CHARACTERIZATION OF HISTONES AND THEIR POSTTRANSLATIONAL MODIFICATIONS BY MASS SPECTROMETRY

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By

Lanhao Yang, M.S.

The Ohio State University
2009

Dissertation Committee:

Professor Michael A. Freitas, Advisor
Professor Heather C. Allen
Professor Ross E. Dalbey

Approved by

Advisor
Graduate Program in Chemistry
ABSTRACT

In this dissertation, mass spectrometry-based approaches were developed to characterize histones and their posttranslational modifications (PTMs). Histones are the primary protein component of chromatin. Histone tails are subject to a variety of PTMs including acetylation, methylation, and phosphorylation. These PTMs play critical roles in the regulation of chromatin remodeling and gene transcription. Therefore there is a need for better approaches to characterize histones and their PTMs. Yeast are chosen as the model for my studies due to many of their beneficial properties.

Methods for yeast histone preparation, separation, and LC-MS were developed and described in Chapter 2. Through extensive washing of nuclei prior to lysis and acid extraction, yeast histones were extracted from yeast whole cells with a high yield and purity. The isolated yeast histones were then separated by use of 16.5% Tris-tricine-SDS gels and characterized by LC-MS with trifluoroacetic acid (TFA) as the ion-pairing agent with minimal TFA effects. By using the methods developed in Chapter 2, histone acetylation and methylation patterns in yeast were studied (Chapter 3). We found that acetylation and trimethylation have different effects on retention time in reversed-phase liquid chromatography. The former shifts retention time, while the latter does not. By using retention time, acetylation and trimethylation were unambiguously distinguished.
even on a low-resolution mass spectrometer. In Chapter 4, we examined five JmjC domain-containing proteins (Ecm5, Gis1, Rph1, Jhd1, and Jhd2) in yeast for demethylase activities. Lysine methylation used to be considered irreversible. By using mass spectrometry-based approaches, we found that these proteins (except for Ecm5) displayed demethylase activity. In particular, Rph1 exhibited specificity toward H3K36 di- and trimethylation and Jhd2 had specificity toward H3K4 trimethylation. Cross-talk between methylation and acetylation was assessed in Chapter 5 by altering the expression of the JmjC proteins. Our results show that H3 acetylation changes significantly in all strains expect for ecm5Δ. LC-MS/MS results indicate that Gis1 and Jhd1 facilitate K56Ac, while Rph1 decreases the level of K56Ac. Taking into consideration that the JmjC proteins are histone demethylases, this study implies that H3 methylation is correlated with acetylation on H2B1, H2B2, and H3.
Dedicated to my parents
ACKNOWLEDGMENTS

When I look back at my life so far, I come to realize that God has blessed me all the way here in every aspect of my life. I could not praise Him enough for all of the blessings He has poured out upon me. He blessed me with my current advisor, Dr. Michael A. Freitas, my whole-hearted loving family, and He also brought helpful and sincere labmates and friends around me.

I am grateful for Dr. Freitas. Under his supervision, I grew up not only in science but in character. His passions for science and his optimistic and graceful character have been a great influence on my life.

Behind the scenes, my parents and brother are always with me, supporting me, encouraging me, and cheering for me. It is they who keep reminding me of my role as an international student. Without their love, I could not have come this far.

I would like to thank Dr. Pang-Hung Hsu, Dr. Xiaodan Su, Dr. Chen Ren, and Dr. Mitchell Meade for their helpful discussions and patient teaching, and Dr. Hua Xu for generous encouragement, pleasant collaboration, and permission for using MassMatrix. I also want to thank all the other lab group members for their friendship and discussions: John P. Shapiro, Liwen Wang, Bei Zhao, Jia You, Kelly DiRienzo, Justin Staubli, Xiaoyan Guan, Fatemeh Talebian, and Fatima Bube.
Many thanks to all of my collaborators: Dr. Ming-Daw Tsai, Dr. Guido Marcucci, Dr. Shunjun Liu, Dr. Mark Parthun, Dr. Esther M. M. Bulloch, Dr. Shengjiang Tu, Dr. Richard Fishel, Dr. Naduparambil K. Jacob, Dr. Chung-Lin Liao, and Amy Knapp for providing cell lines, generous use of their equipments, and helpful discussions and insights.

