UNIVERSITY OF CALIFORNIA

RIVERSIDE

Application of Mass Spectrometry on Quantitative Proteomics and Histone Post-Translational Modifications

Doctor of Philosophy

in

Chemistry

by

Lei Xiong

March 2011

Dissertation Committee:

Dr. Yinsheng Wang, Chairperson
Dr. Cynthia K. Larive
Dr. Wenwan Zhong

The Dissertation of Lei Xiong is approved:

Committee Chairperson

University of California, Riverside

ACKNOWLEGEMENTS

First and foremost, I would like to express my deep gratitude to my doctoral advisor Prof. Yinsheng Wang, for giving me this great opportunity to be his student, and the source of knowledge, guidance and support over the past years. His respected altitudes toward research, working hard, keeping eyes open for the forefront of science, as well as the talent of seeing the big picture, will benefit me in all my life. His guidance and encouragement of my work were inspiring, and have made the time I have spent at UC Riverside enjoyable. He guides me to grow up and be strong. Without his help, I would not have been here.

I would like to express my sincere thanks to the members of my dissertation committee, Prof. Cynthia Larive and Prof. Wenwan Zhong for their help, advice and evaluation of this work. Their constructive suggestions and generous help are truly appreciated. I would also like to thank Prof. Quan "Jason" Cheng, Prof. Ryan Julian and Prof. Michael Pirrung for all the helpful discussion and patient teaching since the first year in my graduate school.

I would like to thank the staff members in Chemistry department, Center of Plant Cell Biology and Core Instrumentation Facility, Dr. Songqing Pan for his training and help with QSTAR MALDI mass spectrometer, Dr. David Carter for his training in confocal spectroscopy, Mr. Ron New for his help with Voyager-DE STR MALD-TOF.

I am grateful to meet and work together with such many brilliant cooperators all over the country. Special thanks go to Prof. Kangling Zhang from Loma Linda University for his generous guidance and great suggestions, Dr. Xinning Jiang From University of California San Diego for setting up protein quantification system for the proteomic projects; Prof. Eric U. Selker and Dr. Keyur K. Adhvaryu from University of Oregon for providing *Neurospora* samples and discussions, Dr. Lihua Jiang from Thermo Fisher Scientific for offering helps with LTQ-Orbitrap ETD analysis in the histone modification project.

I am so happy to have had many excellent colleagues and friends here. I would like to thank Dr. Haibo Qiu for his help for the complete process of quantitative proteomics, Dr. Bifeng Yuan for his help for molecular biology techniques, Dr. Lijie Men, Dr. Qingchun Zhang, Dr. Huachuan Cao, Dr. Yong Jiang and Dr. Hongxia Wang for their help with cell culture and LC-MS/MS when I started my research here, Dr. Xiaoli Dong for her help on protein quantification, Dr. Jing Zhang, Dr. Liyan Ping, Dr. Congfang Lai, Yongsheng Xiao and Fan Zhang for all the valuable discussions and friendship, and all other group members: Dr. Jin Wang, Dr. John Prins, Dr. Changjun You, Jianshuang Wang, Mario Vargas, Renee Williams, Nisana Andersen, Candace R. Guerrero, Ashley Swanson, Qian Cai, Lijuan Fu, Lei Guo and Shuo Liu. A very special thank to my dear friends, near or far, Jilong Kuang, Xiaoyue Wang, Qingyu Sun, Jun Wang and Shijing Cheng, for all the precious moments we shared.

Finally, thanks, my dear dad, mom and Meng. I love you, forever.

COPYRIGHT ACKNOWLEDGEMENTS

The text and figures in Chapter 2, in part or full, are a reprint of the material as it appears in *J. Proteome Res.*, **2010**, 9, 1129–1137. The coauther (Dr. Yinsheng Wang) listed in that publication directed and supervised the research that froms the basis of this chapter.

The text and figures in Chapter 3, in part or full, are a reprint of the material as it appears in *J. Proteome Res.*, **2010**, 9, 6007–1015. The coauther (Dr. Yinsheng Wang) listed in that publication directed and supervised the research that froms the basis of this chapter.

The text and figures in Chapter 4, in part or full, are a reprint of the material as it appears in *Biochemistry,* **2010**, 49, 5236-5343. The coauther (Dr. Yinsheng Wang) listed in that publication directed and supervised the research that froms the basis of this chapter.

The text and figures in Chapter 5, in part or full, are a reprint of the material as it appears in *Int. J. Mass Spectrom.,* **2010**, in press. The coauther (Dr. Yinsheng Wang) listed in that publication directed and supervised the research that froms the basis of this chapter.

The text and figures in Chapter 6, in part or full, are a reprint of the material as it appears in *J. Am. Soc. Mass Spectrom.,* **2009**, 20, 1172-1181. The coauther (Dr. Yinsheng Wang) listed in that publication directed and supervised the research that froms the basis of this chapter.

ABSTRACT OF THE DISSERTATION

Application of Mass Spectrometry on Quantitative Proteomics and Histone Post-Translational Modifications

by

Lei Xiong

Doctor of Philosophy, Graduate Program in Chemistry
University of California, Riverside, March 2011
Dr. Yinsheng Wang, Chairperson

Mass spectrometry (MS), coupled with highly developed sample preparation and separation techniques, has served as one of the most powerful analytical technique. In this dissertation, we focus on the application of MS in two important areas, quantitative proteomics and protein post-translation modification (PTM) analysis.

In Chapters 2 and 3, we report the use of mass spectrometry together with stable isotope labeling by amino acids in cell culture (SILAC) for the comparative study of protein expression in HL-60 and K562 cells that were untreated or treated with a clinically relevant concentration of arsenite or imatinib. Our results revealed that, among more than 1000 quantified proteins, 56 and 73 proteins including many important enzymes had significantly altered levels of expression by arsenite and imatinib treatment, respectively. With the down-stream pathway analysis, our results provided potential biomarkers for monitoring the therapeutic intervention of leukemia

and offered important new knowledge for gaining insight into the molecular mechanisms of action of anti-cancer drugs.

In Chapters 4 and 5, we obtained relatively comprehensive mappings of core histone post-translational modifications in two important model eukaryotic organisms, *Neurospora crassa* and *Schizosaccharomyces pombe*. We used several mass spectrometric techniques, coupled with HPLC separation and multiple protease digestion, to identify the methylation, acetylation and phosphoration sites in core histones. Our analysis provides potentially comprehensive pictures of core histones PTMs, which serve as foundations for future studies on the function of histone PTMs in these organisms.

In Chapter 6, we observed an unusual discrepancy between MALDI-MS/MS and ESI-MS/MS on the methylation of trimethyllysine-containing peptides. It turned out that the discrepancy could be attributed to an unusual methyl group migration from the side chain of trimethyllysine to the C-terminal arginine residue during peptide fragmentation. The results highlighted that caution should be exerted while MS/MS of singly charged ions is employed to interrogate the PTMs of trimethyllysine-containing peptides.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

differences with the lysine being tri-methylated or acetylated (Calcd. ΔMass, ac/me3). The two mass deviations (M.D.) between the Measured ΔMass and Calcd. ΔMass were further calculated, with one being markedly smaller than the other. The modification type at the target lysine could be determined as the one with smaller deviation.

LIST OF SUPPORTING INFORMATION

CHAPTER 1

General Overview

1.1 Introduction: Application of mass spectrometry - from single protein characterization to proteomics

The development of techniques for protein identification/quantification and structure determination has become increasingly important in biological research. Being able to define a protein's amino acid sequence can provide a link between the protein and its coding genes, thereby establishing a relationship between cellular physiology and genetics [1]. On the other hand, a protein's post-translational modifications (PTMs) ultimately determine and modulate its activity, localization, turnover rate, and interactions with other proteins [2]. Therefore, the ability to identify not only a protein and its amino acid sequence, but also its post-translational modifications can be extremely beneficial to biological research.

Mass spectrometry (MS), coupled with mordern sample preparation and separation techniques, has become one of the most powerful techniques for protein characterization. MS measures mass-to-charge ratio (*m/z*), yielding the molecular weight and the fragmentation pattern of peptides derived from proteins. Therefore, this technique can be used to determine the complete covalent structure of individual proteins by determining the protein sequence and all modifications that change the molecular weight.

The developing trend towards mass spectrometry as the technique of choice for identifying and probing covalent structure of proteins was accelerated by the advancement in genomics. Genomics utilized high-throughput techniques for the analysis of biological systems, which provided complete sequencing of the genome. Sequence databases derived from genomic techniques are crucial for the rapid and accurate identification of peptides by mass spectrometry. Proteomics is the term used to describe the protein complement of genomics and involves the systematic analysis of all proteins in a cell or tissue [1].

Although proteomics has recently enjoyed tremendous success, it remains a rapidly developing and open-ended endeavor. The high complexity of protein compositions in the real biological samples and the extremely low abundance of many important proteins remains a significant technical challenge in proteomic techniques. For example, it has been estimated that more than 10^6 different proteins spanning a concentration range greater than 10 orders of magnitude are present in human serum [1]. In addition to the technical challenges arising from sample complexity, an ideal proteomic technique should provide not only the identification and quantification of proteins, but also the determination of a protein's localization, interactions, activity level and function [4]. These challenges have led to the development of novel mass spectrometers with different ionization methods and ion fragmentation strategies. With these recent developments in mass spectrometry, in combination with new sample preparation and quantification methods, as well as various sensitive

biotechnologies for downstream event analysis, proteomics has evolved as a powerful tool for solving cellular biological and real pharmaceutical problems systemically. In this chapter, I will review MS-based strategies to study quantitative proteomics and protein PTMs, as well as their applications in understanding the mechanisms of action of anti-cancer agents and core histone analysis.

1.2 MS-based quantitative proteomics

Proteomics is aimed toward providing a detailed description of the structure, function and control of biological systems in health and disease [2]. Classic proteomic techniques that rely on quantification methods utilizing dyes, fluorophores, or radioactivity have provided very good sensitivity, linearity and dynamic range. However, they suffer from two significant shortcomings: First, they require high-resolution protein separation typically provided by two- dimensional gel electrophoresis, which is labor-intensive and limits its applicability to abundant and soluble proteins; second, they do not reveal the identity of the underlying protein [3, 4]. Both of these drawbacks can be overcome by modern LC-MS/MS techniques.

In order to achieve accurate quantification, various proteomic workflows have been designed. One major approach is based on stable isotope labeling techniques [5, 6]. Since a mass spectrometer has the ability to recognize the mass difference between the isotopically labeled and unlabeled forms of a peptide, quantification can be achieved by comparing their respective signal intensities. Recently, label-free

strategies have emerged for quantitative proteomics, by comparing the direct mass spectrometric signal intensity for given peptides, or using the number of acquired spectra as the protein amount indicator [7]. These methods all have their particular strengths and weaknesses, but they can be applied to the study of biological systems on a proteomic scale. In this section, I will summarize several widely used isotope labeling approaches for MS-based quantitative proteomics and their application in anti-cancer research.

1.2.1 Isotope labeling strategies

Several stable isotope labeling strategies have been developed for the MS-based comparative analysis of differential protein expression between normal and cancer cells (or between drug-treated and untreated cancer cells). Isotope labeling approaches are based on the use of stable isotope-coded chemical reagents or stable isotope-labeled nutrients in cell culture to differentially label proteins or peptides in order to compare their relative abundances in different samples [8, 9]. Once proteins or peptides are differentially labeled with stable isotopes, they can be differentiated from their natural counterpart using MS or MS/MS based on their characteristic mass shifts. The relative abundance of labeled peptide pairs can be determined by comparing the signal intensities of isotope-labeled and unlabeled peptides in the mass spectra.

1.2.1.1 Stable isotope labeling by amino acids in cell culture (SILAC)

Stable isotope labeling by amino acids in cell culture (SILAC) is a metabolic labeling method. Mammalian cells cannot synthesize a number of amino acids; therefore, these essential amino acids must be supplied in cell culture medium to support cell growth. If an isotope-labeled analog of an amino acid is supplied instead of its naturally occurring counterpart, it will be incorporated into newly synthesized proteins [10].

Figure 1.1 shows an outline for a standard SILAC experiment. Cells are cultured in separate culture flasks containing media with heavy or light isotope-labeled amino acids (typically lysine or arginine). As cells proliferate, they begin to incorporate the isotope-labeled analogs of lysine or arginine. After a certain number of cell doublings, the naturally occurring amino acid will be completely replaced by its isotope-labeled analog. Both the heavy and light cells are then harvested, and the extracts mixed directly at certain ratios, usually in equal amounts. Protein mixtures can be further fractionated by HPLC or SDS-PAGE, and digested to peptides. The peptide mixture is then subjected to LC-MS/MS analysis and the quantification is based on the MS signal intensity of the peptide pair.

SILAC is a very simple, efficient and reproducible technique, which can facilitate almost complete heavy isotope incorporation. In addition, SILAC introduces heavy isotopes during cell culture, and light and heavy samples are mixed prior to various steps of sample manipulation, thereby minimizing differential sample loss during protein extraction, fractionation and digestion. Therefore, SILAC is very

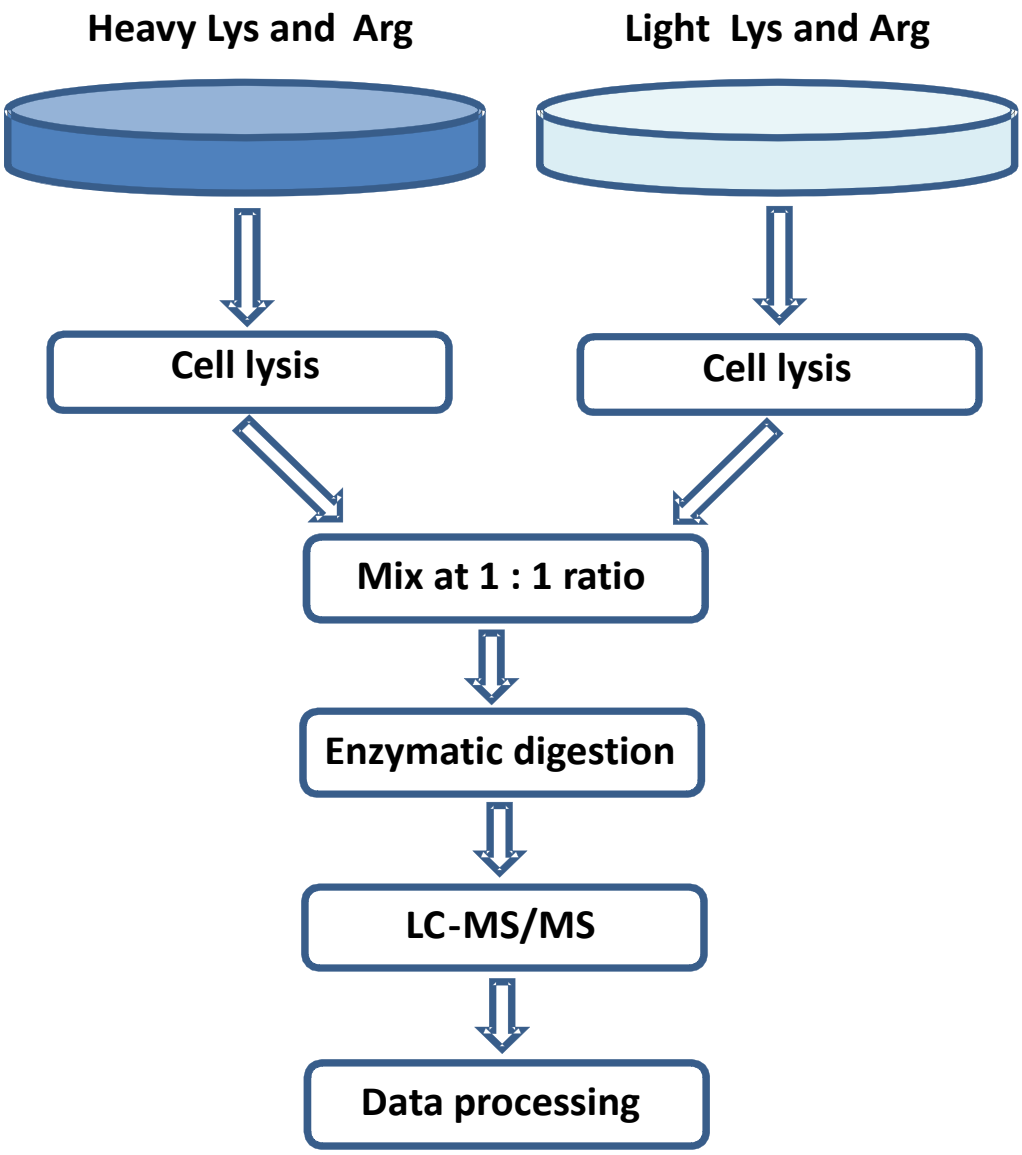


Figure 1.1 General procedures of SILAC experiment for quantitative proteomics.

suitable for the comparative study of protein expression in cells with or without drug treatments since accurate experimental results can be obtained with minimal bias, allowing for relative quantitation of small changes in protein abundance [5, 10]. In addition to global proteome quantification, SILAC can also be used for functional and subcellular proteome quantification [5, 11]. For example, SILAC is suitable for quantitative proteomic studies, since it can be combined with immobilized metal affinity chromatography (IMAC) and/or immunoprecipitation to enrich phosphopeptides [15].

1.2.1.2 Isotope-coded affinity tag (ICAT)

In addition to metabolic labeling, chemical tagging strategies serve as efficient approaches for quantitative proteomic analysis, in which isotope-coded affinity tag (ICAT) is among the most successful strategies. The ICAT reagent, developed by Aebersold and co-workers [12], consists of three elements: an affinity tag (biotin); a light or heavy isotope labeled poly(ethylene glycol) linker containing either hydrogens or deuterons; and a thiol reactive group (Figure 1.2A). The outline for a typical ICAT experiment is shown in Figure 1.2B. Two protein samples are separately labeled with either light or heavy tags, through the reaction between the free thiol groups of cysteine and the reactive group on the tag. The samples are combined and digested into peptide fragments. The tagged cysteine-containing peptides are then isolated by avidin affinity chromatography and eluted for LC-MS/MS analysis.

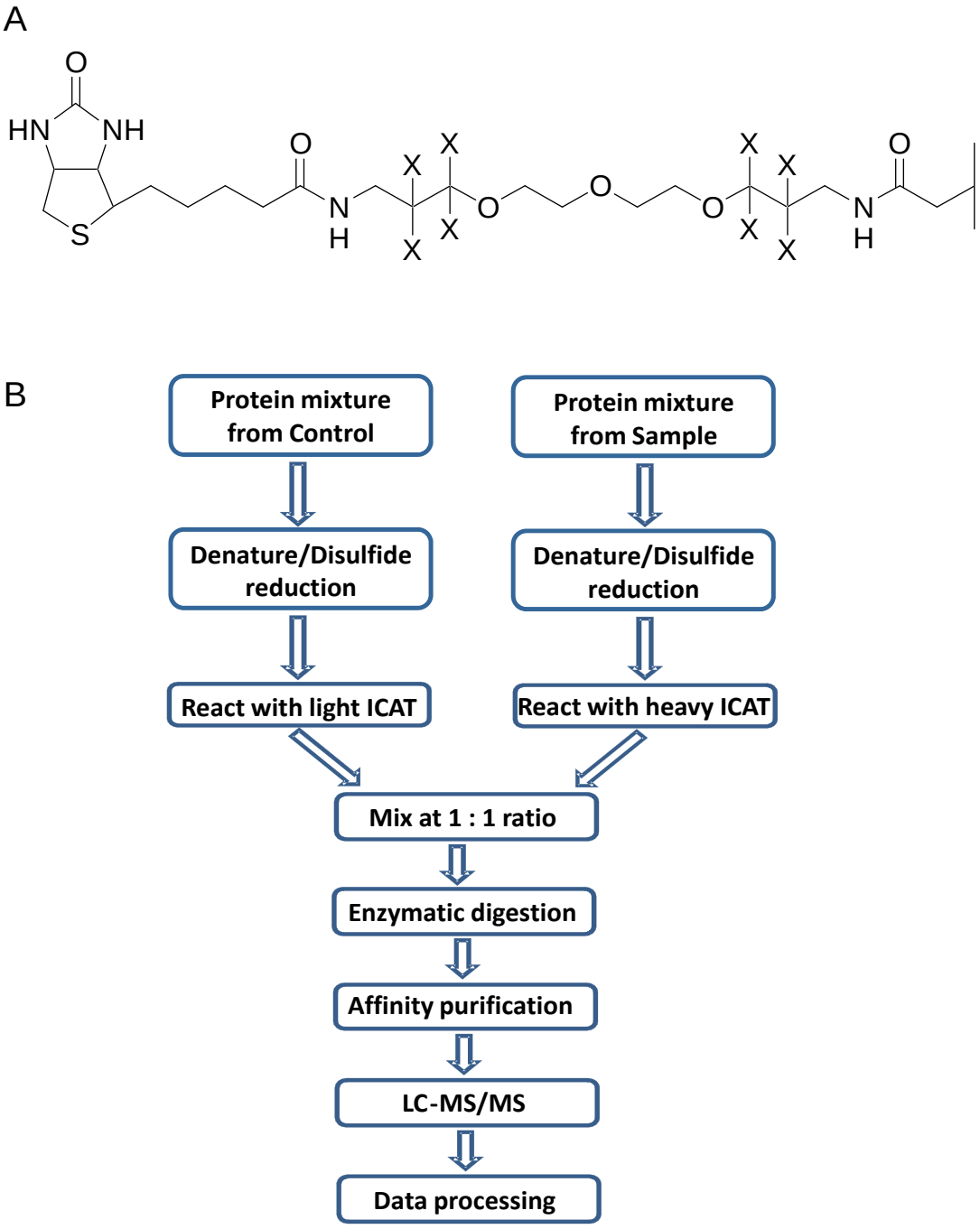


Figure 1.2 (A) The structure of ICAT reagent. (B) General procedures of ICAT experiment for quantitative proteomics.

The affinity purification of labeled peptides significantly reduces the sample complexity due to the selectivity of avidin toward biotin-attached, cysteine-containing peptides. The quantification is based on the MS signal intensities of the peptide pairs.

Since its introduction, ICAT has been widely used for MS-based quantitative analysis[13]. However, Regnier and co-workers [14] observed that the hydrogen- and deuterium-labeled peptides could be separated during HPLC, which severely affected the quantification accuracy. Moreover, the MS fragmentation of the labeled peptide is also hampered by the large moiety of the biotin tag. However, these problems have been overcome by the application of a ^{13}C-coded reagent in lieu of the deuterium-labeled tag and the introduction of an acid-cleavable site between biotin and the thiol-reactive group (cICAT) which allows for the removal of biotin prior to MS analysis [15, 16].

1.2.1.3 Isobaric tag for relative and absolute quantitation (iTRAQ)

The isobaric tag for relative and absolute quantitation (iTRAQ) strategy was developed by Pappin and colleagues [17] at Applied Biosystems. Figure 1.3.A shows the structure of the multiplexed isobaric tag. The complete molecule consists of a reporter group (based on *N*-methylpiperazine), a mass balance group (carbonyl) and a peptide-reactive group (NHS ester). When treated with a peptide, the tag forms an amide linkage to any peptide amine group located on N-terminal or lysine side chain. These amide linkages fragment in a similar fashion as peptide backbone by CID. The

balance group is lost after fragmentation, while the charge is retained by the reporter

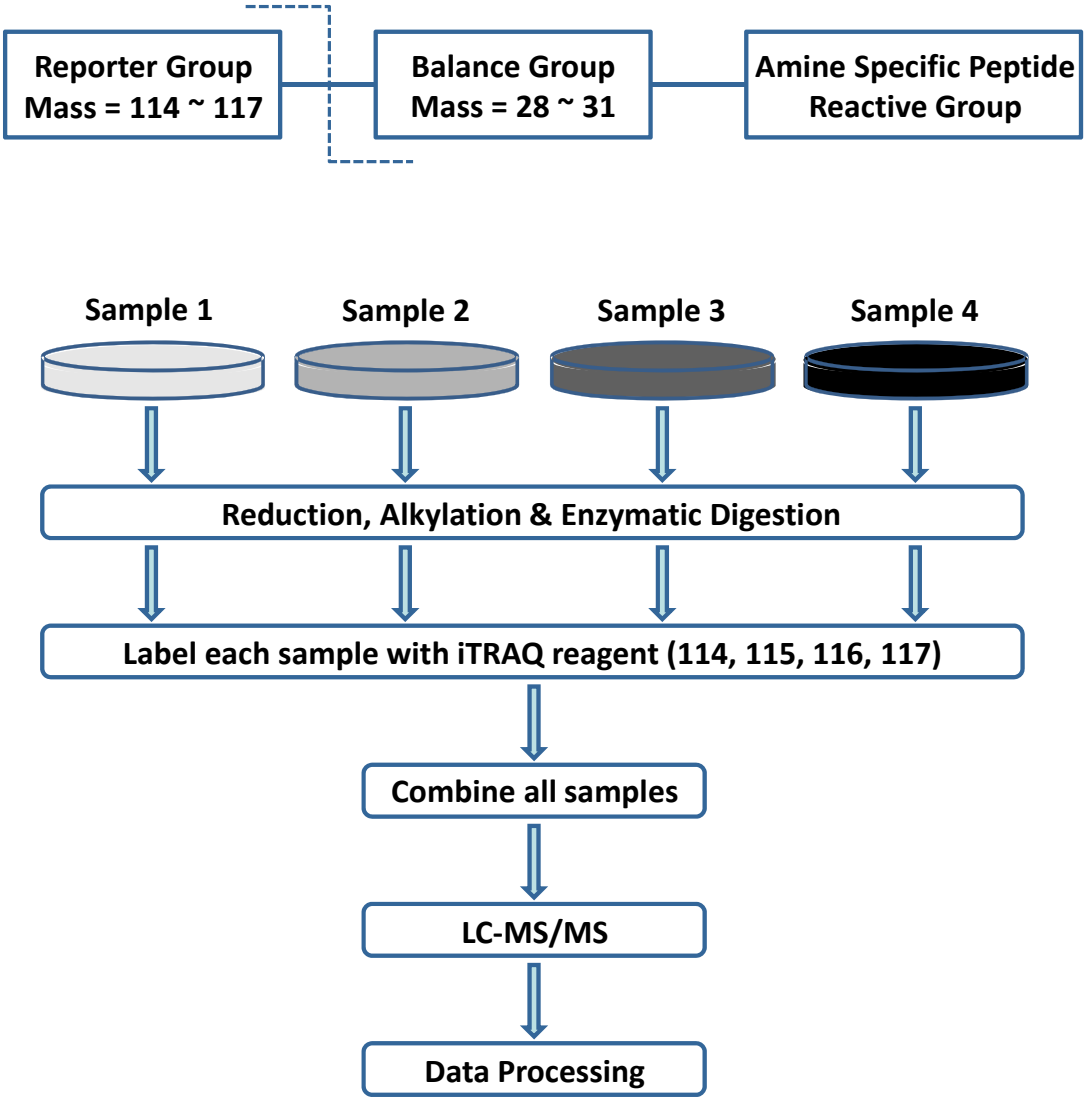


Figure 1.3 (A) The structure of iTRAQ reagent. (B) General procedures of iTRAQ experiment for quantitative proteomics.

group fragment. The reporter group includes four isotope-coded variants ranges in mass from *m/z* 114.1 to 117.1, which allows for simultaneous quantification of relative protein expression of up to four experimental conditions. Unlike ICAT, iTRAQ quantification is based on the comparison of signal intensity of the isotopic reporter group fragments in MS/MS.

There are several advantages of using iTRAQ. For instance, the isobaric tags allow for the quantification of multiple samples simultaneously without increasing MS complexity at the peptide level; therefore, more MS/MS spectra can be acquired for protein identification and quantification. Recently, a new version of iTRAQ 8-plex labeling regents was introduced by Applied Biosystems which can be employed for simultaneous quantification of eight samples. Thermo Fisher Scientific has also released a similar product, TMT isobaric mass tagging kit [18], which can provide simultaneous quantification of up to 6 different samples. In addition to relative quantification, iTRAQ can be used for absolute quantification by using tagged standard peptides in the multiplex mixtures [17].

1.2.1.4 Stable isotope dimethyl labeling

In 2009, Boersema and coworkers [19] developed a novel MS-based stable isotopic dimethyl labeling approach, which incorporates the stable isotope directly at the peptide level for the first time. For this technique, protein samples are first

digested with proteases and the derived peptides of different samples are then labeled with isotopomeric dimethyl labels (Fig 1.4). By using combinations of several isotopomers of formaldehyde and cyanoborohydride, peptide masses can be altered by a minimum of 4 Da. All primary amines in a peptide mixture are converted to dimethylamines except the rare occurrence of an N-terminal proline. The labeled samples are mixed and simultaneously analyzed by LC-MS whereby the mass difference of the dimethyl labels is used to compare the peptide abundance in different samples. Stable isotope dimethyl labeling is a reliable, cost-effective and undemanding procedure and can be easily automated and applied in high-throughput proteomics experiments.

1.2.1.5 Other MS-based quantitative strategies

In addition to the strategies described above, there are many other approaches that have been designed for quantitative proteomic studies, including ^{18}O-labeling [20], ^{15}N-labeling [21], mass-coded abundance tagging (MCAT) [22] and label-free quantitative strategies [7].

Both ^{18}O- and ^{15}N-labelings are metabolic labeling strategies. In quantitative proteomics, ^{18}O-labeling is processed during the protease digestion step, in which digestion buffer is prepared by either ^{18}O labeled H_2O or ^{16}O labeled (normal) H_2O. Two ^{18}O or ^{16}O atoms are incorporated universally into the carboxyl termini of all peptides during the proteolytic cleavage of proteins in the complex mixture, to induce

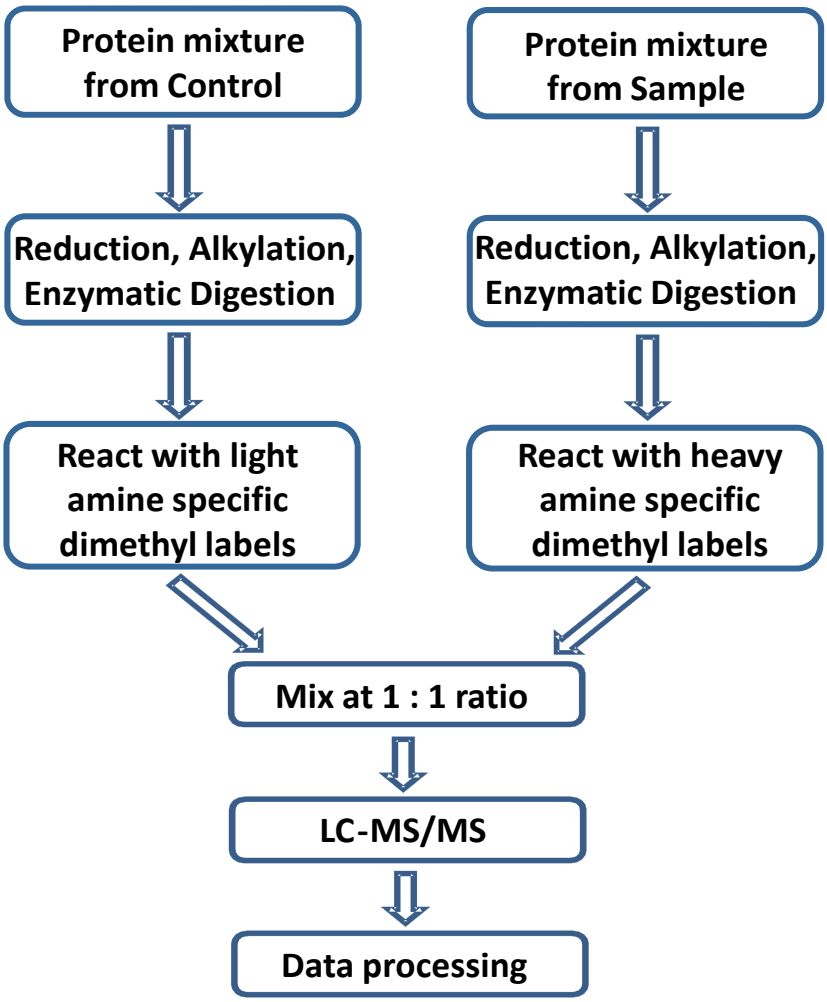


Figure 1.4 General procedures of stable isotope dimethyl labeling for quantitative proteomics.

4 Da mass difference between the light and heavy peptides. While in the case of ^{15}N-labeling, ^{15}N labeled NH_4Cl is used in cell culture media preparation. After several circles of cell doubling, the heavy isotope is distributed over the whole protein, resulting in an undefined mass difference between the labeled and unlabeled proteins.

MCAT is another widely used chemical labeling strategy. It relies on the selective and quantitative guanidination of the ε-amino group of C-terminal lysine residues of tryptic peptides at high pH with *O*-methylisourea. For protein quantification, two portions of protein samples are firstly digested with trypsin. One sample is then treated with MCAT reagent which adds 42 Da to the mass of the peptides, while the other one remains untreated. The two samples are combined at 1:1 ratio and subjected to LC-MS/MS analysis. MCAT offers marked advantages: It is simple, economic, and effective for complex sample analysis.

The label-free quantification strategy can be based on either the direct comparison of the MS signal intensity of two parallel protein digestion samples, or the setup of MRM transitions for specific parent and fragment ions. Compared with isotope labeling, the label-free strategy is simpler to handle and cost-effective, but suffers from poor quantification accuracy.

1.2.2 Applications of proteomics in cancer research

The genome sequencing is only the beginning of the quest to understand the functionality of cells, tissues, and organs, both in health and disease. Together with

advances in bioinformatics, the emerging technologies with proteomics and functional genomics have moved our targets from single molecules to complex biological systems, and elicited major advances in human disease and medical research.

Cancer, a complex disease that affects a significant fraction of the population, is foreseen as a prime target for the new technologies [23]. Although tremendous advances have been made in our understanding of the molecular basis of cancer, substantial gaps remain in our understanding of disease pathogenesis and in developing effective strategies for early diagnosis and treatment. Proteomic technologies, which could overcome some of the limitations of other approaches, are expected to play a key role in the study of cancer treatment [24]. Particularly promising areas of research include delineation of altered protein expression, the development of novel biomarkers for diagnosis and early detection of disease, the identification of new targets for therapeutics, and the potential for accelerating drug development through more effective strategies to evaluate therapeutic effects and toxicity [25]. In this section, I will summarize the major contributions of quantitative proteomics toward cancer research.

1.2.2.1 Proteomic analysis for bio-marker detection

Ultimately, the early detection of cancer is crucial for its control and prevention [26]. Although advances in conventional diagnostic strategies, such as mammography and prostate-specific antigen (PSA) testing, have provided some

improvements in disease detection, they still cannot provide the sensitivity and specificity that are required for early-stage disease detection.

For decades, biomarkers have become important tools for cancer detection and monitoring. Traditionally, proteins are selected for investigation as biomarkers for particular cancers because of their track record as biomarkers and their function or protein family relationships [27]. Thousands of publications have explored the potential use of individual proteins or collections of proteins as cancer biomarkers and have produced promising results. Traditional methods are based on antibody-related techniques such as Western blot analysis or 2-DE for individual protein validation. Quantitative proteomics can enhance dramatically the efficiency of biomarker discovery through the quantification of the entire proteome in cells, tissues and plasma and more importantly, provide better diagnostic performance through the combination of multiple biomarkers [28, 29]. Quantitative proteomic analysis combines various isotope-labeling strategies and mature mass spectrometric techniques with high sensitivity and mass accuracy. It allows for the identification and quantification of the majority of proteins in the proteome of normal and cancer cells, with the exception of proteins with extremely low abundance. Comparing protein expression levels between healthy and cancer cells can provide valuable insights into the cellular mechanisms of cancer and allow for the determination of biomarkers for specific types of cancer. In addition, MS-based proteomic studies can be used for the detection and quantification of cancer-specific PTMs in proteins, such as altered

protein glycosylation, which can be used as potential biomarkers for cancer detection [33].

1.2.2.2 Proteomic analysis for anti-cancer drug development

The development of anti-cancer drugs is a time-consuming and labor-intensive process that evolves through several sequential phases including target identification, validation, drug design, lead identification, lead characterization, clinical candidate selection, pre-clinical testing, and clinical testing [30]. The validation of drug efficiency and understanding the drug's mechanism of action, represents a very important part of the drug development process. With the advent of proteomic approaches, it is now feasible to select the most promising drug targets and outline entire molecular pathways for efficient medicinal intervention. The ability to quantify the entire proteome of cancer cells that are either untreated or treated with specific anti-cancer reagents, allows for identification of protein targets whose expression are significantly perturbed by drugs. With metabolic pathway analysis and application of related biological techniques, relationships between the perturbed proteins and the drug-perturbed biological pathways can be determined. The obtained information affords important insights into the cellular response towards drug treatment and allows for the design of novel therapeutic strategies.

It is worth noting that proteomic approaches can be efficiently used to study drug resistance in cancer. Drugs are usually given systematically and tumors can be

located in parts of the body into which drugs do not easily penetrate and can be protected by local environments due to increased tissue hydrostatic pressure or altered tumor vasculature [31]. In addition, drug resistance has become one of the main concerns for successful anti-cancer treatments. Research on drug resistance in cancer has focused on cellular resistance due to either the specific nature and genetic background of the cancer cell itself or the genetic changes that follow toxic chemotherapy. Proteomic analysis can provide evidence of protein expression level changes in the drug-resistant cancer cells and has been applied to study the drug resistance problems associated with many widely used anti-cancer reagents including Imatinib [32, 33], cisplatin [34], antiquated pentavalent antimonials [Sb(V)] [35], etc.

1.3 Core histone proteins

The genomic DNA of eukaryotic cells is organized in subunits called nucleosomes, the basic repeating element of chromatin [36]. Each nucleosome consists of 146 bp of DNA wrapped around an octameric core histone complex comprising of two H2A-H2B dimers flanking a $(H3\text{-}H4)_2$ tetramer [37]. Additionally, linker histone H1 binds to DNA in between nucleosomes thereby adding further structure to the chromatin polymers. All core histones have a basic N-terminal domain extending out from the core particle, a globular domain organized by the histone fold, and a C-terminal tail. The histone N-terminal tails are involved in the establishment of chromatin structural states, whereas their histone fold domains mediate

histone-histone and histone-DNA interactions.

Core histones have recently attracted attention for their potential dynamic roles in the regulation of gene expression and other DNA template-related events through chromatin remodeling [38]. A central mechanism for chromatin remodeling is the covalent modifications of histones, mostly located on the N-terminal portions of histones that extend beyond the surface of the nucleosomes. In recent years, mass spectrometry has become increasingly used in the identification and quantification of histone modifications, and the discovery of novel histone modifications. In this section, I summarize the common PTMs of core histones and the application of mass spectrometry on protein PTM analysis.

1.3.1 Post-translational modifications (PTMs) of core histones

To date, the most studied histone PTMs include lysine acetylation, methylation, sumoylation, ubiquitination, arginine methylation, and serine and threonine phosphorylation. These modifications can affect the interactions of nucleosomes with transacting factors, and are thought to play important roles in the assembly and disassembly of chromatin states, ultimately controlling the accessibility of DNA for important cellular processes including transcription, replication, gene silencing and DNA repair. The combinatorial modification profiles of histones suggest that the modification sites can also act as binding platforms for specific proteins that 'read' these particular marks leading to active or silenced genomic regions, as articulated by

the ‘histone code’ hypothesis [39].

1.3.1.1 Histone Acetylation

Among the common histone PTMs, histone acetylation has been most studied and appreciated [40]. Compelling evidence supports that acetylation of specific lysine residues in the amino termini of the core histones plays a fundamental role in transcriptional regulation, nucleosome assembly, etc. The discovery of enzymes responsible for bringing about the steady-state balance of this modification also serves as a very important and popular research direction [41]. The sites of acetylation include at least four highly conserved lysines in histone H4 (K5, K8, K12 and K16) [40], five in histone H3 (K9, K14, K18, K23 and K27) [42], as well as less conserved sites in histone H2A and H2B [43-45].

Histone acetylation is known to affect transcription in vivo [40]. Acetylation neutralizes the positively charged lysine residues of the histone N termini, thereby decreasing their affinity for DNA. It might allow the termini to be displaced from the nucleosome, causing the nucleosomes to unfold and increasing access to transcription factors [37, 46, 47]. Transcription-linked acetylation shows a preference for H3 K14 in vitro [48], although an expanded set of lysine residues is likely to be used in vivo [49, 50]. It has been reported that the H3 lysine residues outside the preferred K14 are important for histone-binding specificity [41] and H4 lysine acetylation can trigger orderly nucleosomal disruption near TATA elements conducive to efficient transcription [51].

Histone acetylation is also associated with biological processes apart from transcription. During DNA replication, histones H3 and H4 are brought to replicating chromatin in a pre-acetylated state that becomes erased after replication is completed and the newly assembled chromatin matures [52]. The sites of deposition-related H4 acetylation, K5 and K12, are highly conserved [53, 54], and K9 in H3 appears to have a more dominant role in histone deposition and chromatin assembly in some organisms [48, 52, 54].

1.3.1.2 Histone Methylation

The major methylation sites within histone tails are the basic amino acid side chains of lysine and arginine residues [55, 56]. In vivo, methylated lysines can be found either in a mono-, di- or trimethylated state, whereas arginines can be either mono- or di-methylated [57]. Histone methylation has been thought to function as more of a long-term epigenetic information storage mechanism involved in propagating specific gene expression patterns in order to maintain cellular "identity and memory" [58]. On the other hand, multiple lines of evidence support that histone methylation could be dynamically controlled [57].

Histone lysine methylation occurs on histones H3 and H4. So far, methylation has been found to be conserved on six lysine residues encompassing K4, K9, K27, K36, K79 in H3 and K20 in H4. Unlike acetylation, which generally correlates with transcriptional activation, histone lysine methylation can signal either activation or

repression, depending on the site of the modification [56]. The different numbers of methyl groups added on the same site can also lead to different outcomes for certain processes. Unlike acetylation, which generally correlates with transcriptional activation, histone lysine methylation can signal either activation or repression, depending on the modification sites [56]. K9 methylation plays an important role in euchromatic gene silencing and K27 methylation is related with Hox gene silencing, whereas K4, K9 and K79 methylation is associated with transcriptional activation [59-61]. H4 K20 methylation serves as a marker of mammalian heterochromatin [62].

There are three main forms of arginine methylation identified in eukaryotes, monomethylarginines (MMA), asymmetric dimethylarginines (aDMA) and symmetric dimethylarginines (sDMA). The enzymes that catalyze this process have been divided into two types with the type-I protein arginine methyltransferase catalyzing the formation of MMA and aDMA, whereas the type-II enzymes catalyze the formation of sDMA [56]. In histones, R3 in H4 was found to be methylated by PRMT1 [63, 64]. Significantly, methylation on H4R3 facilitates acetylation on histone H4K8 and H4K12 [63]. Consistent with a role in facilitating lysine acetylation, the methyltransferase activity of PRMT1 was found to stimulate transcription [63]. R2, R17 and R26 in H3 were recently shown to be methylated by co-activator-interacting protein (CARM1), which are correlated with cell fate and potency [65-67].

1.3.1.3 Histone Phosphorylation

Histone phosphorylation is another important type of PTM that has been widely studied in recent years. Phosphorylation, particularly that of histone H3, has long been implicated in chromosome condensation during mitosis [68, 69]. The mitosis-specific phosphorylation of H3 occurs at serine 10 [70], 28 [71] and threonine 11 [72]. More recently, histone H2A was found phosphorylated at T119 in mitosis and meiosis, which serves as a prerequisite for acetylation of H3 K14 and H4 K5 [73].

Accumulating evidence indicated that the phosphorylation of H3 S10 is also directly correlated with the induction of immediate-early genes such as *c-jun*, *c-fos* and *c-myc*; thus it plays an important role in transcriptional activation [74-76]. It has been proven that there exists a cross-talk between H3 S10 phosphorylation and H3 acetylation and methylation. S10 phosphorylation can enhance H3 K14 acetylation [77, 78] and inhibit H3 K9 acetylation and methylation [79, 80]. In addition, phosphorylation of H2B S14 and H4 S1 has been linked to chromatin compaction during apoptosis and cellular response to DNA double-strand break [81-83].

1.3.2 Application of mass spectrometry for core histone PTM analysis

Identification and characterization of protein PTMs have traditionally relied on modification-specific antibodies in different immunoassay methods (e.g. western blotting and immunofluorescence). However, these methods could not provide protein sequencing information and the use of antibodies presents some potential problems, such as cross-reactivity between different modification sites, variable specificity

through interference by neighboring modifications within the recognized sequence [84]. Protein micro-sequencing also proved to be highly successful in the identification of histone post-translational modifications. However, it is a fairly cumbersome technique that requires relatively large amount of highly purified samples.

Nowadays, mass spectrometry is rapidly becoming a key analytical technique for examining protein PTMs as it reveals covalent modifications via detection of modification-specific changes of peptide molecular weight [85]. It provides direct information about the site and types of modifications, differentiates isobaric modifications (e.g. acetylation vs. tri-methylation) [86], and allows for quantitative analysis [38]. For these reasons, mass spectrometry is extensively used to study core histone PTMs.

1.3.2.1 Identification of core histone PTMs

Identification of types and sites of histone modification by MS can be accomplished by a variety of related MS methods. Core histones can be extracted from cells or tissues by nuclei isolation and acid extraction-based strategies. Each protein can be purified through reverse-phase high-performance liquid chromatography (RP-HPLC), and subjected to bottom-up or top-down MS/MS analysis.

Bottom-up approaches, including peptide mass fingerprinting (to determine molecular weight) and tandem MS (to determine peptide sequence), involve the

cleavage of a protein into peptides by enzymatic digestions followed by mass spectrometric analysis. Collisionally induced dissociation (CID) is typically used for MS/MS fragmentation. In top-down approaches, intact proteins, instead of digested peptides, were subjected to MS for fragmentation by electron capture dissociation (ECD) or electron transfer dissociation (ETD), which are capable of dissociating intact proteins.

Top-down strategies have been increasingly used to sequence proteins, map PTMs and study their cross-talk, it suffers from the protein mass limit and fragmentation efficiency. In contrast, bottom-up strategies are more mature and widely used for PTM identification. With the combination of digestion with multiple proteases and a sensitive instrument, high or even full sequence coverage of the whole protein can be obtained. CID provides efficient fragmentation for the medium-sized peptides and ETD can be applied to fragment relatively large peptides. With the observation of complete series of fragment ions, we can easily determine the modification site. Modern mass spectrometers also provide measurement with high mass accuracy; therefore, it unambiguously differentiates isobaric modifications (e.g. acetylation vs. tri-methylation).

The global PTM mappings for the modern organisms including mammals [38], yeast [87], plants [88], etc. have been completed, providing solid foundations for further studies on the regulation and functions of histone modifications.

1.3.2.2 Quantification of core histone PTMs

Along with the rapid development of mass spectrometers and mass spectrometric methodologies, quantitative analysis of histone modifications is now a routine operation [89]. The aforementioned mass spectrometry-based quantification methods are applied in the analysis of histone modification using LC-MS/MS in conjunction with stable isotope labeling strategies including SILAC, ICAT, iTRAQ, etc. The modified heavy- and light-labeled peptides are different in molecular weight, but share almost identical ionization efficiency.

Several label-free methods for PTM quantification have also been developed. Owing to their simplicity and low cost, they are employed by many research groups. For instance, proteolytic peptides can be used as internal standards to quantify modified peptides, by normalizing the MS signal of the modified peptides against the signal of internal standard peptide; alternatively peptide signals can be compared directly after ionization efficiency correction [90]. Recently, LC-MS/MS measurement with multiple-reaction-monitoring (MRM) has been applied to quantify protein PTMs. With well defined transitions and chromatographic profiling, histone modifications can be quantified efficiently by MRM measurements [91].

1.4 Scope of the Dissertation

In this dissertation, we focus on the application of mass spectrometry on two important areas, quantitative proteomics and histone post-translational modification

analysis. We studied the perturbation of protein expression in human leukemia cells upon treatment with anti-cancer drugs, the goal of which is to discover novel molecular mechanisms of action of these drugs. In addition, we mapped comprehensively the PTMs of core histone proteins in two important eukaryotic organisms and discovered novel methyl group migration during the fragmentation of singly charged trimethyllysine-containing peptides.

In Chapters 2 and 3, we report the use of mass spectrometry together with stable isotope labeling by amino acids in cell culture (SILAC) for the comparative study of protein expression in HL-60 and K562 cells that were untreated or treated with a clinically relevant concentration of arsenite or imatinib. Our results revealed that, among the more than 1000 quantified proteins, 56 and 73 proteins including many important enzymes had significantly altered levels of expression induced by arsenite and imatinib treatment, respectively. During arsenite treatment, drug-induced growth inhibition of HL-60 cells was found to be rescued by treatment with palmitate, the final product of fatty acid synthase, supporting that arsenite exerts its cytotoxic effect, in part, via suppressing the expression of fatty acid synthase and inhibiting the endogenous production of fatty acid. In the imatinib treatment project, we found, by assessing alteration in the acetylation level in histone H4 upon imatinib treatment, that the imatinib-induced hemoglobinization and erythroid differentiation in K562 cells are associated with global histone H4 hyperacetylation.

In Chapters 4 and 5, we described relatively comprehensive mappings of core

histone post-translational modifications in two important eukaryotic model organisms, *Neurospora crassa* and *Schizosaccharomyces pombe*. We used several mass spectrometric techniques, coupled with HPLC separation and multiple protease digestion, to identify the methylation, acetylation and phosphorylation sites in core histones. Electron transfer dissociation (ETD) was employed to fragment the heavily modified long N-terminal peptides. Moreover, accurate mass measurement of fragment ions allowed for unambiguous differentiation of acetylation from tri-methylation. Many modification sites conserved in other organisms were identified in these two organisms. In addition, some unique modification sites were found for the first time in core histones.

In Chapter 6, we observed an unusual discrepancy between MALDI-MS/MS and ESI-MS/MS on the methylation of trimethyllysine-containing peptides with residues 9-17 from human histone H3 and residues 73-83 from yeast histone H3. It turned out that the discrepancy could be attributed to an unusual methyl group migration from the side chain of trimethyllysine to the C-terminal arginine residue during peptide fragmentation, and this methyl group transfer occurred only for singly charged ions, but not for doubly charged ions. The methyl group transfer argument received its support from the results on the studies of the fragmentation of the ESI- or MALDI-produced singly charged ions of several synthetic trimethyllysine-bearing peptides.

References:

1. Anderson, N. L.; Anderson, N. G., The human plasma proteome - History, character, and diagnostic prospects. *Mol. Cell. Proteomics* **2002,** 1, 845-867.

2. Patterson, S. D.; Aebersold, R. H., Proteomics: the first decade and beyond. *Nature Genet.* **2003,** 33, 311-323.

3. Bantscheff, M.; Schirle, M.; Sweetman, G.; Rick, J.; Kuster, B., Quantitative mass spectrometry in proteomics: a critical review. *Anal. Bioanal. Chem.* **2007,** 389, 1017-1031.

4. Santoni, V.; Molloy, M.; Rabilloud, T., Membrane proteins and proteomics: Un amour impossible? *Electrophoresis* **2000,** 21, 1054-1070.

5. Ong, S. E.; Mann, M., Mass spectrometry-based proteomics turns quantitative. *Nat. Chem. Biol.* **2005,** 1, 252-262.

6. Julka, S.; Regnier, F., Quantification in proteomics through stable isotope coding: A review. *J. Proteome Res.* **2004,** 3, 350-363.

7. Ono, M.; Shitashige, M.; Honda, K.; Isobe, T.; Kuwabara, H.; Matsuzuki, H.; Hirohashi, S.; Yamada, T., Label-free quantitative proteomics using large peptide data sets generated by nanoflow liquid chromatography and mass spectrometry. *Mol. Cell. Proteomics* **2006,** 5, 1338-1347.

8. Leitner, A.; Lindner, W., Chemistry meets proteomics: The use of chemical tagging reactions for MS-based proteomics. *Proteomics* **2006,** 6, 5418-5434.

9. Beynon, R. J.; Pratt, J. M., Metabolic labeling of proteins for proteomics. *Mol. Cell. Proteomics* **2005,** 4, 857-872.

10. Ong, S. E.; Blagoev, B.; Kratchmarova, I.; Kristensen, D. B.; Steen, H.; Pandey, A.; Mann, M., Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* **2002,** 1, 376-386.

11. Mann, M., Functional and quantitative proteomics using SILAC. *Nat. Rev. Mol. Cell Biol.* **2006,** 7, 952-958.

12. Gygi, S. P.; Rist, B.; Gerber, S. A.; Turecek, F.; Gelb, M. H.; Aebersold, R., Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **1999,** 17, 994-999.

13. Leitner, A.; Lindner, W., Current chemical tagging strategies for proteome analysis by mass spectrometry. *J. Chromatogr. B* **2004,** 813, 1-26.

14. Zhang, R. J.; Sioma, C. S.; Wang, S. H.; Regnier, F. E., Fractionation of isotopically labeled peptides in quantitative proteomics. *Anal. Chem.* **2001,** 73, 5142-5149.

15. Hansen, K. C.; Schmitt-Ulms, G.; Chalkley, R. J.; Hirsch, J.; Baldwin, M. A.; Burlingame, A. L., Mass spectrometric analysis of protein mixtures at low levels using cleavable C-13-isotope-coded affinity tag and multidimensional chromatography. *Mol. Cell. Proteomics* **2003,** 2, 299-314.

16. Li, J. X.; Steen, H.; Gygi, S. P., Protein profiling with cleavable isotope-coded affinity tag (cICAT) reagents - The yeast salinity stress response. *Mol. Cell. Proteomics* **2003,** 2, 1198-1204.

17. Ross, P. L.; Huang, Y. L. N.; Marchese, J. N.; Williamson, B.; Parker, K.; Hattan, S.; Khainovski, N.; Pillai, S.; Dey, S.; Daniels, S.; Purkayastha, S.; Juhasz, P.; Martin, S.; Bartlet-Jones, M.; He, F.; Jacobson, A.; Pappin, D. J., Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* **2004,** 3, 1154-1169.

18. Thompson, A.; Schafer, J.; Kuhn, K.; Kienle, S.; Schwarz, J.; Schmidt, G.; Neumann, T.; Hamon, C., Tandem mass tags: A novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal. Chem.* **2003,** 75, 1895-1904.

19. Boersema, P. J.; Raijmakers, R.; Lemeer, S.; Mohammed, S.; Heck, A. J. R., Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nat. Protoc.* **2009,** 4, 484-494.

20. Yao, X. D.; Freas, A.; Ramirez, J.; Demirev, P. A.; Fenselau, C., Proteolytic O-18 labeling for comparative proteomics: Model studies with two serotypes of adenovirus. *Anal. Chem.* **2001,** 73, 2836-2842.

21. Oda, Y.; Huang, K.; Cross, F. R.; Cowburn, D.; Chait, B. T., Accurate quantitation of protein expression and site-specific phosphorylation. *Proc. Natl. Acad. Sci. U. S. A.* **1999,** 96, 6591-6596.

22. Cagney, G.; Emili, A., De novo peptide sequencing and quantitative profiling of complex protein mixtures using mass-coded abundance tagging. *Nat. Biotechnol.* **2002,** 20, 163-170.

23. Celis, J. E.; Gromov, P., Proteomics in translational cancer research: Toward an integrated approach. *Cancer Cell* **2003,** 3, 9-15.

24. Petricoin, E. F.; Zoon, K. C.; Kohn, E. C.; Barrett, J. C.; Liotta, L. A., Clinical proteomics: Translating benchside promise into bedside reality. *Nat. Rev. Drug Discov.* **2002,** 1, 683-695.

25. Hanash, S., Disease proteomics. *Nature* **2003,** 422, 226-232.

26. Wulfkuhle, J. D.; Liotta, L. A.; Petricoin, E. F., Proteomic applications for the early detection of cancer. *Nat. Rev. Cancer* **2003,** 3, 267-275.

27. Hanash, S. M.; Pitteri, S. J.; Faca, V. M., Mining the plasma proteome for cancer biomarkers. *Nature* **2008,** 452, 571-579.

28. Li, J. N.; Zhang, Z.; Rosenzweig, J.; Wang, Y. Y.; Chan, D. W., Proteomics and bioinformatics approaches for identification of serum biomarkers to detect breast cancer. *Clin. Chem.* **2002,** 48, 1296-1304.

29. Rosenwald, A.; Wright, G.; Chan, W. C.; Connors, J. M.; Campo, E.; Fisher, R. I.; Gascoyne, R. D.; Muller-Hermelink, H. K.; Smeland, E. B.; Staudt, L. M.; Lymphoma Leukemia Mol Profiling, P., The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *N. Engl. J. Med.* **2002,** 346, 1937-1947.

30. Kraljevic, S.; Sedic, M.; Scott, M.; Gehrig, P.; Schlapbach, R.; Pavelic, K., Casting light on molecular events underlying anti-cancer drug treatment: What can be seen from the proteomics point of view? *Cancer Treat. Rev.* **2006,** 32, 619-629.

31. Szakacs, G.; Paterson, J. K.; Ludwig, J. A.; Booth-Genthe, C.; Gottesman, M. M., Targeting multidrug resistance in cancer. *Nat. Rev. Drug Discov.* **2006,** 5, 219-234.

32. Pocaly, M.; Lagarde, V.; Etienne, G.; Dupouy, M.; Lapaillerie, D.; Claverol, S.; Vilain, S.; Bonneu, M.; Turcq, B.; Mahon, F. X.; Pasquet, J. M., Proteomic analysis of an imatinib-resistant K562 cell line highlights opposing roles of heat shock cognate 70 and heat shock 70 proteins in resistance. *Proteomics* **2008,** 8, 2394-2406.

33. Ferrari, G.; Pastorelli, R.; Buchi, F.; Spinelli, E.; Gozzini, A.; Bosi, A.; Santini, V., Comparative proteomic analysis of chronic myelogenous leukemia cells: Inside the mechanism of imatinib resistance. *J. Proteome Res.* **2007,** 6, 367-375.

34. Castagna, A.; Antonioli, P.; Astner, H.; Hamdan, M.; Righetti, S. C.; Perego, P.; Zunino, F.; Righetti, P. G., A proteomic approach to cisplatin resistance in the cervix squamous cell carcinoma cell line A431. *Proteomics* **2004,** 4, 3246-3267.

35. Vergnes, B.; Gourbal, B.; Girard, I.; Sundar, S.; Drummelsmith, J.; Ouellette, M., A proteomics screen implicates HSP83 and a small kinetoplastid calpain- related protein in drug resistance in Leishmania donovani clinical field isolates by modulating drug-induced programmed cell death. *Mol. Cell. Proteomics* **2007,** 6, 88-101.

36. Garcia, B. A.; Shabanowitz, J.; Hunt, D. F., Characterization of histones and their post-translational modifications by mass spectrometry. *Curr. Opin. Chem. Biol.* **2007,** 11, 66-73.

37. McGhee, J. D.; Felsenfeld, G., Nucleosome structure. In *Snell, E. E.*, 1980; pp P1115-1156.

38. Beck, H. C.; Nielsen, E. C.; Matthiesen, R.; Jensen, L. H.; Sehested, M.; Finn, P.; Grauslund, M.; Hansen, A. M.; Jensen, O. N., Quantitative proteomic analysis of posttranslational modifications of human histones. *Mol. Cell. Proteomics* **2006,** 5, 1314-1325.

39. Jenuwein, T.; Allis, C. D., Translating the histone code. *Science* **2001,** 293, 1074-1080.

40. Grunstein, M., Histone acetylation in chromatin structure and transcription. *Nature* **1997,** 389, 349-352.

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

42. Thorne, A. W.; Kmiciek, D.; Mitchelson, K.; Sautiere, P.; Cranerobinson, C., Patterns of histone acetylation. *Eur. J. Biochem.* **1990,** 193, 701-713.

43. Xiong, L.; Adhvaryu, K. K.; Selker, E. U.; Wang, Y. S., Mapping of Lysine Methylation and Acetylation in Core Histones of Neurospora crassa. *Biochemistry* **2010,** 49, 5236-5243.

44. Beck, H. C.; Nielsen, E. C.; Matthiesen, R.; Jensen, L. H.; Sehested, M.; Finn, P.; Grauslund, M.; Hansen, A. M.; Jensen, O. N., Quantitative proteomic analysis of post-translational modifications of human histones. *Mol Cell Proteomics* **2006,** 5, 1314-25.

45. Zhang, K.; Sridhar, V. V.; Zhu, J.; Kapoor, A.; Zhu, J. K., Distinctive core histone post-translational modification patterns in Arabidopsis thaliana. *PLoS One* **2007,** 2, e1210.

46. Norton, V. G.; Marvin, K. W.; Yau, P.; Bradbury, E. M., Nucleosome linking number change controlled by acetylation of histone-H3 and histone-H4. *J. Biol. Chem.* **1990,** 265, 19848-19852.

47. Lee, D. Y.; Hayes, J. J.; Pruss, D.; Wolffe, A. P., A positive role for histone acetylation in transcription factor access to nucleosomal DNA. *Cell* **1993,** 72, 73-84.

48. Kuo, M. H.; Brownell, J. E.; Sobel, R. E.; Ranalli, T. A.; Cook, R. G.; Edmondson, D. G.; Roth, S. Y.; Allis, C. D., Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines. *Nature* **1996,** 383, 269-272.

49. Grant, P. A.; Eberharter, A.; John, S.; Cook, R. G.; Turner, B. M.; Workman, J. L., Expanded lysine acetylation specificity of Gcn5 in native complexes. *J. Biol. Chem.* **1999,** 274, 5895-5900.

50. Zhang, W. Z.; Bone, J. R.; Edmondson, D. G.; Turner, B. M.; Roth, S. Y., Essential and redundant functions of histone acetylation revealed by mutation of target lysines and loss of the Gcn5p acetyltransferase. *Embo J.* **1998,** 17, 3155-3167.

51. Wan, J. S.; Mann, R. K.; Grunstein, M., Yeast histone H3 and H4 N-termini function through different GAL1 regulatory elements to repress and activate transcription. *Proc. Natl. Acad. Sci. U. S. A.* **1995,** 92, 5664-5668.

52. Turner, B. M.; Oneill, L. P., Histone acetylation in chromatin and chromosomes. *Semin. Cell Biol.* **1995,** 6, 229-236.

53. Allis, C. D.; Chicoine, L. G.; Richman, R.; Schulman, I. G., Depostion-related histone acetylation in micronuclei of conjugating tetrahymena. *Proc. Natl. Acad. Sci. U. S. A.* **1985,** 82, 8048-8052.

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

55. Kouzarides, T., Histone methylation in transcriptional control. *Curr. Opin. Genet. Dev.* **2002,** 12, 198-209.

56. Zhang, Y.; Reinberg, D., Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Genes Dev.* **2001,** 15, 2343-2360.

57. Bannister, A. J.; Schneider, R.; Kouzarides, T., Histone methylation: Dynamic or static? *Cell* **2002,** 109, 801-806.

58. Lachner, M.; O'Sullivan, R. J.; Jenuwein, T., An epigenetic road map for histone lysine methylation. *J. Cell Sci.* **2003,** 116, 2117-2124.

59. Ng, H. H.; Robert, F.; Young, R. A.; Struhl, K., Targeted recruitment of set1 histone methylase by elongating pol II provides a localized mark and memory of recent transcriptional activity. *Mol. Cell* **2003,** 11, 709-719.

60. Xiao, T. J.; Hall, H.; Kizer, K. O.; Shibata, Y.; Hall, M. C.; Borchers, C. H.; Strahl, B. D., Phosphorylation of RNA polymerase II CTD regulates H3 methylation in yeast. *Genes Dev.* **2003,** 17, 654-663.

61. Krogan, N. J.; Kim, M.; Tong, A.; Golshani, A.; Cagney, G.; Canadien, V.; Richards, D. P.; Beattie, B. K.; Emili, A.; Boone, C.; Shilatifard, A.; Buratowski, S.; Greenblatt, J., Methylation of histone H3 by Set2 in Saccharomyces cerevisiae is linked to transcriptional elongation by RNA polymerase II. *Mol. Cell. Biol.* **2003,** 23, 4207-4218.

62. Schotta, G.; Lachner, M.; Sarma, K.; Ebert, A.; Sengupta, R.; Reuter, G.; Reinberg, D.; Jenuwein, T., A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev.* **2004,** 18, 1251-1262.

63. Wang, H. B.; Huang, Z. Q.; Xia, L.; Feng, Q.; Erdjument-Bromage, H.; Strahl, B. D.; Briggs, S. D.; Allis, C. D.; Wong, J. M.; Tempst, P.; Zhang, Y., Methylation of histone H4 at arginine 3 facilitating transcriptional activation by nuclear hormone receptor. *Science* **2001,** 293, 853-857.

64. Strahl, B. D.; Briggs, S. D.; Brame, C. J.; Caldwell, J. A.; Koh, S. S.; Ma, H.; Cook, R. G.; Shabanowitz, J.; Hunt, D. F.; Stallcup, M. R.; Allis, C. D., Methylation

of histone H4 at arginine 3 occurs in vivo and is mediated by the nuclear receptor coactivator PRMT1. *Curr. Biol.* **2001,** 11, 996-1000.

65. Bannister, A. J.; Kouzarides, T., Reversing histone methylation. *Nature* **2005,** 436, 1103-1106.

66. Schurter, B. T.; Koh, S. S.; Chen, D.; Bunick, G. J.; Harp, J. M.; Hanson, B. L.; Henschen-Edman, A.; Mackay, D. R.; Stallcup, M. R.; Aswad, D. W., Methylation of histone H3 by coactivator-associated arginine methyltransferase 1. *Biochemistry* **2001,** 40, 5747-5756.

67. Torres-Padilla, M. E.; Parfitt, D. E.; Kouzarides, T.; Zernicka-Goetz, M., Histone arginine methylation regulates pluripotency in the early mouse embryo. *Nature* **2007,** 445, 214-218.

68. Bradbury, E. M., Reversible histone modifications and the chromosome cell-cycle. *Bioessays* **1992,** 14, 9-16.

69. Koshland, D.; Strunnikov, A., Mitotic chromosome condensation. *Annu. Rev. Cell Dev. Biol.* **1996,** 12, 305-333.

70. Gurley, L. R.; Danna, J. A.; Barham, S. S.; Deaven, L. L.; Tobey, R. A., Histone phosphorylation and chromatin structure during mitosis in Chinese-hamster cells. *Eur. J. Biochem.* **1978,** 84, 1-15.

71. Goto, H.; Tomono, Y.; Ajiro, K.; Kosako, H.; Fujita, M.; Sakurai, M.; Okawa, K.; Iwamatsu, A.; Okigaki, T.; Takahashi, T.; Inagaki, M., Identification of a novel

phosphorylation site on histone H3 coupled with mitotic chromosome condensation. *J. Biol. Chem.* **1999,** 274, 25543-25549.

72. Preuss, U.; Landsberg, G.; Scheidtmann, K. H., Novel mitosis-specific phosphorylation of histone H3 at Thr11 mediated by Dlk/ZIP kinase. *Nucleic Acids Res.* **2003,** 31, 878-885.

73. Cullen, C. F.; Brittle, A. L.; Ito, T.; Ohkura, H., The conserved kinase NHK-1 is essential for mitotic progression and unifying acentrosomal meiotic spindles in Drosophila melanogaster. *J. Cell Biol.* **2005,** 171, 593-602.

74. Mahadevan, L. C.; Willis, A. C.; Barratt, M. J., Rapid histone H3 phosphorylation in resonse to growth-factors, phorbol esters, okadaic acid, and protien-synthesis inhibitors. *Cell* **1991,** 65, 775-783.

75. Thomson, S.; Mahadevan, L. C.; Clayton, A. L., MAP kinase-mediated signalling to nucleosomes and immediate-early gene induction. *Semin. Cell Dev. Biol.* **1999,** 10, 205-214.

76. Chadee, D. N.; Hendzel, M. J.; Tylipski, C. P.; Allis, C. D.; Bazett-Jones, D. P.; Wright, J. A.; Davie, J. R., Increased Ser-10 phosphorylation of histone H3 in mitogen-stimulated and oncogene-transformed mouse fibroblasts. *J. Biol. Chem.* **1999,** 274, 24914-24920.

77. Cheung, P.; Tanner, K. G.; Cheung, W. L.; Sassone-Corsi, P.; Denu, J. M.; Allis, C. D., Synergistic coupling of histone H3 phosphorylation and acetylation in response to epidermal growth factor stimulation. *Mol. Cell* **2000,** 5, 905-915.

78. Lo, W. S.; Trievel, R. C.; Rojas, J. R.; Duggan, L.; Hsu, J. Y.; Allis, C. D.; Marmorstein, R.; Berger, S. L., 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, 917-926.

79. Edmondson, D. G.; Davie, J. K.; Zhou, J.; Mirnikjoo, B.; Tatchell, K.; Dent, S. Y. R., Site-specific loss of acetylation upon phosphorylation of histone H3. *J. Biol. Chem.* **2002,** 277, 29496-29502.

80. Rea, S.; Eisenhaber, F.; O'Carroll, N.; Strahl, B. D.; Sun, Z. W.; Schmid, M.; Opravil, S.; Mechtler, K.; Ponting, C. P.; Allis, C. D.; Jenuwein, T., Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* **2000,** 406, 593-599.

81. Ajiro, K., Histone H2B phosphorylation in mammalian apoptotic cells - An association with DNA fragmentation. *J. Biol. Chem.* **2000,** 275, 439-443.

82. Cheung, W. L.; Ajiro, K.; Samejima, K.; Kloc, M.; Cheung, P.; Mizzen, C. A.; Beeser, A.; Etkin, L. D.; Chernoff, J.; Earnshaw, W. C.; Allis, C. D., Apoptotic phosphorylation of histone H2B is mediated by mammalian sterile twenty kinase. *Cell* **2003,** 113, 507-517.

83. Fernandez-Capetillo, O.; Allis, C. D.; Nussenzweig, A., Phosphorylation of histone H2B at DNA double-strand breaks. *J. Exp. Med.* **2004,** 199, 1671-1677.

84. Cheung, P., Generation and characterization of antibodies directed against di-modified histones, and comments on antibody and epitope recognition. In

Chromatin and Chromatin Remodeling Enzymes, Pt B, Academic Press Inc: San Diego, 2004; Vol. 376, pp 221-234.

85. Mann, M.; Jensen, O. N., Proteomic analysis of post-translational modifications. *Nat. Biotechnol.* **2003,** 21, 255-261.

86. Zhang, K. L.; Yau, P. M.; Chandrasekhar, B.; New, R.; Kondrat, R.; Imai, B. S.; Bradbury, M. E., Differentiation between peptides containing acetylated or tri-methylated lysines by mass spectrometry: An application for determining lysine 9 acetylation and methylation of histone H3. *Proteomics* **2004,** 4, 1-10.

87. Sinha, I.; Wiren, M.; Ekwall, K., Genome-wide patterns of histone modifications in fission yeast. *Chromosome Res.* **2006,** 14, 95-105.

88. Zhang, K. L.; Sridhar, V. V.; Zhu, J. H.; Kapoor, A.; Zhu, J. K., Distinctive Core Histone Post-Translational Modification Patterns in Arabidopsis thaliana. *PLoS One* **2007,** 2.

89. Pan, S.; Aebersold, R.; Chen, R.; Rush, J.; Goodlett, D. R.; McIntosh, M. W.; Zhang, J.; Brentnall, T. A., Mass Spectrometry Based Targeted Protein Quantification: Methods and Applications. *J. Proteome Res.* **2009,** 8, 787-797.

90. Zeng, J. M.; Dunlop, R. A.; Rodgers, K. J.; Davies, M. J., Evidence for inactivation of cysteine proteases by reactive carbonyls via glycation of active site thiols. *Biochem. J.* **2006,** 398, 197-206.

91. Darwanto, A.; Curtis, M. P.; Schrag, M.; Kirsch, W.; Liu, P.; Xu, G. L.; Neidigh, J. W.; Zhang, K. L., A Modified "Cross-talk" between Histone H2B Lys-120 Ubiquitination and H3 Lys-79 Methylation. *J. Biol. Chem.* **2010,** 285, 21868-21876.

CHAPTER 2

Quantitative Proteomic Analysis Reveals the Perturbation of Multiple Cellular Pathways in HL-60 Cells Induced by Arsenite Treatment

Introduction

Arsenic is ubiquitously present in the environment from both natural and anthropogenic sources, especially in groundwater [1]. Arsenic contamination in groundwater has become a wide-spread public health problem in recent years, causing serious arsenic poisoning to a large population. Chronic arsenic exposure has been associated with increased incidence of various human diseases including atherosclerosis, diabetes, and cancers [2-4]. Despite being a human carcinogen, arsenic trioxide (As_2O_3) has also been used successfully for the clinical remission of acute promyelocytic leukemia (APL) patients [5-7], including those who are resistant to all-*trans* retinoic acid [8]; in 2001, FDA approved the use of arsenic trioxide for APL treatment.

