonged incubation with any of the HDACi that we tested.

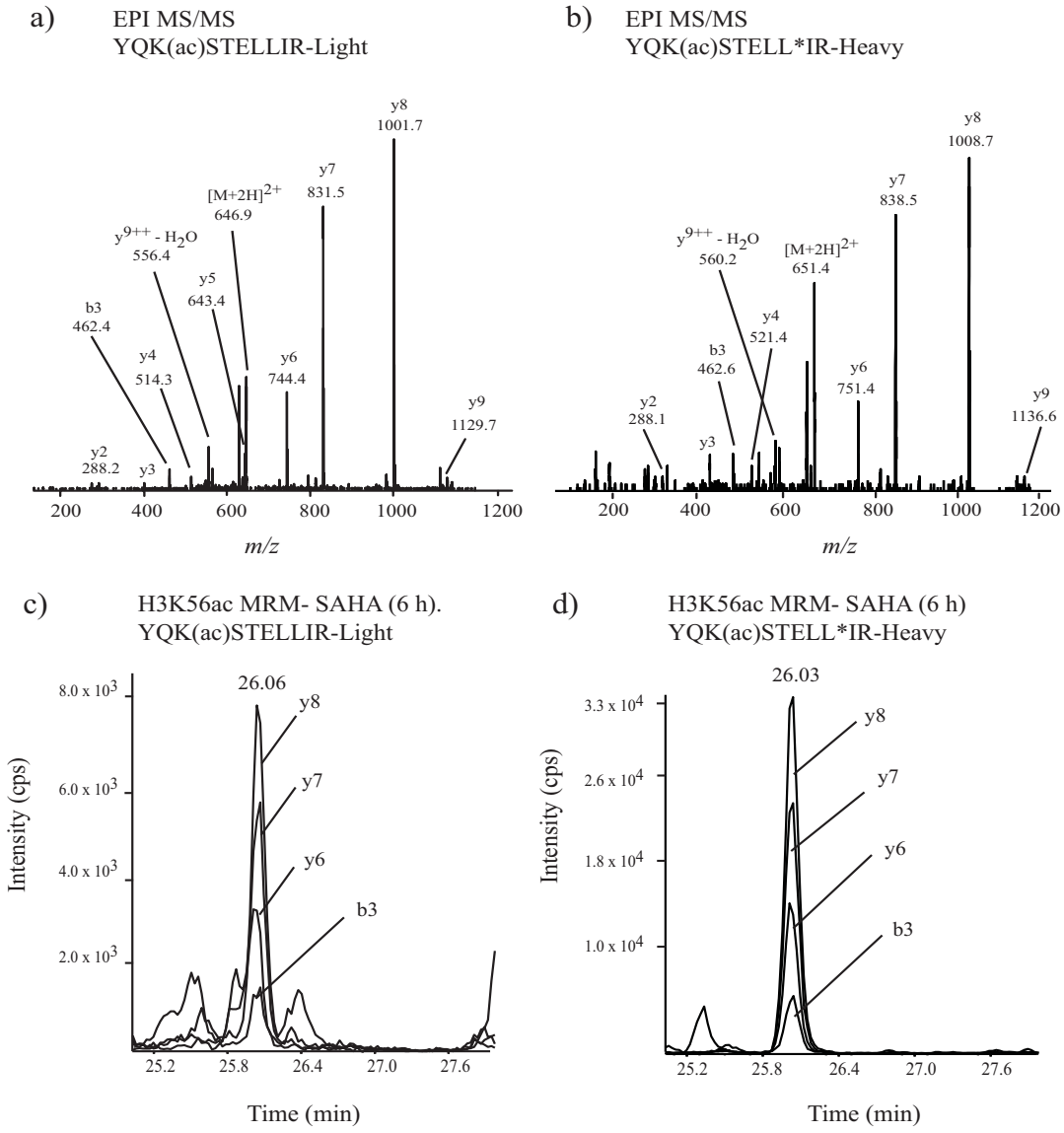


Figure 5.6: Identification and quantitation of H3K56ac from K562 cells using targeted MRM on an AB Sciex 4000 Q-Trap instrument. Tandem MS spectra of the light (a) and heavy (b) YQK_{56}(ac)STELLIR peptide standards at 10 µg/mL using enhanced product ion (EPI) scan. LC-MS/MS analysis (MRM) of the native H3K56ac peptide (c) and the [$^{13}C_6{}^{15}N_1$-isoleucine]-labelled YQK_{56}(ac)STELLI*R internal standard peptide (d) from a digest of histones isolated from K562 cells treated with SAHA for 6h.

We also performed high-resolution LC-MS analyses on the LTQ-Orbitrap XL mass spectrometer to determine the stoichiometry of H3K56ac in the same samples. The

absolute amount of H3K56ac in transformed cells was determined using a calibration curve spiked at levels ranging from 5 to 1000 pg/mL, containing a constant amount of the [$^{13}C_6{}^{15}N_1$]-labelled internal standard. In order to mimic, as closely as possible, the reaction conditions used to process human histones prior to MS, recombinant yeast histone H3 was propionylated and digested first, then spiked with increasing amounts of the synthetic peptide standard. The rationale was to generate reaction mixtures containing similar sets of peptides and chemical species as those produced during analysis of human cell-derived histones. Although recombinant yeast histone H3 is less complex and devoid of PTM [49], the conditions of propionylation and trypsin digestion are similar, which helps tackle challenges to quantitation, such as chemicals that co-elute with the H3K56ac peptide of interest and might suppress its ionization. This is particularly important in this case because of the very low abundance of the H3K56ac peptide (see below). The first residue of the yeast H3 tryptic peptide containing K56 is a Phe, instead of a Tyr residue in human H3. Hence, H3K56 tryptic peptides derived from yeast histones do not interfere with our quantitative analysis. Our calibration curve showed excellent linearity (r^2 = 0.9993) over the range of concentrations examined (figure 5.7a). We detected the authentic H3K56ac peptide at a level approaching the instrumental detection limit in all the human histone samples. An example of the precursor ion mass spectrum for the H3K56ac peptide derived from K562 cells treated with SAHA for 24h is shown in figure 5.7b. The mass measurement obtained for the light and heavy labelled doubly protonated peptides were within 5 ppm of those calculated from their corresponding theoretical masses. An MS/MS spectrum that confirms the identity of the precursor ion corresponding to the human K56ac peptide is shown in figure 5.7c. Similar to what was described above for the MRM analyses, the H3K56ac stoichiometry was determined from the ratio of the peak height of the doubly-protonated K56ac tryptic peptide divided by the sum of the peak heights for the K56ac and K56pr peptides. The results from high-resolution LC-MS analyses are shown in table V-1, together with the absolute amount of tryptic peptide detected. The amount of H3K56ac present in all conditions ranged from 150 to 300 attomols, while the acetylation stoichiometry varied from 0.05-0.08%. No significant change in the stoichiometry of H3K56ac was observed under any of the conditions examined (table V-1). Thus, based on two different quantitative MS

experiments, we find that the stoichiometry of H3K56ac (approximately 0.05%) is considerably lower than previously reported (1%) [43, 47]. In addition, unlike other sites of acetylation, H3K56ac did not increase in response to several HDACi (table V-1). This was very intriguing because two of the HDACi that we used in our MS experiments, namely sodium butyrate and nicotinamide, were previously shown to increase H3K56ac based on immunoblotting [43, 46].

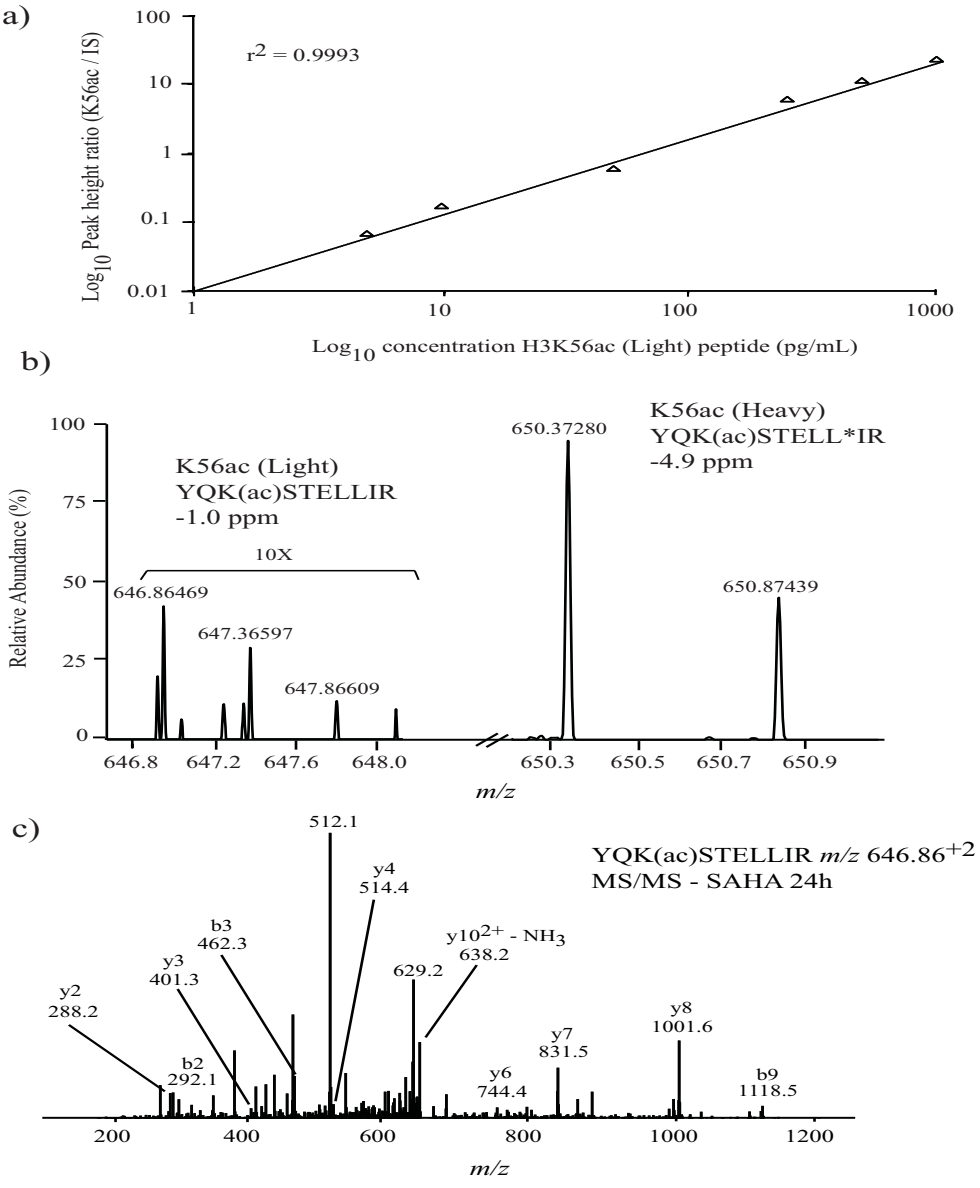


Figure 5.7: LC-MS/MS analysis of H3K56ac from K562 cells using an LTQ-Orbitrap XL mass spectrometer. (a) Standards (n=1, see experimental section) were prepared by mixing solutions of increasing concentrations (from 5 to 1000 pg/mL) of [^{12}C-

isoleucine]-containing K56-acetylated peptide with a fixed amount (100 pg/mL) of internal standard (IS) peptide (YQK_{56}(ac)STELLI*R), where I* is the [$^{13}C_6{}^{15}N_1$]-labelled isoleucine residue). The *y*-axis is the peak height abundance ratio (as determined by LC-MS) of the YQK_{56}(ac)STELLIR peptide (containing only the [^{12}C] isotope) to that of the [$^{13}C_6{}^{15}N_1$]-labelled internal standard (IS) peptide. (b) Precursor ion mass spectrum (from LC-MS/MS analysis of histones derived from K562 cells treated with 1μM SAHA for 24h) showing the detection of the native YQK_{56}(ac)STELLIR peptide along with the [$^{13}C_6{}^{15}N_1$-isoleucine]-labelled internal standard. (c) MS/MS spectrum of precursor *m/z* 646.86^{2+} from the analysis shown in (b). This spectrum confirms the identity of the natural peptide containing H3K56ac.

This prompted us to investigate the specificity of two commercially available antibodies that were previously used to detect H3K56ac in immunoblots [43, 45]. For these experiments, the same histone samples were analyzed either by MS or immunoblotting. The amounts of acid-extracted histones typically used in immunoblots are sufficiently large to be readily detected by Ponceau S staining after histone transfer [43, 45]. Under these conditions, we found that, as previously reported [43, 45, 46], histones derived from HeLa cells treated with sodium butyrate showed a strong increase in signal with antibodies supposedly specific for H3K56ac (figure 5.8a). However, with equal amounts of histones loaded, the same antibodies detected a considerable signal in histones derived from yeast *rtt109*Δ mutant cells (figure 5.8b). In yeast, Rtt109 is the main enzyme that acetylates H3K56 and, consistent with this, H3K56ac is severely compromised, as judged by MS analysis of histones derived from *rtt109*Δ mutant cells [31]. These observations, combined with the fact that MS did not reveal an increase in H3K56ac after extended periods of incubation with several HDACi (table V-1), suggest that the two commercial antibodies raised against H3K56 acetylated peptides might react against acetylated lysine residues other than H3K56 in human cells. However, further experiments would need to be performed to reach a definitive conclusion. For example, a peptide competition assay utilizing H3K9ac and H3K56ac peptides present in different stoichiometries blotted against the H3K56ac antibody would an important experiment to conduct. Cross reactivity with another site with similar sequence homology could be one such explanation; one possible candidate would be H3K9ac, given that the stoichiometry of H3K9ac increases from 3% in control to 33% in SAHA-treated K562 cells. These results may help resolve some of the conflicting data concerning the physiological roles

of H3K56ac in human cells [43-48]. In addition, our results demonstrate the importance of combining MS and immunoblotting data to address whether antibodies raised against certain histone modifications are likely to be modification-specific or not. This is particularly important for studies of vertebrate cells where it is not currently possible to mutate specific histone residues in order to rigorously prove antibody specificity.

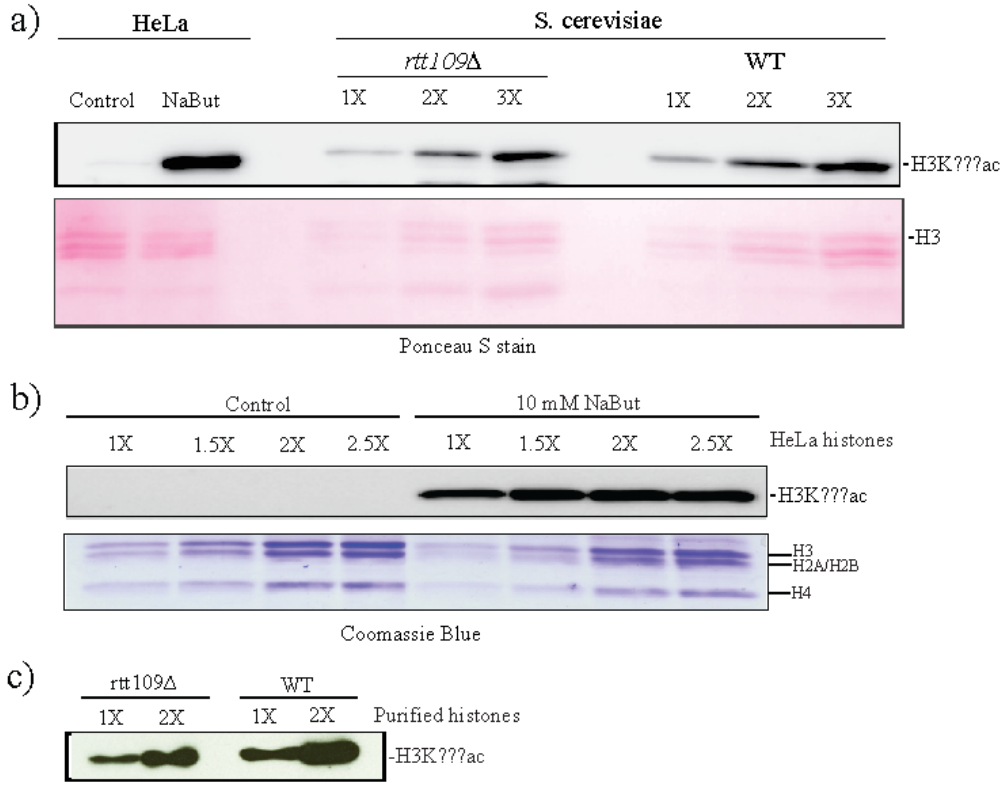


Figure 5.8: Antibodies supposedly specific for H3K56ac can cross-react against other acetylated lysine residues in histone H3. (a) Immunoblot of core histones purified from control and butyrate-treated HeLa cells (NaBut), as well as WT and *rtt109*Δ yeast cells probed with an antibody from Upstate/Millipore (07-677). As judged by Ponceau S staining, similar amounts of core histones were present on the blot prior to probing. (b) Immunoblot of core histones purified from control and butyrate-treated HeLa cells probed with the Epitomics antibody 2134-1. As judged from a Coomassie Blue-stained gel loaded with the same amounts of histones as the immunoblot, similar amounts of histones from control and butyrate-treated cells were immunoblotted cells. (c) Immunoblot of core histones purified from yeast cells lacking the H3K56 acetyltransferase Rtt109 (*rtt109*Δ) or WT cells. In panels (a) to (c), the "H3K???ac" label indicates that the immunoblotting signal generated by these antibodies cannot be ascribed to a specific acetylated lysine(s).

Table V-1: Stoichiometry and amounts of H3K56ac detected in transformed cells exposed to different HDACi (n = 1).