Thanks are also given to Dr. Kari Green-Church, Dr. Liwen Zhang, and Nan Kleinholz for helping with LC-MS, LC-MS/MS, and MALDI-TOF MS experiments. Thanks are also given to Frances Ziegler, Jennifer Hambach, Judith Brown, and Dr. David J. Hart for their faithful services.

Finally, I would like to thank all of the brothers and sisters in All Nations Christian Fellowship. It is they who add flavor, taste, and comfort to my life, especially in difficult times. Special thanks to Dr. Heather C. Allen, Dr. Ross E. Dalbey, and Dr. Kirk A. Mykytyn for their assistance and efforts for my graduation.

The study was funded by The Ohio State University, Dreyfus Foundation, the National Institutes of Health (CA110496, CA107106, and CA101956).
VITA

August 23, 1977 ............................................. Born – Xingyang, Henan, P. R. China

July, 2000 ....................................................... B. S. Henan Normal University
Department of Chemistry,
Xinxiang, P. R. China

July, 2003 ....................................................... M. S. Wuhan University
Department of Chemistry
Wuhan, P. R. China

July, 2003-July, 2004 ..................................... National Research and Development Center
of 863 Projects
Shenzhen, P. R. China

July, 2004 – Present ....................................... Graduate Teaching and Research Assistant
Department of Chemistry
The Ohio State University

PUBLICATIONS

Research Publications


FIELDS OF STUDY

Major Field: Chemistry
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract ............................................................................................................................... ii</td>
</tr>
<tr>
<td>Dedication .......................................................................................................................... iv</td>
</tr>
<tr>
<td>Acknowledgment ................................................................................................................ v</td>
</tr>
<tr>
<td>Vita .................................................................................................................................... vii</td>
</tr>
<tr>
<td>List of Tables ................................................................................................................... xiii</td>
</tr>
<tr>
<td>List of Figures .................................................................................................................. xiv</td>
</tr>
<tr>
<td>List of Abbreviation ......................................................................................................... xix</td>
</tr>
</tbody>
</table>

Chapters:

1. Introduction ....................................................................................................................1
   1.1 Histones and Chromatin Structure ...........................................................................1
      1.1.1 Background ........................................................................................................ 1
      1.1.2 Histone Posttranslational Modifications ......................................................... 3
      1.1.3 Chromatin Remodeling Enzymes ..................................................................... 5
   1.2 Strategies to Investigate Histone Posttranslational Modifications .......................9
      1.2.1 Traditional Approaches .................................................................................... 9
      1.2.2 Mass Spectrometry-Based Approaches ............................................................. 11
         1.2.2.1 Ionization Sources of Mass Spectrometers ................................................ 13
         1.2.2.2 Liquid Chromatography-Mass Spectrometry (LC-MS) ............................. 16
         1.2.2.3 Tandem Mass Spectrometry ..................................................................... 17
         1.2.2.4 Peptide Mass Fingerprinting (PMF) ............................................................ 20
         1.2.2.5 Acetic Acid-Urea Polyacrylamide Gel Electrophoresis (AU-PAGE) and Acetic Acid Triton-Polyacrylamide Gel Electrophoresis (AUT-AGE) ......................................................... 20
      1.2.2.6 Quantitative Analysis of Histones and Their Posttranslational Modifications ................................................................................................................................. 23
   1.3 Objective .................................................................................................................... 27
2. Characterization of Yeast Histones and Their Posttranslational Modifications ........31
   2.1 Introduction ...........................................................................................................31
   2.2 Experimental .........................................................................................................34
      2.2.1 Bovine Thymus Histone Extraction ..................................................................34
   2.2.2 Yeast Histone Extraction ..................................................................................35
      2.2.3 Liquid Chromatography-Mass Spectrometry ................................................35
      2.2.4 Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis (SDS-PAGE) Separation of Histones ..................................................37
      2.2.5 Liquid Chromatography-Tandem Mass Spectrometry ..................................38
   2.3 Results and Discussion .........................................................................................38
      2.3.1 Yeast Histone Extraction .................................................................................38
      2.3.2 Yeast Histone Variants ...................................................................................40
      2.3.3 Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE) Separation of Yeast Histones ........................................45
      2.3.4 Identification of Yeast Histone Separation Pattern from Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis (SDS-PAGE)..................................................................................................................48
      2.3.5 Optimization of Liquid Chromatography-Mass Spectrometry Histone Profiling..................................................................................................................52
      2.3.6 Optimization of Liquid Chromatography-Mass Spectrometry for Separation of Yeast Histone Isoforms .............................................................57
   2.4 Conclusions ...........................................................................................................63