The cellular responses toward arsenite treatment have been extensively studied over the years. Arsenite has been reported to induce the formation of micronuclei and sister chromatid exchanges [9, 10]. It can bind to cysteine sulfhydryl groups in proteins [11], stimulate the formation of oxyradicals [12], inhibit DNA repair, and modulate DNA and histone methylation in mammalian cells [13]. In addition, microarray technique revealed

that over one hundred genes in human fibroblast cells were induced or repressed by arsenite treatment [14]. However, microarray analysis does not provide information about the translational regulation of gene expression, which often exhibits a poor correlation with transcript levels owing to the different kinetics of protein translation and turnover [15].

Mass spectrometry (MS)-based proteomics allows for the identification and quantification of a large number of proteins in complex samples. Two-dimensional gel electrophoresis (2-DE) is a traditional technique for studying the effects of drug treatments on protein expression. In 2-DE, quantification is achieved by recording differences in the stained spot intensities of proteins derived from two states of cell populations or tissues [16]. A number of proteomic studies underlying the effect of treatments with different anti-cancer drugs, such as cisplatin [17, 18], etoposide [19] and all-*trans* retinoic acid [20], have been performed by using mass spectrometry for protein identification and 2-DE for protein quantification. The combination of 2-DE with mass spectrometry has also been used previously for assessing the arsenite- or arsenic trioxide-induced alterations in protein expression in rat lung epithelial cells [21], TK6 human lymphoblastoid cells [22], immortalized human keratinocytes [23], NB4 human promyelocytic leukemia cells [24], and U266 human multiple myeloma cells [25].

Other than 2-DE, several stable isotope labeling strategies, such as isotope-coded affinity tag [26], isobaric tags for relative and absolute quantitation [27] and stable isotope labeling by amino acids in cell culture (SILAC) [28], have been developed

for MS-based analysis of differential protein expression. Among these isotope-labeling strategies, SILAC is a metabolic labeling method, which is simple, efficient, and can facilitate almost complete heavy isotope incorporation. SILAC is very suitable for the comparative study of protein expression in cells with and without drug treatments; accurate results could be obtained with minimal bias, allowing for relative quantification of small changes in protein abundance [28]. In this context, Wiseman et al. [29] used SILAC together with LC-MS/MS and examined the arsenite-induced alterations in the subunit composition of the 26S human proteasome in HEK 293T cells; they found that arsenite treatment led to a 50-fold reduction in the proteasome-associated TRP32.

In the present study, we employed LC-MS/MS, together with SILAC, to assess quantitatively the perturbation of protein expression in cultured HL-60 human acute promyelocytic leukemia cells upon arsenite treatment. We were able to quantify a total of 1067 proteins in both forward and reverse SILAC measurements, among which 56 were significantly altered upon arsenite treatment. The MS quantification results for several target proteins were verified by Western blotting analysis. The identification of proteins perturbed by arsenite treatment sets a stage for understanding the biological pathways affected by arsenite treatment.

Experimental

Materials

Heavy lysine and arginine ([$^{13}C_6$,$^{15}N_2$]-L-lysine and [$^{13}C_6$,$^{15}N_4$]-L-arginine) were purchased from Cambridge Isotope Laboratories (Andover, MA). All chemicals unless otherwise noted were from Sigma (St. Louis, MO).

The rabbit anti-histone H2A, H3, actin and mouse anti-histone H4 antibodies were purchased from Abcam (Cambridge, MA). The mouse anti-histone H2B was from MBL International (Woburn, MA). The rabbit anti-fatty acid synthase was from Cell Signaling (Danvers, MA). HRP-conjugated goat anti-rabbit and anti-mouse IgG secondary antibodies were obtained from Abcam and Santa Cruz Biotechnology (Santa Cruz, CA), respectively.

Cell Culture

HL-60 cells, obtained freshly from ATCC (Manassas, VA), were cultured in Iscove's modified minimal essential medium (IMEM) supplemented with 10% fetal bovine serum (FBS, Invitrogen, Carlsbad, CA) and penicillin (100 IU/mL). Cells were maintained in a humidified atmosphere with 5% CO_2 at 37ºC, with medium renewal at every 2 or 3 days depending on cell density. For SILAC experiments, the IMEM medium without L-lysine or L-arginine was custom-prepared according to the ATCC formulation. The complete light and heavy IMEM media were prepared by the addition of light or heavy lysine and arginine, along with dialyzed FBS (Invitrogen), to the above lysine, arginine-depleted medium. The HL-60 cells were cultured in

heavy IMEM medium for at least 5 cell doublings to achieve complete isotope incorporation.

Arsenite Treatment and Cell Lysate Preparation

HL-60 cells, at a density of approximately 7.5×10^5 cells/mL, were collected by centrifugation at 300 g and at 4ºC for 5 min, washed twice with ice-cold phosphate-buffered saline (PBS) to remove FBS, and resuspended in FBS-free heavy or light media. In forward SILAC experiment, the cells cultured in light medium were treated with 5 μM arsenite (Sigma) for 24 hrs, whereas the cells cultured in heavy medium were untreated. Reverse SILAC experiments were also performed where the cells cultured in the heavy and light medium were treated with arsenite and mock-treated, respectively (Figure 2.1). After 24 hrs, the light and heavy cells were collected by centrifugation at 300 g, and washed three times with ice-cold PBS.

The cell pellets were resuspended in the CelLytic™ M cell lysis buffer (Sigma) for 30 min with occasional vortexing. Cell lysates were centrifuged at 12,000 g at 4°C for 30 min, and the resulting supernatants were collected. To the supernatant was subsequently added a protease inhibitor cocktail (Sigma), and the protein concentrations of the cell lysates were determined by using Quick Start Bradford Protein Assay kit (Bio-Rad, Hercules, CA).

SDS-PAGE Separation and In-gel Digestion

The light and heavy cell lysates were combined at 1:1 ratio (w/w), denatured by boiling in Laemmli loading buffer for 5 min and separated by a 12% SDS-PAGE with 4% stacking gel. The gel was stained with Coomassie blue; after destaining, the gel was cut into 20 bands, in-gel reduced with dithiothreitol and alkylated with iodoacetamide. The proteins were digested in-gel with trypsin (Promega, Madison, WI) for overnight, after which peptides were extracted from gels with 5% acetic acid in H_2O and in CH_3CN/H_2O (1:1, v/v). The resulting peptide mixtures were dried and stored at –20°C for further analysis.

Western Blotting

For Western blotting analysis, lysates of the control and arsenite-treated HL-60 cells were prepared following the same procedures as described above. After SDS-PAGE separation, proteins were transferred to a nitrocellulose membrane under standard conditions. A basic transfer buffer containing 10 mM $NaHCO_3$ and 3 mM Na_2CO_3 (pH 9.9) was used for immunoblotting histones. After protein transfer, the membranes were blocked with 5% non-fat milk in PBS-T buffer [PBS solution containing 0.1% (v/v) Tween-20, pH 7.5] for 7 hrs and incubated subsequently with primary antibodies for overnight at optimized dilution ratios. The membranes were washed five times (10 min each) with fresh changes of PBS-T at room temperature. After washing, the membranes were incubated with horseradish peroxidase (HRP)-conjugated secondary antibodies at room temperature for 50 min. The

membranes were washed thoroughly with PBS-T for five times (10 min each). The secondary antibody was detected by using ECL Advance Western Blotting Detection Kit (GE Healthcare) and visualized with HyBlot CL autoradiography film (Denville Scientific Inc., Metuchen, NJ).

Fatty Acid Synthase Inhibition and Cell Viability Assay

The palmitate-BSA complex was prepared following a previously published method [30]. Bovine serum albumin (BSA, 5 g) was dissolved in 25 mL of 0.9% NaCl solution with pH being adjusted to 7.4 by NaOH at 5ºC, while 4.016 mg sodium palmitate (Sigma) was dissolved in 15 mL of the same NaCl solution at 60ºC. The BSA solution was then added to the hot palmitate solution, and the resulting BSA-palmitate mixture was stirred extensively and brought to a volume of 40 mL at a palmitate/BSA molar ratio of 1:5.

HL-60 cells were washed twice with ice-cold PBS to remove FBS, resuspended in FBS-free IMEM medium, and seeded in 6-well plates at a density of ~ 3×10^5 cells/mL. To the cultured cells were added cerulenin or arsenite solutions until their concentrations reached 2.5 μg/mL and 5 μM, respectively. The palmitate-BSA complex solution was added subsequently to the wells containing the control, cerulenin-, or arsenite-treated cells until palmitate concentration was 40 or 80 μM. After 24 or 48 hrs of treatment, cells were stained with trypan blue, and counted on a hemocytometer to measure cell viability.

LC-MS/MS for Protein Identification and Quantification

Online LC-MS/MS analysis was performed on an Agilent 6510 Q-TOF system coupled with an Agilent HPLC-Chip Cube MS interface (Agilent Technologies, Santa Clara, CA). The sample injection, enrichment, desalting, and HPLC separation were carried out automatically on the Agilent HPLC Chip with an integrated trapping column (160 nL) and a separation column (Zorbax 300SB-C18, 75 μm×150 mm, 5 μm in particle size). The peptide mixture was first loaded onto the trapping column with a solvent mixture of 0.1% formic acid in CH_3CN/H_2O (2:98, v/v) at a flow rate of 4 μL/min, which was delivered by an Agilent 1200 capillary pump. The peptides were then separated with a 90-min linear gradient of 2-60% acetonitrile in 0.1% formic acid and at a flow rate of 300 nL/min, which was delivered by an Agilent 1200 Nano pump.

The Chip spray voltage (VCap) was set as 1950 V and varied depending on chip conditions. The temperature and flow rate of the drying gas were set at 325ºC and 4 L/min, respectively. Nitrogen was used as the collision gas, and the collision energy followed an equation with a slope of 3 V/100 Da and an offset of 2.5 V. MS/MS experiments were carried out in the data-dependent scan mode with a maximum of five MS/MS scans following each MS scan. The m/z ranges for MS and MS/MS were 300-2000 and 60-2000, and the acquisition rates were 6 and 3 spectra/s, respectively.

Data Processing

Agilent MassHunter workstation software (Version B.01.03) was used to extract the MS and MS/MS data. The data were converted to m/z Data files with MassHunter Qualitative Analysis. Mascot Server 2.2 (Matrix Science, London, UK) was used for protein identification by searching the m/z Data files against the weekly-updated NCBI database. The maximum number of miss-cleavage for trypsin was set as one per peptide. Cysteine carbamidomethylation was set as a fixed modification. Methionine oxidation as well as lysine (+8 Da) and arginine (+10 Da) mass shifts introduced by heavy isotope labeling were considered as variable modifications. The mass tolerances for MS and MS/MS were 50 ppm and 0.6 Da, respectively. Peptides identified with individual scores at or above the Mascot-assigned homology score ($p < 0.01$ and individual peptide score > 30) were considered as specific peptide sequences. The false discovery rates (FDR) determined by decoy database search were less than 0.95%.

Protein quantification was carried out manually and was based on the average ratios of peptide pairs, which was obtained from the intensities of peaks of precursor ions of light and heavy peptides. For those proteins identified with less than 5 peptides, the abundance ratios for ions of all peptide pairs were used for the quantification; for those proteins identified with more than 5 peptides, five peptide pairs with the highest scores were chosen to determine the protein ratios. The ratio

obtained for each individual protein was then normalized against the average ratio for all quantified proteins. This "multi-point" normalization strategy assumes that the ratios for the majority of proteins are not affected by the treatment, facilitating the use of the average ratio of all quantified proteins for re-scaling the data. This has been widely employed to eliminate the inaccuracy during sample mixing introduced by protein quantification with the Bradford assay [31, 32]. The quantification was based on three independent SILAC and LC-MS/MS experiments, which included two forward and one reverse SILAC labeling, and the proteins reported here could be quantified in both forward and reverse SILAC experiments. Some peptides identified in only 1 or 2 trials of QTOF analysis could be quantified in all three trials, where the accurate mass of peptide ions, retention time, and the numbers of isotope-labeled lysine and/or arginine in the peptide were employed as criteria to locate the light/heavy peptide pairs for the quantification.

Results

Arsenite Treatment, Protein Identification and Quantification

To gain insights into the molecular pathways perturbed by arsenite treatment, we employed SILAC combined with LC-MS/MS to assess the arsenite-induced differential expression of the whole proteome of HL-60 cells. In this context, clinical pharmacokinetic analyses indicate that the peak plasma arsenite concentration is in the low [33] to high [8, 34] micromolar range for APL patients treated with arsenite. We

observed, based on trypan blue exclusion assay, a less than 5% cell death after a 24-hr treatment with 5 μM arsenite, whereas approximately 20% cells were dead if the cells were treated with 7.5 μM arsenite. Thus, we decided to employ 5 μM arsenite for the subsequent experiments to minimize the apoptosis-induced alteration in protein expression.

To obtain reliable results, we carried out the SILAC experiments in triplicate and both forward and reverse SILAC labeling was performed (Figure 2.1A, see also Materials and Methods). Figure 2.2 shows example results for the quantification of the peptide GTPLISPLIK from fatty acid synthase, which reveals clearly the down-regulation of this protein in both forward and reverse SILAC experiments results were obtained from forward, and "Set 3" results were from reverse SILAC experiments (Figure 2.2A&B). The peptide sequence was confirmed by MS/MS analysis (Figure 2.2C&D). In addition, we employed serum-free medium in all SILAC-related cell culture experiments to avoid the interactions between arsenite and proteins in the FBS. After cell lysis, SDS-PAGE fractionation, in-gel digestion, LC-MS/MS analysis and database search, we were able to identify 1401 proteins, among which 1380 could be quantified.

Among the quantified proteins, 953 could be quantified in all three measurements, 182 could be quantified in two measurements, and another 245 could be quantified in only one measurement (Figure 2.1B). We include here only the quantification results for those proteins that could be quantified in all three

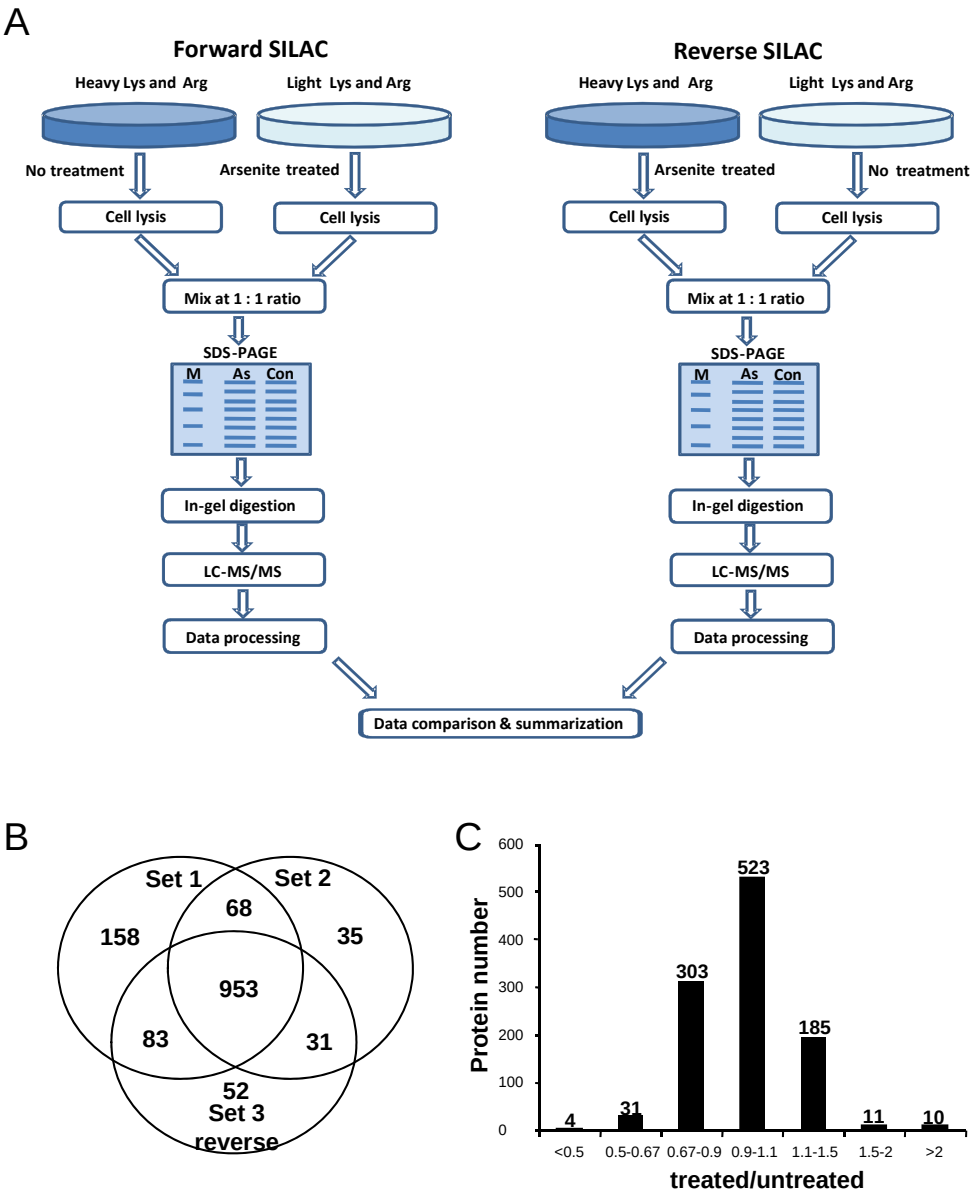


Figure 2.1 Forward- and reverse-SILAC combined with LC-MS/MS for the comparative analysis of protein expression in HL-60 cells upon arsenite treatment (A). Shown in (B) and (C) are a summary of the number of proteins quantified from three independent SILAC experiments and the distribution of expression ratios (treated/untreated) for the proteins quantified, respectively. "Set 1" and "Set 2"

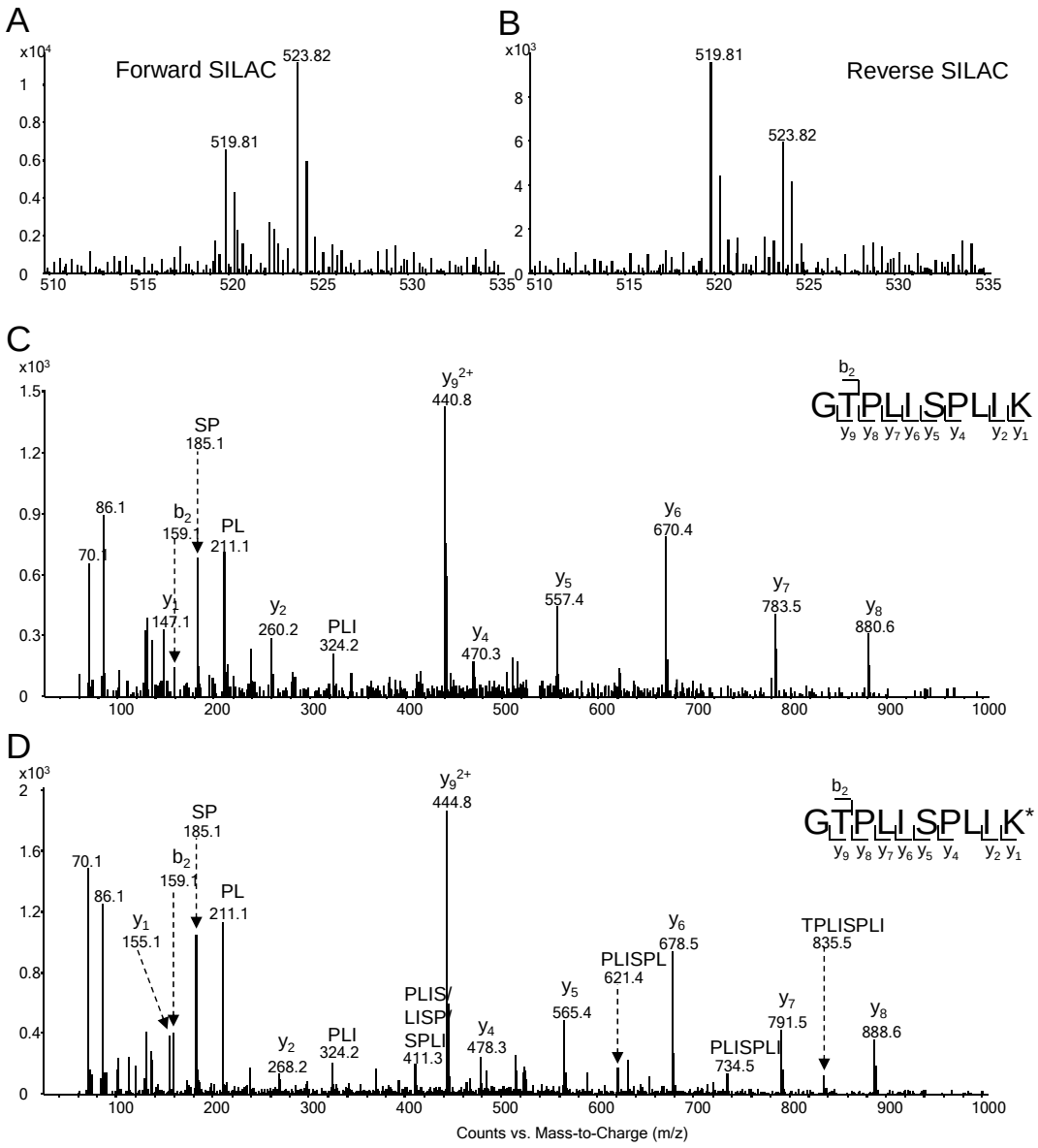


Figure 2.2 Example ESI-MS and MS/MS data revealed the arsenite-induced down-regulation of fatty acid synthase. Shown are the MS for the $[M+2H]^{2+}$ ions of FAS peptide GTPLISPLIK and GTPLISPLIK* ('K*' represents the heavy lysine) from the forward (A) and reverse (B) SILAC samples. Depicted in (C) and (D) are the MS/MS for the $[M+2H]^{2+}$ ions of GTPLISPLIK and GTPLISPLIK*, respectively.

experiments or in two experiments which include both the forward and reverse SILAC. Together, this gives quantifiable results for 1067 proteins.

The distribution of the changes in protein expression levels induced by arsenite treatment is shown in Figure 2.1C. Among the quantified proteins, 56 display significant changes upon arsenite treatment (the ratio of treated/untreated was greater than 1.5 or less than 0.67), with 21 and 35 being up- and down-regulated, respectively. The quantification results for the proteins with significant changes are summarized in Table 2.1, and the detailed information about the quantified peptides and ratios for each measurement are listed in Table S2.1.

Histone proteins are up-regulated upon arsenite treatment

Among the differentially expressed proteins, all four core histones and linker histone H1 were markedly up-regulated in arsenite-treated HL-60 cells (Table S2.2). In this respect, histone proteins bear many sites of post-translational modifications (PTMs), particularly on their N-terminal tails [35]. Arsenite treatment may perturb the PTMs of histone proteins, which may give rise to inaccurate quantifications of their expression levels. Indeed the exposure of A549 human lung carcinoma cells to arsenite was found to increase H3K9 dimethylation and H3K4 trimethylation while decreasing H3K27 trimethylation [36]. To avoid the inaccurate quantification of histones introduced by arsenite-induced change in PTMs, we chose to use those peptides that do not contain any known PTMs for the quantification [35]. Consistent with what we

Table 2.1 Proteins quantified with more than 1.5 fold changes, with GI numbers, protein names, average ratios and S.D. listed (Peptides used for the quantification of individual proteins are listed in Table S2.2).

GI Number	Protein Name	Ratio (treated/untreated)
A. Histone and HMG proteins		
11321591	HMG-2	1.65±0.26
968888	HMG-1	1.85±0.07
223582	histone H4	2.41±0.09
386772	histone H3	2.80±0.42
356168	histone H1b	2.76±0.33
1568557	histone H2B	2.98±0.45
510990	histone H2A	4.21±0.54
B. Translation-related proteins		
181969	elongation factor 2	0.54±0.09
38202255	threonyl-tRNA synthetase	0.60±0.12
19353009	Similar to elongation factor 2b	0.62±0.06
4506707	ribosomal protein S25	0.65±0.11
C. hnRNPs		
386547	d(TTAGGG)n-binding protein B39	0.50±0.22
119597533	hnRNP U, isoform CRA_b	0.59±0.01
55958547	hnRNP K	0.66±0.15
D. Enzymes		
531202	spermidine synthase	0.57±0.18
41584442	fatty acid synthase	0.61±0.10
66392203	NME1-NME2 protein	0.64±0.18
1230564	Gu protein	0.64±0.14
2661039	α-enolase	0.65±0.07
190281	protein phosphatase I α subunit	0.66±0.02
4507789	ubiquitin-conjugating enzyme E2L 3 isoform 1	0.67±0.06
5231228	ribonuclease T2 precursor	1.53±0.21
4758504	hydroxysteroid (17-β) dehydrogenase 10 isoform 1	1.56±0.02
4503143	cathepsin D	1.65±0.08
186461558	neutrophil elastase	1.95±0.01

627372	α-mannosidase precursor	2.05±0.24
4504437	heme oxygenase 1	2.32±0.78
12654715	TXNDC5 protein	3.14±0.35
E. Others		
21755073	unnamed protein product	0.42±0.02
119587276	hCG19802, isoform CRA_a	0.47±0.20
35570	unnamed protein product	0.50±0.07
13477237	ZNF607 protein	0.52±0.05
194374111	unnamed protein product	0.57±0.11
35844	unnamed protein product	0.58±0.15
431422	Ran/TC4 binding protein	0.60±0.22
114645930	nucleosome assembly protein 1-like 1 isoform 9	0.60±0.02
2580550	DEAD box, X isoform	0.61±0.10
122168	HLA class I histocompatibility antigen, B-58 α chain	0.61±0.20
181486	DNA-binding protein B	0.63±0.07
193788267	unnamed protein product	0.63±0.08
801893	leucine-rich PPR-motif containing protein	0.63±0.10
5107666	importin β	0.63±0.20
1235727	unnamed protein product	0.64±0.10
40225729	FUBP1 protein	0.65±0.08
119571409	hCG1643342, isoform CRA_a	0.65±0.03
386777	transplantation antigen	0.66±0.05
1136741	KIAA0002	0.66±0.20
23712	myoblast antigen 24.1D5	0.66±0.08
13569879	acidic (leucine-rich) nuclear phosphoprotein 32 family, member E isoform 1	0.67±0.08
4506773	S100 calcium-binding protein A9	1.53±0.04
9955206	Rho GDP-dissociation factor 2	1.55±0.19
8037945	prothymosin α	1.56±0.02
119239	bone marrow proteoglycan	1.71±0.08
28375485	unnamed protein product	1.81±0.22
4506191	proteasome β10 subunit proprotein	2.21±0.10
32111	unnamed protein product	4.90±0.96

found from SILAC and LC-MS/MS analysis, Western blotting results also revealed that all core histones were expressed at higher levels in arsenite-treated than in control HL-60 cells (Figure S2.1).

Although the mechanisms through which the histone proteins are up-regulated upon arsenite treatment remain unknown, many studies showed that arsenite could induce chromosome damage and modulate DNA methylation in mammalian cells [9, 13]. The substantially increased expression of histones might be reflective of the considerable change in chromatin structure induced by arsenite treatment.

Modest down-regulation of heterogeneous nuclear ribonuclear proteins (hnRNPs) and proteins involved in translation

Aside from the considerable upregulation of histone proteins, arsenite treatment also led to a systematic down-regulation of hnRNPs and several important groups of proteins involved in translation (Table S2.3). In this context, hnRNPs, ribosomal proteins, translation initiation proteins, and translation elongation proteins were all modestly down-regulated upon arsenite treatment (Table S2.3); the average ratios (treated/untreated) for these four groups of proteins are 0.85, 0.88, 0.82 and 0.82, respectively. These results are in accordance with the previous findings [24, 37] and with the growth inhibition induced by arsentite treatment (*vide infra*).

Arsenite induced the down-regulation of fatty acid synthase (FAS)

Our LC-MS/MS results showed that FAS was down-regulated by approximately 40% upon arsenite treatment. The LC-MS/MS quantification result was validated by Western blotting analysis, which revealed that arsenite treatment gives rise to the dose-dependant decrease in the expression of FAS (Figure 2.3A). FAS is the sole protein in the human genome capable of reductive synthesis of long-chain fatty acids from acetyl-coenzyme A (acetyl-CoA), malonyl-CoA and NADPH [38]. It has been reported that FAS is highly expressed in human carcinomas [39]. Inhibitors of FAS are selectively toxic to cell lines derived from human malignancies, supporting that cancer cells rely on endogenous fatty acid synthesis for survival and inhibition of FAS may afford an effective route for cancer treatment and prevention [39].

We reason that the decreased expression of FAS may account partly for the cytotoxic effect of arsenite. If this is the case, the arsenite-induced growth inhibition of HL-60 cells should be rescued by palmitate, the end product of FAS. To test this, we assessed whether the proliferation of HL-60 cells is perturbed by arsenite treatment and how this perturbation is affected by externally added palmitate. It turned out that arsenite treatment inhibits the proliferation of HL-60 cells, and palmitate indeed protects, in a dose-dependent fashion, HL-60 cells from arsenite-induced growth inhibition (Figure 2.3B&C). It was previously observed that the cytotoxicity of cerulenin, a common FAS inhibitor [40, 41], could be abolished by co-administration with exogenous fatty acid [42]. We also observed that the survival of

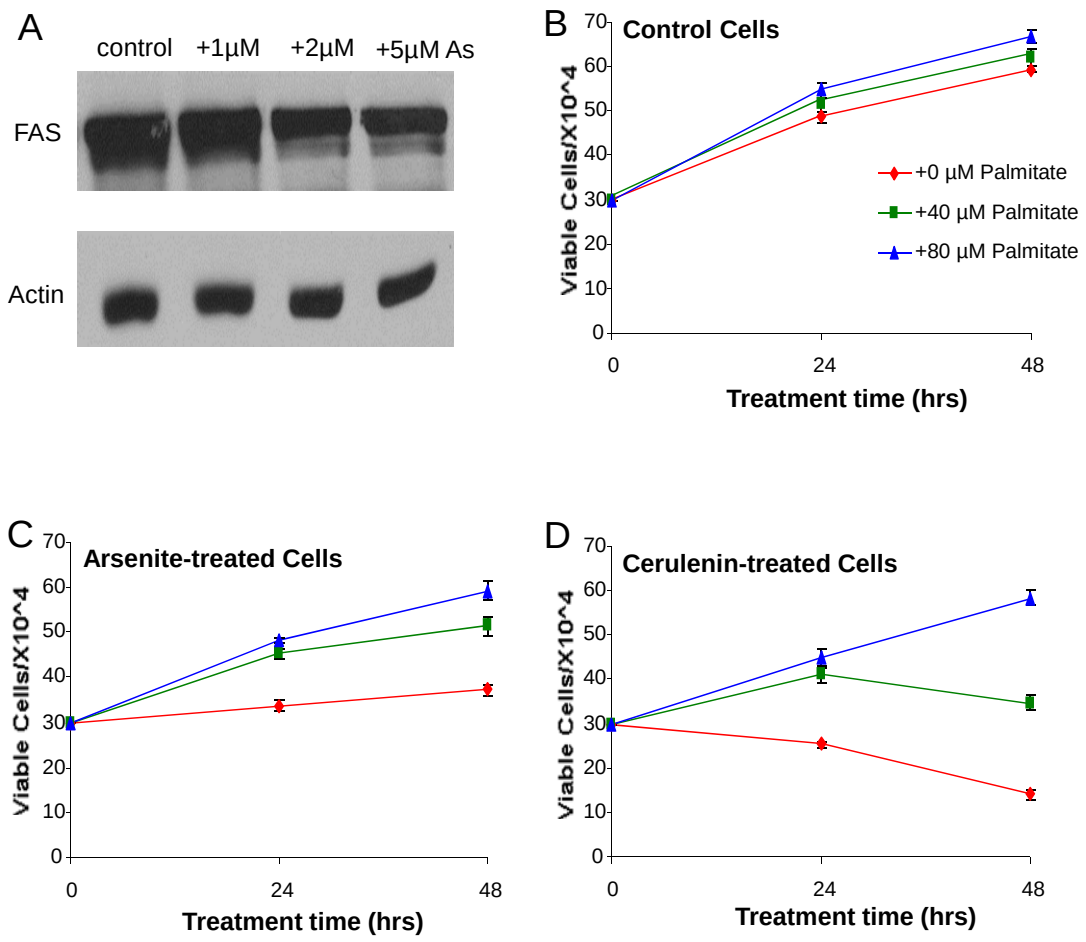


Figure 2.3 Western blotting analysis of Fatty acid synthase (FAS) with lysates of untreated HL-60 cells ("control") and HL-60 cells that are treated with 1 μM, 2 μM, or 5 μM of arsenite for 24 hrs (A), and actin was used as the loading control. The viabilities of HL-60 cells after 24 and 48 hrs of treatment with 0, 40, 80 μM palmitate alone (B), or together with 5 μM arsenite (C) or 2.5 μg/mL cerulenin (D).

HL-60 cells is compromised upon incubation with cerulenin, which can again be rescued by palmitate (Figure 2.3D). This result corroborates with the arsenite-induced down-regulation of FAS and underscores that arsenite may induce cytotoxic effect by inhibiting endogenous fatty acid synthesis through suppressing FAS expression. FAS expression has been found to be controlled by many pathways including Akt [43], and by tumor suppressors and oncogenes, which encompass p53, p63, p73, and H-ras [44, 45]. Further study is needed to unravel the mechanisms involved in the arsenite-mediated down-regulation of FAS.

Arsenite induced the alteration in expression of other important enzymes

Arsenite treatment also gave rise to considerable changes in the expression levels of some other important enzymes, including neutrophil elastase, cathepsin D, α-enolase, α-mannosidase, etc. (Table 2.1). These proteins play pivotal roles in different cellular pathways, and we would like to discuss some of them in detail.

Neutrophil elastase (NE) is the key protease involved in the cleavage of PML-RARα fusion protein in mouse and human APL cells [46]. The PML-RARα fusion protein generated by the t(15;17) translocation, which is associated with APL, initiated APL when expressed in the early myeloid compartment [46]. Arsenite was shown to induce the degradation of the PML-RARα fusion protein in NB4-S1 cells, which was attributed to the destabilization of lysosome and the subsequent release of hydrolytic enzyme (i.e., cathepsin L) to the cytosol [47]. We observed that NE was

up-regulated by approximately 2 fold upon arsenite treatment. Viewing that neutrophil elastase is the dominant PML-RARα cleaving activity in human APL cells [46], our result suggests that arsenite treatment may result in the degradation of the PML-RARα protein through stimulating the expression of NE.

Cathepsin D (CatD) is a lysosomal aspartic protease found in neutrophils and monocytes [48]. CatD is an important cell death mediator [49], and fibroblasts from CatD-deficient mice display more resistance toward etoposide- and adriamycin-induced apoptosis than fibroblasts from the wild-type littermates [50]. Upon arsenite treatment, CatD was up-regulated by 60% (Table 2.1). The arsenite-induced up-regulation of CatD may also contribute, in part, to the cytotoxic effect of arsenite.

Aside from the enzymes discussed above, there are many other important proteins that are significantly up- or down-regulated by arsenite in HL-60 cells (Table 2.1). For instance, we found that prothymosin α was substantially up-regulated upon arsenite treatment. This protein is known to compete with transcription factor Nrf2 for binding to the same domain on Keap1, thereby releasing Nrf2 from the Nrf2-Keap1 inhibitory complex [51]. The liberation of Nrf2 can result in the up-regulation of genes involved in defense against oxidative stress and electrophilic attack by binding to the anti-oxidant response elements in promoters of these genes [51-53]. Consistent with this notion, we indeed observed that arsenite treatment led to substantial up-regulation of heme oxygenase 1 (Table 2.1), a Nrf2 target gene [53], and arsenite is known to induce the formation of reactive oxygen species [12].

Discussion and Conclusions

Arsenite is an established human carcinogen [54]; on the other hand, it is also a clinically successful anticancer drug for APL treatment [7]. In this study, we employed SILAC, together with LC-MS/MS, and assessed quantitatively the perturbation of protein expression in HL-60 leukemic cells induced by arsenite treatment. Our results revealed that the treatment led to the up- or down-regulation of many important proteins, including histones, fatty acid synthase, neutrophil elastase, cathepsin D, etc. In addition, most ribosomal proteins, translation initiation factors, translation elongation factors, and hnRNPs were modestly down-regulated upon the treatment.

We also confirmed the up- and down-regulation of several proteins by Western blotting analysis. Among the proteins whose expressions are perturbed by arsenite, the down-regulation of fatty acid synthase is of particular importance. In this context, normal adult cells acquire fatty acids mainly from dietary sources and rarely rely on *de novo* fatty acid synthesis because nutritional fatty acids inhibit strongly the expression of the genes involved in fatty acid synthesis [55]. Cancer cells, however, are no longer sensitive to this nutritional signal and instead depend on endogenous fatty acid synthesis to provide vital structural lipids that are required for survival and proliferation [42]. Therefore, fatty acid synthase is substantially up-regulated in many types of tumors and inhibition of this enzyme has been suggested for cancer treatment [39]. Our cell survival data support that the arsenite-induced growth inhibition of HL-60 cells can be rescued by palmitate, the final product of fatty acid synthase. This finding

parallels the growth inhibition induced by cerulenin, which is a specific noncompetitive inhibitor of the β-ketoacyl transferase activity of FAS [40, 41]. Thus, our result underscored that the inhibition of endogenous fatty acid synthase may constitute a novel mechanism for arsenite-induced cytotoxic effect.

Further studies about the implications of arsenite-induced perturbation of other proteins may also lead to the discovery of additional molecular pathways that are altered by arsenite and contribute to the arsenite-induced cytotoxicity. The pharmacoproteomic profiling could constitute a valuable tool for the identification of drug-responsive biomarkers for arsenite treatment and establish a molecular basis for developing novel and more effective therapeutic approaches for the treatment of APL and other human cancers.

References:

1. National Research Council, *Arsenic in Drinking Water*. National Academy Press: Washington DC, 1999; and 2001 update.

2. Abernathy, C. O.; Liu, Y. P.; Longfellow, D.; Aposhian, H. V.; Beck, B.; Fowler, B.; Goyer, R.; Menzer, R.; Rossman, T.; Thompson, C.; Waalkes, M., Arsenic: Health effects, mechanisms of actions, and research issues. *Environ. Health Perspect.* **1999,** 107, 593-597.

3. Morales, K. H.; Ryan, L.; Kuo, T. L.; Wu, M. M.; Chen, C. J., Risk of internal cancers from arsenic in drinking water. *Environ. Health Perspect.* **2000,** 108, 655-661.

4. Cantor, K. P.; Lubin, J. H., Arsenic, internal cancers, and issues in inference from studies of low-level exposures in human populations. *Toxicol. Appl. Pharmacol.* **2007,** 222, 252-7.

5. Chen, G. Q.; Shi, X. G.; Tang, W.; Xiong, S. M.; Zhu, J.; Cai, X.; Han, Z. G.; Ni, J. H.; Shi, G. Y.; Jia, P. M.; Liu, M. M.; He, K. L.; Niu, C.; Ma, J.; Zhang, P.; Zhang, T. D.; Paul, P.; Naoe, T.; Kitamura, K.; Miller, W.; Waxman, S.; Wang, Z. Y.; deThe, H.; Chen, S. J.; Chen, Z., Use of arsenic trioxide (As_2O_3) in the treatment of acute promyelocytic leukemia (APL) .1. As_2O_3 exerts dose-dependent dual effects on APL cells. *Blood* **1997,** 89, 3345-3353.

6. Chen, G. Q.; Zhu, J.; Shi, X. G.; Ni, J. H.; Zhong, H. J.; Si, G. Y.; Jin, X. L.; Tang, W.; Li, X. S.; Xong, S. M.; Shen, Z. X.; Sun, G. L.; Ma, J.; Zhang, P.; Zhang, T.

D.; Gazin, C.; Naoe, T.; Chen, S. J.; Wang, Z. Y.; Chen, Z., In vitro studies on cellular and molecular mechanisms of arsenic trioxide (As_2O_3) in the treatment of acute promyelocytic leukemia: As_2O_3 induces NB4 cell apoptosis with downregulation of bcl-2 expression and modulation of PML-RAR alpha/PML proteins. *Blood* **1996,** 88, 1052-1061.

7. Zhu, J.; Chen, Z.; Lallemand-Breitenbach, V.; de The, H., How acute promyelocytic leukaemia revived arsenic. *Nat. Rev. Cancer* **2002,** 2, 705-13.

8. Shen, Z. X.; Chen, G. Q.; Ni, J. H.; Li, X. S.; Xiong, S. M.; Qiu, Q. Y.; Zhu, J.; Tang, W.; Sun, G. L.; Yang, K. Q.; Chen, Y.; Zhou, L.; Fang, Z. W.; Wang, Y. T.; Ma, J.; Zhang, P.; Zhang, T. D.; Chen, S. J.; Chen, Z.; Wang, Z. Y., Use of arsenic trioxide (As_2O_3) in the treatment of acute promyelocytic leukemia (APL): II. Clinical efficacy and pharmacokinetics in relapsed patients. *Blood* **1997,** 89, 3354-60.

9. Lerda, D., Sister-chromatid exchange (Sce) among individuals chronically exposed to arsenic in drinking-water. *Mutat. Res.* **1994,** 312, 111-120.

10. Gonsebatt, M. E.; Vega, L.; Salazar, A. M.; Montero, R.; Guzman, P.; Blas, J.; DelRazo, L. M.; GarciaVargas, G.; Albores, A.; Cebrian, M. E.; Kelsh, M.; OstroskyWegman, P., Cytogenetic effects in human exposure to arsenic. *Mutat. Res.* **1997,** 386, 219-228.

11. Yan, H.; Wang, N.; Weinfeld, M.; Cullen, W. R.; Le, X. C., Identification of arsenic-binding proteins in human cells by affinity chromatography and mass spectrometry. *Analytical chemistry* **2009,** 81, 4144-52.

12. Liu, S. X.; Athar, M.; Lippai, I.; Waldren, C.; Hei, T. K., Induction of oxyradicals by arsenic: implication for mechanism of genotoxicity. *Proc. Natl. Acad. Sci. USA* **2001,** 98, 1643-8.

13. Kitchin, K. T., Recent advances in arsenic carcinogenesis: Modes of action, animal model systems, and methylated arsenic metabolites. *Toxicol. Appl. Pharmacol.* **2001,** 172, 249-261.

14. Yih, L. H.; Peck, K.; Lee, T. C., Changes in gene expression profiles of human fibroblasts in response to sodium arsenite treatment. *Carcinogenesis* **2002,** 23, 867-876.

15. Wilkins, M. R.; Sanchez, J. C.; Williams, K. L.; Hochstrasser, D. F., Current challenges and future applications for protein maps and post-translational vector maps in proteome projects. *Electrophoresis* **1996,** 17, 830-838.

16. Klose, J.; Kobalz, U., 2-Dimensional Electrophoresis of Proteins - an Updated Protocol and Implications for a Functional-Analysis of the Genome. *Electrophoresis* **1995,** 16, 1034-1059.

17. Castagna, A.; Antonioli, P.; Astner, H.; Hamdan, M.; Righetti, S. C.; Perego, P.; Zunino, F.; Righetti, P. G., A proteomic approach to cisplatin resistance in the cervix squamous cell carcinoma cell line A431. *Proteomics* **2004,** 4, 3246-3267.

18. Sinha, P.; Poland, J.; Kohl, S.; Schnolzer, M.; Helmbach, H.; Hutter, G.; Lage, H.; Schadendorf, D., Study of the development of chemoresistance in melanoma cell lines using proteome analysis. *Electrophoresis* **2003,** 24, 2386-2404.

19. Urbani, A.; Poland, J.; Bernardini, S.; Bellincampi, L.; Biroccio, A.; Schnolzer, M.; Sinha, P.; Federici, G., A proteomic investigation into etoposide chemo-resistance of neuroblastoma cell lines. *Proteomics* **2005,** 5, 796-804.

20. Bertagnolo, V.; Grassilli, S.; Bavelloni, A.; Brugnoli, F.; Piazzi, M.; Candiano, G.; Petretto, A.; Benedusi, M.; Capitani, S., Vav1 modulates protein expression during ATRA-induced maturation of APL-derived promyelocytes: A proteomic-based analysis. *J. Proteome Res.* **2008,** 7, 3729-3736.

21. Lau, A. T.; He, Q. Y.; Chiu, J. F., A proteome analysis of the arsenite response in cultured lung cells: evidence for in vitro oxidative stress-induced apoptosis. *Biochem. J.* **2004,** 382, 641-50.

22. Tapio, S.; Danescu-Mayer, J.; Asmuss, M.; Posch, A.; Gomolka, M.; Hornhardt, S., Combined effects of gamma radiation and arsenite on the proteome of human TK6 lymphoblastoid cells. *Mutat. Res.* **2005,** 581, 141-52.

23. Berglund, S. R.; Santana, A. R.; Li, D.; Rice, R. H.; Rocke, D. M.; Goldberg, Z., Proteomic analysis of low dose arsenic and ionizing radiation exposure on keratinocytes. *Proteomics* **2009,** 9, 1925-38.

24. Zheng, P. Z.; Wang, K. K.; Zhang, Q. Y.; Huang, Q. H.; Du, Y. Z.; Zhang, Q. H.; Xiao, D. K.; Shen, S. H.; Imbeaud, S.; Eveno, E.; Zhao, C. J.; Chen, Y. L.; Fan, H. Y.; Waxman, S.; Auffray, C.; Jin, G.; Chen, S. J.; Chen, Z.; Zhang, J., Systems analysis of transcriptome and proteome in retinoic acid/arsenic trioxide-induced cell

differentiation/apoptosis of promyelocytic leukemia. *Proc. Natl. Acad. Sci. USA* **2005,** 102, 7653-8.

25. Ge, F.; Lu, X. P.; Zeng, H. L.; He, Q. Y.; Xiong, S.; Jin, L., Proteomic and functional analyses reveal a dual molecular mechanism underlying arsenic-induced apoptosis in human multiple myeloma cells. *J. Proteome Res.* **2009,** 8, 3006-19.

26. Gygi, S. P.; Rist, B.; Gerber, S. A.; Turecek, F.; Gelb, M. H.; Aebersold, R., Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **1999,** 17, 994-999.

27. Ross, P. L.; Huang, Y. L. N.; Marchese, J. N.; Williamson, B.; Parker, K.; Hattan, S.; Khainovski, N.; Pillai, S.; Dey, S.; Daniels, S.; Purkayastha, S.; Juhasz, P.; Martin, S.; Bartlet-Jones, M.; He, F.; Jacobson, A.; Pappin, D. J., Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* **2004,** 3, 1154-1169.

28. Ong, S. E.; Blagoev, B.; Kratchmarova, I.; Kristensen, D. B.; Steen, H.; Pandey, A.; Mann, M., Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* **2002,** 1, 376-386.

29. Wiseman, R. L.; Chin, K. T.; Haynes, C. M.; Stanhill, A.; Xu, C. F.; Roguev, A.; Krogan, N. J.; Neubert, T. A.; Ron, D., Thioredoxin-related Protein 32 is an arsenite-regulated Thiol Reductase of the proteasome 19 S particle. *J. Biol. Chem.* **2009,** 284, 15233-45.

30. Van Harken, D. R.; Dixon, C. W.; Heimberg, M., Hepatic lipid metabolism in experimental diabetes. V. The effect of concentration of oleate on metabolism of triglycerides and on ketogenesis. *J. Biol. Chem.* **1969,** 244, 2278-85.

31. Romijn, E. P.; Christis, C.; Wieffer, M.; Gouw, J. W.; Fullaondo, A.; van der Sluijs, P.; Braakman, I.; Heck, A. J. R., Expression clustering reveals detailed coexpression patterns of functionally related proteins during B cell differentiation - A proteomic study using a combination of one-dimensional gel electrophoresis, LC-MS/MS, and stable isotope labeling by amino acids in cell culture (SILAC). *Mol. Cell. Proteomics* **2005,** 4, 1297-1310.

32. Uitto, P. M.; Lance, B. K.; Wood, G. R.; Sherman, J.; Baker, M. S.; Molloy, M. P., Comparing SILAC and two-dimensional gel electrophoresis image analysis for profiling urokinase plasminogen activator signaling in ovarian cancer cells. *J. Proteome Res.* **2007,** 6, 2105-2112.

33. Zhang, T. D.; Chen, G. Q.; Wang, Z. G.; Wang, Z. Y.; Chen, S. J.; Chen, Z., Arsenic trioxide, a therapeutic agent for APL. *Oncogene* **2001,** 20, 7146-53.

34. Douer, D.; Tallman, M. S., Arsenic trioxide: new clinical experience with an old medication in hematologic malignancies. *J. Clin. Oncol.* **2005,** 23, 2396-410.

35. Lennartsson, A.; Ekwall, K., Histone modification patterns and epigenetic codes. *Biochim. Biophys. Acta* **2009,** 1790, 863-8.

36. Zhou, X.; Sun, H.; Ellen, T. P.; Chen, H. B.; Costa, M., Arsenite alters global histone H3 methylation. *Carcinogenesis* **2008,** 29, 1831-1836.

37. Othumpangat, S.; Kashon, M.; Joseph, P., Sodium arsenite-induced inhibition of eukaryotic translation initiation factor 4E (eIF4E) results in cytotoxicity and cell death. *Mol. Cell. Biochem.* **2005,** 279, 123-131.

38. Wakil, S. J., Fatty-acid synthase, a proficient multifunctional enzyme. *Biochemistry* **1989,** 28, 4523-4530.

39. Kuhajda, F. P., Fatty acid synthase and cancer: New application of an old pathway. *Cancer Res.* **2006,** 66, 5977-5980.

40. Funabashi, H.; Kawaguchi, A.; Tomoda, H.; Omura, S.; Okuda, S.; Iwasaki, S., Binding site of cerulenin in fatty acid synthetase. *J. Biochem.* **1989,** 105, 751-5.

41. Omura, S., The antibiotic cerulenin, a novel tool for biochemistry as an inhibitor of fatty acid synthesis. *Bacteriol. Rev.* **1976,** 40, 681-97.

42. Pizer, E. S.; Wood, F. D.; Pasternack, G. R.; Kuhajda, F. P., Fatty acid synthase (FAS): A target for cytotoxic antimetabolites in HL60 promyelocytic leukemia cells. *Cancer Research* **1996,** 56, 745-751.

43. Furuta, E.; Pai, S. K.; Zhan, R.; Bandyopadhyay, S.; Watabe, M.; Mo, Y. Y.; Hirota, S.; Hosobe, S.; Tsukada, T.; Miura, K.; Kamada, S.; Saito, K.; Iiizumi, M.; Liu, W.; Ericsson, J.; Watabe, K., Fatty acid synthase gene is up-regulated by hypoxia via activation of Akt and sterol regulatory element binding protein-1. *Cancer Res.* **2008,** 68, 1003-11.

44. Yang, Y. A.; Han, W. F.; Morin, P. J.; Chrest, F. J.; Pizer, E. S., Activation of fatty acid synthesis during neoplastic transformation: role of mitogen-activated protein kinase and phosphatidylinositol 3-kinase. *Exp. Cell Res.* **2002,** 279, 80-90.

45. D'Erchia, A. M.; Tullo, A.; Lefkimmiatis, K.; Saccone, C.; Sbisa, E., The fatty acid synthase gene is a conserved p53 family target from worm to human. *Cell Cycle* **2006,** 5, 750-8.

46. Lane, A. A.; Ley, T. J., Neutrophil elastase cleaves PML-RAR alpha and is important for the development of acute promyelocytic leukemia in mice. *Cell* **2003,** 115, 305-318.

47. Kitareewan, S.; Roebuck, B. D.; Demidenko, E.; Sloboda, R. D.; Dmitrovsky, E., Lysosomes and trivalent arsenic treatment in acute promyelocytic leukemia. *J. Natl. Cancer Instit.* **2007,** 99, 41-52.

48. Benes, P.; Vetvicka, V.; Fusek, M., Cathepsin D--many functions of one aspartic protease. *Crit. Rev. Oncol. Hematol.* **2008,** 68, 12-28.

49. Jaattela, M.; Cande, C.; Kroemer, G., Lysosomes and mitochondria in the commitment to apoptosis: a potential role for cathepsin D and AIF. *Cell Death Differ.* **2004,** 11, 135-6.

50. Wu, G. S.; Saftig, P.; Peters, C.; El-Deiry, W. S., Potential role for cathepsin D in p53-dependent tumor suppression and chemosensitivity. *Oncogene* **1998,** 16, 2177-83.

51. Karapetian, R. N.; Evstafieva, A. G.; Abaeva, I. S.; Chichkova, N. V.; Filonov, G. S.; Rubtsov, Y. P.; Sukhacheva, E. A.; Melnikov, S. V.; Schneider, U.; Wanker, E. E.; Vartapetian, A. B., Nuclear oncoprotein prothymosin α is a partner of Keap1: implications for expression of oxidative stress-protecting genes. *Mol. Cell. Biol.* **2005,** 25, 1089-99.

52. Hayes, J. D.; McMahon, M., Molecular basis for the contribution of the antioxidant responsive element to cancer chemoprevention. *Cancer Lett.* **2001,** 174, 103-13.

53. Ishii, T.; Itoh, K.; Takahashi, S.; Sato, H.; Yanagawa, T.; Katoh, Y.; Bannai, S.; Yamamoto, M., Transcription factor Nrf2 coordinately regulates a group of oxidative stress-inducible genes in macrophages. *J. Biol. Chem.* **2000,** 275, 16023-9.

54. Snow, E. T.; Sykora, P.; Durham, T. R.; Klein, C. B., Arsenic, mode of action at biologically plausible low doses: what are the implications for low dose cancer risk? *Toxicol. Appl. Pharmacol.* **2005,** 207, 557-64.

55. Sul, H. S.; Wang, D., Nutritional and hormonal regulation of enzymes in fat synthesis: studies of fatty acid synthase and mitochondrial glycerol-3-phosphate acyltransferase gene transcription. *Annu. Rev. Nutr.* **1998,** 18, 331-51.

Supporting Information for Chapter 2

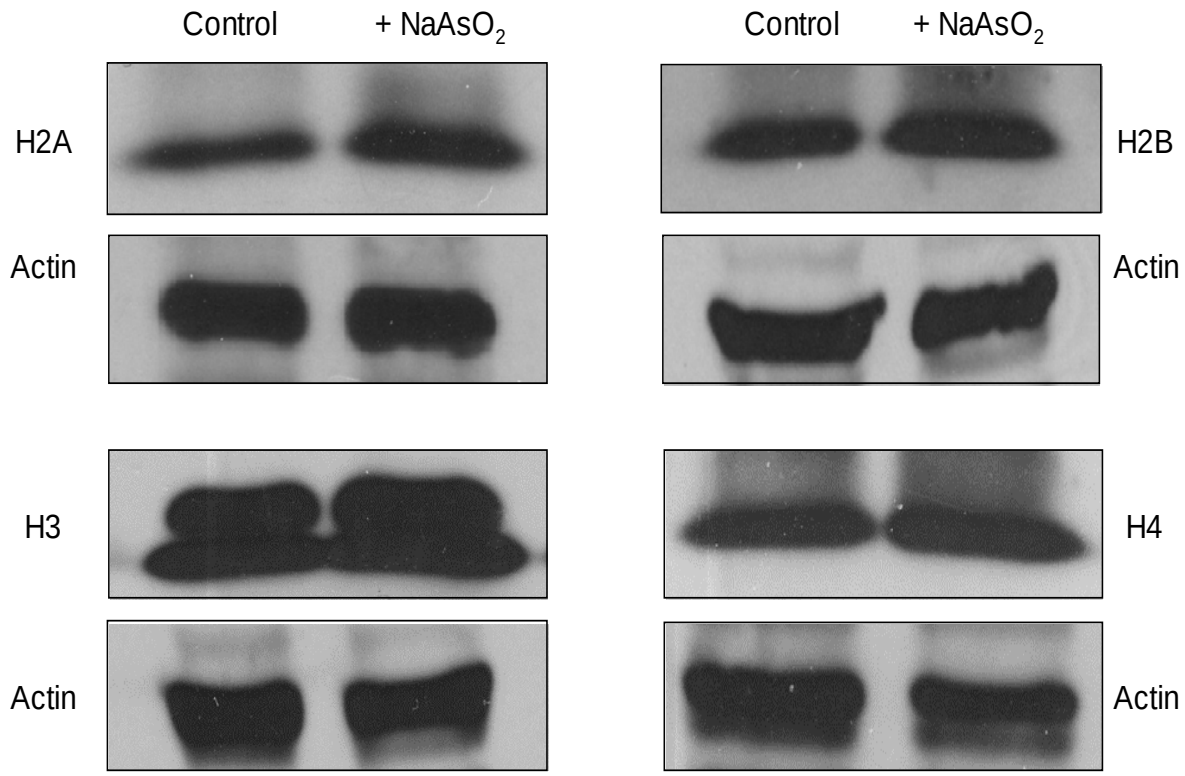


Figure S2.1 Western blotting analysis of histones H2A, H2B, H3, H4 with lysates of arsenite-treated ("+$NaAsO_2$") and untreated ("Control") HL-60 cells. Actin was used as the loading control. Quantification results (based on Western analysis) showed that histones H2A, H2B, H3 and H4 were upregulated by 1.56, 1.55, 1.52, and 1.44 folds, respectively.

Table S2.1 The detailed quantification results for proteins with substantial changes.

Protein Name	forward	forward	reverse	Pep. #	Ave.	S. D.
A. Histone and HMG proteins						
gi|11321591 high-mobility group box 2 K.YEKDIAAYR.A K.LKEKYEKDIAAYR.A	1.634	1.921	1.396	2	1.65	0.263
gi|968888 HMG-1 K.IKGEHPGLSIGDVAK.K K.LKEKYEKDIAAYR.A K.KHPDASVNFSEFSK.K	1.821	1.924	1.798	3	1.847	0.067
gi|223582 histone H4 R.DAVTYTEHAK.R R.ISGLIYEETR.G K.TVTAMDVVYALKR.Q K.VFLENVIRDAVTYTEHAK.R K.VFLENVIR.D	2.332	2.388	2.505	5	2.408	0.088
gi|386772 histone H3 K.STELLIR.K R.KLPFQR.L R.KLPFQR.L	2.64	2.479	3.279	3	2.799	0.423
gi|356168 histone H1b K.ATGPPVSELITK.A K.ALAAGGYDVEKNNSR. K.ALAAAGYDVEK.N -.SETAPAAPAAPAPAEKTPVKK. K K.GTLVQTK.G	2.929	2.963	2.381	5	2.758	0.327
gi|1568557 histone H2B R.LLLPGELAK.H R.EIQTAVR.L K.AMGIMNSFVNDIFER.I R.KESYSVYVYK.V K.QVHPDTGISSK.A	2.577	2.888	3.465	5	2.977	0.451

gi|510990 histone H2A K.VTIAQGGVLPNIQAVLLPK.K R.HLQLAIR.N R.HLQLAIRNDEELNK.L R.NDEELNKLLGK.V R.AGLQFPVGR.V	4.592	4.429	3.595	5	4.205	0.535
B. Translation related proteins						
gi|38202255 threonyl-tRNA synthetase K.WPFWLSPR.Q R.NELSGALTGLTR.V	0.552	0.728	0.512	2	0.597	0.115
gi|181969 elongation factor 2 R.ETVSEESNVLCLSK.S K.AYLPVNESFGFTADLR.S K.EGIPALDNFLDKL.-	0.447	0.545	0.636	3	0.543	0.094
gi|19353009 Similar to Elongation factor 2b R.VFSGLVSTGLK.V R.TFCQLILDPIFK.V	0.55	0.65	0.651	2	0.617	0.058
gi|4506707 ribosomal protein S25 R.AALQELLSK.G	0.552	0.615	0.771	1	0.646	0.113
C. hnRNP						
gi|386547 d(TTAGGG)n-binding protein B39 .IFVGGLSPDTPEEK.I	0.652	0.592	0.25	1	0.498	0.217
gi|119597533 heterogeneous nuclear ribonucleoprotein U, isoform CRA_b K.EKPYFPIPEEYTFIQNVPLED R.V	0.583	0.582	0.602	1	0.589	0.012
gi|55958547 heterogeneous nuclear ribonucleoprotein K R.TDYNASVSVPDSSGPER.I R.NTDEMVELR.I R.SRNTDEMVELR.I	0.773	0.724	0.489	3	0.662	0.152

D. Enzymes						
gi|531202 spermidine synthase R.KVLIIGGGDGGVLR.E K.VLIIGGGDGGVLR.E	0.466	0.465	0.773	2	0.568	0.178
gi|41584442 fatty acid synthase R.GTPLISPLIK.W R.GYAVLGGER.G K.TGTVSLEVR.V R.LQVVDQPLPVR.G R.DNLEFFLAGIGR.L	0.704	0.612	0.512	5	0.609	0.096
gi|66392203 NME1-NME2 protein R.GLVGEIIKR.F K.DRPFFAGLVK.Y K.EHYVDLKDRPFFAGLVK.Y R.TFIAIKPDGVQR.G K.FMQASEDLLK.E	0.704	0.441	0.78	5	0.642	0.178
gi|1230564 Gu protein R.APQVLVLAPTR.E K.TFSFAIPLIEK.L	0.743	0.693	0.489	2	0.642	0.135
gi|2661039 alpha enolase K.EGLELLK.T R.IGAEVYHNLK.N R.YISPDQLADLYK.S R.AAVPSGASTGIYEALELR.D K.LAMQEFMILPVGAANFR.E	0.564	0.689	0.691	5	0.648	0.073
gi|190281 protein phosphatase I alpha subunit R.EIFLSQPILLELEAPLK.I	0.683	0.657	0.643	1	0.661	0.02
gi|4507789 ubiquitin-conjugating enzyme E2L 3 isoform 1 K.ELEEIR.K R.ADLAEEYSK.D R.IEINFPAEYPFKPPK.I K.GQVCLPVISAENWKPATK R.NIQVDEANLLTWQGLIVPDN PPYDK.G	0.614	0.657	0.734	5	0.668	0.061

gi|5231228 ribonuclease T2 precursor R.ELDLNSVLLK.L HGTCAAQVDALNSQK.K K.LQIKPSINYYQVADFK.D	1.314	1.567	1.72	3	1.533	0.205
gi|4758504 hydroxysteroid (17-beta) dehydrogenase 10 isoform 1 R.LVGQGASAVLLDLPNSGGEAQAK.K	1.541		1.574	1	1.557	0.023
gi|4503143 cathepsin D R.VGFAEAAR.L R.VGFAEAARL.-	1.574	1.734	1.655	2	1.654	0.08
gi|186461558 neutrophil elastase R.QVFAVQR.I R.VVLGAHNLSR.R	1.935	1.951	1.957	2	1.948	0.011
gi|627372 alpha-mannosidase precursor R.HLVLLDTAQAAAAGHR.L	2.291	2.056	1.811	1	2.053	0.24
gi|4504437 heme oxygenase (decyclizing) 1 R.TEPELLVAHAYTR.Y R.YLGDLSGGQVLK.K	3.203	1.713	2.054	2	2.323	0.781
gi|12654715 TXNDC5 protein R.GYPTLLLFR.G R.GYPTLLWFR.D R.DLESLREYVESQLQR.T K.ALAPTWEQLALGLEHSETVK.I	3.327	3.353	2.732	4	3.138	0.351
E. Others						
gi|21755073 unnamed protein product K.MDLNSEQAEQLER.I	0.442	0.416	0.403	1	0.42	0.02
gi|119587276 hCG19802, isoform CRA_a	0.621	0.551	0.243	1	0.472	0.201

M.AEVQVLVLDGR.G						
gi|35570 unnamed protein product K.NLDDGIDDER.L R.IVATKPLYVALAQR.K	0.552	0.521	0.411	2	0.495	0.074
gi|13477237 ZNF607 protein R.FDAGELITQR.E K.SLVARFDAGELITQR.E	0.555		0.486	2	0.52	0.049
gi|56462484 MHC class I antigen R.AYLEGLCVEWLR.R R.FIAVGYVDDTQFVR.F R.APWIEQEGPEYWDGETR.N R.DGEDQTQDTELVETRPAGDR .T	0.703	0.56	0.453	4	0.572	0.125
gi|194374111 unnamed protein product M.YGCDVGPDGR.L K.THVTHHPISDHEATLR.C	0.663	0.603	0.456	2	0.573	0.107
gi|35844 unnamed protein product K.ELSLAGNELGDEGAR.L	0.442	0.551	0.739	1	0.577	0.151
gi|114645930 PREDICTED: nucleosome assembly protein 1-like 1 isoform 9 K.GIPEFWLTVFK.N K.FYEEVHDLER.K R.LDGLVETPTGYIESLPR.V	0.583	0.615	0.599	3	0.599	0.016
gi|431422 Ran/TC4 Binding Protein R.FLNAENAQK.F K.TLEEDEEELFK.M	0.673	0.769	0.358	2	0.6	0.215
gi|2580550 dead box, X isoform K.DLLDLLVEAK.Q R.SFLLDLLNATGK.D K.TAAFLLPILSQIYSDGPGEAL R.A	0.602	0.71	0.512	3	0.608	0.099
gi|122168 HLA class I	0.731	0.724	0.375	2	0.61	0.204

histocompatibility antigen, B-58 α chain R.DGEDQTQDTELVETRPAGDR.T R.VAEQLR.A						
gi|181486 DNA-binding protein B R.EDGNEEDKENQGDETQGQQPPQR.R	0.592	0.716	0.587	1	0.632	0.073
gi|193788267 unnamed protein product R.AIRLELQGPR.G K.NLPYKVTQDELKEVFEDAAEIR.L K.EVFEDAAEIR.L	0.681	0.658	0.538	3	0.626	0.077
gi|801893 leucine-rich PPR-motif containing protein K.TVLDQQQTPSR.L K.VIEPQYFGLAYLFR.K	0.547	0.595	0.742	2	0.628	0.101
gi|5107666 importin β R.VAALQNLVK.I R.VLANPGNSQVAR.V K.SNEILTAIIQGMR.K K.LAATNALLNSLEFTK.A R.AAVENLPTFLVELSR.V	0.59	0.453	0.849	5	0.631	0.201
gi|1235727 unnamed protein product K.EPLVDVVDPK.Q R.ATLYVTAIEDR.Q	0.562		0.71	2	0.636	0.104
gi|40225729 FUBP1 protein R.ITGDPYKVQQAK.E R.LLDQIVEK.G K.IQIAPDSGGLPER.S	0.706		0.589	1	0.647	0.083
gi|119571409 hCG1643342, isoform CRA_a K.LLEPVLLLGKER	0.682	0.635	0.627	1	0.648	0.03

gi|386777 transplantation antigen R.FDSDAASPR.G K.THVTHHPLSDHEATLR.C	0.716	0.636	0.618	2	0.657	0.052
gi|1136741 KIAA0002 K.AIADTGANVVVTGGK.V K.FAEAFEAIPR.A K.TVGATALPR.L	0.731	0.812	0.438	3	0.661	0.197
gi|23712 myoblast antigen 24.1D5 R.AVVIVDDR.G R.FATHAALSVR.N R.NLSPYVSNELLEEAFSQFGPI ER.A	0.574	0.704	0.71	3	0.663	0.077
gi|13569879 acidic (leucine-rich) nuclear phosphoprotein 32 family, member E isoform 1 K.SLDLFNCEITNLEDYR.E K.DLSTVEALQNLK.N K.IKDLSTVEALQNLK.N	0.616	0.753	0.629	3	0.666	0.076
gi|4506773 S100 calcium-binding protein A9 K.ELPGFLQSGK.D	1.506	1.571	1.499	1	1.525	0.040
gi|9955206 Rho GDP-dissociation factor 2 K.TLLGDGPVVTDPK.A K.APNVVVTR.L R.LTLVCESAPGPITMDLTGDLE ALKK.E	1.417	1.46	1.773	3	1.55	0.194
gi|8037945 prothymosin alpha K.EVVEEAENGR.D	1.541		1.574	1	1.557	0.023
gi|119239 Bone marrow proteoglycan R.GNLVSIHNFNINYR.I	1.626	1.785	1.733	1	1.714	0.081
gi|28375485 unnamed protein product K.LKPLEVELR.R	1.917	1.952	1.548	1	1.806	0.224

gi|4506191 proteasome beta 10 subunit proprotein R.LPFTALGSGQDAALAVLEDR.F	2.15	2.145	2.323	1	2.206	0.101
gi|32111 unnamed protein product R.AGLQFPVGR.V	5.999	4.425	4.272	1	4.899	0.956

Table S2.2 The detailed quantification results for histone proteins.