Condition	cal. conc. (pg/mL)	amount inj. (fg)	attomols detected	K56ac stoichio. (%)	K56ac stoichio. (%)
				MRM	Orbitrap
K562 CTL 0h	10.3	207	160	0.03	0.05
K562 CTL 12h	12.8	255	198	0.03	0.05
K562 CTL 24h	14.4	289	224	0.03	0.05
K562 SAHA 6h Rx.	13.3	266	206	0.03	0.05
K562 SAHA 24h Rx.	10.6	211	164	0.03	0.05
K562 MS-275 6h Rx.	10.9	218	169	0.03	0.05
K562 MS-275 24h Rx.	19.2	384	297	0.03	0.05
K562 Nicotinic acid 12h Rx	13.0	260	201	0.03	0.05
K562 Nicotinamide 12h Rx	10.2	205	158	0.03	0.05
K562 Nicotinamide 24h Rx	14.5	291	225	0.03	0.07
HeLa Butyrate 24h Rx	18.4	367	284	0.03	0.08

In separate experiments, we also monitored histone H4K91 acetylation, a modification previously reported to be involved in chromatin assembly [6]. We searched for the peptide containing this lysine residue, K_{79}TVTAMDVVYALK$_{91}$R in histones from both control and HDACi treated K562 cells. Precursor and fragment ion pairs were selected from the five most abundant transitions observed for both K91ac and its propionylated counterpart K91pr to determine the corresponding acetylation stoichiometry (see table A.12 in appendix 1). The extracted ion chromatogram of the unmodified H4K91 peptide (both doubly and triply charged states) yielded an intense peak for the K91pr tryptic peptide in all MRM transitions monitored (see figure A.3 in the appendix). In contrast, the extracted ion chromatogram of the K91ac tryptic peptide showed no detectable signal for any of the MRM transitions monitored under any of the conditions tested (see figure A.4 in appendix 1). We estimate that the acetylation stoichiometry of H4K91 is below 0.05%, which corresponds to the detection limit in this LC-MS/MS experiment.

5.5 Conclusion

Our two-pronged approach to identify global changes in histone acetylation that arise from HDAC inhibition has enabled us to uncover two unexpected findings about the dynamics of histone acetylation in human cells. First, HDACi with either hydroxamic acid (SAHA) or benzamide (MS-275) as inhibitory moieties trigger hyperacetylation of H3 and H4 preferentially in transformed cell lines compared with normal diploid fibroblasts. This observation may well explain why SAHA and other HDACi exert more pronounced effects (such as cell cycle arrest, differentiation or apoptosis) on cancer cells. This finding raises fascinating questions for future investigation, such as the underlying molecular basis of the differential susceptibility of normal and cancer cells to HDACi. Second, the development of a robust MS-based quantitative assay to determine the stoichiometry of histone H3K56ac in control and HDACi-treated cells has revealed some highly unexpected properties of this modification in human cells that are distinct from H3K56ac in yeast. In *S. cerevisiae* and other yeasts, H3K56ac is very abundant during S-phase and G2 (essentially all new H3 molecules are K56-acetylated) and deacetylated genome-wide prior to the subsequent round of S-phase [4, 50]. This is clearly not the case for H3K56ac in human cells. In contrast to the acetylation of many other residues in H3/H4 N-terminal tails [51], it is puzzling that there is no turnover of H3K56ac that can be blocked by either class I/II or class III HDACi in human cells. The very low stoichiometry of H3K56ac in human cells and the lack of turnover make it unlikely that this modification can enhance genome-wide deposition of newly synthesized histone molecules. Our findings certainly do not rule out the possibility that H3K56ac may have an important function in human cells. A stoichiometry of 0.03% corresponds to roughly one K56-acetylated H3 molecule every 150 nucleosomes (or roughly every 30Kb, assuming an average nucleosome repeat length of 200bp). Given the size of the genome of a diploid human cell (roughly $6x10^9$ bp), there are approximately $2x10^5$ K56-acetylated H3 molecules per cell. Despite the absence of turnover, some of these molecules may occur at specific chromosomal loci where H3K56ac could play an important function, such as the regulation of gene expression proposed from studies of H3K56ac in embryonic stem cells [47]. Overall, our findings demonstrate the importance of

quantitative MS to determine the stoichiometry of specific protein modifications under a variety of physiological conditions, and subsequently use this information to formulate hypotheses regarding the plausible function(s) of these modifications.

5.6 Acknowledgment

This work was supported by funding from the Collaborative Health Research Program of the National Science and Engineering Research Council (NSERC) and the Canadian Institutes for Health Research (CIHR) to A.V. and P.T and the Canadian Cancer Research Society to G.F. The Institute for Research in Immunology and Cancer receives infrastructure support from the Canadian Center of Excellence in Commercialization and Research, the Canadian Foundation for Innovation, and the Fonds de la Recherche en Santé du Québec.

5.7 References

1. Davey, C.A., et al., *Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution.* J Mol Biol, 2002. **319**(5): p. 1097-113.

2. Luger, K., *Dynamic nucleosomes.* Chromosome Res, 2006. **14**(1): p. 5-16.

3. Kouzarides, T., *Chromatin modifications and their function.* Cell, 2007. **128**(4): p. 693-705.

4. Masumoto, H., et al., *A role for cell-cycle-regulated histone H3 lysine 56 acetylation in the DNA damage response.* Nature, 2005. **436**(7048): p. 294-8.

5. van Leeuwen, F., P.R. Gafken, and D.E. Gottschling, *Dot1p modulates silencing in yeast by methylation of the nucleosome core.* Cell, 2002. **109**(6): p. 745-56.

6. Ye, J., et al., *Histone H4 lysine 91 acetylation a core domain modification associated with chromatin assembly.* Mol Cell, 2005. **18**(1): p. 123-30.

7. Strahl, B.D. and C.D. Allis, *The language of covalent histone modifications.* Nature, 2000. **403**(6765): p. 41-5.

8. Morales, V., et al., *Chromatin structure and dynamics: functional implications.* Biochimie, 2001. **83**(11-12): p. 1029-39.

9. Bolden, J.E., M.J. Peart, and R.W. Johnstone, *Anticancer activities of histone deacetylase inhibitors.* Nat Rev Drug Discov, 2006. **5**(9): p. 769-84.

10. Yoshida, M., et al., *Potent and specific inhibition of mammalian histone deacetylase both in vivo and in vitro by trichostatin A.* J Biol Chem, 1990. **265**(28): p. 17174-9.

11. Halkidou, K., et al., *Upregulation and nuclear recruitment of HDAC1 in hormone refractory prostate cancer.* Prostate, 2004. **59**(2): p. 177-89.

12. Minucci, S., et al., *Histone deacetylases: a common molecular target for differentiation treatment of acute myeloid leukemias?* Oncogene, 2001. **20**(24): p. 3110-5.

13. 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.

14. Lee, M.J., et al., *Histone deacetylase inhibitors in cancer therapy.* Curr Opin Oncol, 2008. **20**(6): p. 639-49.

15. Prince, H.M., M.J. Bishton, and S.J. Harrison, *Clinical studies of histone deacetylase inhibitors.* Clin Cancer Res, 2009. **15**(12): p. 3958-69.

16. Stimson, L., et al., *HDAC inhibitor-based therapies and haematological malignancy.* Ann Oncol, 2009. **20**(8): p. 1293-302.

17. Bots, M. and R.W. Johnstone, *Rational combinations using HDAC inhibitors.* Clin Cancer Res, 2009. **15**(12): p. 3970-7.

18. Dokmanovic, M., C. Clarke, and P.A. Marks, *Histone deacetylase inhibitors: overview and perspectives.* Mol Cancer Res, 2007. **5**(10): p. 981-9.

19. Altucci, L., et al., *RAR and RXR modulation in cancer and metabolic disease.* Nat Rev Drug Discov, 2007. **6**(10): p. 793-810.

20. Mann, B.S., et al., *FDA approval summary: vorinostat for treatment of advanced primary cutaneous T-cell lymphoma.* Oncologist, 2007. **12**(10): p. 1247-52.

21. Campas-Moya, C., *Romidepsin for the treatment of cutaneous T-cell lymphoma.* Drugs Today (Barc), 2009. **45**(11): p. 787-95.

22. Beckers, T., et al., *Distinct pharmacological properties of second generation HDAC inhibitors with the benzamide or hydroxamate head group.* Int J Cancer, 2007. **121**(5): p. 1138-48.

23. Stimson, L. and N.B. La Thangue, *Biomarkers for predicting clinical responses to HDAC inhibitors.* Cancer Lett, 2009. **280**(2): p. 177-83.

24. Smith, C.M., et al., *Mass spectrometric quantification of acetylation at specific lysines within the amino-terminal tail of histone H4.* Anal Biochem, 2003. **316**(1): p. 23-33.

25. Garcia, B.A., et al., *Chemical derivatization of histones for facilitated analysis by mass spectrometry.* Nat Protoc, 2007. **2**(4): p. 933-8.

26. Johnson, L., et al., *Mass spectrometry analysis of Arabidopsis histone H3 reveals distinct combinations of post-translational modifications.* Nucleic Acids Res, 2004. **32**(22): p. 6511-8.

27. Syka, J.E., et al., *Novel linear quadrupole ion trap/FT mass spectrometer: performance characterization and use in the comparative analysis of histone H3 post-translational modifications.* J Proteome Res, 2004. **3**(3): p. 621-6.

28. Beck, H.C., et al., *Quantitative proteomic analysis of post-translational modifications of human histones.* Mol Cell Proteomics, 2006. **5**(7): p. 1314-25.

29. Lee, A.Y., et al., *Quantitative Analysis of Histone Deacetylase-1 Selective Histone Modifications by Differential Mass Spectrometry.* J Proteome Res, 2008. **7**: p. 5177-86.

30. Matthiesen, R., et al., *VEMS 3.0: algorithms and computational tools for tandem mass spectrometry based identification of post-translational modifications in proteins.* J Proteome Res, 2005. **4**(6): p. 2338-47.

31. Drogaris, P., et al., *Comprehensive profiling of histone modifications using a label-free approach and its applications in determining structure-function relationships.* Anal Chem, 2008. **80**(17): p. 6698-707.

32. Shechter, D., et al., *Extraction, purification and analysis of histones.* Nat Protoc, 2007. **2**(6): p. 1445-57.

33. Larson, E.M., et al., *A new, simple, nonradioactive, nontoxic in vitro assay to monitor corneal endothelial cell viability.* Invest Ophthalmol Vis Sci, 1997. **38**(10): p. 1929-33.

34. Dyer, P.N., et al., *Reconstitution of nucleosome core particles from recombinant histones and DNA.* Methods Enzymol, 2004. **375**: p. 23-44.

35. Benson, L.J., et al., *Modifications of H3 and H4 during chromatin replication, nucleosome assembly, and histone exchange.* J Biol Chem, 2006. **281**(14): p. 9287-96.

36. Loyola, A., et al., *PTMs on H3 variants before chromatin assembly potentiate their final epigenetic state.* Mol Cell, 2006. **24**(2): p. 309-16.

37. Sobel, R.E., et al., *Conservation of deposition-related acetylation sites in newly synthesized histones H3 and H4.* Proc Natl Acad Sci U S A, 1995. **92**(4): p. 1237-41.

38. Verreault, A., et al., *Nucleosome assembly by a complex of CAF-1 and acetylated histones H3/H4.* Cell, 1996. **87**(1): p. 95-104.

39. Anoopkumar-Dukie, S., et al., *Resazurin assay of radiation response in cultured cells.* Br J Radiol, 2005. **78**(934): p. 945-7.

40. Song, O.K., et al., *An Nalpha-acetyltransferase responsible for acetylation of the N-terminal residues of histones H4 and H2A.* J Biol Chem, 2003. **278**(40): p. 38109-12.

41. Marzluff, W.F., et al., *The human and mouse replication-dependent histone genes.* Genomics, 2002. **80**(5): p. 487-98.

42. Orsi, G.A., P. Couble, and B. Loppin, *Epigenetic and replacement roles of histone variant H3.3 in reproduction and development.* Int J Dev Biol, 2009. **53**(2-3): p. 231-43.

43. Das, C., et al., *CBP/p300-mediated acetylation of histone H3 on lysine 56.* Nature, 2009. **459**(7243): p. 113-7.

44. Miller, K.M., et al., *Human HDAC1 and HDAC2 function in the DNA-damage response to promote DNA nonhomologous end-joining.* Nat Struct Mol Biol, 2010. **17**(9): p. 1144-51.

45. Tjeertes, J.V., K.M. Miller, and S.P. Jackson, *Screen for DNA-damage-responsive histone modifications identifies H3K9Ac and H3K56Ac in human cells.* EMBO J, 2009. **28**(13): p. 1878-89.

46. Vempati, R.K., et al., *p300-mediated acetylation of histone H3 lysine 56 functions in DNA damage response in mammals.* J Biol Chem, 2010. **285**(37): p. 28553-64.

47. Xie, W., et al., *Histone h3 lysine 56 acetylation is linked to the core transcriptional network in human embryonic stem cells.* Mol Cell, 2009. **33**(4): p. 417-27.

48. Yuan, J., et al., *Histone H3-K56 acetylation is important for genomic stability in mammals.* Cell Cycle, 2009. **8**(11): p. 1747-53.

49. Luger, K., et al., *Characterization of nucleosome core particles containing histone proteins made in bacteria.* J Mol Biol, 1997. **272**(3): p. 301-11.

50. Celic, I., et al., *The sirtuins hst3 and Hst4p preserve genome integrity by controlling histone h3 lysine 56 deacetylation.* Curr Biol, 2006. **16**(13): p. 1280-9.

51. Waterborg, J.H., *Dynamics of histone acetylation in vivo. A function for acetylation turnover?* Biochem Cell Biol, 2002. **80**(3): p. 363-78.

6. Conclusion

6.1 Conclusion and Final Thesis Overview

Histones play a vital role in the biology of the cell. Their role includes that of DNA packaging "agents", and their PTMs help regulate important cellular processes (i.e., transcription, damage repair). The deregulation of these PTMs can directly affect and sustain diseases such as cancer. Hence, characterizing, analyzing, and quantifying the extent of these PTMs is important in understanding how they are implicated in disease, as well as to how to exploit them therapeutically. The main objective of this Ph.D. project was to develop analytical MS-based strategies to analyze histones to further understand their involvement in cell biology. This objective was successfully completed and documented in the preceding chapters of this thesis.

The first step was to develop a rugged and reproducible workflow that could be applied to any sample set. An in-depth literature search revealed that most, if not all, research groups used a generic proteomics approach to analyze histones; the few that developed more specialized workflows provided no statistical analysis to validate their results and findings. This lack of analytically tested and verified results in many published papers prompted us to develop an analytically sound strategy for histone analysis, and became the basis for the first publication presented in this thesis (chapter three).

Our approach was a dual, two-pronged strategy. Histones extracted from cells were profiled by high resolution mass spectrometry to generate a molecular weight profile. Using histones from wild type and mutant yeasts strains (*asf1*Δ and *rtt*109Δ) as a biological model, we profiled and examined closely each core intact histone from the three different sources. These two mutants were chosen because the deletion of the two genes encoding the ASF1 histone chaperone and RTT109 acetyltransferase is known to catalyze the acetylation of H3K56 [1]. The results obtained were consistent with our expectations. Core histones H2A, H2B, and H4 did not demonstrate any significant changes in their intact mass profiles, while histone H3 had obvious changes in the abundance and mass shifting between wild type and mutants. This rapid MS profiling

demonstrated the utility of performing such an experiment; instead of spending time on analyzing histones that showed no significant changes, we channeled our research efforts to analyze and determine the changes occurring in histone H3 between wild-type and mutants. Histone H3 became the primary candidate for the next step of our dual strategy, i.e., bottom-up sequencing. Since this was our first report and major study involving histones, we decided to analyze all core histones to demonstrate proof of concept.

Sample preparation for histones was the next aspect to be investigated and developed. Since a cellular extract contains a great number of proteins besides histones, some type of fractionation step was necessary to remove unwanted proteins, and isolate individual histones. Two methods were available for this purpose: SDS-PAGE and HPLC. Separation of histones by gels and subsequent staining was not chosen because of limited loading and resolution, and the possibility of creating chemical artifacts [2, 3]. Using both narrow and standard bore HPLC with conventional reversed-phase C18 media, we were able to separate core histones from co-extracted proteins, as well as from each other.

The protein digestion method posed a serious challenge. It was evident that a generic tryptic digestion procedure would not suffice for histone proteins. Endoproteinase Arg-C was initially used to perform the digestion to limit proteolysis only to arginine residues. However, the use of this enzyme also resulted in non-specific cleavage at lysine residues, and was not very reproducible between each digestion. Our preference was to use trypsin based protocols due to its efficiency and reproducibility. After an extensive literature search, the only technique found that was customized for preparing histones for bottom-up sequencing was Hunt's method using chemical derivatization with propionic anhydride [4]. One initial concern about using this method was the possibility that the addition of a propionyl group could also occur *in-vivo*, as shown recently by Liu in histone H3 at lysine 23 in mammalian cells [5], albeit at low abundance. We decided to search for this PTM in our preliminary experiments. In-vivo propionylation was never detected at any significant level. Our initial experiments using this method resulted in a high extent of side reactions, i.e., amidation and esterification. Modifications to the

original method resulted in the elimination of carboxylic acid esters being formed, and minimizing amidation of the same functional group. Our optimized method resulted in almost complete chemically derivatization of all lysine residues. Validation experiments using HPLC fractionation, chemical labeling, tryptic digestion and MS analysis resulted in only a 15% CV for the intensity measurements of all ions detected from histone H3 extracts. Spiking experiments using different ratios of H3K56ac demonstrated excellent correlation between the observed and calculated fold-changes ($r^2 = 0.995$). Using the H3K56ac peptide, we were able to detect changes over a two orders of magnitude with a high degree of confidence (< 10% CV for all spiking levels). These figures of merit provided us the analytical confidence in our workflow to detect and quantify changes in histone PTM from complex matrices.

LC-MS/MS analysis was performed on extracts derived from all four core histones. The raw data was rapidly processed by the bioinformatics mining procedure outlined in section 2.4. Scatter plots displaying peptide intensities from wild type and mutant yeast strains were constructed from the data exported into Microsoft Excel. These graphical representations allow a researcher to make a rapid assessment of the differential PTM expression. Peptides from histones H2A, H2B, and H4 exhibited little change between wild type and either *asf1*Δ or *rtt*109Δ. Only one data point stood out, and tandem MS revealed this to be H3K56ac. The results reinforced our expectations that the observations detected were real, and we could accurately attribute a quantifiable value to these changes.