3. Unambiguous Determination of Isobaric Modifications by LC-MS and High-Mass Accuracy.............................................................................................................65
   3.1 Introduction ...........................................................................................................65
   3.2 Experimental .........................................................................................................68
      3.2.1 Yeast Histone Extraction ................................................................................68
      3.2.2 Yeast Histone H3 In-Gel Digestion ..................................................................69
      3.2.3 Nano-Liquid Chromatography-Tandem Mass Spectrometry (nano-LC-MS/MS).............................................................................................................69
   3.3 Results and Discussion .........................................................................................70
      3.3.1 Posttranslational Modification Patterns of Yeast Histone H3 .........................70
3.3.2  Retention Time Shift on an FT-ICR Mass Spectrometer......................79
3.3.3  Unambiguous Determination of Acetylation and Trimethylation
       at Low-Mass Accuracy ...........................................................................90
3.4  Conclusions..........................................................................................91

4.  Identification of Histone Demethylases in *Saccharomyces cerevisiae*........96
  4.1  Introduction.........................................................................................96
  4.2  Experimental.....................................................................................100
    4.2.1  Yeast Strains and Cell Culture.................................................100
    4.2.2  Histone Extraction and Mass Spectrometry Analysis................100
    4.2.3  Assays of *In Vitro* Activities................................................101
    4.2.4  Growth Phenotype Assays ......................................................101
  4.3  Results..............................................................................................102
    4.3.1  Global Changes of Histone H3 Modifications in JmjC Knockout
           Strains .........................................................................................102
    4.3.2  High-Resolution Mass Spectrometry Analysis of *S. cerevisiae*
           Histone Methylation Patterns......................................................104
    4.3.3  Methylation Changes at Specific Sites in the Knockout Strains......107
    4.3.4  Methylation Changes in *RPH1, JHD1, and JHD2*
           Overexpression Strains ...............................................................112
    4.3.5  *In Vitro* Enzyme Assays of Rph1 Demethylase Activity..............117
    4.3.6  Overexpression of *RPH1* Causes a Growth Defect Upon UV
           Irradiation....................................................................................121
    4.3.7  *RPH1* and *JHD1* Overexpression Strains are Slightly Resistant to
           6-Azauracil....................................................................................122
  4.4  Discussion..........................................................................................124
    4.4.1  JmjC Domain Encodes Histone Demethylase Activities...............124
    4.4.2  Rph1 is an H3K36Me3 Demethylase Involved in DNA Damage
           Response and Transcription Elongation .......................................125
    4.4.3  Jhd2 is an H3K4 Demethylase.....................................................127
    4.4.4  Possible H3K79 Demethylases....................................................128
    4.4.5  Yeast JmjC Proteins in Transcription .........................................128
  4.5  Conclusions........................................................................................128
5. Label-Free Mass Spectrometry Determination of Histone Cross-Talk in Saccharomyces cerevisiae

5.1 Introduction

5.2 Experimental
  5.2.1 Yeast Cell Culture and Media
  5.2.2 Histone Preparation
  5.2.3 Liquid Chromatography-Mass Spectrometry (LC-MS) Analysis
  5.2.4 Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS) Analysis

5.3 Results and Discussion
  5.3.1 Global Changes of Histone H3 Acetylation
  5.3.2 Enzyme Digestion of Histones
  5.3.3 Quantification of H3K56Ac by Spectral Counting
  5.3.4 Quantification of H3K56Ac by Peak Area Abundance
  5.3.5 Global Changes of Histone H4 Acetylation
  5.3.6 Global Changes of Histone H2B Acetylation
  5.3.7 Global Changes of Histone H2A Acetylation