Experiments	Histones	Peptide Ratio	Protein Ratio	S. D.
	gi|223582 histone H4			
Set 1	K.VFLENVIR.D	2.28	2.33	0.16
	R.ISGLIYEETR.G	2.51		
	K.TVTAMDVVYALK.R	2.20		
Set 2	R.DAVTYTEHAK.R	2.40	2.39	0.26
	R.ISGLIYEETR.G	2.18		
	K.TVTAMDVVYALKR.Q	2.40		
	K.VFLENVIRDAVTYTEHAK.R	2.83		
	K.VFLENVIR.D	2.09		
	R.DNIQGITKPAIR.R	2.42		
Set 3	K.VFLENVIR.D	2.36	2.51	0.19
	R.ISGLIYEETR.	2.24		
	K.VFLENVIRDAVTYTEHAK.R	2.66		
	K.TVTAMDVVYALKR.Q	2.58		
	K.VFLENVIR.D	2.68		
	gi|510990 histone H2A			
Set 1	R.HLQLAIR.N	2.36	4.59	1.96
	R.VTIAQGGVLPNIQAVLLPK.K	5.43		
	R.AGLQFPVGR.V	5.99		
Set 2	K.VTIAQGGVLPNIQAVLLPK.K	5.86	4.43	1.86
	R.HLQLAIR.N	3.84		
	R.HLQLAIRNDEELNK.L	3.74		
	R.NDEELNKLLGK.V	2.01		
	R.AGLQFPVGR.V	6.70		
Set 3	R.AGLQFPVGR.V	3.56	3.60	0.38
	R.AGLQFPVGR.V	4.06		
	R.HLQLAIR.N	3.90		
	R.HLQLAIRNDEELNK.L	3.16		
	R.HLQLAIRNDEELNKLLGK.V	3.30		
	gi|1568551 histone H2B			
Set 1	K.ESYSIYVYK.V	2.62	2.58	0.04
	R.LLLPGELAK.H	2.55		

	K.QVHPDTGISSK.A	2.56		
Set 2	R.LLLPGELAK.H	2.86	2.89	0.24
	R.EIQTAVR.L	2.67		
	K.AMGIMNSFVNDIFER.I	2.52		
	R.KESYSVYVYK.V	3.21		
	K.QVHPDTGISSK.A	3.00		
	K.ESYSVYVYK.V	2.85		
	K.AMGIMNSFVNDIFER.I	2.71		
	R.KESYSVYVYK.V	3.21		
	K.ESYSIYVYK.V	2.96		
Set 3	K.AMGIMNSFVNDIFER.I	4.21	3.47	0.90
	K.AMGIMNSFVNDIFERIAGEASR.L	4.50		
	K.AMGIMNSFVNDIFER.I	2.92		
	R.EIQTAVR.L	3.39		
	R.KESYSVYVYK.V	2.31		
	gi|386772 histone H3			
Set 1	K.STELLIR.K	2.50	2.64	0.20
	K.DIQLAR.H	2.78		
Set 2	K.STELLIR.K	2.52	2.48	0.06
	R.KLPFQR.L	2.44		
Set 3	K.STELLIR.K	3.06	3.28	0.35
	K.DIQLAR.H	3.68		
	R.YRPGTVALR.E	3.10		
	gi|356168 histone H1b			
Set 1	R.SGVSLAALK.K	2.45	2.93	0.84
	K.ATGPPVSELITK.A	2.439		
	K.ALAAAGYDVEK.N	3.899		
Set 2	K.ATGPPVSELITK.A	2.342	2.96	0.94
	K.ALAAGGYDVEKNNSR.	2.566		
	K.ALAAAGYDVEK.N	1.974		
	.SETAPAAPAAPAPAEKTPVKK.K	3.959		
	K.GTLVQTK.G	3.975		
Set 3	K.ALAAAGYDVEK.N	2.281	2.38	0.14
	K.ATGPPVSELITK.A	2.481		

Table S2.3 Quantification results for hnRNPs, translation initiation factors, elongation factors and ribosomal proteins.

hnRNPs

GI Number	forward(1)	forward(2)	reverse(3)	Pep. #	Average	S. D.
gi|386547	0.652	0.592	0.25	1	0.498	0.217
gi|133254	1.024	0.962	1.208	5	1.065	0.128
gi|58761496	0.875		0.891	3	0.883	0.011
gi|14110407	0.771	0.963	1.009	3	0.914	0.126
gi|4826760	0.805	0.855	0.992	5	0.884	0.097
gi|542850	1.116	0.982	0.93	5	1.009	0.096
gi|5031753	0.752	0.855	0.709	3	0.772	0.075
gi|55958547	0.773	0.724	0.489	3	0.662	0.152
gi|55958543	0.935	1.01	1.207	3	1.051	0.140
gi|5031755	0.861	0.973	0.764	3	0.866	0.105
gi|119597533	0.583	0.582	0.602	1	0.589	0.012
gi|109082737	0.812	1.061	1.182	5	1.018	0.189

GI Number	forward(1)	forward(2)	reverse(3)	Pep. #	Average	S. D.
gi|119610575	0.646	0.693	0.697	3	0.679	0.028
gi|33356163	0.745	0.59	0.762	3	0.699	0.095
gi|485388	0.646	0.718	0.775	2	0.713	0.065
gi|4503529	0.859	0.715	0.796	5	0.79	0.072
gi|4261795		1	0.708	1	0.854	0.206
gi|4503545	0.757	0.821	0.889	2	0.822	0.066
gi|4503519	1.078	1.186	0.783	1	1.016	0.209
gi|7705433	0.895		1.034		0.965	0.099

GI Number	forward(1)	forward(2)	reverse(3)	Pep. #	Average	S. D.
gi|181969	0.447	0.545	0.636	3	0.54	0.09
gi|19353009	0.55	0.65	0.651	2	0.62	0.06
gi|4503475	0.684	0.652	0.726	1	0.69	0.04
gi|7108915	0.685	0.677	0.703	3	0.69	0.01
gi|4503471	0.684	0.652	0.841	2	0.73	0.10
gi|62897525	0.789	0.722	0.68	3	0.73	0.06
gi|4503481	0.729	0.637	0.834	2	0.73	0.10
gi|119602640	0.729	0.743	0.874	5	0.78	0.08
gi|194239729	0.797	0.764	0.907	5	0.82	0.08
gi|927065	0.789	0.827	0.909	3	0.84	0.06
gi|170785039	0.759	0.954	0.839	3	0.85	0.10
gi|4503483	0.862	0.811	0.956	5	0.88	0.07
gi|181965		0.897	0.731	1	0.81	0.12

gi|119594429	0.977	0.998	0.956	3	0.98	0.02
gi|1220311	0.911	0.829	1.209	3	0.98	0.20
gi|38522	0.911	0.868	1.209	5	1.00	0.19
gi|704416	1.149	1.043	0.98	5	1.06	0.09
gi|25453472	1.105	1.243	1.013	5	1.12	0.12

GI Number	**forward(1)**	**forward(2)**	**reverse(3)**	**Pep. #**	**Average**	**S. D.**
gi|4506707	0.552	0.615	0.771	1	0.646	0.113
gi|4506669	0.625	0.706	0.729	2	0.687	0.055
gi|4506661	0.702	0.649	0.734	2	0.695	0.043
gi|550021	0.458	0.815	0.829	4	0.701	0.210
gi|7513316	0.758	0.652	0.726	5	0.712	0.054
gi|119587829		0.634	0.835	3	0.735	0.142
gi|4506685	0.789	0.586	0.845	2	0.740	0.136
gi|74723863	0.76	0.923	0.558	3	0.747	0.183
gi|5032051	0.793	0.763	0.769	3	0.775	0.016
gi|36138	0.816	0.763	0.766	2	0.782	0.030
gi|15431301	0.767	0.782	0.801	4	0.783	0.017
gi|89573879	0.726	0.942	0.719	1	0.796	0.127
gi|119572747	0.726	0.942	0.719	1	0.796	0.127
gi|14591909	0.843	0.744	0.826	3	0.804	0.053
gi|4506667	0.76	0.923	0.742	5	0.808	0.100
gi|4506697	0.81	0.93	0.724	5	0.821	0.103
gi|7705813	0.774	0.826	0.886	5	0.829	0.056
gi|14165469	0.763	0.795	0.928	5	0.829	0.088
gi|6912634	0.82	1.048	0.621	2	0.830	0.214
gi|113417504	0.777	0.804	0.913	3	0.831	0.072
gi|73959145	0.987	0.791	0.719	1	0.832	0.139
gi|119615473	0.875	0.847	0.783	3	0.835	0.047
gi|119605048	0.774	0.762	0.97	3	0.835	0.117
gi|4506597	0.777	0.819	0.913	5	0.836	0.070
gi|292435	0.756	0.826	0.941	5	0.841	0.093
gi|113429348	0.801	0.911	0.812	2	0.841	0.061
gi|4432750	0.866	0.833	0.836	1	0.845	0.018
gi|7765076	0.971		0.726	5	0.849	0.173
gi|114606879	0.866	0.905	0.813	3	0.861	0.046
gi|4506613	0.826	0.75	1.009	2	0.862	0.133
gi|495126	0.766	0.975	0.845	3	0.862	0.106
gi|6677809	0.921	0.879	0.804	3	0.868	0.059
gi|4506723		0.953	0.792	5	0.873	0.114
gi|119592221		0.863	0.884	3	0.874	0.015
gi|4506607	0.921	0.937	0.799	2	0.886	0.075

gi|4506715	0.717	0.928	1.015	3	0.887	0.153
gi|36142	0.766	0.918	0.982	1	0.889	0.111
gi|14141193	0.833	0.922	0.916	5	0.890	0.050
gi|553640	1.073	0.723	0.904	1	0.900	0.175
gi|4506691	0.704	1.115	0.885	5	0.901	0.206
gi|4506679	0.992	0.849	0.873	3	0.905	0.077
gi|4506703	0.804	1.019	0.891	3	0.905	0.108
gi|119575011	0.804	1.019	0.891	3	0.905	0.108
gi|21104402	0.914	0.905	0.906	5	0.908	0.005
gi|4506743	0.91	0.922	0.905	5	0.912	0.009
gi|119572749	1.035	0.924	0.785	5	0.915	0.125
gi|23491733	0.865	0.901	1.021	5	0.929	0.082
gi|4506695	0.972		0.897	3	0.935	0.053
gi|15431303	0.84	0.942	1.051	4	0.944	0.106
gi|62896495	1.094	0.762	1.085	3	0.980	0.189
gi|4506635	1.047	0.904	0.996	1	0.982	0.072
gi|4506681	1.169	0.912	0.931	5	1.004	0.143
gi|13366090	1.054	1.142	0.849	2	1.015	0.150
gi|73948356	0.964	1.038	1.075	5	1.026	0.057
gi|1655596	1.078	1.037	1.012	3	1.042	0.033
gi|36146	1.216	1.037	0.922	5	1.058	0.148
gi|414587	1.008	1.062	1.179	5	1.083	0.087
gi|4506609	1.058	1.063	1.306	5	1.142	0.142
gi|4432754		1.296	1.153	1	1.225	0.101
gi|30047125	1.186	1.304	1.245	5	1.245	0.059
						Average Ratio
						0.878

CHAPTER 3

Global Proteome Quantification for Discovering Imatinib-induced Perturbation of Multiple Biological Pathways in K562 Human Chronic Myeloid Leukemia Cells

Introduction

Chronic myeloid leukemia (CML) is a hematopoietic stem cell disorder, which arises from a translocation between the long arms of chromosomes 9 and 22 [1, 2]. The translocation renders the *Bcr-Abl* fusion gene where the *Bcr* gene on chromosome 22 is linked with the proto-oncogene *Abl* on chromosome 9 [2]. The fusion gene encodes the Bcr-Abl tyrosine kinase, which is constitutively active and leads to uncontrolled growth [2]. The Bcr-Abl kinase activates several signaling pathways such as the Ras/mitogen-activated protein kinase, signal transducer and activator of transcription 5, and phosphatidylinositol 3 kinase/Akt pathways; enhances nuclear factor κB (NF-κB) activity; up-regulates the level of Bcl-X_L; and suppresses the mitochondrial pathway of apoptosis [3].

Imatinib mesylate, which is marketed by Novartis as Gleevec in the US, has emerged as the leading compound to treat patients with CML [2]. As a selective tyrosine kinase inhibitor, imatinib associates directly with the ATP-binding site and inhibits the kinase activity of Bcr-Abl. Upon imatinib treatment, the Bcr-Abl protein is rapidly dephosphorylated and becomes inactive, thereby interrupting the constitutive activation of signaling cascades, arresting cell cycle progression, and

triggering apoptosis [4] . Despite demonstrating remarkable clinical efficacy against chronic-phase CML, the outcome after imatinib therapy in the accelerated and blastic phases of CML is unacceptably poor [5], mostly owing to the emergence of mutations in the Bcr-Abl kinase domain that may inhibit binding of imatinib to the kinase domain. Thus, the discovery of novel targets of imatinib could contribute significantly to our understanding of the mechanisms of the anti-cancer functions of the drug and the development of resistance to imatinib among CML patients.

There have been a few studies on the imatinib-induced perturbation in global protein expression [6-8], in which 2-dimensional gel electrophoresis (2-DE) coupled with tandem mass spectrometry (MS/MS) was employed for protein identification and quantification. Other than 2-DE, several stable isotope-labeling strategies [9], especially stable isotope labeling by amino acids in cell culture (SILAC) [10], have been developed for MS-based differential protein expression analysis. SILAC is more efficient than 2-DE in the quantification of the whole proteome and in the detection of relatively small changes in protein abundance. In this context, Liang et al. [11] used SILAC together with LC-MS/MS and examined the imatinib-induced alterations of the Bcr-Abl kinase in CML cells.

In the present study, we employed LC-MS/MS, along with SILAC, to assess quantitatively the imatinib-induced alteration in protein expression in the *Bcr-Abl*-positive human K562 cells. We were able to quantify a total of 1344 proteins, among which 73 were significantly changed upon imatinib treatment in both forward

and reverse SILAC measurements. The identification of proteins perturbed by imatinib treatment sets a stage for understanding the biological pathways affected by imatinib.

Experimental

Materials

Heavy lysine and arginine ([$^{13}C_6$,$^{15}N_2$]-L-lysine and [$^{13}C_6$,$^{15}N_4$]-L-arginine) were purchased from Cambridge Isotope Laboratories (Andover, MA). All reagents unless otherwise noted were from Sigma (St. Louis, MO).

Cell Culture

K562 cells (ATCC, Manassas, VA) were cultured in Iscove's modified minimal essential medium (IMEM) supplemented with 10% fetal bovine serum (FBS, Invitrogen, Carlsbad, CA) and penicillin (100 IU/mL). Cells were maintained in a humidified atmosphere with 5% CO_2 at 37ºC, with medium renewal at every 2-3 days. For SILAC experiments, the IMEM medium without L-lysine or L-arginine was custom-prepared following the ATCC formulation. The complete light and heavy IMEM media were prepared by the addition of light or heavy lysine and arginine, together with dialyzed FBS, to the above lysine, arginine-depleted medium. The K562 cells were cultured in heavy IMEM medium for at least 5 cell doublings to achieve complete isotope incorporation.

Imatinib Treatment and Cell Lysate Preparation

K562 cells, at a density of approximately 7.5×10^5 cells/mL, were collected by centrifugation at 300 g and at 4ºC for 5 min, washed twice with ice-cold phosphate-buffered saline (PBS) to remove FBS, and resuspended in the fresh serum-free heavy or light media. In forward SILAC experiment, the cells cultured in light medium were treated with 1 μM imatinib for 24 hrs and the cells cultured in heavy medium were untreated. Reverse SILAC experiments were also performed where the cells cultured in the heavy and light media were treated with imatinib and mock-treated, respectively (Figure 3.1). After 24 hrs, the light and heavy cells were collected by centrifugation at 300 g, and washed three times with ice-cold PBS.

The cell pellets were resuspended in the CelLytic™ M cell lysis buffer for 30 min with occasional vortexing. Cell lysates were centrifuged at 12,000 g at 4°C for 30 min, and the resulting supernatants were collected. To the supernatant was subsequently added a protease inhibitor cocktail, and the protein concentrations of the cell lysates were determined by using Quick Start Bradford Protein Assay kit (Bio-Rad, Hercules, CA).

SDS-PAGE Separation and In-gel Digestion

The light and heavy cell lysates were combined at 1:1 ratio (w/w), denatured by boiling in Laemmli loading buffer for 5 min and separated by a 12% SDS-PAGE

with 4% stacking gel. The gel was stained with Coomassie blue; after destaining, the gel was cut into 20 slices, reduced in-gel with dithiothreitol (DTT) and alkylated with iodoacetamide (IAM). The proteins were digested in-gel with trypsin (Promega, Madison, WI) for overnight, after which peptides were extracted from the gels with 5% acetic acid in H_2O and subsequently with 5% acetic acid in CH_3CN/H_2O (1:1, v/v). The resultant peptide mixtures were dried and stored at –20°C for further analysis.

Benzidine Staining

The untreated and imatinib-treated K562 cells were collected without removing the media and mixed, at 1:1 (v/v) ratio for 4 min, with an aqueous solution containing 0.2% benzidine dihydrochloride, 0.6% H_2O_2 and 0.5 M acetic acid. The cells were spotted to a hemocytometer and pictures were taken using a Nikon Eclipse TI microscope (Melville, NY).

Histone Extraction

Core histones were obtained following previously reported procedures [12]. In brief, the untreated and imatinib-treated 562 cells were harvested by centrifugation at 500 g. The cell pellets were subsequently washed with a 5-mL lysis buffer containing 0.25 M sucrose, 0.01 M $MgCl_2$, 0.5 mM PMSF, 50 mM Tris (pH 7.4), and 0.5% Triton X-100, then resuspended in 5 mL of the same buffer and kept at 4ºC overnight. The histones were extracted from the cell lysate with 0.4 M sulfuric acid by

incubating at 4°C for 4 hr with continuous vortexing, precipitated with cold acetone, centrifuged, dried and redissolved in water.

HPLC Purification and Protease Digestion

Core histones were purified by HPLC on an Agilent 1100 system (Agilent Technologies, Santa Clara, CA) as described previously [12]. A 4.6×250 mm C4 column (Grace Vydac, Hesperia, CA) was used. The wavelength for the UV detector was set at 220 nm. The flow rate was 0.8 mL/min, and a 60-min linear gradient of 30-60% acetonitrile in 0.1% trifluoroacetic acid (TFA) was employed.

Histones H4 and H3 were digested with Asp-N and Arg-C in buffers containing 50 mM sodium phosphate (pH 8.0) and 100 mM NH_4HCO_3 (pH 8.0), respectively. The digestion was carried out with a protein/enzyme ratio of 20:1 (w/w) at 37ºC overnight. The peptide mixtures were subjected directly to LC-MS/MS analysis, or to a further peptide fractionation by HPLC and then analyzed by MALDI-MS (See below).

Mass Spectrometry

MALDI mass spectra were acquired on a Voyager DE STR MALDI-TOF mass spectrometer (Applied Biosystems, Framingham, MA) in positive reflection mode. The mass spectrometer was equipped with a pulsed nitrogen laser operated at 337 nm with 3 ns duration pulses. The acceleration voltage, grid voltage, and delayed

extraction time were set as 20 kV, 65%, and 190 ns, respectively. Each mass spectrum was acquired from an average of 100 laser shots.

Online LC-MS/MS analysis was carried out on an Agilent 6510 Q-TOF system with an Agilent HPLC-Chip Cube MS interface (Agilent Technologies, Santa Clara, CA). The sample injection, enrichment, desalting, and HPLC separation were carried out automatically on the Agilent HPLC Chip with an integrated trapping column (160 nL) and a separation column (Zorbax 300SB-C18, 75 μm×150 mm, 5 μm in particle size). The peptide or protein sample was first loaded onto the trapping column with a solvent mixture of 0.1% formic acid in CH_3CN/H_2O (2:98, v/v) at a flow rate of 4 μL/min, which was delivered by an Agilent 1200 capillary pump. The peptides were then separated using a 120-min linear gradient of 2-35% acetonitrile, while intact histones were eluted with a 30-min linear gradient of 40-70% acetonitrile in 0.1% formic acid and at a flow rate of 300 nL/min, which was delivered by an Agilent 1200 Nano pump.

The Chip spray voltage (VCap) was set at 1950 V and varied with chip conditions. The temperature and flow rate of the drying gas were 325ºC and 4 L/min, respectively. Nitrogen was employed as the collision gas, and the collision energy followed an equation with a slope of 3 V/100 Da and an offset of 2.5 V. MS/MS experiments were carried out in the data-dependent scan mode where a maximum of five MS/MS scans were acquired following each MS scan. The m/z ranges for MS and MS/MS were 300-2000 and 60-2000, and the data acquisition rates were 6 and 3

spectra/s, respectively.

Data Processing

The LC-MS/MS raw data were searched against human IPI protein database (version 3.21) and its reverse complement using TurboSEQUEST with Bioworks 3.2 (Thermo Fisher Scientific, San Jose, CA) for protein identification. Cysteine carbamidomethylation was set as a fixed modification. Methionine oxidation (+16 Da) as well as lysine (+8 Da) and arginine (+10 Da) mass shifts introduced by heavy isotope labeling were considered as variable modifications. Peptide filters with appropriate cross-correlation (Xcorr $\geq 1.9, \geq 2.4, \geq 3.5$ for peptide ions that are singly, doubly, and triply charged) and delta correlation ($\Delta C_n \geq 0.1$) scores were used to sort the search results. The protein false discovery rate was less than 1%.

Census, developed by Yates and coworkers [13, 14], was employed for protein quantification. The TurboSEQUEST search results were first filtered using DTASelect [15] and ion chromatograms were generated for peptide ions based on their *m/z* values. Peptide ion intensity ratios were subsequently calculated in Census from peak areas found in each pair of extracted-ion chromatograms. The ratio measurement results were filtered by setting thresholds of Determinant Factor as 0.5 and Outlier p-Value as 0.01.

The ratio obtained for each individual protein was then normalized against the average ratio for all quantified proteins. In this "multi-point" normalization strategy, it

was assumed that the ratios for the majority of proteins were not perturbed by imatinib treatment, facilitating the use of the average ratio of all quantified proteins to re-scale the data. This has been widely employed to remove the inaccuracy during sample mixing introduced by protein quantification with the Bradford assay [16, 17]. Some peptides identified by TurboSEQUEST for only 1 or 2 sets of SILAC samples could also be manually found, in the LC-MS/MS data for the remaining set(s) of SILAC samples, and quantified. In this context, the Chip HPLC provided excellent reproducibility in retention time, and most of the time the difference in elution time for a peptide among different runs was within 2 min, though occasionally the difference could be up to 5 min. In addition, the mass accuracy afforded by the Q-TOF mass spectrometer is within 20 ppm with external calibration. Therefore, the accurate *m/z* values of peptide ions (within 20 ppm) and HPLC retention time (within 5 min variation) were employed as criteria to locate the light/heavy peptide pairs for the quantification. Only those proteins with fold changes >1.5 and quantified in at least 2 sets (including both forward and reverse) of SILAC measurements were reported as significantly changed proteins.

Results

Imatinib Treatment, Protein Identification and Quantification

To gain insights into the molecular pathways perturbed by imatinib treatment, we used SILAC in combination with LC-MS/MS to assess the imatinib-induced

differential expression of the whole proteome of K562 cells. The standard dose of imatinib used in the clinical treatment of CML is 400 mg/day, which gives plasma concentrations at baseline and 3 hr after administration at 0.5-4 and 1.5-8 μM, respectively [18, 19]. We observed, based on trypan blue exclusion assay, a less than 5% cell death after a 24-hr treatment of K562 cells with 1 μM imatinib, whereas greater than 20% cells were dead if treated with 1.5 μM imatinib for 24 hr. Thus, we chose 1 μM imatinib in subsequent experiments to minimize the apoptosis-induced alteration in protein expression.

To obtain reliable results, we carried out the SILAC experiments in triplicate and both forward and reverse SILAC labeling was performed (Figure 3.1A, see also Materials and Methods). Figure 3.2 shows representative mass spectrometric results for the identification and quantification of the peptide MNVEEAGGEALGR from hemoglobin ε subunit, which reveals clearly the up-regulation of this protein in both forward and reverse SILAC experiments (Figure 3.2A&B). The peptide sequence was confirmed by MS/MS analysis (Figure 3.2C&D).

We were able to quantify a total of 1344 proteins (Table S3.1), and the distribution of changes in protein expression levels induced by imatinib treatment is shown in Figure 3.1B. The average relative standard deviation in the expression ratios determined for all quantified proteins based on different tryptic peptides of the protein

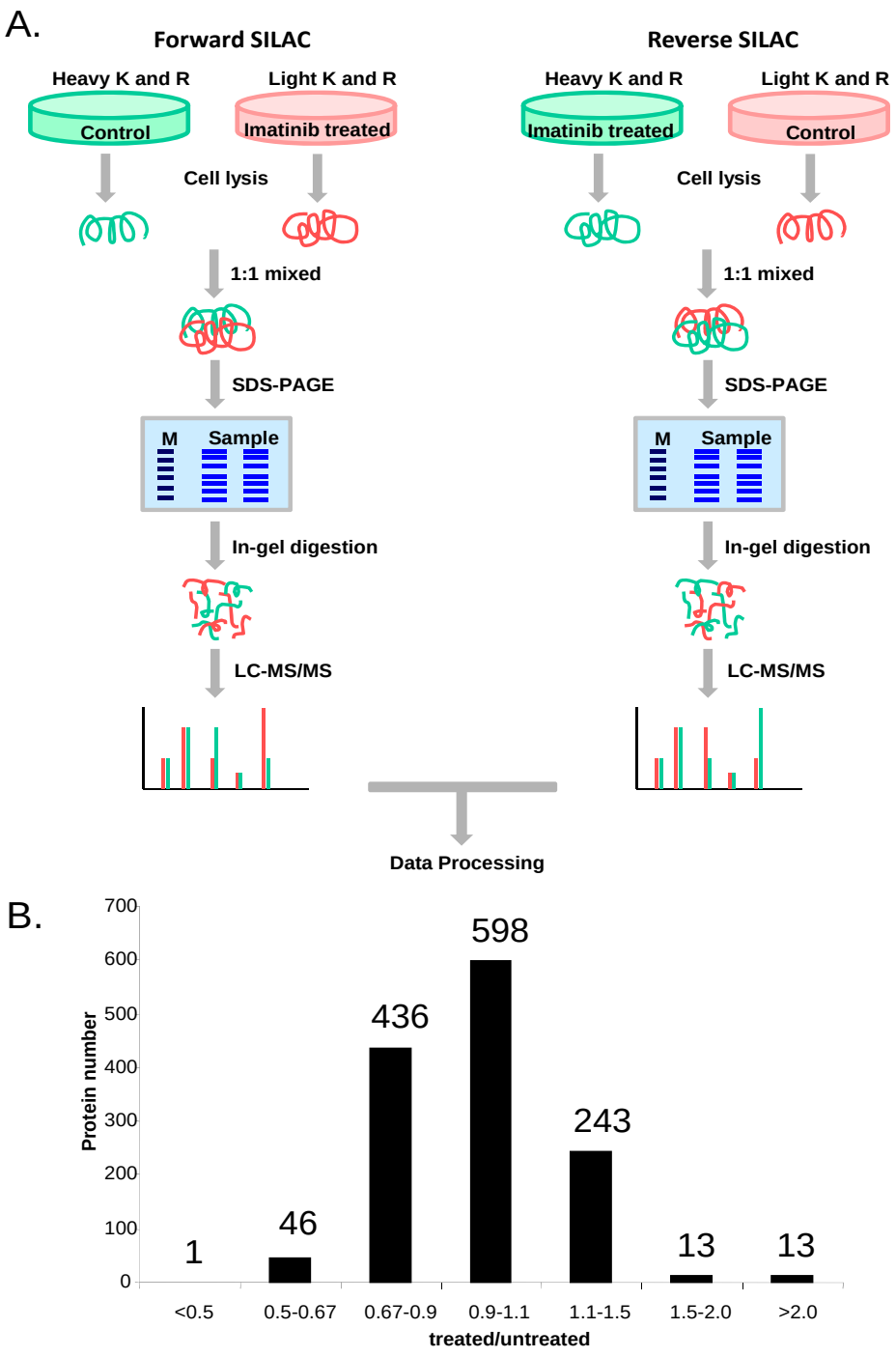


Figure 3.1 (A) Forward- and reverse-SILAC combined with LC-MS/MS for the comparative analysis of protein expression in K562 cells upon imatinib treatment. (B) The distribution of expression ratios (treated/untreated) for the quantified proteins.

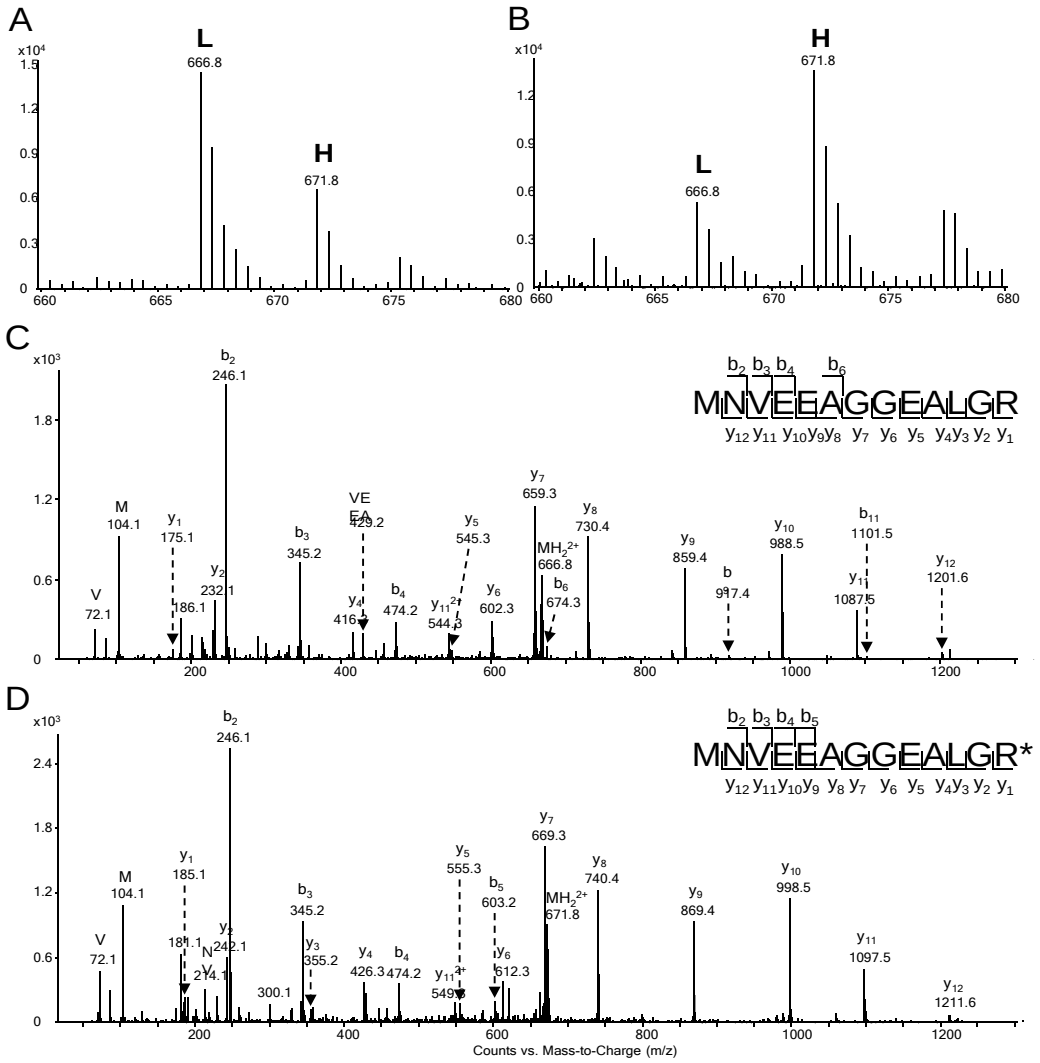


Figure 3.2 Representative ESI-MS and MS/MS data revealing the imatinib-induced up-regulation of hemoglobin ε subunit. Shown are the MS for the $[M+2H]^{2+}$ ions of hemoglobin ε peptide MNVEEAGGEALGR and MNVEEAGGEALGR* ('R*' designates the heavy arginine) from the forward (A) and reverse (B) SILAC samples. Light and heavy peptides are labeled as "L" and "H", respectively. The ratios determined from the forward and reverse SILAC experiments were 2.17 and 2.45 (treated/untreated), respectively. Depicted in (C) and (D) are the MS/MS for the $[M+2H]^{2+}$ ions of MNVEEAGGEALGR and MNVEEAGGEALGR*, respectively.

from a single SILAC experiment was within 9%. In addition, the average relative standard deviation in the expression ratios for all the quantified proteins from the results of three sets of SILAC measurements was 15%. Thus, we considered those proteins with expression levels being altered by more than 1.5 fold as significantly changed proteins induced by imatinib treatment. Those significantly changed proteins that could be quantified in at least two, which encompassed at least one forward and one reverse, sets of SILAC experiments were considered for the subsequent discussion. Together, this gives quantifiable results for 73 proteins being significantly altered upon imatinib exposure, with 26 and 47 being up- and down-regulated, respectively (Table 3.1, and the detailed information about the quantified peptides and ratios for each measurement is listed in Table S3.2).

Imatinib induced erythroid differentiation in K562 cells

Among the differentially expressed proteins, all the identified hemoglobins as well as γ-globin were markedly up-regulated in the drug-treated K562 cells (Table 3.1). Hemoglobinization is the overarching feature of erythroid differentiation, during which less specialized blood cells are differentiated into erythrocytes. We also performed benzidine staining and observed more than 30% benzidine-positive cells upon a 24-hr treatment with 1 μM imatinib (Figure S3.1), which is consistent with previous findings [20, 21]. The up-regulation of hemoglobin proteins, together with benzidine staining result, supported unambiguously the imatinib-induced erythroid

Table 3.1. Proteins quantified with greater than 1.5 fold changes, with IPI numbers, protein names, average ratios and S.D. The listed S.D. values were calculated based on the ratios obtained from three biological replicates. Peptides used for the quantification of individual proteins are listed in Table S3.2.

IPI Number	Protein name	Peptide number	Average±S.D.
	Erythroid differentiation-related proteins		
IPI00410714.4	Hemoglobin α subunit	6	2.36±0.76
IPI00220706.9	Hemoglobin γ1 subunit	4	2.21±0.83
IPI00217471.2	Hemoglobin ε subunit	5	2.75±0.77
IPI00554676.1	Hemoglobin γ2 subunit	1	2.84±0.58
IPI00217473.4	Hemoglobin ζ subunit	12	3.02±1.18
IPI00030809.1	γ-G globin	6	2.34±0.55
IPI00027776.6	Ferrochelatase, mitochondrial precursor	5	1.54±0.22
IPI00093057.6	Coproporphyrinogen III oxidase, mitochondrial precursor	5	1.61±0.31
IPI00028160.1	Splice isoform 1 of porphobilinogen deaminase	8	1.83±0.07
IPI00008475.1	Hydroxymethylglutaryl-CoA synthase, cytoplasmic	12	0.67±0.21
IPI00642144.2	Aldehyde dehydrogenase 1 family, member A1	2	1.59±0.24
	Translation-related proteins		
IPI00021266.1	60S ribosomal protein L23a	9	0.63±0.13
IPI00299573.11	60S ribosomal protein L7a	14	0.65±0.15
IPI00013415.1	40S ribosomal protein S7	8	0.65±0.19
IPI00219155.4	60S ribosomal protein L27	2	0.66±0.09
IPI00030179.3	60S ribosomal protein L7	6	0.66±0.15
IPI00037619.2	Ribosomal protein L39 variant	1	0.67±0.05
IPI00176692.7	Heterogeneous nuclear ribonucleoprotein A1	4	0.63±0.24
IPI00029266.1	Small nuclear ribonucleoprotein E	2	0.64±0.23
IPI00013068.1	Eukaryotic translation initiation factor 3 subunit 6	2	0.57±0.17
IPI00016910.1	Eukaryotic translation initiation factor 3 subunit 8	11	0.66±0.07
IPI00219678.2	Eukaryotic translation initiation factor 2 subunit 1	2	0.67±0.10
	Histones		

IPI00021924.1	Histone H1x	4	1.59±0.20
IPI00219038.8	Histone H3.3	1	2.11±0.22
IPI00217465.4	Histone H1.2	10	2.26±0.47
IPI00003935.5	Histone H2B	7	2.35±0.70
IPI00171611.5	Histone H3.1	4	2.61±0.41
IPI00026272.1	Histone H2A	4	2.75±0.70
IPI00453473.5	Histone H4	8	3.06±0.82
	Other Enzymes		
IPI00221108.4	Thymidylate synthase	2	0.52±0.04
IPI00010349.1	Alkyldihydroxyacetonephosphate synthase, peroxisomal precursor	2	0.56±0.08
IPI00291669.3	Ubiquitin-like domain containing CTD phosphatase 1	2	0.62±0.07
IPI00305166.1	Succinate dehydrogenase [ubiquinone] flavoprotein subunit, mitochondrial precursor	1	0.62±0.04
IPI00017895.2	Splice isoform 1 of glycerol-3-phosphate dehydrogenase, mitochondrial precursor	3	0.64±0.07
IPI00026260.1	Nucleoside diphosphate kinase B	5	0.65±0.22
IPI00017617.1	Probable ATP-dependent RNA helicase DDX5	13	0.66±0.10
IPI00220373.3	Insulin-degrading enzyme	2	0.67±0.05
IPI00010157.1	S-adenosylmethionine synthetase isoform type-2	4	0.67±0.01
IPI00000728.3	Splice isoform 1 of ubiquitin carboxyl-terminal hydrolase 15	5	0.67±0.10
IPI00295741.4	Cathepsin B precursor	6	1.50±0.22
IPI00218568.6	Pterin-4-α-carbinolamine dehydratase	1	1.58±0.11
IPI00337541.3	NAD(P) transhydrogenase, mitochondrial precursor	1	1.91±0.03
IPI00449049.4	Poly [ADP-ribose] polymerase 1	7	2.43±0.42
	Other Proteins		
IPI00032561.1	Calcium-binding protein 39	2	0.40±0.06
IPI00002214.1	Importin α2 subunit	6	0.52±0.02
IPI00011631.5	Centromere/kinetochore protein zw10 homolog	1	0.55±0.08
IPI00005045.1	ATP-binding cassette sub-family F member 2	5	0.57±0.04
IPI00026182.4	F-actin capping protein α2 subunit	2	0.58±0.10
IPI00218505.6	Protein FAM112B	2	0.59±0.03
IPI00337602.4	56 kDa protein	1	0.59±0.09
IPI00218292.2	Splice isoform short of ubiquitin fusion degradation protein	4	0.60±0.23
IPI00306043.1	Splice isoform 1 of YTH domain protein 2	1	0.60±0.01

IPI00009010.3	UPF0315 protein AD-001	2	0.61±0.06
IPI00743544.1	15 kDa protein	1	0.61±0.02
IPI00220113.1	Splice isoform 2 of microtubule-associated protein 4	11	0.61±0.04
IPI00183294.3	Nuclear pore complex protein Nup214	7	0.61±0.16
IPI00221035.3	Splice isoform 1 of transcription factor BTF3	3	0.61±0.08
IPI00329625.2	Cell cycle progression 2 protein isoform 1	2	0.62±0.24
IPI00009057.1	Splice isoform A of Ras-GTPase-activating protein-binding protein 2	6	0.63±0.04
IPI00012535.1	DnaJ homolog subfamily A member 1	3	0.64±0.16
IPI00027493.1	4F2 cell-surface antigen heavy chain	5	0.65±0.09
IPI00186290.5	Elongation factor 2	31	0.66±0.03
IPI00012479.1	Putative nascent polypeptide-associated complex subunit α-like protein	1	0.66±0.06
IPI00013495.1	Splice isoform 2 of ATP-binding cassette sub-family F	2	0.66±0.04
IPI00023748.3	Nascent polypeptide-associated complex α subunit	7	0.67±0.04
IPI00397828.2	20 kDa protein	1	0.67±0.02
IPI00444704.2	Splice isoform 2 of G-rich sequence factor 1	1	0.67±0.01
IPI00219229.2	U6 snRNA-associated Sm-like protein LSm3	2	0.67±0.10
IPI00156689.3	Synaptic vesicle membrane protein VAT-1 homolog	6	1.50±0.10
IPI00217357.2	Cell division cycle and apoptosis regulator protein 1	2	1.51±0.11
IPI00024145.1	Splice isoform 1 of voltage-dependent anion-selective channel 2	4	1.57±0.03
IPI00220835.6	Protein transport protein Sec61 β subunit	1	1.57±0.03
IPI00744503.1	17 kDa protein	1	1.60±0.17

differentiation of K562 cells.

Apart from the observation of the up-regulation of hemoglobins and γ-globin, we also found three important enzymes required for heme biosynthesis, i.e., porphobilinogen deaminase, coproporphyrinogen III oxidase and ferrochelatase, to be substantially up-regulated upon imatinib treatment (Table 3.1). Porphobilinogen deaminase (PBGD) converts porphobilinogen to hydroxymethylbilane, coproporphyrinogen III oxidase oxidizes coproporphyrinogen III to protoporphyrinogen IX, and ferrochelatase catalyzes the final step of heme biosynthesis, which converts protoporphyrin IX into heme (Figure S3.2). Thus, the hemoglobinization in K562 cells could stem, in part, from the enhanced heme biosynthesis induced by the over-expression of these three enzymes.

Imatinib induced histone hyperacetylation in K562 cells

Viewing that many inducers of erythroid differentiation, which include sodium butyrate and valproic acid, were reported to act as histone deacetylase inhibitors and induce histone hyperacetylation in leukemia cells [22, 23], we reason that the imatinib-induced erythroid differentiation of K562 cells might also involve histone hyperacetylation. To test this, we extracted core histones from K562 cells that are treated with imatinib or mock-treated, and subjected the intact histone H4 and its constituent N-terminal peptide to MS analyses. It turned out that we indeed observed increased histone H4 acetylation based on the LC-MS analysis of the intact protein

(Figure S3.3A & B). Likewise, we found a greater than 25% increase in the levels of di- and tri-acetylation of the N-terminal peptide upon treatment with imatinib (the MALDI-MS for the Asp-N-produced H4 N-terminal peptide $_{1}$SGRGKGGKGLGKGGAKRHRKVLR$_{23}$ is shown in Figure 3.3). LC-MS/MS for the mono-, di- and tri-acetylated peptides revealed the acetylation of the N-terminus (for all three acetylated forms of the peptide), K16 (for both the di- and tri-acetylated peptide), and K8 or K12 (for the triacetylated peptide, Figure S3.4).

Imatinib affects the expression of histones and proteins involved in translation

Aside from the considerable upregulation of hemoglobins and related proteins, imatinib treatment also led to significant down-regulation of some important proteins associated with translational machinery, including ribosomal proteins, translation initiation factors and nuclear ribonucleoproteins (Table 3.1). These results are in accordance with the previous findings [24, 25] and with the imatinib-induced growth inhibition.

All four core histones and linker histone H1 were markedly up-regulated in imatinib-treated K562 cells (Table 3.1). The substantially increased expression of histones might reflect the considerable change in chromatin structure induced by imatinib treatment. In this regard, we chose to use only those peptides that do not bear any known post-translational modifications (PTMs) for the quantification [26]. This avoids the inaccurate quantification emanating from imatinib-induced perturbation in

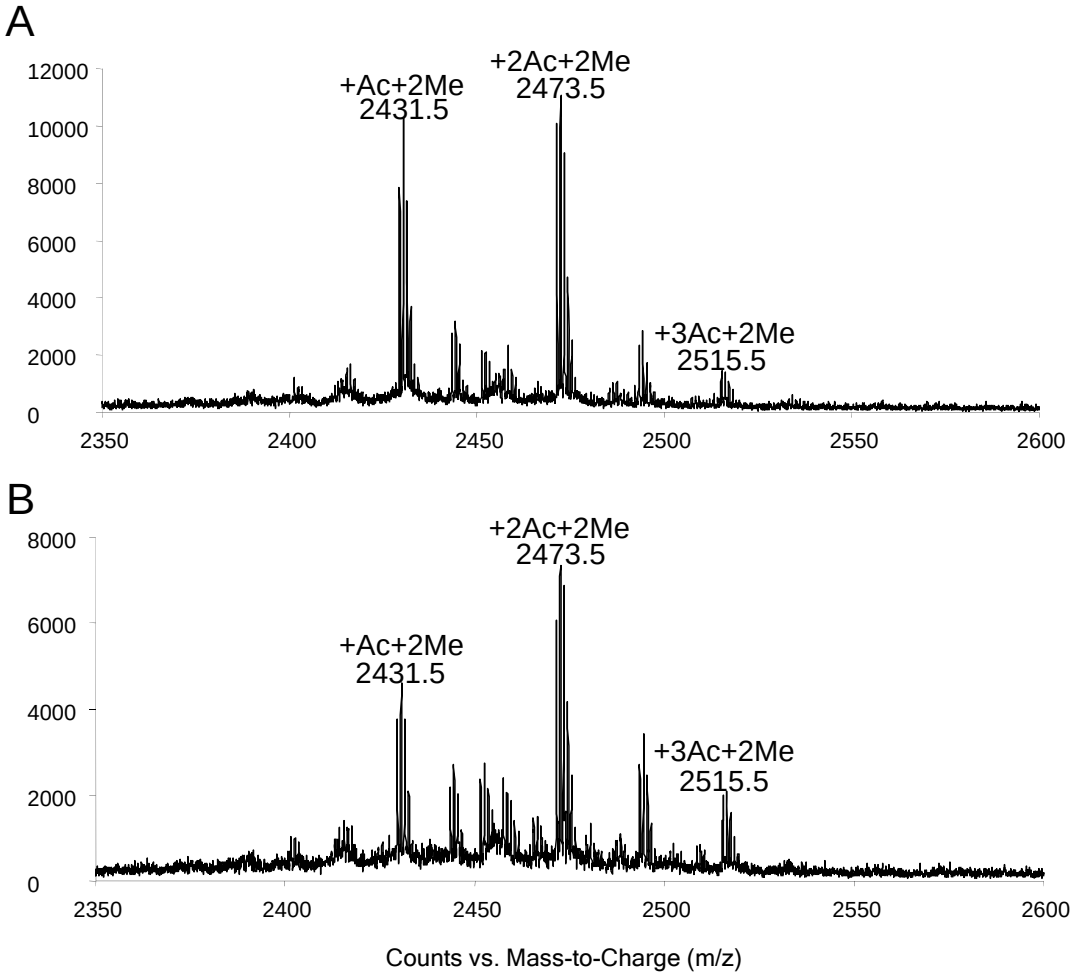


Figure 3.3 MALDI-MS for the Asp-N-produced peptide $_{1}$SGRGKGGKGLGKGGAKRHRKVLR$_{23}$ of histone H4 extracted from control (A) and imatinib-treated (B) K562 cells. The peptide with dimethylation and mono-, di-, tri-acetylation were labeled, while unmodified as well as mono- and tri-methylated peptides could also be observed as low-intensity peaks.

histone PTMs (e.g. H4 hyperacetylation).

Imatinib induced the alteration in expression of other important enzymes

Imatinib treatment also gave rise to considerable changes in the expression levels of some other important enzymes, including thymidylate synthase (TS), *S*-adenosylmethionine synthetase (AdoMetS), mitochondrial glycerol-3-phosphate dehydrogenase (GPD2), NAD(P) transhydrogenase, poly(ADP-ribose) polymerase 1, etc. (Table 3.1). These proteins play pivotal roles in different cellular pathways, and we would like to discuss some of them in detail.

TS is a folate-dependent enzyme that catalyzes the reductive methylation of 2′-deoxyuridine-5′-monophosphate to generate thymidine-5′-monophosphate and the latter is subsequently phosphorylated to thymidine-5′-triphosphate, which is an essential precursor for DNA synthesis [27]. We observed that TS was down-regulated by approximately 2 fold upon imatinib treatment. The inhibition of TS results in a lack of thymidine-5′-triphosphate and an accumulation of 2′-deoxyuridine-5′-monophosphate, which may be subsequently converted to 2′-deoxyuridine-5′-triphosphate and incorporated into DNA and induces DNA strand breaks and cell death. In this vein, Lee and coworkers [28] reported that a histone deacetylase inhibitor trichostatin A enhanced 5-fluorouracil cytotoxicity by down-regulating TS in human cancer cells. The diminished expression of TS might be attributed to the imatinib-induced histone hyperacetylation.

Poly(ADP-ribosyl)ation (PAR) catalyzed by poly(ADP-ribose) polymerase (PARP) is a major mediator of cell death after exposure to a variety of DNA damaging agents [29]. PARP1 is rapidly activated by DNA strand breaks and this enzyme catalyzes the transfer of the ADP-ribose moiety from the co-enzyme NAD^+ to a number of nuclear acceptor proteins [29]. Imatinib treatment was found to lead to a rapid increase in PAR which preceded the loss of integrity of mitochondrial membrane and DNA fragmentation; inhibition of PAR in imatinib-treated cells partially prevented cell death [30]. Therefore, the imatinib-induced up-regulation of PARP1 may contribute to the growth inhibition and apoptosis induction in K562 cells through stimulating PAR.

Discussion and Conclusions

Imatinib mesylate is considered the most effective and relatively safe drug for the treatment of chronic phase of CML. In this study, we employed SILAC, together with LC-MS/MS, and assessed quantitatively the perturbation of protein expression in K562 human CML cells by imatinib. Although a modest number of proteins were quantified in the present study, which is mainly due to the limitation in instrument sensitivity, the expression of many important proteins was found to be perturbed by imatinib. Our results revealed that the imatinib treatment led to the up-regulation of all the hemoglobinization-related proteins and the perturbation in the expression of many important enzymes.

Upon imatinib treatment, all the quantified hemoglobins and γ-globin were markedly up-regulated, and benzidine staining provided direct evidence supporting the imatinib-induced hemoglobinization and erythroid differentiation of K562 cells. More importantly, three key enzymes for heme biosynthesis were found up-regulated by 1.5-2 folds, demonstrating the enhanced endogenous hemoglobin biosynthesis. Viewing that many well-known histone deacetylase inhibitors could lead to differentiation of leukemia cells [22, 23], our quantitative proteomic results led us to predict and confirm that imatinib treatment gave rise to histone hyperacetylation in K562 cells. Although the imatinib-induced hemoglobinization has been reported previously [20, 21], the results from the present study paints the most complete picture for this pathway, by providing quantitative measurements of many proteins including hemoglobins and enzymes involved in heme biosynthesis, and demonstrating the ability of imatinib to induce histone hyperacetylation. The latter perturbation in histone acetylation epigenetic mark might be important in the imatinib-induced differentiation of CML cells.

Aside from proteins involved with erythroid differentiation, we found that many other important proteins were up- or down-regulated upon imatinib treatment. Future studies about the biological implications of the perturbation in expression of these proteins may also facilitate the discovery of additional molecular pathways that are altered by imatinib. This may contribute to an improved understanding of the imatinib-induced cytotoxicity and the development of resistance toward this drug.

Taken together, the pharmacoproteomic profiling could constitute a valuable tool for the identification of drug-responsive biomarkers and for the establishment of a molecular basis for developing novel and more effective approaches for the therapeutic intervention of human CML.

References:

1. Rowley, J. D., New consistent chromosomal abnormality in chronic myelogenous leukemia identified by quinacrine fluorescence and giemsa staining. *Nature* **1973,** 243, 290-293.

2. Sherbenou, D. W.; Druker, B. J., Applying the discovery of the Philadelphia chromosome. *J. Clin. Invest.* **2007,** 117, 2067-74.

3. Chu, S.; Holtz, M.; Gupta, M.; Bhatia, R., BCR/ABL kinase inhibition by imatinib mesylate enhances MAP kinase activity in chronic myelogenous leukemia $CD34^+$ cells. *Blood* **2004,** 103, 3167-3174.

4. Horita, M.; Andreu, E. J.; Benito, A.; Arbona, C.; Sanz, C.; Benet, I.; Fernandez-Luna, J. L., Blockade of the Bcr-Abl kinase activity induces apoptosis of chronic myelogenous leukemia cells by supressing STAT5-dependent expression of Bcl-x_L. *J. Exp. Hematol.* **2000,** 191, 977-784.

5. Savage, D. G.; Antman, K. H., Drug therapy: Imatinib mesylate - A new oral targeted therapy. *N. Eng. J. Med.* **2002,** 346, 683-693.

6. Balabanov, S.; Gontarewicz, A.; Ziegler, P.; Hartmann, U.; Kammer, W.; Copland, M.; Brassat, U.; Priemer, M.; Hauber, I.; Wilhelm, T.; Schwarz, G.; Kanz, L.; Bokemeyer, C.; Hauber, J.; Holyoake, T. L.; Nordheim, A.; Brummendorf, T. H., Hypusination of eukaryotic imtiation factor 5A (eIF5A): a novel therapeutic target in BCR-ABL-positive leukemias identified by a proteomics approach. *Blood* **2007,** 109, 1701-1711.

7. Smith, D. L.; Evans, C. A.; Pierce, A.; Gaskell, S. J.; Whetton, A. D., Changes in the proteome associated with the action of Bcr-Abl tyrosine kinase are not related to transcriptional regulation. *Mol. Cell. Proteomics* **2002,** 1, 876-884.

8. Unwin, R. D.; Sternberg, D. W.; Lu, Y. N.; Pierce, A.; Gilliland, D. G.; Whetton, A. D., Global effects of BCR/ABL and TEL/PDGFR beta expression on the proteome and phosphoproteome - Identification of the rho pathway as a target of BCR/ABL. *J. Biol. Chem.* **2005,** 280, 6316-6326.

9. Ong, S. E.; Mann, M., Mass spectrometry-based proteomics turns quantitative. *Nat. Chem. Biol.* **2005,** 1, 252-62.

10. Ong, S. E.; Blagoev, B.; Kratchmarova, I.; Kristensen, D. B.; Steen, H.; Pandey, A.; Mann, M., Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* **2002,** 1, 376-386.

11. Liang, X. Q.; Hajivandi, M.; Veach, D.; Wisniewski, D.; Clarkson, B.; Resh, M. D.; Pope, R. M., Quantification of change in phosphorylation of BCR-ABL kinase and its substrates in response to imatinib treatment in human chronic myelogenous leukemia cells. *Proteomics* **2006,** 6, 4554-4564.

12. Xiong, L.; Ping, L. Y.; Yuan, B. F.; Wang, Y. S., Methyl group migration during the fragmentation of singly charged ions of trimethyllysine-containing peptides: Precaution of using MS/MS of singly charged ions for interrogating peptide methylation. *J. Am. Soc. Mass Spectrom.* **2009,** 20, 1172-1181.

13. Park, S. K.; Venable, J. D.; Xu, T.; Yates, J. R., A quantitative analysis software tool for mass spectrometry-based proteomics. *Nat. Methods* **2008,** 5, 319-322.

14. McClatchy, D. B.; Liao, L. J.; Park, S. K.; Venable, J. D.; Yates, J. R., Quantification of the synaptosomal proteome of the rat cerebellum during post-natal development. *Genome Res.* **2007,** 17, 1378-1388.

15. Tabb, D. L.; McDonald, W. H.; Yates, J. R., DTASelect and contrast: Tools for assembling and comparing protein identifications from shotgun proteomics. *J. Proteome Res.* **2002,** 1, 21-26.

16. Uitto, P. M.; Lance, B. K.; Wood, G. R.; Sherman, J.; Baker, M. S.; Molloy, M. P., Comparing SILAC and two-dimensional gel electrophoresis image analysis for profiling urokinase plasminogen activator signaling in ovarian cancer cells. *J. Proteome Res.* **2007,** 6, 2105-2112.

17. Romijn, E. P.; Christis, C.; Wieffer, M.; Gouw, J. W.; Fullaondo, A.; van der Sluijs, P.; Braakman, I.; Heck, A. J. R., Expression clustering reveals detailed coexpression patterns of functionally related proteins during B cell differentiation - A proteomic study using a combination of one-dimensional gel electrophoresis, LC-MS/MS, and stable isotope labeling by amino acids in cell culture (SILAC). *Mol. Cell. Proteomics* **2005,** 4, 1297-1310.

18. Horikoshi, A.; Takei, K.; Sawada, S., Relationship between daily dose of imatinib per square meter and its plasma concentration in patients with chronic-phase chronic myeloid leukemia (CML). *Leuk. Res.* **2007,** 31, 574-575.

19. Tsutsumi, Y.; Kanamori, H.; Yamato, H.; Ehira, N.; Miura, T.; Kawamura, T.; Obara, S.; Masauzi, N.; Tanaka, J.; Imamura, M.; Asaka, M., Monitoring of plasma imatinib concentration for the effective treatment of CML patients. *Leuk. Res.* **2004,** 28, 1117-1118.

20. Fang, G.; Kim, C. N.; Perkins, C. L.; Ramadevi, N.; Winton, E.; Wittmann, S.; Bhalla, K. N., CGP57148B (STI-571) induces differentiation and apoptosis and sensitizes Bcr-Abl-positive human leukemia cells to apoptosis due to antileukemic drugs. *Blood* **2000,** 96, 2246-53.

21. Jacquel, A.; Herrant, M.; Legros, L.; Belhacene, N.; Luciano, F.; Pages, G.; Hofman, P.; Auberger, P., Imatinib induces mitochondria-dependent apoptosis of the Bcr-Abl positive K562 cell line and its differentiation towards the erythroid lineage. *FASEB J.* **2003,** 17, 2160-2162.

22. Gurvich, N.; Tsygankova, O. M.; Meinkoth, J. L.; Klein, P. S., Histone deacetylase is a target of valproic acid-mediated cellular differentiation. *Cancer Res.* **2004,** 64, 1079-86.

23. Wei, G. H.; Zhao, G. W.; Song, W.; Hao, D. L.; Lv, X.; Liu, D. P.; Liang, C. C., Mechanisms of human gamma-globin transcriptional induction by apicidin involves p38 signaling to chromatin. *Biochem. Biophys. Res. Commun.* **2007,** 363, 889-94.

24. Prabhu, S.; Saadat, D.; Zhang, M.; Halbur, L.; Fruehauf, J. P.; Ong, S. T., A novel mechanism for Bcr-Abl action: Bcr-Abl-mediated induction of the eIF4F translation initiation complex and mRNA translation. *Oncogene* **2007,** 26, 1188-1200.

25. Ly, C.; Arechiga, A. F.; Melo, J. V.; Walsh, C. M.; Ong, S. T., Bcr-Abl kinase modulates the translation regulators ribosomal protein S6 and 4E-BP1 in chronic myelogenous leukemia cells via the mammalian target of rapamycin. *Cancer Res.* **2003,** 63, 5716-5722.

26. Kouzarides, T., Chromatin modifications and their function. *Cell* **2007,** 128, 693-705.

27. Carreras, C. W.; Santi, D. V., The catalytic mechanism and structure of thymidylate synthase. *Ann. Rev. Biochem.* **1995,** 64, 721-762.

28. Lee, J. H.; Park, J. H.; Jung, Y.; Kim, J. H.; Jong, H. S.; Kim, T. Y.; Bang, Y. J., Histone deacetylase inhibitor enhances 5-fluorouracil cytotoxicity by down-regulating thymidylate synthase in human cancer cells. *Mol. Cancer Ther.* **2006,** 5, 3085-3095.

29. D'Amours, D.; Desnoyers, S.; D'Silva, I.; Poirier, G. G., Poly(ADP-ribosyl)ation reactions in the regulation of nuclear functions. *Biochem. J.* **1999,** 342, 249-268.

30. Moehring, A.; Wohlbold, L.; Aulitzky, W. E.; van der Kuip, H., Role of poly(ADP-ribose) polymerase activity in imatinib mesylate-induced cell death. *Cell Death Differ.* **2005,** 12, 627-636.

Supporting Information for Chapter 3

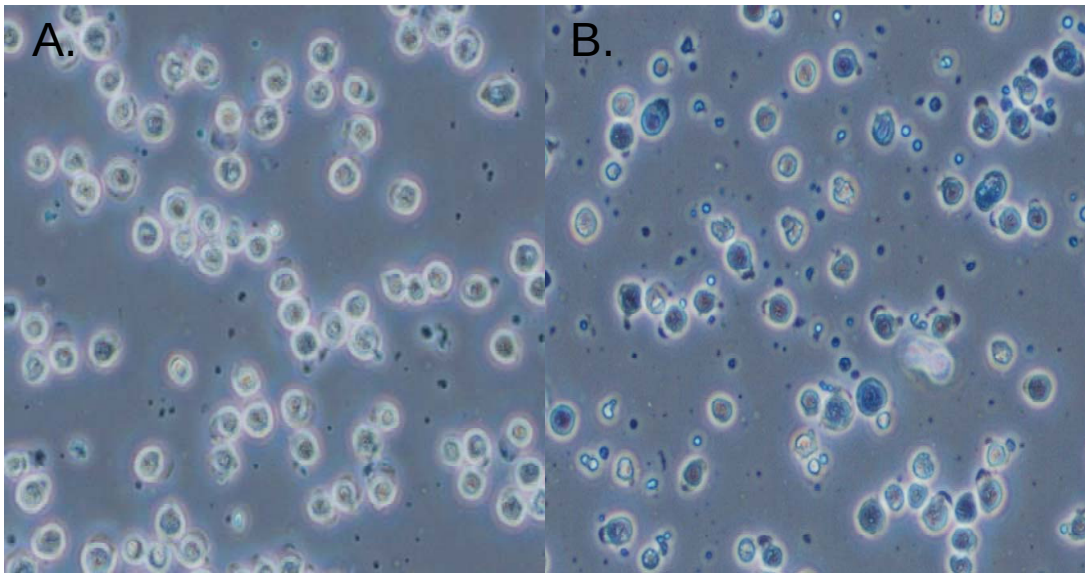


Figure S3.1 Light microscopic images for control (A) and imatinib-treated (B) K562 cells after benzidine staining.

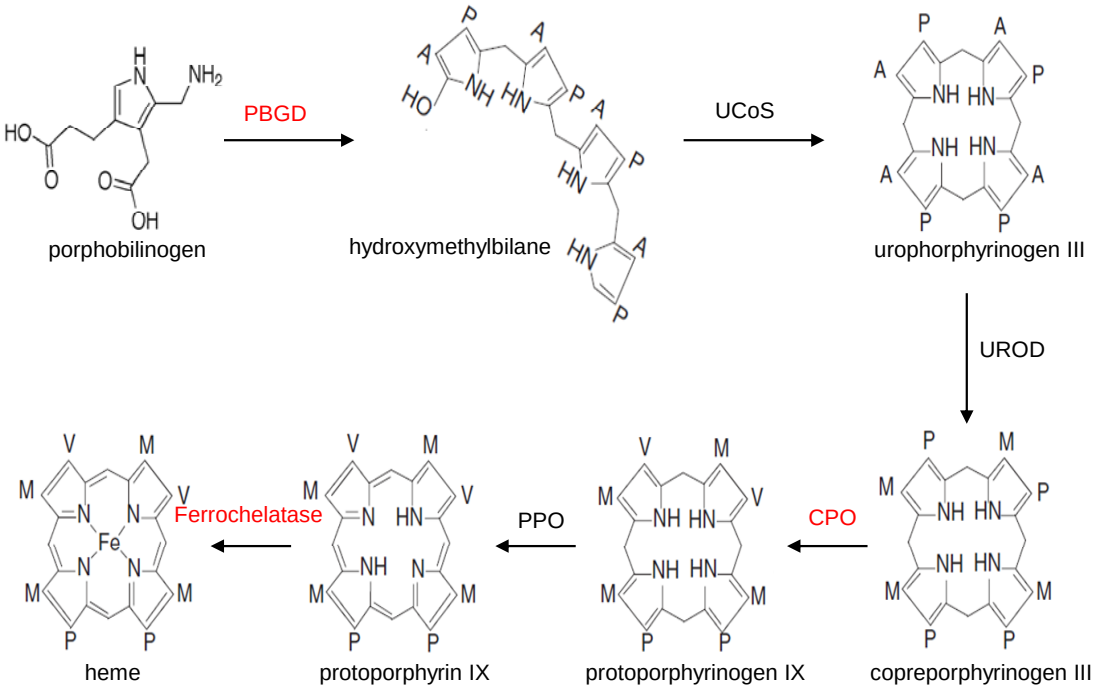


Figure S3.2 The heme biosynthesis pathway. Enzymes catalyzing the individual reactions in this pathway are indicated. The three enzymes found, by SILAC experiments, to be altered upon imatinib treatment are highlighted in red. Abbreviations: PBGD, porphobilinogen deaminase; UCoS, uroporphyrinogen III cosynthase; UROD, uroporphyrinogen decarboxylase; CPO, coproporphyrinogen oxidase; PPO, protoporphyrinogen oxidase; side chains: A, $-CH_2COO-$; P, $-CH_2CH_2COO-$; V, $-CH=CH$; M, $-CH_3$.

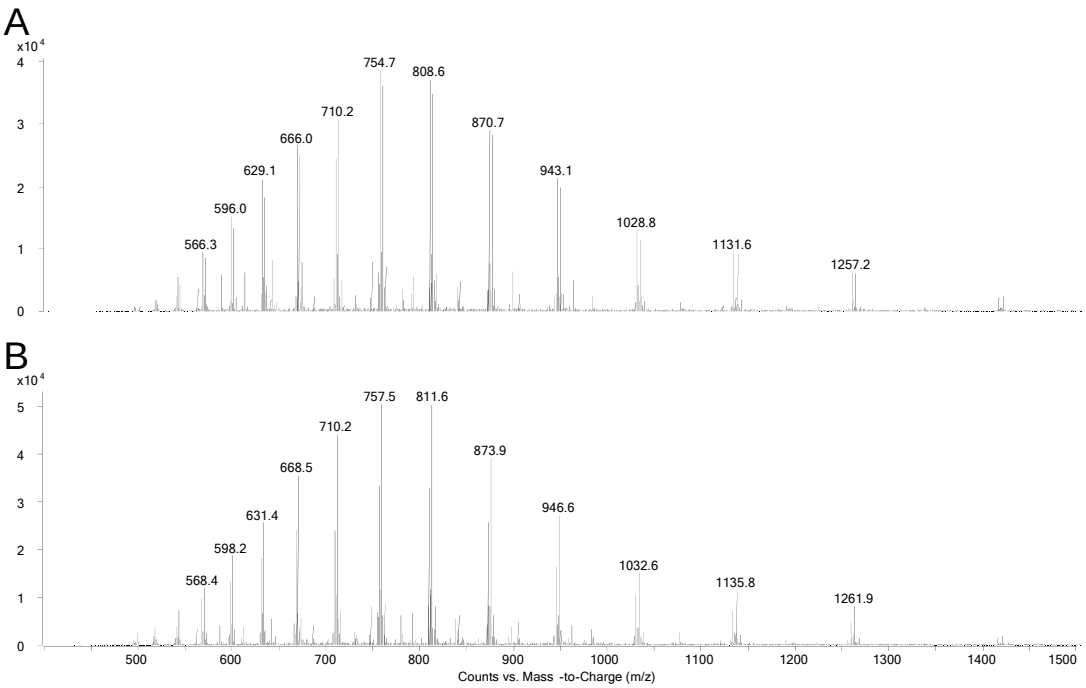


Figure S3.3 LC-MS of the intact histone H4 extracted from control (A) and imatinib-treated (B) K562 cells. The spectra were acquired on the Agilent QTOF mass spectrometer.

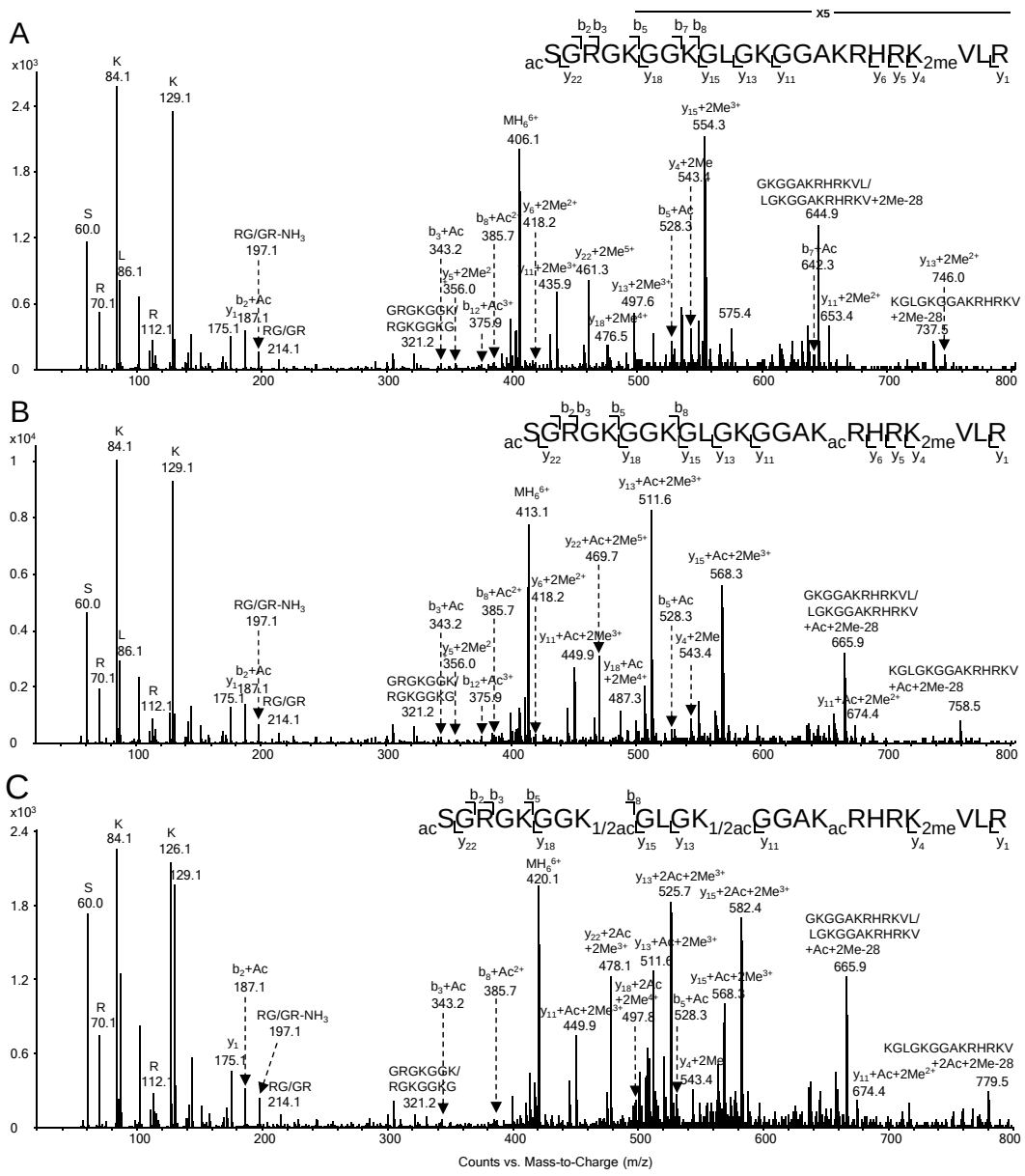


Figure S3.4 ESI-LC-MS/MS for the Asp-N-produced mono- (A), di- (B), and tri-acetylated (C) $_{1}$SGRGKGGKGLGKGGAKRHRKVLR$_{23}$ of histone H4 extracted from K562 cells. This peptide is also dimethylated on K20. "ac" and "2me" designate the acetylation and dimethylation sites; "1/2ac" listed in the peptide sequence represents the sites (K5 and K8) that are partially acetylated.

Table S3.1 Detailed results for all the proteins quantified in this study. "S.D." represents standard deviations calculated based on results from three cycles of SILAC measurements.