The application of MRM to quantify histone PTM in this study is one of the very first reported in the literature. MRM was used to determine the stoichiometry of H3K56ac in wild type vs. *rtt*109Δ mutant yeast strain with higher precision. When comparing histones derived from wild type and mutant strains, H3K56ac was not present genome-wide, but in a sub-population (17%) of all nucleosomes in wild type cells; this PTM was drastically reduced to 0.4% in the mutant strain upon deletion of the gene encoding the RTT109 acetyltransferase. *S. cerevisiae* was the first species in which we were able to detect this modification. This PTM is present in significant abundance in

yeast histones, whose appearance is cell-cycle dependant, and plays a crucial role in the DNA damage response [6]. The same sample preparation protocol and MRM method was used in a collaborative project to assess the stoichiometry of H3K56ac in a similar but pathogenic organism, i.e., *Candida albicans* [7]. This was the second organism in which we were able to detect H3K56ac with a relatively high stoichiometry (20 to 27%). We were able to assess the stoichiometry of H3K56ac in different experimental conditions with a high degree of precision (CV < 5% over triplicate injections). A more profound understanding of this PTM led to its exploitation by our collaborators as a therapeutic and anti-infective target. Inhibition of the Sirtuin HDAC Hst3p enzyme by nicotinamide treatment lead to an increase in H3K56ac, and inhibited the growth of *C. albicans*. The virulence of this organism was reduced by selectively modulating H3K56ac expression in a mouse *in vivo* infection model.

The development of MS analytical strategies for histone profiling continued on the AB Sciex 4000 Q-trap, and become the basis for the second publication presented in this thesis (chapter four). The potential of doing charge separation is an attractive and beneficial attribute that can aid in the detection of analytes present in complex biological matrices. Techniques such as high field asymmetric wave form ion mobility spectrometry (FAIMS) [8] and T-wave ion mobility [9] have been incorporated into various instrumental platforms, providing an added dimension of separation. The application of the EMC scan for smaller analytes such as peptides had already been reported by Hopfgartner [10]. Hence, we decided to apply the same scan mode to enhance the detection of a more challenging class of analytes, i.e., intact proteins. Our initial LC-MS experiments went beyond the limits of the instrument and failed. The successful implementation of EMC required modifications to be made to the method table that controlled various voltages along the ion path and within the LIT, transferring a limited range of ions into the linear trap in a sequential fashion, and applying the charge separation (MCS) voltage to selectively remove singly charged species. This modified or targeted EMC scan not only reduced space charge effects, but also resulted in a 4-fold increase in signal-to-noise, as well as a 5-fold increase in resolution. The targeted EMC (tEMC) mode provided added resolution approaching that of TOF instruments in the

intact histone profiles generated. LC-MS analysis using tEMC of DT40 histones derived from wild type and DOT1L mutant helped in the discrimination of closely spaced peaks and many histone isoforms. DOT1L is a methyltransferase that modifies predominantly lysine 79 in histone H3. Thus, in DOT1L deficient cells, a reduction of lysine 79 methylation was to be expected. Intact histone H3 profiles between wild type and mutant showed subtle changes in the abundance of isoforms peaks, indicating a difference was present between the two strains. A true top down strategy was attempted, whereby whole protein ions were introduced into the LIT and fragmented by either multistage activation, or EPI scan, similar to Meluzzi's work on an LTQ instrument [11]. However, we were not able to induce and detect any fragments from intact histones using the 4000 Q-trap. As an alternative, we applied our bottom-up strategy and protocol with MRM detection on the same platform. Once again, MRM proved to be very versatile and useful; the change in methylation was localized at H3K79, whose mono-methylation was down-regulated by 240-fold in DOT1L deficient cells. This study was a good example of how software manipulation can increase the analytical potential of a mass spectrometer, enabling both intact profiling and peptide detection on the same instrumental platform.

HDAC inhibitors have proven to be a promising epigenetic drug strategy for cancer therapy. Upon exposure to these molecules, histones become hyperacetylated, and silenced tumor suppressor genes reactivated. We applied the methodology developed in chapter three to study the effects of two clinically relevant HDAC inhibitors in both normal and cancer cells. This study became the basis for the third publication presented in this thesis (chapter five). Intact histone profiling proved to be once again very useful. In K562 cancer cells treated with either SAHA or MS-275, significant differences in the mass profiles were shown for both histone H3 and H4 when compared to non-treated K562 cells. The same effects were not observed in WI-38 or IMR90 normal cells treated with SAHA. This result was very surprising, and is one of the few demonstrations of the effects of HDACi in normal cells reported in the literature. These inhibitors can selectively affect cancerous cells and induce hyperacetylated states, without perturbing the functions of normal cells. This is in contrast to older therapies, such as oxazaphosphorine-based drugs and alkylating agents (e.g. cyclosphosphamide) which

have inherent toxicity in cancer chemotherapy [12]. This apparent selective effect of HDACi is very desirable in drug design, and can be exploited by medicinal chemists in the discovery of next generation HDACi with less side effects and toxicity.

Only histones H3 and H4 in K562 cells were selected for the bottom-up strategy, since these histones were the only two to demonstrate significant changes. The instrumental advantages of the LTQ-Orbitrap XL mass spectrometer equipped with both CID and ETD capability were used to probe deeper into the histone proteome. Our analytical protocol enabled the high sequence coverage for both histones H3 and H4 (63% and 75%, respectively). Peptide sequencing for histone H3 revealed nine unique peptides, in multiply modified forms, both on the N-terminus and globular domain. Only a handful of the modified peptides exhibited a significant fold-change. These included known acetylation sites at lysine 9, 14, 18, and 23. Both SAHA and MS-275 inhibited HDAC enzymes that catalyzed these acetylation "hotspots" primarily on the N-terminus. Interestingly, histone H3 methylation was also affected by exposure to both HDACi's, but to a lesser extent. Most notably was the dimethylation present concurrently at both lysine 27 and 36. Again, this is one of the first reports of histone methylation being upregulated in cells exposed to HDACi. The overexpression of this peptide was observed in several instances after exposure to both SAHA and MS-275, with a similar fold-change. It is perplexing why such modifications would be upregulated simultaneously on the same peptide, since H3 lysine 27 methylation is a mark of transcriptionally silent chromatin [13], while H3 lysine 36 methylation is a mark of active chromatin [14]. The function (if any) of this over-expressed dual modification remains to be explored and understood.

Peptide sequencing for histone H4 revealed seven unique peptides in multiply modified forms. Similar effects were observed for histone H4 in response to HDACi treatment. The acetylation of four different sites on the H4 N-terminus was up-regulated at different extents upon exposure to either SAHA or MS-275. The concurrent hyperacetylation of the tetra-acetylated H4 N-terminus at lysines 5, 8, 12, and 16 indicated that HDACi inhibit several enzymes at the same time that catalyze this

modification at these sites. The challenge in HDACi drug design remains in selectively inhibiting specific deacetylases that target removal of acetyl groups, rather than inhibiting several HDACs at the same time. This challenge is further compounded by many of these enzymes being structurally homologous and are inhibited by the same mechanism [15].

A number of publications have reported the presence of H3K56ac in human cells. Up to this point, in our experiments using higher organisms (mouse, fruitfly, chicken, human stem cells), we had no success in detecting this modification. Unlike yeast histones, H3K56ac in mammalian cells is present in very low abundance. Based on extrapolation from our analytical determinations using MRM, the stoichiometry of this PTM was well below 0.1% compared to the unmodified (free K56) peptide. It is difficult to speculate the real function of this PTM at such a low stoichiometry in mammalian cells. Furthermore, some research groups have made claims about this modification based on immunoblotting data, with no MS data to confirm and support their findings. For example, Jackson reported the down-regulation of H3K56ac in response to DNA damage [16], while Yuan argued that the modification plays a crucial role in genomic stability and the DNA damage response in mammals [17]. Our experiments using similar antibodies that were supposedly specific against H3K56ac demonstrated cross reactivity with other up-regulated acetylations sites; hence, this showed the inherent risk of drawing erroneous conclusions using only data generated from immunoblots. Using our sample preparation and MRM techniques, we screened for this PTM in human cells from a variety of sources including K562 and HeLa. Our analytically rugged and proven MS methodology was able to detect and quantitate this low abundant PTM in human cells at stoichiometric levels below 0.1%. Using high resolution MS combined with absolute quantitation with an isotopically labeled internal standard, we determined the amounts of this PTM to be at the attomole level in mammalian cells, and did not change significantly after HDACi treatment. With this information, it is unlikely that the same PTM in both yeast and mammalian cells has the same function. It is difficult to conclude what is the real purpose of this PTM in mammalian cells; this question remains to be answered.

6.2 Future Perspectives

Research into histone biology is a dynamic and perpetually changing field. New methods, sample preparation, and instrumental techniques are in constant development to learn more insight about how these proteins and their modifications affect cell biology and disease. One such trend is the use of an alternative enzyme (e.g. Asp-N) to generate larger peptides, and the use of ETD to sequence and localize PTM, as reported by Coon in the large scale analysis of histone H4 from human stem cells [18]. This "middle down" strategy could be used in parallel with our approach to not only detect more PTM, but also gain insight in the connectivity and cross-talk of the modifications.

Gaseous phase sample introduction devices such as FAIMS have been shown to possess great analytical benefits. This technique can be an orthogonal alternative to detecting hydrophilic peptides that would not retain on conventional C_{18} phases using conventional nano-flow HPLC. Instead of developing new separation conditions on different stationary phases (e.g. HILIC and ion exchange sorbents) which take a considerable amount of time, ion mobility separations can be developed and performed quite rapidly to characterize poorly retained peptides. In a collaborative study, we were able to use FAIMS to detect, quantify, and sequence a small hydrophilic histone H4 tetrapeptide (K_{20}VLR) that failed to be retained on our C_{18} trapping and analytical columns [19]. Hence, we believe that the use of FAIMS and other ion mobility techniques will also become another important tool in the MS detection of histone peptides.

Targeted proteomic strategies using MRM pioneered by Aebersold's group will be adopted by more groups, as the benefits of MRM technologies become more apparent and recognized. The ability to monitor hundreds of MRM transitions in a single HPLC run will provide histone researchers a powerful means to detect and quantity many differently modified peptides accurately, rapidly, with high sensitivity. Peptides demonstrating a high up-regulation between a control and test condition (i.e. HDACi treatment) can serve as biomarkers and monitored routinely in a rapid MRM assay. This

MS-driven analysis has several advantages over classical antibody-based assays. Some of these advantages include a non-biased approach for profiling modified peptides, precise site localization and quantitative capability, and highly accurate stoichiometric assessments. With emerging tools and new ones to be discovered, a researcher is only limited by his dedication and imagination.

6.3 References

1. Han, J., et al., *Acetylation of lysine 56 of histone H3 catalyzed by RTT109 and regulated by ASF1 is required for replisome integrity.* J Biol Chem, 2007. **282**(39): p. 28587-96.

2. Gharib, M., et al., *Artifactual sulfation of silver-stained proteins: implications for the assignment of phosphorylation and sulfation sites.* Mol Cell Proteomics, 2009. **8**(3): p. 506-18.

3. Perdivara, I., et al., *Mass spectrometric identification of oxidative modifications of tryptophan residues in proteins: chemical artifact or post-translational modification?* J Am Soc Mass Spectrom, 2010. **21**(7): p. 1114-7.

4. Garcia, B.A., et al., *Chemical derivatization of histones for facilitated analysis by mass spectrometry.* Nat Protoc, 2007. **2**(4): p. 933-8.

5. Liu, B., et al., *Identification and characterization of propionylation at histone H3 lysine 23 in mammalian cells.* J Biol Chem, 2009. **284**(47): p. 32288-95.

6. Ozdemir, A., et al., *Histone H3 lysine 56 acetylation: a new twist in the chromosome cycle.* Cell Cycle, 2006. **5**(22): p. 2602-8.

7. Wurtele, H., et al., *Modulation of histone H3 lysine 56 acetylation as an antifungal therapeutic strategy.* Nat Med, 2010. **16**(7): p. 774-80.

8. Saba, J., et al., *Enhanced sensitivity in proteomics experiments using FAIMS coupled with a hybrid linear ion trap/Orbitrap mass spectrometer.* J Proteome Res, 2009. **8**(7): p. 3355-66.

9. Giles, K., et al., *Applications of a travelling wave-based radio-frequency-only stacked ring ion guide.* Rapid Commun Mass Spectrom, 2004. **18**(20): p. 2401-14.

10. Tschappat, V., et al., *The application of 2-D dual nanoscale liquid chromatography and triple quadrupole-linear ion trap system for the identification of proteins.* J Sep Sci, 2005. **28**(14): p. 1704-11.

11. Meluzzi, D., et al., *Top-down mass spectrometry on low-resolution instruments: characterization of phosphopantetheinylated carrier domains in polyketide and non-ribosomal biosynthetic pathways.* Bioorg Med Chem Lett, 2008. **18**(10): p. 3107-11.

12. Giraud, B., et al., *Oxazaphosphorines: new therapeutic strategies for an old class of drugs.* Expert Opin Drug Metab Toxicol, 2010. **6**(8): p. 919-38.

13. Pasini, D., et al., *Characterization of an antagonistic switch between histone H3 lysine 27 methylation and acetylation in the transcriptional regulation of Polycomb group target genes.* Nucleic Acids Res, 2010. **38**(15): p. 4958-69.

14. Bannister, A.J., et al., *Spatial distribution of di- and tri-methyl lysine 36 of histone H3 at active genes.* J Biol Chem, 2005. **280**(18): p. 17732-6.

15. Zheng, Y.G., et al., *Chemical regulation of epigenetic modifications: opportunities for new cancer therapy.* Med Res Rev, 2008. **28**(5): p. 645-87.

16. Tjeertes, J.V., K.M. Miller, and S.P. Jackson, *Screen for DNA-damage-responsive histone modifications identifies H3K9Ac and H3K56Ac in human cells.* EMBO J, 2009. **28**(13): p. 1878-89.

17. Yuan, J., et al., *Histone H3-K56 acetylation is important for genomic stability in mammals.* Cell Cycle, 2009. **8**(11): p. 1747-53.

18. Phanstiel, D., et al., *Mass spectrometry identifies and quantifies 74 unique histone H4 isoforms in differentiating human embryonic stem cells.* Proc Natl Acad Sci U S A, 2008. **105**(11): p. 4093-8.

19. Fitzgerald, J., et al., *Regulation of the DNA Damage Response and Gene Expression by the Dot1L Histone Methyltransferase and the 53Bp1 Tumour Suppressor (accepted).* PLOS One, 2010.