5.4 Conclusions

6. Summary

References

Appendices

Appendix A Supplementary Information for Chapter 2
Appendix B Supplementary Information for Chapter 4
Appendix C Protocols
  C.1 In-Gel Digestion
  C.2 Preparation for MALDI-TOF MS Experiment
  C.3 AU-PAGE Separation of Histones
  C.4 Histone Separation by Use of 16.5% Tris-Tricine-SDS Gels
C.5 Human Histone Extraction.................................................................225
C.6 Cell Growth of \textit{S. cerevisiae} ........................................................227
C.7 Yeast Histone Extraction .................................................................230
C.8 Histone Extraction from Bovine Calf Thymus ...............................233
C.9 Histone Extraction from \textit{Drosophila} .............................................235
C.10 Cell Culture and Histone Extraction from \textit{Tetrahymena} .............237
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Mass shift and occurring residues of PTMs on histones..................</td>
</tr>
<tr>
<td>1.2</td>
<td>HDAC inhibitors in clinical trials..........................................</td>
</tr>
<tr>
<td>2.1</td>
<td>Reversed-phase gradient used for LC-MS (LCT)................................</td>
</tr>
<tr>
<td>2.2</td>
<td>Reversed-phase gradient used for LC-MS (Q-TOF).............................</td>
</tr>
<tr>
<td>2.3</td>
<td>Yeast histone variants ..................................................................</td>
</tr>
<tr>
<td>2.4</td>
<td>Sequence variations in yeast histone H2B1 and H2B2........................</td>
</tr>
<tr>
<td>2.5</td>
<td>Yeast histone acetylation sites ..............................................</td>
</tr>
<tr>
<td>2.6</td>
<td>LC-MS/MS identification of yeast histone separation pattern on a SDS-PAGE gel</td>
</tr>
</tbody>
</table>
**LIST OF FIGURES**

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Histone arrangement in nucleosomes</td>
<td>2</td>
</tr>
<tr>
<td>1.2</td>
<td>The histone acetylation-deacetylation cycle</td>
<td>7</td>
</tr>
<tr>
<td>1.3</td>
<td>A protein phosphorylation-dephosphorylation cycle</td>
<td>7</td>
</tr>
<tr>
<td>1.4</td>
<td>Mechanisms of lysine demethylation by LSD1 and JmjC proteins</td>
<td>9</td>
</tr>
<tr>
<td>1.5</td>
<td>Typical proteomic workflow for histone characterization</td>
<td>12</td>
</tr>
<tr>
<td>1.6</td>
<td>A schematic of an ESI source</td>
<td>14</td>
</tr>
<tr>
<td>1.7</td>
<td>(A) Charge Residue Model; (B) Ion Desorption Model</td>
<td>14</td>
</tr>
<tr>
<td>1.8</td>
<td>A schematic of MALDI ion source</td>
<td>15</td>
</tr>
<tr>
<td>1.9</td>
<td>CID fragmentation</td>
<td>19</td>
</tr>
<tr>
<td>1.10</td>
<td>Peptide fragmentation nomenclature</td>
<td>19</td>
</tr>
<tr>
<td>1.11</td>
<td>AU-PAGE separation of yeast histones</td>
<td>22</td>
</tr>
<tr>
<td>1.12</td>
<td>Strategies for quantitative proteomic profiling</td>
<td>24</td>
</tr>
<tr>
<td>1.13</td>
<td>Overview of confounding effects due to internal standard addition at different points in sample preparation</td>
<td>27</td>
</tr>
<tr>
<td>2.1</td>
<td>An LC trace of yeast histones obtained from an LTQ-Orbitrap mass spectrometer</td>
<td>40</td>
</tr>
<tr>
<td>2.2</td>
<td>Yeast histone separation on a 16.5% Tris-tricine-SDS gel</td>
<td>48</td>
</tr>
</tbody>
</table>
2.3 LC-MS spectra of bovine histones from a Q-TOF mass spectrometer with a cone voltage of 21 V ..........................................................53
2.4 An LC trace of bovine histones from a Q-TOF mass spectrometer ...........54
2.5 LC-MS spectra of wild type yeast histones from a Q-TOF mass spectrometer ........................................................................56
2.6 LC-MS spectra of bovine histones from an LCT mass spectrometer ..........59
2.7 Deconvoluted LC-MS spectra of bovine histones from an LCT mass spectrometer .................................................................60
2.8 LC-MS spectra of yeast histones from an LCT mass spectrometer ..........61
2.9 LC-MS/MS spectra of precursor ion at 639.44 $^{2+}$ ................................62
3.1 Full MS results of peptide $^{9}$KSTGGKAPR$^{17}$ from an LTQ FT-ICR mass spectrometer ......................................................................75
3.2 Tandem MS results and LC traces of peptide $^{9}$KSTGGKAPR$^{17}$ from an LTQ FT-ICR mass spectrometer ......................................................76
3.3 Full MS results of peptide $^{73}$EIAQDFKTDLR$^{83}$ from an LTQ FT-ICR mass spectrometer ...............................................................77
3.4 Tandem MS results and LC traces of peptide $^{73}$EIAQDFKTDLR$^{83}$ from an LTQ FT-ICR mass spectrometer ......................................................78
3.5 LC traces and spectra of peptide $^{10}$STGGKAPR$^{17}$ from an LTQ FT-ICR mass spectrometer ...............................................................82
3.6 LC traces and spectra of peptide $^{18}$KQLASK$^{23}$ from an LTQ FT-ICR mass spectrometer ...............................................................83
3.7 LC traces and spectra of peptide $^{53}$RFQKSTELLIR$^{63}$ from an LTQ FT-ICR mass spectrometer ...............................................................84
3.8 LC traces and spectra of peptide $^{3}$TKQTAR$^{8}$ from an LTQ FT-ICR mass spectrometer ...............................................................85
3.9 LC traces and spectra of peptide $^{70}$LVREIAQDFKTDLR$^{83}$ from an LTQ FT-ICR mass spectrometer........................................................................................................86