IPI number	Protein name	Peptide number	Average	S.D.
IPI00032561.1	Calcium-binding protein 39	2	0.396	0.060
IPI00002214.1	Importin alpha-2 subunit	6	0.521	0.015
IPI00221108.4	Thymidylate synthase	2	0.522	0.043
IPI00011631.5	Centromere/kinetochore protein zw10 homolog	1	0.493	0.082
IPI00010349.1	Alkyldihydroxyacetonephosphate synthase, peroxisomal precursor	2	0.557	0.075
IPI00013068.1	Eukaryotic translation initiation factor 3 subunit 6	2	0.567	0.166
IPI00005045.1	ATP-binding cassette sub-family F member 2	5	0.569	0.036
IPI00026182.4	F-actin capping protein alpha-2 subunit	2	0.580	0.102
IPI00218505.6	Protein FAM112B	2	0.585	0.032
IPI00337602.4	56 kDa protein	1	0.588	0.090
IPI00218292.2	Splice Isoform Short of Ubiquitin fusion degradation protein	4	0.598	0.226
IPI00306043.1	Splice Isoform 1 of YTH domain protein 2	1	0.600	0.011
IPI00009010.3	UPF0315 protein AD-001	2	0.606	0.055
IPI00743544.1	15 kDa protein	1	0.607	0.023
IPI00220113.1	Splice Isoform 2 of Microtubule-associated protein 4	11	0.609	0.036
IPI00183294.3	Nuclear pore complex protein Nup214	7	0.610	0.156
IPI00221035.3	Splice Isoform 1 of Transcription factor BTF3	3	0.612	0.078
IPI00291669.3	Ubiquitin-like domain containing CTD phosphatase 1	2	0.618	0.071
IPI00305166.1	Succinate dehydrogenase [ubiquinone] flavoprotein subunit, mitochondrial precursor	1	0.621	0.044
IPI00329625.2	Cell cycle progression 2 protein isoform 1	2	0.621	0.244
IPI00021266.1	60S ribosomal protein L23a	9	0.629	0.128
IPI00009057.1	Splice Isoform A of Ras-GTPase-activating protein-binding protein 2	6	0.631	0.043
IPI00176692.7	Heterogeneous nuclear ribonucleoprotein A1	4	0.634	0.243
IPI00029266.1	Small nuclear ribonucleoprotein E	2	0.637	0.233
IPI00017895.2	Splice Isoform 1 of Glycerol-3-phosphate	3	0.637	0.068

	dehydrogenase, mitochondrial precursor			
IPI00012535.1	DnaJ homolog subfamily A member 1	3	0.641	0.158
IPI00299573.11	60S ribosomal protein L7a	14	0.646	0.154
IPI00026260.1	Nucleoside diphosphate kinase B	5	0.648	0.224
IPI00013415.1	40S ribosomal protein S7	8	0.650	0.191
IPI00027493.1	4F2 cell-surface antigen heavy chain	5	0.653	0.085
IPI00016910.1	Eukaryotic translation initiation factor 3 subunit 8	11	0.655	0.069
IPI00186290.5	Elongation factor 2	31	0.656	0.030
IPI00012479.1	FKSG17	1	0.661	0.062
IPI00219155.4	60S ribosomal protein L27	2	0.662	0.088
IPI00013495.1	Splice Isoform 2 of ATP-binding cassette sub-family F	2	0.663	0.038
IPI00030179.3	60S ribosomal protein L7	6	0.663	0.154
IPI00017617.1	Probable ATP-dependent RNA helicase DDX5	13	0.664	0.095
IPI00023748.3	Nascent polypeptide-associated complex alpha subunit	7	0.666	0.043
IPI00220373.3	Insulin-degrading enzyme	2	0.668	0.053
IPI00219678.2	Eukaryotic translation initiation factor 2 subunit 1	2	0.668	0.102
IPI00397828.2	20 kDa protein	1	0.668	0.023
IPI00444704.2	Splice Isoform 2 of G-rich sequence factor 1	1	0.668	0.008
IPI00037619.2	Ribosomal protein L39 variant	1	0.668	0.050
IPI00219229.2	U6 snRNA-associated Sm-like protein LSm3	2	0.668	0.100
IPI00008475.1	Hydroxymethylglutaryl-CoA synthase, cytoplasmic	12	0.669	0.208
IPI00010157.1	S-adenosylmethionine synthetase isoform type-2	4	0.669	0.010
IPI00000728.3	Splice Isoform 1 of Ubiquitin carboxyl-terminal hydrolase 15	5	0.669	0.103
IPI00005658.3	Ubiquitin-like protein 4A	2	0.670	0.042
IPI00023406.1	Cytochrome c-type heme lyase	2	0.673	
IPI00176527.1	OTTHUMP00000018178	3	0.673	
IPI00018262.1	Acidic leucine-rich nuclear phosphoprotein 32 family member C	2	0.675	0.209
IPI00034006.1	Tyrosine-protein phosphatase non-receptor type 23	1	0.676	
IPI00373924.1	PREDICTED: similar to 40S ribosomal protein S3a isoform	2	0.676	0.020
IPI00418471.5	Vimentin	16	0.677	0.042
IPI00215965.1	Heterogeneous nuclear ribonucleoprotein A1 isoform b	3	0.678	0.087

IPI00012795.3	Eukaryotic translation initiation factor 3 subunit 2	2	0.679	0.058
IPI00000494.5	60S ribosomal protein L5	2	0.679	0.177
IPI00180292.5	Splice Isoform 5 of Brain-specific angiogenesis inhibitor 1-associated-protein 2	2	0.680	
IPI00220365.4	Eukaryotic translation initiation factor 4 gamma, 1 isoform 4	14	0.682	0.040
IPI00012079.1	Eukaryotic translation initiation factor 4B	2	0.682	0.058
IPI00215719.5	60S ribosomal protein L18	2	0.685	0.179
IPI00293474.1	Eucaryotic translation initiation factor 4G isoform 1	2	0.688	0.016
IPI00007764.3	Splice Isoform 1 of Hematological and neurological expressed 1 protein	2	0.690	
IPI00430812.4	Zinc finger protein 9	2	0.693	0.140
IPI00025244.1	Zinc-finger protein ZPR1	1	0.694	
IPI00031812.2	Nuclease sensitive element-binding protein 1	3	0.697	0.052
IPI00002203.5	BRCA2 and CDKN1A-interacting protein isoform BCCIPbeta	2	0.697	0.040
IPI00014424.1	Elongation factor 1-alpha 2	4	0.697	0.026
IPI00075558.7	PREDICTED: similar to 60S ribosomal protein L7a	2	0.698	0.184
IPI00024650.1	Monocarboxylate transporter 1	1	0.698	
IPI00011118.2	Ribonucleoside-diphosphate reductase M2 subunit	2	0.699	
IPI00012726.4	Splice Isoform 1 of Polyadenylate-binding protein 4	4	0.702	0.071
IPI00022228.1	Vigilin	2	0.704	0.183
IPI00009335.1	Brain protein 16	3	0.704	0.029
IPI00215637.4	ATP-dependent RNA helicase DDX3X	7	0.706	0.166
IPI00010700.2	Splice Isoform 1 of Large proline-rich protein BAT2	2	0.706	0.030
IPI00292020.3	Spermidine synthase	6	0.712	0.123
IPI00026833.4	Adenylosuccinate synthetase isozyme 2	5	0.712	0.220
IPI00215780.4	40S ribosomal protein S19	7	0.712	0.081
IPI00159072.3	ROD1 regulator of differentiation 1	2	0.713	0.172
IPI00329791.7	Probable ATP-dependent RNA helicase DDX46	6	0.713	
IPI00000875.5	Elongation factor 1-gamma	16	0.714	0.078
IPI00005416.2	NICE-4 protein	2	0.714	0.014
IPI00030910.1	GPI-anchored protein p137	5	0.715	0.154
IPI00005792.1	Poly(A) binding protein, nuclear 1	2	0.716	
IPI00016912.1	Tetratricopeptide repeat protein 1	2	0.716	

IPI00025447.6	EEF1A1 protein	4	0.717	0.038
IPI00641950.3	Lung cancer oncogene 7	10	0.718	0.069
IPI00024933.3	60S ribosomal protein L12	8	0.718	0.095
IPI00176854.2	13 kDa protein	3	0.718	0.050
IPI00026781.2	Fatty acid synthase	46	0.719	0.133
IPI00008524.1	Splice Isoform 1 of Polyadenylate-binding protein 1	21	0.719	0.037
IPI00023860.1	Nucleosome assembly protein 1-like 1	2	0.720	0.102
IPI00003865.1	Splice Isoform 1 of Heat shock cognate 71 kDa protein	92	0.720	0.053
IPI00647650.2	Eukaryotic translation initiation factor 3 subunit 3	3	0.720	0.115
IPI00012750.3	40S ribosomal protein S25	3	0.720	0.079
IPI00029557.3	GrpE protein homolog 1, mitochondrial precursor	2	0.722	0.113
IPI00012772.7	60S ribosomal protein L8	8	0.723	0.145
IPI00377261.1	Splice Isoform 1 of Far upstream element-binding protein 3	3	0.723	
IPI00011457.1	WUGSC:H_RG054D04.1 protein	19	0.725	0.017
IPI00414860.5	60S ribosomal protein L37a	3	0.726	
IPI00020850.3	Serine/threonine-protein phosphatase 2A 55 kDa regulatory subunit B beta isoform	2	0.726	
IPI00014919.1	UTP--glucose-1-phosphate uridylyltransferase 1	3	0.726	
IPI00155940.1	Importin alpha-7 subunit	2	0.730	
IPI00029012.1	Eukaryotic translation initiation factor 3 subunit 10	7	0.730	0.061
IPI00003918.5	60S ribosomal protein L4	3	0.731	0.067
IPI00030243.1	Splice Isoform 1 of Proteasome activator complex subunit 3	2	0.733	0.061
IPI00306708.3	T-lymphokine-activated killer cell-originated protein kinase	3	0.733	0.297
IPI00654777.2	Eukaryotic translation initiation factor 3 subunit 5	4	0.734	0.161
IPI00023461.2	Myeloid\/lymphoid or mixed-lineage leukemia (Trithorax homolog, Drosophila)\; translocated to, 4	3	0.734	0.115
IPI00221092.7	40S ribosomal protein S16	15	0.735	0.048
IPI00240909.1	PREDICTED: similar to eukaryotic translation initiation	2	0.736	0.073
IPI00741181.1	PREDICTED: similar to Heat shock protein HSP 90-beta	2	0.737	
IPI00019407.1	Sterol-4-alpha-carboxylate 3-dehydrogenase,	3	0.737	

	decarboxylating			
IPI00328082.3	Threonyl-tRNA synthetase-like 2	1	0.737	
IPI00419919.4	60S ribosomal protein L29	2	0.740	0.041
IPI00397963.3	9 kDa protein	2	0.741	
IPI00301936.3	ELAV-like protein 1	10	0.741	0.245
IPI00396980.3	PREDICTED: similar to 60S ribosomal protein L35	2	0.741	0.048
IPI00011913.1	Heterogeneous nuclear ribonucleoprotein A0	3	0.741	0.092
IPI00221091.8	40S ribosomal protein S15a	7	0.741	
IPI00019770.3	Ubiquitin-like protein fubi and ribosomal protein S30 precursor	7	0.742	0.014
IPI00176686.3	PREDICTED: similar to 60S ribosomal protein L6 (TAX-responsive enhancer element binding protein 107) (TAXREB107) (Neoplasm-related protein C140) isoform 1	2	0.743	0.100
IPI00031618.2	DNA-damage inducible protein 2	2	0.743	0.128
IPI00003505.3	Splice Isoform 1 of Thyroid receptor-interacting protein 13	2	0.743	0.128
IPI00022462.1	Transferrin receptor protein 1	8	0.743	0.048
IPI00002821.3	60S ribosomal protein L14	8	0.744	0.034
IPI00014263.1	Eukaryotic translation initiation factor 4H isoform 1	5	0.744	0.018
IPI00290460.3	Eukaryotic translation initiation factor 3 subunit 4	4	0.744	0.051
IPI00376798.3	Ribosomal protein L11	5	0.745	0.078
IPI00006181.1	Eukaryotic translation initiation factor 3 subunit 7	4	0.746	0.175
IPI00007144.1	60S ribosomal protein L26-like 1	3	0.746	0.030
IPI00016868.2	Hypothetical protein KIAA0683	2	0.746	
IPI00419880.5	40S ribosomal protein S3a	6	0.747	0.075
IPI00166711.4	FLJ00369 protein (Fragment)	2	0.748	
IPI00002255.4	Lipopolysaccharide-responsive and beige-like anchor protein	2	0.748	
IPI00329389.6	DNA-binding protein TAXREB107	8	0.749	0.106
IPI00219005.2	FK506-binding protein 4	6	0.751	0.101
IPI00027438.2	Flotillin-1	2	0.752	
IPI00029750.1	Splice Isoform 1 of 40S ribosomal protein S24	3	0.755	
IPI00020513.2	Zyxin	2	0.755	0.072
IPI00301263.2	CAD protein	10	0.756	0.050
IPI00175098.5	PREDICTED: similar to ribosomal protein L10	1	0.758	
IPI00024781.1	Small EDRK-rich factor 2	2	0.758	

IPI00019383.1	Galactokinase	4	0.759	
IPI00002212.1	Splice Isoform B of Serine/threonine-protein kinase 24	3	0.759	
IPI00178440.2	Elongation factor 1-beta	1	0.759	0.086
IPI00026138.3	PREDICTED: similar to ribosomal protein S3a isoform 1	6	0.760	0.038
IPI00247583.4	60S ribosomal protein L21	3	0.762	0.006
IPI00168262.1	CDNA PSEC0241 fis, clone NT2RP3000234, moderately similar to Homo sapiens cerebral cell adhesion molecule mRNA	1	0.763	
IPI00013296.1	40S ribosomal protein S18	9	0.764	0.098
IPI00007675.5	Cytoplasmic dynein 1 light intermediate chain 1	2	0.764	0.159
IPI00064328.3	Protein arginine methyltransferase 5 isoform b	3	0.765	0.145
IPI00025329.1	60S ribosomal protein L19	12	0.767	0.021
IPI00032158.2	Splice Isoform 2 of NMDA receptor-regulated protein 1	2	0.767	0.012
IPI00022143.3	Splice Isoform 1 of Protein FAM62A	13	0.768	0.019
IPI00465361.3	60S ribosomal protein L13	7	0.769	0.035
IPI00031523.3	Heat shock protein 86 (Fragment)	9	0.769	0.057
IPI00550021.3	60S ribosomal protein L3	4	0.769	
IPI00021808.3	Histidyl-tRNA synthetase	5	0.769	
IPI00007402.2	Importin-7	6	0.770	0.197
IPI00305692.4	Thioredoxin-like protein 1	3	0.770	0.197
IPI00219512.1	Splice Isoform 2 of Ubiquitin carboxyl-terminal hydrolase isozyme L5	2	0.770	0.177
IPI00022254.4	Splice Isoform 1 of Autophagy-related protein 3	2	0.771	0.139
IPI00025491.1	Eukaryotic initiation factor 4A-I	11	0.772	0.021
IPI00302925.3	T-complex protein 1 subunit theta	26	0.772	0.060
IPI00221325.3	Ran-binding protein 2	2	0.772	0.179
IPI00017763.4	Nucleosome assembly protein 1-like 4	3	0.772	0.044
IPI00219673.5	Glutathione S-transferase kappa 1	5	0.772	0.216
IPI00216153.6	40S ribosomal protein S15	2	0.773	0.022
IPI00025419.5	PREDICTED: similar to 40S ribosomal protein S3a (V-fos transformation effector protein) isoform 3	7	0.774	0.015
IPI00329745.4	130 kDa leucine-rich protein	43	0.774	0.190
IPI00008529.1	60S acidic ribosomal protein P2	10	0.774	0.055
IPI00013881.5	Heterogeneous nuclear ribonucleoprotein H1	2	0.775	0.185
IPI00216237.4	60S ribosomal protein L36	6	0.775	0.035

IPI00413324.5	60S ribosomal protein L17	7	0.775	0.044
IPI00000643.1	BAG family molecular chaperone regulator 2	3	0.775	
IPI00456889.2	PREDICTED: similar to Phosphoglycerate mutase 1	2	0.776	
IPI00299214.5	Thymidine kinase, cytosolic	2	0.779	0.195
IPI00013184.1	N-terminal acetyltransferase complex ARD1 subunit homolog A	7	0.779	0.029
IPI00219153.3	60S ribosomal protein L22	4	0.780	0.086
IPI00218407.5	Fructose-bisphosphate aldolase B	2	0.780	
IPI00304417.6	Isocitrate dehydrogenase [NAD] subunit beta, mitochondrial precursor	2	0.780	
IPI00735515.1	PREDICTED: similar to chaperonin containing TCP1, subunit 8	2	0.780	
IPI00412579.5	60S ribosomal protein L10a	16	0.780	0.029
IPI00027270.1	60S ribosomal protein L26	2	0.781	
IPI00012202.1	Methylosome protein 50	2	0.781	
IPI00333010.6	SR-related CTD associated factor 6	3	0.781	
IPI00290142.5	CTP synthase	4	0.781	0.110
IPI00008530.1	60S acidic ribosomal protein P0	12	0.782	0.108
IPI00009943.2	Tumor protein, translationally-controlled 1	3	0.784	0.152
IPI00328840.7	THO complex subunit 4	5	0.785	0.054
IPI00410693.3	Splice Isoform 1 of Plasminogen activator inhibitor 1 RNA-binding protein	6	0.786	0.145
IPI00027569.1	Heterogeneous nuclear ribonucleoprotein C-like 1	3	0.786	0.179
IPI00154473.4	Elongation factor G 1, mitochondrial precursor	2	0.786	0.185
IPI00026271.4	40S ribosomal protein S14	13	0.787	0.043
IPI00156232.1	15 kDa protein	2	0.787	
IPI00056494.4	60S ribosomal protein L36a-like	2	0.787	
IPI00001453.2	Alpha-internexin	3	0.787	
IPI00006408.3	ENOS interacting protein	2	0.787	
IPI00008809.1	GAGE-7 protein	2	0.787	
IPI00300620.3	Interferon-induced transmembrane protein 1	2	0.787	
IPI00001589.1	Mitochondrial import inner membrane translocase subunit Tim13	2	0.787	
IPI00013721.2	Serine/threonine-protein kinase PRP4 homolog	2	0.787	
IPI00182533.4	60S ribosomal protein L28	12	0.789	
IPI00015018.1	Inorganic pyrophosphatase	5	0.789	0.034
IPI00013917.1	40S ribosomal protein S12	10	0.789	0.002
IPI00030706.1	Activator of 90 kDa heat shock protein ATPase	2	0.789	0.176

	homolog 1			
IPI00026302.3	60S ribosomal protein L31	5	0.790	0.044
IPI00004208.4	PREDICTED: similar to eukaryotic translation initiation factor 3, subunit 6 interacting protein	3	0.791	#DIV/0!
IPI00179964.5	Splice Isoform 1 of Polypyrimidine tract-binding protein 1	2	0.791	0.131
IPI00006640.3	Serpin I2 precursor	2	0.791	0.258
IPI00329200.6	Importin beta-3	4	0.793	0.143
IPI00335061.6	Pyrroline-5-carboxylate reductase family, member 2	2	0.793	0.115
IPI00334587.1	Splice Isoform 2 of Heterogeneous nuclear ribonucleoprotein A/B	3	0.793	0.157
IPI00398121.5	Hypothetical protein DKFZp779B0247	2	0.794	
IPI00397611.1	PREDICTED: similar to ribosomal protein L13 isoform 1	13	0.794	0.071
IPI00252294.2	PREDICTED: similar to 60S ribosomal protein L35	4	0.795	0.085
IPI00032879.1	Adenylate kinase isoenzyme 6	2	0.795	
IPI00396435.2	DEAH (Asp-Glu-Ala-His) box polypeptide 15	2	0.795	
IPI00386189.2	Splice Isoform 1 of NMDA receptor-regulated protein 1	3	0.795	0.127
IPI00008433.3	40S ribosomal protein S5	5	0.795	0.060
IPI00008438.1	40S ribosomal protein S10	12	0.795	0.048
IPI00035167.7	PREDICTED: similar to ribosomal protein L10a	3	0.796	0.060
IPI00009328.2	Probable ATP-dependent RNA helicase DDX48	2	0.797	0.117
IPI00550746.4	Nuclear migration protein nudC	6	0.798	0.015
IPI00221300.2	Translation initiation factor eIF-2B alpha subunit	5	0.798	0.101
IPI00219160.2	60S ribosomal protein L34	2	0.798	
IPI00022648.2	Eukaryotic translation initiation factor 5	3	0.799	0.031
IPI00106642.4	Dihydropyrimidinase-like 2	2	0.799	
IPI00397774.2	PREDICTED: similar to laminin receptor 1	3	0.799	0.115
IPI00017596.2	Microtubule-associated protein RP/EB family member 1	3	0.799	0.065
IPI00014238.2	Lysyl-tRNA synthetase	5	0.799	0.142
IPI00165506.4	Polymerase delta-interacting protein 2	2	0.800	0.088
IPI00017870.1	OTTHUMP00000021786	4	0.800	
IPI00017963.1	Small nuclear ribonucleoprotein Sm D2	12	0.800	
IPI00334775.5	Hypothetical protein DKFZp761K0511	86	0.801	0.066
IPI00006377.2	PNAS-110	2	0.801	

IPI00005605.4	Splice Isoform 1 of Protein NDRG3	2	0.801	
IPI00030275.5	Heat shock protein 75 kDa, mitochondrial precursor	21	0.801	0.150
IPI00556589.2	PREDICTED: similar to 40S ribosomal protein S28	8	0.801	0.071
IPI00646304.3	Peptidylprolyl isomerase B precursor	8	0.802	
IPI00008293.4	PREDICTED: similar to 40S ribosomal protein S7	7	0.803	0.028
IPI00004534.3	Phosphoribosylformylglycinamidine synthase	10	0.803	0.123
IPI00013485.3	40S ribosomal protein S2	10	0.804	0.046
IPI00183169.5	ACF7 protein	3	0.804	
IPI00024821.1	26S proteasome non-ATPase regulatory subunit 14	3	0.805	0.202
IPI00008552.6	Thioredoxin-like protein 2	2	0.805	0.050
IPI00154451.5	Splice Isoform 1 of MMS19-like protein	2	0.806	
IPI00748920.1	Splicing coactivator subunit SRm300	15	0.808	0.016
IPI00003519.1	116 kDa U5 small nuclear ribonucleoprotein component	10	0.808	0.087
IPI00216106.3	Splice Isoform 3 of Putative GTP-binding protein PTD004	2	0.808	0.101
IPI00216592.1	Splice Isoform C1 of Heterogeneous nuclear ribonucleoproteins C1/C2	4	0.809	0.217
IPI00016572.1	Small nuclear ribonucleoprotein G	1	0.810	
IPI00395694.1	Splice Isoform 2 of Transportin 3	6	0.810	0.164
IPI00394699.1	Ribosomal protein homolog PD-1	7	0.811	0.058
IPI00221093.6	40S ribosomal protein S17	4	0.811	0.051
IPI00176696.1	PREDICTED: similar to 40S ribosomal protein S26	2	0.813	
IPI00399036.1	PREDICTED: similar to 40S ribosomal protein SA (p40)	2	0.813	
IPI00290566.1	T-complex protein 1 subunit alpha	16	0.814	0.076
IPI00025333.4	Splice Isoform 3 of Anamorsin	1	0.814	0.238
IPI00218606.6	40S ribosomal protein S23	6	0.815	
IPI00396370.5	Splice Isoform 1 of Eukaryotic translation initiation factor 3 subunit 9	7	0.815	0.117
IPI00011253.3	40S ribosomal protein S3	6	0.815	0.048
IPI00419373.1	Splice Isoform 1 of Heterogeneous nuclear ribonucleoprotein A3	2	0.816	0.065
IPI00414127.3	Ran-specific GTPase-activating protein	2	0.817	0.053

IPI00010720.1	T-complex protein 1 subunit epsilon	22	0.817	0.070
IPI00304612.8	60S ribosomal protein L13a	10	0.817	0.024
IPI00005969.1	F-actin capping protein alpha-1 subunit	3	0.817	0.067
IPI00011916.1	Multisynthetase complex auxiliary component p38	2	0.818	0.129
IPI00024364.2	Transportin-1	2	0.819	0.135
IPI00018140.3	Splice Isoform 1 of Heterogeneous nuclear ribonucleoprotein Q	5	0.820	0.109
IPI00294701.1	CDK-activating kinase assembly factor MAT1	2	0.820	
IPI00550821.2	Splice Isoform 1 of Cleavage and polyadenylation specificity factor 7	2	0.820	
IPI00303207.3	ATP-binding cassette sub-family E member 1	4	0.820	0.070
IPI00297779.6	T-complex protein 1 subunit beta	35	0.820	0.110
IPI00015029.1	Prostaglandin E synthase 3	7	0.821	0.027
IPI00304596.3	Non-POU domain-containing octamer-binding protein	8	0.821	0.018
IPI00386854.5	HNRPA2B1 protein	16	0.821	0.231
IPI00418497.1	Splice Isoform 2 of Import inner membrane translocase subunit TIM50, mitochondrial precursor	2	0.822	0.445
IPI00106491.3	mRNA turnover protein 4 homolog	2	0.824	0.013
IPI00299085.2	Melanoma-associated antigen C1	2	0.824	0.043
IPI00062151.4	PREDICTED: similar to ribosomal protein L15	2	0.824	
IPI00026958.4	Splice Isoform Short of NADPH:adrenodoxin oxidoreductase, mitochondrial precursor	2	0.825	0.306
IPI00006362.1	Splice Isoform 2 of Endothelial differentiation-related factor 1	4	0.826	
IPI00169434.3	Grb10 interacting GYF protein 2	2	0.826	
IPI00306301.1	Pyruvate dehydrogenase E1 component alpha subunit, somatic form, mitochondrial	3	0.826	
IPI00218342.9	C-1-tetrahydrofolate synthase, cytoplasmic	3	0.827	0.078
IPI00031691.1	60S ribosomal protein L9	4	0.827	0.103
IPI00373913.2	PREDICTED: KIAA1447 protein	3	0.828	0.007
IPI00183526.5	NCL protein	10	0.828	0.071
IPI00396378.3	Splice Isoform B1 of Heterogeneous nuclear ribonucleoproteins A2/B1	15	0.829	0.145
IPI00006482.1	Splice Isoform Long of Sodium/potassium-transporting ATPase alpha-1 chain precursor	6	0.830	0.146

IPI00010153.5	60S ribosomal protein L23	2	0.830	0.042
IPI00398958.1	PREDICTED: similar to 40S ribosomal protein SA	12	0.831	0.105
IPI00027193.2	Splice Isoform 2 of Chloride intracellular channel protein 5	2	0.831	0.224
IPI00012442.1	Ras-GTPase-activating protein-binding protein 1	2	0.831	0.170
IPI00306369.3	NOL1/NOP2/Sun domain family 2 protein	5	0.833	0.167
IPI00006246.1	Cell cycle progression 2 protein	2	0.833	
IPI00218829.8	G1 to S phase transition protein 1 homolog	5	0.833	
IPI00064765.2	60S ribosomal protein L10-like	6	0.833	0.052
IPI00171438.2	Thioredoxin domain-containing protein 5 precursor	1	0.833	0.086
IPI00221088.4	40S ribosomal protein S9	2	0.834	0.101
IPI00022305.4	MSTP017	2	0.834	0.039
IPI00163782.2	Splice Isoform 2 of Far upstream element-binding protein 1	12	0.834	0.217
IPI00009922.3	DC50	7	0.835	
IPI00429191.2	Eukaryotic peptide chain release factor subunit 1	2	0.835	0.049
IPI00003021.1	Sodium/potassium-transporting ATPase alpha-2 chain precursor	5	0.836	0.150
IPI00182289.5	40S ribosomal protein S29	4	0.836	0.016
IPI00329633.5	Threonyl-tRNA synthetase, cytoplasmic	9	0.836	0.138
IPI00031801.4	Splice Isoform 1 of DNA-binding protein A	2	0.838	0.301
IPI00025273.1	Splice Isoform Long of Trifunctional purine biosynthetic protein adenosine-3	6	0.838	0.106
IPI00026202.1	60S ribosomal protein L18a	3	0.839	
IPI00169288.3	Splice Isoform 2 of Retinal dehydrogenase 2	6	0.839	0.076
IPI00216587.8	40S ribosomal protein S8	4	0.840	0.070
IPI00419979.2	p21-activated kinase 2	2	0.840	
IPI00376817.1	PREDICTED: similar to 60S ribosomal protein L32	2	0.840	
IPI00375370.1	SEC13-like 1 isoform a	2	0.840	
IPI00329351.5	P60, 60-kDa heat shock protein, HSP60 (Fragment)	2	0.840	
IPI00018627.1	CDNA FLJ13194 fis, clone NT2RP3004378, weakly similar to Drosophila melanogaster separation anxiety protein	3	0.842	0.170
REV_37799_Protein	no description	2	0.842	0.269

IPI00029534.1	Amidophosphoribosyltransferase precursor	1	0.842	0.209
IPI00023048.3	Elongation factor 1-delta	7	0.843	0.136
IPI00010415.2	Splice Isoform 1 of Cytosolic acyl coenzyme A thioester hydrolase	2	0.843	
IPI00295485.1	Heat shock 70 kDa protein 4L	7	0.844	
IPI00302927.5	T-complex protein 1 subunit delta	34	0.845	0.049
IPI00012493.1	40S ribosomal protein S20	4	0.845	0.078
IPI00006252.3	Multisynthetase complex auxiliary component p43	7	0.846	0.138
IPI00221089.4	40S ribosomal protein S13	12	0.846	0.066
IPI00290770.3	Chaperonin containing TCP1, subunit 3 isoform b	33	0.846	0.041
IPI00218993.1	Splice Isoform Beta of Heat-shock protein 105 kDa	1	0.846	0.148
IPI00000684.4	Splice Isoform AGX2 of UDP-N-acetylhexosamine pyrophosphorylase	3	0.846	0.067
IPI00002149.1	GTP-binding protein SAR1b	2	0.846	0.027
IPI00022744.5	Splice Isoform 1 of Importin-alpha re-exporter	3	0.846	0.169
IPI00301204.2	Retinol dehydrogenase 13	2	0.847	0.185
IPI00025178.3	Breast carcinoma amplified sequence 2	2	0.847	
IPI00295098.3	Signal recognition particle receptor beta subunit	1	0.848	0.031
IPI00398135.2	PREDICTED: similar to 60S ribosomal protein L27a	3	0.849	0.102
IPI00022970.4	Nucleoprotein TPR	6	0.849	0.156
IPI00012200.1	MGC2477 protein	3	0.850	
IPI00017342.1	Rho-related GTP-binding protein RhoG precursor	5	0.850	
IPI00006052.3	Prefoldin subunit 2	4	0.850	0.082
IPI00011603.2	26S proteasome non-ATPase regulatory subunit 3	2	0.851	0.224
IPI00029731.8	60S ribosomal protein L35a	3	0.851	
IPI00027626.2	T-complex protein 1 subunit zeta	25	0.851	0.080
IPI00021840.1	40S ribosomal protein S6	11	0.852	0.114
IPI00397431.3	PREDICTED: similar to monoacylglycerol O-acyltransferase 2	5	0.852	0.122
IPI00000873.3	Valyl-tRNA synthetase	17	0.853	0.144
IPI00003783.1	Dual specificity mitogen-activated protein kinase kinase 2	2	0.853	0.103
IPI00015351.1	AD039	1	0.854	
IPI00001091.3	AFG3-like protein 2	3	0.855	
IPI00293817.3	Gamma-soluble NSF attachment protein	3	0.855	
IPI00031583.2	Hypothetical protein DKFZp451D234	2	0.855	

IPI00009946.4	Mitochondrial import receptor subunit TOM34	6	0.855	
IPI00103415.1	Signal transducer and activator of transcription 5B	4	0.855	
IPI00306332.4	60S ribosomal protein L24	12	0.855	0.019
IPI00444262.1	CDNA FLJ45706 fis, clone FEBRA2028457, highly similar to Nucleolin	15	0.857	0.078
IPI00019599.1	Splice Isoform 1 of Ubiquitin-conjugating enzyme E2 variant 1	3	0.857	
IPI00013847.4	Ubiquinol-cytochrome-c reductase complex core protein I, mitochondrial precursor	2	0.857	0.109
IPI00001734.2	Splice Isoform 1 of Phosphoserine aminotransferase	6	0.857	0.261
IPI00373857.1	PREDICTED: similar to laminin receptor 1	2	0.858	0.058
IPI00644127.1	Isoleucyl-tRNA synthetase, cytoplasmic	2	0.860	0.094
IPI00219913.9	Ubiquitin carboxyl-terminal hydrolase 14	5	0.860	0.210
IPI00082831.3	130 kDa protein	2	0.861	0.201
IPI00171903.1	Heterogeneous nuclear ribonucleoprotein M isoform a	29	0.861	0.030
IPI00017297.1	Matrin-3	31	0.861	0.090
IPI00472102.3	Heat shock protein 60	64	0.861	0.111
IPI00219034.2	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subun	2	0.862	
IPI00026167.2	NHP2-like protein 1	2	0.862	
IPI00396437.3	Splice Isoform 2 of Drebrin-like protein	2	0.862	
IPI00015139.4	PREDICTED: similar to Heterogeneous nuclear ribonucleoprotein A1 isoform 1	8	0.862	0.027
IPI00177008.1	IMP dehydrogenase/GMP reductase family protein	1	0.863	0.102
IPI00217223.1	Multifunctional protein ADE2	18	0.863	0.068
IPI00376005.1	Eukaryotic initiation factor 5A isoform I variant A	2	0.863	0.178
IPI00291510.3	Inosine-5'-monophosphate dehydrogenase 2	12	0.864	0.039
IPI00302850.4	Small nuclear ribonucleoprotein Sm D1	5	0.864	0.079
IPI00021187.3	RuvB-like 1	4	0.864	0.018
IPI00009104.6	RuvB-like 2	3	0.864	0.154
IPI00376667.3	PREDICTED: similar to 40S ribosomal protein S3a (V-fos transformation effector protein) isoform 1	3	0.865	0.124
IPI00337385.1	Splice Isoform 1 of Pre-mRNA-processing factor 40 homolog A	3	0.865	0.191
IPI00170935.1	Leucine-rich repeat-containing protein 47	3	0.865	0.141

IPI00076042.2	Short heat shock protein 60 Hsp60s2	6	0.865	0.067
IPI00148061.3	L-lactate dehydrogenase A-like 6A	2	0.865	0.076
IPI00645194.1	Integrin beta 1 isoform 1A precursor	2	0.865	
IPI00019812.1	Serine/threonine-protein phosphatase 5	2	0.865	
IPI00234368.1	Splice Isoform Short of Adenosine kinase	9	0.865	
IPI00012048.1	Nucleoside diphosphate kinase A	8	0.866	0.122
IPI00024719.1	Histone acetyltransferase type B catalytic subunit	10	0.866	0.075
IPI00219156.6	60S ribosomal protein L30	9	0.866	0.081
IPI00217030.9	40S ribosomal protein S4, X isoform	23	0.867	0.124
IPI00299254.3	Eukaryotic translation initiation factor 5B	5	0.867	0.175
IPI00019600.2	Ubiquitin-conjugating enzyme E2 variant 2	1	0.868	0.284
IPI00395627.3	Splice Isoform 1 of Calcyclin-binding protein	19	0.869	0.105
IPI00026513.5	Ribose 5-phosphate isomerase A	2	0.869	0.171
IPI00018288.1	DNA-directed RNA polymerase II 33 kDa polypeptide	2	0.870	
IPI00027423.2	Protein phosphatase 1, catalytic subunit, alpha isoform	2	0.870	
IPI00219617.4	Ribose-phosphate pyrophosphokinase II	2	0.870	
IPI00061178.1	RNA binding motif protein, X-linked-like 1	1	0.870	
IPI00216393.1	Splice Isoform Non-brain of Clathrin light chain A	4	0.870	
IPI00334190.4	Stomatin-like protein 2	2	0.870	0.011
IPI00156374.5	Splice Isoform 1 of Importin-4	10	0.872	0.085
IPI00640703.3	Splice Isoform 1 of Exportin-5	2	0.873	0.093
IPI00180675.4	Tubulin alpha-3 chain	3	0.873	0.373
IPI00052885.5	18 kDa protein	2	0.873	0.218
IPI00219740.3	Splice Isoform 2 of DNA replication licensing factor MCM7	4	0.873	0.067
IPI00023647.2	Hypothetical protein DKFZp451P021	4	0.874	0.269
IPI00295992.4	Splice Isoform 2 of ATPase family AAA domain-containing protein 3A	2	0.874	0.038
IPI00014151.3	26S proteasome non-ATPase regulatory subunit 6	2	0.874	0.073
IPI00219616.6	Ribose-phosphate pyrophosphokinase I	2	0.874	0.164
IPI00030876.4	Protein diaphanous homolog 1	6	0.876	0.149
IPI00420014.2	U5 small nuclear ribonucleoprotein 200 kDa helicase	10	0.876	
IPI00182757.9	Splice Isoform 1 of Protein KIAA1967	6	0.876	0.121
REV_13527_Protein		2	0.876	
IPI00027430.1	Leukosialin precursor	2	0.876	

IPI00221178.1	Splice Isoform 2 of Tumor protein D54	5	0.876	0.149
IPI00012578.1	Importin alpha-4 subunit	4	0.877	0.046
IPI00292894.3	Hypothetical protein LOC55720	5	0.877	
IPI00163505.2	Splice Isoform 1 of RNA-binding region-containing protein 39	3	0.877	
IPI00219649.3	Splice Isoform 2 of Cystathionine beta-synthase	2	0.877	
IPI00329332.1	Syntaxin-12	2	0.877	
IPI00293126.1	Tubulin-specific chaperone B	2	0.877	
IPI00301058.4	Vasodilator-stimulated phosphoprotein	5	0.877	
IPI00216049.1	Splice Isoform 1 of Heterogeneous nuclear ribonucleoprotein K	29	0.877	0.122
IPI00294536.1	Serine-threonine kinase receptor-associated protein	4	0.878	0.093
IPI00018873.1	Splice Isoform 1 of Nicotinamide phosphoribosyltransferase	2	0.878	0.033
IPI00019927.2	26S proteasome non-ATPase regulatory subunit 7	10	0.878	0.131
IPI00008240.2	Methionyl-tRNA synthetase	5	0.879	0.142
IPI00017334.1	Prohibitin	10	0.879	0.054
IPI00445401.2	Splice Isoform 2 of HECT, UBA and WWE domain-containing protein 1	5	0.880	0.235
IPI00220480.7	60 kDa protein	2	0.880	0.137
IPI00005589.1	PREDICTED: similar to 60S ribosomal protein L32	2	0.880	
IPI00011937.1	Peroxiredoxin-4	3	0.880	0.133
IPI00007750.1	Tubulin alpha-1 chain	33	0.881	0.134
IPI00175108.4	PREDICTED: similar to heterogeneous nuclear ribonucleoprotein K isoform a isoform 2	3	0.881	0.113
IPI00292753.7	GTPase activating protein and VPS9 domains 1	4	0.882	0.112
IPI00216184.2	Splice Isoform 2 of Phosphatidylinositol-binding clathrin assembly protein	3	0.882	0.009
IPI00402104.5	15 kDa protein	2	0.882	0.160
IPI00105598.3	Proteasome 26S non-ATPase subunit 11 variant (Fragment)	8	0.883	0.131
IPI00215918.2	ADP-ribosylation factor 4	2	0.883	
IPI00031570.1	Probable UPF0334 kinase-like protein C1orf57	2	0.883	
IPI00029079.5	GMP synthase	27	0.884	0.082
IPI00031820.2	Phenylalanyl-tRNA synthetase alpha chain	2	0.884	0.077
IPI00221354.1	Splice Isoform Short of RNA-binding protein FUS	4	0.885	0.047

IPI00008575.3	Splice Isoform 1 of KH domain-containing, RNA-binding, signal transduction-associated protein 1	6	0.885	0.012
IPI00011569.1	Acetyl-CoA carboxylase 1	2	0.885	
IPI00000279.2	Hypothetical protein LEREPO4	2	0.885	
IPI00005657.1	Prefoldin subunit 6	2	0.885	
IPI00012340.1	Splicing factor, arginine/serine-rich 9	2	0.885	
IPI00218775.2	FK506-binding protein 5	2	0.886	0.285
IPI00219217.2	L-lactate dehydrogenase B chain	2	0.886	0.038
IPI00030702.1	Splice Isoform 1 of Isocitrate dehydrogenase [NAD] subunit alpha, mitochondrial precursor	4	0.888	0.198
IPI00005948.1	Hypothetical protein LOC84245 isoform 1	2	0.889	0.188
IPI00019880.1	47 kDa heat shock protein precursor	2	0.889	0.216
IPI00005537.2	39S ribosomal protein L12, mitochondrial precursor	3	0.890	0.089
IPI00007163.1	U6 snRNA-associated Sm-like protein LSm7	4	0.890	
IPI00011274.1	Heterogeneous nuclear ribonucleoprotein D-like	11	0.891	0.019
IPI00031836.3	Developmentally-regulated GTP-binding protein 1	3	0.891	0.014
IPI00010204.1	Splicing factor, arginine/serine-rich 3	2	0.891	0.086
IPI00290461.3	Eukaryotic translation initiation factor 3 subunit 1	3	0.891	0.114
IPI00027397.3	Splice Isoform 1 of Hematological and neurological expressed 1-like protein	4	0.893	
IPI00299000.4	Proliferation-associated protein 2G4	9	0.893	0.004
IPI00374686.3	PREDICTED: similar to heterogeneous nuclear ribonucleoprotein A3 isoform 1	4	0.894	
IPI00289535.6	Mitochondrial-processing peptidase beta subunit, mitochondrial precursor	2	0.894	0.239
IPI00018465.1	T-complex protein 1 subunit eta	8	0.894	0.013
IPI00549248.3	Nucleophosmin	3	0.894	
IPI00219330.2	Splice Isoform 5 of Interleukin enhancer-binding factor 3	2	0.894	0.247
IPI00219525.9	6-phosphogluconate dehydrogenase, decarboxylating	5	0.894	0.205
IPI00029764.1	Splicing factor 3A subunit 3	2	0.895	0.091
IPI00301434.4	BolA-like protein 2 isoform a	3	0.895	0.109
IPI00013146.1	Mitochondrial 28S ribosomal protein S22	3	0.895	0.056
IPI00243220.2	132 kDa protein	3	0.895	0.156
IPI00011284.1	Catechol O-methyltransferase	3	0.896	0.073

IPI00033130.3	Ubiquitin-like 1-activating enzyme E1A	6	0.897	0.144
IPI00006510.1	Tubulin beta-1 chain	5	0.897	0.034
IPI00027448.3	ATP synthase subunit g, mitochondrial	2	0.898	0.016
IPI00026964.1	Ubiquinol-cytochrome c reductase iron-sulfur subunit, mitochondrial precursor	2	0.898	0.145
IPI00176698.2	12 kDa protein	7	0.898	0.203
IPI00026689.4	Hypothetical protein DKFZp686L20222	3	0.898	0.183
IPI00024664.1	Splice Isoform Long of Ubiquitin carboxyl-terminal hydrolase 5	2	0.898	0.233
IPI00032460.3	U6 snRNA-associated Sm-like protein LSm2	4	0.899	0.102
IPI00007927.3	Splice Isoform 1 of Structural maintenance of chromosome 2-like 1 protein	1	0.899	0.217
IPI00021016.3	Splice Isoform 1 of Elongation factor Ts, mitochondrial precursor	2	0.900	0.011
IPI00029485.2	Splice Isoform p150 of Dynactin-1	7	0.901	0.231
IPI00009253.1	Alpha-soluble NSF attachment protein	5	0.901	0.141
IPI00021926.2	26S protease regulatory subunit S10B	6	0.902	0.143
IPI00013894.1	Stress-induced-phosphoprotein 1	17	0.902	0.041
IPI00549861.2	CDNA FLJ13369 fis, clone PLACE1000610, weakly similar to MSN5 PROTEIN	3	0.902	0.046
IPI00002520.1	Serine hydroxymethyltransferase, mitochondrial precursor	3	0.903	0.179
IPI00465260.3	GARS protein	11	0.903	0.215
IPI00218847.4	Splice Isoform 2 of Low molecular weight phosphotyrosine protein phosphatase		0.904	0.175
IPI00016339.4	Ras-related protein Rab-5C	12	0.904	0.078
IPI00009305.1	Glucosamine-6-phosphate isomerase	2	0.904	0.081
IPI00003881.5	Heterogeneous nuclear ribonucleoprotein F	3	0.905	0.112
IPI00021700.3	Proliferating cell nuclear antigen	5	0.905	0.127
IPI00013877.2	Splice Isoform 1 of Heterogeneous nuclear ribonucleoprotein H3	4	0.906	0.162
IPI00005705.1	Splice Isoform Gamma-1 of Serine/threonine-protein phosphatase PP1-gamma catalytic subunit	4	0.906	0.148
IPI00007812.1	Vacuolar ATP synthase subunit B, brain isoform	2	0.906	0.174
IPI00026670.3	Transcription elongation factor B polypeptide 2	1	0.906	
IPI00014898.1	Splice Isoform 1 of Plectin-1	2	0.906	0.147
IPI00020984.1	Calnexin precursor	4	0.907	0.051
IPI00026154.1	Glucosidase 2 beta subunit precursor	2	0.907	0.082

IPI00185374.3	26S proteasome non-ATPase regulatory subunit 12	4	0.907	0.079
IPI00029744.1	Single-stranded DNA-binding protein, mitochondrial precursor	13	0.907	0.293
IPI00003949.1	Ubiquitin-conjugating enzyme E2 N	5	0.908	0.051
IPI00006612.1	Clathrin coat assembly protein AP180	2	0.908	
IPI00000060.1	Guanine nucleotide-binding protein G(I)/G(S)/G(O) gamma-5-like subunit	2	0.908	
IPI00027230.3	Endoplasmin precursor	23	0.908	0.049
IPI00017303.1	DNA mismatch repair protein Msh2	2	0.909	0.303
IPI00024524.3	Putatative 28 kDa protein	2	0.909	
IPI00215888.3	Signal recognition particle 72 kDa protein	1	0.909	
IPI00290305.3	TP53-regulating kinase	2	0.909	
IPI00001890.8	Coatomer subunit gamma	2	0.909	0.176
IPI00456750.2	Niban-like protein	2	0.910	0.118
IPI00643041.2	GTP-binding nuclear protein Ran	2	0.910	0.143
IPI00007765.5	Stress-70 protein, mitochondrial precursor	49	0.910	0.015
IPI00376503.1	Pyrroline-5-carboxylate reductase 1 isoform 2	8	0.912	0.111
IPI00029740.1	Sorting nexin 3 isoform a	3	0.912	
IPI00396321.1	Leucine-rich repeat-containing protein 59	6	0.913	0.050
IPI00010201.4	26S proteasome non-ATPase regulatory subunit 8	2	0.914	0.054
IPI00012074.3	Heterogeneous nuclear ribonucleoprotein R	10	0.915	0.083
IPI00003377.1	Splice Isoform 1 of Splicing factor, arginine/serine-rich 7	8	0.915	0.066
IPI00000877.1	150 kDa oxygen-regulated protein precursor	2	0.916	0.120
IPI00000874.1	Peroxiredoxin-1	3	0.916	0.155
IPI00554777.1	Asparagine synthetase	2	0.916	0.113
IPI00013949.1	Small glutamine-rich tetratricopeptide repeat-containing protein A	3	0.917	0.033
IPI00022694.3	Splice Isoform Rpn10A of 26S proteasome non-ATPase regulatory subunit 4	6	0.917	0.118
IPI00026268.2	Guanine nucleotide-binding protein G(I)/G(S)/G(T) beta s	4	0.917	0.103
IPI00465315.5	Cytochrome c	2	0.917	0.181
IPI00219335.9	60S ribosomal protein L3-like	2	0.917	
IPI00025277.5	Programmed cell death protein 6	1	0.917	
IPI00000948.3	Transducin beta-like 2 protein	2	0.917	
IPI00018398.4	26S protease regulatory subunit 6A	4	0.918	0.140
IPI00016768.2	L-lactate dehydrogenase A-like 6B	1	0.919	

IPI00219077.3	LTA4H protein	2	0.919	
IPI00020008.1	NEDD8 precursor	2	0.919	
REV_1501_Protein	no description	1	0.919	
IPI00465439.4	Fructose-bisphosphate aldolase A	7	0.919	0.130
IPI00220642.6	14-3-3 protein gamma	11	0.919	0.075
IPI00291946.7	Ubiquitin carboxyl-terminal hydrolase 10	2	0.919	0.350
IPI00003362.2	Hypothetical protein	37	0.919	0.122
IPI00294627.1	Splice Isoform 2 of Splicing factor 1	2	0.919	0.206
IPI00024620.5	E(Y)2 homolog	2	0.920	0.112
IPI00375380.3	proteasome 26S non-ATPase subunit 13 isoform 2	3	0.921	0.101
IPI00028888.1	Splice Isoform 1 of Heterogeneous nuclear ribonucleoprotein D0	2	0.921	0.189
IPI00301277.1	Heat shock 70 kDa protein 1L	7	0.921	0.165
IPI00438229.1	Splice Isoform 1 of Transcription intermediary factor 1-beta	8	0.921	0.095
IPI00300371.3	Splicing factor 3B subunit 3	2	0.922	0.050
IPI00007755.2	Ras-related protein Rab-21	5	0.922	0.059
IPI00011134.1	Heat shock 70 kDa protein 7 (Fragment)	2	0.922	0.187
IPI00021290.4	ATP-citrate synthase	27	0.922	0.094
IPI00383751.1	CALRETICULIN (Calcium binding protein)	2	0.923	0.153
IPI00550882.2	Pyrroline-5-carboxylate reductase 1	1	0.923	0.125
IPI00027285.1	Splice Isoform SM-B' of Small nuclear ribonucleoprotein-associated proteins B and B'	3	0.923	0.100
IPI00006167.1	Protein phosphatase 2C isoform gamma	3	0.923	0.132
IPI00010796.1	Protein disulfide-isomerase precursor	6	0.924	0.009
IPI00031697.1	Transmembrane protein 109 precursor	3	0.924	
IPI00220644.7	Splice Isoform M1 of Pyruvate kinase isozymes M1/M2	26	0.925	0.023
IPI00646917.1	Cleavage and polyadenylation specificity factor 5	2	0.925	0.078
IPI00023640.2	Programmed cell death protein 5	15	0.926	0.207
IPI00293350.3	Translin-associated protein X	2	0.926	0.035
IPI00069750.1	Fuse-binding protein-interacting repressor isoform a	3	0.926	0.044
IPI00004454.2	Dolichyl-phosphate mannosyltransferase polypeptide 3 isoform 1	2	0.926	
IPI00031556.5	Splicing factor U2AF 65 kDa subunit	5	0.926	
IPI00554786.2	Thioredoxin reductase 1, cytoplasmic precursor	2	0.926	0.079
IPI00003479.2	Mitogen-activated protein kinase 1	10	0.926	0.192

IPI00012007.5	Adenosylhomocysteinase	9	0.927	0.128
IPI00027255.1	Myosin light chain 1, slow-twitch muscle A isoform	2	0.927	
IPI00293434.1	Signal recognition particle 14 kDa protein	6	0.927	0.170
IPI00005978.7	Splicing factor, arginine/serine-rich 2	4	0.927	0.132
IPI00009904.1	Protein disulfide-isomerase A4 precursor	11	0.928	0.078
IPI00382412.3	Branched-chain-amino-acid aminotransferase, cytosolic	4	0.928	0.088
IPI00216171.2	Gamma-enolase	7	0.928	0.041
IPI00001639.2	Importin beta-1 subunit	15	0.929	0.036
IPI00177728.3	Cytosolic nonspecific dipeptidase	3	0.929	
IPI00002803.5	Serine/threonine-protein kinase N1	3	0.929	
IPI00010460.1	AH receptor-interacting protein	2	0.929	0.136
IPI00383581.3	Splice Isoform 1 of Neutral alpha-glucosidase AB precursor	2	0.930	0.031
IPI00166768.2	TUBA6 protein	2	0.930	0.077
IPI00179291.4	Cytosolic aminopeptidase P	2	0.930	0.331
IPI00218371.3	Ribose-phosphate pyrophosphokinase III	3	0.931	0.061
IPI00220038.1	Splice Isoform B of Arsenite-resistance protein 2	2	0.932	0.016
IPI00020042.2	Splice Isoform 1 of 26S protease regulatory subunit 6B	8	0.933	0.183
IPI00031461.1	Rab GDP dissociation inhibitor beta	8	0.933	0.111
IPI00180637.4	TIP41, TOR signalling pathway regulator-like isoform 1	2	0.933	0.131
IPI00294178.1	Serine/threonine-protein phosphatase 2A 65 kDa regulator	5	0.933	0.081
IPI00027434.1	Rho-related GTP-binding protein RhoC precursor	5	0.933	0.105
IPI00017964.1	Small nuclear ribonucleoprotein Sm D3	6	0.933	0.025
IPI00007752.1	Tubulin beta-2C chain	2	0.934	0.098
IPI00220637.4	Seryl-tRNA synthetase	20	0.934	0.052
IPI00008527.3	60S acidic ribosomal protein P1	2	0.934	0.250
IPI00024968.1	BM022	3	0.935	
IPI00219483.1	Splice Isoform 2 of U1 small nuclear ribonucleoprotein 70 kDa	2	0.935	
IPI00305442.3	Ubiquitin associated domain containing 1	3	0.935	
IPI00032003.1	Emerin	2	0.935	0.012
IPI00011675.1	Splice Isoform Sp100-HMG of Nuclear autoantigen Sp-100	2	0.935	0.341
IPI00220991.2	Splice Isoform 2 of AP-2 complex subunit beta-1	2	0.936	0.088

IPI00215638.5	ATP-dependent RNA helicase A	20	0.936	0.075
IPI00250297.3	L-aminoadipate-semialdehyde dehydrogenase-phosphopantetheinyl transferase	3	0.936	0.010
IPI00140420.4	Staphylococcal nuclease domain-containing protein 1	15	0.936	0.050
IPI00013890.1	14-3-3 protein sigma	2	0.937	
IPI00028055.4	Transmembrane emp24 domain-containing protein 10 precursor	2	0.937	
IPI00011454.1	Splice Isoform 2 of Neutral alpha-glucosidase AB precursor	3	0.937	0.034
IPI00021435.2	26S protease regulatory subunit 7	2	0.937	0.120
IPI00007188.4	ADP/ATP translocase 2	9	0.937	0.039
IPI00018768.1	Translin	3	0.937	0.108
IPI00025796.3	NADH-ubiquinone oxidoreductase 30 kDa subunit, mitochondrial precursor	2	0.938	0.147
IPI00013871.1	Ribonucleoside-diphosphate reductase large subunit	2	0.938	0.117
IPI00013808.1	Alpha-actinin-4	8	0.938	0.042
IPI00011654.2	Tubulin beta-2 chain	15	0.938	0.017
IPI00027834.3	Heterogeneous nuclear ribonucleoprotein L isoform a	5	0.940	0.251
IPI00465256.3	GTP:AMP phosphotransferase mitochondrial	4	0.940	
IPI00024644.5	Phosphoribosyl transferase domain containing 1 variant	2	0.940	
IPI00334159.6	Von Hippel-Lindau binding protein 1	2	0.940	0.121
IPI00023919.4	26S protease regulatory subunit 8	3	0.940	0.184
IPI00289334.1	Splice Isoform 1 of Filamin-B	11	0.941	0.126
IPI00220301.4	Peroxiredoxin-6	38	0.941	0.169
IPI00295400.1	Tryptophanyl-tRNA synthetase	5	0.942	0.041
IPI00013769.1	Alpha-enolase, lung specific	10	0.943	0.067
IPI00143753.3	PREDICTED: U2-associated SR140 protein	3	0.943	0.071
IPI00003734.1	Putative S100 calcium-binding protein H_NH0456N16.1	3	0.943	0.024
IPI00096066.2	Succinyl-CoA ligase [GDP-forming] beta-chain, mitochondrial	4	0.943	0.171
IPI00004968.1	Pre-mRNA-splicing factor 19	4	0.943	0.139
IPI00009480.1	COP9 signalosome complex subunit 8	2	0.943	
IPI00291165.6	Polyribonucleotide nucleotidyltransferase 1, mitochondrial	4	0.943	

IPI00640981.2	Retinoblastoma-associated factor 600	2	0.943	
IPI00009235.1	Translocon-associated protein gamma subunit	2	0.943	
IPI00180954.4	Cold-inducible RNA-binding protein	4	0.944	
IPI00010105.1	Eukaryotic translation initiation factor 6	5	0.944	0.181
IPI00027107.5	Tu translation elongation factor, mitochondrial	25	0.945	0.105
IPI00178352.4	Splice Isoform 1 of Filamin-C	31	0.945	0.134
IPI00219018.6	Glyceraldehyde-3-phosphate dehydrogenase	40	0.945	0.154
IPI00013679.1	Splice Isoform DUT-M of Deoxyuridine 5'-triphosphate nucleotidohydrolase, mitochondrial precursor	8	0.947	0.248
IPI00172460.3	Splice Isoform 3 of Adenylate kinase isoenzyme 2, mitochondrial	14	0.947	0.125
IPI00216105.1	Splice Isoform 2 of Putative GTP-binding protein PTD004	3	0.947	0.122
IPI00246058.5	PDCD6IP protein	8	0.947	0.165
IPI00221222.6	Activated RNA polymerase II transcriptional coactivator p15	3	0.947	0.052
IPI00025366.4	Citrate synthase, mitochondrial precursor	8	0.948	0.184
IPI00013174.2	Splice Isoform 1 of RNA-binding protein 14	2	0.948	
IPI00298558.2	Programmed cell death protein 10	2	0.948	0.273
IPI00401264.5	Thioredoxin domain-containing protein 4 precursor	5	0.949	0.284
IPI00294398.1	Splice Isoform 1 of Short chain 3-hydroxyacyl-CoA dehydrogenase, mitochondrial precursor	2	0.949	0.248
IPI00024661.4	Protein transport protein Sec24C	2	0.949	0.058
IPI00411706.1	Esterase D	1	0.950	0.092
IPI00453476.2	PREDICTED: similar to Phosphoglycerate mutase 1 (Phosphoglycerate mutase isozyme B) (PGAM-B) (BPG-dependent PGAM 1) isoform 1	16	0.950	0.045
IPI00004416.1	Charged multivesicular body protein 2a	7	0.951	
IPI00299904.3	Splice Isoform 1 of DNA replication licensing factor MCM7	5	0.951	0.011
IPI00219622.2	Proteasome subunit alpha type 2	1	0.952	0.024
IPI00304925.3	Heat shock 70 kDa protein 1	15	0.952	0.094
IPI00220362.4	10 kDa heat shock protein, mitochondrial	14	0.952	0.109
IPI00033030.2	Adhesion-regulating molecule 1 precursor	2	0.952	0.127
IPI00293655.3	ATP-dependent RNA helicase DDX1	11	0.952	
IPI00009949.1	Proteasome inhibitor PI31 subunit	2	0.952	

IPI00005613.2	Splicing factor U2AF 35 kDa subunit	2	0.952	
IPI00044779.1	TC4 protein	7	0.952	0.043
IPI00291922.2	Proteasome subunit alpha type 5	8	0.953	0.107
IPI00215790.5	60S ribosomal protein L38	5	0.953	0.241
IPI00016342.1	Ras-related protein Rab-7	10	0.954	0.047
IPI00024911.1	Endoplasmic reticulum protein ERp29 precursor	5	0.955	0.099
IPI00005198.2	Interleukin enhancer-binding factor 2	5	0.955	0.179
IPI00005102.3	Spermine synthase	3	0.956	0.008
IPI00217661.3	Splice Isoform 2 of Protein raver-1	3	0.957	
IPI00412259.2	PREDICTED: similar to ATP-dependent DNA helicase II, 70 kDa subunit (Lupus Ku autoantigen protein p70) (Ku70) (70 kDa subunit of Ku antigen) (Thyroid-lupus autoantigen) (TLAA) (CTC box binding factor 75 kDa subunit) (CTCBF) (CTC75) isoform 1	6	0.957	0.116
IPI00032831.4	Synaptosomal-associated protein 29	4	0.958	0.082
IPI00023004.6	Eukaryotic translation initiation factor 1A, Y-chromosomal	2	0.959	
IPI00032892.1	Geranylgeranyl pyrophosphate synthetase	4	0.959	0.148
IPI00023542.5	Transmembrane emp24 protein transport domain containing 9	2	0.960	
IPI00003527.4	Ezrin-radixin-moesin-binding phosphoprotein 50	1	0.962	
IPI00555915.1	Heat shock protein 90Bf	1	0.962	
IPI00217053.6	Hypothetical protein	2	0.962	
IPI00155168.2	Protein tyrosine phosphatase, receptor type, C	2	0.962	
IPI00165261.5	Sec1 family domain-containing protein 1	2	0.962	
IPI00008167.1	Sodium/potassium-transporting ATPase beta-3 chain	5	0.962	
IPI00170796.1	Splice Isoform 1 of Vacuolar protein sorting 29	2	0.962	
IPI00256684.1	Splice Isoform B of AP-2 complex subunit alpha-1	3	0.962	
IPI00029601.4	Src substrate cortactin	1	0.962	
IPI00004324.1	Trafficking protein particle complex subunit 3	2	0.962	
IPI00305383.1	Ubiquinol-cytochrome-c reductase complex core protein 2,	8	0.962	
IPI00553131.2	UDP-glucose 4-epimerase	3	0.962	
IPI00009268.1	Aminoacylase-1	2	0.962	0.010
IPI00465430.5	70 kDa protein	9	0.962	0.103
IPI00328343.7	Spliceosome RNA helicase BAT1	2	0.962	0.053

IPI00004839.1	Crk-like protein	3	0.964	0.242
IPI00025252.1	Protein disulfide-isomerase A3 precursor	28	0.964	0.006
IPI00293867.6	D-dopachrome decarboxylase	4	0.964	0.081
IPI00010896.2	Chloride intracellular channel protein 1	10	0.964	0.022
IPI00030783.1	Signal transducer and activator of transcription 5A	6	0.965	0.129
IPI00220766.3	Lactoylglutathione lyase	3	0.965	0.081
IPI00006378.3	Coiled-coil domain-containing protein 72	3	0.965	
IPI00024971.1	Oxysterol-binding protein 1	3	0.965	0.055
IPI00006935.2	Eukaryotic translation initiation factor 5A-2	4	0.966	0.160
IPI00456966.2	26 kDa protein	6	0.966	0.021
IPI00419237.2	Leucine aminopeptidase	3	0.966	0.118
IPI00456635.1	Splice Isoform 1 of Unc-13 homolog D	6	0.967	0.204
IPI00012268.3	26S proteasome non-ATPase regulatory subunit 2	7	0.967	0.097
IPI00106509.2	Splice Isoform 4 of Heterogeneous nuclear ribonucleoprotein A/B	7	0.968	0.152
IPI00004860.2	Arginyl-tRNA synthetase	4	0.968	0.089
IPI00299155.5	Proteasome subunit alpha type 4	5	0.969	0.076
IPI00032826.1	Hsc70-interacting protein	4	0.969	0.085
IPI00029623.1	Proteasome subunit alpha type 6	11	0.969	0.048
IPI00005159.3	Actin-like protein 2	1	0.970	0.071
IPI00184195.2	19 kDa protein	2	0.971	0.148
IPI00008274.5	Adenylyl cyclase-associated protein 1	11	0.971	0.132
IPI00328987.3	Bystin	5	0.971	
IPI00399318.1	Epsilon subunit of coatomer protein complex isoform b	3	0.971	
IPI00006440.4	Mitochondrial ribosomal protein S7	2	0.971	
IPI00025239.2	NADH-ubiquinone oxidoreductase 49 kDa subunit, mitochondrial precursor	2	0.971	
IPI00028481.1	Ras-related protein Rab-8A	2	0.971	
IPI00100197.3	Splice Isoform 1 of NSFL1 cofactor p47	3	0.971	
IPI00464952.2	Splicing factor arginine/serine-rich 11	2	0.971	
IPI00021428.1	Actin, alpha skeletal muscle	4	0.972	0.062
IPI00644079.2	Heterogeneous nuclear ribonucleoprotein U isoform a	2	0.972	
IPI00165579.6	PP856	3	0.972	
IPI00043363.3	Prickle-like protein 1	2	0.972	
IPI00017373.1	Replication protein A 14 kDa subunit	1	0.972	
IPI00008219.1	UV excision repair protein RAD23 homolog A	4	0.972	

IPI00385562.1	Quaking homolog, KH domain RNA binding isoform HQK-7	2	0.972	0.196
IPI00102936.3	Splice Isoform 2 of Signal recognition particle 68 kDa protein	2	0.972	0.151
IPI00298363.2	Far upstream element-binding protein 2	7	0.973	0.031
IPI00171199.4	Splice Isoform 2 of Proteasome subunit alpha type 3	5	0.973	0.037
IPI00303318.2	Protein FAM49B	3	0.973	0.205
IPI00017448.1	40S ribosomal protein S21	3	0.973	0.221
IPI00293464.5	DNA damage-binding protein 1	2	0.974	0.028
IPI00009032.1	Lupus La protein	15	0.974	0.106
IPI00300074.3	Phenylalanyl-tRNA synthetase beta chain	1	0.974	0.164
IPI00019894.2	Splice Isoform 2 of Ubiquitin-conjugating enzyme E2-25 kDa	1	0.975	0.031
IPI00166865.3	Similar to mouse 1500009M05Rik protein	2	0.975	
IPI00007074.4	Tyrosyl-tRNA synthetase, cytoplasmic	5	0.976	0.105
IPI00009071.2	Splice Isoform 3 of FUS-interacting serine-arginine-rich protein 1	4	0.976	0.007
IPI00289499.3	Bifunctional purine biosynthesis protein PURH	18	0.976	0.025
IPI00550308.1	RNA-binding protein 12	2	0.976	0.071
IPI00306960.3	Asparaginyl-tRNA synthetase, cytoplasmic	10	0.976	0.323
IPI00216298.5	Thioredoxin	3	0.977	
IPI00023510.1	Ras-related protein Rab-5A	2	0.977	0.016
IPI00007052.6	Mitochondria fission 1 protein	4	0.977	0.279
IPI00022830.3	Splice Isoform 2 of NSFL1 cofactor p47	2	0.978	0.204
IPI00305668.6	Mitochondrial 28S ribosomal protein S6	2	0.978	
IPI00479786.2	KH-type splicing regulatory protein	2	0.978	0.037
IPI00219129.8	Ribosyldihydronicotinamide dehydrogenase	4	0.979	0.101
IPI00179330.6	Ubiquitin and ribosomal protein S27a precursor	18	0.979	0.071
IPI00002460.2	Splice Isoform 1 of Annexin A7	2	0.980	0.157
IPI00014474.1	A-kinase anchor protein 8	2	0.980	
IPI00022018.1	Dolichol-phosphate mannosyltransferase	2	0.980	
IPI00306516.1	Import inner membrane translocase subunit TIM44, mitocho	4	0.980	
IPI00395887.4	Thioredoxin domain-containing protein 1 precursor	2	0.980	
IPI00045921.1	TOB3	2	0.980	
IPI00016613.2	CSNK2A1 protein	2	0.981	0.124
IPI00374975.2	Probable phosphoglycerate mutase 4	2	0.981	0.049

IPI00031517.1	DNA replication licensing factor MCM6	3	0.982	0.147
IPI00025091.1	40S ribosomal protein S11	12	0.982	0.205
IPI00013508.5	Alpha-actinin-1	7	0.982	0.030
IPI00219365.2	Moesin	19	0.982	0.084
IPI00015602.1	Mitochondrial precursor proteins import receptor	3	0.982	0.165
IPI00024993.4	Enoyl-CoA hydratase, mitochondrial precursor	8	0.983	0.135
IPI00435950.1	FP944	2	0.983	
IPI00171664.2	Nucleoporin 43kDa	2	0.983	
IPI00000690.1	Splice Isoform 1 of Programmed cell death protein 8, mitochondrial precursor	6	0.983	0.097
IPI00016832.1	Splice Isoform Short of Proteasome subunit alpha type 1	1	0.983	0.057
IPI00302176.5	Splice Isoform 1 of H/ACA ribonucleoprotein complex subunit 1	3	0.983	0.068
IPI00140827.2	PREDICTED: similar to SMT3 suppressor of mif two 3 homolog 2	2	0.984	0.198
IPI00003348.2	Guanine nucleotide-binding protein G(I)/G(S)/G(T) beta subunit 2	5	0.984	0.044
IPI00031522.2	Trifunctional enzyme alpha subunit, mitochondrial precursor	27	0.984	0.325
IPI00009901.1	Nuclear transport factor 2	1	0.984	
IPI00022314.1	Superoxide dismutase [Mn], mitochondrial precursor	2	0.984	0.043
IPI00000816.1	14-3-3 protein epsilon	14	0.985	0.035
IPI00384051.4	Proteasome activator complex subunit 2	8	0.986	0.064
IPI00008557.3	Insulin-like growth factor 2 mRNA binding protein 1	5	0.986	0.025
IPI00026216.4	Puromycin-sensitive aminopeptidase	6	0.986	0.126
IPI00169280.2	Aflatoxin B1 aldehyde reductase member 4 (Fragment)	2	0.987	0.236
IPI00010154.3	Rab GDP dissociation inhibitor alpha	8	0.987	0.020
IPI00013452.8	Glutamyl-prolyl tRNA synthetase	11	0.988	0.111
IPI00155723.3	DEAH (Asp-Glu-Ala-His) box polypeptide 9 isoform 2	5	0.988	
IPI00025512.2	Heat-shock protein beta-1	15	0.988	0.056
IPI00550363.2	Transgelin-2	3	0.989	0.034
IPI00217236.3	Tubulin-specific chaperone A	9	0.989	0.041
IPI00550689.3	HSPC117 protein	3	0.990	0.028
IPI00022793.4	Trifunctional enzyme beta subunit, mitochondrial	3	0.990	0.325

	precursor			
IPI00026519.1	Peptidyl-prolyl cis-trans isomerase, mitochondrial precursor	2	0.990	
IPI00100460.2	Aspartyl-tRNA synthetase 2	2	0.990	
IPI00029264.3	Cytochrome c1, heme protein, mitochondrial precursor	6	0.990	
IPI00162199.8	DKFZP434B0335 protein	1	0.990	
IPI00260769.3	Hypothetical protein	4	0.990	
IPI00641574.1	Isoleucine-tRNA synthetase	2	0.990	
IPI00554769.1	Mutant GSTP1	2	0.990	
IPI00329600.3	Probable saccharopine dehydrogenase	5	0.990	
IPI00514380.1	Pyrroline-5-carboxylate reductase family, member 2	2	0.990	
IPI00027233.3	SCO1 protein homolog, mitochondrial precursor	3	0.990	
IPI00001676.7	Splice Isoform 2 of Nuclear protein localization protein	2	0.990	
IPI00168235.4	CDNA FLJ33352 fis, clone BRACE2005087, weakly similar to PRE-MRNA SPLICING HELICASE BRR2	6	0.990	0.014
IPI00012066.1	Poly(rC)-binding protein 2 isoform b	5	0.991	0.044
IPI00010740.1	Splice Isoform Long of Splicing factor, proline- and glutamine-rich	16	0.991	0.125
IPI00399212.2	PREDICTED: similar to Ran-specific GTPase-activating protein	5	0.992	0.251
IPI00013895.1	Protein S100-A11	2	0.992	
IPI00302592.2	Filamin A, alpha	42	0.993	0.012
IPI00165360.4	3-mercaptopyruvate sulfurtransferase	2	0.993	0.287
IPI00376317.3	Autoantigen RCD8	2	0.995	0.022
IPI00382470.2	Heat shock protein HSP 90-alpha 2	20	0.995	0.063
IPI00018349.5	DNA replication licensing factor MCM4	6	0.996	0.087
IPI00026665.1	Glutaminyl-tRNA synthetase	2	0.998	0.168
IPI00375704.1	Proteasome subunit beta type (Fragment)	8	0.998	0.067
IPI00106668.4	Splice Isoform A of Mannose-6-phosphate receptor-binding protein	6	0.999	0.159
IPI00291928.7	Ras-related protein Rab-14	5	0.999	0.113
IPI00020436.3	Ras-related protein Rab-11B	11	0.999	0.098
IPI00413451.1	Hypothetical protein DKFZp686I04222	2	0.999	0.119
IPI00015148.3	Ras-related protein Rap-1b precursor	3	1.000	0.097
IPI00025019.3	Proteasome subunit beta type 1	9	1.000	0.122

IPI00172594.3	39S ribosomal protein L28, mitochondrial precursor	3	1.000	
IPI00019912.2	Peroxisomal multifunctional enzyme type 2	2	1.000	
IPI00165230.1	Splice Isoform 1 of DAZ-associated protein 1	3	1.000	
IPI00298887.5	Splice Isoform 1 of Signal transducer and activator of transcription 3	2	1.000	
IPI00022202.3	Splice Isoform A of Phosphate carrier protein, mitochondial	5	1.000	
IPI00029048.2	Tubulin--tyrosine ligase-like protein 12	5	1.000	
IPI00293276.9	Macrophage migration inhibitory factor	10	1.000	0.091
IPI00031169.1	Ras-related protein Rab-2A	3	1.001	0.040
IPI00020599.1	Calreticulin precursor	6	1.001	0.063
IPI00016513.5	Ras-related protein Rab-10	2	1.001	0.141
IPI00215914.4	ADP-ribosylation factor 1	3	1.001	0.126
IPI00299571.4	CDNA FLJ45525 fis, clone BRTHA2026311, highly similar to Protein disulfide isomerase A6	2	1.002	0.029
IPI00009149.2	Splice Isoform 2 of Suppressor of G2 allele of SKP1	2	1.002	0.087
IPI00291175.6	Splice Isoform 1 of Vinculin	51	1.002	0.010
IPI00002966.1	Heat shock 70 kDa protein 4	8	1.002	0.031
IPI00289819.4	Cation-independent mannose-6-phosphate receptor precursor	1	1.002	0.103
IPI00219447.6	Serine/threonine-protein kinase PAK 2	5	1.003	0.259
IPI00034319.2	Splice Isoform A of Protein CutA precursor	2	1.003	0.086
IPI00218570.5	Phosphoglycerate mutase 2	10	1.004	0.071
IPI00010414.3	PDZ and LIM domain protein 1	2	1.004	0.352
IPI00022881.1	Splice Isoform 1 of Clathrin heavy chain 2	7	1.004	0.155
IPI00441344.1	Beta-galactosidase precursor	8	1.004	
IPI00020906.1	Inositol monophosphatase	1	1.004	
IPI00168184.7	CDNA FLJ34068 fis, clone FCBBF3001918, highly similar to SERINE/THREONINE PROTEIN PHOSPHATASE 2A, 65 kDa REGULATORY SUBUNIT A, ALPHA ISOFORM	6	1.004	0.121
IPI00012069.1	NAD(P)H dehydrogenase [quinone] 1	2	1.005	0.152
IPI00003420.1	Splice Isoform 1 of Microtubule-associated protein RP/EB family member 2	2	1.005	0.158
IPI00030770.1	Splice Isoform 1 of Down syndrome critical region protein 2	1	1.006	0.166