Appendix 1: Supplementary Information and Figures

Table A.1: Histone H3 volcano plot intensity data CTL vs. SAHA 6h Rx

Peptide m/z	Peptide mass	Charge	Uniprot ID	Protein description	Sequence	PTM/Chemical Modification	CTL-1	CTL-2	CTL-3	SAHA 6h-1	SAHA 6h-2	SAHA 6h-3	Avg CTL	Avg SAHA	fold change CTL/SAHA	fold change SAHA/CTL
380.7196	759.4236	2	P68431/P84243	Histone H3.1/H3.3	TKQTAR	K4pr	192735	263714	300090	297881	344462	292742	252180	311695	0.81	1.24
387.7283	773.4409	2	P68431/P84243	Histone H3.1/H3.3	TKQTAR	K4pr+me1	4792703	5684494	6785001	7912475	8144964	10629199	5754066	8895546	0.65	1.55
507.2910	1012.5663	2	P68431/P84243	Histone H3.1/H3.3	KSTGGKAPR	K9pr+K14pr	23700420	27640555	30296523	16315685	16984009	18063249	27212499	17120981	1.59	0.63
514.2999	1026.5841	2	P68431/P84243	Histone H3.1/H3.3	KSTGGKAPR	K9pr+me1+K14pr	11507683	14361273	16159885	6847960	7426358	8543901	14009614	7606073	1.84	0.54
493.2925	984.5694	2	P68431/P84243	Histone H3.1/H3.3	KSTGGKAPR	K9me2+K14pr	7133480	6262590	10366055	1773399	6892073	8689263	7920708	7790668	1.02	0.98
500.3002	998.5848	2	P68431/P84243	Histone H3.1/H3.3	KSTGGKAPR	K9me3+K14pr	650159	579439	1032156	121397	796166	1173713	753918	697092	1.08	0.92
500.2832	998.5508	2	P68431/P84243	Histone H3.1/H3.3	KSTGGKAPR	K9pr+K14ac	15983285	19084819	21146942	25161247	30253902	32053440	18738349	29156196	0.64	1.56
493.2749	984.5342	2	P68431/P84243	Histone H3.1/H3.3	KSTGGKAPR	K9ac+K14ac	636572	847846	1090938	7187669	7895249	10432328	858452	8505082	0.10	**9.91**
507.2901	1012.5646	2	P68431/P84243	Histone H3.1/H3.3	KSTGGKAPR	K9me1+K14ac	179952	228089	325360	588777	798111	866959	244467	751282	0.33	**3.07**
540.2837	1078.5517	2	P68431/P84243	Histone H3.1/H3.3	KSTGGKAPR	K9me3+S10phos+K14pr	231268	187582	220375	129014	196383	207261	213075	177553	1.20	0.83
549.8357	1097.6557	2	P68431/P84243	Histone H3.1/H3.3	KQLATKAAR	K18pr+K23pr	96618395	96664152	100853698	71223467	76629021	82335957	98045415	76729482	1.28	0.78
542.8275	1083.6394	2	P68431/P84243	Histone H3.1/H3.3	KQLATKAAR	K18pr+K23ac	41196100	47703971	48074084	61817919	65985992	69476600	45658052	65760170	0.69	1.44
556.8430	1111.6704	2	P68431/P84243	Histone H3.1/H3.3	KQLATKAAR	K18pr+me1+K23pr	444104	359782	421224	262885	327452	330743	408370	307027	1.33	0.75
535.8195	1069.6233	2	P68431/P84243	Histone H3.1/H3.3	KQLATKAAR	K18ac+K23ac	3520626	3900144	4596917	20591760	20161430	20844000	4005896	20532397	0.20	**5.13**
516.8015	1031.5874	2	P68431/P84243	Histone H3.1/H3.3	YRPGTVALR	none	1353574	1372508	1317450	1278659	1242896	1221855	1347844	1247803	1.08	0.93
653.8736	1305.7316	2	P68431/P84243	Histone H3.1/H3.3	YQKSTELLIR	K56pr	80770721	74555581	64099017	80293262	73075908	68201479	73141773	73856883	0.99	1.01
422.7577	843.4998	2	P68431/P84243	Histone H3.1/H3.3	KLPFQR	K64pr	59519868	50061343	42624602	49877241	50008556	49237441	50735271	49707746	1.02	0.98
696.3619	1390.7082	2	P68431/P84243	Histone H3.1/H3.3	EIAQDFKTDLR	K79pr	85202197	80486516	74650782	87750254	79019707	75428111	80113165	80732690	0.99	1.01
703.3710	1404.7264	2	P68431/P84243	Histone H3.1/H3.3	EIAQDFKTDLR	K79pr+me1	354743	400964	386305	440185	393732	437320	380671	423746	0.90	1.11
682.3646	1362.7136	2	P68431/P84243	Histone H3.1/H3.3	EIAQDFKTDLR	K79me2	573092	497089	419574	471790	549250	561100	496585	527380	0.94	1.06
720.9172	1439.8188	2	P68431/P84243	Histone H3.1/H3.3	VTIMPKDIQLAR	K122pr	59685861	56665423	44875998	54989012	45874233	40403891	53742427	47089045	1.14	0.88
728.9145	1455.8133	2	P68431/P84243	Histone H3.1/H3.3	VTIMPKDIQLAR	M120ox + K122pr	89300550	83054880	82094630	90663720	86884010	88378560	84816687	88642097	0.96	1.05
801.4611	1600.9066	2	P68431	Histone H3.1	KSAPATGGVKKPHR	K27pr+K36pr+K37pr	935406	1009471	994296	1293741	1380662	1507974	979724	1394126	0.70	1.42
808.4690	1614.9223	2	P68431	Histone H3.1	KSAPATGGVKKPHR	K27pr+me1+K36pr+K37pr	2083318	1463196	1796979	1464506	1628806	1713104	1781164	1602139	1.11	0.90
539.3156	1614.9234	3	P68431	Histone H3.1	KSAPATGGVKKPHR	K27pr+me1+K36pr+K37pr	13204758	12078743	13336185	11375332	13596138	14090210	12873229	13020560	0.99	1.01
787.4642	1572.9127	2	P68431	Histone H3.1	KSAPATGGVKKPHR	K27me2+K36pr+K37pr	1288890	1471089	1752451	1677169	1526114	1627701	1504143	1610328	0.93	1.07
794.4714	1586.9272	2	P68431	Histone H3.1	KSAPATGGVKKPHR	K27me3+K36pr+K37pr	527600	387889	587641	581436	690253	577068	501043	616253	0.81	1.23
529.9815	1586.9211	3	P68431	Histone H3.1	KSAPATGGVKKPHR	K27me3+K36pr+K37pr	7828952	10274631	11283457	10889302	12024530	11422445	9795680	11445426	0.86	1.17
515.9797	1544.9158	3	P68431	Histone H3.1	KSAPATGGVKKPHR	K27me2+K36me2+K37pr	691731	1017532	1721796	1326207	2007111	1712807	1143686	1682042	0.68	1.47
387.2379	1544.9204	4	P68431	Histone H3.1	KSAPATGGVKKPHR	K27me2+K36me2+K37pr	89642	202333	345846	450379	619266	568162	212607	545936	0.39	**2.57**
815.4775	1628.9394	2	P68431	Histone H3.1	KSAPATGGVKKPHR	K27pr+me1+K36pr+me1+K37pr	272541	226104	242785	142647	234960	187389	247144	188332	1.31	0.76
543.9866	1628.9364	3	P68431	Histone H3.1	KSAPATGGVKKPHR	K27pr+me1+K36pr+me1+K37pr	1954278	2079490	1744268	1149368	1678902	1781018	1926012	1536429	1.25	0.80
794.4709	1586.9261	2	P68431	Histone H3.1	KSAPATGGVKKPHR	K27pr+me1+K36me2+K37pr	2210028	3162552	3574318	2818872	3101063	3687072	2982299	3202336	0.93	1.07
397.7394	1586.9261	4	P68431	Histone H3.1	KSAPATGGVKKPHR	K27pr+me1+K36me2+K37pr	308837	632602	612579	486179	528326	583414	518006	532640	0.97	1.03
794.4716	1586.9275	2	P68431	Histone H3.1	KSAPATGGVKKPHR	K27me2 + K36me1 + K37pr	224580	231079	309051	236265	267773	313249	254904	272429	0.94	1.07
534.6534	1600.9369	3	P68431	Histone H3.1	KSAPATGGVKKPHR	K27me3 + K36me1 + K37pr	665578	1434768	1471778	1129548	1546011	1889856	1190708	1521805	0.78	1.28
401.2434	1600.9421	4	P68431	Histone H3.1	KSAPATGGVKKPHR	K27me1 + K36me3 +K37pr	179991	272334	351079	234456	301306	420628	267802	318797	0.84	1.19
809.4569	1616.8981	2	P84243	Histone H3.3	KSAPSTGGVKKPHR	K27pr+K36pr+K37pr	203400	135196	145585	145785	130530	167237	161393	147851	1.09	0.92
816.4657	1630.9158	2	P84243	Histone H3.3	KSAPSTGGVKKPHR	K27pr+me1+K36pr+K37pr	185520	108634	126750	123758	106303	137260	140301	122441	1.15	0.87
535.3149	1602.9214	3	P84243	Histone H3.3	KSAPSTGGVKKPHR	K27me1 + K36me2 + K37pr	415023	445586	389926	251899	274846	243643	416845	256796	1.62	0.62
539.9850	1616.9315	3	P84243	Histone H3.3	KSAPSTGGVKKPHR	K27me1 + K36me3 +K37pr	309899	447661	428555	401394	538517	565512	395371	501808	0.79	1.27
549.3166	1644.9263	3	P84243	Histone H3.3	KSAPSTGGVKKPHR	K27me1 + K36me1 + K37pr	410604	412196	360048	245683	313975	322021	394282	293893	1.34	0.75
535.3129	1602.9153	3	P84243	Histone H3.3	KSAPSTGGVKKPHR	K27me3 + K36pr+ K37pr	499713	716374	1002578	773060	1082451	1143635	739555	999715	0.74	1.35
530.6413	1588.9004	3	P84243	Histone H3.3	KSAPSTGGVKKPHR	K27me2 + K36pr+ K37pr	384185	769925	859174	636745	827252	784661	671095	749553	0.90	1.12

pr= propionyl group added by chemical reaction
ac = invivo acetylation
me1= invivo monomethyl.
me2= invivo dimethyl.
me3= invivo trimethyl.
phos= phosphorylation
ox= oxidation

bold values= significant differential expression

Table A.2: Histone H3 volcano plot intensity data CTL vs. SAHA 24h Rx

Peptide m/z	Peptide mass	Charge	Uniprot ID	Protein desciption	Sequence	PTM/Chemical Modification	CTL-1	CTL-2	CTL-3	SAHA 24h-1	SAHA 24h-2	SAHA 24h-3	Avg CTL	Avg SAHA	fold change CTL/SAHA	fold change SAHA/CTL
380.7196	759.4236	2	P68431/P84243	Histone H3.1/H3.3	TKQTAR	K4pr	239490	307771	349890	195266	198523	171374	299050	188388	1.59	0.63
387.7283	773.4409	2	P68431/P84243	Histone H3.1/H3.3	TKQTAR	K4pr+me1	5955350	6634160	7910983	12306839	14189708	16882630	6833498	14459725	0.47	**2.12**
507.2910	1012.5663	2	P68431/P84243	Histone H3.1/H3.3	KSTGGKAPR	K9pr+K14pr	29449833	32258257	35324282	26742922	27567830	28841180	32344124	27717310	1.17	0.86
514.2999	1026.5841	2	P68431/P84243	Histone H3.1/H3.3	KSTGGKAPR	K9pr+me1+K14pr	14299296	16760504	18841645	11985932	14741342	16011139	16633815	14246138	1.17	0.86
493.2925	984.5694	2	P68431/P84243	Histone H3.1/H3.3	KSTGGKAPR	K9me2+K14pr	8863969	7308834	12086318	7063150	9909281	9827620	9419707	8933350	1.05	0.95
500.3002	998.5848	2	P68431/P84243	Histone H3.1/H3.3	KSTGGKAPR	K9me3+K14pr	797358	676242	1165882	566893	688057	879845	879827	711598	1.24	0.81
500.2832	998.5508	2	P68431/P84243	Histone H3.1/H3.3	KSTGGKAPR	K9pr+K14ac	19860622	22273178	24656312	36691113	39119636	40694440	22263370	38835063	0.57	1.74
493.2749	984.5342	2	P68431/P84243	Histone H3.1/H3.3	KSTGGKAPR	K9ac+K14ac	790996	989489	1271981	8424484	8989578	11167978	1017489	9527347	0.11	**9.36**
507.2901	1012.5646	2	P68431/P84243	Histone H3.1/H3.3	KSTGGKAPR	K9me1+K14ac	1491352	1834893	2239743	3135539	3734626	3687388	1855329	3519184	0.53	1.90
540.2837	1078.5517	2	P68431/P84243	Histone H3.1/H3.3	KSTGGKAPR	K9me3+S10phos+K14pr	287370	218920	256946	213502	253124	255776	254412	240801	1.06	0.95
549.8357	1097.6557	2	P68431/P84243	Histone H3.1/H3.3	KQLATKAAR	K18pr+K23pr	131719500	120278400	118115800	78189300	88326000	88529800	123371233	85015033	1.45	0.69
542.8275	1083.6394	2	P68431/P84243	Histone H3.1/H3.3	KQLATKAAR	K18pr+K23ac	56679300	55673518	56052058	73142160	82611399	83182243	56134959	79645267	0.70	1.42
556.8430	1111.6704	2	P68431/P84243	Histone H3.1/H3.3	KQLATKAAR	K18pr+me1+K23pr	551838	419888	491127	296001	383850	422753	487618	367535	1.33	0.75
535.8195	1069.6233	2	P68431/P84243	Histone H3.1/H3.3	KQLATKAAR	K18ac+K23ac	4374684	4528128	5359783	21429660	24478920	25096980	4754198	23668520	0.20	**4.98**
516.8015	1031.5874	2	P68431/P84243	Histone H3.1/H3.3	YRPGTVALR	none	1681933	1601803	1536083	1770120	1608937	1478215	1606606	1619090	0.99	1.01
653.8736	1305.7316	2	P68431/P84243	Histone H3.1/H3.3	YQKSTELLIR	K56pr	100364645	87011027	74736355	80385114	76096931	78262273	87370676	78248106	1.12	0.90
422.7577	843.4998	2	P68431/P84243	Histone H3.1/H3.3	KLPFQR	K64pr	73958612	58424719	49698226	56316160	56699484	57754619	60693852	56923421	1.07	0.94
696.3619	1390.7082	2	P68431/P84243	Histone H3.1/H3.3	EIAQDFKTDLR	K79pr	105871140	93932799	87039203	87398222	85720090	88937485	95614381	87351933	1.09	0.91
703.3710	1404.7264	2	P68431/P84243	Histone H3.1/H3.3	EIAQDFKTDLR	K79pr+me1	440799	467950	450413	949959	856861	828974	453054	878598	0.52	1.94
682.3646	1362.7136	2	P68431/P84243	Histone H3.1/H3.3	EIAQDFKTDLR	K79me2	712117	580133	489203	727823	729341	754011	593818	737058	0.81	1.24
720.9172	1439.8188	2	P68431/P84243	Histone H3.1/H3.3	VTIMPKDIQLAR	K122pr	74164874	66132093	52323245	54628582	52088130	50900356	64206737	52539023	1.22	0.82
728.9145	1455.8133	2	P68431/P84243	Histone H3.1/H3.3	VTIMPKDIQLAR	M120ox + K122pr	110963700	96930240	95718370	100942800	101785200	107210400	101204103	103312800	0.98	1.02
801.4611	1600.9066	2	P68431	Histone H3.1	KSAPATGGVKKPHR	K27pr+K36pr+K37pr	1162323	1178116	1159301	1406580	1236564	1551302	1166580	1398149	0.83	1.20
808.4690	1614.9223	2	P68431	Histone H3.1	KSAPATGGVKKPHR	K27pr+me1+K36pr+K37pr	2588703	1707641	2095190	2571383	3101198	3026899	2130512	2899827	0.73	1.36
539.3156	1614.9234	3	P68431	Histone H3.1	KSAPATGGVKKPHR	K27pr+me1+K36pr+K37pr	16408060	14096649	12208500	18729716	23794722	26148693	14237736	22891044	0.62	1.61
787.4642	1572.9127	2	P68431	Histone H3.1	KSAPATGGVKKPHR	K27me2+K36pr+K37pr	2654528	1987621	1998410	3551710	3484515	2703077	2213520	3246434	0.68	1.47
794.4714	1586.9272	2	P68431	Histone H3.1	KSAPATGGVKKPHR	K27me3+K36pr+K37pr	655590	452690	685161	970860	1129663	1043543	597813	1048022	0.57	1.75
529.9815	1586.9211	3	P68431	Histone H3.1	KSAPATGGVKKPHR	K27me3+K36pr+K37pr	9728154	12381600	13148740	13321140	15493773	15837900	11752831	14884271	0.79	1.27
515.9797	1544.9158	3	P68431	Histone H3.1	KSAPATGGVKKPHR	K27me2+K36me2+K37pr	859536	1187524	2007531	2581764	3409437	2796969	1351530	2929390	0.46	**2.17**
387.2379	1544.9204	4	P68431	Histone H3.1	KSAPATGGVKKPHR	K27me2+K36me2+K37pr	111388	236135	361877	772620	1133445	1399607	236467	1101891	0.21	**4.66**
815.4775	1628.9394	2	P68431	Histone H3.1	KSAPATGGVKKPHR	K27pr+me1+K36pr+me1+K37pr	338656	263878	283076	237969	403821	326939	295203	322910	0.91	1.09
543.9866	1628.9364	3	P68431	Histone H3.1	KSAPATGGVKKPHR	K27pr+me1+K36pr+me1+K37pr	2428360	2426895	2033732	1850940	3255798	3015601	2296329	2707446	0.85	1.18
794.4709	1586.9261	2	P68431	Histone H3.1	KSAPATGGVKKPHR	K27pr+me1+K36me2+K37pr	2746152	3690896	4167482	4054260	4778016	4840712	3534843	4557663	0.78	1.29
397.7394	1586.9261	4	P68431	Histone H3.1	KSAPATGGVKKPHR	K27pr+me1+K36me2+K37pr	383757	738287	714237	772744	933491	852086	612094	852774	0.72	1.39
794.4716	1586.9275	2	P68431	Histone H3.1	KSAPATGGVKKPHR	K27me2+K36me1+K37pr	279060	269684	360339	401348	426499	421702	303028	416516	0.73	1.37
801.4790	1600.9424	2	P68431	Histone H3.1	KSAPATGGVKKPHR	K27me3 + K36me1 + K37pr	166046	101421	102919	122267	136022	229883	123462	162724	0.76	1.32
534.6534	1600.9369	3	P68431	Histone H3.1	KSAPATGGVKKPHR	K27me3 + K36me1 + K37pr	827039	1674464	1716022	2002308	3019908	3273166	1405842	2765127	0.51	1.97
520.6379	1558.8904	3	P68431	Histone H3.1	KSAPATGGVKKPHR	K27me2+K36ac+K37pr	363110	337565	590571	967071	1174955	1298904	430415	1146977	0.38	**2.66**
401.2434	1600.9421	4	P68431	Histone H3.1	KSAPATGGVKKPHR	K27me1 + K36me3 +K37pr	223655	317831	409341	531566	590834	599586	316942	573996	0.55	1.81
809.4569	1616.8981	2	P84243	Histone H3.3	KSAPSTGGVKKPHR	K27pr+K36pr+K37pr	252742	157782	169745	273759	257107	259055	193423	263307	0.73	1.36
816.4657	1630.9158	2	P84243	Histone H3.3	KSAPSTGGVKKPHR	K27pr+me1+K36pr+K37pr	230525	126783	147784	190635	209579	223310	168364	207841	0.81	1.23
535.3149	1602.9214	3	P84243	Histone H3.3	KSAPSTGGVKKPHR	K27me1 + K36me2 + K37pr	515702	520027	454634	612276	464342	518996	496788	531871	0.93	1.07
539.9850	1616.9315	3	P84243	Histone H3.3	KSAPSTGGVKKPHR	K27me1 + K36me3 +K37pr	385076	522448	499674	560102	696120	701301	469066	652508	0.72	1.39
549.3166	1644.9263	3	P84243	Histone H3.3	KSAPSTGGVKKPHR	K27me1 + K36me1 + K37pr	510211	481058	419798	356213	658246	576091	470356	530183	0.89	1.13
535.3129	1602.9153	3	P84243	Histone H3.3	KSAPSTGGVKKPHR	K27me3 + K36pr + K37pr	620937	825189	1168958	1489320	1816992	1924914	871694	1743742	0.50	**2.00**
530.6413	1588.9004	3	P84243	Histone H3.3	KSAPSTGGVKKPHR	K27me2 + K36pr + K37pr	477383	898550	1001756	431075	757080	764280	792563	650812	1.22	0.82

pr= propionyl group added by chemical reaction
ac = invivo acetylation
me1= invivo monomethyl.
me2= invivo dimethyl.
me3= invivo trimethyl.
phos= phosphorylation
ox= oxidation