3.10 LC traces of peptide $^{27}$KSAPSTGGVKKPHR$^{40}$ from an LTQ FT-ICR mass spectrometer ........................................................................................................87

3.11 Full MS results of peptide $^{27}$KSAPSTGGVKKPHR$^{40}$ from an LTQ FT-ICR mass spectrometer ........................................................................................................88

3.12 Tandem MS results of peptide $^{27}$KSAPSTGGVKKPHR$^{40}$ from an LTQ FT-ICR mass spectrometer ........................................................................................................89

3.13 Tandem MS results and LC traces of peptide $^{27}$KSAPSTGGVK$^{36}$ from an LCQ Deca XP+ ion trap mass spectrometer ................................................................................92

3.14 LC traces and spectra of peptide $^{1}$SGRGKGGKGLGKGGAKR$^{17}$ from an LCQ Deca XP+ ion trap mass spectrometer ................................................................................93

3.15 Tandem MS results and LC traces of peptide $^{73}$EIAQDFKTDLR$^{83}$ from an LCQ Deca XP+ ion trap mass spectrometer ................................................................................94

4.1 LC-MS spectra of H3 ..............................................................................................105

4.2 Global changes of histone H3 methylation in JmjC knockout strains ..............106

4.3 MALDI-TOF MS spectra of $^{3}$TKQTAR$^{8}$ ......................................................................109

4.4 LC-MS/MS spectra of peptides $^{28}$SAPSTGGVKKPHR$^{40}$ and $^{73}$EIAQDFKTDLR$^{83}$ ........................................................................................................110

4.5 Identification of demethylation sites in knockout strains by MS ......................111

4.6 Global changes of histone H3 methylation in JmjC overexpression strains ......114

4.7 Identification of demethylation sites in overexpression strains by MS ............115

4.8 LC-MS spectra of full length and truncated histone H3 from WT, RPH1 and JHD2 ......................................................................................................................116

4.9 In vitro enzyme assays of GST-Rph1(1-373) ..................................................................119

xvii
4.10 Control experiments for \textit{in vitro} enzyme assays of GST-Rph1 (1-373) ..........120

4.11 Growth phenotypes of the JmjC overexpression strains ........................................123

5.1 H3K56Ac changes in knockout and overexpression strains in response to
JmjC proteins quantified by spectral counting ..............................................................144

5.2 H3K56Ac changes in knockout and overexpression strains in response to
JmjC proteins quantified by peak abundance of all K56 related peptides .................146

5.3 H3K56Ac changes in knockout and overexpression strains in response to
JmjC proteins quantified by peak abundance of peptide $^{54}$FQKSTELLIR$^{63}$ ........147

5.4 H3K56Ac changes in knockout and overexpression strains in response to
JmjC proteins quantified by peak abundance of peptide $^{53}$RFQKSTELLIR$^{63}$ ................148

5.5 H3K56Ac changes in knockout and overexpression strains in response to
JmjC proteins quantified by peak abundance of peptides
$^{53}$RFQKSTELLIR$^{63}$ and $^{54}$FQKSTELLIR$^{63}$ .........................................................149