IPI00549189.3	Thimet oligopeptidase	2	1.007	0.019
IPI00220014.2	Isopentenyl-diphosphate delta isomerase	2	1.007	0.299
IPI00478861.3	55 kDa protein	4	1.008	0.111
IPI00100160.3	Splice Isoform 1 of Cullin-associated NEDD8-dissociated protein 1	10	1.008	0.209
IPI00012345.2	Splice Isoform SRP55-1 of Splicing factor, arginine/serine-rich 6	2	1.009	
IPI00024175.3	Splice Isoform 1 of Proteasome subunit alpha type 7	2	1.010	0.024
IPI00218200.7	B-cell receptor-associated protein 31	6	1.010	0.316
IPI00007068.1	Actin-related protein 3-beta isoform 1	2	1.010	
IPI00419307.2	Alpha isoform of regulatory subunit A, protein phosphata	6	1.010	
IPI00218782.2	Capping protein (Actin filament) muscle Z-line, beta	2	1.010	
IPI00395425.3	HSPC260	2	1.010	
IPI00170877.1	Mitochondrial ribosomal protein L10 isoform b	3	1.010	
IPI00328293.2	Serine\/arginine repetitive matrix 1	2	1.010	
IPI00031608.1	Spermatogenesis associated 5-like 1	2	1.010	
IPI00020194.1	Splice Isoform Short of TATA-binding protein-associated factor 2N	3	1.010	
IPI00413895.2	Golgi autoantigen, golgin subfamily A member 2	2	1.010	0.014
IPI00514082.1	Isoleucine-tRNA synthetase	2	1.011	0.160
IPI00295857.6	Coatomer subunit alpha	4	1.011	0.059
IPI00298520.3	Hypothetical protein DKFZp686M09245	2	1.012	0.152
IPI00072377.1	Splice Isoform 1 of Protein SET	7	1.012	0.128
IPI00000792.1	Quinone oxidoreductase	11	1.012	0.042
IPI00178083.2	29 kDa protein	5	1.012	0.026
IPI00215610.2	55 kDa erythrocyte membrane protein	7	1.013	0.171
IPI00028004.2	Proteasome subunit beta type 3	8	1.013	0.089
IPI00013939.3	Splice Isoform 1 of Replication protein A 32 kDa subunit	3	1.013	0.088
IPI00030911.2	Vesicle-associated membrane protein 8	3	1.013	
IPI00182856.10	Splice Isoform 3 of WD-repeat protein 1	2	1.014	0.017
IPI00170692.4	Vesicle-associated membrane protein-associated protein A isoform 2	3	1.014	0.172
IPI00000567.1	HSPC269 (Fragment)	2	1.015	
IPI00007694.4	Protein phosphatase methylesterase-1	5	1.015	0.136
IPI00006558.3	Splice Isoform 1 of SH3 domain GRB2-like	2	1.015	

	protein B1			
IPI00217561.1	Splice Isoform Beta-1C of Integrin beta-1 precursor	2	1.015	
IPI00018206.3	Aspartate aminotransferase, mitochondrial precursor	5	1.015	0.019
IPI00216318.4	14-3-3 protein beta/alpha	9	1.016	0.105
IPI00216691.4	Profilin-1	17	1.016	0.060
IPI00298961.3	Exportin-1	3	1.017	0.086
IPI00291412.1	Ca(2+)/calmodulin-dependent protein kinase phosphatase	3	1.017	0.231
IPI00099730.3	Splicing coactivator subunit SRm300	3	1.018	0.168
IPI00014230.1	Complement component 1, Q subcomponent-binding protein, mitochondrial precursor	8	1.018	0.011
IPI00145540.7	PREDICTED: similar to tropomyosin 3 isoform 2	2	1.018	0.080
IPI00217966.6	Lactate dehydrogenase A	1	1.019	0.218
IPI00395769.2	Splice Isoform Heart of ATP synthase gamma chain, mitochondrial precursor	2	1.019	0.050
IPI00010779.3	Tropomyosin 4	2	1.020	0.112
IPI00026087.1	Barrier-to-autointegration factor	2	1.020	0.145
IPI00026530.4	ERGIC-53 protein precursor	2	1.020	0.291
IPI00013234.2	IARS protein	10	1.020	0.074
IPI00215919.4	ADP-ribosylation factor 5	2	1.020	
IPI00456429.2	Ubiquitin and ribosomal protein L40 precursor	2	1.020	
IPI00018146.1	14-3-3 protein theta	7	1.020	0.126
IPI00185146.4	Importin-9	4	1.020	
IPI00658000.1	Insulin-like growth factor 2 mRNA binding protein 3	2	1.020	
IPI00182180.1	OTU domain-containing protein 6B	2	1.020	
IPI00098827.1	Placental protein 25	2	1.020	
IPI00003870.1	Putative ATP-dependent Clp protease proteolytic subunit, mitochondrial	3	1.020	
IPI00008380.1	Serine/threonine-protein phosphatase 2A catalytic subuni	4	1.020	
IPI00006025.1	Splice Isoform 1 of Squamous cell carcinoma antigen recognized by T-cells 3	9	1.020	
IPI00006451.5	Vesicle-fusing ATPase	2	1.020	
IPI00298547.3	Protein DJ-1	10	1.021	0.049
IPI00026546.1	Platelet-activating factor acetylhydrolase IB beta	2	1.021	0.080

	subunit			
IPI00000861.1	Splice Isoform 1 of LIM and SH3 domain protein 1	4	1.021	0.200
IPI00017367.5	Radixin	8	1.022	0.079
IPI00007682.2	Vacuolar ATP synthase catalytic subunit A, ubiquitous isoform	5	1.023	0.108
IPI00456969.1	Dynein heavy chain, cytosolic	14	1.023	0.358
IPI00215901.1	Adenylate kinase 2 isoform a	3	1.023	0.178
IPI00003588.1	Eukaryotic translation elongation factor 1 epsilon-1	3	1.023	0.100
IPI00185919.3	Splice Isoform 1 of La-related protein 1	4	1.024	0.197
IPI00163187.9	Fascin	14	1.024	0.098
IPI00218493.6	Hypoxanthine-guanine phosphoribosyltransferase	3	1.024	0.083
IPI00418262.3	Fructose-bisphosphate aldolase C	2	1.025	0.120
IPI00022745.1	Diphosphomevalonate decarboxylase	6	1.026	
REV_25025_Protein	no description	6	1.026	
IPI00216256.2	Splice Isoform 2 of WD-repeat protein 1	1	1.026	
IPI00013214.1	DNA replication licensing factor MCM3	7	1.026	0.015
IPI00011126.6	26S protease regulatory subunit 4	4	1.026	0.079
IPI00183920.1	BA395L14.12	5	1.028	0.233
IPI00294879.1	Ran GTPase-activating protein 1	2	1.028	0.435
IPI00299095.1	Sorting nexin-2	2	1.028	0.063
IPI00063234.1	PRKAR2A protein	2	1.028	0.011
IPI00303882.2	Splice Isoform B of Mannose-6-phosphate receptor-binding protein 1	2	1.029	0.174
IPI00008603.1	Actin, aortic smooth muscle	2	1.029	0.165
IPI00219953.5	Cytidylate kinase	4	1.030	0.104
IPI00070643.6	Splice Isoform Long of FAS-associated factor 1	1	1.031	0.015
IPI00003815.1	Rho GDP-dissociation inhibitor 1	7	1.031	0.236
IPI00641706.1	46 kDa protein	1	1.031	
IPI00005160.2	Actin-related protein 2/3 complex subunit 1B	2	1.031	
IPI00011200.4	D-3-phosphoglycerate dehydrogenase	2	1.031	
IPI00301107.5	Importin-11	2	1.031	
IPI00295625.2	Succinyl-CoA ligase [GDP-forming] alpha-chain, mitochondrial	7	1.031	
IPI00220739.2	Membrane-associated progesterone receptor component 1	1	1.031	0.203
IPI00021263.3	14-3-3 protein zeta/delta	18	1.031	0.060

IPI00296913.1	ADP-sugar pyrophosphatase	3	1.031	0.190
IPI00027175.1	Sorcin	3	1.033	0.060
IPI00294911.1	Succinate dehydrogenase [ubiquinone] iron-sulfur protein, mitochondrial precursor	2	1.033	0.241
IPI00470467.4	NADPH--cytochrome P450 reductase	2	1.034	0.208
IPI00216008.3	Splice Isoform Long of Glucose-6-phosphate 1-dehydrogenase	11	1.034	0.073
IPI00304692.1	Heterogeneous nuclear ribonucleoprotein G	5	1.034	0.327
IPI00018671.1	Dual specificity protein phosphatase 3	2	1.035	
IPI00027382.1	Splice Isoform 2 of Serine/threonine-protein kinase PAK 3	2	1.035	0.154
IPI00183274.2	Sorting nexin 1 isoform c	2	1.035	0.088
IPI00007321.1	Splice Isoform 1 of Acyl-protein thioesterase 1	3	1.035	0.080
IPI00002134.3	26S proteasome non-ATPase regulatory subunit 5	9	1.035	0.180
IPI00027444.1	Leukocyte elastase inhibitor	13	1.035	0.180
IPI00479877.4	4-trimethylaminobutyraldehyde dehydrogenase	5	1.036	
IPI00006663.1	Aldehyde dehydrogenase, mitochondrial precursor	2	1.036	
IPI00003921.2	Splice Isoform 1 of Protein 4.1	1	1.036	0.106
IPI00016621.6	AP-2 complex subunit alpha-2	1	1.037	0.099
IPI00418823.3	PREDICTED: similar to lactate dehydrogenase A-like 6B	2	1.037	0.258
IPI00013297.1	28 kDa heat- and acid-stable phosphoprotein	2	1.037	0.207
IPI00021812.1	Neuroblast differentiation-associated protein AHNAK (Fragment)	18	1.037	0.257
IPI00219029.2	Aspartate aminotransferase, cytoplasmic	7	1.037	0.106
IPI00026089.3	Splicing factor 3B subunit 1	7	1.038	0.020
IPI00024919.3	Thioredoxin-dependent peroxide reductase, mitochondrial precursor	9	1.039	0.037
IPI00022977.1	Creatine kinase B-type	13	1.039	0.026
IPI00017341.2	Splicing factor 3A subunit 2	2	1.039	0.207
IPI00029267.1	U2 small nuclear ribonucleoprotein B"	5	1.040	0.084
IPI00103994.4	Leucyl-tRNA synthetase, cytoplasmic	4	1.040	0.119
IPI00005162.2	Actin-related protein 2/3 complex subunit 3	2	1.040	
IPI00239077.4	Histidine triad nucleotide-binding protein 1	11	1.041	0.106
IPI00027497.4	Glucose-6-phosphate isomerase	22	1.041	0.041
IPI00060200.3	Aldose 1-epimerase	2	1.042	
IPI00386354.1	CDNA FLJ14048 fis, clone HEMBA1006650, weakly similar to ARP2/3 COMPLEX 20 KD	3	1.042	

	SUBUNIT			
IPI00005087.1	Tropomodulin-3	3	1.042	
IPI00024067.3	Clathrin heavy chain 1	22	1.042	0.024
IPI00018931.6	Vacuolar protein sorting 35	8	1.042	0.001
IPI00479186.4	Pyruvate kinase 3 isoform 1	5	1.043	0.110
IPI00037448.3	Glyoxylate reductase/hydroxypyruvate reductase	7	1.043	0.115
IPI00291006.1	Malate dehydrogenase, mitochondrial precursor	4	1.043	0.062
IPI00034049.1	Splice Isoform 1 of Regulator of nonsense transcripts 1	3	1.044	0.130
IPI00012837.1	Kinesin heavy chain	6	1.045	0.128
IPI00015955.5	Exosome complex exonuclease RRP46	2	1.045	0.226
IPI00296053.3	Fumarate hydratase, mitochondrial precursor	6	1.045	0.036
IPI00021327.3	Splice Isoform 1 of Growth factor receptor-bound protein 2	3	1.045	
IPI00440493.2	ATP synthase alpha chain, mitochondrial precursor	25	1.045	0.031
IPI00062206.1	DEAD (Asp-Glu-Ala-Asp) box polypeptide 39, isoform 2	2	1.046	0.105
IPI00033494.3	Myosin regulatory light chain	2	1.046	0.146
IPI00395865.4	Histone-binding protein RBBP7	2	1.047	
IPI00328415.10	NADH-cytochrome b5 reductase	2	1.047	0.121
IPI00219291.3	Splice Isoform 2 of ATP synthase f chain, mitochondrial	4	1.047	
IPI00021439.1	Actin, cytoplasmic 1	28	1.048	0.120
IPI00643920.2	Transketolase	11	1.049	0.113
IPI00021347.1	Ubiquitin-conjugating enzyme E2 L3	3	1.050	0.023
IPI00003856.1	Vacuolar ATP synthase subunit E	2	1.050	0.111
IPI00017672.4	Hypothetical protein FLJ25678	5	1.051	0.025
IPI00303476.1	ATP synthase beta chain, mitochondrial precursor	9	1.051	0.031
IPI00022774.2	Transitional endoplasmic reticulum ATPase	28	1.051	0.028
IPI00007423.1	Acidic leucine-rich nuclear phosphoprotein 32 family member B	9	1.051	0.180
IPI00646779.2	TUBB6 protein	2	1.052	0.325
IPI00001663.1	Splice Isoform 1 of Serine protease HTRA2, mitochondrial	2	1.053	
IPI00028006.1	Proteasome subunit beta type 2	2	1.053	0.218
IPI00169383.2	Phosphoglycerate kinase 1	25	1.054	0.074
IPI00219446.4	Phosphatidylethanolamine-binding protein	11	1.054	0.068
IPI00025054.1	Splice Isoform Long of Heterogenous nuclear	2	1.055	0.071

	ribonucleoprotein U			
IPI00218693.7	Adenine phosphoribosyltransferase	9	1.055	0.091
IPI00103467.3	Aldehyde dehydrogenase X, mitochondrial precursor	2	1.058	
IPI00025974.3	Charged multivesicular body protein 4b	2	1.058	
IPI00604756.2	Nuclear receptor-binding protein	2	1.058	
IPI00009790.1	6-phosphofructokinase type C	3	1.058	0.227
IPI00147874.1	Sialic acid synthase	4	1.059	0.026
IPI00334627.2	Similar to annexin A2 isoform 1	4	1.059	0.190
IPI00016610.2	Poly(rC)-binding protein 1	12	1.060	0.014
IPI00016786.1	Splice Isoform 2 of Cell division control protein 42 homolog precursor	6	1.061	0.101
IPI00012998.2	Splice Isoform 1 of Cleavage and polyadenylation specificity factor 6	2	1.062	0.115
IPI00218848.4	ATP synthase e chain, mitochondrial	2	1.063	0.161
IPI00027443.4	Cysteinyl-tRNA synthetase isoform c	3	1.064	0.160
IPI00026215.1	Flap endonuclease 1	7	1.064	
IPI00010090.1	Glutamate--cysteine ligase regulatory subunit	4	1.064	
IPI00006213.1	Pericentriol material 1	2	1.064	
IPI00413587.2	Splice Isoform 1 of BH3-interacting domain death agonist	6	1.064	
IPI00024502.2	Ubiquilin-4	2	1.064	
IPI00000051.4	Prefoldin subunit 1	2	1.065	0.277
IPI00010270.1	Ras-related C3 botulinum toxin substrate 2 precursor	4	1.065	0.065
IPI00337741.4	Acylamino-acid-releasing enzyme	4	1.065	0.105
IPI00029468.1	Alpha-centractin	2	1.066	0.212
IPI00552186.2	Hypothetical protein DKFZp313O211	3	1.066	0.050
IPI00166749.3	Mitochondrial-processing peptidase alpha subunit, mitochondrial precursor	1	1.068	0.135
IPI00012011.5	Cofilin-1	21	1.068	0.038
IPI00479359.6	villin 2	2	1.068	0.174
IPI00098902.4	Oxoglutarate (alpha-ketoglutarate) dehydrogenase (lipoamide) isoform 1 precursor	2	1.068	
IPI00013862.7	Thymidylate kinase	4	1.069	
IPI00030919.3	Mitogen-activated protein kinase kinase 1-interacting protein 1	3	1.070	
IPI00258804.1	ATPase inhibitory factor 1 isoform 3 precursor	2	1.070	
IPI00000811.2	Proteasome subunit beta type 6 precursor	6	1.070	0.070

IPI00216951.2	Aspartyl-tRNA synthetase	2	1.070	0.300
IPI00220834.7	ATP-dependent DNA helicase 2 subunit 2	2	1.070	0.105
IPI00328319.7	Histone-binding protein RBBP4	5	1.071	0.057
IPI00179953.2	Splice Isoform 1 of Nuclear autoantigenic sperm protein	8	1.073	0.091
IPI00004397.1	Ras-related protein Ral-B	2	1.074	0.079
IPI00002324.4	Methionine adenosyltransferase II, beta isoform 1	2	1.074	0.143
IPI00465028.6	Triosephosphate isomerase	5	1.074	0.073
IPI00008418.5	Splice Isoform 1 of Diablo homolog, mitochondrial precursor	2	1.075	0.021
IPI00165393.1	Acidic leucine-rich nuclear phosphoprotein 32 family member E	5	1.075	0.117
IPI00020672.4	Splice Isoform 1 of Dipeptidyl-peptidase 3	5	1.075	0.061
IPI00009841.4	CDNA FLJ31747 fis, clone NT2RI2007377, highly similar to RNA-BINDING PROTEIN EWS	2	1.075	
IPI00168317.1	Hypothetical protein FLJ90806	2	1.075	
IPI00005050.1	Mitochondrial 28S ribosomal protein S14	2	1.075	
IPI00032139.1	Serpin B9	4	1.075	
IPI00221106.5	Splicing factor 3B subunit 2	8	1.076	0.087
IPI00419585.8	Peptidyl-prolyl cis-trans isomerase A	3	1.077	0.089
IPI00216697.2	Ankyrin 1 isoform 1	4	1.077	0.050
IPI00025849.1	Acidic leucine-rich nuclear phosphoprotein 32 family member A	11	1.077	0.105
IPI00248321.8	PREDICTED: similar to peptidylprolyl isomerase A isoform 1	2	1.077	0.123
IPI00383046.2	Hypothetical protein LOC134147	2	1.079	0.031
IPI00301518.5	Mob4B protein	1	1.079	
IPI00514832.2	Tripartite motif protein 56	1	1.079	
IPI00020495.1	Mitochondrial 28S ribosomal protein S36	7	1.080	
IPI00007102.3	CGI-150 protein	2	1.080	0.085
IPI00215884.3	Splicing factor, arginine/serine-rich 1	4	1.080	0.309
IPI00025084.3	Calpain small subunit 1	3	1.081	0.029
IPI00003217.3	Proteasome subunit beta type 7 precursor	2	1.082	0.095
IPI00032830.1	Splice Isoform 1 of Oligoribonuclease, mitochondrial precursor (Fragment)	2	1.082	0.082
IPI00030229.4	UDP-galactose-4-epimerase	2	1.084	0.112
IPI00305304.3	LAG1 longevity assurance homolog 2	2	1.084	0.075
IPI00397526.1	Myosin-10	4	1.084	0.213

IPI00017726.1	Hydroxyacyl-Coenzyme A dehydrogenase, type II isoform 1	10		
IPI00027442.3	Alanyl-tRNA synthetase	3	1.084	0.166
IPI00013122.1	Hsp90 co-chaperone Cdc37	3	1.086	0.038
IPI00099871.1	39S ribosomal protein L40, mitochondrial precursor	2	1.087	
IPI00154975.3	DnaJ homolog subfamily C member 9	3	1.087	
IPI00023086.3	Mitochondrial ribosomal protein L15	3	1.087	
IPI00007928.4	Pre-mRNA-processing-splicing factor 8	2	1.087	
IPI00028083.1	Translation initiation factor eIF-2B beta subunit	3	1.087	
IPI00019385.1	Translocon-associated protein delta subunit precursor	2	1.087	
IPI00291419.5	Acetyl-CoA acetyltransferase, cytosolic	2	1.088	0.072
IPI00550069.2	Ribonuclease inhibitor	4	1.089	0.208
IPI00290614.7	Endonuclease G, mitochondrial precursor	2	1.090	0.194
IPI00038378.4	E-1 enzyme	2	1.091	0.318
IPI00451401.2	Splice Isoform 2 of Triosephosphate isomerase	11	1.092	0.023
IPI00018188.3	Splice Isoform 1 of Adaptin ear-binding coat-associated protein 2	2	1.092	0.102
IPI00219757.12	Glutathione S-transferase P	14	1.093	0.022
IPI00221328.2	Chloride intracellular channel protein 2	3	1.093	0.053
IPI00006865.2	Vesicle trafficking protein SEC22b	8	1.094	
IPI00005719.1	RAB1A, member RAS oncogene family	6	1.095	0.205
IPI00073603.3	Peptidyl-prolyl cis-trans isomerase	2	1.095	0.219
IPI00027341.1	Macrophage capping protein	2	1.095	0.219
IPI00003269.1	Hypothetical protein LOC345651	2	1.097	0.226
IPI00304082.7	ISOC1 protein	3	1.098	0.072
IPI00304814.3	Phospholipid hydroperoxide glutathione peroxidase, mitochondrial precursor	2	1.099	
IPI00470610.3	Pyrroline-5-carboxylate reductase 2	3	1.099	
IPI00299086.3	Syntenin-1	3	1.099	
IPI00025815.1	TAR DNA-binding protein 43	4	1.099	
IPI00017451.1	Splicing factor 3 subunit 1	5	1.099	0.120
IPI00168388.1	Splice Isoform 1 of Signal recognition particle 68 kDa protein	3	1.099	0.033
IPI00037283.2	Splice Isoform 5 of Dynamin-1-like protein	2	1.099	0.034
IPI00303954.3	Cytochrome b5 outer mitochondrial membrane precursor	4	1.100	
IPI00027165.3	Splice Isoform R-type of Pyruvate kinase	10	1.101	0.107

	isozymes R/L			
IPI00063827.1	Splice Isoform 1 of Abhydrolase domain-containing protein 14B	2	1.101	0.150
IPI00003438.1	DnaJ (Hsp40) homolog, subfamily C, member 8	2	1.102	0.089
IPI00030116.1	Splice Isoform 1 of Phosphoacetylglucosamine mutase	2	1.102	0.345
IPI00383071.1	RcTPI1 (Fragment)	15	1.103	0.038
IPI00001159.9	GCN1-like protein 1	9	1.104	0.083
IPI00024157.1	FK506-binding protein 3	2	1.104	0.089
IPI00017855.1	Aconitate hydratase, mitochondrial precursor	3	1.105	0.050
IPI00017510.3	Cytochrome c oxidase subunit 2	4	1.105	
IPI00026185.5	Splice Isoform 1 of F-actin capping protein beta subunit	7	1.107	0.062
IPI00007189.1	Splice Isoform 1 of Cell division control protein 42 homolog precursor	2	1.110	0.271
IPI00219219.2	Galectin-1	5	1.110	0.052
IPI00165467.5	55 kDa protein	4	1.111	
IPI00017381.1	Activator 1 37 kDa subunit	2	1.111	
IPI00023503.3	Cell division protein kinase 3	2	1.111	
IPI00008453.3	Coronin-1C	2	1.111	
IPI00014938.2	Nuclear protein Hcc-1	7	1.111	
IPI00010080.2	Serine/threonine-protein kinase OSR1	2	1.111	
IPI00217952.6	Splice Isoform 1 of Glucosamine--fructose-6-phosphate aminotransferase [isomerizing] 1	2	1.111	
IPI00021983.1	Splice Isoform 1 of Nicastrin precursor	2	1.111	
IPI00176678.8	RcADH5 (Fragment)	3	1.112	0.052
IPI00514212.2	Mps one binder kinase activator-like 1A	1	1.114	0.170
IPI00019755.3	Glutathione transferase omega-1	9	1.114	0.076
IPI00550488.2	TALDO1 protein	5	1.114	0.066
IPI00031708.1	Fumarylacetoacetase	5	1.116	0.092
IPI00184330.5	DNA replication licensing factor MCM2	10	1.116	0.133
IPI00010706.1	Glutathione synthetase	2	1.119	0.101
IPI00009368.3	Sideroflexin-1	2	1.119	0.062
IPI00010810.1	Electron transfer flavoprotein alpha-subunit, mitochondrial precursor	8	1.119	0.128
IPI00030154.1	Proteasome activator complex subunit 1	8	1.120	0.042
IPI00015911.1	Dihydrolipoyl dehydrogenase, mitochondrial precursor	12	1.120	0.074

IPI00000230.5	Tropomyosin 1 alpha chain isoform 2	2	1.120	0.167
IPI00293721.3	Aflatoxin B1 aldehyde reductase member 3	2	1.122	
IPI00418169.2	Annexin A2 isoform 1	4	1.123	0.079
IPI00020956.1	Hepatoma-derived growth factor	2	1.123	0.198
IPI00004902.1	Electron-transfer-flavoprotein, beta polypeptide isoform 1	9	1.123	0.013
IPI00385156.1	D-myo-inositol-3-phosphate synthase	2	1.124	0.026
IPI00217920.4	Aldehyde dehydrogenase 16 family, member A1	2	1.124	
IPI00008986.1	Large neutral amino acids transporter small subunit 1	2	1.124	
IPI00005158.1	Lon protease homolog, mitochondrial precursor	8	1.124	
IPI00027252.6	Prohibitin-2	11	1.124	
IPI00305152.5	SEC31L1 protein	1	1.124	
IPI00017469.1	Sepiapterin reductase	2	1.124	
IPI00220421.2	Splice Isoform 2 of Centaurin-delta 2	2	1.124	
IPI00001146.1	U6 snRNA-associated Sm-like protein LSm6	2	1.124	
IPI00030363.1	Acetyl-CoA acetyltransferase, mitochondrial precursor	5	1.124	0.129
IPI00004503.5	LAMP1 protein	3	1.124	0.094
IPI00024871.1	Core-binding factor, beta subunit isoform 1	1	1.125	
IPI00014808.1	Platelet-activating factor acetylhydrolase IB gamma subunit	3	1.125	0.035
IPI00000015.2	Splicing factor, arginine/serine-rich 4	2	1.125	0.373
IPI00009342.1	Ras GTPase-activating-like protein IQGAP1	1	1.127	0.180
IPI00015550.6	Prothymosin alpha	2	1.128	0.248
IPI00410214.1	Splice Isoform 1 of 3'(2'),5'-bisphosphate nucleotidase	2	1.128	0.130
IPI00291005.7	Malate dehydrogenase, cytoplasmic	1	1.128	0.053
IPI00294158.1	Biliverdin reductase A precursor	4	1.129	0.074
IPI00022430.1	Glyceraldehyde-3-phosphate dehydrogenase, testis-specific	2	1.132	
IPI00030144.1	Peptidyl-prolyl cis-trans isomerase	3	1.133	0.130
IPI00220487.3	ATP synthase, H+ transporting, mitochondrial F0 complex, subunit d isoform a	6	1.133	0.187
IPI00017592.1	Leucine zipper-EF-hand-containing transmembrane protein 1, mitochondrial precursor	4	1.136	0.101
IPI00014439.3	Dihydropteridine reductase	2	1.136	0.020
IPI00647400.1	68 kDa protein	2	1.136	0.086
IPI00016634.1	Protein C20orf11	2	1.136	

IPI00014718.4	RcDNAJ9 (Fragment)	4	1.136	
IPI00033022.2	Splice Isoform 1 of Dynamin-2	3	1.136	
IPI00006684.3	Splice Isoform 4 of Apoptosis inhibitor 5	3	1.136	
IPI00012585.1	Beta-hexosaminidase beta chain precursor	2	1.140	0.140
IPI00008964.3	Ras-related protein Rab-1B	6	1.141	0.003
IPI00744692.1	Transaldolase	11	1.143	0.094
IPI00019450.2	Splice Isoform Long of 60-kDa SS-A/Ro ribonucleoprotein	3	1.144	0.188
IPI00010471.3	Plastin-2	4	1.144	0.122
IPI00013004.1	Splice Isoform 1 of Pyridoxal kinase	2	1.146	0.068
IPI00219669.4	Carbonic anhydrase-related protein	8	1.146	0.125
IPI00003406.2	Drebrin	2	1.148	0.166
IPI00011285.1	Calpain-1 catalytic subunit	7	1.149	
IPI00328715.4	Protein LYRIC	2	1.149	
IPI00024466.1	UDP-glucose ceramide glucosyltransferase-like 1 isoform	2	1.149	
IPI00016801.1	Glutamate dehydrogenase 1, mitochondrial precursor	6	1.152	0.124
IPI00007346.1	Peptidyl-prolyl cis-trans isomerase H	3	1.152	
IPI00003168.1	Phosphoribosyl pyrophosphate synthetase-associated protein 2	4	1.153	
IPI00021828.1	Cystatin B	5	1.154	0.013
IPI00329801.11	Annexin A5	7	1.157	0.078
IPI00018871.1	ADP-ribosylation factor-like protein 8B	5	1.160	0.054
IPI00456887.2	PREDICTED: similar to heterogeneous nuclear ribonucleoprotein U	5	1.160	0.170
IPI00007058.1	Coronin-1B	3	1.163	0.124
IPI00023530.6	Cell division protein kinase 5	1	1.163	
IPI00027485.3	Eukaryotic translation initiation factor 4E	2	1.163	
IPI00414384.1	Hydroxysteroid dehydrogenase-like protein	2	1.163	
IPI00002412.1	Palmitoyl-protein thioesterase 1 precursor	3	1.163	
IPI00014177.3	Septin-2	2	1.163	
IPI00216348.1	Splice Isoform 2C of Cytoplasmic dynein 1 intermediate chain 2	1	1.163	
IPI00022597.1	NEDD8-conjugating enzyme Ubc12	6	1.164	
IPI00419258.3	High mobility group protein B1	4	1.164	0.092
IPI00007611.1	ATP synthase O subunit, mitochondrial precursor	8	1.165	0.213
IPI00012382.2	U1 small nuclear ribonucleoprotein A	4	1.168	0.228
IPI00033600.1	Yeast sds22 homolog	2	1.171	0.067

IPI00218733.5	Superoxide dismutase	5	1.171	0.107
IPI00168479.2	Apolipoprotein A-I binding protein precursor	2	1.172	0.132
IPI00024989.7	Protein-L-isoaspartate (D-aspartate) O-methyltransferase	7	1.174	0.072
IPI00027350.2	Peroxiredoxin-2	1	1.174	0.078
IPI00011268.2	RNA binding protein (Fragment)	3	1.174	
IPI00555956.1	Proteasome subunit beta type 4 precursor	5	1.174	0.359
IPI00007019.1	Peptidyl-prolyl cis-trans isomerase-like 1	1	1.175	
IPI00008982.1	Splice Isoform Long of Delta 1-pyrroline-5-carboxylate synthetase	2	1.175	0.072
IPI00018452.1	Copine-1	2	1.175	0.106
IPI00023529.1	Cell division protein kinase 6	4	1.175	
IPI00299608.3	Splice Isoform 1 of 26S proteasome non-ATPase regulatory subunit 1	4	1.175	0.360
IPI00298994.5	Talin-1	37	1.176	0.033
IPI00001960.3	Chloride intracellular channel protein 4	6	1.176	
IPI00010402.2	Hypothetical protein	2	1.176	
IPI00294619.1	Protein TFG	2	1.176	
IPI00179589.2	Myotrophin	4	1.177	0.012
IPI00005668.4	Aldo-keto reductase family 1 member C2	3	1.177	0.044
IPI00300567.1	Splice Isoform 1 of 3,2-trans-enoyl-CoA isomerase, mitochondrial precursor	6	1.178	0.154
IPI00301109.3	Splice Isoform 1 of Inorganic pyrophosphatase 2, mitochondrial precursor	4	1.179	0.383
IPI00216699.1	Splice Isoform 2 of Unc-112-related protein 2	3	1.181	0.098
IPI00645078.1	Ubiquitin-activating enzyme E1	5	1.184	0.063
IPI00002745.1	Cathepsin Z precursor	2	1.184	0.345
IPI00292140.4	Caspase-3 precursor	2	1.186	
IPI00479997.3	Stathmin	14	1.186	0.047
IPI00002048.2	PREDICTED: aldo-keto reductase, truncated	2	1.187	0.089
IPI00294578.1	Splice Isoform 1 of Protein-glutamine gamma-glutamyltransferase 2	2	1.188	0.333
IPI00010120.4	C-terminal-binding protein 2	2	1.190	
IPI00183503.7	Novel protein	3	1.190	
IPI00025874.2	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase 67 kDa subunit precursor	3	1.194	0.326
IPI00642211.3	Aminopeptidase B	1	1.194	0.062
IPI00641319.1	Ubiquitin-activating enzyme E1	2	1.194	0.064
IPI00026119.6	Ubiquitin-activating enzyme E1	6	1.195	0.066

IPI00030357.5	Dihydrofolate reductase	2	1.198	0.194
IPI00465436.3	Catalase	8	1.200	0.066
IPI00107531.1	Splice Isoform 3 of DNA repair protein RAD50	4	1.202	0.379
IPI00552452.1	Ubiquitin-activating enzyme E1	2	1.204	0.027
IPI00329742.1	Fumarylacetoacetate hydrolase domain containing 2A	2	1.205	
IPI00297635.5	SA hypertension-associated homolog isoform 1	2	1.205	
IPI00300026.5	Sulfotransferase 1A1	2	1.205	
IPI00297477.3	U2 small nuclear ribonucleoprotein A'	4	1.205	
IPI00218667.2	Stathmin-2	3	1.205	0.027
IPI00007001.1	39S ribosomal protein L11, mitochondrial precursor	2	1.206	
IPI00604707.3	Dihydrolipoamide S-acetyltransferase (Fragment)	2	1.206	0.002
IPI00217143.2	SDHA protein	2	1.214	0.127
IPI00060181.1	EF-hand domain-containing protein 2	3	1.216	0.330
IPI00008454.1	DnaJ homolog subfamily B member 11 precursor	2	1.220	
IPI00013774.1	Histone deacetylase 1	2	1.220	
IPI00465373.1	Kynurenine aminotransferase III isoform 1	2	1.220	
IPI00017256.5	Ras suppressor protein 1	2	1.220	
IPI00032134.1	Serpin B8	2	1.220	
IPI00023785.5	Splice Isoform 1 of Probable ATP-dependent RNA helicase	2	1.220	
IPI00306825.3	Splice Isoform 1 of Tumor protein D54	2	1.220	
IPI00296337.2	Splice Isoform 1 of DNA-dependent protein kinase catalytic subunit	16	1.225	0.104
IPI00174852.4	TRF-proximal protein homolog	2	1.227	
IPI00021997.1	CREG1 protein precursor	2	1.228	0.150
IPI00157820.2	Splice Isoform 2 of Thioredoxin reductase 2, mitochondrial precursor	2	1.229	
IPI00374732.3	PREDICTED: similar to peptidylprolyl isomerase A isoform 1	2	1.229	0.174
IPI00181135.4	Splice Isoform B of Branched-chain-amino-acid aminotransferase, mitochondrial precursor	1	1.232	0.368
IPI00014198.2	Exosome complex exonuclease RRP42	2	1.235	
IPI00465006.1	Hypothetical protein DKFZp686D19113	2	1.235	
IPI00017412.1	Splice Isoform 1 of Activator 1 40 kDa subunit	2	1.235	
IPI00025156.4	Splice Isoform 1 of STIP1 homology and U box-containing protein 1	3	1.235	
IPI00005979.2	Splice Isoform 1 of Translation initiation factor	3	1.235	

	eIF-2B			
IPI00305978.4	Aflatoxin B1 aldehyde reductase member 2	3	1.238	0.270
REV_20262_Protein	no description	2	1.239	
IPI00024913.1	Splice Isoform Long of ES1 protein homolog, mitochondrial precursor	5	1.240	0.029
IPI00007926.1	c-Myc-responsive protein Rcl	3	1.246	0.079
IPI00219299.2	Talin-2	5	1.250	0.027
IPI00220503.9	Dynactin 2	4	1.250	
IPI00005040.1	Medium-chain specific acyl-CoA dehydrogenase, mitochondrial precursor	1	1.251	0.244
IPI00029133.4	ATP synthase B chain, mitochondrial precursor	4	1.256	0.367
IPI00010860.1	Splice Isoform p27-L of 26S proteasome non-ATPase regulatory subunit 9	2	1.257	0.075
IPI00300086.3	Nicotinate-nucleotide pyrophosphorylase	10	1.258	0.034
IPI00010865.1	Casein kinase II subunit beta	1	1.266	
IPI00017283.2	Isoleucyl-tRNA synthetase, mitochondrial precursor	6	1.266	
IPI00303568.3	Prostaglandin E synthase 2	1	1.266	
IPI00014587.1	Splice Isoform Brain of Clathrin light chain A	1	1.266	
IPI00026516.1	Succinyl-CoA:3-ketoacid-coenzyme A transferase 1, mitochondrial precursor	2	1.268	0.110
IPI00171798.1	Metastasis-associated protein MTA2	4	1.269	0.223
IPI00216319.2	14-3-3 protein eta	3	1.272	0.047
IPI00218918.4	Annexin A1	31	1.278	0.097
IPI00427330.2	Shwachman-Bodian-Diamond syndrome protein	2	1.278	0.473
IPI00219575.5	Bleomycin hydrolase	3	1.282	
IPI00028109.1	Dpy-30-like protein	2	1.282	
REV_20581_Protein	no description	2	1.282	
IPI00018246.4	Splice Isoform 1 of Hexokinase-1	3	1.282	
IPI00169400.1	Splice Isoform 1 of Mitochondrial 28S ribosomal protein S5	2	1.282	
IPI00010271.3	Splice Isoform A of Ras-related C3 botulinum toxin substrate 1 precursor	2	1.290	
IPI00448095.3	L-xylulose reductase	2	1.290	
IPI00220740.1	Splice Isoform 2 of Nucleophosmin	36	1.296	0.321
IPI00031107.1	HSDL2 protein	4	1.299	
IPI00301994.5	Hypothetical protein DKFZp434N062	3	1.299	

IPI00009747.1	Lanosterol synthase	2	1.299	
IPI00465044.2	Protein RCC2	1	1.299	
IPI00104050.3	Thyroid hormone receptor-associated protein 3	2	1.301	0.072
IPI00029997.1	6-phosphogluconolactonase	2	1.303	0.468
IPI00003031.1	Hypothetical protein FLJ23469	2	1.303	
IPI00025318.1	SH3 domain-binding glutamic acid-rich-like protein	3	1.307	
IPI00019502.2	Myosin-9	20	1.309	0.403
IPI00414320.1	Annexin A11	2	1.313	0.229
IPI00001352.1	21 kDa protein	3	1.315	
IPI00219841.6	DNA ligase 1	2	1.316	
IPI00181231.9	Hypothetical protein LOC55037	2	1.316	
IPI00186711.3	Plectin 1 isoform 6	2	1.316	
IPI00171412.1	Splice Isoform 1 of Sulfatase-modifying factor 2 precursor	3	1.316	
IPI00003482.1	2,4-dienoyl-CoA reductase, mitochondrial precursor	7	1.318	0.247
IPI00384992.6	Myosin light polypeptide 4	2	1.318	0.121
IPI00011229.1	Cathepsin D precursor	3	1.319	0.106
IPI00029733.1	Aldo-keto reductase family 1 member C1	2	1.326	0.215
IPI00018342.5	Adenylate kinase isoenzyme 1	3	1.329	0.141
IPI00219910.1	Flavin reductase	18	1.329	0.076
IPI00021338.1	Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex, mitochondrial precursor	2	1.330	0.092
IPI00020530.1	Thioesterase superfamily member 2	4	1.331	0.176
IPI00010219.1	Spindle pole body component 25	3	1.333	
IPI00013475.1	Tubulin, beta 2	4	1.333	
IPI00003925.5	Splice Isoform 1 of Pyruvate dehydrogenase E1 component subunit beta, mitochondrial precursor	4	1.339	0.186
IPI00011416.2	Enoyl Coenzyme A hydratase 1, peroxisoma	3	1.339	0.167
IPI00011107.2	Isocitrate dehydrogenase [NADP], mitochondrial precursor	5	1.342	0.179
IPI00024915.2	Peroxiredoxin-5, mitochondrial precursor	2	1.344	0.148
IPI00013860.3	3-hydroxyisobutyrate dehydrogenase, mitochondrial precursor	2	1.347	0.108
IPI00002459.4	Annexin VI isoform 2	5	1.357	
IPI00301489.3	Uroporphyrinogen decarboxylase	6	1.360	0.096
IPI00335168.8	Myosin, light polypeptide 6, alkali, smooth	2	1.364	0.322

	muscle and non-muscle isoform 1			
IPI00219097.3	High mobility group protein B2	2	1.375	0.075
IPI00022334.1	Ornithine aminotransferase, mitochondrial precursor	4	1.385	0.031
IPI00304435.3	Protein NipSnap1	3	1.389	
IPI00017305.2	Ribosomal protein S6 kinase alpha-1	3	1.389	
IPI00028031.2	Splice Isoform 1 of Very-long-chain specific acyl-CoA dehydrogenase, mitochondrial precursor	3	1.395	0.369
IPI00061507.1	Hypothetical protein C11orf54	2	1.400	
IPI00220271.2	Alcohol dehydrogenase	8	1.400	0.224
IPI00017184.2	EH-domain-containing protein 1	2	1.405	0.023
IPI00179700.3	High mobility group AT-hook 1 isoform a	2	1.408	
IPI00218235.4	Dehydrogenase/reductase (SDR family) member 2 isoform 2		1.416	
IPI00060627.1	Hypothetical protein LOC115098	2	1.429	
IPI00328257.4	Splice Isoform A of AP-1 complex subunit beta-1	3	1.429	
IPI00021389.1	Copper chaperone for superoxide dismutase	2	1.430	0.027
IPI00031681.1	Cell division protein kinase 2	1	1.449	
IPI00177817.4	Splice Isoform SERCA2A of Sarcoplasmic/endoplasmic reticulum calcium ATPase 2	2	1.449	
IPI00179713.5	Insulin-like growth factor 2 mRNA binding protein 2	2	1.471	
IPI00059762.5	Lysophospholipase-like 1	1	1.471	
IPI00011494.3	Methyltransferase-like protein 1 isoform c	2	1.471	
IPI00294380.4	Phosphoenolpyruvate carboxykinase (GTP) family protein	3	1.471	
IPI00291930.5	Splice Isoform 1 of Epsin-4	1	1.471	
IPI00216308.4	Voltage-dependent anion-selective channel protein 1	3	1.471	
IPI00155601.1	Protein LRP16	4	1.493	0.460
IPI00216125.6	SRP9 protein	2	1.495	
IPI00020127.1	Replication protein A 70 kDa DNA-binding subunit	2	1.496	
IPI00156689.3	Synaptic vesicle membrane protein VAT-1 homolog	6	1.500	0.097
IPI00295741.4	Cathepsin B precursor	6	1.503	0.223
IPI00217357.2	Cell division cycle and apoptosis regulator protein	2	1.507	0.111

	1			
IPI00027776.6	Ferrochelatase, mitochondrial precursor	5	1.537	0.215
IPI00024145.1	Splice Isoform 1 of Voltage-dependent anion-selective chanel 2	4	1.566	0.030
IPI00220835.6	Protein transport protein Sec61 beta subunit	1	1.570	0.025
IPI00218568.6	Pterin-4-alpha-carbinolamine dehydratase	1	1.581	0.108
IPI00021924.1	Histone H1x	4	1.588	0.200
IPI00642144.2	Aldehyde dehydrogenase 1 family, member A1	2	1.592	0.238
IPI00744503.1	17 kDa protein	1	1.598	0.174
IPI00093057.6	Coproporphyrinogen III oxidase, mitochondrial precursor	5	1.609	0.314
IPI00028160.1	Splice Isoform 1 of Porphobilinogen deaminase	8	1.830	0.070
IPI00337541.3	NAD(P) transhydrogenase, mitochondrial precursor	1	1.905	0.026
IPI00219038.8	Histone H3.3	1	2.111	0.219
IPI00220706.9	Hemoglobin gamma-1 subunit	4	2.210	0.830
IPI00217465.4	Histone H1.2	10	2.255	0.470
IPI00030809.1	Gamma-G globin	6	2.344	0.552
IPI00003935.5	Histone H2B	7	2.347	0.695
IPI00410714.4	Hemoglobin alpha subunit	6	2.357	0.756
IPI00449049.4	Poly [ADP-ribose] polymerase 1	7	2.425	0.421
IPI00171611.5	Histone H3.1	4	2.608	0.407
IPI00217471.2	Hemoglobin epsilon subunit	5	2.745	0.765
IPI00026272.1	Histone H2A	4	2.749	0.699
IPI00554676.1	Hemoglobin gamma-2 subunit	1	2.837	0.583
IPI00217473.4	Hemoglobin zeta subunit	12	3.025	1.182
IPI00453473.5	Histone H4	8	3.062	0.818

Table S3.2 Detailed quantification results for proteins with substantial changes.

IPI number	Protein name	Peptide number	Forward 1	Reverse 2	Reverse 3	Average
IPI00032561.1	Calcium-binding protein 39 R.DLKRPAQQE.A K.VFVANPNKTQPILDILLK.N	2	0.353		0.438	0.396
IPI00002214.1	Importin alpha-2 subunit R.QDQIQQVVNHGLVPFLVSVLSKADFK.T K.TGVVPQLVKLLGASELPIVTPALR.A K.LLGASELPIVTPALR.A K.TGVVPQLVK.L R.NKNPAPPIDAVEQILPTLVR.L R.EKQPPIDNIIR.A	6	0.505	0.535	0.524	0.521
IPI00221108.4	Thymidylate synthase -.PVAGSELPR.R -.PVAGSELPRRPLPPAAQER.D	2	0.485	0.513	0.569	0.522
IPI00011631.5	Centromere/kinetochore protein zw10 homolog K.AMGTLLNTAISEVIGK.I	1		0.493	0.609	0.493
IPI00010349.1	Alkyldihydroxyacetonephosphate synthase, peroxisomal precursor K.NIYGNIEDLVVHIK.M K.GFDPNQLSVATLLFEGDREK.V	2	0.504	0.642	0.524	0.557
IPI00013068.1	Eukaryotic translation initiation factor 3 subunit 6 K.NLYSDDIPHALR.E R.TTVVAQLKQLQAETEPIVK.M	2	0.685		0.450	0.567

IPI00005045.1	ATP-binding cassette sub-family F member 2 R.VALVGPNGAGK.S R.ILHGLGFTPAMQR.K K.SMLLSAIGKR.E K.LLTGELLPTDGMIR.K R.EVPIPEHIDIYHLTR.E	5	0.568	0.534	0.605	0.569
IPI00026182.4	F-actin capping protein alpha-2 subunit K.FIIHAPPGEFNEVFNDVR .L K.FTITPSTTQVVGILK.I	2	0.508		0.652	0.580
IPI00218505.6	Protein FAM112B K.LATCPFNAR.H K.SLPYVLPWKNNGNA.Q	2	0.611	0.595	0.550	0.585
IPI00337602.4	56 kDa protein R.ESVLTATSILNNPIVK.A	1		0.651	0.524	0.588
IPI00218292.2	Splice Isoform Short of Ubiquitin fusion degradation protein R.LNITYPMLFK.L R.AFSGSGNRLDGK.K R.FVAFSGEGQSLR.K K.IIMPPSALDQLSR.L	4	0.758		0.438	0.598
IPI00306043.1	Splice Isoform 1 of YTH domain protein 2 K.LGSTEVASNVPK.V	1	0.588	0.611	0.602	0.600
IPI00009010.3	UPF0315 protein AD-001 R.ICPVEFNPNFVAR.M R.GIPNMLLSEEETE.S	2	0.660	0.549	0.609	0.606
IPI00743544.1	15 kDa protein K.LYTLVTYVTVTTFK.N	1		0.624	0.591	0.607

IPI00220113.1	Splice Isoform 2 of Microtubule-associated protein 4 K.NVVLPTETEVAPAKDVTLLKETER.A K.TTTLSGTAPAAGVVPSRVK.A R.LATNTSAPDLKNVR.S K.KPMSLASGLVPAAPPK.R K.VGSLDNVGHLPAGGAVK.T R.ASPSKPASAPASR.S K.TTTAAAVASTGPSSR.S R.SPSTLLPK.K K.KPTSAKPSSTTPR.L R.LATNTSAPDLK.N K.TEAAATTRKPESNAVTK.T	11	0.568	0.629	0.630	0.609
IPI00183294.3	Nuclear pore complex protein Nup214 R.TPSIQPSLLPHAAPFAK.S K.TPHPVLTPVAANQAK.Q K.QGSLINSLKPSGPTPASGQLSSGDKASGTAK.I K.ASSTSLTSTQPTK.T K.TSGVPSGFNFTAPPVLGK.H R.IFDSPEELPKER.S K.AAPGPGPSTFSFVPPSK.A	7	0.439	0.743	0.650	0.610
IPI00221035.3	Splice Isoform 1 of Transcription factor BTF3 K.VQASLAANTFTITGHAETK.Q K.QLTEMLPSILNQLGADSLTSLR.R K.QLTEMLPSILNQLGADS	3	0.559	0.702	0.577	0.612

	LTSLRR.L					
IPI00291669.3	Ubiquitin-like domain containing CTD phosphatase 1 K.TLTGVLPER.Q R.GLIDVKPLGVIWGK.F	2	0.699	0.568	0.588	0.618
IPI00305166.1	Succinate dehydrogenase [ubiquinone] flavoprotein subunit, mitochondrial precursor R.GVIALCIEDGSIHR.I	1		0.590	0.652	0.621
IPI00329625.2	Cell cycle progression 2 protein isoform 1 K.LLGSLYALGIPK.A K.ELQSVEQEVR.W	2	0.794		0.449	0.621
IPI00021266.1	60S ribosomal protein L23a R.LAPDYDALDVANK.I R.LAPDYDALDVANKIGI.I K.KEAPAPPKAEAK.A K.EAPAPPKAEAK.A R.TSPTFR.R K.KLYDIDVAK.V K.LYDIDVAKVNTLIRPDGEK.K K.VNTLIRPDGEK.K K.VNTLIRPDGEKK.A	9	0.613	0.764	0.510	0.629
IPI00009057.1	Splice Isoform A of Ras-GTPase-activating protein-binding protein 2 R.VEAKPEVQSQPPR.V K.GVGGKLPNFGFVVFDDSEPVQR.I R.VEAKPEVQSQPPRVR.E K.VLSLNFSECHTK.I R.VEAKPEVQSQPPRVR.E R.ILIAKPIMFR.G	6	0.629	0.675	0.589	0.631

IPI00176692.7	PREDICTED: similar to Heterogeneous nuclear ribonucleoprotein A1 R.EDSQRPGAHLTVK.K R.GFAFVTFDDHDSVDKIVIQK.Y R.AVSREDSQRPGAHLTVK.K R.GFAFVTFDDHDSVDK.I	4	0.775	0.773	0.353	0.634
IPI00029266.1	Small nuclear ribonucleoprotein E K.VMVQPINLIFR.Y R.IMLKGDNITLLQSVS.N	2	0.368	0.763	0.780	0.637
IPI00017895.2	Splice Isoform 1 of Glycerol-3-phosphate dehydrogenase, mitochondrial precursor R.VIFFLPWK.M R.LAFLNVQAAEEALPR.I R.YGAATANYMEVVSLLK.K	3	0.573	0.708	0.630	0.637
IPI00012535.1	DnaJ homolog subfamily A member 1 K.NVVHQLSVTLEDLYNGATR.K K.QISQAYEVLSDAK.K R.TIVITSHPGQIVK.H	3	0.813	0.608	0.502	0.641
IPI00299573.11	60S ribosomal protein L7a K.KVAPAPAVVK.K K.VAPAPAVVK.K R.LKVPPAINQFTQALDR.Q R.QTATQLLK.L R.AGVNTVTTLVENK.K R.AGVNTVTTLVENKK.A K.VPPAINQFTQALDR.Q K.TCTTVAFTQVNSEDKG	14	0.758	0.711	0.470	0.646

	ALAK.L R.HWGGNVLGPK.S K.VAPAPAVVKKQEAK.K R.LKVPPAINQFTQALDRQ TATQLLK.L K.ELATKL.G R.RHWGGNVLGPK.S R.KTCTTVAFTQVNSEDKG ALAK.L					
IPI00026260.1	Nucleoside diphosphate kinase B K.YMNSGPVVAMVWEGL NVVK.T K.SCAHDWVY.E K.SAEKEISLWFKPEELVD YK.S R.NIIHGSDSVK.S K.DRPFFPGLVK.Y	5		0.489	0.806	0.648
IPI00013415.1	40S ribosomal protein S7 R.KAIIIFVPVPQLK.S K.AIIIFVPVPQLK.S R.TLTAVHDAILEDLVFPSE IVGK.R R.TLTAVHDAILEDLVFPSE IVGKR.I K.DVNFEFPEFQ.L R.SRTLTAVHDAILEDLVF PSEIVGKR.I K.KLTGKDVNFEFPEFQ.L K.LTGKDVNFEFPEFQ.L	8	0.714	0.435	0.801	0.650
IPI00027493.1	4F2 cell-surface antigen heavy chain R.IGDLQAFQGHGAGNLA GLK.G R.LLTSFLPAQLLR.L K.GQSEDPGSLLSLFR.R K.ADLLLSTQPGREEGSPL ELER.L	5	0.704	0.700	0.556	0.653

	R.LKLEPHEGLLLRFPYA.A					
IPI00016910.1	Eukaryotic translation initiation factor 3 subunit 8 K.SIVDKEGVPR.F R.AKELLGQGLLLR.S K.VWDLFPEADKVR.T K.ELLGQGLLLR.S R.TEPTAQQNLALQLAEKL GSLVENNER.V K.RLDEEEEDNEGGEWER VR.G K.LGSLVENNER.V K.GTEITHAVVIKKLNEILQ AR.G K.LNEILQAR.G K.DAHNALLDIQSSGR.A R.TEPTAQQNLALQLAEKL GSLVENNERVFDHK.Q	11	0.680	0.708	0.577	0.655
IPI00186290.5	Elongation factor 2 R.YVEPIEDVPCGNIVGLV GVDQFLVK.T R.IMGPNYTPGKK.E K.KEDLYLKPIQR.T K.EDLYLKPIQR.T K.TGTITTFEHAHNMR.V R.TILMMGR.Y R.RWLPAGDALLQMITIHL PSPVTAQK.Y R.VMKFSVSPVVR.V R.WLPAGDALLQMITIHLP SPVTAQK.Y K.FSVSPVVR.V R.CELLYEGPPDDEAAMGI K.S R.VAVEAKNPADLPK.L K.SCDPKGPLMMYISK.M K.GPLMMYISK.M R.VFSGLVSTGLK.V R.YFDPANGKFSK.S	31	0.649	0.688	0.630	0.656

	R.TFCQLILDPIFK.V K.QFAEMYVAK.F R.IKPVLMMNKMDR.A K.FAAKGEGQLGPAER.A R.ALLELQLEPEELYQTFQ R.I R.IKPVLMMNK.M K.GEGQLGPAER.A K.STAISLFYELSENDLNFI KQSK.D K.STAISLFYELSENDLNFI K.Q K.AGIIASAR.A K.STLTDSLVCKAGIIASAR .A K.STLTDSLVCK.A R.NMSVIAHVDHGKSTLT DSLVCK.A R.NMSVIAHVDHGKSTLT DSLVCKAGIIASAR.A -.VNFTVDQIR.A					
IPI00012479.1	FKSG17 K.NILFVITKPDVYK.S	1	0.617		0.705	0.661
IPI00219155.4	60S ribosomal protein L27 K.VVLVLAGR.Y R.YSVDIPLDK.T	2	0.565	0.738	0.682	0.662
IPI00013495.1	Splice Isoform 2 of ATP-binding cassette sub-family F K.TLLIVSHDQGFLDDVCT DIIHLDAQR.L R.FTFPDPPPLSPPVLGLHG VTFGYQGQKPLFK.N	2	0.662	0.625	0.701	0.663
IPI00030179.3	60S ribosomal protein L7 K.EVPAVPETLK.K K.EVPAVPETLKK.K R.KAGNFYVPAEPK.L	6	0.709	0.789	0.491	0.663

	K.AGNFYVPAEPK.L K.ASINMLR.I K.SVNELIYKR.G					
IPI00017617.1	Probable ATP-dependent RNA helicase DDX5 K.QVSDLISVLR.E R.RTAQEVETYRR.S K.LLQLVEDR.G R.GDGPICLVLAPTR.E R.ELAQQVQQVAAEYCR.A K.STCIYGGAPK.G R.GVEICIATPGR.L R.LIDFLECGKTNLRR.T R.DWVLNEFKHGKAPILIA TDVASR.G R.EANQAINPK.L K.QVSDLISVLREANQAIN PK.L K.APILIATDVASR.G K.TGTAYTFFTPNNIKQVS DLISVLR.E	13	0.585	0.639	0.769	0.664
IPI00023748.3	Nascent polypeptide-associated complex alpha subunit K.NILFVITKPDVYKSPASD TYIVFGEAK.I K.SPASDTYIVFGEAK.I K.IEDLSQQAQLAAAEKFK .V R.ALKNNSNDIVNAIMELT. M K.SPASDTYIVFGEAKIEDL SQQAQLAAAEKFK.V K.IEDLSQQAQLAAAEK.F K.DIELVMSQANVSR.A	7	0.641	0.715	0.641	0.666
IPI00220373.3	Insulin-degrading enzyme R.ESLDDLTNLVVK.L	2		0.705	0.630	0.668

IPI00219678.2	Eukaryotic translation initiation factor 2 subunit 1 R.ADIEVACYGYEGIDAVK EALR.A K.VVTDTDETELAR.Q	2	0.642	0.581	0.780	0.668
IPI00397828.2	20 kDa protein R.LGGPEAGLGEYLFER.L	1	0.651	0.658	0.694	0.668
IPI00444704.2	Splice Isoform 2 of G-rich sequence factor 1 R.GLPFQANAQDIINFFAPL KPVR.I	1		0.674	0.662	0.668
IPI00037619.2	Ribosomal protein L39 variant K.QNRPIPQWIR.M	1	0.694	0.699	0.611	0.668
IPI00219229.2	U6 snRNA-associated Sm-like protein LSm3 R.GDGVVLVAPPLR.V R.GDGVVLVAPPLRV.G	2	0.725	0.552	0.728	0.668
IPI00008475.1	Hydroxymethylglutaryl-CoA synthase, cytoplasmic R.IGVFSYGSGLAATLYSL K.V R.IGVFSYGSGLAATLYSL KVTQDATPGSALDKITASL CDLK.S K.VTQDATPGSALDKITAS LCDLK.S R.TGVAPDVFAENMK.L R.RPTPNDDTLDEGVGLVH SNIATEHIPSPAK.K R.NNLSYDCIGRLEVGTETI IDKSK.S R.LEVGTETIIDKSK.S R.PTGGVGAVALLIGPNAP LIFER.G K.ITASLCDLK.S	12	0.671	0.875	0.459	0.669

	K.LSIQCYLSALDR.C R.MLLNDFLNDQNR.D K.ASSELFSQK.T					
IPI00010157.1	S-adenosylmethionine synthetase isoform type-2 K.AVVPAKYLDEDTIYHLQPSGR.F R.FVIGGPQGDAGLTGR.K K.NFDLRPGVIVR.D R.DLDLKKPIYQR.T	4	0.676	0.674	0.658	0.669
IPI00000728.3	Splice Isoform 1 of Ubiquitin carboxyl-terminal hydrolase 15 R.FSKADTIDTIEKEIR.K R.KIFSIPDEKETR.L R.TLEVYLVR.M R.NNTEDKLYNLLLLR.M K.AAYVLFYQR.Q	5	0.787	0.622	0.598	0.669
IPI00156689.3	Synaptic vesicle membrane protein VAT-1 homolog K.VVTYGMANLLTGPK.R R.LLALYNQGHIKPHIDSVWPFEK.V K.VLLVPGPEKE.N R.PAAPPAPGPGQLTLR.L R.ACGLNFADLMAR.Q K.LQSRPAAPPAPGPGQLTLR.L	6	1.538	1.573	1.390	1.500
IPI00295741.4	Cathepsin B precursor K.LPASFDAREQWPQCPTIK.E K.LPASFDAR.E R.EQWPQCPTIK.E R.EQWPQCPTIKEIR.D K.ILRGQDHCGIESEVVAGIPR.T R.GQDHCGIESEVVAGIPR.	6	1.587	1.672	1.250	1.503

	T					
IPI00217357.2	Cell division cycle and apoptosis regulator protein 1 R.IQTLPNQNQSQTQPLLK.T K.AGLLQPPVR.I	2	1.429		1.586	1.507
IPI00027776.6	Ferrochelatase, mitochondrial precursor R.IGGGSPIKIWTSK.Q K.LLDELSPNTAPHKYYIGFR.Y R.YVHPLTEEAIEEMERDGLER.A R.GDPYPQEVSATVQK.V K.ALADLVHSHIQSNELCSK.Q	5	1.754	1.532	1.324	1.537
IPI00024145.1	Splice Isoform 1 of Voltage-dependent anion-selective chanel 2 R.NNFAVGYR.T K.YQLDPTASISAK.V K.VNNSSLIGVGYTQTLRPGVK.L K.LTLSALVDGK.S	4	1.587		1.545	1.566
IPI00220835.6	Protein transport protein Sec61 beta subunit -.PGPTPSGTNVGSSGR.S	1		1.587	1.552	1.570
IPI00218568.6	Pterin-4-alpha-carbinolamine dehydratase R.LSAEERDQLLPNLR.A	1	1.470	1.587	1.685	1.581
IPI00021924.1	Histone H1x K.YSQLVVETIRR.L K.ALVQNDTLLQVK.G R.GAPAAATAPAPTAHK.A R.RGAPAAATAPAPTAHK.	4	1.786	1.385	1.592	1.588

	A					
IPI00642144.2	Aldehyde dehydrogenase 1 family, member A1 R.QAFQIGSPWR.T R.YCAGWADKIQGR.T	2	1.587	1.832	1.357	1.592
IPI00744503.1	17 kDa protein K.VLTSLGDATK.H	1	1.721		1.474	1.598
IPI00093057.6	Coproporphyrinogen III oxidase, mitochondrial precursor R.ATSLGRPEEEEDELAHR.C R.AVVPSYIPLVK.K R.GIGGIFFDDLDSPSKEEVFR.F R.IESILMSLPLTAR.W K.EAEILEVLR.H	5	1.639	1.907	1.282	1.609
IPI00028160.1	Splice Isoform 1 of Porphobilinogen deaminase R.IQTDSVVATLK.A R.ENPHDAVVFHPK.F K.TLETLPEKSVVGTSSLR.R R.GPQLAAQNLGISLANLLLSK.G R.KFPHLEFR.S R.KLDEQQEFSAIILATAGLQR.M R.AKDQDILDLVGVLHDPETLLR.C R.NIPRGPQLAAQNLGISLANLLLSK.G	8	1.887	1.850	1.752	1.830
IPI00337541.3	NAD(P) transhydrogenase, mitochondrial precursor R.VALSPAGVQNLVK.Q	1	1.887		1.923	1.905

IPI00219038.8	Histone H3.3 R.FQSAAIGALQEASEAYL VGLFEDTNLCAIHAK.R	1	2.083	2.343	1.907	2.111
IPI00220706.9	Hemoglobin gamma-1 subunit -.GHFTEEDKATITSLWGK. V -.GHFTEEDKATITSLWGK VNVEDAGGETLGR.L K.MVTAVASALSSR.Y K.MVTAVASALSSRY.H	4	2.000	3.125	1.506	2.210
IPI00217465.4	Histone H1.2 K.SLVSKGTLVQTK.G R.KASGPPVSELITK.A K.ASGPPVSELITK.A K.ASGPPVSELITKAVAAS KER.S R.SGVSLAALK.K R.SGVSLAALKK.A R.SGVSLAALKKALAAAG YDVEK.N K.KALAAAGYDVEKNNSR .I K.ALAAAGYDVEK.N K.ALAAAGYDVEKNNSR.I	10	2.632	1.728	2.404	2.255
IPI00030809.1	Gamma-G globin (Fragment) K.VNVEDAGGETLGR.L R.LLVVYPWTQR.F R.LLVVYPWTQRFFDSFGN LSSASAIMGNPK.V R.FFDSFGNLSSASAIMGNP K.V K.VLTSLGDAIK.H K.EFTPEVQASWQK.M	6	2.128	2.971	1.934	2.344
IPI00003935.5	Histone H2B K.AMGIMNSFVNDIFER.I R.STITSREIQTAVR.L R.EIQTAVRLLLPGELAK.H	7	3.030	2.370	1.641	2.347

	R.EIQTAVRLLLPGELAKH AVSEGTK.A R.LLLPGELAKHAVSEGTK .A K.QVHPDTGISSK.A R.EIQTAVR.L					
IPI00410714.4	Hemoglobin alpha subunit -.VLSPADKTNVK.A K.AAWGKVGAHAGEYGA EALER.M K.VGAHAGEYGAEALER. M R.MFLSFPTTK.T K.TYFPHFDLSHGSAQVK. G K.VADALTNAVAHVDDM PNALSALSDLHAHK.L	6	2.222	3.171	1.677	2.357
IPI00449049.4	Poly [ADP-ribose] polymerase 1 R.TTNFAGILSQGLR.I K.TLGDFAAEYAK.S K.GFSLLATEDKEALKK.Q K.SLQELFLAHILSPWGAE VKAEPVEVVAPR.G K.KFYPLEIDYGQDEEAVK .K K.LTVNPGTK.S K.MIFDVESMKK.A	7	2.632	2.703	1.941	2.425
IPI00171611.5	histone H3.1 R.YRPGTVALR.E K.STELLIR.K R.LVREIAQDFKTDLR.F R.EIAQDFKTDLR.F	4	3.030	2.575	2.218	2.608
IPI00217471.2	Hemoglobin epsilon subunit K.AAVTSLWSK.M K.AAVTSLWSKMNVEEAG GEALGR.L	5	2.174	3.613	2.447	2.745

	K.MNVEEAGGEALGR.L R.LLVVYPWTQRFFDSFGN LSSPSAILGNPK.V R.FFDSFGNLSSPSAILGNP K.V					
IPI00026272.1	Histone H2A R.VTIAQGGVLPNIQAVLL PK.K R.HLQLAIRNDEELNK.L R.NDEELNKLLGKVTIAQG GVLPNIQAVLLPK.K R.AGLQFPVGR.V	4	3.333	1.975	2.938	2.749
IPI00554676.1	Hemoglobin gamma-2 subunit K.MVTGVASALSSR.Y	1	2.778	3.448	2.286	2.837
IPI00217473.4	Hemoglobin zeta subunit R.TIIVSMWAK.I K.ISTQADTIGTETLER.L K.ISTQADTIGTETLERLFL SHPQTK.T R.LFLSHPQTK.T K.TYFPHFDLHPGSAQLR.A R.AHGSKVVAAVGDAVK. S K.VVAAVGDAVK.S K.SIDDIGGALSK.L K.SIDDIGGALSKLSELHAY ILR.V K.SIDDIGGALSKLSELHAY ILRVDPVNFK.L K.LSELHAYILR.V R.VDPVNFK.L	12	2.778	4.311	1.987	3.025
IPI00453473.5	Histone H4 R.DNIQGITKPAIR.R R.DNIQGITKPAIRR.L R.GVLKVFLENVIRDAVTY TEHAK.R R.ISGLIYEETR.G	8	4.000	2.690	2.496	3.062

R.KTVTAMDVVYALKR.Q
K.VFLENVIR.D
K.VFLENVIRDAVTYTEHA
K.R
K.TVTAMDVVYALKR.Q

CHAPTER 4

Mapping of Lysine Methylation and Acetylation in Core Histones of *Neurospora crassa*

Introduction

The eukaryotic nucleosome, the fundamental unit of chromatin, plays an important role in packaging and organizing the genetic material [1]. Each nucleosome consists of 146 bp of DNA wrapped around an octameric core histone complex comprising of two H2A-H2B dimers flanking a $(H3\text{-}H4)_2$ tetramer [1-2]. All core histones have a basic N-terminal domain, a globular domain organized by the histone fold and a C-terminal tail [1-2]. The histone N-terminal tails, which extend out from the core particle, are involved in the establishment of chromatin structural states, whereas their histone fold domains mediate histone-histone and histone-DNA interactions [2]. Core histones are susceptible to a variety of post-translational modifications (PTMs), which include methylation, acetylation, phosphorylation, ubiquitination, SUMOylation and ADP-ribosylation. These modifications, which occur mainly on the N-terminal tails [3], can affect the interactions of nucleosomes with transacting factors, and are thought to play a role in the assembly and disassembly of chromatin states, ultimately controlling the accessibility of DNA for important cellular processes including transcription, replication, gene silencing and DNA repair [4]. Distinct modifications of the histone tails can recruit specific chromatin-binding proteins, and

modifications on the same or different histone tails may be interdependent and generate various combinations on any individual nucleosome [5].

Mass spectrometry has been widely used for assessing histone PTMs. It provides direct information about the sites and types of modifications, differentiates isobaric modifications (e.g., acetylation vs. tri-methylation) [6], and allows for quantitative analysis [7]. PTM information obtained by mass spectrometric analysis facilitates genome-wide functional studies, which typically involve chromatin immunoprecipitation (ChIP) using antibodies recognizing specifically modified histones [8].

Neurospora crassa is a convenient model eukaryote, showing genomic features lacking in similar eukaryotes (e.g. yeast) including DNA methylation and certain histone modifications (e.g. H3K27 methylation). Studies using this organism have contributed significantly to the fundamental understanding of circadian rhythms, DNA methylation, genome defense systems, mitochondrial protein import, post-transcriptional gene silencing, DNA repair and other processes [9]. Being a multicellular filamentous fungus, *Neurospora* has also provided a system to study cellular differentiation and development [10]. However, no comprehensive investigation of the PTMs of *Neurospora* core histones has been reported.

In the present study, we extracted core histones from *Neurospora crassa* and achieved a systematic mapping of histone methylation and acetylation, with the combination of digestion with various proteases and analyses with multiple types of

mass spectrometers. Our results allowed for the identification of acetylation and methylation sites of lysine residues that were found previously in *Arabidopsis thaliana*, *Saccharomyces cerevisiae* and humans. More importantly, we identified several unique acetylation and methylation sites in core histones of *Neurospora*. Our analysis on core histone PTMs provides a foundation for examining the regulation of histone modifications and for genome-wide functional studies in this model organism.

Experimental

Extraction of core histones from Neurospora crassa

Neurospora crassa was cultured as described previously [11] and stored at -80 ºC. The core histones were obtained by using procedures reported for nuclei isolation by Emmett et al. [12] and histone extraction by Goff [13] with some modifications. Briefly, frozen *Neurospora* tissue (5 g) was ground to fine powder with a pestle in a cold mortar under liquid nitrogen. To the powder was subsequently added 20 mL nuclei isolation buffer containing 0.3 M sucrose, 40 mM $HaHSO_3$, 25 mM Tris-HCl (pH 7.4), 10 mM $MgSO_4$, 0.5 mM EDTA, 0.5% NP40, and the suspension was stirred vigorously. Cells were disrupted by intermittent exposure of the homogenate to sonication. The resulting mixture was filtered through two layers of silk in a Buchner funnel to remove whole cells and cell debris.

The above filtrate was centrifuged at 7,500g for 10 min and the supernatant was removed. To the precipitate was subsequently added 50 mL nuclei isolation

buffer, and the resulting mixture was gently shaked for 30 min and centrifuged again. The precipitate, which contained the nuclei, was resuspended in 1-mL ice-cold $NaHSO_3$ solution (25 mM, pH 7.2), followed by brief sonication (5 s). Hydrochloric acid was added immediately into the suspension to a final concentration of 0.3 M, and the resulting mixture was incubated at 4°C for 1 hr with continuous vortexing on an automatic vortexing machine. The histones in the supernatant were precipitated with cold acetone, centrifuged, dried and redissolved in water.

HPLC separation and protease digestion

Core histones were isolated by HPLC on an Agilent 1100 system (Agilent Technologies, Santa Clara, CA) as described previously [14]. A 4.6×250 mm C4 column (Grace Vydac, Hesperia, CA) was used. The wavelength for the UV detector was 220 nm. The flow rate was 0.8 mL/min, and a 60-min linear gradient of 30-60% acetonitrile in 0.1% trifluoroacetic acid (TFA) was employed.

In order to obtain high sequence coverage of the proteins, purified histones were digested separately with several proteases, including trypsin, Arg-C, Glu-C, Asp-N, and chymotrypsin. A protein/enzyme ratio of 50:1 (w/w) was employed for trypsin and 20:1 for other proteases. The different buffers used for the digestions were 100 mM NH_4HCO_3 (pH 8.0) for trypsin, Arg-C or Glu-C; 50 mM sodium phosphate (pH 8.0) for Asp-N; and 100 mM Tris-HCl (pH 7.8) along with 10 mM $CaCl_2$ for chymotrypsin. The digestion was carried out overnight at room temperature for

chymotrypsin and at 37ºC for other proteases. The peptide mixtures were subjected directly to LC-MS/MS analysis, or to a further peptide fractionation by HPLC and then analyzed by MALDI-MS/MS on a Q-STAR instrument or ESI-MS/MS on an LTQ-Orbitrap (See below).

Peptide fractionation was performed on the same HPLC system with a Zorbax SB-C18 capillary column (0.5×150 mm, 5 μm in particle size, Agilent Technologies), and a 60-min gradient of 2-60% acetonitrile in 0.6% acetic acid was used. The flow rate was 10 μL/min.

Mass spectrometry

MALDI-MS/MS measurements were performed on a QSTAR XL quadrupole/time-of-flight instrument equipped with an o-MALDI ion source (Applied Biosystems, Foster City, CA). The laboratory collision energy applied for MS/MS varied from 50 to 75 eV depending on peptide sequences and modification levels. The collision gas was nitrogen.

LC-MS/MS experiments were performed on three different instruments, including a 6510 QTOF LC/MS system with HPLC-Chip Cube MS interface (Agilent Technologies), an LTQ linear ion trap mass spectrometer and an LTQ-Orbitrap XL mass spectrometer with electron transfer dissociation (ETD) capability (Thermo Electron Co., San Jose, CA). The same 60-min linear gradient of 2-60% acetonitrile in 0.1% formic acid was applied for peptide separation.

In the 6510 QTOF LC/MS system, the sample enrichment, desalting, and HPLC separation were carried out automatically on the Agilent HPLC-Chip with an integrated trapping column (40 nL) and a separation column (Zorbax 300SB-C18, 75 μm × 150 mm, 5 μm in particle size). The Chip spray voltage (VCap) was set at 1950 V and varied depending on chip conditions. MS/MS experiments were carried out in either the data-dependent scan mode or the pre-selected ion mode. The width for precursor ion selection was 4 *m/z* units. The temperature and flow rate for the drying gas were 325ºC and 4 L/min, respectively. Nitrogen was used as collision gas, and collision energy followed a linear equation with a slope of 3 V per 100 m/z units and an offset of 2.5 V. The raw data obtained in the data-dependent scan mode were converted to Mascot generic format files, and submitted to the Mascot Database search engine (Matrix Science, Boston, MA) for protein and PTM identification. For LC-MS/MS analysis on the LTQ, peptides were separated with a Zorbax SB-C18 capillary column (0.5×150 mm, 5 μm in particle size, Agilent Technologies), and the mobile phases were delivered by the Agilent 1100 capillary HPLC pump at a flow rate of 6 μL/min. Helium was employed as the collision gas, and the normalized collision energy was 30%. The spray voltage was 4.5 kV, and the temperature for the ion transport tube was 275ºC.