bold values= significant differential expression

Table A.3: Histone H3 volcano plot intensity data CTL vs. MS-275 6h Rx

Peptide m/z	Peptide mass	Charge	Uniprot ID	Protein desciption	Sequence	PTM/Chemical Modification	CTL-1	CTL-2	CTL-3	MS275 6h-1	MS275 6h-2	MS275 6h-3	Avg CTL	Avg MS275	fold change CTL/MS275	fold change MS275/CTL
380.7196	759.4236	2	P68431/P84243	Histone H3.1/H3.3	TKQTAR	K4pr	246705	297932	363983	344734	279094	242167	302873	288665	1.05	0.95
387.7283	773.4409	2	P68431/P84243	Histone H3.1/H3.3	TKQTAR	K4pr+me1	6134764	6422061	8203483	7635133	9892976	10860314	6920102	9462808	0.73	1.37
507.2910	1012.5663	2	P68431/P84243	Histone H3.1/H3.3	KSTGGKAPR	K9pr+K14pr	30337054	31226933	36630358	25766030	26961508	29063597	32731448	27263712	1.20	0.83
514.2999	1026.5841	2	P68431/P84243	Histone H3.1/H3.3	KSTGGKAPR	K9pr+me1+K14pr	14730085	16224657	19538294	12322463	14503890	16090604	16831012	14305652	1.18	0.85
493.2925	984.5694	2	P68431/P84243	Histone H3.1/H3.3	KSTGGKAPR	K9me2+K14pr	9131009	7075164	12533197	5285923	8816525	9129434	9579790	7743961	1.24	0.81
500.3002	998.5848	2	P68431/P84243	Histone H3.1/H3.3	KSTGGKAPR	K9me3+K14pr	821380	654622	738001	443978	695544	581955	738001	573825	1.29	0.78
500.2832	998.5508	2	P68431/P84243	Histone H3.1/H3.3	KSTGGKAPR	K9pr+K14ac	20458953	21561086	25567952	27253429	34241414	36634189	22529330	32709677	0.69	1.45
493.2749	984.5342	2	P68431/P84243	Histone H3.1/H3.3	KSTGGKAPR	K9ac+K14ac	814825	957854	1319011	3955053	5670499	5558367	1030564	5061306	0.20	**4.91**
507.2901	1012.5646	2	P68431/P84243	Histone H3.1/H3.3	KSTGGKAPR	K9me1+K14ac	230342	257683	393381	503230	600315	685210	293802	596252	0.49	**2.03**
540.2837	1078.5517	2	P68431/P84243	Histone H3.1/H3.3	KSTGGKAPR	K9me3+S10phos+K14pr	296028	211921	266447	249202	361665	443274	258132	351381	0.73	1.36
549.8357	1097.6557	2	P68431/P84243	Histone H3.1/H3.3	KQLATKAAR	K18pr+K23pr	136345630	115932658	122483000	82117758	82284728	96355100	124920429	86919195	1.44	0.70
542.8275	1083.6394	2	P68431/P84243	Histone H3.1/H3.3	KQLATKAAR	K18pr+K23ac	53290823	53893589	58124520	69032962	69582311	79579530	55102977	72731601	0.76	1.32
556.8430	1111.6704	2	P68431/P84243	Histone H3.1/H3.3	KQLATKAAR	K18pr+me1+K23pr	568463	406464	509285	603231	529180	654526	494737	595646	0.83	1.20
535.8195	1069.6233	2	P68431/P84243	Histone H3.1/H3.3	KQLATKAAR	K18ac+K23ac	4506478	4406190	5557955	10995110	11826000	13643200	4823541	12154770	0.40	**2.52**
516.8015	1031.5874	2	P68431/P84243	Histone H3.1/H3.3	YRPGTVALR	none	1732604	1550592	1592878	1951439	1650280	1925532	1625358	1842417	0.88	1.13
653.8736	1305.7316	2	P68431/P84243	Histone H3.1/H3.3	YQKSTELLIR	K56pr	103388283	84229212	77499648	84320741	80920663	81356932	88372381	82199446	1.08	0.93
422.7577	843.4998	2	P68431/P84243	Histone H3.1/H3.3	KLPFQR	K64pr	76186728	56556832	51535762	52520783	49473292	55874798	61426441	52622958	1.17	0.86
696.3619	1390.7082	2	P68431/P84243	Histone H3.1/H3.3	EIAQDFKTDLR	K79pr	109060669	90929689	90257381	89129497	85664018	91460854	96749246	88751456	1.09	0.92
703.3710	1404.7264	2	P68431/P84243	Histone H3.1/H3.3	EIAQDFKTDLR	K79pr+me1	454079	452989	467067	730247	486729	612113	458045	609696	0.75	1.33
682.3646	1362.7136	2	P68431/P84243	Histone H3.1/H3.3	EIAQDFKTDLR	K79me2	733571	561586	507291	849166	830057	938467	600816	872563	0.69	1.45
720.9172	1439.8188	2	P68431/P84243	Histone H3.1/H3.3	VTIMPKDIQLAR	K122pr	76399203	64017795	54257838	62138204	56688943	54165319	64891612	57664155	1.13	0.89
728.9145	1455.8133	2	P68431/P84243	Histone H3.1/H3.3	VTIMPKDIQLAR	M120ox + K122pr	114306650	93831300	99257450	95897200	92242800	97207800	102465133	95115933	1.08	0.93
801.4611	1600.9066	2	P68431	Histone H3.1	KSAPATGGVKKPHR	K27pr+K36pr+K37pr	1197340	1140451	1202165	1407133	1633114	1954827	1179985	1665024	0.71	1.41
808.4690	1614.9223	2	P68431	Histone H3.1	KSAPATGGVKKPHR	K27pr+me1+K36pr+K37pr	2666692	1653046	2172658	2155343	2629962	2724405	2164132	2503237	0.86	1.16
539.3156	1614.9234	3	P68431	Histone H3.1	KSAPATGGVKKPHR	K27pr+me1+K36pr+K37pr	16902378	13645967	16124266	15823064	19743576	22549147	15557537	19371929	0.80	1.25
787.4642	1572.9127	2	P68431	Histone H3.1	KSAPATGGVKKPHR	K27me2+K36pr+K37pr	1649807	1661964	2118821	1810812	2202572	2229587	1810197	2080990	0.87	1.15
794.4714	1586.9272	2	P68431	Histone H3.1	KSAPATGGVKKPHR	K27me3+K36pr+K37pr	675340	438217	710494	639890	935024	924178	608017	833031	0.73	1.37
529.9815	1586.9211	3	P68431	Histone H3.1	KSAPATGGVKKPHR	K27me3+K36pr+K37pr	10021229	11607771	13634900	11821810	13323960	14775819	11754633	13307196	0.88	1.13
515.9797	1544.9158	3	P68431	Histone H3.1	KSAPATGGVKKPHR	K27me2+K36me2+K37pr	885431	1149558	1801607	2324029	2234410	3063074	1278865	2540504	0.50	1.99
387.2379	1544.9204	4	P68431	Histone H3.1	KSAPATGGVKKPHR	K27me2+K36me2+K37pr	114744	205470	361877	638158	606280	1117037	227364	787158	0.29	**3.46**
815.4775	1628.9394	2	P68431	Histone H3.1	KSAPATGGVKKPHR	K27pr+me1+K36pr+me1+K37pr	348859	255441	293542	359562	341592	354712	299281	351955	0.85	1.18
543.9866	1628.9364	3	P68431	Histone H3.1	KSAPATGGVKKPHR	K27pr+me1+K36pr+me1+K37pr	2501518	2349305	2108927	3234439	2857558	3312047	2319917	3134681	0.74	1.35
794.4709	1586.9261	2	P68431	Histone H3.1	KSAPATGGVKKPHR	K27pr+me1+K36me2+K37pr	2828884	3572895	4321570	4381510	4083912	5372010	3574450	4612477	0.77	1.29
397.7394	1586.9261	4	P68431	Histone H3.1	KSAPATGGVKKPHR	K27pr+me1+K36me2+K37pr	395318	714683	740645	968097	786148	841704	616882	865316	0.71	1.40
794.4716	1586.9275	2	P68431	Histone H3.1	KSAPATGGVKKPHR	K27me2+K36me1+K37pr	287467	261062	373662	376164	440122	470886	307397	429057	0.72	1.40
801.4790	1600.9424	2	P68431	Histone H3.1	KSAPATGGVKKPHR	K27me3 + K36me1 + K37pr	92121	98178	106724	121768	158041	172775	99008	150861	0.66	1.52
534.6534	1600.9369	3	P68431	Histone H3.1	KSAPATGGVKKPHR	K27me3 + K36me1 + K37pr	851955	1620930	1779470	2099818	2459808	3231733	1417452	2597120	0.55	1.83
401.2434	1600.9421	4	P68431	Histone H3.1	KSAPATGGVKKPHR	K27me1 + K36me3 + K37pr	230393	307670	424476	423751	333162	565760	320846	440891	0.73	1.37
809.4569	1616.8981	2	P84243	Histone H3.3	KSAPSTGGVKKPHR	K27pr+K36pr+K37pr	260356	152737	176021	163656	188836	228369	196371	193620	1.01	0.99
816.4657	1630.9158	2	P84243	Histone H3.3	KSAPSTGGVKKPHR	K27pr+me1+K36pr+K37pr	237470	122730	153248	190395	224213	201930	171149	205513	0.83	1.20
535.3149	1602.9214	3	P84243	Histone H3.3	KSAPSTGGVKKPHR	K27me1 + K36me2 + K37pr	366750	303718	287817	259986	306698	434750	319428	333811	0.96	1.05
539.9850	1616.9315	3	P84243	Histone H3.3	KSAPSTGGVKKPHR	K27me1 + K36me3 +K37pr	406756	505745	518149	554354	621618	681658	476883	619210	0.77	1.30
549.3166	1644.9263	3	P84243	Histone H3.3	KSAPSTGGVKKPHR	K27me1 + K36me1 + K37pr	525582	465678	435320	662376	615167	673496	475527	650346	0.73	1.37
535.3129	1602.9153	3	P84243	Histone H3.3	KSAPSTGGVKKPHR	K27me3 + K36pr+ K37pr	639644	798807	1212178	1264851	1687176	1773616	883543	1575214	0.56	1.78
530.6413	1588.9004	3	P84243	Histone H3.3	KSAPSTGGVKKPHR	K27me2 +K36pr + K37pr	491765	869823	1038795	677894	827820	1236415	800128	914043	0.88	1.14

pr= propionyl group added by chemical reaction
ac = invivo acetylation
me1= invivo monomethyl.
me2= invivo dimethyl.
me3= invivo trimethyl.
phos= phosphorylation
ox= oxidation

bold values= significant differential expression

Table A.4: Histone H3 volcano plot intensity data CTL vs. MS-275 24h Rx

Peptide m/z	Peptide mass	Charge	Uniprot ID	Protein desciption	Sequence	PTM/Chemical Modification	CTL-1	CTL-2	CTL-3	MS275 24h-1	MS275 24h-2	MS275 24h-3	Avg CTL	Avg MS275	Fold change CTL/MS275	Fold change MS275/CTL
380.7196	759.4236	2	P68431/P84243	Histone H3.1/H3.3	TKQTAR	K4pr	191490	246254	302539	292672	296019	272724	246761	287138	0.86	1.16
387.7283	773.4409	2	P68431/P84243	Histone H3.1/H3.3	TKQTAR	K4pr+me1	4761744	5308116	6840377	9440935	12173809	11697055	5636745	11103933	0.51	1.97
507.2910	1012.5663	2	P68431/P84243	Histone H3.1/H3.3	KSTGGKAPR	K9pr+K14pr	23547324	25810435	30543790	16687406	19750678	22446737	26633850	19628274	1.36	0.74
514.2999	1026.5841	2	P68431/P84243	Histone H3.1/H3.3	KSTGGKAPR	K9pr+me1+K14pr	11433347	13410393	16291775	6315883	7180687	9039319	13711838	7511963	1.83	0.55
493.2925	984.5694	2	P68431/P84243	Histone H3.1/H3.3	KSTGGKAPR	K9me2+K14pr	7087400	5847935	10450658	1314537	4952659	2481888	7795331	2916361	**2.67**	0.37
500.3002	998.5848	2	P68431/P84243	Histone H3.1/H3.3	KSTGGKAPR	K9me3+K14pr	1447746	1810007	1910760	1062851	1547351	1885224	1722838	1498475	1.15	0.87
500.2832	998.5508	2	P68431/P84243	Histone H3.1/H3.3	KSTGGKAPR	K9pr+K14ac	15880038	17821187	21319534	26581038	28963322	33483560	18340253	29675974	0.62	1.62
493.2749	984.5342	2	P68431/P84243	Histone H3.1/H3.3	KSTGGKAPR	K9ac+K14ac	632460	791709	1099842	11589838	15927932	14310088	841337	13942619	0.06	**16.57**
507.2901	1012.5646	2	P68431/P84243	Histone H3.1/H3.3	KSTGGKAPR	K9me1+K14ac	178789	212987	328016	550872	758224	647090	239931	652062	0.37	**2.72**
540.2837	1078.5517	2	P68431/P84243	Histone H3.1/H3.3	KSTGGKAPR	K9me3+S10phos+K14pr	229774	175162	222173	139738	183749	274692	209036	199393	1.05	0.95
549.8357	1097.6557	2	P68431/P84243	Histone H3.1/H3.3	KQLATKAAR	K18pr+K23pr	103298663	95823445	101676822	58222100	62805160	64121971	100266310	61716410	1.62	0.62
542.8275	1083.6394	2	P68431/P84243	Histone H3.1/H3.3	KQLATKAAR	K18pr+K23ac	44781877	44545424	48466443	75050364	78552075	79507305	45931248	77703248	0.59	1.69
556.8430	1111.6704	2	P68431/P84243	Histone H3.1/H3.3	KQLATKAAR	K18pr+me1+K23pr	441235	335961	424662	270985	257580	315891	400619	281485	1.42	0.70
535.8195	1069.6233	2	P68431/P84243	Histone H3.1/H3.3	KQLATKAAR	K18ac+K23ac	3497884	3641910	4634435	28966470	28933470	31926440	3924743	29942127	0.13	**7.63**
516.8015	1031.5874	2	P68431/P84243	Histone H3.1/H3.3	YRPGTVALR	none	1344830	1281632	1328202	1379116	1337485	1227076	1318222	1314559	1.00	1.00
653.8736	1305.7316	2	P68431/P84243	Histone H3.1/H3.3	YQKSTELLIR	K56pr	80248970	69619152	64622165	89449109	82967357	83627927	71496762	85348131	0.84	1.19
422.7577	843.4998	2	P68431/P84243	Histone H3.1/H3.3	KLPFQR	K64pr	59135390	46746711	42972485	53402861	42399202	31615240	49618196	42472434	1.17	0.86
696.3619	1390.7082	2	P68431/P84243	Histone H3.1/H3.3	EIAQDFKTDLR	K79pr	84651821	75157391	75260049	91075345	85056969	85806044	78356420	87312786	0.90	1.11
703.3710	1404.7264	2	P68431/P84243	Histone H3.1/H3.3	EIAQDFKTDLR	K79pr+me1	352451	374416	389458	457327	359803	378706	372108	398612	0.93	1.07
682.3646	1362.7136	2	P68431/P84243	Histone H3.1/H3.3	EIAQDFKTDLR	K79me2	569390	464176	422999	568734	633156	577853	485522	593248	0.82	1.22
720.9172	1439.8188	2	P68431/P84243	Histone H3.1/H3.3	VTIMPKDIQLAR	K122pr	59300312	52913526	45242256	50429594	42118248	40282254	52485365	44276699	1.19	0.84
728.9145	1455.8133	2	P68431/P84243	Histone H3.1/H3.3	VTIMPKDIQLAR	M120ox + K122pr	88723700	77555700	82764650	103236900	101219400	96803200	83014683	100419833	0.83	1.21
801.4611	1600.9066	2	P68431	Histone H3.1	KSAPATGGVKKPHR	K27pr+K36pr+K37pr	929364	942633	1002411	1410557	1846133	1998266	958136	1751652	0.55	1.83
808.4690	1614.9223	2	P68431	Histone H3.1	KSAPATGGVKKPHR	K27pr+me1+K36pr+K37pr	2069860	1366315	1811645	2037439	1865210	2104843	1749273	2002498	0.87	1.14
539.3156	1614.9234	3	P68431	Histone H3.1	KSAPATGGVKKPHR	K27pr+me1+K36pr+K37pr	13119460	11278992	13681700	14684893	15739556	16858458	12693384	15760969	0.81	1.24
787.4642	1572.9127	2	P68431	Histone H3.1	KSAPATGGVKKPHR	K27me2+K36pr+K37pr	1280564	1373686	1766754	1443312	1690173	1694056	1473668	1609180	0.92	1.09
794.4714	1586.9272	2	P68431	Histone H3.1	KSAPATGGVKKPHR	K27me3+K36pr+K37pr	409015	364077	399636	666178	609217	714476	390909	663290	0.59	1.70
529.9815	1586.9211	3	P68431	Histone H3.1	KSAPATGGVKKPHR	K27me3+K36pr+K37pr	7778380	9594334	11369300	11827140	13177620	12751960	9580671	12585573	0.76	1.31
515.9797	1544.9158	3	P68431	Histone H3.1	KSAPATGGVKKPHR	K27me2+K36me2+K37pr	687263	950160	1735848	2274047	2243506	3244070	1124424	2587208	0.43	**2.30**
387.2379	1544.9204	4	P68431	Histone H3.1	KSAPATGGVKKPHR	K27me2+K36me2+K37pr	89063	188936	348668	682921	780654	1073692	208889	845755	0.25	**4.05**
815.4775	1628.9394	2	P68431	Histone H3.1	KSAPATGGVKKPHR	K27pr+me1+K36pr+me1+K37pr	270781	211134	244767	190829	233770	284944	242227	236514	1.02	0.98
543.9866	1628.9364	3	P68431	Histone H3.1	KSAPATGGVKKPHR	K27pr+me1+K36pr+me1+K37pr	1941654	1941804	1758504	1807379	1770828	2078945	1880654	1885717	1.00	1.00
794.4709	1586.9261	2	P68431	Histone H3.1	KSAPATGGVKKPHR	K27pr+me1+K36me2+K37pr	2115643	3679100	4404412	4360005	4698108	5501028	3399718	4853047	0.70	1.43
397.7394	1586.9261	4	P68431	Histone H3.1	KSAPATGGVKKPHR	K27pr+me1+K36me2+K37pr	306842	590717	617578	499648	588896	622270	505046	570271	0.89	1.13
794.4716	1586.9275	2	P68431	Histone H3.1	KSAPATGGVKKPHR	K27me2+K36me1+K37pr	223129	215779	311574	254493	277593	243195	250161	258427	0.97	1.03
801.4790	1600.9424	2	P68431	Histone H3.1	KSAPATGGVKKPHR	K27me3 + K36me1 + K37pr	101723	106626	222558	112976	132999	235810	143636	160595	0.89	1.12
534.6534	1600.9369	3	P68431	Histone H3.1	KSAPATGGVKKPHR	K27me3 + K36me1 + K37pr	661279	1339770	1483790	1563588	1986192	2336308	1161613	1962029	0.59	1.69
401.2434	1600.9421	4	P68431	Histone H3.1	KSAPATGGVKKPHR	K27me1 + K36me3 +K37pr	178829	254302	353945	268341	308288	366166	262359	314265	0.83	1.20
809.4569	1616.8981	2	P84243	Histone H3.3	KSAPSTGGVKKPHR	K27pr+K36pr+K37pr	202086	126244	146773	182503	205282	246076	158368	211287	0.75	1.33
816.4657	1630.9158	2	P84243	Histone H3.3	KSAPSTGGVKKPHR	K27pr+me1+K36pr+K37pr	184322	101441	127784	161398	164063	245643	137849	190368	0.72	1.38
535.3149	1602.9214	3	P84243	Histone H3.3	KSAPSTGGVKKPHR	K27me1 + K36me2 + K37pr	327485	254653	281412	435447	387188	385279	287850	402638	0.71	1.40
539.9850	1616.9315	3	P84243	Histone H3.3	KSAPSTGGVKKPHR	K27me1 + K36me3 +K37pr	315720	418021	432053	396918	503342	478777	388598	459679	0.85	1.18
549.3166	1644.9263	3	P84243	Histone H3.3	KSAPSTGGVKKPHR	K27me1 + K36me1 + K37pr	407952	384904	362986	303417	412371	375185	385280	363658	1.06	0.94
535.3129	1602.9153	3	P84243	Histone H3.3	KSAPSTGGVKKPHR	K27me3 + K36pr + K37pr	496485	660249	1010761	1057770	1170740	1235273	722498	1154594	0.63	1.60
530.6413	1588.9004	3	P84243	Histone H3.3	KSAPSTGGVKKPHR	K27me2 + K36pr + K37pr	381703	718947	866187	821886	1050390	1787136	655612	1219804	0.54	1.86