5.6 LC-MS spectrum of yeast histone H4 .......................................................................151

5.7 Relative changes of H4 acetylation in knockout and overexpression strains
in response to JmjC domains ......................................................................................152

5.8 Yeast H2B1 and H2B2 spectrum .............................................................................153

5.9 H2B1 acetylation changes in knockout and overexpression strains in
response to the JmjC proteins ......................................................................................154

5.10 H2B2 acetylation changes in knockout and overexpression strains in
response to the JmjC proteins ......................................................................................155

5.11 Yeast H2A spectrum ...............................................................................................158

5.12 H2A acetylation changes in knockout and overexpression strains in
response to the JmjC proteins ......................................................................................159

S5.1 Sequence of yeast histone H3 ...............................................................................160

xviii
A.1 An LC trace of yeast histones on LTQ-Orbitrap .......................................................201
A.2 LC-MS spectra of bovine H2B from a Q-TOF mass spectrometer with a cone voltage of 30 V .................................................................202
A.3 LC-MS spectra of bovine histone H2B from a Q-TOF mass spectrometer with a cone voltage of (a) 21V, (b) 25 V, (c) 35 V, and (d) 50 V .......................203
A.4 LC-MS spectra of bovine histones from a Q-TOF mass spectrometer with a cone voltage of 50 V .................................................................204
A.5 LC-MS spectra of bovine histones from a Q-TOF mass spectrometer with a cone voltage of 25 V ...........................................................................205
B.1 The five JmjC domain-containing proteins in S. cerevisiae .....................................207
B.2 LC-MS/MS spectra of peptide \(^{73}\)EIAQDFKTLR\(^{83}\) ...........................................208
B.3 LC-MS/MS spectra of peptide \(^{73}\)EIAQDFK\(_{Me}^{1}\)TDLR\(^{83}\) ....................................209
B.4 LC-MS/MS spectra of peptide \(^{73}\)EIAQDFK\(_{Me}^{2}\)TDLR\(^{83}\) ....................................210
B.5 LC-MS/MS spectra of peptide \(^{73}\)EIAQDFK\(_{Me}^{3}\)TDLR\(^{83}\) ....................................211
B.6 LC-MS/MS spectra of peptide \(^{28}\)SAPSTGGVK\(_{Me}^{3}\)KPHR\(^{40}\) ..............................212
B.7 LC-MS/MS spectra of peptide \(^{28}\)SAPSTGGVK\(_{Me}^{3}\)KPHR\(^{40}\) ..............................213
B.8 Western blot of overexpressed JmjC proteins .......................................................214
B.9 Growth phenotype assays of yeast cells ...............................................................214
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABC</td>
<td>ammonium bicarbonate</td>
</tr>
<tr>
<td>Ac</td>
<td>acetylation</td>
</tr>
<tr>
<td>Ac0</td>
<td>unmodified</td>
</tr>
<tr>
<td>Ac1</td>
<td>monoacetylated/monoacetylation</td>
</tr>
<tr>
<td>Ac2</td>
<td>diacetylated/diacetylation</td>
</tr>
<tr>
<td>Ac3</td>
<td>triacetylated/triacetylation</td>
</tr>
<tr>
<td>Ac4</td>
<td>tetraacetylated/tetraacetylation</td>
</tr>
<tr>
<td>Ac5</td>
<td>pentaacetylated/pentaacetylation</td>
</tr>
<tr>
<td>ACN</td>
<td>acetonitrile</td>
</tr>