ETD spectra were acquired on an LTQ-Orbitrap (Thermo Electron Co., San Jose, CA). Both online LC-MS/MS and offline direct infusion analyses were performed. In brief, during the online LC-MS/MS analysis, samples were redissolved

in 15 μL of 0.1% formic acid and loaded to a Biobasic C18 Picofrit capillary column (75 μm ×100 mm, 15 μm in particle size, New Objective, Woburn, MA) at a flow rate 0.3 μL/min. The ETD reaction time was 100 ms. In the offline analysis, the pre-fractionated H2B and H4 N-terminal peptides were redissolved in 20 μL 50:50 acetonitrile:H_2O with 0.1% formic acid. Samples were directly infused. Different ETD reaction time was used to optimize peptide fragmentation.

Results

Identification of PTMs in histones H2B and H2A

Neurospora crassa is a convenient model for multicellular eukaryotes, and it has been frequently used for studying the regulation of various cellular processes [10]. To improve the foundation of information on chromatin structure and function in *Neurospora*, we initiated a systematic investigation of the PTMs of core histones in this organism. We first extracted core histones from *Neurospora* tissues and fractionated individual core histones by using reverse-phase HPLC. The core histones were eluted in the order of H2B, H4, H2A, and H3 (Figure S4.1). We then digested the core histones individually with different proteases and analyzed the peptide mixtures with LC-MS/MS on various instrument platforms to obtain high sequence coverage and achieve unambiguous PTM assignment. The detected PTM sites are summarized in Figure 4.1 and the sequence coverage is shown in Figure S4.2.

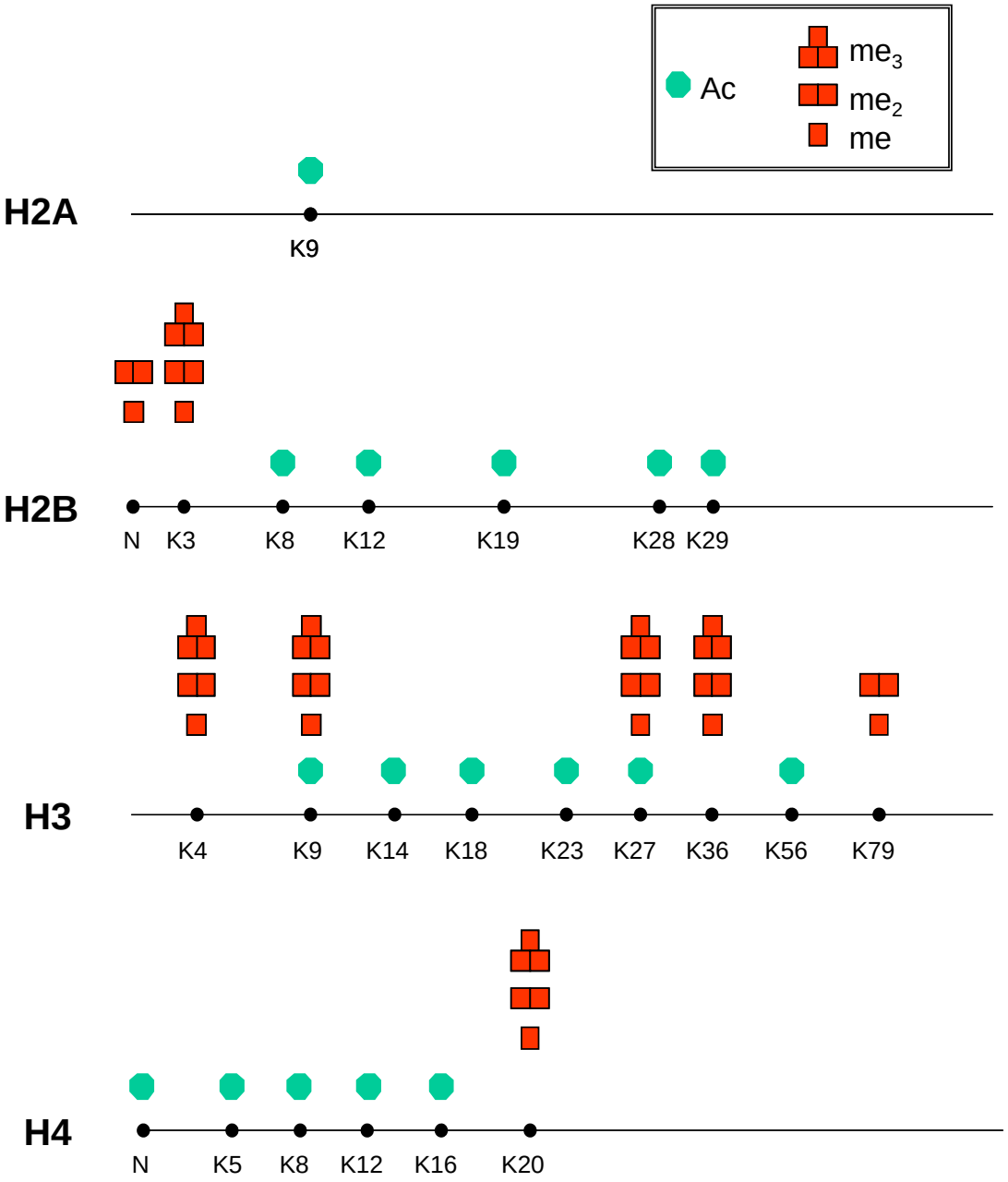


Figure 4.1 Summaries of the detected PTMs of *Neurospora* core histones. The modified residues are labeled, and "N" represents N terminus. Acetylation is designated with solid octagon, and mono-, di-, and tri-methylation are represented by one-, two, and three square boxes, respectively.

The isolated histone H2B was digested with trypsin, Asp-N and Glu-C separately to obtain peptides in appropriate lengths and good sequence coverage. The digestion mixtures are subjected subsequently to LC-MS/MS analysis, and the acquired mass spectra were searched with Mascot search engine and the results manually verified. A sequence coverage of 100% was reached and multiple modification sites were identified.

Acetylation of histone H2B was reported for several organisms, including human, budding yeast, and *Arabidopsis* [15]. Here we identified many acetylation sites, including K7 and K12, which appear acetylated among different organisms, and K19, K28 and K29, which have not been reported.

Aside from acetylation, we also observed the methylation of K3 and the N-terminus of *Neurospora* H2B. Example mass spectra for the H2B N-terminal peptide $_{1}$PPKPADKKPASK$_{12}$ are depicted in Figure 4.2 and Figure S4.3. In this regard, the positive-ion ESI-MS (Figure 4.2A) reveals the presence of one acetyl group and 2, 3, 4, or 5 methyl groups in this peptide segment. The MS/MS of this group of peptides were all obtained with the selected-ion monitoring (SIM) mode of analysis. The modification sites could be easily determined from fragment ions, with the consideration of mass shifts introduced by PTMs, e.g. 14.0157 Da for monomethylation and 42.0106 Da for acetylation. MS/MS results revealed that the N-terminus was mono- or di-methylated, K3 was mono-, di- or tri-methylated, and K7 was completely acetylated. In the MS/MS of the di-methylated and mono-acetylated

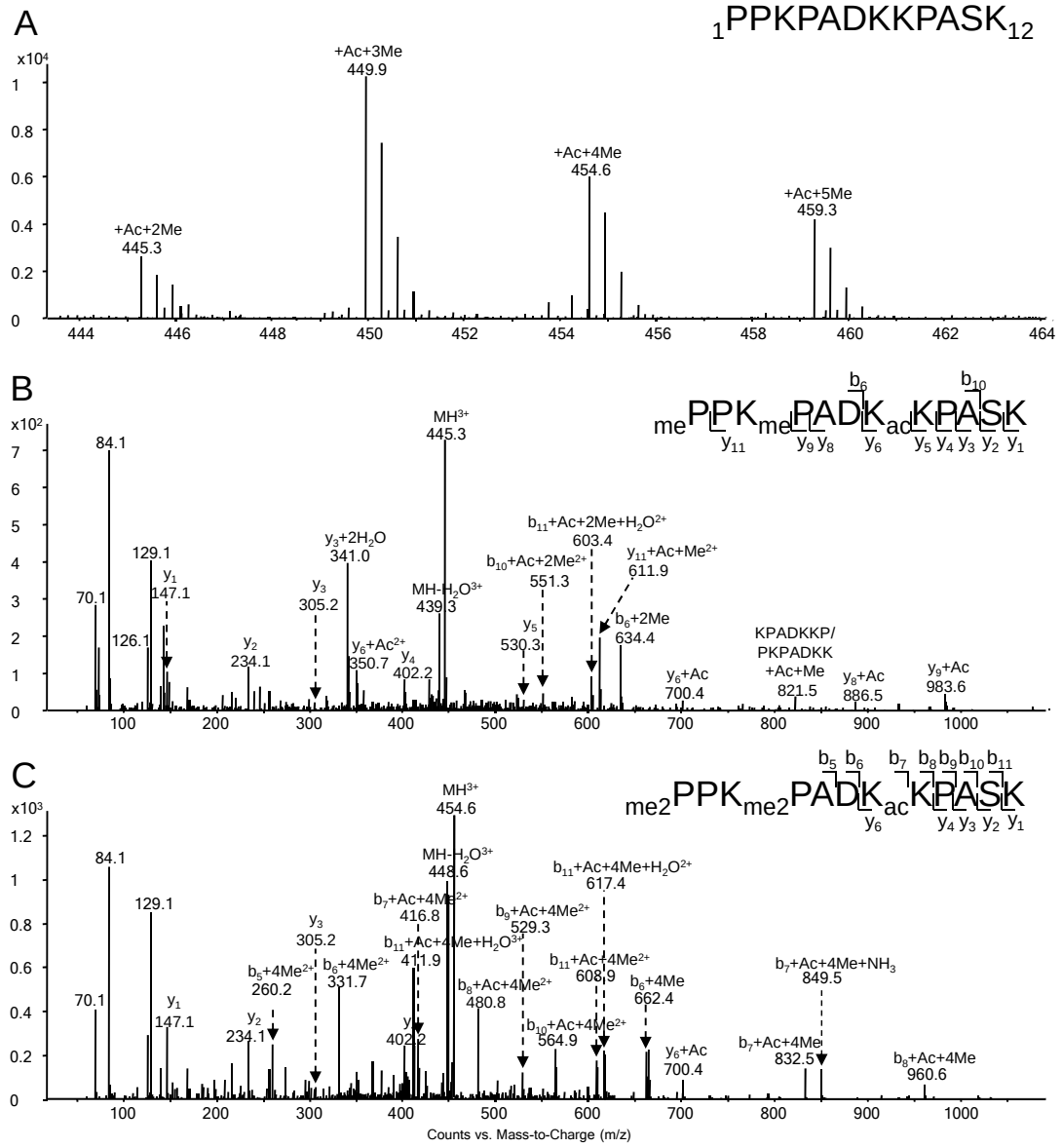


Figure 4.2 (A) ESI-MS of N-terminal tryptic peptide $_{1}$PPKPADKKPASK$_{12}$ of histone H2B extracted from *Neurospora*. (B-C) The MS/MS of the di-methylated, mono-acetylated (B) and tetra-methylated, mono-acetylated (C) H2B peptide with residues 1-12 obtained by Q-TOF analysis.

peptide (Figure 4.2B), the existence of the b_6+2Me, b_{10}+Ac+2Me and y_{11}+Ac+Me ions supports the mono-methylation on the N-terminus and K3. We also observed y_4, y_5, y_6+Ac ions, providing evidence for the K7 acetylation. Similar methylation of N-terminal proline has been observed in H2B from *Drosophila melanogaster* [16] and gonads of the starfish *Asterias rubens* [17]. Along this line, N-terminal alanine methylation was found for *Tetrahymena* histone H2B [18] and *Arabidopsis* histone H2B variants HTB-9 and HTB-11 [19]. Moreover, N-terminal α-methylation of RCC1 protein is believed to promote stable association of RCC1 with chromatin through DNA binding in an α-methylation-dependent manner [20]. This is required for correct spindle assembly and chromosome segregation during mitosis. Presence of methylation at the N-terminus of H2B suggests that this modification may have an unknown significant function.

Differentiation of tri-methylation from acetylation is essential in PTM studies of histones [6]. A typical method to distinguish these two modifications is based on the immonium ion with *m/z* 126.1 from acetylated lysine and the neutral loss of a trimethylamine [$N(CH_3)_3$] from tri-methyl lysine-containing precursor and fragment ions [6]. However, the method does not work effectively when dealing with peptides containing more than one acetylation or tri-methylation sites, such as the H2B N-terminal peptide with multiple lysines being modified. Here we differentiate these two modifications based on subtle difference in mass increase of the lysine residue induced by the acetylation and tri-methylation. For instance, the measured mass

difference between the b_6 and b_7 ions observed in Figure 4.2C for the peptide segment housing residues 1-12 was 832.5055-662.3982 = 170.1073 Da. This mass difference matches much more closely with the calculated mass difference with the consideration of K7 acetylation (170.1056 Da, with a relative deviation of 9.9 ppm) than that with the consideration of K7 tri-methylation (170.1420 Da, with a relative deviation of 204 ppm). Thus, K7 is acetylated. All the acetylation and tri-methylation sites of *Neurospora* core histones were unambiguously established in this way, and more sample results are displayed in Table S4.1.

The presence of an acetylated lysine can block the trypsin cleavage of its C-terminal side amide bond. With many lysine residues being acetylated, tryptic cleavage of H2B gives rise to a very long N-terminal peptide. Traditional collisionally induced dissociation (CID) cannot provide a complete series of fragment ions due to its poor ability to cleave the backbone of large peptides, rendering it difficult to identify modification sites. For instance, the exact numbers of methyl groups on the N-terminus and K3 cannot be delineated unambiguously based on the CID-produced MS/MS on the tetra-methylated peptide with residues 1-12 (Figure 4.2C). To overcome this limitation, we applied ETD, which can afford efficient cleavage along the backbone of long peptides or even intact proteins [21], to analyze those large N-terminal peptides. Figure 4.3 shows the ETD MS/MS of the $[M+5H]^{5+}$ ion of $_{1}$PPKPADKKPASKAPATASKAPEKK$_{24}$ with the N-terminus and K3 being di-methylated, as well as with K7, K12 and K19 being acetylated. Nearly complete

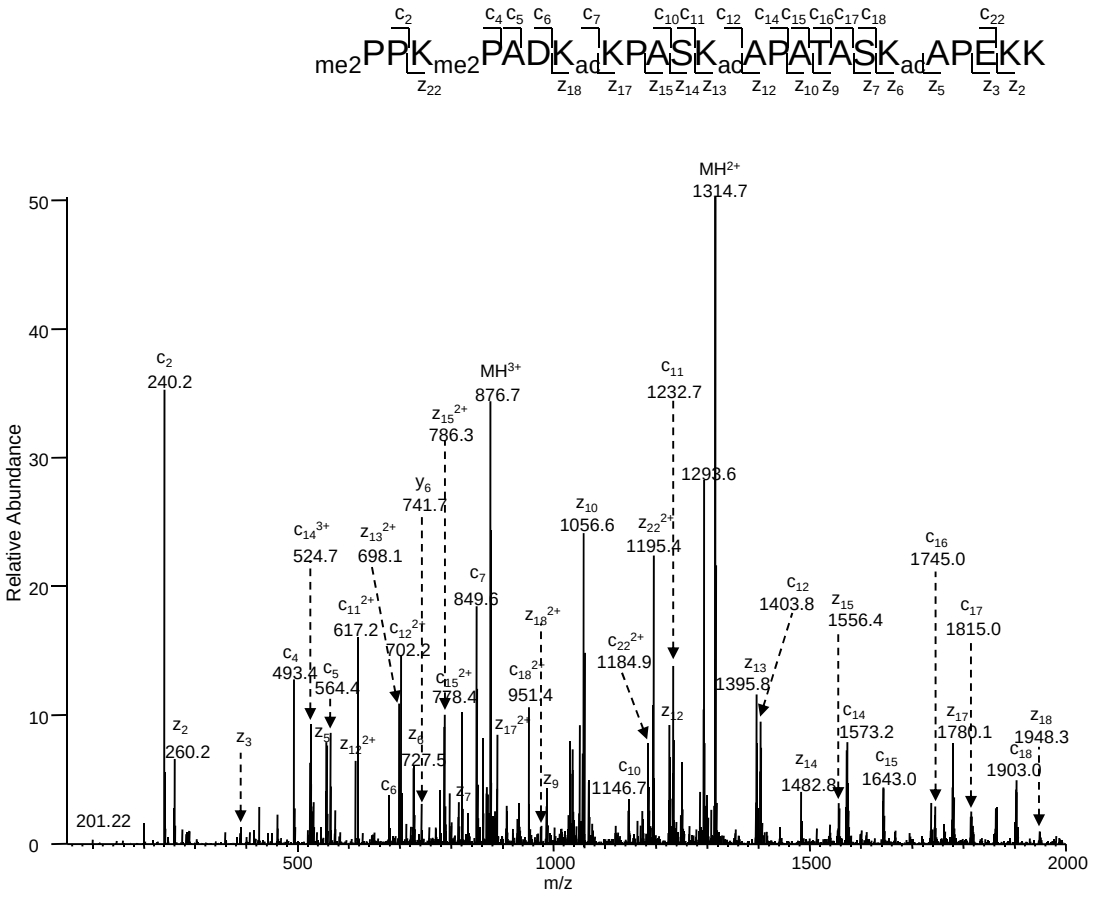


Figure 4.3 The ETD MS/MS of the *Neurospora* H2B N-terminal peptide with residues 1-24 with the N-terminus and K3 being dimethylated, and with K7, K12, K19 being acetylated.

series of c and z ions were formed from ETD. The observation of c_2+2Me and c_4+4Me ions reveals the dimethylation of the N terminus and K3. The presence of the c_6+4Me, c_7+4Me+Ac, z_{17}+2Ac, and z_{18}+3Ac ions supported the K7 acetylation, while the existence of the c_{11}+4Me+Ac, c_{12}+4Me+2Ac, z_{12}+Ac, z_{13}+2Ac ions demonstrated the K12 acetylation. Along this line, the acetylation of K19 was identified based on the observation of the z_5 and z_6+Ac ions. K28 and K29 were also found to be acetylated heterogeneously in *Neurospora* H2B, as supported by the coexistence of y_5, y_5+Ac and b_4, b_4+Ac ions in the MS/MS of the Asp-N-produced peptide $_{25}$DAGKKTAASG$_{34}$ (Figure S4.4A).

Purified H2A was digested individually with trypsin, Arg-C or chymotrypsin under optimized conditions, and analyzed by LC-MS/MS. We found that only K9 was acetylated (MS/MS for the $[M+2H]^{2+}$ ion of $_6$SGGKASGSKNAQSR$_{19}$ is shown in Figure S4.4B, which reveals the formation of the b_4+Ac, y_{10} and y_{11}+Ac ions).

Identification of PTMs in histone H3

The HPLC-purified histone H3 was digested with Arg-C and Glu-C, and the digestion mixture was analyzed by LC-MS/MS directly, or further fractionated by HPLC and analyzed by MALDI-MS/MS. All the conserved methylation and acetylation sites reported in other organisms were identified in *Neurospora*, including methylation at K4, K9, K27, K36, K79 and acetylation at K9, K14, K18, K23, K27 and K56. K4 was found to be mono-, di- and tri-methylated, whereas K79 was found

to be mono- and di-methylated in *Neurospora* (Figure S4.5A&B). In the MS/MS of the $[M + 2H]^{2+}$ ion (*m/z* 366.7) of the tri-methylated peptide $_{3}$TKQTAR$_{8}$ (Figure S4.5C), we observed the b_3+3Me, b_4+3Me, and y_4 ions, supporting the observed tri-methylation of K4. MS/MS of the dimethylated peptide segment $_{73}$EIAQDFKSDLR$_{83}$ displayed the presence of a complete series of y ions, revealing di-methylation of K79 (Figure S4.5D). Figure S4.5E illustrates the fragmentation of the precursor ion at *m/z* 646.9, which is consistent with the acetylation of K56 in peptide $_{54}$YQKSTELLIR$_{63}$. In the MS/MS of the triply charged peptide $_{9}$KSTGGKAPRKQLASKAAR$_{26}$, K9, K14, K18 and K23 were determined to be acetylated (Figure S4.5F), with the observation of b_2+Ac ion for supporting K9 acetylation, the b_5+Ac and b_7+2Ac ions for K14 acetylation, as well as y_1, y_5+Ac, y_{11}+2Ac ions for K18 and K23 acetylation.

It is worth noting that K9 and K27 were found either methylated or acetylated. The methylated and acetylated peptides exhibited different retention time on the reverse-phase column used for LC-MS/MS analysis; the retention times for the methylated and acetylated peptides were approximately 4 and 8 min, respectively (Figure S4.6). Thus, methylated and acetylated peptides could be well isolated, and unambiguous MS/MS spectra could be obtained. Figure 4.4A shows the ESI-MS of the triply charged ions of the unmodified, mono-, di-, and tri-methylated peptides with residues $_{9}$KSTGGKAPR$_{17}$, while Figure 4.4B shows the MS for the acetylated peptide with same sequence. In the MS/MS of the $[M+3H]^{3+}$ ions of the tri-methylated and

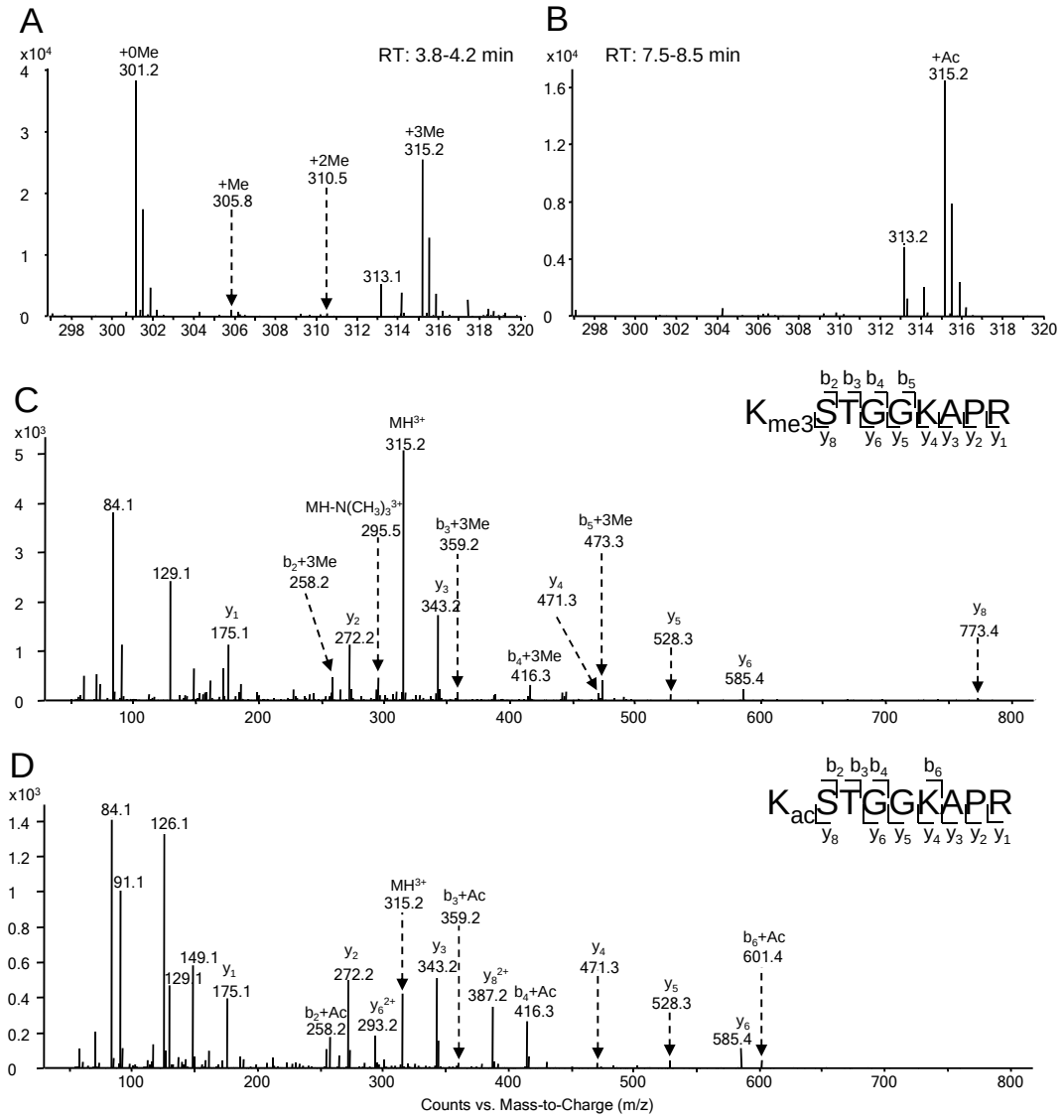


Figure 4.4 Positive-ion ESI-MS of the Arg-C-produced *Neurospora* H3 peptide $_{9}$KSTGGKAPR$_{17}$ with K9 being methylated (A) or acetylated (B). Shown in (C) and (D) are the MS/MS, obtained on the Q-TOF mass spectrometer, of the tri-methylated and acetylated peptides with residues 9-17.

acetylated peptides, a series of tri-methylated and acetylated b ions and unmodified y ions were observed, supporting the conclusion that K9 was tri-methylated and acetylated, respectively (Figure 4.4C&D). The presence of the immonium ion at *m/z* 126, the neutral loss of $N(CH_3)_3$ (59 Da) from the precursor ion, as well as the mass difference between tri-methylation and acetylation helped differentiate these two types of isobaric modifications.

K27 was also found to be mono-, di- or tri-methylated or acetylated in *Neurospora* (Figure S4.7A, B). The MS/MS of the peptides $_{27}$KSAPSTGGVKKPHR$_{42}$ with different modification levels were obtained by the selected-ion monitoring. The MS/MS of the $[M+4H]^{4+}$ ion at *m/z* 384.2 eluted at 16 and 18 min revealed the tri-methylation on K36 as well as the tri-methylation or acetylation on K27. In this context, the presence of the y_7 + 3me, y_8+ 3me, y_9 + 3me, y_{10}+ 3me and y_{11} + 3me, along with the observation of the y_1, y_2, and y_4 ions, demonstrates the tri-methylation on K36 (Figure S4.7C, D). On the other hand, the acetylation and tri-methylation on K27 are manifested by the presence of b_2 and b_3 ions bearing an acetyl group and three methyl groups, respectively (Figure S4.7C, D).

We also attempted to examine the phosphorylation of core histones by enriching phosphorylated peptides using TiO_2-coated magnetic beads [Phosphopeptide enrichment kit (PerkinElmer, Waltham, MA)] [22]. We were able to detect very low level of phosphorylation of H3 S10 in a phosphatase-deficient

(PP1-deficient) *Neurospora* strain (Figure 4.S8) [11a]; H3 S10 phosphorylation in wild-type *Neurospora* was, however, below the detection limit of the instruments.

Identification of PTMs in histone H4

Purified H4 was digested by trypsin and Asp-N separately and subjected to LC- or MALDI-MS/MS analysis. All the modifications were located on the N-terminal segment, similar to the situation observed for other organisms. Asp-N digestion produced a long N-terminal peptide with residues $_{1}$TGRGKGGKGLGKGGAKRHRKILR$_{23}$ containing all the modification sites in H4, which include acetylation at the N-terminus and at K5, K8, K12 and K16, along with methylation at K20 (Figure S4.9A). A LTQ-Orbitrap with ETD provided high-quality tandem mass spectra of the $[M+6H]^{6+}$ ions of the peptide with a nearly complete series of c and z ions. Figure S4.9B shows an ETD MS/MS of the di-acetylated and tri-methylated peptide with residues 1-23. The formation of the mono-acetylated small c ions and z_{20}+3Me+Ac ions suggests N-terminal acetylation. Additionally, the presence of c_{15}+Ac, c_{16}+2Ac, z_7+3Me, z_8+3Me+Ac ions, and c_{19}+2Ac, c_{20}+2Ac+3Me, z_3, z_4+3Me ions supports the conclusions of K16 acetylation and K20 tri-methylation, respectively.

LC-MS/MS with the QTOF mass spectrometer of the tryptic digestion mixture led to the identification of relatively low levels of acetylation on K5, K8 and K12, which cannot be identified from corresponding analyses of the Asp-N digestion

mixture of histone H4. In the MS/MS of the doubly charged peptide $_4$GKGGKGLGK$_{12}$, K5 and K8 were determined to be acetylated, as exemplified by the observations of the b_2+Ac, b_5+2Ac, y_4, y_5+Ac, and y_7+Ac ions (Figure S4.10A). MS/MS of the doubly charged ion of the di-acetylated peptide $_9$GLGKGGAKR$_{17}$ revealed the formation of b_3, b_4+Ac, y_5+Ac and y_6+2Ac ions supporting the K12 acetylation, and the presence of the y_1 and y_2+Ac ions suggesting the acetylation of K16 (Figure S4.10B).

Discussion and Conclusions

We used a combination of various mass spectrometric methods including MALDI-TOF, LC-MS/MS with CID and ETD, coupled with different protease digestions and HPLC purification to identify histone methylation and acetylation sites in *Neurospora crassa*. We provide a thorough mapping of lysine methylation and acetylation for core histones in this organism. Our results show that the core histones from *Neurospora* are extensively acetylated and/or methylated on lysine residues that are also acetylated and/or methylated in mammals, budding yeast and plants. Moreover, some novel modifications were detected for the first time. The methylation and acetylation sites found in H2B *Neurospora* include N-terminal and K3 methylation and K7, K12, K19, K28, K29 acetylation. H2A was shown to have acetylation on K9. H3 was found to bear methylation on K4, K9, K27, K36 and K79, together with acetylation on K9, K14, K18, K23, K27 and K56. Finally, H4 was

found to sport methylation of K20 and acetylation of N-terminus, K5, K8, K12, and K16 (summarized in Figure 4.1). However, it is necessary to mention that some low levels of modification in *Neurospora* might not be detected even with the combination of multiple protease digestions and various mass spectrometric measurements, and so is true with the histone modifications reported in other organisms. A detailed comparison of PTMs of core histones in different organisms is summarized in Table 4.1 [7, 15a, 23].

As depicted in Table 4.1, we found conserved acetylation of H2B at K7 and K12 (the corresponding residues in other organisms are K6 and K11, respectively. See Table 4.1), and unique acetylation of *Neurospora* K19, K28, K29. It has been reported that acetylated lysine residues in yeast H2B activate the transcription of genes involved in NAD biosynthesis and vitamin metabolism [24]. Thus, it will be interesting to assess the role of *Neurospora* H2B acetylation in transcription activation. The K3 methylation that we observed in H2B appears to be novel. H2A was found to be only acetylated on K9 in *Neurospora*, as in humans.

In *Neurospora* H3, we found all the commonly conserved N-terminal modifications including methylation of K4, K9, K27 and K36, as well as acetylation of K9, K14, K18, K23 and K27. K4 was mostly unmodified, but approximately 10% of the K4 peptide was mono-, di- or tri-methylated. K9 was predominantly tri-methylated. K27 and K36 were mono-, di- and tri-methylated. The methylated and acetylated peptides housing the same residues had different retention times during

Table 4.1 Comparison of core histone methylation and acetylation among different organisms including *Neurospora crassa, Saccharomyces cerevisiae*, *Arabidopsis thaliana* and *Homo sapiens* [7, 15a, 23].

Histones	Modifications	Modification sites			
		N. crassa	*S. cerevisiae*	*A. thaliana*	*H. sapiens*
H2B	Methylation	N terminus, K3			
	Acetylation	K7, K12, K19, K28, K29	K3, K6, K11, K16, K21, K22	K6, K11, K27, K32	K5, K11, K12, K15, K16, K23
H2A	Acetylation	K9	K4, K7	K5, K144	K5, K9
H3	Methylation	K4, K9, K27, K36, K79	K4, K36, K79	K4, K9, K27, K36	K4, K9, K27, K36, K79
	Acetylation	K9, K14, K18, K23, K27, K56	K9, K14, K18, K23, K27, K56	K9, K14, K18, K23	K9, K14, K18, K23, K27
H4	Methylation	K20	K20		K20
	Acetylation	N terminus, K5, K8, K12, K16	N terminus, K5, K8, K12, K16	K5, K8, K12, K16	N terminus, K5, K8, K12, K16

HPLC separation, allowing the modification types to be easily determined. Moreover, we observed the acetylation on K56.

H3K79 is mono-, di- or tri-methylated in many mammalian and non-mammalian cell lines and in budding yeast. However, K79 was found only mono- and di-methylated in *Neurospora*. It would be interesting to explore the functional implications of the lack of K79 tri-methylation, since the H3 K79 methyltransferase, DOT1, is involved in telomeric silencing [25]. In addition, it was worth noting that K64 of H3 of mammalian cells was recently found to be tri-methylated; this methylation is associated with heterochromatin, and it is lost during developmental reprogramming [26]. We monitored specifically the fragmentation of the peptide containing this putative modification by MS/MS but did not find this modification in *Neurospora*. It is possible that the level of this modification is below the detection limits of the method; conceivably, it could also have been lost during extraction process.

Histone H4 acetylation sites, including the N-terminus, and residues K5, K8, K12 and K16, are conserved in almost all organisms, including *Neurospora*. They play important roles in many processes including transcriptional activation, DNA double strand break repair and cellular lifespan regulation [27]. H4 K20 methylation, conserved in almost all multi-cellular organisms, was also found in *Neurospora*, in the form of mono-, di- and mainly tri-methylation. While its role in heterochromatin silencing and DNA damage response has been extensively studied in humans,

Drosophila melanogaster and *Schizosaccharomyces pombe* [28], it is also important to study its function in *Neurospora*.

In summary, a systematic mapping of histone methylation and acetylation in *Neurospora crassa* was obtained by mass spectrometric analyses. The rigorous identification of modification sites provides a foundation for further studies on the regulation and functions of histone modifications in this model organism.

References:

1. McGhee, J. D.; Felsenfeld, G., Nucleosome structure. *Annu. Rev. Biochem.* 1980, *49*, 1115-56.

2. Luger, K.; Mader, A. W.; Richmond, R. K.; Sargent, D. F.; Richmond, T. J., Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* 1997, *389* (6648), 251-60.

3. (a) Strahl, B. D.; Allis, C. D., The language of covalent histone modifications. *Nature* 2000, *403* (6765), 41-5; (b) Berger, S. L., Histone modifications in transcriptional regulation. *Curr. Opin. Genet. Dev.* 2002, *12* (2), 142-148.

4. (a) Geiman, T. M.; Robertson, K. D., Chromatin remodeling, histone modifications, and DNA methylation-how does it all fit together? *J. Cell. Biochem.* 2002, *87* (2), 117-25; (b) Nacheva, G. A.; Guschin, D. Y.; Preobrazhenskaya, O. V.; Karpov, V. L.; Ebralidse, K. K.; Mirzabekov, A. D., Change in the pattern of histone binding to DNA upon transcriptional activation. *Cell* 1989, *58* (1), 27-36; (c) Irvine, R. A.; Lin, I. G.; Hsieh, C. L., DNA methylation has a local effect on transcription and histone acetylation. *Mol. Cell. Biol.* 2002, *22* (19), 6689-96; (d) Peterson, C. L.; Laniel, M. A., Histones and histone modifications. *Curr. Biol.* 2004, *14* (14), R546-R551; (e) Cheung, P.; Allis, C. D.; Sassone-Corsi, P., Signaling to chromatin through histone modifications. *Cell* 2000, *103* (2), 263-271.

5. (a) Jenuwein, T.; Allis, C. D., Translating the histone code. *Science* 2001, *293* (5532), 1074-80; (b) Klose, R. J.; Zhang, Y., Regulation of histone methylation by

demethylimination and demethylation. *Nat. Rev. Mol. Cell. Biol.* 2007, *8* (4), 307-18; (c) Fillingham, J.; Greenblatt, J. F., A histone code for chromatin assembly. *Cell* 2008, *134* (2), 206-208.

6. Zhang, K.; Yau, P. M.; Chandrasekhar, B.; New, R.; Kondrat, R.; Imai, B. S.; Bradbury, M. E., Differentiation between peptides containing acetylated or tri-methylated lysines by mass spectrometry: an application for determining lysine 9 acetylation and methylation of histone H3. *Proteomics* 2004, *4* (1), 1-10.

7. Beck, H. C.; Nielsen, E. C.; Matthiesen, R.; Jensen, L. H.; Sehested, M.; Finn, P.; Grauslund, M.; Hansen, A. M.; Jensen, O. N., Quantitative proteomic analysis of post-translational modifications of human histones. *Mol. Cell. Proteomics* 2006, *5* (7), 1314-25.

8. (a) Bernstein, B. E.; Humphrey, E. L.; Erlich, R. L.; Schneider, R.; Bouman, P.; Liu, J. S.; Kouzarides, T.; Schreiber, S. L., Methylation of histone H3 Lys 4 in coding regions of active genes. *Proc. Natl. Acad. Sci. USA* 2002, *99* (13), 8695-700; (b) Kurdistani, S. K.; Tavazoie, S.; Grunstein, M., Mapping global histone acetylation patterns to gene expression. *Cell* 2004, *117* (6), 721-33; (c) Ng, H. H.; Robert, F.; Young, R. A.; Struhl, K., Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol. Cell* 2003, *11* (3), 709-19.

9. Galagan, J. E.; Calvo, S. E.; Borkovich, K. A.; Selker, E. U.; Read, N. D.; Jaffe, D.; FitzHugh, W.; Ma, L. J.; Smirnov, S.; Purcell, S.; Rehman, B.; Elkins, T.; Engels,

R.; Wang, S.; Nielsen, C. B.; Butler, J.; Endrizzi, M.; Qui, D.; Ianakiev, P.; Bell-Pedersen, D.; Nelson, M. A.; Werner-Washburne, M.; Selitrennikoff, C. P.; Kinsey, J. A.; Braun, E. L.; Zelter, A.; Schulte, U.; Kothe, G. O.; Jedd, G.; Mewes, W.; Staben, C.; Marcotte, E.; Greenberg, D.; Roy, A.; Foley, K.; Naylor, J.; Stange-Thomann, N.; Barrett, R.; Gnerre, S.; Kamal, M.; Kamvysselis, M.; Mauceli, E.; Bielke, C.; Rudd, S.; Frishman, D.; Krystofova, S.; Rasmussen, C.; Metzenberg, R. L.; Perkins, D. D.; Kroken, S.; Cogoni, C.; Macino, G.; Catcheside, D.; Li, W.; Pratt, R. J.; Osmani, S. A.; DeSouza, C. P.; Glass, L.; Orbach, M. J.; Berglund, J. A.; Voelker, R.; Yarden, O.; Plamann, M.; Seiler, S.; Dunlap, J.; Radford, A.; Aramayo, R.; Natvig, D. O.; Alex, L. A.; Mannhaupt, G.; Ebbole, D. J.; Freitag, M.; Paulsen, I.; Sachs, M. S.; Lander, E. S.; Nusbaum, C.; Birren, B., The genome sequence of the filamentous fungus Neurospora crassa. *Nature* 2003, *422* (6934), 859-68.

10. Davis, R. H.; Perkins, D. D., Timeline: Neurospora: a model of model microbes. *Nat. Rev. Genet.* 2002, *3* (5), 397-403.

11. (a) Adhvaryu, K. K.; Selker, E. U., Protein phosphatase PP1 is required for normal DNA methylation in Neurospora. *Genes Dev.* 2008, *22* (24), 3391-6; (b) Adhvaryu, K. K.; Morris, S. A.; Strahl, B. D.; Selker, E. U., Methylation of histone H3 lysine 36 is required for normal development in Neurospora crassa. *Eukaryot. Cell* 2005, *4* (8), 1455-64; (c) Selker, E. U.; Fritz, D. Y.; Singer, M. J., Dense nonsymmetrical DNA methylation resulting from repeat-induced point mutation in Neurospora. *Science* 1993, *262* (5140), 1724-8.

12. Emmett, N.; Williams, C. M.; Frederick, L.; Williams, L. S., Arginyl-transfer ribonucleic acid and synthetase of Neurospora crassa. *Mycologia* 1972, *64* (3), 499-509.

13. Goff, C. G., Histones of Neurospora crassa. *J. Biol. Chem.* 1976, *251* (13), 4131-8.

14. Xiong, L.; Ping, L.; Yuan, B.; Wang, Y., Methyl group migration during the fragmentation of singly charged ions of trimethyllysine-containing peptides: precaution of using MS/MS of singly charged ions for interrogating peptide methylation. *J. Am. Soc. Mass Spectrom.* 2009, *20* (6), 1172-81.

15. (a) Zhang, K.; Sridhar, V. V.; Zhu, J.; Kapoor, A.; Zhu, J. K., Distinctive core histone post-translational modification patterns in *Arabidopsis thaliana*. *PLoS One* 2007, *2* (11), e1210; (b) Zhang, K. L., Characterization of acetylation of Saccharomyces cerevisiae H2B by mass spectrometry. *Int. J. Mass Spectrom.* 2008, *278* (1), 89-94; (c) Clayton, A. L.; Hebbes, T. R.; Thorne, A. W.; Cranerobinson, C., Histone Acetylation and Gene Induction in Human-Cells. *FEBS Lett.* 1993, *336* (1), 23-26; (d) Grimes, S. R.; Henderson, N., Acetylation of rat testis histones H2B and Th2B. *Dev. Biol.* 1984, *101* (2), 516-521.

16. Desrosiers, R.; Tanguay, R. M., Methylation of Drosophila histones at proline, lysine, and arginine residues during heat shock. *J. Biol. Chem.* 1988, *263* (10), 4686-92.

17. Martinage, A.; Briand, G.; Van Dorsselaer, A.; Turner, C. H.; Sautiere, P., Primary structure of histone H2B from gonads of the starfish *Asterias rubens*. Identification of an N-dimethylproline residue at the amino-terminal. *Eur. J. Biochem.* 1985, *147* (2), 351-9.

18. Nomoto, M.; Kyogoku, Y.; Iwai, K., N-Trimethylalanine, a novel blocked N-terminal residue of *Tetrahymena* histone H2B. *J. Biochem.* 1982, *92* (5), 1675-8.

19. Bergmuller, E.; Gehrig, P. M.; Gruissem, W., Characterization of post-translational modifications of histone H2B-variants isolated from Arabidopsis thaliana. *J. Proteome Res.* 2007, *6* (9), 3655-68.

20. Chen, T.; Muratore, T. L.; Schaner-Tooley, C. E.; Shabanowitz, J.; Hunt, D. F.; Macara, I. G., N-terminal alpha-methylation of RCC1 is necessary for stable chromatin association and normal mitosis. *Nat. Cell Biol.* 2007, *9* (5), 596-603.

21. (a) Syka, J. E.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F., Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc. Natl. Acad. Sci. USA* 2004, *101* (26), 9528-33; (b) Mikesh, L. M.; Ueberheide, B.; Chi, A.; Coon, J. J.; Syka, J. E.; Shabanowitz, J.; Hunt, D. F., The utility of ETD mass spectrometry in proteomic analysis. *Biochim. Biophys. Acta* 2006, *1764* (12), 1811-22.

22. Larsen, M. R.; Thingholm, T. E.; Jensen, O. N.; Roepstorff, P.; Jorgensen, T. J., Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Mol. Cell. Proteomics* 2005, *4* (7), 873-86.

23. (a) Garcia, B. A.; Hake, S. B.; Diaz, R. L.; Kauer, M.; Morris, S. A.; Recht, J.; Shabanowitz, J.; Mishra, N.; Strahl, B. D.; Allis, C. D.; Hunt, D. F., Organismal differences in post-translational modifications in histones H3 and H4. *J Biol Chem* 2007, *282* (10), 7641-55; (b) Sinha, I.; Wiren, M.; Ekwall, K., Genome-wide patterns of histone modifications in fission yeast. *Chromosome Research* 2006, *14* (1), 95-105.

24. Parra, M. A.; Kerr, D.; Fahy, D.; Pouchnik, D. J.; Wyrick, J. J., Deciphering the roles of the histone H2B N-terminal domain in genome-wide transcription. *Mol. Cell. Biol.* 2006, *26* (10), 3842-52.

25. Ng, H. H.; Feng, Q.; Wang, H. B.; Erdjument-Bromage, H.; Tempst, P.; Zhang, Y.; Struhl, K., Lysine methylation within the globular domain of histone H3 by Dot1 is important for telomeric silencing and Sir protein association. *Genes Dev.* 2002, *16* (12), 1518-1527.

26. Daujat, S.; Weiss, T.; Mohn, F.; Lange, U. C.; Ziegler-Birling, C.; Zeissler, U.; Lappe, M.; Schubeler, D.; Torres-Padilla, M. E.; Schneider, R., H3K64 trimethylation marks heterochromatin and is dynamically remodeled during developmental reprogramming. *Nat. Struct. Mol. Biol.* 2009, *16* (7), 777-81.

27. (a) Dang, W.; Steffen, K. K.; Perry, R.; Dorsey, J. A.; Johnson, F. B.; Shilatifard, A.; Kaeberlein, M.; Kennedy, B. K.; Berger, S. L., Histone H4 lysine 16 acetylation regulates cellular lifespan. *Nature* 2009, *459* (7248), 802-7; (b) Bird, A. W.; Yu, D. Y.; Pray-Grant, M. G.; Qiu, Q.; Harmon, K. E.; Megee, P. C.; Grant, P. A.; Smith, M.

M.; Christman, M. F., Acetylation of histone H4 by Esa1 is required for DNA double-strand break repair. *Nature* 2002, *419* (6905), 411-5.

28. (a) Fang, J.; Feng, Q.; Ketel, C. S.; Wang, H. B.; Cao, R.; Xia, L.; Erdjument-Bromage, H.; Tempst, P.; Simon, J. A.; Zhang, Y., Purification and functional characterization of SET8, a nucleosomal histone H4-lysine 20-specific methyltransferase. *Curr. Biol.* 2002, *12* (13), 1086-1099; (b) Rice, J. C.; Nishioka, K.; Sarma, K.; Steward, R.; Reinberg, D.; Allis, C. D., Mitotic-specific methylation of histone H4 Lys 20 follows increased PR-Set7 expression and its localization to mitotic chromosomes. *Genes Dev.* 2002, *16* (17), 2225-2230; (c) Sanders, S. L.; Portoso, M.; Mata, J.; Bahler, J.; Allshire, R. C.; Kouzarides, T., Methylation of histone H4 lysine 20 controls recruitment of Crb2 to sites of DNA damage. *Cell* 2004, *119* (5), 603-614; (d) Schotta, G.; Lachner, M.; Sarma, K.; Ebert, A.; Sengupta, R.; Reuter, G.; Reinberg, D.; Jenuwein, T., A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev.* 2004, *18* (11), 1251-1262.

Supporting Information for Chapter 4

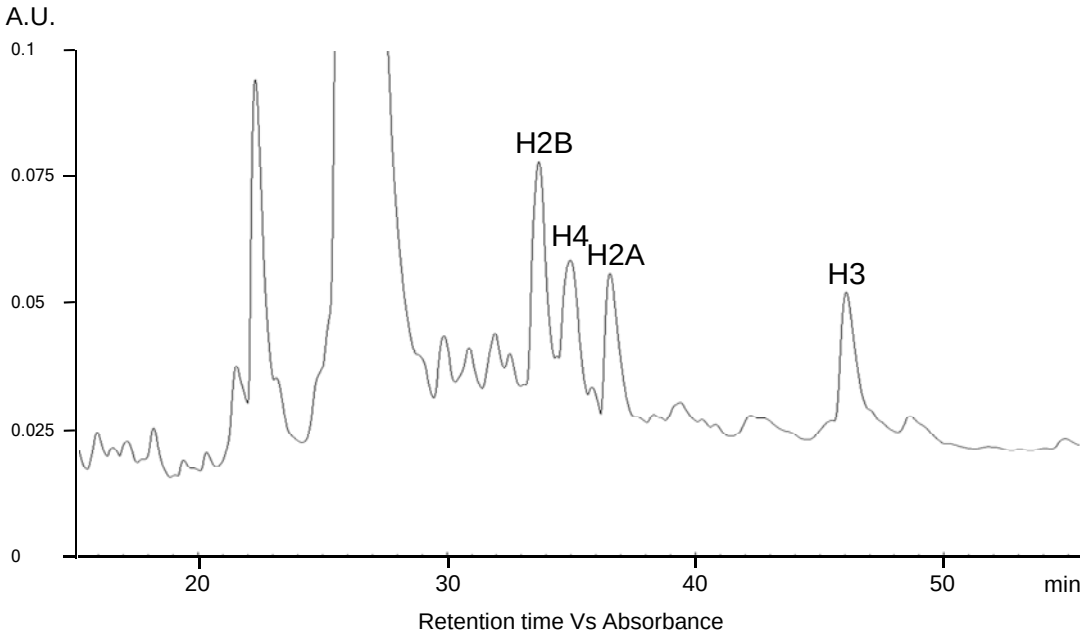


Figure S4.1 The HPLC chromatogram for the separation of *Neurospora* core histones. Histones were eluted in the order of H2B, H4, H2A and H3.

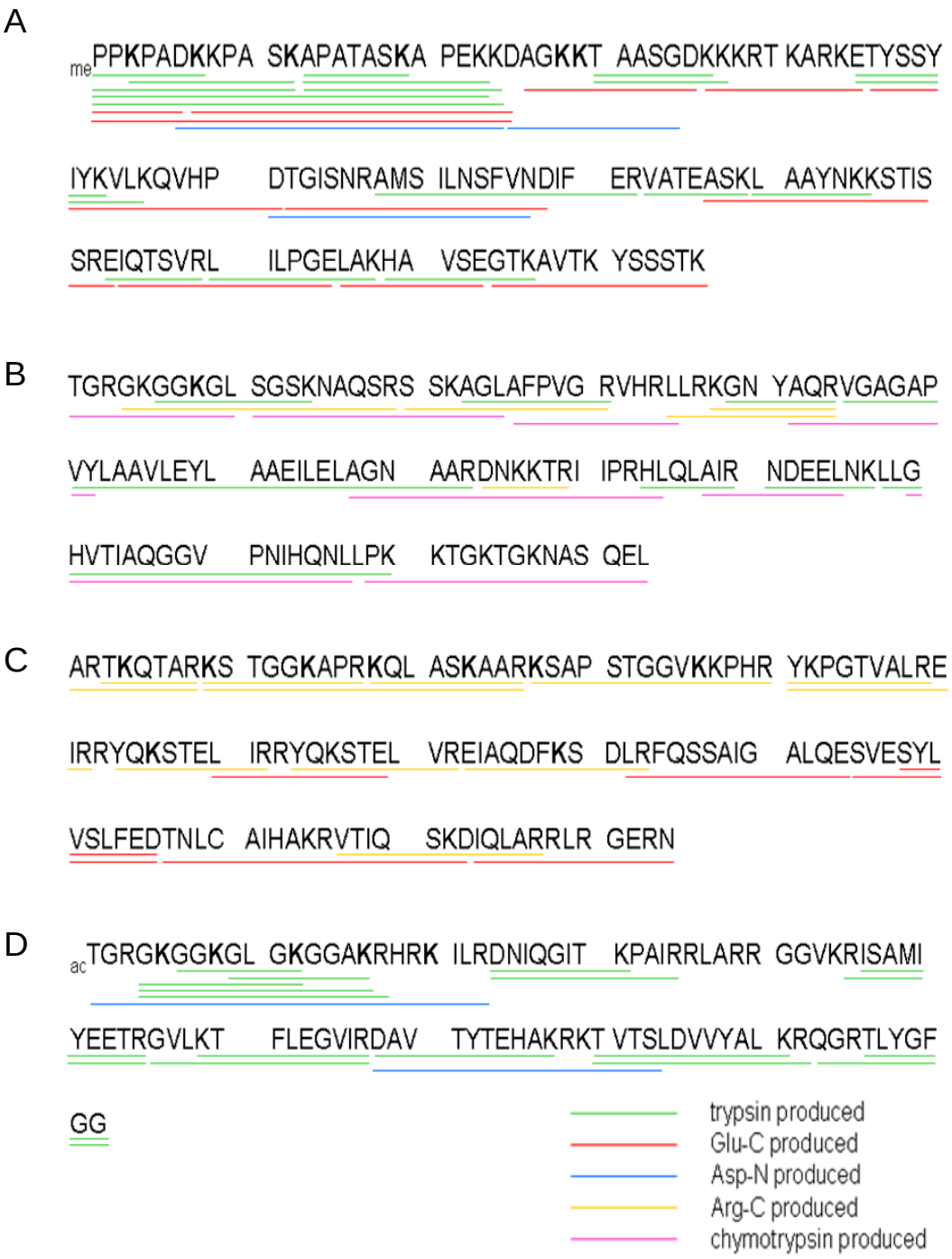


Figure S4.2 Sequence coverage for the *Neurospora* core histones based on MS/MS analyses. The sequences for *Neurospora* core histones were obtained from Swissprot, and the identified peptides produced by different proteases were underlined with different colors. The identified modification sites are shown in bold.

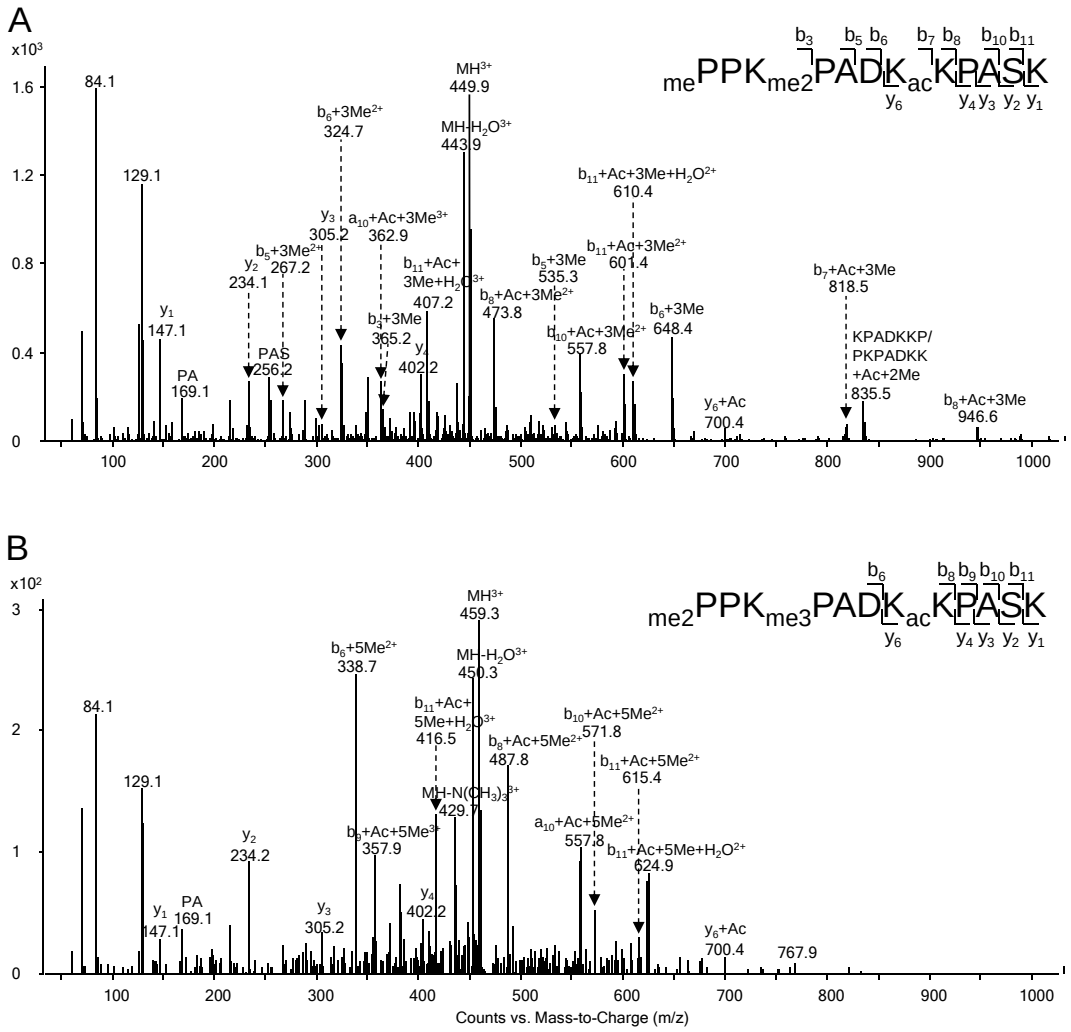


Figure S4.3 The MS/MS of the tri-methylated, mono-acetylated (A) and penta-methylated, mono-acetylated (B) tryptic peptide containing residues 1-12 of *Neurospora* H2B. The data were obtained on the Agilent Q-TOF mass spectrometer.

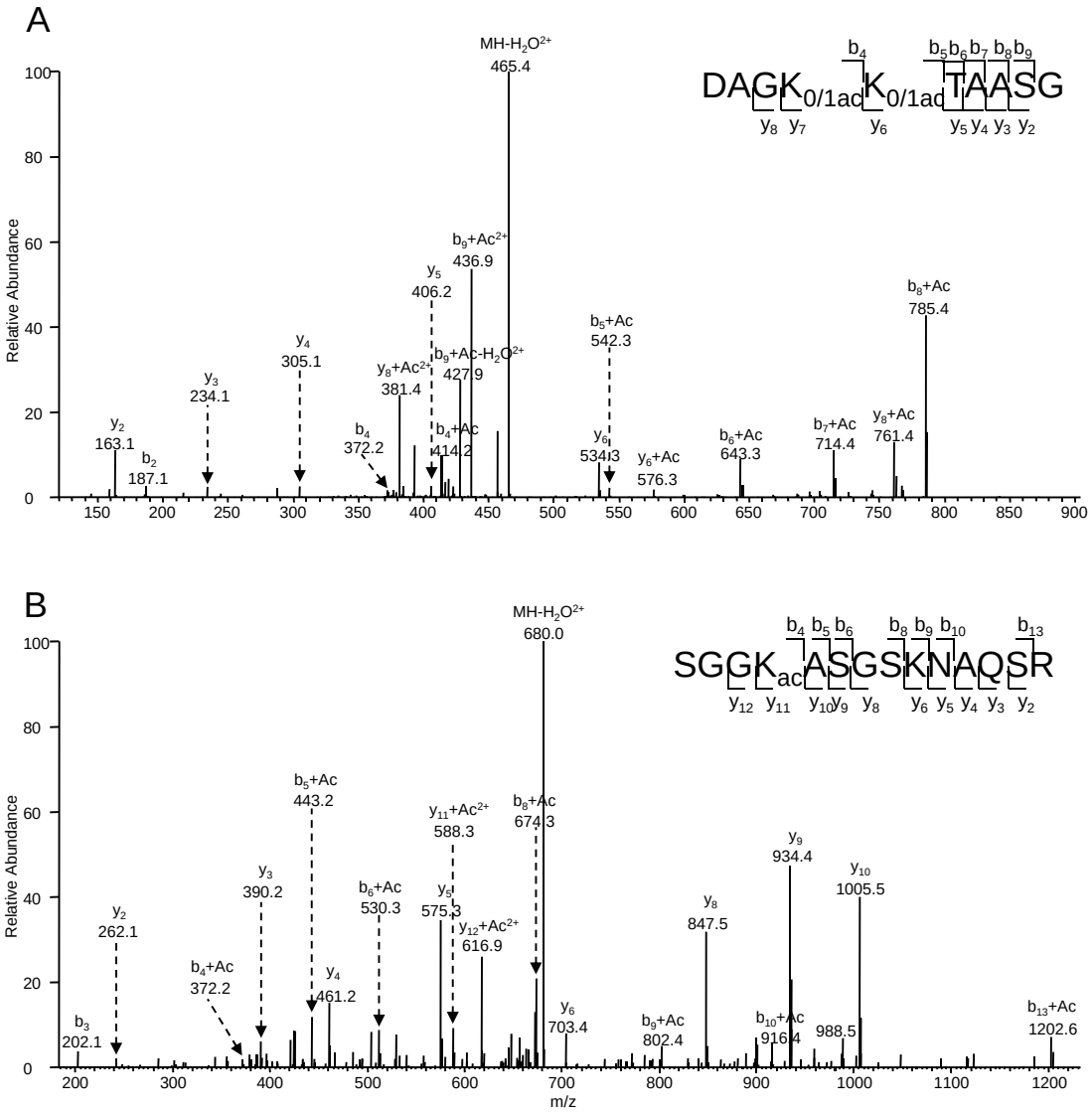
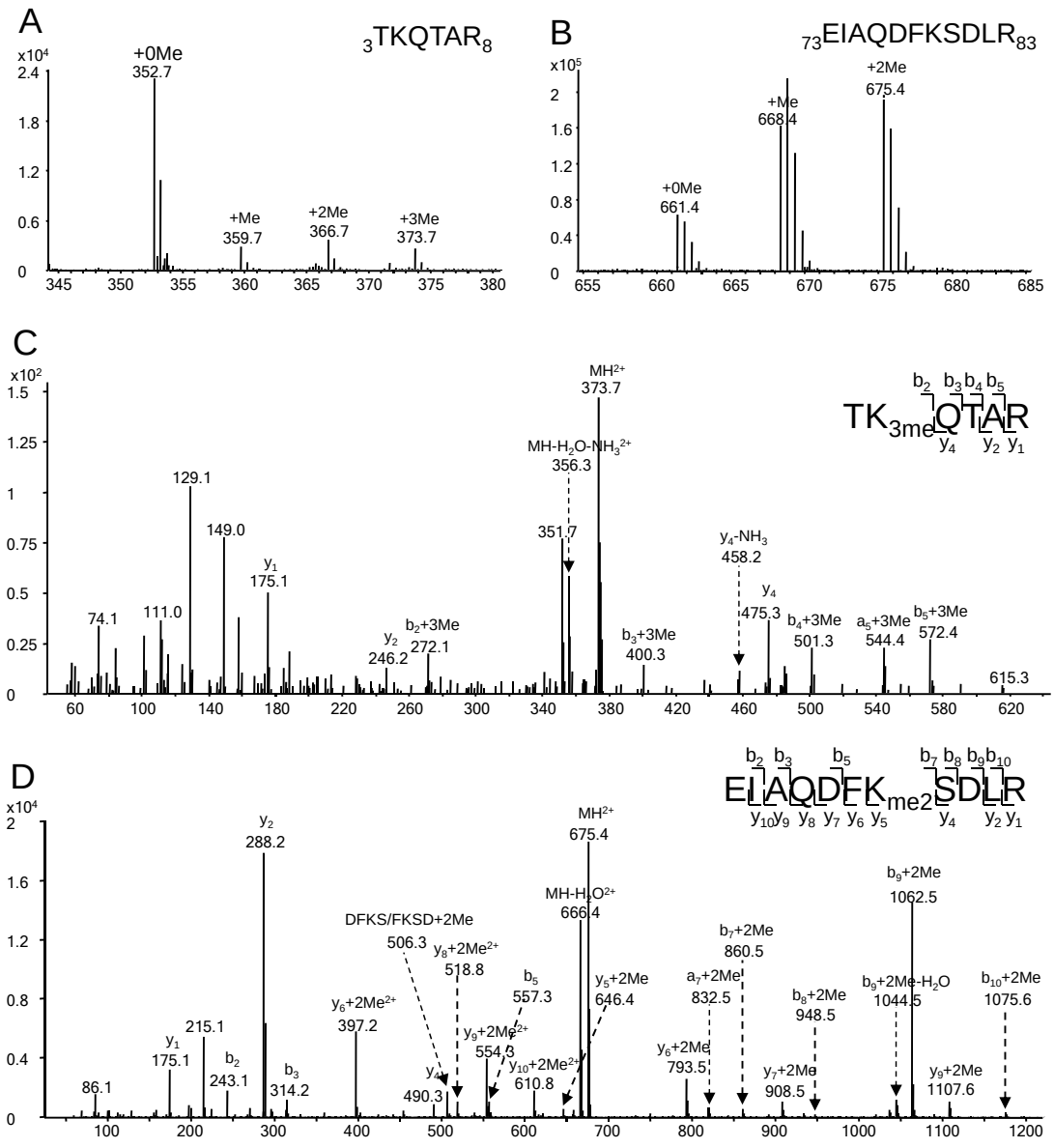


Figure S4.4 The MS/MS of the Asp-N produced peptide containing residues 25-34 of *Neurospora* histone H2B (A) and the Arg-C-produced peptide containing residues 6-19 of *Neurospora* histone H2A (B). The data were acquired on LTQ.


A
$_{3}$TKQTAR$_{8}$
x10^{4}
+0Me
352.7
+Me
359.7
+2Me
366.7
+3Me
373.7
B
$_{73}$EIAQDFKSDLR$_{83}$
x10^{5}
+0Me
661.4
+Me
668.4
+2Me
675.4
C
x10^{2}
MH^{2+}
373.7
MH-H$_2$O-NH$_3^{2+}$
356.3
351.7
129.1
149.0
y$_1$
175.1
74.1
111.0
y$_2$
246.2
b$_2$+3Me
272.1
b$_3$+3Me
400.3
y$_4$-NH$_3$
458.2
y$_4$
475.3
b$_4$+3Me
501.3
a$_5$+3Me
544.4
b$_5$+3Me
572.4
615.3
b$_2$ b$_3$ b$_4$ b$_5$
TK$_{3me}$QTAR
y$_4$ y$_2$ y$_1$
D
x10^{4}
y$_2$
288.2
MH^{2+}
675.4
MH-H$_2$O^{2+}
666.4
DFKS/FKSD+2Me
506.3
y$_8$+2Me^{2+}
518.8
b$_5$
557.3
y$_5$+2Me
646.4
y$_6$+2Me^{2+}
397.2
y$_9$+2Me^{2+}
554.3
y$_{10}$+2Me^{2+}
610.8
y$_4$
490.3
215.1
y$_1$
175.1
b$_2$
243.1
b$_3$
314.2
86.1
b$_7$+2Me
860.5
a$_7$+2Me
832.5
b$_9$+2Me
1062.5
b$_8$+2Me
948.5
b$_9$+2Me-H$_2$O
1044.5
b$_{10}$+2Me
1075.6
y$_6$+2Me
793.5
y$_7$+2Me
908.5
y$_9$+2Me
1107.6
b$_2$ b$_3$ b$_5$ b$_7$ b$_8$ b$_9$b$_{10}$
EIAQDFK$_{me2}$SDLR
y$_{10}$y$_9$ y$_8$ y$_7$ y$_6$ y$_5$ y$_4$ y$_2$ y$_1$

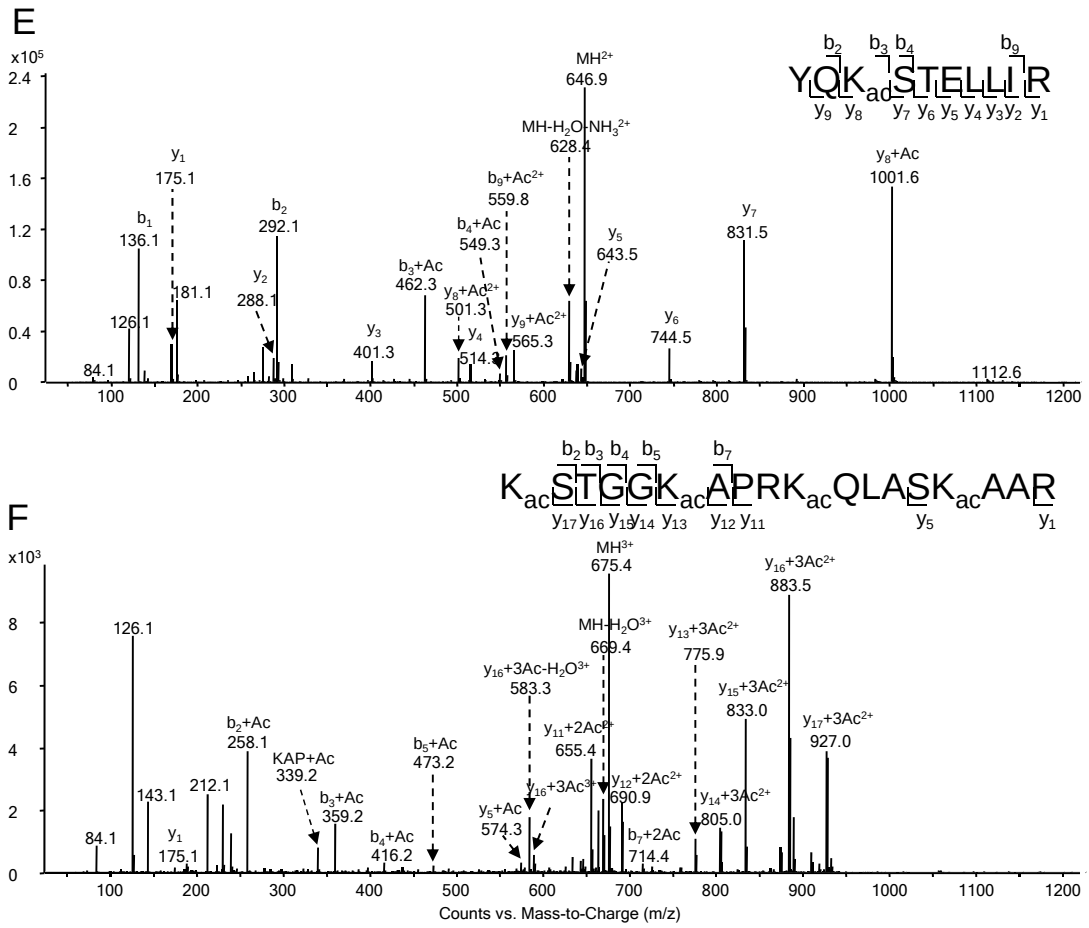


Figure S4.5 The MS of the Arg-C-produced *Neurospora* H3 peptides with residues 3-8 (A) and 73-83 (B). The MS/MS of the H3 peptides containing residues 3-8 with K4 being tri-methylated (C), residues 73-83 with K79 being di-methylated (D), residues 54-63 with K56 being acetylated (E), and residues 9-26 with K9, K14, K18 and K23 being acetylated (F). All data were acquired on the Q-TOF instrument.

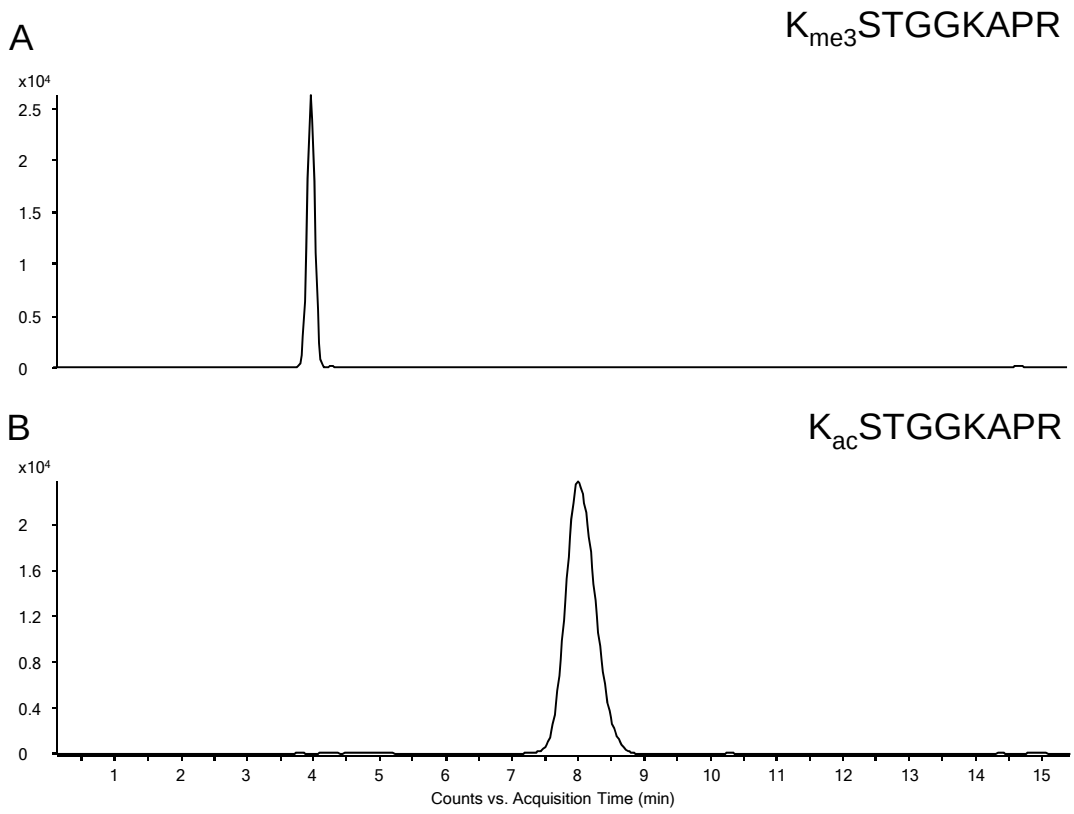


Figure S4.6 The selected-ion chromatograms (SICs) of the triply-charged ions of the tri-methylated (A) and acetylated (B) histone H3 peptide with residues 9-17. These two peptides displayed different retention times during LC-MS/MS analysis on the Q-TOF mass spectrometer.