pr= propionyl group added by chemical reaction
ac = invivo acetylation
me1= invivo monomethyl.
me2= invivo dimethyl.
me3= invivo trimethyl.
phos= phosphorylation
ox= oxidation

bold values= significant differential expression

Table A.5: Histone H4 volcano plot intensity data CTL vs. SAHA 6h Rx

Peptide m/z	Peptide mass	Charge	Uniprot ID	Protein desciption	Sequence	PTM/Chemical modification	CTL-1	CTL-2	CTL-3	SAHA 6h-1	SAHA 6h-2	SAHA 6h-3	Avg CTL	Avg SAHA	fold change CTL/SAHA	fold change SAHA/CTL
747.9414	1493.8672	2	P62805	Histone H4	$GK_5GGK_8GLGK_{12}GGAK_{16}R$	K5pr-K8pr-K12pr-K16pr	46181593	56557915	48164281	29397382	28735108	24910804	50301263	27681098	1.82	0.55
494.2913	1479.8504	3	P62805	Histone H4	$GK_5GGK_8GLGK_{12}GGAK_{16}R$	Mono-acetyl: K16ac or K12ac	9161871	8024100	7091263	10293531	10450402	9892890	8092411	10212274	0.79	1.26
733.9251	1465.8345	2	P62805	Histone H4	$GK_5GGK_8GLGK_{12}GGAK_{16}R$	Di-acetyl: K8ac,K12ac or K12ac,K16ac	6624350	6405618	6982682	20710501	19819755	18212158	6670883	19580805	0.34	**2.94**
489.6194	1465.8348	3	P62805	Histone H4	$GK_5GGK_8GLGK_{12}GGAK_{16}R$	Di-acetyl: K8ac,K16ac	1823874	1730691	1837160	6251017	6207279	6005672	1797242	6154656	0.29	**3.42**
726.9175	1451.8193	2	P62805	Histone H4	$GK_5GGK_8GLGK_{12}GGAK_{16}R$	Tri-acetyl: K5pr-K8ac-K12ac-K16ac	1579110	1464774	1334279	11490814	9316265	10893965	1459388	10567015	0.14	**7.24**
719.9112	1437.8067	2	P62805	Histone H4	$GK_5GGK_8GLGK_{12}GGAK_{16}R$	Tetra-acetyl: K5ac-K8ac-K12ac-K16ac	452345	454885	517277	9358078	7114344	7952371	474836	8141598	0.06	**17.15**
650.0613	1947.1604	3	P62805	Histone H4	KVLRDNIQGITKPAIR	K20pr+me1 + K31pr	142167	141956	113679	106872	106729	97859	132601	103820	1.28	0.78
636.0575	1905.1490	3	P62805	Histone H4	KVLRDNIQGITKPAIR	K20me2 + K31pr	414647	438809	466277	422895	455532	539637	439911	472688	0.93	1.07
691.3959	1380.7762	2	P62805	Histone H4	DNIQGITKPAIR	K31pr	52743600	52808644	54296981	58676800	52958880	60475800	53283075	57370493	0.93	1.08
698.4020	1394.7883	2	P62805	Histone H4	DNIQGITKPAIR	K31pr+me1	16090	16442	14157	15298	13613	21594	15563	16835	0.92	1.08
590.8152	1179.6147	2	P62805	Histone H4	ISGLIYEETR	none	5264101	5302053	6349053	7188693	7641365	9111473	5638402	7980510	0.71	1.42
721.9421	1441.8685	2	P62805	Histone H4	GVLKVFLENVIR	K59pr	7584565	10108180	12290338	10793492	12015959	15586549	9994361	12798667	0.78	1.28
673.8383	1345.6609	2	P62805	Histone H4	DAVTYTEHAKR	K77pr	1075113	1095700	1209936	1194584	1157469	1385552	1126916	1245868	0.90	1.11
853.9816	1705.9476	2	P62805	Histone H4	KTVTAMDVVYALKR	K79pr + K91pr	5919022	7592056	7105709	6053409	8806120	5835337	6872262	6898289	1.00	1.00
861.9783	1721.9409	2	P62805	Histone H4	KTVTAMDVVYALKR	K79pr + M84ox + K91pr	1941660	3394112	2870491	1930656	3070080	4657422	2735421	3219386	0.85	1.18

pr= propionyl group added by chemical reaction
ac = invivo acetylation
me1= invivo monomethyl.
me2= invivo dimethyl.
ox= oxidation

bold values= significant differential expression

Table A.6: Histone H4 volcano plot intensity data CTL vs. SAHA 24h Rx

Peptide m/z	Peptide mass	Charge	Uniprot ID	Protein desciption	Sequence	PTM/Chemical modification	CTL-1	CTL-2	CTL-3	SAHA 24h-1	SAHA 24h-2	SAHA 24h-3	Avg CTL	Avg SAHA	fold change CTL/SAHA	fold change SAHA/CTL
747.9414	1493.8672	2	P62805	Histone H4	$GK_5GGK_8GLGK_{12}GGAK_{16}R$	K5pr-K8pr-K12pr-K16pr	51430805	58421865	53196155	36149473	40448140	32417064	54349608	36338226	1.50	0.67
494.2913	1479.8504	3	P62805	Histone H4	$GK_5GGK_8GLGK_{12}GGAK_{16}R$	Mono-acetyl: K16ac or K12ac	10203252	8288546	7615769	11325566	16230922	12843358	8702522	13466615	0.65	1.55
733.9251	1465.8345	2	P62805	Histone H4	$GK_5GGK_8GLGK_{12}GGAK_{16}R$	Di-acetyl: K8ac,K12ac or K12ac,K16ac	7377304	6616724	6982682	20001538	20241807	17118279	6992237	19120541	0.37	**2.73**
489.6194	1465.8348	3	P62805	Histone H4	$GK_5GGK_8GLGK_{12}GGAK_{16}R$	Di-acetyl: K8ac,K16ac	2031184	1787729	1837160	5948596	6577134	4854373	1885358	5793367	0.33	**3.07**
726.9175	1451.8193	2	P62805	Histone H4	$GK_5GGK_8GLGK_{12}GGAK_{16}R$	Tri-acetyl: K5pr-K8ac-K12ac-K16ac	1758598	1513048	1432969	9875552	8592513	8275090	1568205	8914385	0.18	**5.68**
719.9112	1437.8067	2	P62805	Histone H4	$GK_5GGK_8GLGK_{12}GGAK_{16}R$	Tetra-acetyl: K5ac-K8ac-K12ac-K16ac	503761	469876	594787	4556669	3540912	3339763	522808	3812448	0.14	**7.29**
650.0613	1947.1604	3	P62805	Histone H4	KVLRDNIQGITKPAIR	K20pr+me1 + K31pr	158327	146635	130713	176004	151632	234000	145225	187212	0.78	1.29
636.0575	1905.1490	3	P62805	Histone H4	KVLRDNIQGITKPAIR	K20me2 + K31pr	461777	453271	439746	758075	549850	994249	451598	767391	0.59	1.70
691.3959	1380.7762	2	P62805	Histone H4	DNIQGITKPAIR	K31pr	58738680	54549031	53298010	57744480	64356630	62789580	55528574	61630230	0.90	1.11
698.4020	1394.7883	2	P62805	Histone H4	DNIQGITKPAIR	K31pr+me1	18100	17900	14970	16200	12300	19700	16990	16067	1.06	0.95
590.8152	1179.6147	2	P62805	Histone H4	ISGLIYEETR	none	5862443	5476790	5565559	8957828	8200006	9252943	5634931	8803592	0.64	1.56
721.9421	1441.8685	2	P62805	Histone H4	GVLKVFLENVIR	K59pr	8446662	10441310	14131948	12174065	7761776	13170150	11006640	11035330	1.00	1.00
673.8383	1345.6609	2	P62805	Histone H4	DAVTYTEHAKR	K77pr	1197315	1131810	1209936	1607943	1322178	1527075	1179687	1485732	0.79	1.26
853.9816	1705.9476	2	P62805	Histone H4	KTVTAMDVVYALKR	K79pr + K91pr	6591805	7842263	6059672	7114824	7508762	8241676	6831247	7621754	0.90	1.12
861.9783	1721.9409	2	P62805	Histone H4	KTVTAMDVVYALKR	K79pr + M84ox + K91pr	1583618	3575018	2690912	3401052	3296829	2633400	2616516	3110427	0.84	1.19

pr= propionyl group
added by chemical
reaction
ac = invivo acetylation
me1= invivo monomethyl.
me2= invivo dimethyl.
ox= oxidation

bold values= significant
differential expression

Table A.7: Histone H4 volcano plot intensity data CTL vs. MS-275 6h Rx

Peptide m/z	Peptide mass	Charge	Uniprot ID	Protein desciption	Sequence	PTM/Chemical Modification	CTL-1	CTL-2	CTL-3	MS275 6h-1	MS275 6h-2	MS275 6h-3	Avg CTL	Avg MS275	fold change CTL/MS275	fold change MS275/CTL
747.9414	1493.8672	2	P62805	Histone H4	$GK_5GGK_8GLGK_{12}GGAK_{16}R$	K5pr-K8pr-K12pr-K16pr	53935894	71783190	72860495	52508994	42901600	41278330	66193193	45562975	1.45	0.69
494.2913	1479.8504	3	P62805	Histone H4	$GK_5GGK_8GLGK_{12}GGAK_{16}R$	Mono-acetyl: K16ac or K12ac	9279381	9571888	8635725	13909229	9451856	8850030	9162331	10737038	0.85	1.17
733.9251	1465.8345	2	P62805	Histone H4	$GK_5GGK_8GLGK_{12}GGAK_{16}R$	Di-acetyl: K8ac,K12ac or K12ac,K16ac	7736637	8129997	8148473	22489078	23454278	24623564	8005036	23522307	0.34	**2.94**
489.6194	1465.8348	3	P62805	Histone H4	$GK_5GGK_8GLGK_{12}GGAK_{16}R$	Di-acetyl: K8ac,K16ac	1888068	1845874	1837160	6516784	6304362	6321439	1857034	6380862	0.29	**3.44**
726.9175	1451.8193	2	P62805	Histone H4	$GK_5GGK_8GLGK_{12}GGAK_{16}R$	Tri-acetyl: K5pr-K8ac-K12ac-K16ac	1634689	1562259	1624882	9276528	7238000	9575091	1607277	8696540	0.18	**5.41**
719.9112	1437.8067	2	P62805	Histone H4	$GK_5GGK_8GLGK_{12}GGAK_{16}R$	Tetra-acetyl: K5ac-K8ac-K12ac-K16ac	528298	577339	620732	4560635	3554682	3772251	575456	3962522	0.15	**6.89**
650.0613	1947.1604	3	P62805	Histone H4	KVLRDNIQGITKPAIR	K20pr+me1 + K31pr	166038	180171	136415	98732	102732	90592	160875	97352	1.65	0.61
636.0575	1905.1490	3	P62805	Histone H4	KVLRDNIQGITKPAIR	K20me2 + K31pr	484270	556935	499532	304043	353230	329639	513579	328970	1.56	0.64
691.3959	1380.7762	2	P62805	Histone H4	DNIQGITKPAIR	K31pr	59770276	67024622	63345772	57370484	49501029	56467619	63380224	54446377	1.16	0.86
698.4020	1394.7883	2	P62805	Histone H4	DNIQGITKPAIR	K31pr+me1	18791	20868	14157	20142	21841	19472	17939	20485	0.88	1.14
590.8152	1179.6147	2	P62805	Histone H4	ISGLIYEETR	none	6147991	6729355	6721095	8558475	8370813	8558457	6532813	8495915	0.77	1.30
721.9421	1441.8685	2	P62805	Histone H4	GVLKVFLENVIR	K59pr	8858081	12829282	14748406	12289069	10398827	11570115	12145256	11419337	1.06	0.94
673.8383	1345.6609	2	P62805	Histone H4	DAVTYTEHAKR	K77pr	16650923	15863487	13810812	24367105	20009098	20563493	15441741	21646565	0.71	1.40
853.9816	1705.9476	2	P62805	Histone H4	KTVTAMDVVYALKR	K79pr + K91pr	6912879	9635822	7023766	10138527	11476050	13880282	7857489	11831620	0.66	1.51
861.9783	1721.9409	2	P62805	Histone H4	KTVTAMDVVYALKR	K79pr + M84ox + K91pr	2267682	4307800	3415300	3020693	2277431	3773484	3330261	3023869	1.10	0.91

pr= propionyl group
added by chemical
reaction
ac = invivo acetylation
me1= invivo monomethyl.
me2= invivo dimethyl.
ox= oxidation

bold values= significant
differential expression

Table A.8: Histone H4 volcano plot intensity data CTL vs. MS-275 24h Rx

Peptide m/z	Peptide mass	Charge	Uniprot ID	Protein description	Sequence	PTM/Chemical Modification	CTL-1	CTL-2	CTL-3	MS275 24h-1	MS275 24h-2	MS275 24h-3	Avg CTL	Avg MS275	fold change CTL/MS275	fold change MS275/CTL
747.9414	1493.8672	2	P62805	Histone H4	$GK_5GGK_8GLGK_{12}GGAK_{16}R$	K5pr-K8pr-K12pr-K16pr	53701639	57263682	57496681	16550297	28196324	22168488	56154001	22305036	**2.52**	0.40
494.2913	1479.8504	3	P62805	Histone H4	$GK_5GGK_8GLGK_{12}GGAK_{16}R$	Mono-acetyl: K16ac or K12ac	8224943	8558127	7450371	12460425	10071015	12268297	8077814	11599912	0.70	1.44
733.9251	1465.8345	2	P62805	Histone H4	$GK_5GGK_8GLGK_{12}GGAK_{16}R$	Di-acetyl: K8ac,K12ac or K12ac,K16ac	7703035	6485551	7285189	28250788	25209874	32313104	7157925	28591255	0.25	**3.99**
489.6194	1465.8348	3	P62805	Histone H4	$GK_5GGK_8GLGK_{12}GGAK_{16}R$	Di-acetyl: K8ac,K16ac	1888068	1845874	1837160	8648348	6755281	8447779	1857034	7950469	0.23	**4.28**
726.9175	1451.8193	2	P62805	Histone H4	$GK_5GGK_8GLGK_{12}GGAK_{16}R$	Tri-acetyl: K5pr-K8ac-K12ac-K16ac	1634689	1562259	1452735	19353376	16774806	22183310	1549894	19437164	0.08	**12.54**
719.9112	1437.8067	2	P62805	Histone H4	$GK_5GGK_8GLGK_{12}GGAK_{16}R$	Tetra-acetyl: K5ac-K8ac-K12ac-K16ac	526003	460561	614144	10612387	13935881	13076695	533569	12541654	0.04	**23.51**
650.0613	1947.1604	3	P62805	Histone H4	KVLRDNIQGITKPAIR	K20pr+me1 +K31pr	165317	143728	134967	35369	21793	48924	148004	35362	**4.19**	0.24
636.0575	1905.1490	3	P62805	Histone H4	KVLRDNIQGITKPAIR	K20me2	429241	468013	475140	281543	231728	356938	457465	290070	1.58	0.63
691.3959	1380.7762	2	P62805	Histone H4	DNIQGITKPAIR	K31pr	61332180	53467625	59483519	49640932	57757032	65911948	58094441	57769971	1.01	0.99
698.4020	1394.7883	2	P62805	Histone H4	DNIQGITKPAIR	K31pr+me1	19096	16992	14211	14968	15969	18448	16767	16462	1.02	0.98
590.8152	1179.6147	2	P62805	Histone H4	ISGLIYEETR	none	6121289	5368216	5744752	7418526	8797575	8645263	5744752	8287121	0.69	1.44
721.9421	1441.8685	2	P62805	Histone H4	GVLKVFLENVIR	K59pr	8819609	10234317	14591866	11945549	9632478	13154548	11215264	11577525	0.97	1.03
673.8383	1345.6609	2	P62805	Histone H4	DAVTYTEHAKR	K77pr	1250180	1109373	1209936	1112630	1840743	1889780	1189830	1614384	0.74	1.36
853.9816	1705.9476	2	P62805	Histone H4	KTVTAMDVVYALKR	K79pr + K91pr	6127352	8097329	6059672	9754685	9803245	12660134	6761451	10739355	0.63	1.59
861.9783	1721.9409	2	P62805	Histone H4	KTVTAMDVVYALKR	K79pr + M84ox + K91pr	2257833	3436466	2724491	3530059	4052780	5437646	2806263	4340162	0.65	1.55

pr= propionyl group
added by chemical
reaction
ac = invivo acetylation
me1= invivo monomethyl.
me2= invivo dimethyl.
ox= oxidation