<tr>
<td>APS</td>
<td>ammonium persulfate</td>
</tr>
<tr>
<td>AU-PAGE</td>
<td>acetic acid-urea polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>AUT-PAGE</td>
<td>acetic acid Triton-polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>BCA</td>
<td>bicinchoninic acid</td>
</tr>
<tr>
<td>βME</td>
<td>β-mercaptoethanol</td>
</tr>
<tr>
<td>CHCA</td>
<td>α-cyano-4-hydroxycinnamic acid</td>
</tr>
<tr>
<td>CAD</td>
<td>collision-activated dissociation</td>
</tr>
<tr>
<td>CID</td>
<td>collision-induced dissociation</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>CoA</td>
<td>coenzyme A</td>
</tr>
<tr>
<td>2DE</td>
<td>two-dimensional gel electrophoresis</td>
</tr>
<tr>
<td>DTT</td>
<td>1,4-dithiothreitol</td>
</tr>
<tr>
<td>ECD</td>
<td>electron capture dissociation</td>
</tr>
<tr>
<td>E. coli</td>
<td><em>Escherichia coli</em></td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>ESI</td>
<td>electrospray ionization</td>
</tr>
<tr>
<td>ETD</td>
<td>electron transfer dissociation</td>
</tr>
<tr>
<td>FT-ICR</td>
<td>Fourier-transform-ion cyclotron resonance</td>
</tr>
<tr>
<td>GST</td>
<td>glutathione-S-transferase</td>
</tr>
<tr>
<td>h</td>
<td>hour</td>
</tr>
<tr>
<td>HAT</td>
<td>histone acetyl-transferase</td>
</tr>
<tr>
<td>HDAC</td>
<td>histone deacetylase</td>
</tr>
<tr>
<td>HEPES</td>
<td>4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid</td>
</tr>
<tr>
<td>I.D.</td>
<td>inner diameter</td>
</tr>
<tr>
<td>JmjC</td>
<td>Jumonji C</td>
</tr>
<tr>
<td>kV</td>
<td>Kilovolt</td>
</tr>
<tr>
<td>LC</td>
<td>liquid chromatography</td>
</tr>
<tr>
<td>LC-MS</td>
<td>liquid chromatography-mass spectrometry</td>
</tr>
<tr>
<td>LC-MS/MS</td>
<td>liquid chromatography-tandem mass spectrometry</td>
</tr>
<tr>
<td>Term</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>------------------------------------------</td>
</tr>
<tr>
<td>LSD1</td>
<td>Lysine Specific Demethylase 1</td>
</tr>
<tr>
<td>NIB buffer</td>
<td>nuclear isolation buffer</td>
</tr>
<tr>
<td>M</td>
<td>molarity</td>
</tr>
<tr>
<td>MALDI</td>
<td>matrix assisted laser desorption</td>
</tr>
<tr>
<td>MALDI-TOF</td>
<td>matrix assisted laser desorption/ionization-time of flight</td>
</tr>
<tr>
<td>Me</td>
<td>methylation</td>
</tr>
<tr>
<td>Me0</td>
<td>unmodified</td>
</tr>
<tr>
<td>Me1</td>
<td>monomethylated/monomethylation</td>
</tr>
<tr>
<td>Me2</td>
<td>dimethylated/dimethylation</td>
</tr>
<tr>
<td>Me3</td>
<td>trimethylated/trimethylation</td>
</tr>
<tr>
<td>MES</td>
<td>2-morpholin-4-ylethanesulfonic acid</td>
</tr>
<tr>
<td>Min</td>
<td>minute</td>
</tr>
<tr>
<td>MS</td>
<td>mass spectrometry</td>
</tr>
<tr>
<td>MW</td>
<td>molecular weight</td>
</tr>
<tr>
<td>NP-40</td>
<td>Nonidet P-40</td>
</tr>
<tr>
<td>PBS</td>
<td>phosphate buffered saline</td>
</tr>
<tr>
<td>PIPES</td>
<td>2-[4-(2-sulfoethyl)piperazin-1-yl]ethanesulfonic acid</td>
</tr>
<tr>
<td>PMF</td>
<td>peptide mass fingerprinting</td>
</tr>
<tr>
<td>PMSF</td>
<td>phenylmethysulfonyl fluoride</td>
</tr>
<tr>
<td>PTM</td>
<td>posttranslational modification</td>
</tr>
</tbody>
</table>

xxii
OE overexpression
RP-HPLC reversed-phase-high pressure liquid chromatography
RPLC reversed-phase liquid chromatography
Rpm revolutions per minute
RT retention time
SC-Uracil synthetic complete-Uracil
*S. cerevisiae* *Saccharomyces cerevisiae*
SD synthetic defined
SDS sodium dodecyl sulfate
SDS-PAGE sodium dodecyl sulfate-polyacrylamide gel electrophoresis
TCA trichloroacetic Acid
TFA trifluoroacetic acid
WT wild type
YPD yeast extract, peptone, and dextrose
CHAPTER 1