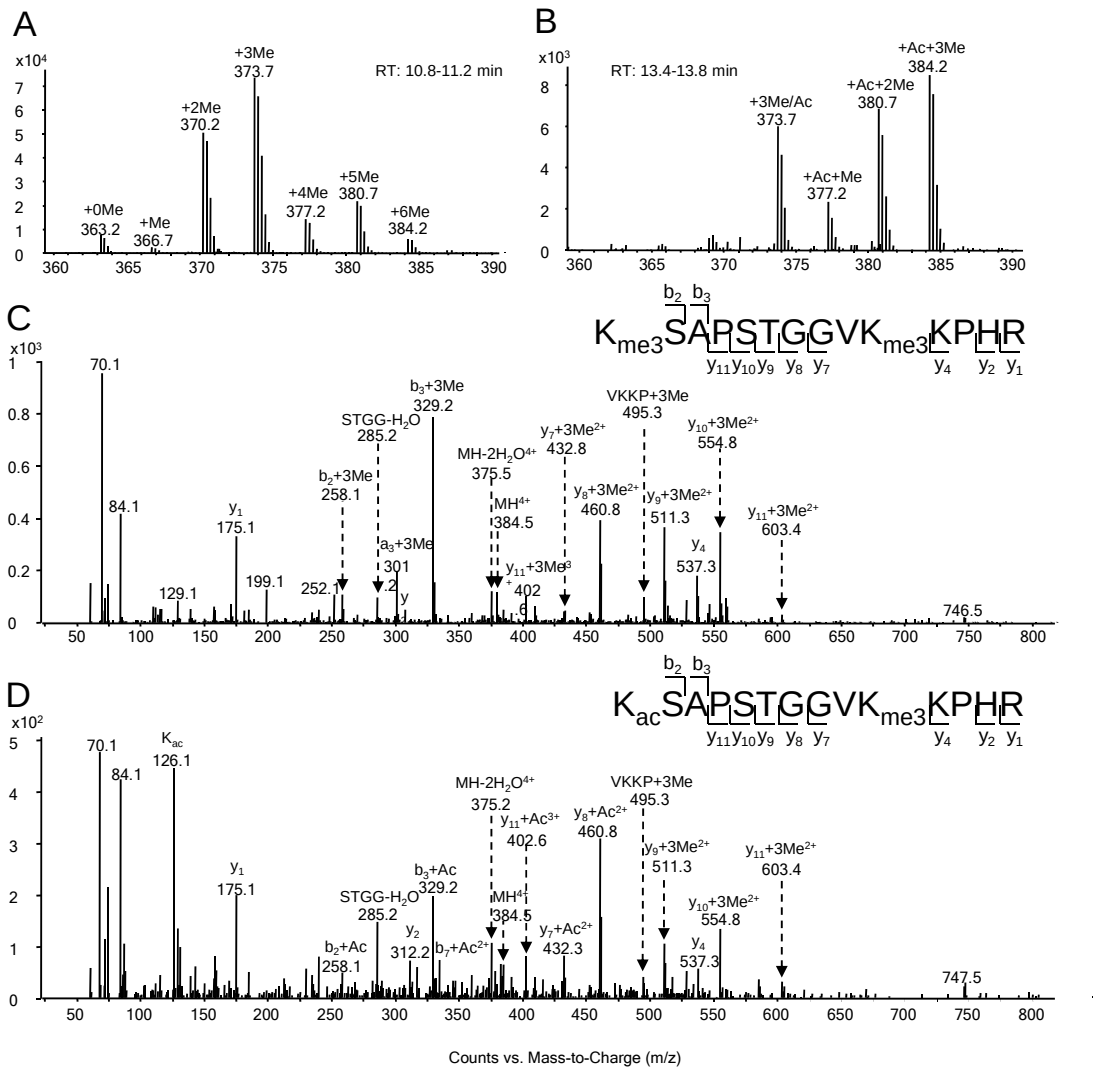


Figure S4.7 The MS of the Arg-C-produced histone H3 peptide bearing residues 27-42 with K36 being tri-methylated and with K27 being tri-methylated (A) or acetylated (B). The MS/MS of the hexa-methylated (C) or tri-methylated, acetylated (D) peptides 27-41 obtained on the Q-TOF.

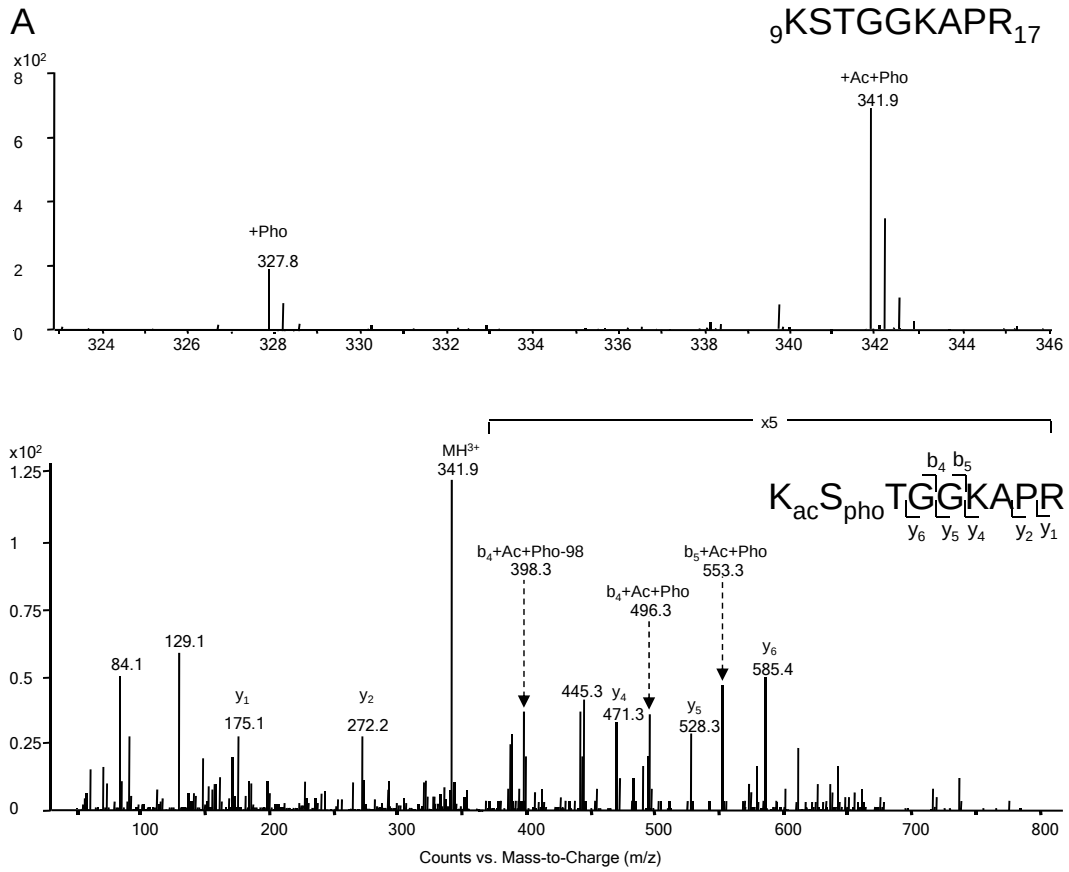


Figure S4.8 (A) The MS of Arg-C-produced phosphorylated histone H3 peptide with residues 9-17. (B) The MS/MS of the peptide containing residues 9-17 with K9 being acetylated and S10 being phosphorylated. The spectra were obtained on the Q-TOF instrument.

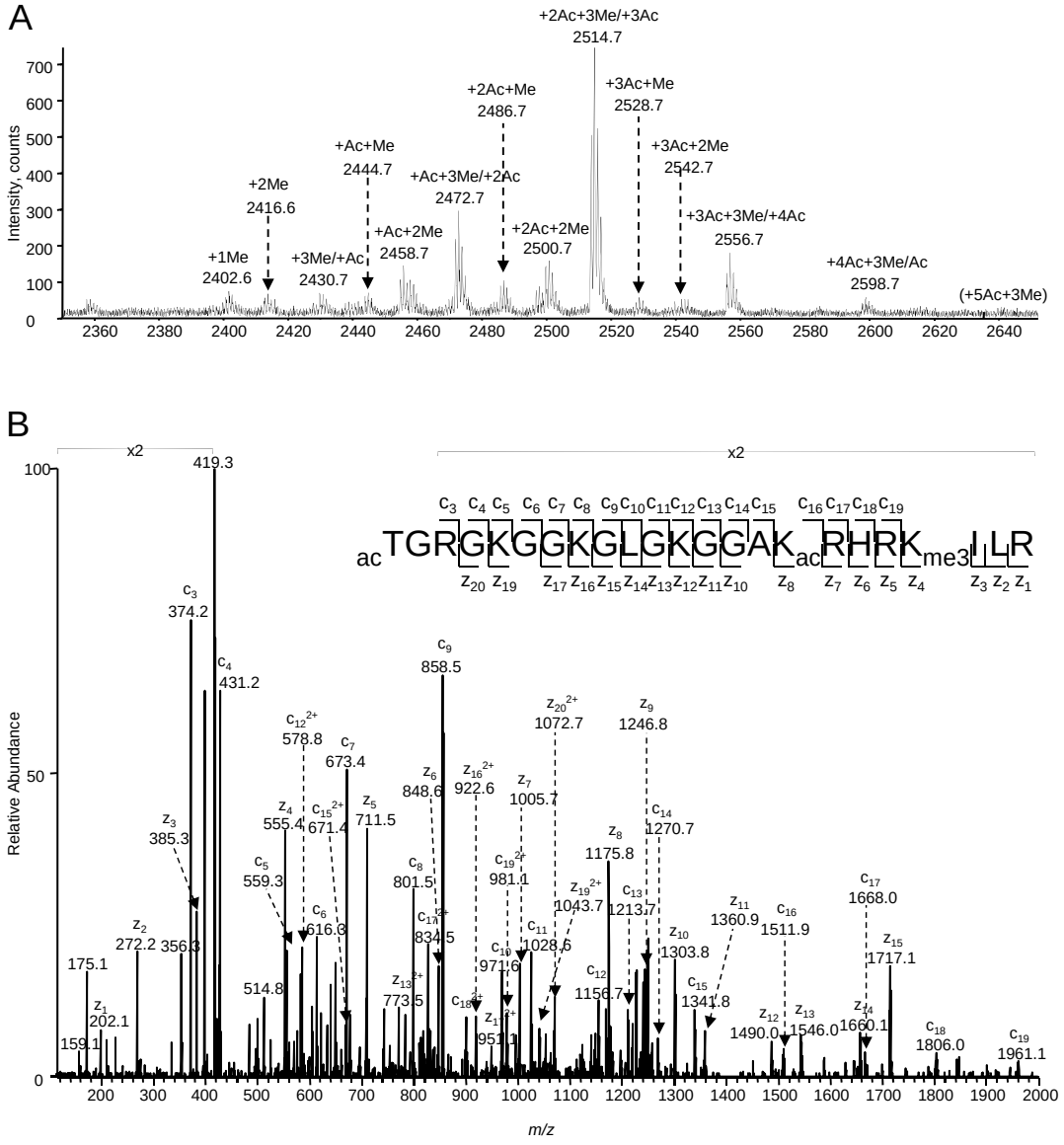


Figure S4.9 (A) The MALDI-MS of the Asp-N-produced *Neurospora* H4 peptide with residues 1-23. (B) The ETD-MS/MS of the same peptide with the N-terminus and K16 being acetylated and K20 being tri-methylated.

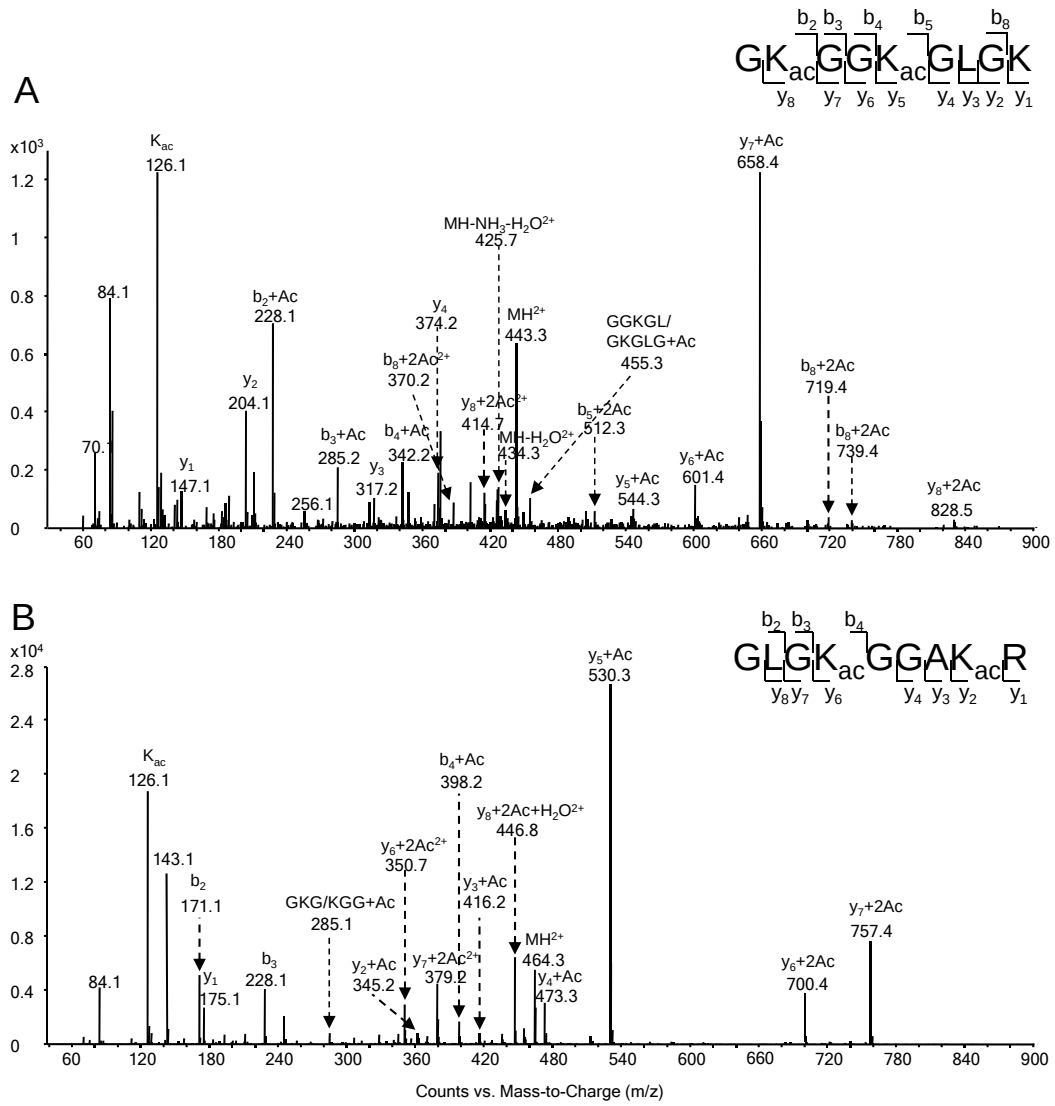


Figure S4.10 The QTOF-produced MS/MS of the trypsin-produced *Neurospora* H4 peptides with residues 4-12 (A) and 9-17 (B) supporting the K5, K8, K12 and K16 acetylation.

CHAPTER 5

Mapping Post-translational Modifications of Histones H2A, H2B and H4 in *Schizosaccharomyces pombe*

Introduction

The eukaryotic nucleosome, which constitutes the fundamental repeating unit of chromatin, plays an important role in packaging and organizing the genetic material [1]. It is comprised of 146 base pairs of DNA wrapped around an octamer of core histone proteins-H2A, H2B, H3, and H4 [1, 2]. Core histone proteins are evolutionarily conserved and consist mainly of flexible amino-terminal tails protruding outward from the nucleosome, and globular carboxy-terminal domains making up the nucleosome scaffold [2]. The histone N-terminal tails are involved in the establishment of chromatin structural states, whereas their histone fold domains mediate histone-histone and histone-DNA interactions [3].

Core histones carry a variety of post-translational modifications (PTMs), which include methylation, acetylation, phosphorylation, ubiquitination, SUMOylation and ADP-ribosylation. These modifications, occurring mainly on the N-terminal tails [4, 5], can affect the interactions of nucleosomes with transacting factors, and are thought to play a role in the assembly and disassembly of chromatin states, ultimately controlling the accessibility of DNA for important cellular processes including transcription, replication, and DNA repair [6-10]. Distinct modifications of the

histone tails can recruit specific chromatin-binding proteins, and modifications on the same or different histone tails may be interdependent and generate various combinations on any individual nucleosome [11-13].

Mass spectrometry has been widely used for assessing histone PTMs. It provides direct information about the sites and types of modifications, differentiates isobaric modifications (e.g., acetylation vs. tri-methylation) [14], and allows for quantitative analysis [15]. New histone modifications identified by mass spectrometry facilitated genome-wide chromatin-related functional studies by using chromatin immunoprecipitation [6, 16-19].

Schizosaccharomyces pombe (*S. pombe*), different from budding yeast *Saccharomyces cerevisiae*, is an excellent model eukaryotic organism as fission yeast for studies on epigenetic regulation [20, 21]. Previous studies on the PTMs of histones in fission yeast and their biological functions were mainly focused on H3. For example, H3 K4 acetylation was found to mediate chromodomain switch which can regulate heterochromatin assembly in fission yeast [22]. K9-methylated histone H3 can bind strongly to the chromodomain of Chp1, which acts upstream of siRNAs during the establishment of centromeric heterochromatin [23]. H3 K56 acetylation plays an important role in DNA damage response [24]. In these previous studies, identification and characterization of PTMs in *S. pombe* histones have all relied on the use of modification-specific antibodies, which may bear some potential problems such as cross-reaction, variable specificity, epitope occlusion, and large consumption of time

and reagents [25]. Very recently, a few mass spectrometry-based investigations of the PTMs of core histones in *S. pombe* have been reported, which concentrated on core histones H3, H4, and a variant of histone H2A (i.e., H2A.Z) [26].

In the present study, we extracted core histones from *S. pombe* and achieved a systematic PTM mapping of histones H2A, H2B and H4, with the combination of digestion with various proteases and analysis with LC-nano-ESI-MS/MS. We report the identification of the conserved PTM sites in *S. pombe* histones that were found previously in other organisms and, more importantly, the unique acetylation sites in H2A and H2B. Our analysis on core histone PTMs sets a stage for examining the regulation of histone modifications and for genome-wide functional studies in fission yeast.

Experimental

Extraction of core histones from S. pombe

S. pombe strain 927 was cultured at 30 °C in a medium containing 0.5% yeast extract, 3% glucose and 0.0225% adenine. Cells were harvested when OD_{600} reached 0.8 and collected by centrifugation at 5000 rpm at 4 ºC for 5 min. The cell pellets were washed with sterile water, resuspended in buffer 1 (0.1 M Tris-HCl, pH 8.0, 0.1 M EDTA, 0.5% 2-mercaptoethanol) and incubated at 30ºC for 10 min. After centrifugation, cells were washed with buffer 2 [1 M sorbitol, 20 mM K_3PO_4, 0.1 mM $CaCl_2$ and 1 mM phenylmethylsulfonyl fluoride (PMSF), pH 6.5] and resuspended in

20 mL of the same buffer containing 10-15 mg zymolyase T20 and incubated at 30ºC for 30 min with gentle shaking to digest the cell wall. The resulting spheroplasts were incubated in an ice-cold nuclei isolation buffer [0.25 M sucrose, 60 mM KCl, 14 mM NaCl, 5 mM $MgCl_2$, 1 mM $CaCl_2$, 15 mM 2-(*N*-morpholino)ethanesulfonic acid, 1 mM PMSF, and 0.8% Triton X-100] on ice for 30 min and homogenized substantially with a mini glass homogenizer. The cell pellets were collected by centrifugation at 4000 rpm for 8 min, and washed again with nuclei isolation buffer, followed by washing three times with buffer A (10 mM Tris, pH 8.0, 0.5% NP-40, 75 mM NaCl, 30 mM sodium butyrate, 1 mM PMSF) and twice with buffer B (10 mM Tris, pH 8.0, 0.4 M NaCl, 30 mM sodium butyrate, 1 mM PMSF). After centrifugation, the resulting pellets were resuspended, with occasional vortexing, in 1 mL 0.4 N sulfuric acid at 4ºC for 1 hr. The core histones in the supernatant were precipitated with cold acetone, centrifuged, dried and redissolved in water.

HPLC separation and protease digestion

Core histones were isolated by HPLC on an Agilent 1100 system (Agilent Technologies, Santa Clara, CA) as described previously [Xiong, 2009 #26]. The wavelength for the UV detector was set at 220 nm. A 4.6×250 mm C4 column (Grace Vydac, Hesperia, CA) was used. The flow rate was 0.8 mL/min, and a 60-min linear gradient of 30-60% acetonitrile in 0.1% trifluoroacetic acid (TFA) was employed.

In order to obtain high sequence coverage, purified histones were digested separately with several proteases, including trypsin, Glu-C, Asp-N, and chymotrypsin. A protein/enzyme ratio of 50:1 (w/w) was employed for trypsin and 20:1 for other proteases. The different buffers used for the digestions were 100 mM NH_4HCO_3 (pH 8.0) for trypsin or Glu-C; 50 mM sodium phosphate (pH 8.0) for Asp-N; and 100 mM Tris-HCl (pH 7.8) along with 10 mM $CaCl_2$ for chymotrypsin. The digestion was carried out overnight at room temperature for chymotrypsin and at 37ºC for other proteases. For the limited tryptic digestion, the same protein/enzyme ratio was used but the incubation time was decreased to 4 hrs. The peptide mixture from the digestion of ~0.2 μg of core histone was subjected directly to LC-MS/MS analysis.

Mass spectrometry

LC-MS/MS experiments were performed on a 6510 QTOF LC/MS system with HPLC-Chip Cube MS interface (Agilent Technologies). The sample enrichment, desalting, and HPLC separation were carried out automatically on the Agilent HPLC-Chip with an integrated trapping column (40 nL) and a separation column (Zorbax 300SB-C18, 75 μm×150 mm, 5 μm in particle size). The flow rates for sample enrichment and peptide separation were 4 μL/min and 0.3 μL/min, respectively. A 60-min linear gradient of 2-35% acetonitrile in 0.1% formic acid was applied for peptide separation. The Chip spray voltage (VCap) was set at 1950 V and varied depending on chip conditions. MS/MS experiments were carried out in either

the data-dependent scan mode or the pre-selected ion mode. The width for precursor ion selection was 4 *m/z* units. The temperature and flow rate for the drying gas were 325ºC and 4 L/min, respectively. Nitrogen was used as collision gas, and collision energy followed a linear equation with a slope of 3 V per 100 m/z units and an offset of 2.5 V. The raw data obtained in the data-dependent scan mode were converted to Mascot generic format files, and submitted to the Mascot database search engine (Matrix Science, Boston, MA) for protein and PTM identification.

Results

To improve the foundation of information on chromatin structure and function in fission yeast, we initiated a systematic investigation of the PTMs of core histones in *S. pombe*. We first extracted core histones from *S. pombe* cells and fractionated individual core histones by using reverse-phase HPLC. Despite multiple attempts using different histone extraction protocols, we were not able to isolate histone H3 from this organism, which might be due to the selective loss of this histone during the extraction processes. The remaining core histones were eluted in the order of H4, H2B and H2A. It is worth noting that, while we were writing up the results of our study, Sinha et al. [27] reported an extraction protocol for core histones from *S. pombe* with the use of glass beads in a beadbeater. With that protocol, these researchers were able to extract core histones including histone H3 from *S. pombe* [27].

Since the PTMs of histones H3 and H4 have been examined by Sinha et al. [27], we decided to focus on the three types of core histones that we were able to isolate and place emphasis on the PTMs of histones H2A and H2B. To this end, we digested the core histones individually with different proteases and analyzed the peptide mixtures with LC-nano-ESI-MS/MS to obtain high sequence coverage and achieve unambiguous PTM assignment. The identified PTM sites are summarized in Figure 5.1 and the sequence coverage is shown in Figure S5.1.

Identification of PTMs in histone H2A

Purified H2A was digested individually with trypsin, Asp-N or chymotrypsin under optimized conditions to obtain peptides in appropriate lengths and good sequence coverage. The digestion mixtures were subjected subsequently to LC-MS/MS analysis, and the acquired mass spectra were searched with Mascot search engine and the results were manually verified.

In H2A, we identified conserved acetylation on K4 and K8, and the unique N-terminal acetylation. Figure 5.2A depicts the MS/MS for the doubly charged H2A N-terminal peptide $_{1}$SGGKSGGKAAVAK$_{13}$with three acetyl groups. The presence of the b_2+Ac, b_3+Ac and y_{10}+2Ac, y_{11}+2Ac, y_{12}+2Ac ions supports unambiguously N-terminal acetylation, and this modification has not been found in other organisms. Additionally, the observation of the b_3+Ac and b_4+2Ac as well as y_9+Ac and y_{10}+2Ac

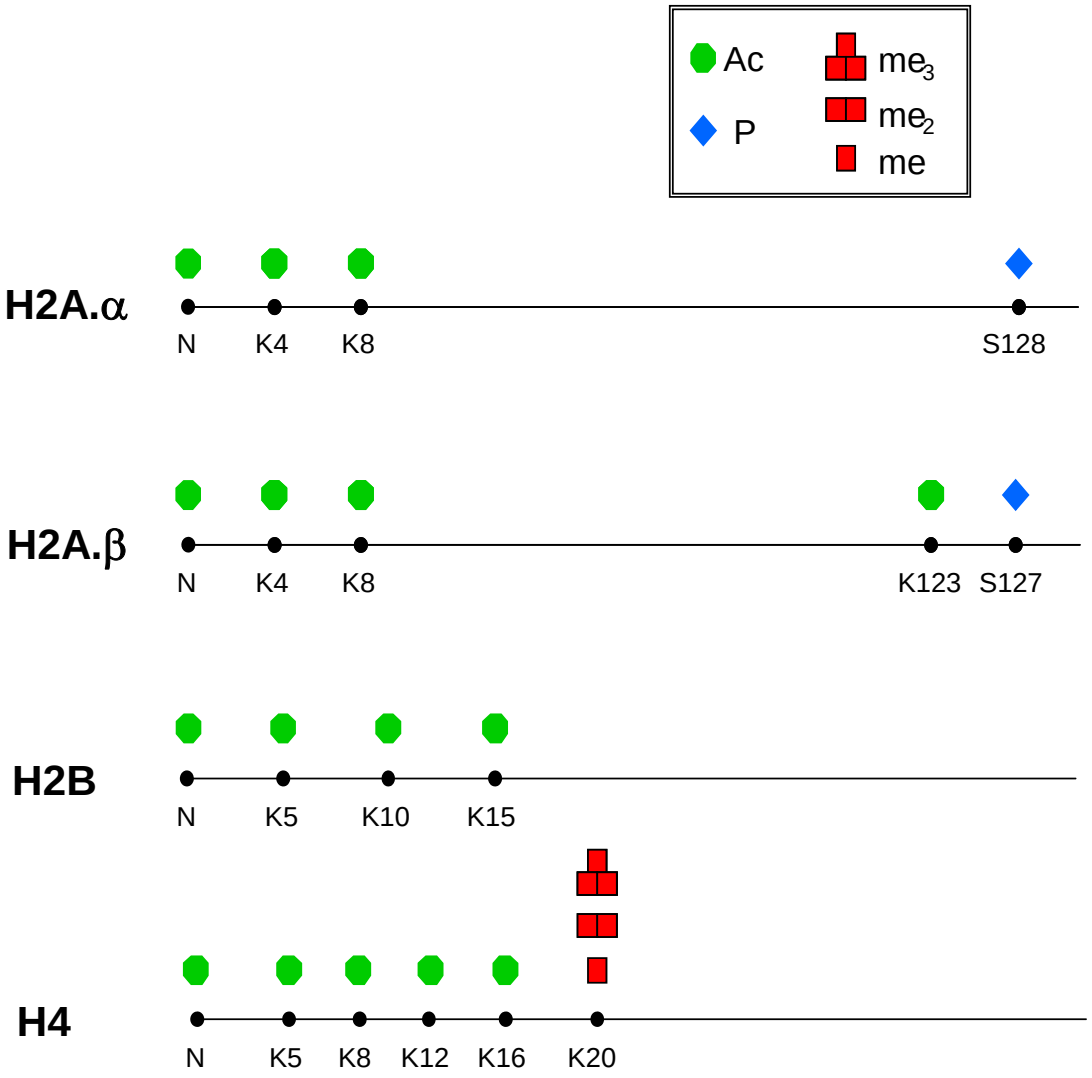


Figure 5.1 Summaries of the detected PTMs of *S. pombe* histones H2A, H2B and H4. The modified residues are labeled, and "N" represents N terminus. Acetylation is designated with solid octagon, phosphorylation is shown with solid diamond, and mono-, di-, and tri-methylation are represented by one-, two, and three square boxes, respectively.

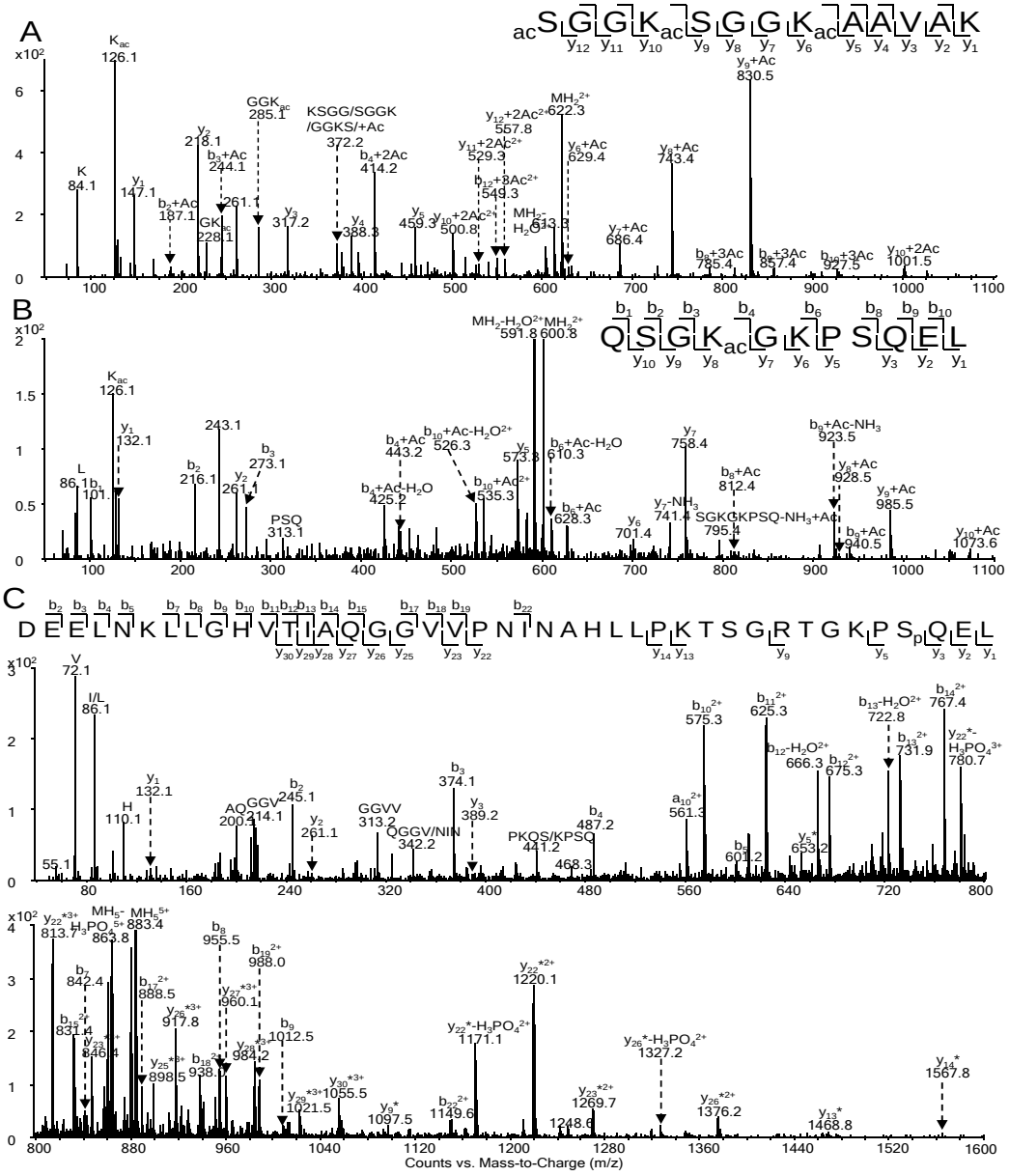


Figure 5.2 ESI-MS/MS of tri-acetylated N-terminal tryptic peptide $_{1}$SGGKSGGKAAVAK$_{13}$ (A), mono-acetylated C-terminal tryptic peptide $_{120}$QSGKGKPSQEL$_{130}$ (B) of histone H2A, and Asp-N-produced phosphorylated peptide $_{91}$DEELNKLLGHVTIAQGGVVPNINAHLLPKTSGRTGKPSQEL$_{131}$ (C) of H2A.α isolated from *S. pombe*, in which b_n* and y_n* designate those fragment ions carrying a phosphorylated residue.

ions suggests K_4 acetylation, whereas the presence of y_5, y_6+Ac, y_7+Ac, and y_8+Ac ions demonstrates K8 acetylation.

More interestingly, we found a novel K123 acetylation in the C-terminal region of H2A.β. In the MS/MS of the H2A.β C-terminal peptide $_{120}$QSGKGKPSQEL$_{130}$ (Figure 5.2B), we observed an almost complete series of b and y ions; the presence of b_1, b_2, b_3, b_4+Ac and y_5, y_6, y_7, y_8+Ac ions provides solid evidence supporting the K123 acetylation.

Differentiation of tri-methylation from acetylation is essential in PTM studies of histones [14]. Besides the typical method based on the immonium ion with *m/z* 126.1 from acetylated lysine and the neutral loss of a trimethylamine (59 Da) from tri-methyl lysine-containing precursor and fragment ions [14], we were able to differentiate, by taking advantage of the high mass accuracy of the QTOF mass spectrometer [28, 29], these modifications based on subtle difference in mass increase of the lysine residue introduced by acetylation and tri-methylation. Taking Figure 5.2B as an example, the measured mass difference between the y_7 and y_9 ions from the peptide segment housing residues 120-130 was 985.5314-758.4048 = 227.1266 Da. This mass difference is in much better agreement with the calculated mass difference with the consideration of K123 acetylation (227.1270 Da, with a mass deviation of 1.8 ppm) than the corresponding mass difference with the consideration of K123 tri-methylation (227.1635 Da, with a mass deviation of 162 ppm). Thus, K123 is

acetylated. All the acetylation and tri-methylation sites of *S. pombe* core histones were unambiguously established with this method, as summarized in Table 5.1.

In addition to acetylation, we observed the phosphorylation of S128 in H2A.α and S127 in H2A.β. The modification sites could again be determined from fragment ions, with the consideration of mass shift +80 Da introduced by phosphorylation and -98 Da from the loss of an H_3PO_4. In the MS/MS of the Asp-N-produced H2A.α peptide $_{91}$DEELNKLLGHVTIAQGGVVPNINAHLLPKTSGRTGKPSQEL$_{131}$ (Figure 5.2C), the existence of y_1, y_2, y_3, y_5*, y_9* ions suggests the S128 phosphorylation ("*" designates those fragmentations carrying a phosphorylated residue), while the similar small y ions formed from the corresponding C-terminal peptide of H2A.β with residues 91-130 support the S127 phosphorylation (Figure S5.2A). It is worth noting that, although we observed the peptide carrying either K123 acetylation or S127 phosphorylation, the same peptide carrying simultaneously K123 acetylation and S127 phosphorylation could not be found.

Identification of PTMs in histone H2B

The isolated histone H2B was digested with trypsin and Glu-C separately and subjected subsequently to LC-MS/MS analysis. A sequence coverage of 100% was reached and multiple conserved acetylation sites including the N-terminus, K5, K10, and K15 were identified. In the MS/MS of the tetra-acetylated peptide $_1$SAAEKKPASKAPAGKAPR$_{18}$ (Figure 5.3), N-terminus was determined to be

Table 5.1 The fragment ion mass comparison of histone peptides to differentiate tri-methylation from acetylation based on MS/MS data acquired on the Agilent 6510 Q-TOF. For each modification site, two b or y ions flanking the modified lysine were chosen. The experimental mass difference (Measured ΔMass) of these two flanking b or y ions was calculated, so were the corresponding theoretical mass differences with the lysine being tri-methylated or acetylated (Calcd. ΔMass, ac/me3). The two mass deviations (M.D.) between the Measured ΔMass and Calcd. ΔMass were further calculated, with one being markedly smaller than the other. The modification type at the target lysine could be determined as the one with smaller deviation.

Peptide	Mod.	Chosen Ions	Measured ΔMass	Calcd. ΔMass,ac	M.D. /ppm	Calcd. ΔMass,me3	M.D. /ppm
H2A: 1-13	N. Ac	b2	187.0698	187.0714	**8.6**	187.1079	203.6
	K4 Ac	y9, y10	170.1053	170.1055	**1.2**	170.1420	215.7
	K8 Ac	y5, y6	170.1071	170.1055	**-9.4**	170.1420	205.1
120-130	K123 Ac	y7, y9	227.1266	227.1270	**1.8**	227.1635	162.4
H2B: 1-18	N. Ac	b2	201.0823	201.0870	**23.4**	201.1235	204.8
	K5 Ac	y13, y14	170.1022	170.1055	**19.4**	170.1422	235.1
	K10 Ac	b9, b10	170.1018	170.1055	**21.8**	170.1422	237.4
	K15 Ac	y3, y5	227.1292	227.1270	**-9.7**	227.1635	151.0
H4: 4-17	K5 Ac	y12, y13	170.1072	170.1055	**10.0**	170.1422	205.7
	K8 Ac	y9, y10	170.1028	170.1055	**15.9**	170.1422	231.6
	K12 Ac	y5, y7	227.1311	227.1270	**-18.1**	227.1635	142.6
	K16 Ac	y1, y2	170.1023	170.1055	**18.8**	170.1422	234.5
1-17	N. Ac	b2	187.0702	187.0714	**6.4**	187.1079	201.5
20-35	K20 3Me	b2	284.2308	284.1969	-119.3	284.2334	**9.1**

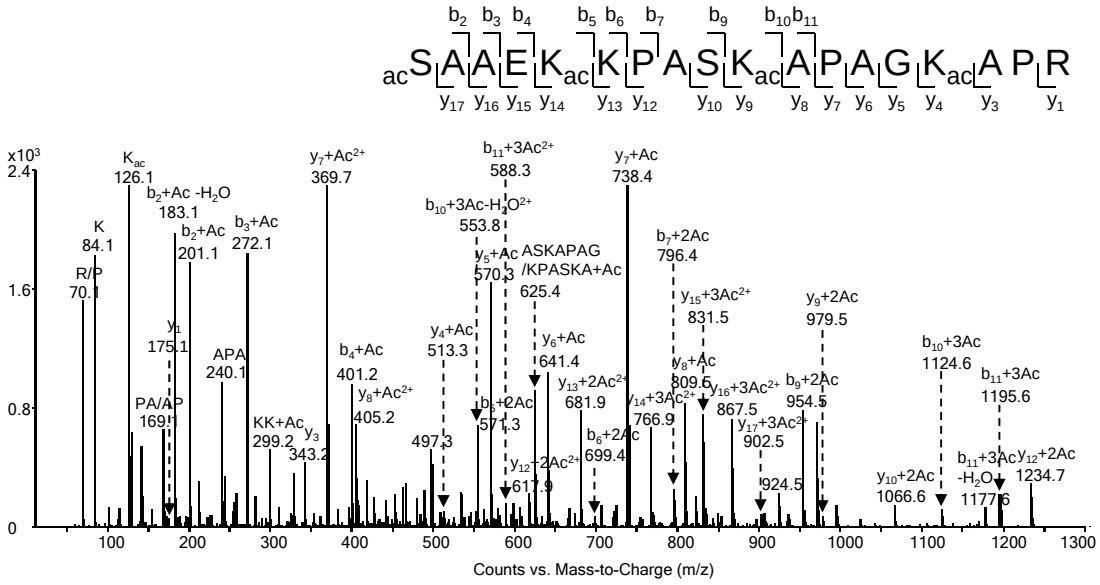


Figure 5.3 The ESI-MS/MS of the *S. pombe* H2B N-terminal peptide $_{1}$SAAEKKPASKAPAGKAPR$_{18}$ with the N-terminus, K5, K10, and K15 being acetylated.

acetylated based on the observation of b_2+Ac and y_{17}+3Ac ions. Mass difference between b_4+Ac and b_5+2Ac, together with that between y_{13}+2Ac and y_{14}+3Ac, supports K5 acetylation. Additionally, the presence of b_9+2Ac and b_{10}+3Ac, along with y_8+Ac and y_9+2Ac ions, demonstrates K10 acetylation. Moreover, K15 acetylation is determined by the formation of y_3 and y_4+Ac ions.

Identification of PTMs in histone H4

Purified histone H4 was digested by trypsin, Asp-N and chymotrypsin separately and subjected to LC-MS/MS analysis. All the modifications were located on the N-terminal segment, similar to the situation observed for other organisms. Upon limited tryptic digestion (i.e., with short incubation time), the miss-cleaved N-terminal peptide $_1$SGRGKGGKGLGKGGAKR$_{17}$ was obtained. Figure 5.4A shows the MS/MS of the di-acetylated peptide with residues 1-17. The formation of b_n+Ac (n = 2, 3, 5-7, 11-13) and y_{16}+Ac ions, but not y_{16}+2Ac ion, supports the N-terminal acetylation, whereas the observation of y_1 and y_2+Ac ions underscores the K16 acetylation. In addition, K5, K8, K12 were all found to be acetylated, as exemplified by the observations of the complete series of b and y ions in the MS/MS of the tetra-acetylated peptide with residues 4-17 (Figure 5.4B).

As conserved in other organisms, K20 in *S. pombe* could be mono-, di- and tri-methylated. In this regard, the positive-ion ESI-MS (Figure S5.3A) reveals the presence of 0, 1, 2 and 3 methyl groups in the peptide $_{20}$KILRDNIQGITKPAIR$_{35}$. The

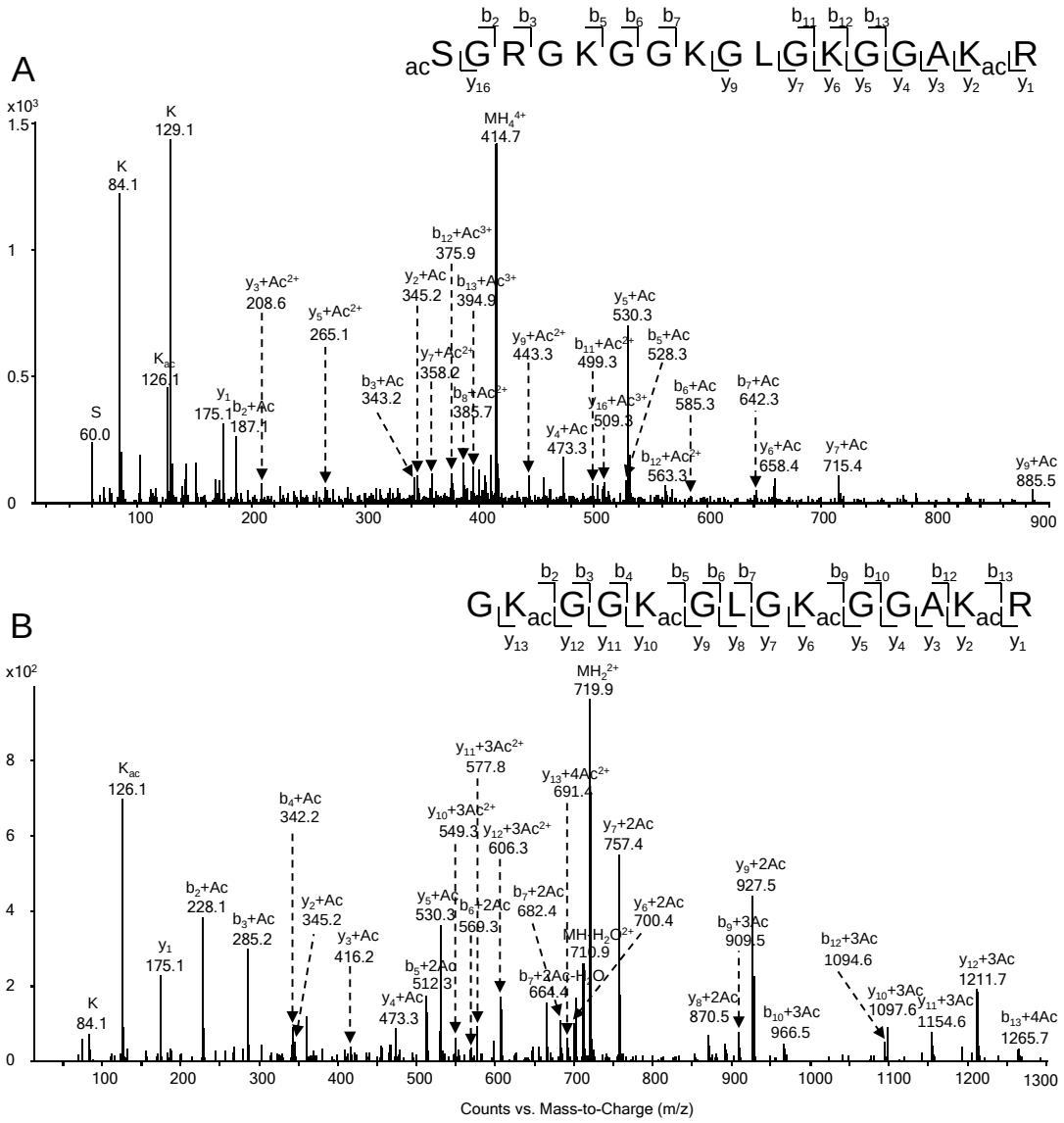


Figure 5.4 The ESI-MS/MS of the *S. pombe* H4 N-terminal peptides, $_{1}$SGRGKGGKGLGKGGAKR$_{17}$ with the N-terminus and K16 being acetylated (A), and $_{4}$GKGGKGLGKGGAKR$_{17}$ with K5, K8, K12 and K16 being acetylated (B).

MS/MS of this group of peptides were all acquired with the selected-ion monitoring (SIM) mode of analysis (Figure S5.3). In the MS/MS of the tri-methylated peptide (Figure S5.3E), the presence of b_2+3Me and y_{14} ions and the abundant fragment ions with the loss of a trimethylamine reveals the tri-methylation on K20. The observation of similar modified small b ions and unmodified large y ions in the MS/MS of the mono- and di-methylated peptides provides solid evidence for K20 mono- and di-methylation, respectively (Figure S5.3C, D). Our above findings with the PTMs of histone H4 are consistent with what was recently reported by Sinha et al. [27].

Discussion and Conclusions

We provided a thorough mapping of PTMs for core histones H2A, H2B and H4 in *S. pombe*, by using LC-nano-ESI-MS/MS coupled with digestion using different proteases. Our work complements the results published very recently by Sinha et al. [27], where the post-translational modifications of core histones H3 and H4 were investigated, and provides a complete picture about the PTMs on core histones in *S. pombe*. The various protease digestions provided high sequence coverage and the accurate mass measurement of fragment ions allowed for unambiguous differentiation of acetylation from tri-methylation. Our results showed that most N-terminal acetylation and methylation sites and H2A C-terminal phosphorylation site are conserved among different organisms including mammals, budding yeast and plants [15, 26, 30-32]. In addition, some novel modifications were detected for the first time.

H2A was found to bear acetylation on N-terminus, K4 and K8, and H2A.β carries an additional acetylation site on K123. In addition, phosphorylation occurs on S128 in H2A.α and S127 in H2A.β. The acetylation sites found in H2B include the N-terminus, K5, K10 and K15. Moreover, H4 was found to carry methylation on K20 and acetylation on the N-terminus, K5, K8, K12, and K16 (Figure 5.1). Nevertheless, it is worth noting that some low levels of modification in *S. pombe* might escape the detection even with the combination of digestion with multiple proteases and the use of a sensitive Q-TOF mass spectrometer with nanospray ionization interface.

Two *S. pombe* H2A isoforms H2A.α and H2A.β are acetylated at K4 and K8, and the similar acetylation was also found in *S. cerevisiae* at K4 and K7. In *S. cerevisiae*, K4 acetylation has been shown to be essential for efficient silencing [33], and H2A K7 acetylation together with H4 N-terminal acetylation was required to maintain chromosome stability [34]. The functions of these acetylations in H2A of *S. pombe* would require further investigation. More interestingly, H2A N-terminal acetylation and H2A.β K123 acetylation were observed here for the first time. It is important to study their functions and crosstalk with other histone modifications. Aside from acetylation, we also observed C-terminal phosphorylation in H2A, which was located at S128 in H2A.α and S127 in H2A.β. These phosphorylations were reported to control Crb2 recruitment at DNA breaks, maintain checkpoint arrest, and influence DNA repair in fission yeast [35].

In *S. pombe* H2B, we found conserved acetylation at K5, K10 and K15. It has been reported that the acetylation of these lysine residues in budding yeast H2B activates the transcription of genes involved in NAD biosynthesis and vitamin metabolism [36]. Thus, it will be interesting to assess the role of H2B acetylation in transcription activation in fission yeast. Moreover, we observed acetylation on the N-terminus of H2B.

Histone H4 acetylation sites, including the N-terminus, and residues K5, K8, K12 and K16, are conserved in almost all organisms, including *S. pombe*. They play important roles in many processes including transcriptional activation, DNA double strand break repair and cellular lifespan regulation [37, 38]. H4 K20 methylation, which is absent in budding yeast, was found in *S. pombe*, in the form of mono-, di- and mainly tri-methylation [39]. It has been reported that this methylation plays important roles in heterochromatin silencing and DNA damage response, similar as H4 in humans and *Drosophila melanogaster* [39-42].

In summary, a systematic PTM mapping of histones H2A, H2B and H4 in *S. pombe* was obtained by mass spectrometric analyses. The rigorous identification of modification sites provides a foundation for further studies on the function of histone PTMs in fission yeast.

References:

1. McGhee, J. D.; Felsenfeld, G., Nucleosome structure. *Annu. Rev. Biochem.* **1980,** 49, 1115-56.

2. Fischle, W.; Wang, Y.; Allis, C. D., Histone and chromatin cross-talk. *Curr. Opin. Cell. Biol.* **2003,** 15, 172-83.

3. Luger, K.; Mader, A. W.; Richmond, R. K.; Sargent, D. F.; Richmond, T. J., Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **1997,** 389, 251-60.

4. Berger, S. L., Histone modifications in transcriptional regulation. *Curr. Opin. Gene. Dev.* **2002,** 12, 142-148.

5. Strahl, B. D.; Allis, C. D., The language of covalent histone modifications. *Nature* **2000,** 403, 41-5.

6. Cheung, P.; Allis, C. D.; Sassone-Corsi, P., Signaling to chromatin through histone modifications. *Cell* **2000,** 103, 263-71.

7. Geiman, T. M.; Robertson, K. D., Chromatin remodeling, histone modifications, and DNA methylation-how does it all fit together? *J. Cell. Biochem.* **2002,** 87, 117-25.

8. Irvine, R. A.; Lin, I. G.; Hsieh, C. L., DNA methylation has a local effect on transcription and histone acetylation. *Mol. Cell. Biol.* **2002,** 22, 6689-96.

9. Nacheva, G. A.; Guschin, D. Y.; Preobrazhenskaya, O. V.; Karpov, V. L.; Ebralidse, K. K.; Mirzabekov, A. D., Change in the pattern of histone binding to DNA upon transcriptional activation. *Cell* **1989,** 58, 27-36.

10. Peterson, C. L.; Laniel, M. A., Histones and histone modifications. *Curr. Biol.* **2004,** 14, R546-R551.

11. Fillingham, J.; Greenblatt, J. F., A histone code for chromatin assembly. *Cell* **2008,** 134, 206-208.

12. Jenuwein, T.; Allis, C. D., Translating the histone code. *Science* **2001,** 293, 1074-80.

13. Klose, R. J.; Zhang, Y., Regulation of histone methylation by demethylimination and demethylation. *Nat. Rev. Mol. Cell. Biol.* **2007,** 8, 307-18.

14. Zhang, K.; Yau, P. M.; Chandrasekhar, B.; New, R.; Kondrat, R.; Imai, B. S.; Bradbury, M. E., Differentiation between peptides containing acetylated or tri-methylated lysines by mass spectrometry: an application for determining lysine 9 acetylation and methylation of histone H3. *Proteomics* **2004,** 4, 1-10.

15. Beck, H. C.; Nielsen, E. C.; Matthiesen, R.; Jensen, L. H.; Sehested, M.; Finn, P.; Grauslund, M.; Hansen, A. M.; Jensen, O. N., Quantitative proteomic analysis of post-translational modifications of human histones. *Mol. Cell. Proteomics* **2006,** 5, 1314-25.

16. Bernstein, B. E.; Humphrey, E. L.; Erlich, R. L.; Schneider, R.; Bouman, P.; Liu, J. S.; Kouzarides, T.; Schreiber, S. L., Methylation of histone H3 Lys 4 in coding regions of active genes. *Proc. Natl. Acad. Sci. U S A* **2002,** 99, 8695-700.

17. Kurdistani, S. K.; Tavazoie, S.; Grunstein, M., Mapping global histone acetylation patterns to gene expression. *Cell* **2004,** 117, 721-33.

18. Meluh, P. B.; Broach, J. R., Immunological analysis of yeast chromatin. *Methods Enzymol.* **1999,** 304, 414-30.

19. Ng, H. H.; Robert, F.; Young, R. A.; Struhl, K., Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol. Cell* **2003,** 11, 709-19.

20. Hendrich, B. D.; Willard, H. F., Epigenetic Regulation of Gene-Expression - the Effect of Altered Chromatin Structure from Yeast to Mammals. *Hum. Mol. Genet.* **1995,** 4, 1765-1777.

21. Verdel, A.; Moazed, D., RNAi-directed assembly of heterochromatin in fission yeast. *Febs. Letters* **2005,** 579, 5872-5878.

22. Xhemalce, B.; Kouzarides, T., A chromodomain switch mediated by histone H3 Lys 4 acetylation regulates heterochromatin assembly. *Genes Dev.* **2010,** 24, 647-52.

23. Schalch, T.; Job, G.; Noffsinger, V. J.; Shanker, S.; Kuscu, C.; Joshua-Tor, L.; Partridge, J. F., High-affinity binding of Chp1 chromodomain to K9 methylated histone H3 is required to establish centromeric heterochromatin. *Mol. Cell* **2009,** 34, 36-46.

24. Xhemalce, B.; Miller, K. M.; Driscoll, R.; Masumoto, H.; Jackson, S. P.; Kouzarides, T.; Verreault, A.; Arcangioli, B., Regulation of histone H3 lysine 56 acetylation in Schizosaccharomyces pombe. *J. Biol. Chem.* **2007,** 282, 15040-7.

25. Garcia, B. A.; Shabanowitz, J.; Hunt, D. F., Characterization of histones and their post-translational modifications by mass spectrometry. *Curr. Opin. Chem. Biol.* **2007,** 11, 66-73.

26. Sinha, I.; Wiren, M.; Ekwall, K., Genome-wide patterns of histone modifications in fission yeast. *Chrom. Res.* **2006,** 14, 95-105.

27. Sinha, I.; Buchanan, L.; Ronnerblad, M.; Bonilla, C.; Durand-Dubief, M.; Shevchenko, A.; Grunstein, M.; Stewart, A. F.; Ekwall, K., Genome-wide mapping of histone modifications and mass spectrometry reveal H4 acetylation bias and H3K36 methylation at gene promoters in fission yeast. *Epigenomics-Uk* **2010,** 2, 377-393.

28. Bahr, U.; Karas, M., Differentiation of 'isobaric' peptides and human milk oligosaccharides by exact mass measurements using electrospray ionization orthogonal time-of-flight analysis. *Rapid Commun Mass Sp* **1999,** 13, 1052-1058.

29. Ono, M.; Shitashige, M.; Honda, K.; Isobe, T.; Kuwabara, H.; Matsuzuki, H.; Hirohashi, S.; Yamada, T., Label-free quantitative proteomics using large peptide data sets generated by nanoflow liquid chromatography and mass spectrometry. *Molecular & Cellular Proteomics* **2006,** 5, 1338-1347.

30. Garcia, B. A.; Hake, S. B.; Diaz, R. L.; Kauer, M.; Morris, S. A.; Recht, J.; Shabanowitz, J.; Mishra, N.; Strahl, B. D.; Allis, C. D.; Hunt, D. F., Organismal

differences in post-translational modifications in histones H3 and H4. *J. Biol. Chem.* **2007,** 282, 7641-55.

31. Schubeler, D.; MacAlpine, D. M.; Scalzo, D.; Wirbelauer, C.; Kooperberg, C.; van Leeuwen, F.; Gottschling, D. E.; O'Neill, L. P.; Turner, B. M.; Delrow, J.; Bell, S. P.; Groudine, M., The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. *Genes Dev.* **2004,** 18, 1263-71.

32. Zhang, K.; Sridhar, V. V.; Zhu, J.; Kapoor, A.; Zhu, J. K., Distinctive core histone post-translational modification patterns in Arabidopsis thaliana. *PLoS One* **2007,** 2, e1210.

33. Wyatt, H. R.; Liaw, H.; Green, G. R.; Lustig, A. J., Multiple roles for Saccharomyces cerevisiae histone H2A in telomere position effect, Spt phenotypes and double-strand-break repair. *Genetics* **2003,** 164, 47-64.

34. Keogh, M. C.; Mennella, T. A.; Sawa, C.; Berthelet, S.; Krogan, N. J.; Wolek, A.; Podolny, V.; Carpenter, L. R.; Greenblatt, J. F.; Baetz, K.; Buratowski, S., The Saccharomyces cerevisiae histone H2A variant Htz1 is acetylated by NuA4. *Genes Dev.* **2006,** 20, 660-5.

35. Nakamura, T. M.; Du, L. L.; Redon, C.; Russell, P., Histone H2A phosphorylation controls Crb2 recruitment at DNA breaks, maintains checkpoint arrest, and influences DNA repair in fission yeast. *Mol. Cell. Biol.* **2004,** 24, 6215-30.

36. Parra, M. A.; Kerr, D.; Fahy, D.; Pouchnik, D. J.; Wyrick, J. J., Deciphering the roles of the histone H2B N-terminal domain in genome-wide transcription. *Mol. Cell. Biol.* **2006,** 26, 3842-52.

37. Bird, A. W.; Yu, D. Y.; Pray-Grant, M. G.; Qiu, Q.; Harmon, K. E.; Megee, P. C.; Grant, P. A.; Smith, M. M.; Christman, M. F., Acetylation of histone H4 by Esa1 is required for DNA double-strand break repair. *Nature* **2002,** 419, 411-5.

38. Dang, W.; Steffen, K. K.; Perry, R.; Dorsey, J. A.; Johnson, F. B.; Shilatifard, A.; Kaeberlein, M.; Kennedy, B. K.; Berger, S. L., Histone H4 lysine 16 acetylation regulates cellular lifespan. *Nature* **2009,** 459, 802-7.

39. Sanders, S. L.; Portoso, M.; Mata, J.; Bahler, J.; Allshire, R. C.; Kouzarides, T., Methylation of histone H4 lysine 20 controls recruitment of Crb2 to sites of DNA damage. *Cell* **2004,** 119, 603-614.

40. Fang, J.; Feng, Q.; Ketel, C. S.; Wang, H. B.; Cao, R.; Xia, L.; Erdjument-Bromage, H.; Tempst, P.; Simon, J. A.; Zhang, Y., Purification and functional characterization of SET8, a nucleosomal histone H4-lysine 20-specific methyltransferase. *Curr. Biol.* **2002,** 12, 1086-1099.

41. Rice, J. C.; Nishioka, K.; Sarma, K.; Steward, R.; Reinberg, D.; Allis, C. D., Mitotic-specific methylation of histone H4 Lys 20 follows increased PR-Set7 expression and its localization to mitotic chromosomes. *Genes Dev.* **2002,** 16, 2225-2230.

42. Schotta, G.; Lachner, M.; Sarma, K.; Ebert, A.; Sengupta, R.; Reuter, G.; Reinberg, D.; Jenuwein, T., A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev.* **2004,** 18, 1251-1262.

Supporting Information for Chapter 5

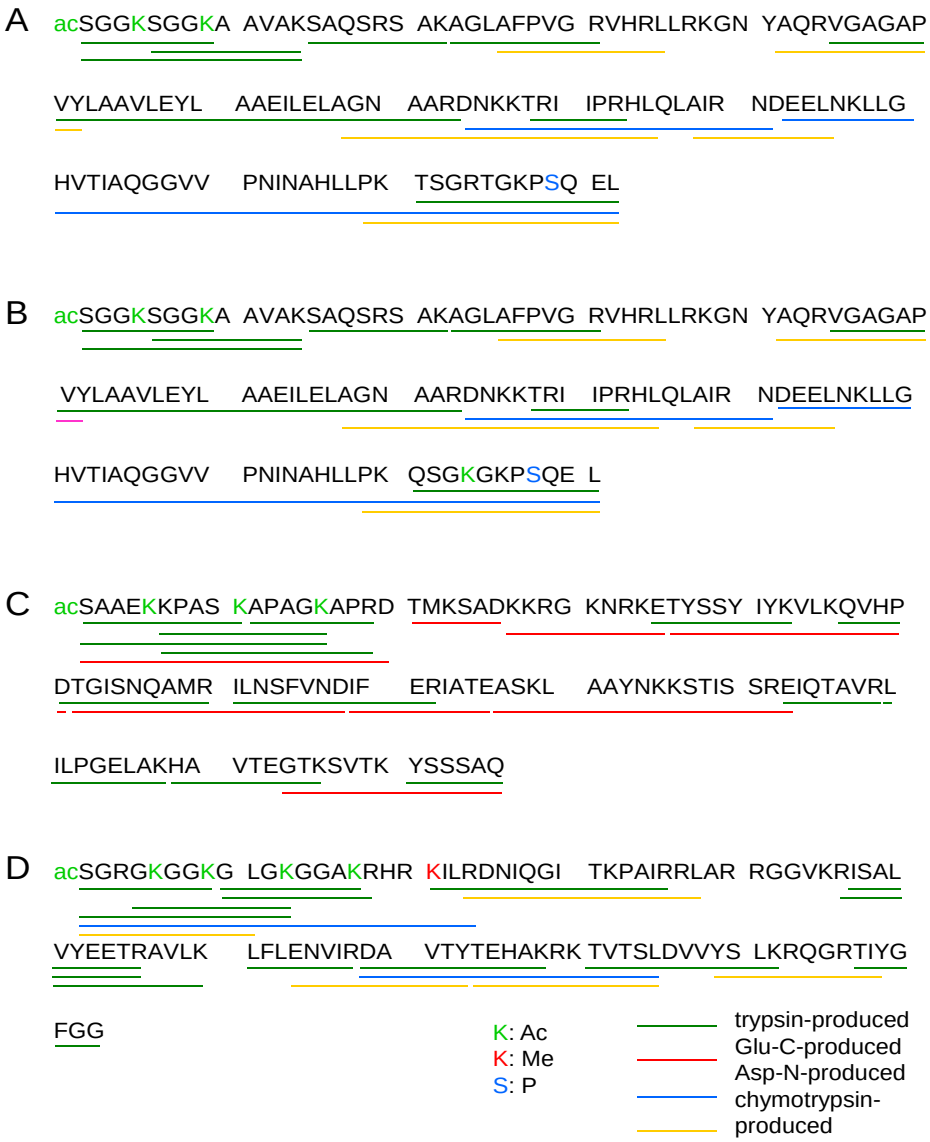


Figure S5.1 Sequence coverage for the *S. pombe* histone H2A.α (A), H2A.β (B), H2B (C), and H4 (D), based on MS/MS analyses. The sequences for *S. pombe* histones were obtained from Swissprot, and the identified peptides produced by different proteases were underlined with different colors. The identified modification sites are shown in different colors as indicated at the bottom of the figure.

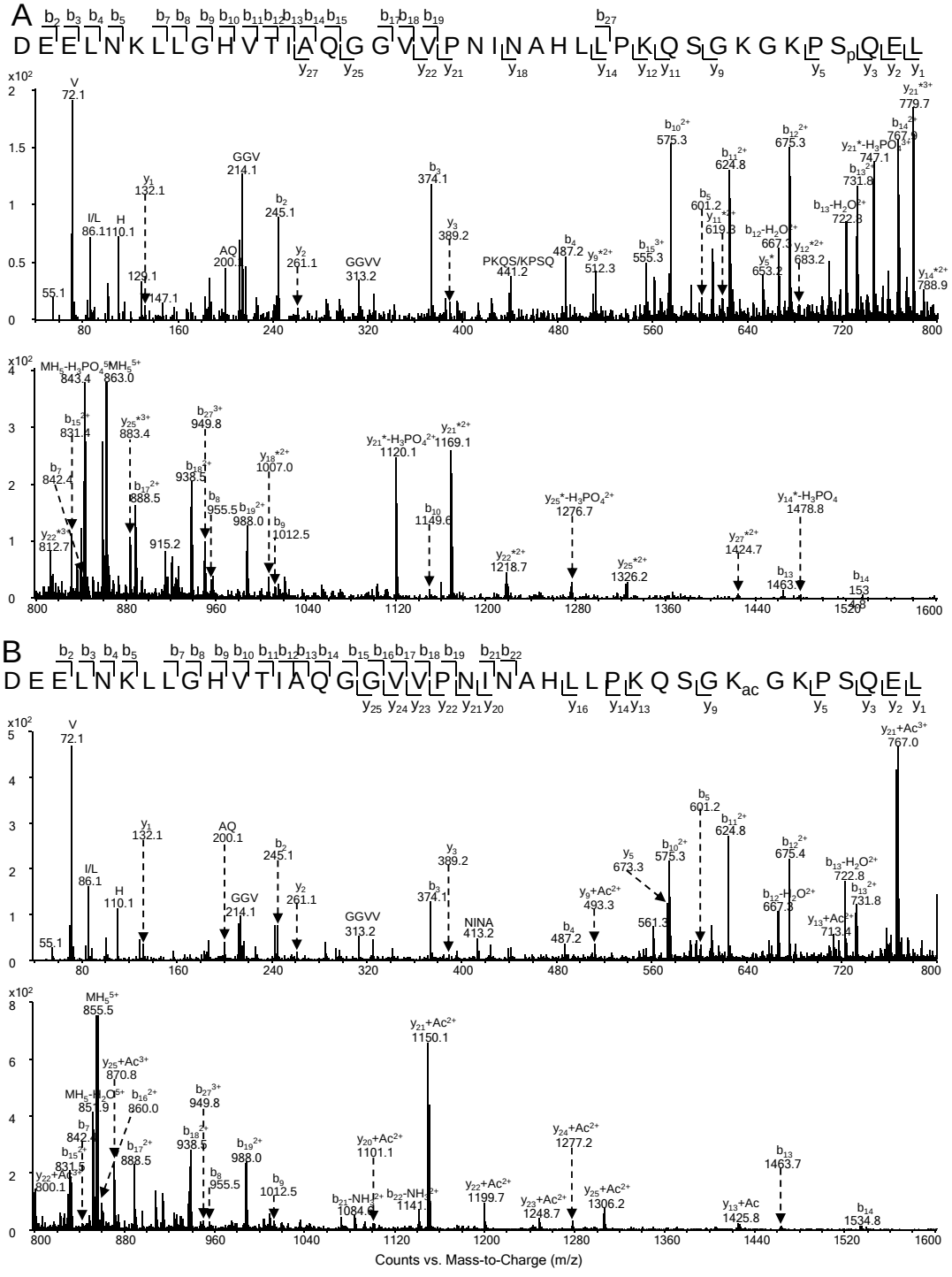
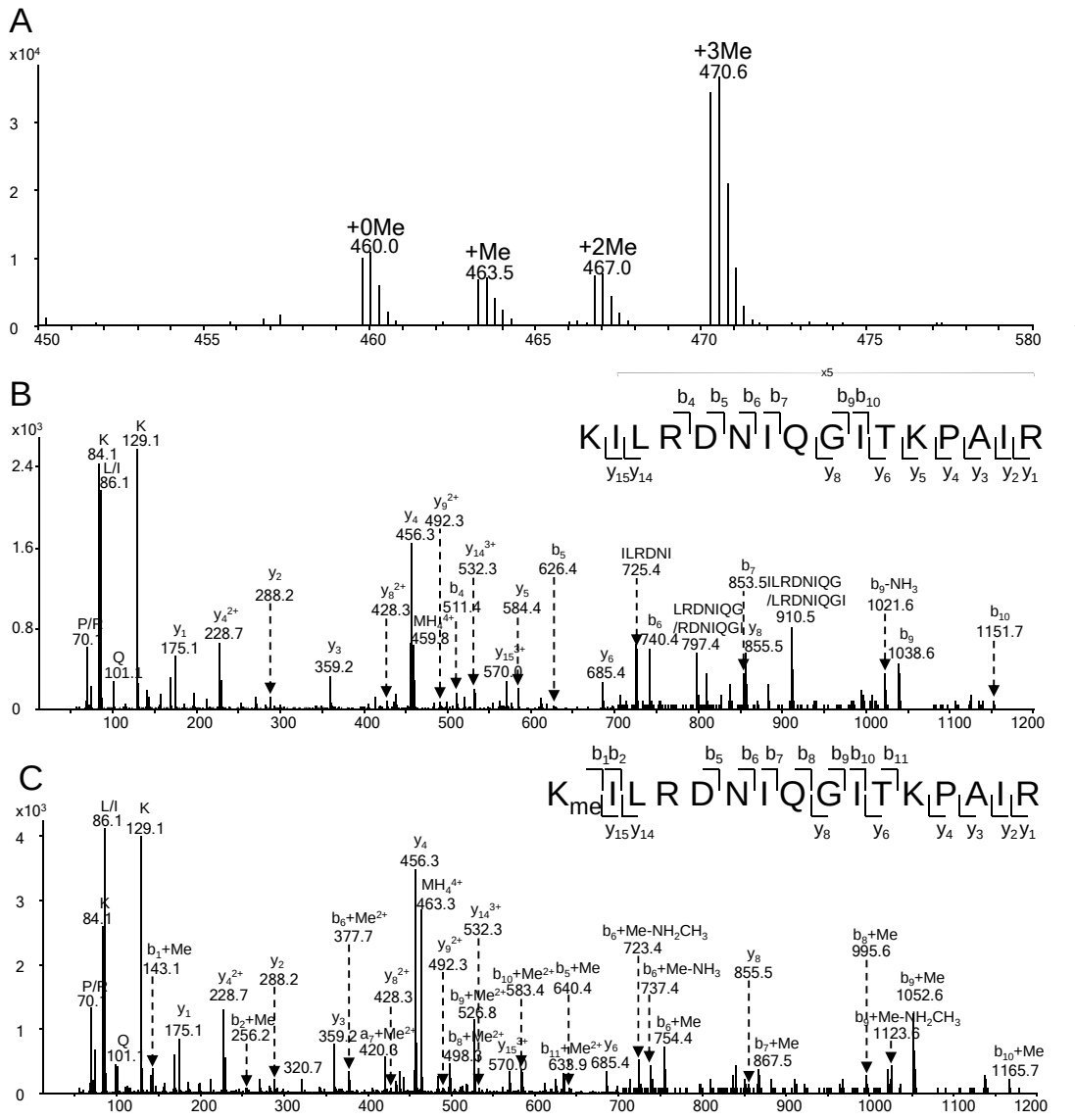


Figure S5.2 ESI-MS/MS of Asp-N-produced phosphorylated (A) and acetylated (B) peptide with residues 90-130 of H2A.β isolated from *S. pombe*, b* and y* designate those fragment ions carrying a phosphorylated residue.