bold values= significant
differential expression

Table A.9: Histone H3 Peptide MS/MS Identifications

H3 peptide ID list

Residue start	Residue end	Experimental m/z	Theoretical m/z	Experimental mass	Theoretical mass	Charge	ppm	Sequence	PTM/Chemical Modification	Mascot score (CID Hi/Low)	Expectation (CID Hi/Low)	Mascot Score (ETD Hi/Low)	Expectation (ETD Hi/Low)
3	8	380.7196	380.7192	759.4236	759.4239	2	0.4	TKQTAR	K4pr	45	4.30E-05		
3	8	387.7283	387.7270	773.4409	773.4395	2	-1.8	TKQTAR	K4pr+me1	46	7.00E-05		
9	17	507.2910	507.2905	1012.5663	1012.5665	2	0.2	KSTGGKAPR	K9pr+K14pr	57	8.30E-06		
9	17	514.2999	514.2984	1026.5841	1026.5822	2	-1.9	KSTGGKAPR	K9pr+me1+K14pr	62	3.50E-06		
9	17	493.2925	493.2931	984.5694	984.5716	2	2.3	KSTGGKAPR	K9me2+K14pr	52	3.80E-05		
9	17	500.3002	500.3009	998.5848	998.5872	2	2.5	KSTGGKAPR	K9me3+K14pr	54	4.20E-05		
9	17	500.2832	500.2827	998.5508	998.5509	2	0.04	KSTGGKAPR	K9pr+K14ac	46	7.00E-05		
9	17	493.2749	493.2749	984.5342	984.5352	2	1.0	KSTGGKAPR	K9ac+K14ac	69	1.20E-07		
9	17	507.2901	507.2905	1012.5646	1012.5665	2	1.9	KSTGGKAPR	K9me1+K14ac	58	7.30E-06		
9	17	540.2837	540.2841	1078.5517	1078.5536	2	1.7	KSTGGKAPR	K9me3+S10phos+K14pr	38	0.00088		
18	26	549.8357	549.8351	1097.6557	1097.6557	2	-0.1	KQLATKAAR	K18pr+K23pr	82	1.20E-07		
18	26	542.8275	542.8273	1083.6394	1083.6400	2	0.6	KQLATKAAR	K18pr+K23ac	67	2.50E-06		
18	26	556.8430	556.8429	1111.6704	1111.6713	2	0.8	KQLATKAAR	K18pr+me1+K23pr	51	0.00015		
18	26	535.8195	535.8195	1069.6233	1069.6244	2	1.0	KQLATKAAR	K18ac+K23ac	72	8.00E-07		
41	49	516.8015	516.8011	1031.5874	1031.5876	2	0.2	YRPGTVALR	none	37	0.00032		
54	63	653.8736	653.8719	1305.7316	1305.7292	2	-1.8	YQKSTELLIR	K56pr	75	4.00E-07		
64	69	422.7577	422.7556	843.4998	843.4966	2	-3.7	KLPFQR	K64pr	32	0.0024		
73	83	696.3619	696.3619	1390.7082	1390.7092	2	0.7	EIAQDFKTDLR	K79pr	82	4.90E-08		
73	83	703.3710	703.3697	1404.7264	1404.7249	2	-1.1	EIAQDFKTDLR	K79pr+me1	59	1.10E-05		
73	83	682.3646	682.3644	1362.7136	1362.7143	2	0.5	EIAQDFKTDLR	K79me2	34	0.0041		
117	128	720.9172	720.9158	1439.8188	1439.8170	2	-1.3	VTIMPKDIQLAR	K122pr	70	7.70E-07		
117	128	728.9145	728.9132	1455.8133	1455.8119	2	-1.0	VTIMPKDIQLAR	M120ox + K122pr	78	9.80E-08		
H3.1 specific peptides													
27	40	801.4611	801.4597	1600.9066	1600.9049	2	-1.1	KSAPATGGVKKPHR	K27pr+K36pr+K37pr	112	8.80E-11		
27	40	808.4690	808.4675	1614.9223	1614.9205	2	-1.1	KSAPATGGVKKPHR	K27pr+me1+K36pr+K37pr	92	9.80E-09		
27	40	539.3156	539.3141	1614.9234	1614.9205	3	-1.8	KSAPATGGVKKPHR	K27pr+me1+K36pr+K37pr			58	4.00E-05
27	40	787.4642	787.4623	1572.9127	1572.9100	2	-1.7	KSAPATGGVKKPHR	K27me2+K36pr+K37pr	92	7.00E-09		
27	40	794.4714	794.4701	1586.9272	1586.9256	2	-1.0	KSAPATGGVKKPHR	K27me3+K36pr+K37pr	62	7.80E-06		
27	40	529.9815	529.9825	1586.9211	1586.9256	3	2.9	KSAPATGGVKKPHR	K27me3+K36pr+K37pr			74	7.50E-07
27	40	515.9797	515.9790	1544.9158	1544.9151	3	-0.5	KSAPATGGVKKPHR	K27me2+K36me2+K37pr			83	5.10E-08
27	40	387.2379	387.2360	1544.9204	1544.9151	4	-3.5	KSAPATGGVKKPHR	K27me2+K36me2+K37pr			63	4.40E-06
27	40	815.4775	815.4754	1628.9394	1628.9362	2	-2.0	KSAPATGGVKKPHR	K27pr+me1+K36pr+me1+K37pr	72	1.10E-06		
27	40	543.9866	543.9860	1628.9364	1628.9362	3	-0.2	KSAPATGGVKKPHR	K27pr+me1+K36pr+me1+K37pr			47	0.00059
27	40	794.4709	794.4701	1586.9261	1586.9256	2	-0.3	KSAPATGGVKKPHR	K27pr+me1+K36me2+K37pr	65	3.50E-06		
27	40	397.7394	397.7387	1586.9261	1586.9256	4	-0.3	KSAPATGGVKKPHR	K27pr+me1+K36me2+K37pr			51	9.80E-05
27	40	794.4716	794.4701	1586.9275	1586.9256	2	-1.2	KSAPATGGVKKPHR	K27me2+K36me1+K37pr	55	4.50E-05		
27	40	801.4790	801.4779	1600.9424	1600.9413	2	-0.7	KSAPATGGVKKPHR	K27me3 + K36me1 +K37pr	54	5.40E-05		
27	40	534.6534	534.6544	1600.9369	1600.9413	3	2.8	KSAPATGGVKKPHR	K27me3 + K36me1 +K37pr			69	2.00E-06
27	40	520.6379	520.6387	1558.8904	1558.8943	3	2.5	KSAPATGGVKKPHR	K27me2 + K36ac + K37pr			49	0.00015
27	40	401.2434	401.2426	1600.9421	1600.9413	4	-0.5	KSAPATGGVKKPHR	K27me1 + K36me3 +K37pr			52	9.00E-05
H3.3 specific peptides													
27	40	809.4569	809.4572	1616.8981	1616.8998	2	1.1	KSAPSTGGVKKPHR	K27pr+K36pr+K37pr	71	1.30E-06		
27	40	816.4657	816.4650	1630.9158	1630.9154	2	-0.2	KSAPSTGGVKKPHR	K27pr+me1+K36pr+K37pr	76	3.20E-07		
27	40	535.3149	535.3141	1602.9214	1602.9205	3	-0.5	KSAPSTGGVKKPHR	K27me1 + K36me2 + K37pr			33	0.0058
27	40	539.9850	539.9860	1616.9315	1616.9362	3	2.9	KSAPSTGGVKKPHR	K27me1 + K36me3 +K37pr			49	0.00014
27	40	549.3166	549.3176	1644.9263	1644.9311	3	2.9	KSAPSTGGVKKPHR	K27me1 + K36me1 + K37pr			50	0.00012
27	40	535.3129	535.3141	1602.9153	1602.9205	3	3.2	KSAPSTGGVKKPHR	K27me3 + K36pr + K37pr			94	5.20E-09
27	40	530.6413	530.6422	1588.9004	1588.9049	3	2.8	KSAPSTGGVKKPHR	K27me2 + K36pr + K37pr			58	2.50E-05

H3.1 Unique peptides: 9
H3.3 Unique peptides: 9

Sequence coverage: 63%

pr= propionyl group
added by chemical
reaction
ac = invivo acetylation
me1= invivo monomethyl.
me2= invivo dimethyl.
me3= invivo trimethyl.
phos= phosphorylation
ox= oxidation

Table A.10: Histone H4 Peptide MS/MS Identifications

H4 peptide ID list

Residue start	Residue end	Experimental m/z	Theoretical m/z	Experimental mass	Theoretical mass	Charge	ppm	Sequence	PTM/Chemical Modification	Mascot score (CID Hi/Low)	Expectation (CID Hi/Low)
4	17	747.9414	747.9412	1493.8672	1493.8678	2	0.4	GKGGKGLGKGGAKR	K5pr-K8pr-K12pr-K16pr	79	1.50E-07
4	17	494.2913	494.2913	1479.8504	1479.8521	3	1.2	GKGGKGLGKGGAKR	K5pr-K8pr-K12pr-**K16ac**	65	3.00E-06
4	17	494.2918	494.2913	1479.8518	1479.8521	3	0.2	GKGGKGLGKGGAKR	K5pr-K8pr-**K12ac**-K16pr	73	5.90E-07
4	17	733.9251	733.9255	1465.8345	1465.8365	2	1.4	GKGGKGLGKGGAKR	K5pr-**K8ac-K12ac**-K16pr	89	1.30E-08
4	17	489.6194	489.6194	1465.8348	1465.8365	3	1.1	GKGGKGLGKGGAKR	K5pr-**K8ac**-K12pr-**K16ac**	37	0.0016
4	17	733.9254	733.9255	1465.8352	1465.8365	2	0.9	GKGGKGLGKGGAKR	K5pr-K8pr-**K12ac-K16ac**	25	0.033
4	17	726.9175	726.9177	1451.8193	1451.8208	2	1.0	GKGGKGLGKGGAKR	K5pr-**K8ac-K12ac-K16ac**	92	4.90E-09
4	17	719.9112	719.9099	1437.8067	1437.8052	2	-1.1	GKGGKGLGKGGAKR	**K5ac-K8ac-K12ac-K16ac**	94	2.10E-09
20	35	650.0613	650.0616	1947.1604	1947.1629	3	1.3	KVLRDNIQGITKPAIR	K20pr+me1 +K31pr	52	0.00038
20	35	636.0575	636.0580	1905.1490	1905.1523	3	1.7	KVLRDNIQGITKPAIR	K20me2 + K31pr	56	9.90E-05
24	35	691.3959	691.3935	1380.7762	1380.7725	2	-2.7	DNIQGITKPAIR	K31pr	68	1.70E-06
24	35	698.4020	698.4013	1394.7883	1394.7881	2	-0.1	DNIQGITKPAIR	K31pr+me1	32	0.0067
46	55	590.8152	590.8140	1179.6147	1179.6135	2	-1.0	ISGLIYEETR	none	76	1.40E-07
56	67	721.9421	721.9401	1441.8685	1441.8656	2	-2.0	GVLKVFLENVIR	K59pr	87	1.50E-08
68	78	673.8383	673.8386	1345.6609	1345.6626	2	1.3	DAVTYTEHAKR	K77pr	74	1.70E-07
79	92	853.9816	853.9791	1705.9476	1705.9436	2	-2.3	KTVTAMDVVYALKR	K79pr + K91pr	108	1.40E-10
79	92	861.9783	861.9765	1721.9409	1721.9385	2	-1.4	KTVTAMDVVYALKR	K79pr + M84ox + K91pr	90	6.90E-09

H4 Unique peptides: 6

Sequence coverage: 75%

pr= propionyl group
added by chemical
reaction
ac = invivo acetylation
me1= invivo monomethyl.
me2= invivo dimethyl.
ox= oxidation

Table A.11: Histone H3K56 peptide MRM transitions

1- H3K56pr
Sequence: YQK(pr)STELLIR
Precursor ion (*m/z*): 653.87^{+2}

MRM transitions:	$653.9^{+2} \rightarrow 1015.6^{+1}$	y8 fragment
	$653.9^{+2} \rightarrow 831.5^{+1}$	y7 fragment
	$653.9^{+2} \rightarrow 744.5^{+1}$	y6 fragment
	$653.9^{+2} \rightarrow 476.3^{+1}$	b3 fragment

2- H3K56ac (light form)
Sequence: YQK(ac)STELLIR
Precursor ion (*m/z*): 646.86^{+2}

MRM transitions:	$646.9^{+2} \rightarrow 1001.6^{+1}$	y8 fragment
	$646.9^{+2} \rightarrow 831.5^{+1}$	y7 fragment
	$646.9^{+2} \rightarrow 744.5^{+1}$	y6 fragment
	$646.9^{+2} \rightarrow 462.2^{+1}$	b3 fragment

3- H3K56ac (heavy form, internal standard)
Sequence: YQK(ac)STELL*IR
Precursor ion (*m/z*): 650.36^{+2}

MRM transitions:	$650.4^{+2} \rightarrow 1008.6^{+1}$	y8 fragment
	$650.4^{+2} \rightarrow 838.5^{+1}$	y7 fragment
	$650.4^{+2} \rightarrow 751.5^{+1}$	y6 fragment
	$650.4^{+2} \rightarrow 462.2^{+1}$	b3 fragment

4- YRPGTVALR (normalization peptide)
Precursor ion (*m/z*): 516.80^{+2}

MRM transitions:	$516.8^{+2} \rightarrow 713.4^{+1}$	y7 fragment
	$516.8^{+2} \rightarrow 616.4^{+1}$	y6 fragment

Table A.12: Histone H4K91 peptide MRM transitions

1- H4K91pr

Sequence: KprTVTAMDVVYALKprR

Precursor ions (*m/z*): 853.98^{+2} and 569.66^{+3}

MRM transitions:	$854.0^{+2} \rightarrow 543.4^{+1}$	y4 fragment
	$854.0^{+2} \rightarrow 706.4^{+1}$	y5 fragment
	$854.0^{+2} \rightarrow 805.5^{+1}$	y6 fragment
	$854.0^{+2} \rightarrow 1019.6^{+1}$	y8 fragment
	$854.0^{+2} \rightarrow 359.2^{+1}$	y2 fragment

MRM transitions:	$569.7^{+3} \rightarrow 543.4^{+1}$	y4 fragment
	$569.7^{+3} \rightarrow 706.4^{+1}$	y5 fragment
	$569.7^{+3} \rightarrow 805.5^{+1}$	y6 fragment
	$569.7^{+3} \rightarrow 1019.6^{+1}$	y8 fragment
	$569.7^{+3} \rightarrow 359.2^{+1}$	y2 fragment

5- H4K91ac

Sequence: KprTVTAMDVVYALKacR

Precursor ions (*m/z*): 846.97^{+2} and 564.98^{+3}

MRM transitions:	$847.0^{+2} \rightarrow 529.4^{+1}$	y4 fragment
	$847.0^{+2} \rightarrow 692.4^{+1}$	y5 fragment
	$847.0^{+2} \rightarrow 791.5^{+1}$	y6 fragment
	$847.0^{+2} \rightarrow 1005.6^{+1}$	y8 fragment
	$847.0^{+2} \rightarrow 345.2^{+1}$	y2 fragment

MRM transitions:	$565.0^{+3} \rightarrow 529.4^{+1}$	y4 fragment
	$565.0^{+3} \rightarrow 692.4^{+}$	y5 fragment
	$565.0^{+3} \rightarrow 791.5^{+1}$	y6 fragment
	$565.0^{+3} \rightarrow 1005.6^{+1}$	y8 fragment
	$565.0^{+3} \rightarrow 345.2^{+1}$	y2 fragment

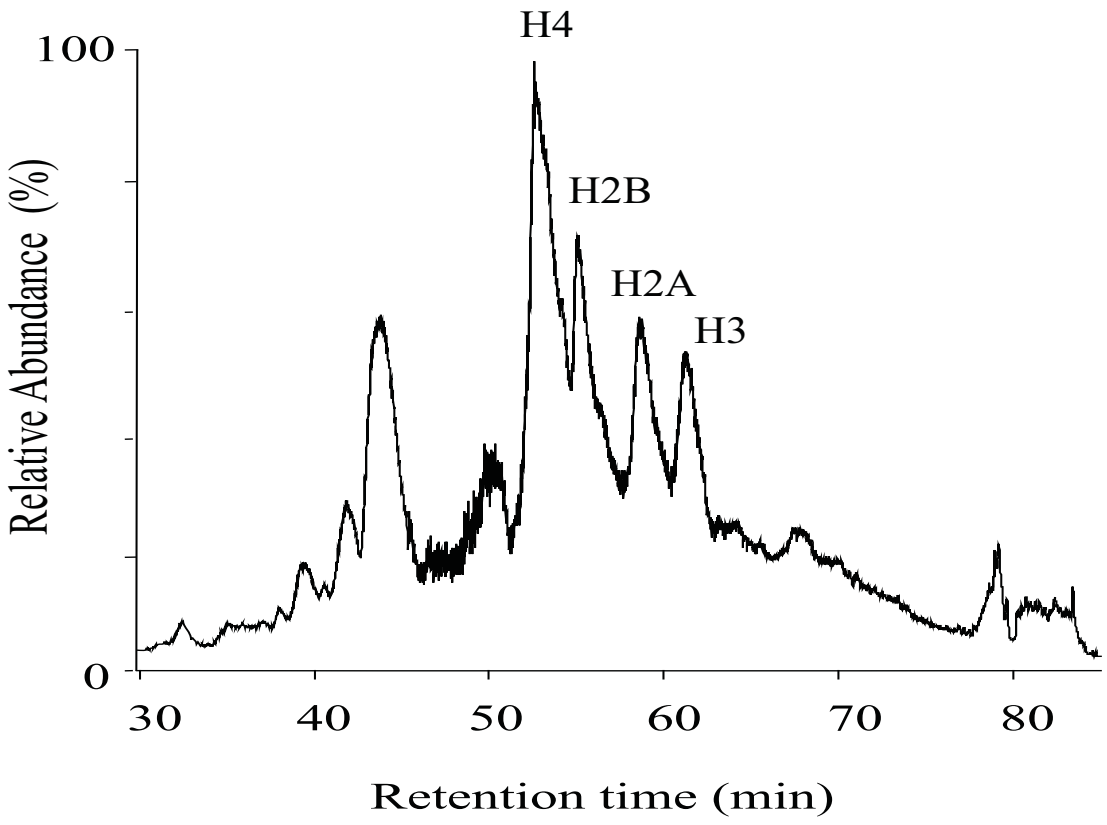


Figure A.1: Total ion chromatogram (TIC) from the LC-MS analysis of intact histones derived from untreated (control) K562 cells. Histone proteins are annotated in the TIC, and eluted between 50 and 65 minutes. Gradient conditions are specified in the main text (see section 5.3.6)