A
+3Me
470.6
+0Me
460.0
+Me
463.5
+2Me
467.0
B
K I L R D N I Q G I T K P A I R
C
K me I L R D N I Q G I T K P A I R

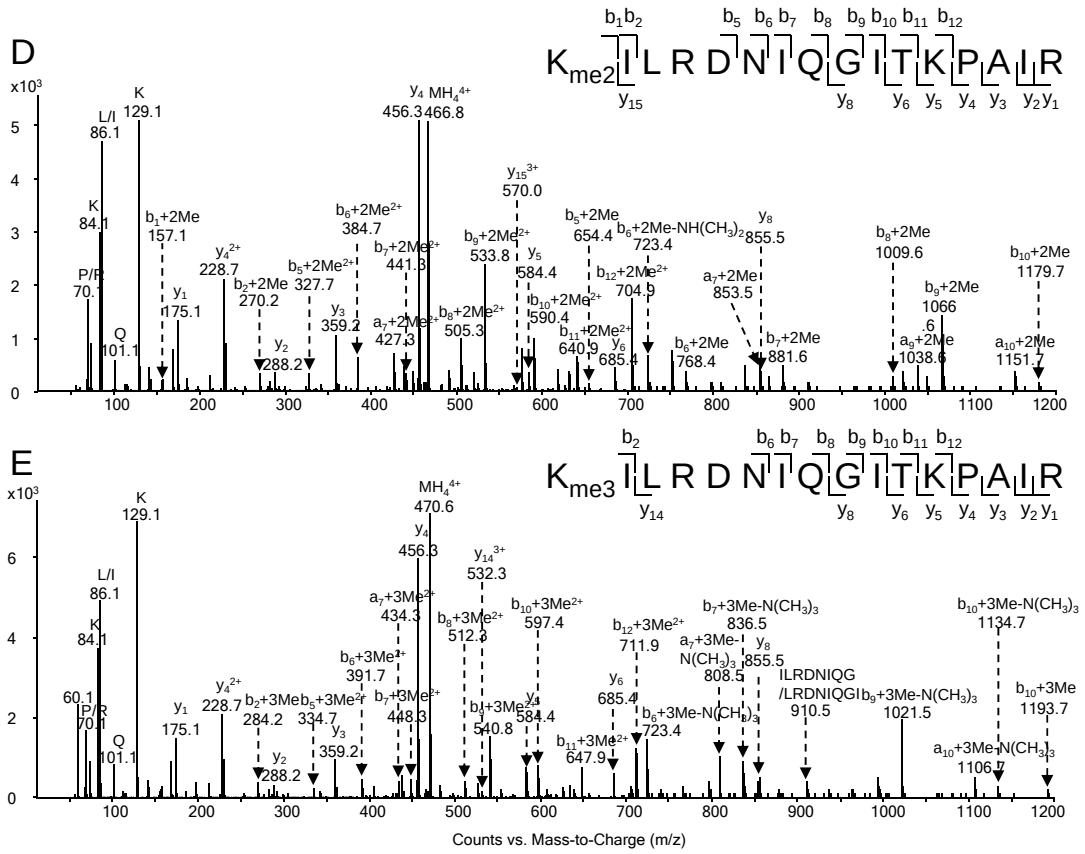


Figure S5.3 Positive-ion ESI-MS (A) of the *S. pombe* H4 peptide $_{20}$KILRDNIQGITKPAIR$_{35}$ with K20 being unmodified or, mono-, di- and tri-methylated, and ESI-MS/MS of the unmodified (B), mono- (C), di- (D) and tri-methylated E) peptide with residues 20-35.

CHAPTER 6

Methyl Group Migration during the Fragmentation of Singly Charged Ions of Trimethyllysine-containing Peptides: Precaution of Using MS/MS of Singly Charged Ions for Interrogating Peptide Methylation

Introduction

The nucleosome core particle consists of a histone octamer around which DNA is wrapped [1]. Two H2A-H2B dimers flank a centrally located $(H3\text{-}H4)_2$ tetramer to afford the histone octamer [1]. The core histones have a similar structure with a basic N-terminal domain, a globular domain organized by the histone fold, and a C-terminal tail. Core histone N-terminal tails, which emerge from the core particle in all directions, are involved in the establishment of a spectrum of chromatin structural states, while their histone fold domains mediate histone-histone and histone-DNA interactions [1].

The core histones are susceptible to an array of post-translational modifications (PTMs), including acetylation, phosphorylation, methylation and ubiquitination [2, 3]. Histone methylation, which occurs on the side chains of lysine and arginine, is the most prominent in histones H3 and H4, and it is associated with transcriptional activation, differentiation, imprinting, and X-inactivation [3, 4]. In general, methylation at H3-K4, H3-K36, and H3-K79 is associated with euchromatin

and gene activation, whereas methylation at H3-K9, H3-K27, and H4-K20 is involved with heterochromatin and repressed genes. Moreover, histone methylation, together with acetylation and phosphorylation, can form a histone code to provide a "mark" to recruit downstream chromatin assembly or modification proteins for chromatin remodeling and transcription activation [3, 4].

MALDI-MS/MS and LC-ESI-MS/MS have been widely used for assessing the PTMs of histones [5, 6]. Together with HPLC separation and enzymatic digestion, they can provide detailed information about the modification sites and levels. Understanding peptide fragmentation is important for investigating the PTMs of proteins by MS/MS with the "bottom-up" strategy. In this context, a peptide can be cleaved, by surface-induced dissociation (SID), collision-induced dissociation (CID), at the amide linkages to afford a series of b and y ions through a 'charge-directed' pathway where cleavage occurs in the vicinity of a charge site. A 'mobile proton transfer' model proposed by Wysocki et al. [7-13] is widely accepted for rationalizing the 'charge-directed' fragmentation of peptides, where a proton is transferred from the peptide N terminus or side chains to the cleavage site (Scheme 1) [14].

Here, we observed an unusual inconsistency between MALDI-MS/MS and ESI-MS/MS when analyzing trimethyllysine-containing peptides from the Arg-C digestion of histone H3 isolated from cultured human and yeast *Saccharomyces cerevisiae* (*S. cerevisiae*) cells. It turned out that a methyl group in trimethyllysine could migrate during the fragmentation of singly charged ions of these peptides.

Based on these findings, we advocate that caution should be exerted while MS/MS of singly charged ions is employed for assessing peptide methylation.

Experimental

Materials

Trimethyllysine-containing synthetic peptides were obtained from Biomatik (Ontario, Canada), and N^G-monomethyl-L-arginine was obtained from Sigma-Aldrich (St. Louis, MO).

Cell culture and protein extraction

Human HL-60 cells [American Type Culture Collection (ATCC)**,** Manassas, VA] were cultured in RPMI-1640 medium (ATCC) with 10% FBS at 37 °C until the cell density reached between 10^5 and 10^6 cells/mL. Human MCF-7 cells (ATCC) were cultured in DMEM medium (ATCC) with 10% FBS at 37 °C until the cells were at 80% confluence level.

The cells were harvested by centrifugation at 5000 rpm. The cell pellets were subsequently washed with a 5-mL buffer containing 0.25 M sucrose, 0.01 M $MgCl_2$, 0.5 mM PMSF, 50 mM Tris (pH 7.4), and 0.5% Triton X-100. The cell pellets were then resuspended in 5 mL of the same buffer and kept at 4 °C overnight [15]. The histones were extracted from the cell lysis mixture with 0.4 N sulfuric acid by

incubating at 4 °C for 4 hrs with continuous vortexing. The histones in the supernatant were precipitated with cold acetone, centrifuged, dried and redissolved in water.

Wild-type BY4742 *Saccharomyces cerevisiae* cells (Open Biosystems, Huntsville, AL) were cultured in a medium containing 1% yeast extract, 2% peptone and 2% glucose, and the cells were harvested when the OD_{600} reached between 1 and 2. Cells were centrifuged at 5000 rpm at 4 ºC for 10 min, and the resulting cell pellets were washed with sterile water and resuspended in a solution bearing 0.1 mM Tris (pH 9.4) and 10 mM DTT. The mixture was incubated at 30 ºC for 15 min with gentle shaking. The cells were recovered by centrifugation, the cell pellets were washed with a solution containing 1.2 M sorbitol and 20 mM HEPES (pH 7.4), and centrifuged again. The resulting cell pellets were resuspended in the same buffer (50 mL) containing 20-30 mg zymolyase and incubated at 30 ºC for 30 min with gentle shaking to digest the cell wall. Cells were subsequently washed twice with ice-cold nuclei isolation buffer (0.25 M sucrose, 60 mM KCl, 14 mM NaCl, 5 mM $MgCl_2$, 1 mM $CaCl_2$, 15 mM MES, 1 mM PMSF, 0.8% Triton X-100), followed by washing three times with buffer A (10 mM Tris, pH 8.0, 0.5% NP-40, 75 mM NaCl, 30 mM sodium butyrate, 1 mM PMSF) and twice with buffer B (10 mM Tris, pH 8.0, 0.4 M NaCl, 30 mM sodium butyrate, 1 mM PMSF) [16, 17]. The cell pellets were centrifuged and resuspended, with occasional vortexing, in 0.4 N sulfuric acid (3 mL) at 4 ºC for 1 hr. The histones in the supernatant were precipitated by cold acetone, centrifuged, dried and redissolved in water.

Histone H3 isolation and digestion

Histone H3 was isolated from the core histone mixture by HPLC on an Agilent 1100 system (Agilent Technologies, Palo Alto, CA) and a 4.6 × 250 mm C4 column (Grace Vydac, Hesperia, CA) was used. The flow rate was 0.8 mL/min, and a 60-min linear gradient of 30-60% acetonitrile in 0.1% trifluoroacetic acid (TFA) was employed. Purified histone H3 was digested with sequencing-grade modified Arg-C (Roche Applied Science, Indianapolis, IN) at a protein/enzyme ratio (w/w) of 20:1 in a 100-mM ammonium bicarbonate buffer at 37 ºC overnight. The digested peptides were further separated on the same HPLC system with a Zorbax SB-C18 capillary column (0.5×150 mm, 5 μm in particle size, Agilent Technologies), and a 60-min gradient of 2-60% acetonitrile in 0.6% acetic acid was used. The flow rate was 10 μL/min.

Mass Spectrometry

MALDI-MS/MS measurements were performed on a QSTAR XL quadrupole/time-of-flight instrument equipped with an o-MALDI ion source (Applied Biosystems, Foster City, CA). The laboratory collision energy applied for MS/MS varied from 50 to 75 eV depending on peptide sequences and modification levels, and the collision gas was nitrogen.

LC-ESI-MS/MS experiments were carried out by coupling directly the effluent from the Zorbax SB-C18 capillary column (0.5×150 mm, 5 μm in particle

size, Agilent Technologies) to an LTQ linear ion trap mass spectrometer (Thermo Electron Co., San Jose, CA). Isolated peptides were subjected to LC-MS/MS analysis, and we used a 60-min linear gradient of 2-60% acetonitrile in 0.6% acetic acid delivered by the Agilent 1100 capillary HPLC pump at a flow rate of 6 μL/min. MS/MS experiments were carried out in either the data-dependent scan mode or the pre-selected ion mode. Helium was employed as the collision gas, and the normalized collision energy was 30%. The width for precursor ion isolation was set at 2.5 (*m*/*z*) with an activation *Q* of 0.25 and an activation time of 30 ms. The spray voltage was 4.5 kV, and the temperature for the heated capillary was 275 ºC.

Results

MALDI-MS/MS suggests the unusual modification of Pro-16 in human histone H3 and the methylation of Arg-83 in yeast histone H3

To assess the PTMs of histone H3, we extracted the protein from cultured human (HL-60 and MCF-7) and yeast *S. cerevisiae* cells, digested it with Arg-C, and subjected the digestion mixtures to MALDI-MS and MS/MS analyses. The MALDI-MS data showed that the peptide with residues 73-83 in human histone H3 was mostly unmodified, though mono- and di-methylation could also be detected (Figure 6.1A). The peptide with residues 73-83 from yeast cells as well as the peptide with residues 9-17 in histone H3 isolated from human cells were mono-, di- and trimethylated (Figure 6.1B&C). On the other hand, the peptide housing residues

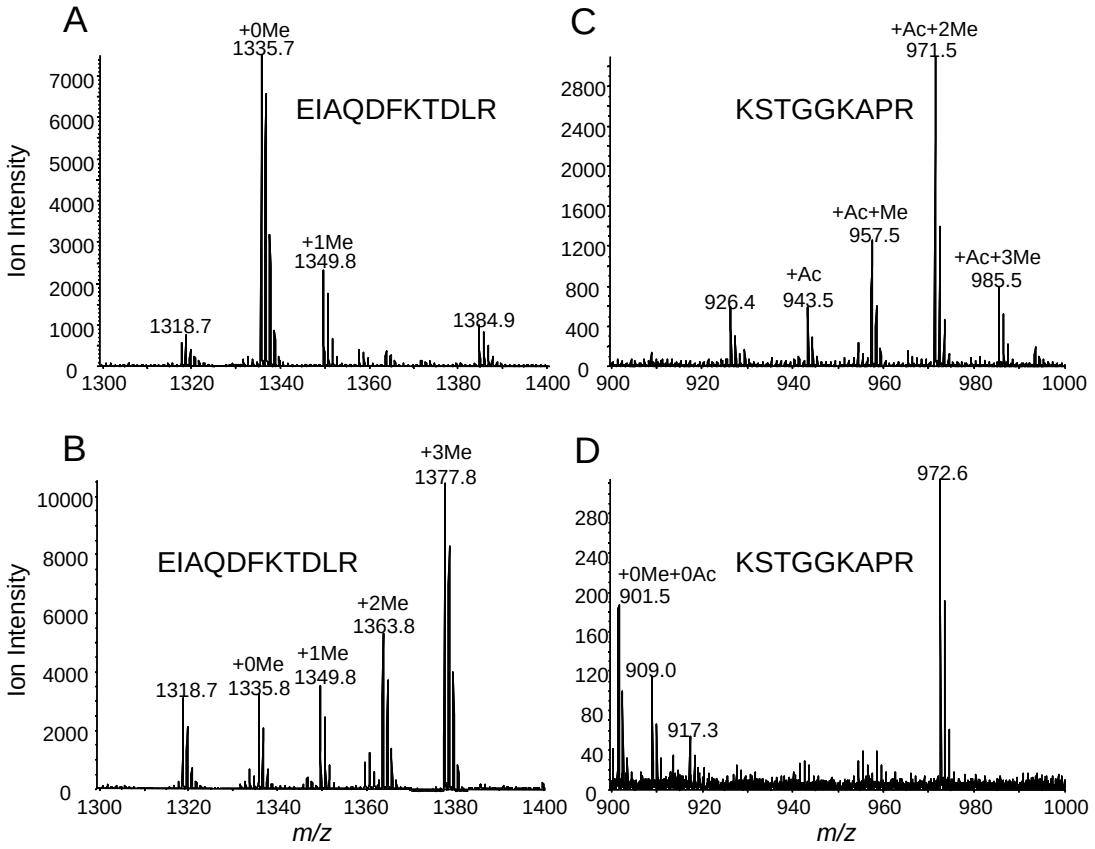


Figure 6.1 MALDI-MS of the Arg-C produced peptide with residues 73-83 in histone H3 isolated from MCF-7 human breast cancer cells (A) and yeast cells (B). Shown in (C) and (D) are the MALDI-MS of the Arg-C produced peptide fragment containing residues 9-17 in histone H3 extracted from MCF-7 cells and yeast cells, respectively.

9-17 is predominantly unmodified in H3 isolated from yeast cells (Figure 6.1D). In this context, it is worth noting that the ion of *m/z* 972.6 (Figure 6.1D) is attributed to the protonated ion of another histone H3 peptide, KQLASKAAR, which was supported by MS/MS analysis (spectrum not shown).

MALDI-MS/MS analysis of the trimethylated forms of the two peptides, however, suggests the modifications of Pro-16 and Arg-83 (Figures 6.2&3). In the product-ion spectrum of the $[M+H]^+$ ions of the trimethylated and monoacetylated peptide KSTGGKAPR (*m/z* 986.4, Figure 6.2), we observed the b_5+2Me, b_6+2Me+Ac, and b_7+2Me+Ac ions, supporting the dimethylation of Lys-9 and the acetylation of Lys-14. We also observed the neutral loss of an $HN(CH_3)_2$ from the b_6+2Me+Ac and b_7+2Me+Ac ions, lending further support for the presence of a dimethyllysine in the peptide. The presence of an acetylated lysine in this peptide is also supported by the observation of the diagnostic ion of *m/z* 126.1 (Figure 6.2) [18]. Additionally, the observation of the y_2+Me and y_3+Me ions and the finding of the b_7+Ac+2Me and b_8+Ac+3Me+H_2O ions appear to support that there is covalent modification of Pro-16 giving rise to a mass increase of 14 Da (Figure 6.2). Monomethylation can lead to a mass increase of 14 Da; however, the methylation of internal proline has not been reported, though methylation of N-terminal proline is known [19].

In the MS/MS of the trimethylated peptide EIAQDFKTDLR (*m/z* 1377.5, Figure 6.3), Lys-79 was found to be di- or trimethylated, which is supported by the mass difference between the y_4 and y_5 ions. A series of y_n+Me (n = 1-2) ions as well

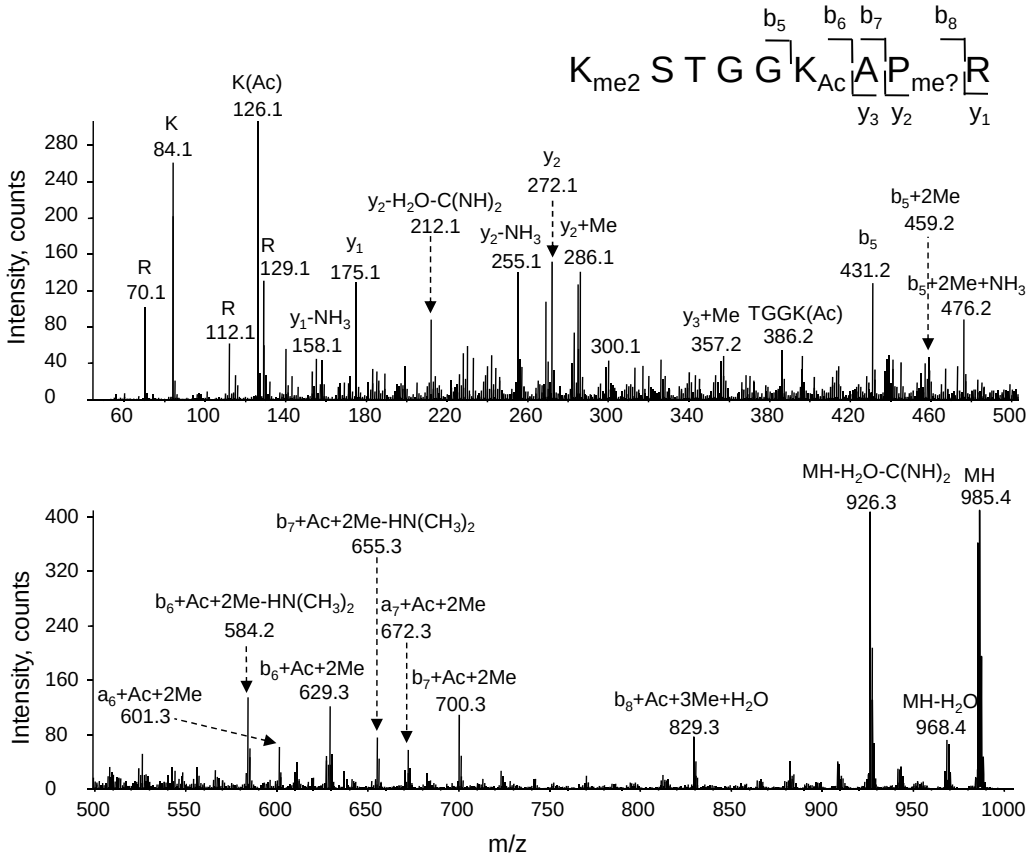


Figure 6.2 MALDI-MS/MS of the singly charged ion of the peptide with human H3 peptide with residues 9-17 that is monoacetylated and trimethylated (*m/z* 985.5).

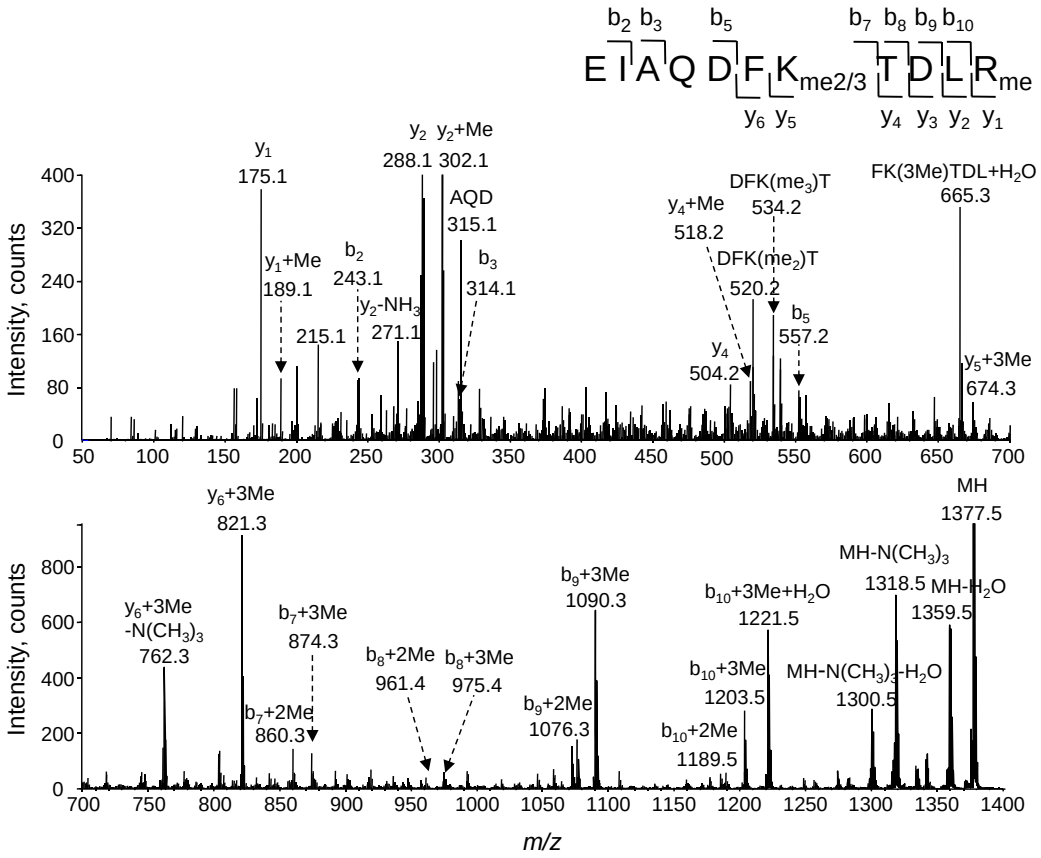


Figure 6.3 MALDI-MS/MS of the singly charged ion (*m/z* 1377.8) of the trimethylated peptide with residues 73-83 in histone H3 extracted from *S. cerevisiae* cells.

as b_m+2Me (m = 8-9) ions were observed, suggesting that Arg-83 might be partially monomethylated. It is worth noting that the assignments of fragment ions were consistent with exact mass measurements of fragment ions based on the MS/MS acquired on the QSTAR instrument (Table 6.1).

We also acquired the MALDI-MS/MS of these two peptides bearing lower level of methylation. Methylation of Pro-16 or Arg-83 could not be detected when the overall methylation level of the peptides dropped. In the MS/MS of the dimethylated peptide with residues 9-17 (Figure S6.1), the presence of the b_n+2Me (n = 3-8) ions supports the dimethylation of Lys-9; the absence of y_2+Me and y_3+Me ions, along with the presence of y_2 and y_3 ions, shows that Pro-16 was not modified (Figure S6.1, and Figure S6.2 depicts the MS/MS in the low-m/z range to better visualize the low-abundance fragments). Together, the above MALDI-MS/MS results suggest that Pro-16 and Arg-83 might be partially modified with a 14-Da mass increase only when these two peptides are trimethylated.

LC-ESI-MS/MS supports the absence of Pro-16 and Arg-83 modification

To further assess the nature of methylation of the above peptides, we subjected the same trimethylated H3 peptides to LC-ESI-MS/MS analyses. Surprisingly, the ESI-MS/MS data did not reveal the modification on Pro-16 or Arg-83, instead they only supported the existence of the trimethylated lysine in these two peptides (Figure 6.4).

Table 6.1 A summary of calculated and measured *m/z* of product ions for the trimethylated human H3 peptide with residues (9-17) and yeast H3 peptide with residues (73-83). The MS/MS was calibrated by using the calculated masses for the y_1 and precursor ions.

Product Ions	Calculated *m/z*	Measured *m/z*	Deviation (p.p.m.)
K(me_3)STGGK(ac)APR			
y_2	272.1717	272.1662	-20
y_2+Me	286.1874	286.1850	-8.3
y_2-NH_3	255.1452	255.1448	-1.5
y_2-H_2O-$C(NH)_2$	212.1395	212.1361	-16
y_3+Me	357.2245	357.2157	-24
b_5	431.2249	431.2207	-9.7
b_5+2Me	459.2562	459.2553	-1.9
b_6+Ac+2Me	629.3618	629.3782	26
b_6+Ac+2Me-$HN(CH_3)_2$	584.3039	584.3103	10
a_6+Ac+2Me	601.3668	601.3788	19
b_7+Ac+2Me	700.3989	700.4117	18
b_7+Ac+2Me-$HN(CH_3)_2$	655.3410	655.3570	24
a_7+Ac+2Me	672.4040	672.4117	11
b_8+Ac+3Me+H_2O	829.4778	829.4677	-12
TGGK(Ac)	386.2034	386.2022	-3.1
MH-$N(CH_3)_3$	926.5055	926.4909	-15
MH-NH_3	968.5524	968.5525	0.1
EIAQDFK(me_3)TDLR			

y_1+Me	189.1347	189.1360	-6.8
y_2	288.2030	288.2006	-8.3
y_2+Me	302.2187	302.2220	10
y_4	504.2776	504.2737	-7.7
y_4+Me	518.2933	518.2992	11
y_5	632.3726	632.3731	0.7
y_5+3Me	674.4196	674.4253	8.4
y_6+3Me	821.4880	821.4685	-23.7
y_6+3Me-N$(CH_3)_3$	762.4145	762.4182	4.8
b_2	243.1339	243.1285	-22
b_8+2Me	961.4990	961.4882	-11
b_8+3Me	975.5147	975.4958	-19
b_9+2Me	1076.5260	1076.5250	-0.9
b_9+3Me	1090.5417	1090.5233	-16
b_{10}+2Me	1189.6100	1189.6324	18
b_{10}+3Me	1203.6257	1203.6505	20
b_{10}+3Me+H_2O	1221.6363	1221.6593	18
AQD	315.1299	315.1267	-10
DFK(me_2)T	520.2766	520.2703	-12
DFK(me_3)T	534.2923	534.2894	-5.4
FK(me_3)TDL	665.3869	665.3765	-16
MH-N$(CH_3)_3$	1318.6638	1318.6801	12
MH-H_2O	1359.7267	1359.7397	9.5

Unlike the MALDI-MS/MS results, the ESI-MS/MS of the doubly charged H3 peptide with residues 9-17 showed the formation of y_1, y_2, y_3, and y_n+Ac (n = 4-8) ions, supporting the acetylation of Lys-14 and the lack of modification on Pro-16 (Figure 6.4A). Additionally, the observation of the neutral loss of a trimethylamine [$N(CH_3)_3$] from both the precursor ion and the b_2+3Me ion reveals the trimethylation of Lys-9 (Figure 6.4A).

ESI-MS/MS of the doubly charged ion of the trimethylated peptide with residues 73-83 of histone H3 isolated from yeast cells also showed the lack of methylation on Arg-83 (Figure 6.4B), instead the formation of abundant y_2, y_6+3Me, b_9+3Me, and b_{10}+3Me ions supports the tri-methylation of Lys-79 and the absence of methylation of Arg-83.

The ESI-MS/MS for the tri-methylated peptides supports unequivocally the lysine trimethylation as previously reported [20, 21] and the absence of modification on Pro-16 or Arg-83, which is in stark contrast with the MALDI-MS/MS results. However, LC-ESI-MS/MS and MALDI-MS/MS data for these peptides with lower modification levels are consistent, which both supported the methylation of lysine residues alone. The interesting discrepancy between LC-ESI-MS/MS and MALDI-MS/MS on the trimethylated peptides calls for further investigation on the nature of methylation in these peptides.

MS/MS analyses of synthetic trimethyllysine-containing peptide EIAQDFK(me_3)TDLR

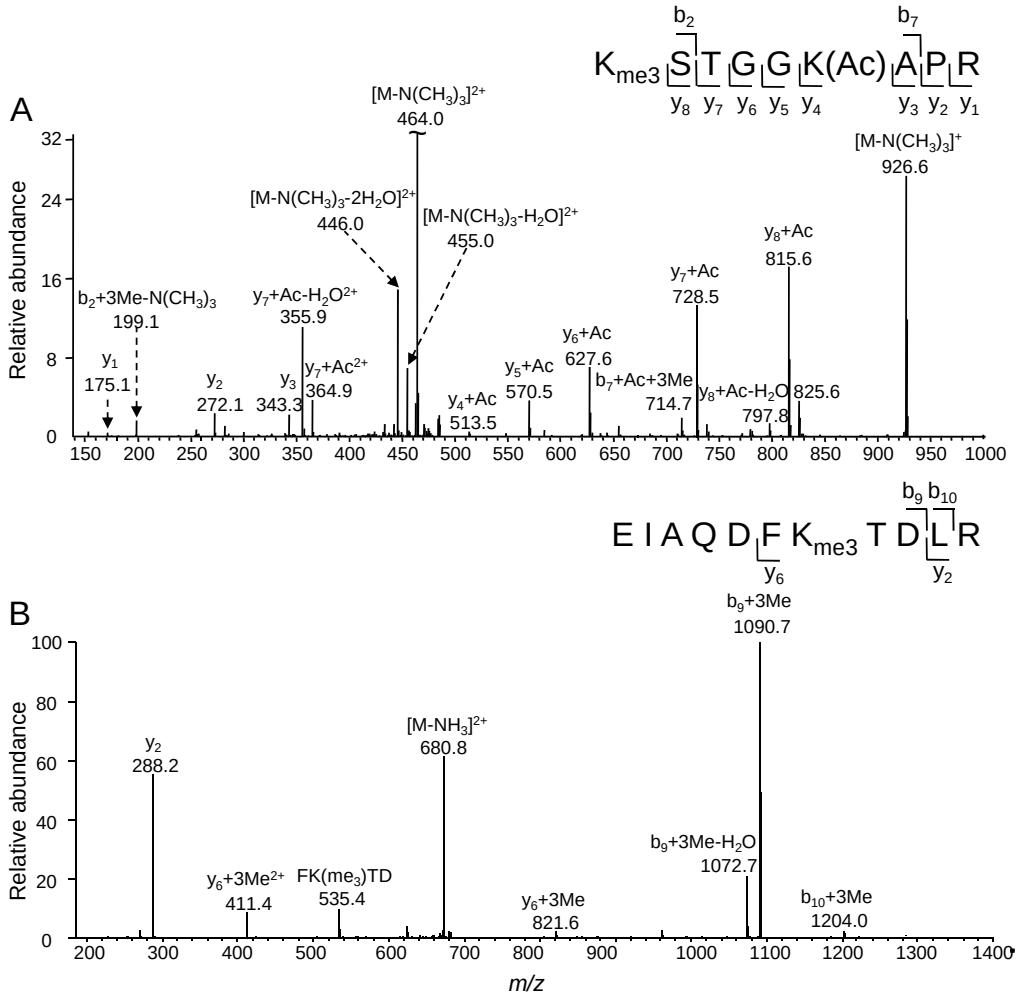


Figure 6.4 MS/MS of the ESI-produced doubly charged ions of tri-methylated peptides with residues 9-17 in histone H3 isolated from MCF-7 cells (A) and with residues 73-83 in histone H3 extracted from *S. cerevisiae* cells (B).

To gain insights into the above disagreement between MALDI- and ESI-MS/MS results, we obtained a synthetic trimethyllyine-containing peptide bearing the same sequence as the yeast H3 peptide with residues 73-83 and subjected it to MALDI- and ESI-MS/MS analysis. The ESI-MS/MS of this peptide supports the lysine trimethylation (spectrum not shown), which is in keeping with the ESI-MS/MS of the peptide of histone H3 isolated from yeast cells. Interestingly, the product-ion spectrum of the MALDI-produced singly charged ion of this trimethyllysine-containing peptide (Figure S6.3) is almost the same as that of the Arg-C produced peptide from histone H3 isolated from yeast cells (Figure 6.3). In this context, a series of y_n+Me (n = 1-2) ions as well as b_m+2Me (m = 7-9) ions were observed, though the C-terminal arginine in this synthetic peptide was not methylated. These results support unambiguously the lack of methylation of Arg-83 in histone H3 in yeast cells and suggest an unusual fragmentation of the MALDI-produced ions of trimethyllysine-containing peptides. In the latter respect, a methyl group may migrate from the side chain of trimethyllysine to the C-terminal side of the peptide during its fragmentation.

Charge state-specific migration of methyl group for trimethyllysine-carrying peptides

To assess whether this unusual fragmentation occurs generally for trimethyllysine-containing peptides, we further obtained two synthetic, alanine-rich

and trimethyllysine-containing peptides, AAAK(me_3)AAK and AAAK(me_3)AAR, and subjected them to MALDI- and ESI-MS/MS analyses.

ESI-MS/MS of the doubly charged ions of the two peptides confirmed unambiguously the trimethylation of the central lysine residue in the peptide, as manifested by the observation of b_n (n = 2-3), b_m+3Me (m = 4-6), y_x (x = 1-3), and y_5+3Me ions (Figure S6.4A&B). However, ions arising from methyl group migration, i.e., b_m+2Me (m=4-6) or y_x+Me (x=1-3) could not be found in the two spectra.

By contrast, fragment ions emanating from the methyl group migration could be readily observed in the MS/MS of the singly charged ions of the two peptides produced by either MALDI or ESI (Figure 6.5 shows the ESI-MS/MS results). For instance, we observed y_n+Me (n =1-3) and a_m+2Me/b_m+2Me (m = 4-6) ions in the MS/MS of the ESI-produced singly charged ion of AAAK(me_3)AAR (Figure 6.5A). While the above a_m+2Me/b_m+2Me ions could be observed in the corresponding spectrum of AAAK(me_3)AAK, the y_n+Me (n=1-3) ions are barely detectable (Figure 6.5B). The different basicities of lysine and arginine, which reside on the C-termini of these two synthetic peptides, may account for their different susceptibilities in the formation of y and y+Me ions. The MALDI-MS/MS of the singly charged ions of the two peptides gave similar results (Data not shown). Together, the above results demonstrated that the interesting methyl group migration can only be detected in the

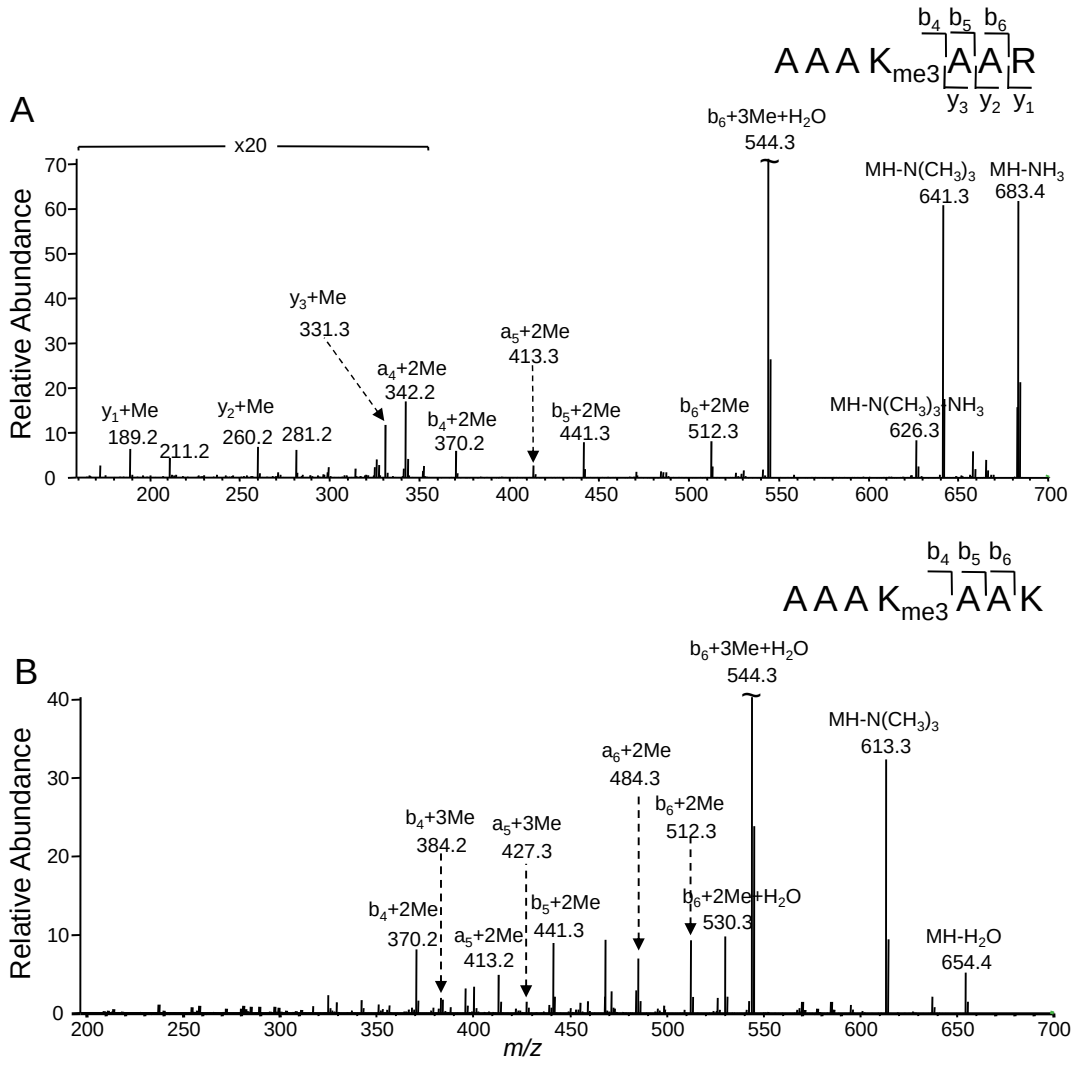


Figure 6.5 MS/MS of the ESI-produced singly charged ions of the synthetic, trimethyllysine-bearing peptides AAAK(me_3)AAR (*m/z* 700.5) (A) and AAAK(me_3)AAK (*m/z* 672.5) (B).

MS/MS of the singly-charged ions of trimethyllysine-containing peptides, regardless of whether the precursor ions are produced by MALDI or ESI.

To further examine the destination of the migrated methyl group, we acquired the MS^3 data, which record the product ions formed from the cleavages of the y_1+Me, y_2+Me, and y_3+Me ions observed in Figure 6.5A. As depicted in Figure 6.6A, the product-ion spectrum emanating from the fragmentation of the y_1+Me ion gives the neutral losses of NH_3 (*m/z* 172), methylamine (CH_3NH_2, *m/z* 158), and monomethylcarbodiimide ($HN=C=NCH_3$, *m/z* 133). The latter two are characteristic neutral losses for the protonated ions of peptides housing an N^G-monomethyl-L-arginine [22-24]. Indeed the MS^3 arising from the further cleavage of y_1+Me shows the formation of all fragment ions that can be found in the product-ion spectrum of the $[M+H]^+$ ion of standard N^G-monomethyl-L-arginine (Figure 6.6A and Figure S6.5). This result supports that the methyl group in the side chain of a trimethyllysine side chain can be migrated to the guanidinium group of arginine. Aside from the ions that can be found in the MS/MS of the standard N^G-monomethyl-L-arginine, the MS^3 reveals the formation of three unique fragment ions, i.e., the ions of *m/z* 157, 130, and 60. The former two ions are attributed to the neutral loss of CH_3OH and the combined neutral losses of NH_3 and carbodiimide, respectively, whereas the ion of *m/z* 60 is attributed to the protonated ion of the guanidinium moiety of arginine. The formation of these three ions is consistent with the view that a portion of the y_1+Me ion carries the methyl group on the C-terminal

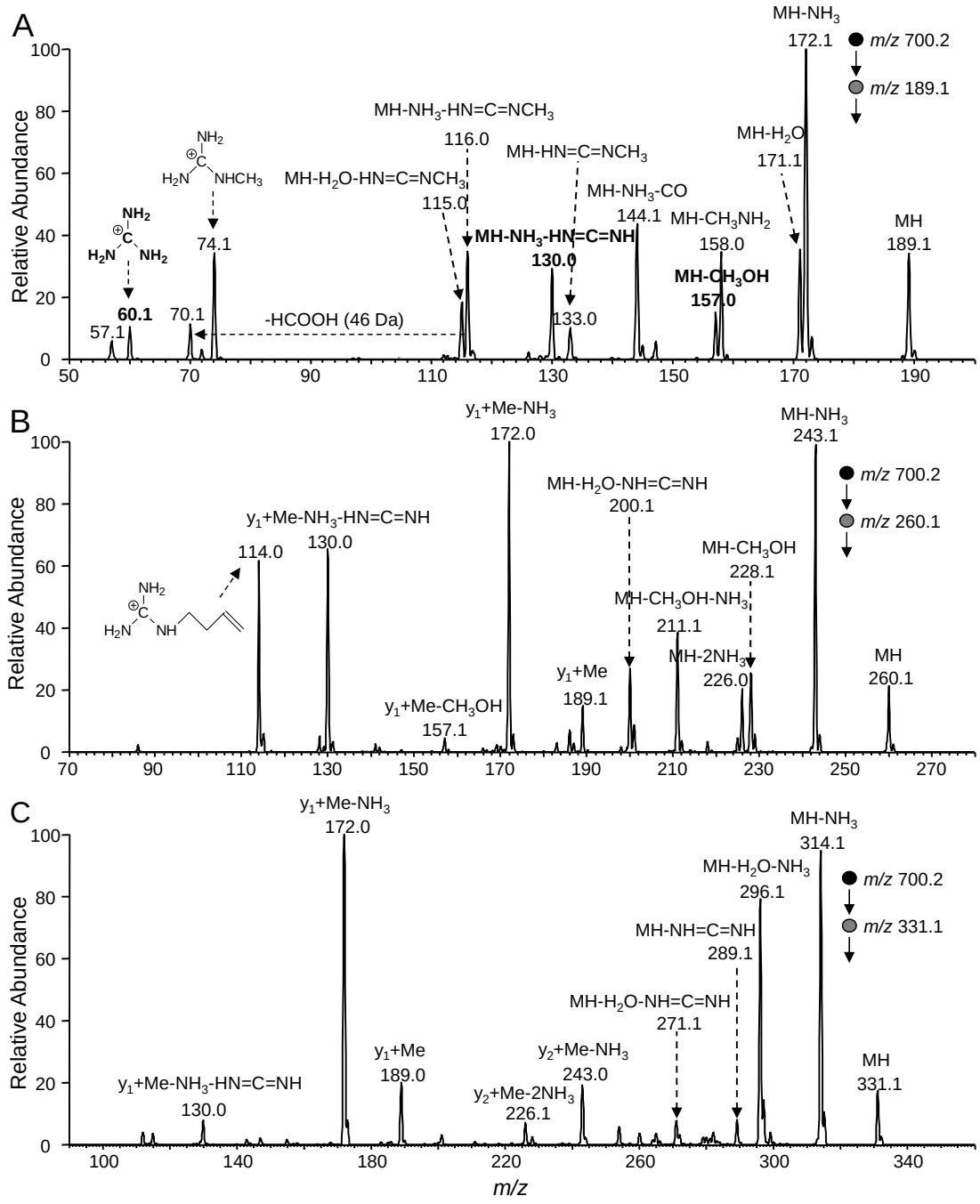


Figure 6.6 MS^3 monitoring the fragmentation of the y_1+Me (A), y_2+Me (B), and y_3+Me (C) ions observed in Figure 6.5A. In Figure 6.6A, the three fragment ions that were not observed in the MS/MS of the $[M+H]^+$ ion of N^G-monomethyl-L-arginine (Figure S6.5) are labeled in bold fonts.

carboxylic acid functionality. The relatively low abundances of these three fragment ions suggest that the methyl group in y_1+Me ion resides mainly on the guanidinium functionality of arginine.

The MS^3 of the y_2+Me and y_3+Me both revealed the formation of abundant y_1+Me (*m/z* 189) and y_1+Me-NH_3 (*m/z* 172) ions, whereas the unmethylated y_1 ion was not detectable (Figure 6.6B&C), strongly suggesting that, in these two fragment ions, the methyl group is located on the C-terminal arginine residue. It is worth noting that the above MS^3 results also lend further support of our assignments of the y_n+Me (n=1-3) ions observed in Figure 6.5A.

Discussion and Conclusions

During the fragmentation of protonated ions of peptides induced by CID or SID, a 'mobile proton' can be transferred to the amide linkage to induce the formation of the b and its complementary y ions [7-13]. This model is successful in rationalizing the formation of b and y ions and the preference of specific chain cleavage sites during the CID or SID of peptides through the 'charge directed' pathway. Here we showed that, during the collisional activation of the singly charged, trimethyllysine-harboring peptides, regardless of whether the ion is produced by MALDI or ESI, a methyl group can migrate from the side-chain of a trimethylysine to the guanidinium side chain or the carboxylic acid moiety of the C-terminal arginine.

This methyl group migration does not occur during the fragmentation of the ESI-produced doubly charged ions of trimethyllysine-containing peptides.

The inconsistency between MALDI-MS/MS and LC-ESI-MS/MS results on the tri-methylated peptides with residues 9-17 and 73-83 is attributed to the peptide charge difference. While the MALDI-produced ions of these peptides carry only one charge, the ESI-formed doubly charged ions of these two peptides adopt a protonated side chain or N-terminus other than the charge located on the side chain of the trimethylated lysine residue. These protons are mobile, thereby fulfilling the requirement for the 'charge-directed' pathway. Peptide bond cleavage arising from the mobile proton transfer is likely to be energetically more favorable than the methyl group migration. Thus, we did not observe fragment ions emanating from the methyl group transfer in the MS/MS of the ESI-produced doubly charged ions of trimethyllysine-containing peptides.

MALDI-MS/MS and LC-ESI-MS/MS for the peptides with lower methylation level are consistent because these peptides carry a 'mobile' proton to facilitate the fragmentation, as observed for the fragmentation of the ESI-produced doubly charged ions of the trimethyllysine-housing peptides where proton transfer takes place.

It is worth emphasizing that the methyl group migration is not essential for the fragmentation of tri-methylated lysine-containing peptides. The formation of y_n (n =1,2 in Figure 6.2, and n = 1-3 in Figures 6.3&S6.3) ions suggests that the cleavage of amide bonds can occur prior to the methyl group transfer process.

We observed an interesting methyl group transfer phenomenon during the fragmentation of singly charged ions of trimethyllysine-containing peptides; a methyl group on the side chain of trimethyllysine can be transferred to the C-terminal residue of the peptide prior to amide bond cleavage. The results from the present study call for special attention to the possibility of this type of methyl group migration while MS/MS of singly charged ions is employed to interrogate the methylation of proteins, including histones.

References:

1. Luger, K.; Mader, A. W.; Richmond, R. K.; Sargent, D. F.; Richmond, T. J., Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **1997,** 389, 251-60.

2. Strahl, B. D.; Allis, C. D., The language of covalent histone modifications. *Nature* **2000,** 403, 41-5.

3. Jenuwein, T.; Allis, C. D., Translating the histone code. *Science* **2001,** 293, 1074-80.

4. Klose, R. J.; Zhang, Y., Regulation of histone methylation by demethylimination and demethylation. *Nat. Rev. Mol. Cell. Biol.* **2007,** 8, 307-18.

5. Zhang, K.; Williams, K. E.; Huang, L.; Yau, P.; Siino, J. S.; Bradbury, E. M.; Jones, P. R.; Minch, M. J.; Burlingame, A. L., Histone acetylation and deacetylation: identification of acetylation and methylation sites of HeLa histone H4 by mass spectrometry. *Mol. Cell Proteomics* **2002,** 1, 500-8.

6. Cocklin, R. R.; Wang, M., Identification of methylation and acetylation sites on mouse histone H3 using matrix-assisted laser desorption/ionization time-of-flight and nanoelectrospray ionization tandem mass spectrometry. *J. Protein Chem.* **2003,** 22, 327-34.

7. Tsaprailis, G.; Somogyi, A.; Nikolaev, E. N.; Wysocki, V. H., Refining the model for selective cleavage at acidic residues in arginine-containing protonated peptides. *Int. J. Mass Spectrom.* **2000,** 196, 467-479.

8. Tsaprailis, G.; Nair, H.; Somogyi, A.; Wysocki, V. H.; Zhong, W. Q.; Futrell, J. H.; Summerfield, S. G.; Gaskell, S. J., Influence of secondary structure on the fragmentation of protonated peptides. *J. Am. Chem. Soc.* **1999,** 121, 5142-5154.

9. Nair, H.; Wysocki, V. H., Are peptides without basic residues protonated primarily at the amino terminus? *Int. J. Mass Spectrom.* **1998,** 174, 95-100.

10. McCormack, A. L.; Somogyi, A.; Dongre, A. R.; Wysocki, V. H., Fragmentation of protonated peptides: surface-induced dissociation in conjunction with a quantum mechanical approach. *Anal. Chem.* **1993,** 65, 2859-72.

11. Dongre, A. R.; Somogyi, A.; Wysocki, V. H., Surface-induced dissociation: an effective tool to probe structure, energetics and fragmentation mechanisms of protonated peptides. *J. Mass Spectrom.* **1996,** 31, 339-50.

12. de Maaijer-Gielbert, J.; Gu, C.; Somogyi, A.; Wysocki, V. H.; Kistemaker, P. G.; Weeding, T. L., Surface-induced dissociation of singly and multiply protonated polypropylenamine dendrimers. *J. Am. Soc. Mass Spectrom.* **1999,** 10, 414-22.

13. Gu, C.; Tsaprailis, G.; Breci, L.; Wysocki, V. H., Selective gas-phase cleavage at the peptide bond C-terminal to aspartic acid in fixed-charge derivatives of Asp-containing peptides. *Anal. Chem.* **2000,** 72, 5804-13.

14. Paizs, B.; Suhai, S., Fragmentation pathways of protonated peptides. *Mass Spectrom. Rev.* **2005,** 24, 508-548.

15. Zhang, K.; Tang, H.; Huang, L.; Blankenship, J. W.; Jones, P. R.; Xiang, F.; Yau, P. M.; Burlingame, A. L., 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, 259-69.

16. Edmondson, D. G.; Smith, M. M.; Roth, S. Y., Repression domain of the yeast global repressor Tup1 interacts directly with histones H3 and H4. *Genes Dev.* **1996,** 10, 1247-1259.

17. Braunstein, M.; Rose, A. B.; Holmes, S. G.; Allis, C. D.; Broach, J. R., Transcriptional silencing in yeast is associated with reduced nucleosome acetylation. *Genes Dev.* **1993,** 7, 592-604.

18. Margueron, R.; Trojer, P.; Reinberg, D., The key to development: interpreting the histone code? *Curr. Opin. Genet. Dev.* **2005,** 15, 163-76.

19. Stock, A.; Clarke, S.; Clarke, C.; Stock, J., N-terminal methylation of proteins: structure, function and specificity. *FEBS Lett.* **1987,** 220, 8-14.

20. Zhang, K.; Sridhar, V. V.; Zhu, J.; Kapoor, A.; Zhu, J. K., Distinctive core histone post-translational modification patterns in *Arabidopsis thaliana*. *PLoS ONE* **2007,** 2, e1210.

21. Miao, F.; Natarajan, R., Mapping global histone methylation patterns in the coding regions of human genes. *Mol. Cell Biol.* **2005,** 25, 4650-61.

22. Zou, Y.; Wang, Y., Tandem mass spectrometry for the examination of the posttranslational modifications of high-mobility group A1 proteins: symmetric and

asymmetric dimethylation of Arg25 in HMGA1a protein. *Biochemistry* **2005,** 44, 6293-301.

23. Thorne, G. C.; Gaskell, S. J., Elucidation of some fragmentations of small peptides using sequential mass spectrometry on a hybrid instrument. *Rapid Commun. Mass Spectrom.* **1989,** 3, 217-21.

24. Zou, Y.; Webb, K.; Perna, A. D.; Zhang, Q.; Clarke, S.; Wang, Y., A mass spectrometric study on the in vitro methylation of HMGA1a and HMGA1b proteins by PRMTs: Methylation specificity, the effect of binding to AT-rich duplex DNA, and the effect of C-terminal phosphorylation. *Biochemistry* **2007,** 46, 7896-906.

Supporting Information for Chapter 6

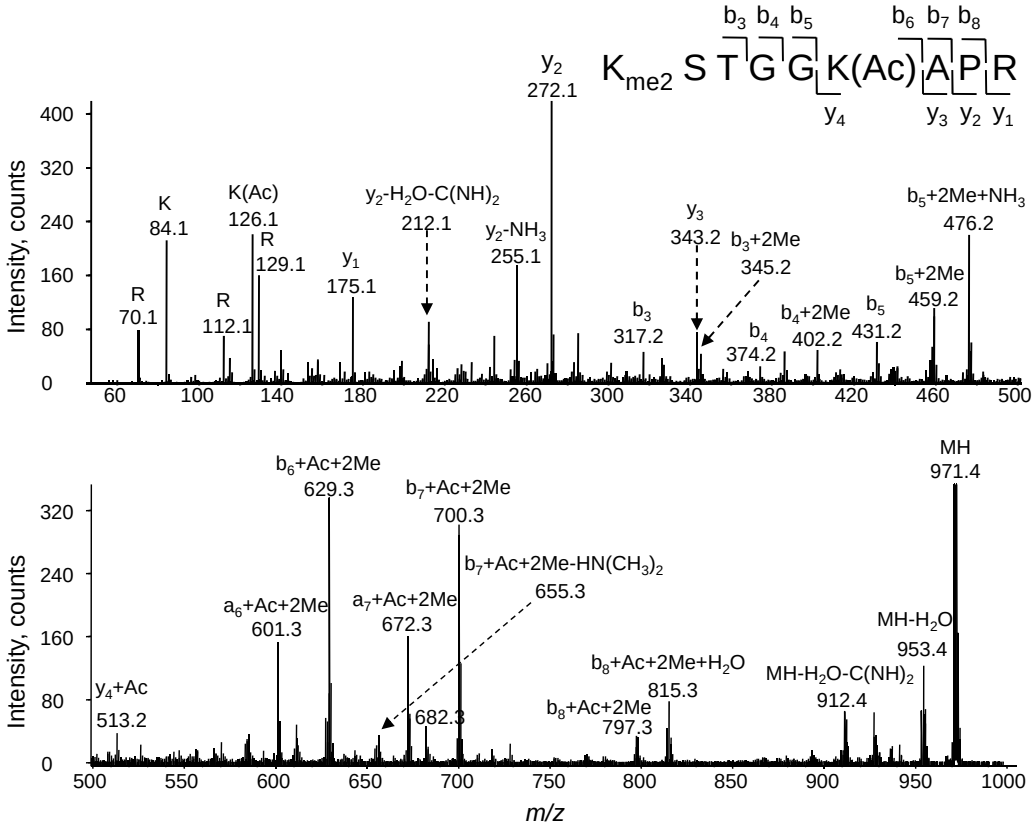


Figure S6.1 MALDI-MS/MS of singly charged ion of the peptide with residues 9-17 in histone H3 isolated from MCF-7 human cells that is both monoacetylated and dimethylated (*m/z* 971.5).

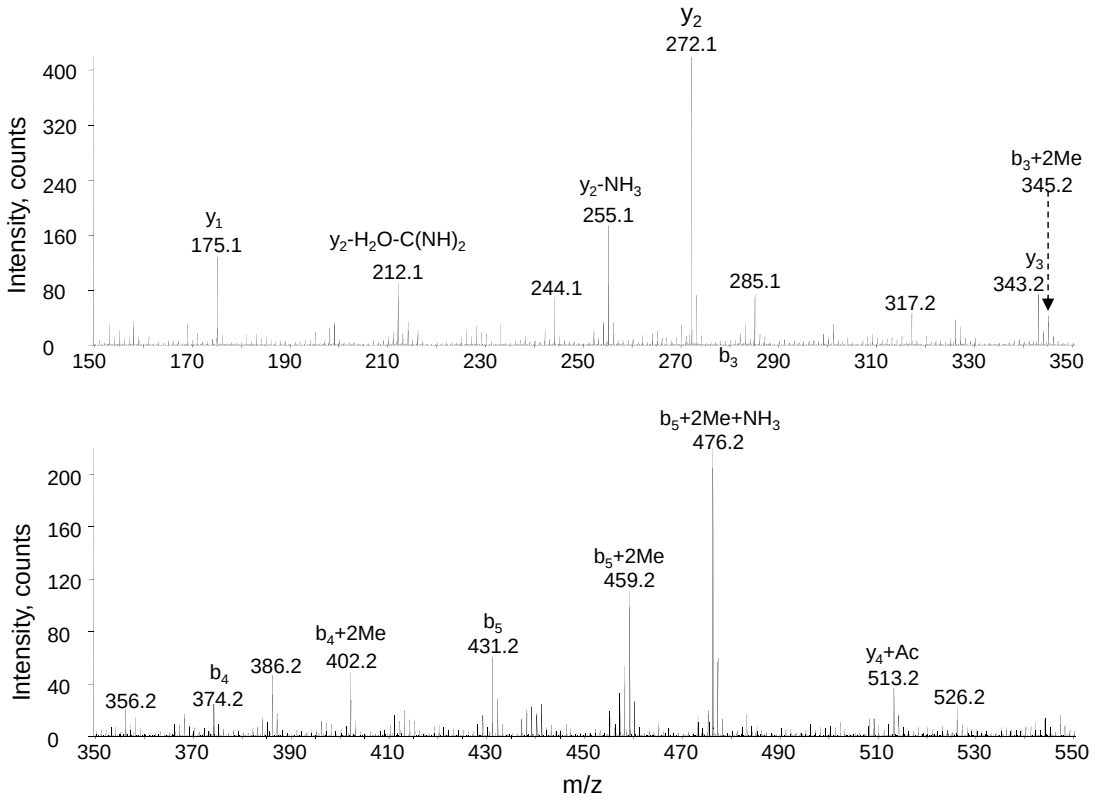


Figure S6.2 A portion of the MS/MS shown in Figure S6.1 is enlarged to better visualize the absence of y_1+Me, y_2+Me, and y_3+Me ions.

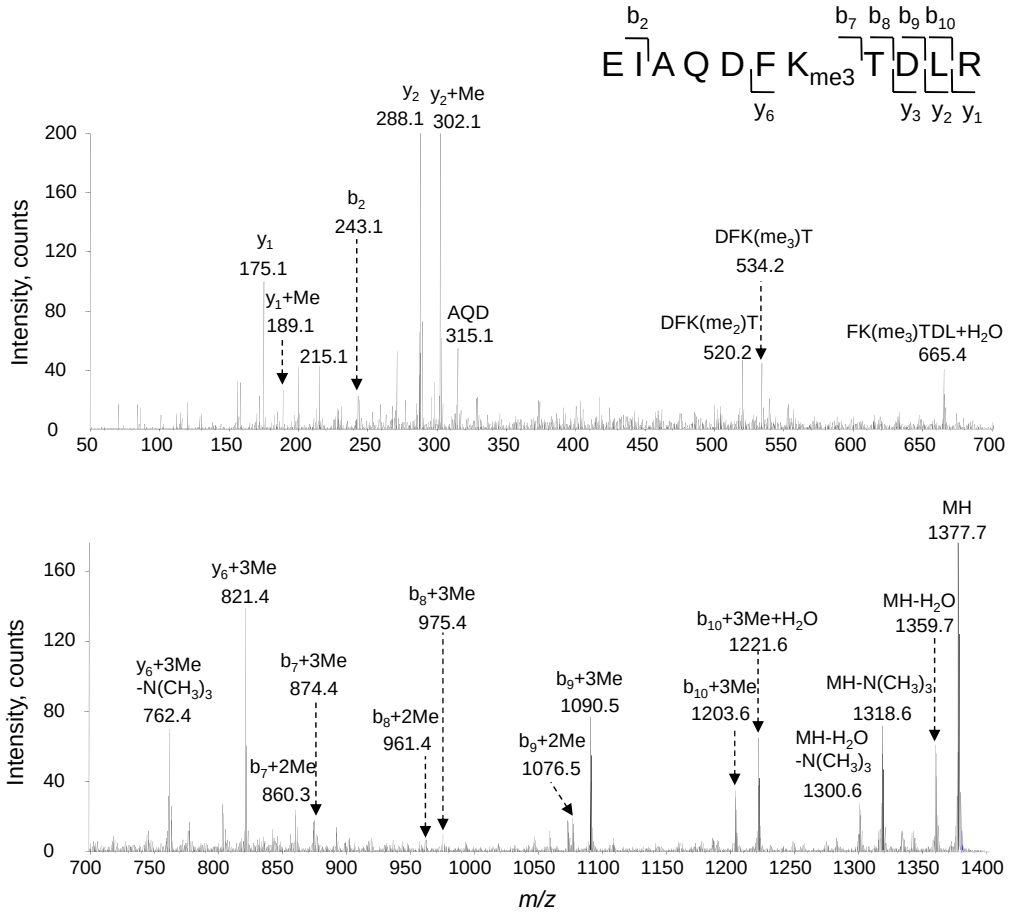


Figure S6.3 MALDI-MS/MS of the singly charged ion (*m/z* 1377.8) of the synthetic, trimethyllysine-carrying peptide EIAQDFK(me$_3$)TDLR.

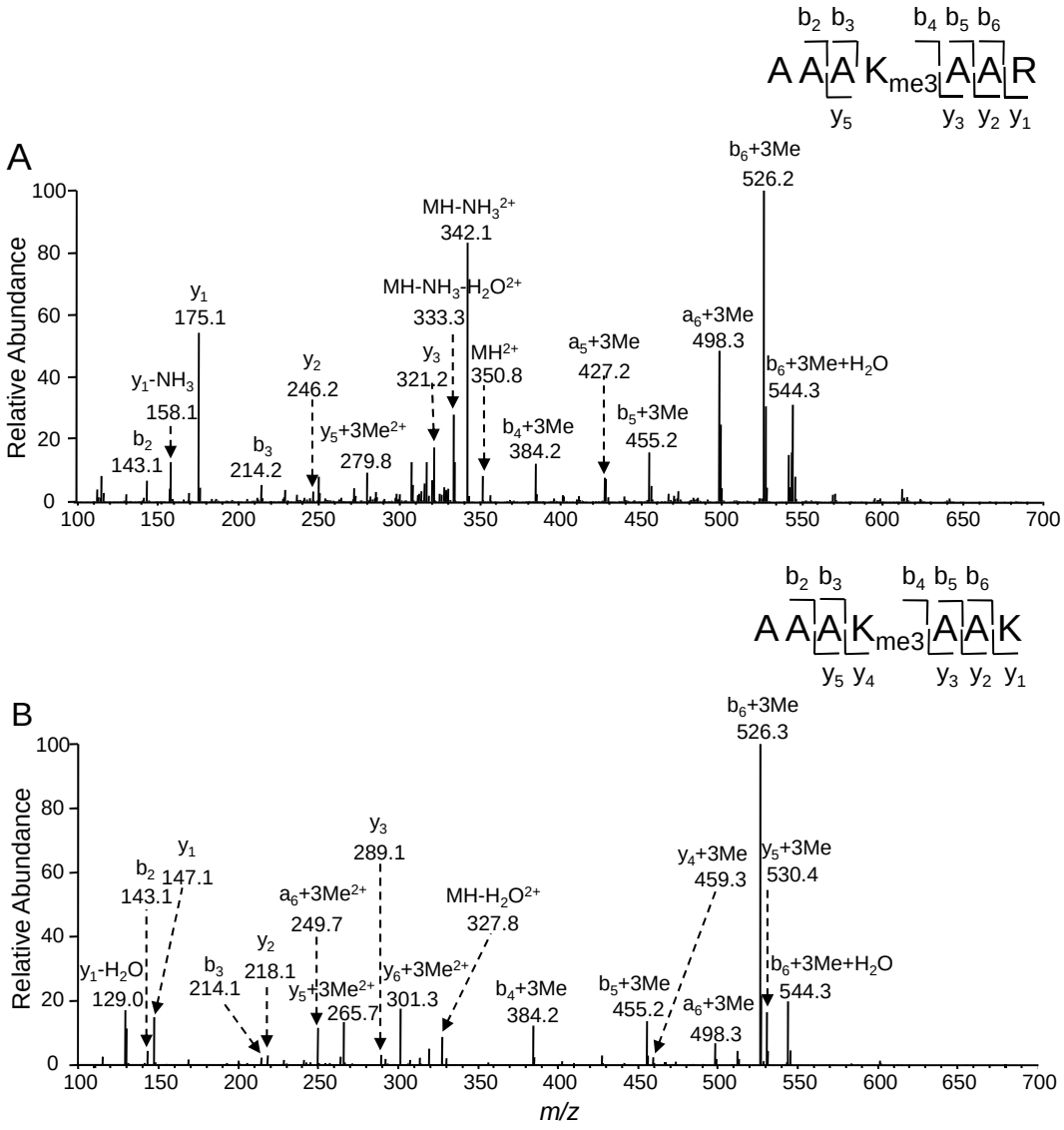


Figure S6.4 MS/MS of the ESI-produced doubly charged ions of the synthetic, trimethyllysine-containing peptides AAAK(me_3)AAR (*m/z* 350.7) (A) and AAAK(me_3)AAK (*m/z* 336.7) (B).

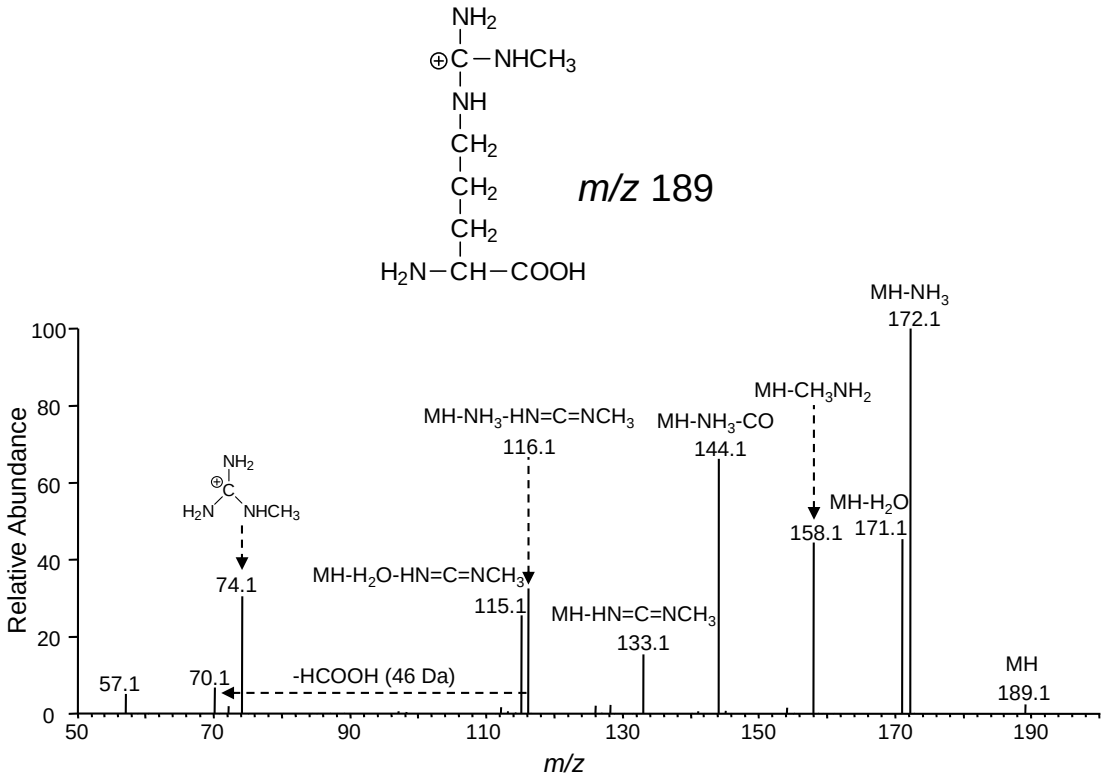


Figure S6.5 Product-ion spectrum of the ESI-produced $[M+H]^+$ ion of N^G-monomethyl-L-arginine. The structure of the monomethylarginine is also shown.

CHAPTER 7

Summary and Conclusions

In this dissertation, we applied mass spectrometry, together with modern sample preparation and separation techniques, to study quantitative proteomics and histone post-translation modifications. We explored the perturbation of protein expression for the discovery of novel molecular mechanisms in human leukemia upon anti-cancer drug treatments, comprehensively mapped the core histone PTMs of two important model eukaryotic organisms, and discovered a novel methyl group migration during the fragmentation of singly charged trimethyllysine-containing peptides.

In Chapters 2 and 3, we report the use of mass spectrometry together with stable isotope labeling by amino acids in cell culture (SILAC) for the comparative study of protein expression in HL-60 and K562 cells that were untreated or treated with a clinically relevant concentration of arsenite or imatinib. Arsenic is ubiquitously present in the environment; it is a known human carcinogen and paradoxically it is also a successful drug for the clinical remission of acute promyelocytic leukemia (APL). Imatinib mesylate, currently marketed by Novartis as Gleevec in the U.S., has emerged as the leading compound to treat the chronic phase of chronic myeloid leukemia (CML), through its inhibition of Bcr-Abl tyrosine kinases, and other cancers. However, resistance to imatinib develops frequently, particularly in late-stage disease.

The cellular responses have been investigated for years; however, the precise mechanisms underlying their cytotoxicity and therapeutic activity remain unclear.

Our results revealed that, among more than 1000 quantified proteins, 56 and 73 proteins including many important enzymes had significantly altered levels of expression by arsenite and imatinib treatment, respectively. During arsenite treatment, drug-induced growth inhibition of HL-60 cells has been confirmed to be rescued by treatment with palmitate, the final product of fatty acid synthase, supporting that arsenite exerts its cytotoxic effect, in part, via suppressing the expression of fatty acid synthase and inhibiting the endogenous production of fatty acid. In the imatinib treatment project, we found, by assessing alteration in the acetylation level in histone H4 upon imatinib treatment, that the imatinib-induced hemoglobinization and erythroid differentiation in K562 cells are associated with global histone H4 hyperacetylation. Overall, these results provided potential biomarkers for monitoring the therapeutic intervention of human leukemia and offered important new knowledge for gaining insight into the molecular mechanisms of action of anti-cancer drugs.

In Chapters 4 and 5, we obtained relatively comprehensive mappings of core histone post-translational modifications in two important model eukaryotic organisms, *Neurospora crassa* and *Schizosaccharomyces pombe*. We used several mass spectrometric techniques, coupled with HPLC separation and multiple protease digestion, to identify the methylation, acetylation and phosphorylation sites in core histones. Electron transfer dissociation (ETD) was employed to fragment the heavily

modified long N-terminal peptides. Moreover, accurate mass measurement of fragment ions allowed for unambiguous differentiation of acetylation from tri-methylation. Many modification sites conserved in other organisms were identified in these two organisms. In addition, some unique modification sites were found for the first time in core histones. Our analysis provides potentially comprehensive pictures of core histones PTMs, which serves as foundation for future studies on the function of histone PTMs in these two model organisms.

In Chapter 6, we observed an unusual discrepancy between MALDI-MS/MS and ESI-MS/MS on the methylation of trimethyllysine-containing peptides with residues 9–17 from human histone H3 and residues 73–83 from yeast histone H3. It turned out that the discrepancy could be attributed to an unusual methyl group migration from the side chain of trimethyllysine to the C-terminal arginine residue during peptide fragmentation, and this methyl group transfer only occurred for singly charged ions, but not for doubly charged ions. The methyl group transfer argument received its support from the results on the studies of the fragmentation of the ESI- or MALDI-produced singly charged ions of several synthetic trimethyllysine-bearing peptides. The results presented in this study highlighted that caution should be exerted while MS/MS of singly charged ions is employed to interrogate the PTMs of trimethyllysine-containing peptides.