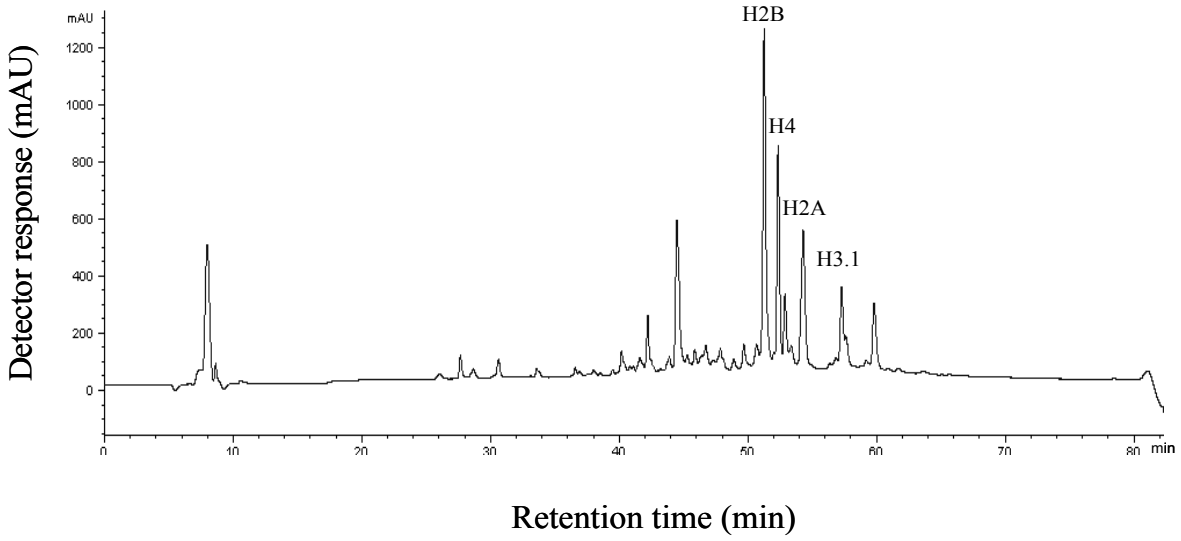


Figure A.2: Reversed-phase C_{18} fractionation of core histones derived from K562 untreated (control) cells. Histone proteins are annotated in the chromatogram. Gradient conditions are specified in the main text (see section 5.3.5)

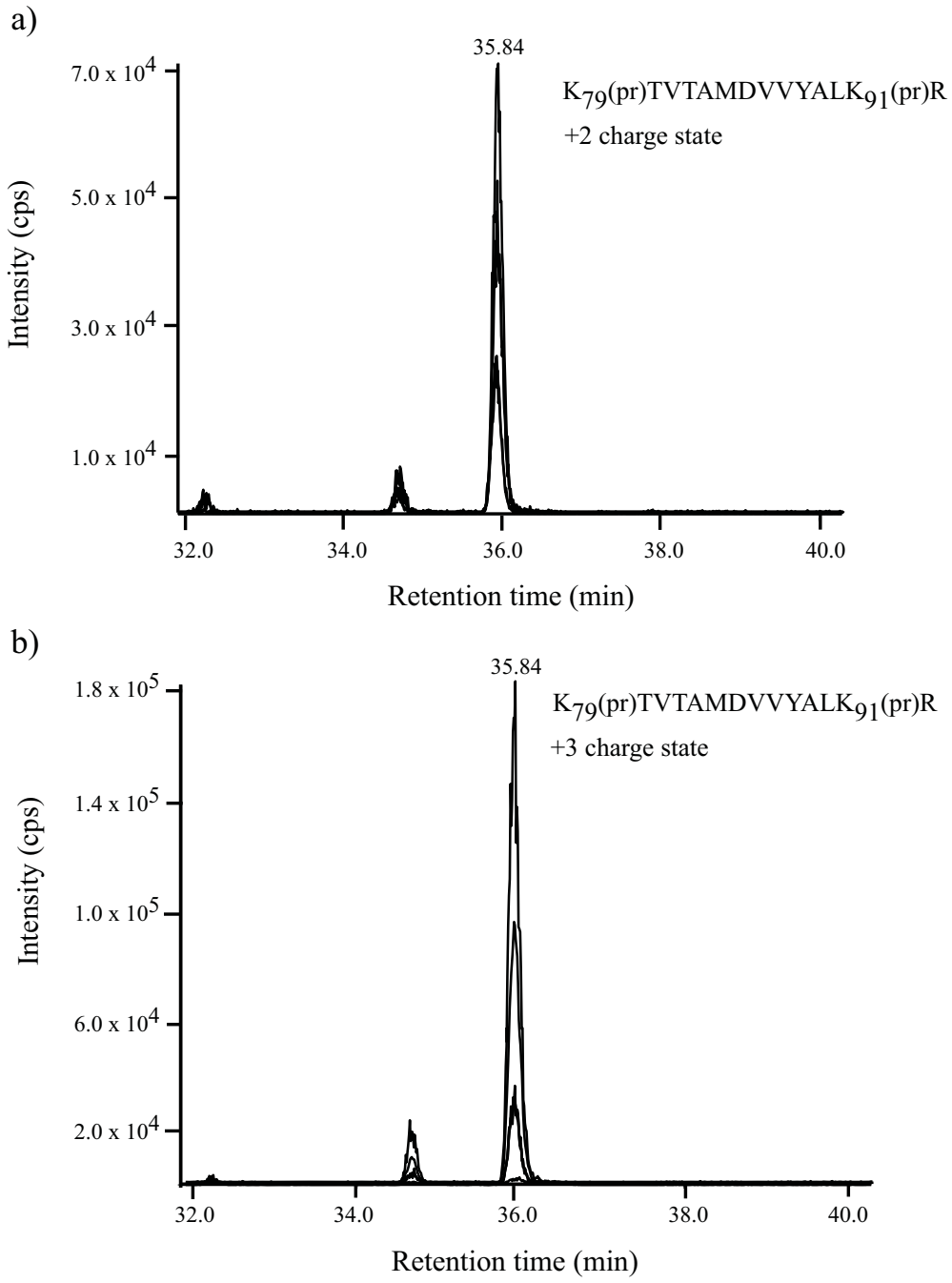


Figure A.3: Targeted MRM analyses of H4K91 peptides in histones derived from K562 cells treated with SAHA for a 24h period. MRM ion transition chromatograms for the free (propionylated) K91 residue are shown above, for both the doubly (a) and triply (b) charged states.

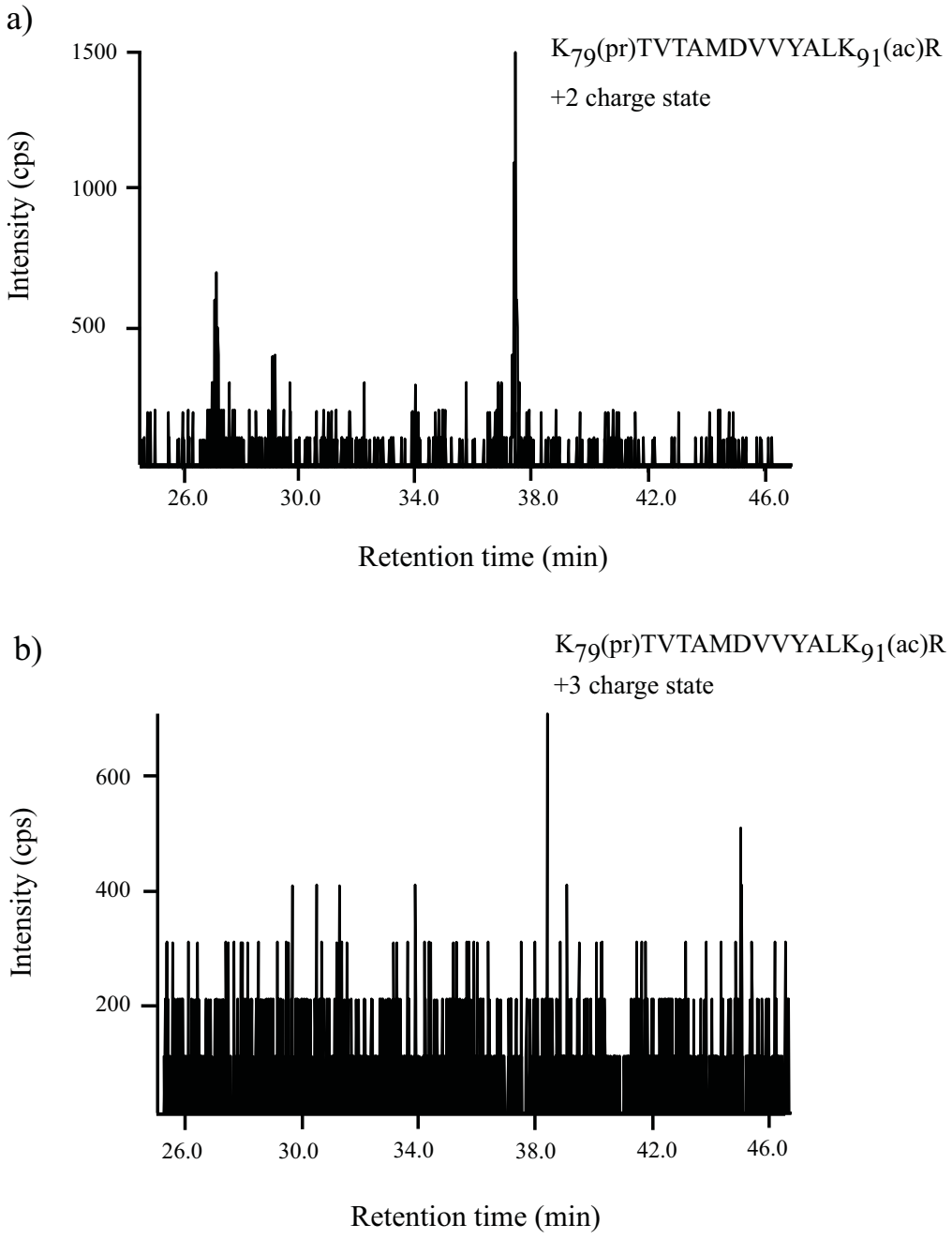


Figure A.4: Targeted MRM analyses of H4K91 peptides in histones derived from K562 cells treated with SAHA for a 24h period. MRM ion transition chromatograms for the acetylated K91 residue are shown above, for both the doubly (a) and triply (b) charged states.

Appendix 2: Scientific Contributions

Publications

Clinically Relevant Histone Deacetylase Inhibitors Enhance Histone H3/H4 Acetylation More Readily in Transformed Cells than in Normal Cells

Paul Drogaris, Valérie Villeneuve, Christelle Pomiès, Eun-Hye Lee, Véronique Bourdeau, Eric Bonneil, Gerardo Ferbeyre, Alain Verreault, and Pierre Thibault.
Journal of Proteome Research, 2010, submitted.

Regulation of the DNA Damage Response and Gene Expression by the Dot1L Histone Methyltransferase and the 53Bp1 Tumour Suppressor

Jennifer Fitzgerald, Sylvie Moureau, Paul Drogaris, Enda O'Connell, Nebiyu Abshiru, Pierre Thibault, Alain Verreault , and Noel Lowndes.
PLOS One, 2010, accepted.

A cross-talk between acetylation and tri-methylation of histone H3 lysine 4 regulates gene expression

Benoit Guillemette, Paul Drogaris, Hsiu-Hsu Sophia Lin, Axel Imhof, Pierre Thibault, Alain Verreault, and Richard J. Festenstein.
PLOS Genetics, 2010, under review.

Structure of the Rtt109-AcCoA/Vps75 complex and implications for chaperone-mediated histone acetylation

Tang, Y., Holbert, M.A., Delgoshaie, N., Wurtele, H., Meeth, K., Yuan, H., Drogaris, P., Thibault, P., Verreault, A., Cole, P.A. and Marmorstein, R.
Structure, 2011, in-press, available online at: doi:10.1016/j.str.2010.12.012

Modulation of Histone H3 Lysine 56 Acetylation as an Antifungal Therapeutic Strategy

Hugo Wurtele, Sarah Tsao, Guylaine Lépine, Alaka Mullick, Jessy Tremblay, Paul Drogaris, Eun-Hye Lee, Pierre Thibault, Alain Verreault, and Martine Raymond.
Nature Medicine, 2010 Jul;16(7): 774-80.

Identification of tandem mass spectra of mixtures of isomeric peptides

Xi Chen, Paul Drogaris, and Marshall Bern.
Journal of Proteome Research, 2010, 9, 3270-3279.

Enhanced protein detection using a trapping mode on a hybrid quadrupole-linear ion trap (Q-Trap)

Paul Drogaris, J.C. Yves Leblanc, Jennifer Fitzgerald, Noel Lowndes, Alain Verreault, and Pierre Thibault.
Analytical Chemistry, 2009, 81 (15), 6300–6309.

Comprehensive Profiling of Histone Modifications using a Label-Free Approach and its Applications in Determining Structure-Function Relationships

Paul Drogaris, Hugo Wurtele, Hiroshi Masumoto, Alain Verreault, and Pierre Thibault.
Analytical Chemistry, 2008, 80, (17), 6698-6707

Presentations

"Enhanced detection of intact proteins by nanoLC-MS using a novel trapping mode on a hybrid linear ion trap spectrometer"
Paul Drogaris, Feng Zhong, J.C. Yves Leblanc, Alain Verreault, Jennifer Fitzgerald, Noel Lowndes, and Pierre Thibault.
Presented at the 57th ASMS conference, Philadelphia, Pennsylvania, 2009.

"Automated on-line sample preparation for profiling changes in histone modifications using a Q-TOF mass spectrometer"
Paul Drogaris, Eric Bonneil, Christelle Pomiès, Kevin Killeen, and Pierre Thibault.
Presented at the 57th ASMS conference, Philadelphia, Pennsylvania, 2009.

"Application of a new Q-TOF mass spectrometer to monitor differential expression of histone modifications in response to HDACi treatment"
Paul Drogaris, Anda Vintiloiu, Christelle Pomiès, Eric Bonneil, Christine Miller, Georges Gauthier, and Pierre Thibault.
Invited speaker, 7th annual CVG conference, Montreal and Toronto, Canada, 2008.

"Profiling histone modifications in response to HDAC inhibitors using FAIMS on an Orbitrap mass spectrometer"
Julian Saba, Paul Drogaris, Christelle Pomiès, Eric Bonneil, and Pierre Thibault.
Presented at the 56th ASMS conference, Denver, Colorado, 2008.

"Application of a new Q-TOF mass spectrometer to monitor differential expression of histone modifications in response to HDACi treatment"
Paul Drogaris, Anda Vintiloiu, Christelle Pomiès, Eric Bonneil, Christine Miller, Georges Gauthier, and Pierre Thibault.
Presented at the 56th ASMS conference, Denver, Colorado, 2008.

"Chip based nano-LC/Q-TOF: application to the identification of histone modifications."
Invited speaker, Agilent Technologies Workshop, ABRF 2008, Salt Lake City, Utah.

"Comprehensive expression profiles for targeted identification of histone modifications and its implications for the determination of structure-function relationships."
Paul Drogaris, Hugo Wurtele, Kevin Eng, Alain Verreault, and Pierre Thibault.
Presented at the 55th ASMS conference, Indianapolis, Indiana, 2007.

"Analysis of Histones and Variants using a Hybrid Linear Mass Spectrometer"
Presented at the 74th ACFAS Scientific Symposium, Mcgill University, Montreal, Canada, 2006.

"LC-MS/MS Method to Rapidly Monitor Plasma and Brain Levels of a Novel Alzheimer Drug Candidate in the Mouse"
Paul Drogaris, Hong Gao, Julie Laurin, and Denis Garceau.
Presented at the 49th ASMS conference, Chicago, Illinois, 2001.

"A Sensitive and Rapid Automated LC/APCI/Tandem MS Assay for the Quantitative Determination of Buprenorphine and Norbuprenorphine in Human Plasma"
P. Patterson, R. Demers, R. DiFabio, P. Drogaris, and J. Marr.
Presented at the 47th ASMS conference, Dallas, Texas, 1999.

"A Sensitive, Rapid Automated 96-well SPE-LC/MS/MS Assay for Oxycodone, Noroxycodone, and Oxymorphone in Human Plasma"
Katrina Fegan, Roger Demers, Luca C. Matassa, and Paul Drogaris.
Presented at the annual AAPS conference, New Orleans, Louisiana, 1999.

"A Quantitative LC/MS/MS Assay for Nefazodone, Hydroxynefazodone, and *m*-CPP in Human Plasma"
Katrina Fegan, Roger Demers, Luca C. Matassa, and Paul Drogaris.
Presented at the annual AAPS conference, New Orleans, Louisiana, 1999.

"A Sensitive, Rapid Automated 96-well SPE-LC/MS/MS Assay for Promethazine in Human Plasma"
Patricia E. Patterson, Luca C. Matassa, Roger Demers, and Paul Drogaris.
Presented at the annual AAPS conference, San Francisco, California, 1998.

"Rapid Automated Protein Precipitation-LC/MS/MS Assay for Pemoline in Human Plasma"
Patricia E. Patterson, Luca C. Matassa, Roger Demers, and Paul Drogaris.
Presented at the annual AAPS conference, San Francisco, California, 1998.

"Rapid Drug Development Using Cassette Dosing and APCI LC-MS/MS Assay from Rat Plasma"
Luca Matassa, Patricia E. Patterson, Paul Drogaris, and Roger Demers.
Presented at the annual AAPS conference, San Francisco, California, 1998.

"Automated 96-well SPE/APCI LC-MS/MS Analysis of Hydromorphone in Human Plasma"
Luca Matassa, Patricia E. Patterson, Paul Drogaris, and Roger Demers.
Presented at the 46th ASMS conference, Orlando, Florida, 1998.

"APCI LC/MS/MS Analysis of Diclofenac in Human Plasma and Urine."
Luca Matassa, Paul Drogaris, and Roger Demers.
Presented at the 46th ASMS conference, Orlando, Florida, 1998.