INTRODUCTION

1.1 HISTONES AND CHROMATIN STRUCTURE

1.1.1 Background

Histones are the primary protein component of chromatin. They include core histones, H3, H4, H2A, H2B, and linker histones, H1. As shown in Figure 1.1A, a tetramer of H3 and H4 and two dimers of H2A and H2B form an octamer and are wrapped by double stranded DNA. The linker histone links the nucleosome into higher order chromatin structures, as shown in Figure 1.1B. The structural details of histone-histone and histone-DNA interactions have been thoroughly investigated. It has been demonstrated that dimers of (H2A/H2B) bind to either side of the (H3/H4)$_2$ tetramer and form four histone dimers linked end-to-end: (H2A/H2B)- (H4/H3)- (H3/H4)- (H2B/H2A)\(^1\). The (H3/H4)$_2$ tetramer is wrapped by >120 base pairs (bp) of DNA forming a stable complex\(^9,10\) and dimers of (H2A/H2B) extend the wrapping of DNA to >160 bp\(^4,9,11\).

The N-termini of all four core histones and the C-terminus of histone H2A are called “tail” domains\(^12\). The histone tails are subject to a large number of posttranslational modifications (PTMs) and play important roles in internucleosomal contact and the self-assembly of chromatin fibers\(^13-18\). It has been suggested that certain
PTMs are related to specific functional and/or conformational states of the chromatin fiber 19, 20.

Figure 1.1 Histone arrangement in nucleosomes. Figure A was adapted from Wheeler, R., 2005 http://en.wikipedia.org.

As shown in Figure 1.1B, the presence of histone H1 requires an octamer of core histones in order to bind DNA. Compared with the interaction between core histones and chromatin, the association of histone H1 with chromatin is much less stable 21 and the binding of H1 partially rearranges the interactions among core histones 22, 23.
1.1.2 Histone Posttranslational Modifications

The histone tails are subject to various PTMs including acetylation, methylation, phosphorylation, and ubiquitination. These modifications usually occur on specific sites as demonstrated in Table 1.1. Efforts have been made to investigate the effects of histone PTMs on chromatin structures and functions. It has been found that histone acetylation helps relax chromatin and activate transcription, while histone deacetylation induces closed chromatin structure and transcription repression. Upon acetylation the chromatin fiber and the interaction of histone tails with nucleosomal DNA are destabilized. Furthermore, histone acetylation can facilitate DNA assembly into nucleosomes. In contrast to acetylation, histone methylation is linked to both transcriptional activation and repression. In Tetrahymena, methylation of K4 on H3 correlates with transcriptionally active nuclei, while methylation of K9 correlates with transcription repression in Drosophila.

Though acetylation and methylation are the most widely studied PTMs, significant progress has been made in the study of other PTMs. It has been demonstrated that phosphorylation of H3 is associated with chromatin compaction and that phosphorylation of H1 weakens the interaction between H1 tails and DNA. In addition, phosphorylation may inhibit transcription. Ubiquitin is a highly conserved 76 amino acid protein that is critical for signaling 26S proteomosal degradation. Ubiquitin also plays key roles in various biological processes, including transcription initiation and elongation, silencing, and DNA repair. Histones can be monoubiquitinylated and polyubiquitinylated. It is believed that ubiquitinylated H2B is related to telomeric...
silencing \(^\text{34}\) and facilitates the association of H1 with the nucleosome \(^\text{35}\). Each of these PTMs plays critical and unique roles in chromatin structure and function. However, they often work in a synergistic fashion. For example, it has been reported that methylation of K4 on H3 requires prior ubiquitination of H2B \(^\text{36}\) and that histone H3 phosphorylation and acetylation act synergistically \(^\text{37}\).

**Table 1.1** Mass shift and modification site for common histone PTMs.

<table>
<thead>
<tr>
<th>Modification</th>
<th>Mass shift (Da)</th>
<th>Residues</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetylation</td>
<td>42.0106</td>
<td>K, N-terminal</td>
</tr>
<tr>
<td>Methylation</td>
<td>14.0156</td>
<td>K, R</td>
</tr>
<tr>
<td>Phosphorylation</td>
<td>79.9663</td>
<td>S, T, Y</td>
</tr>
<tr>
<td>Oxidation</td>
<td>15.9949</td>
<td>M</td>
</tr>
<tr>
<td>Ubiquitination</td>
<td>114.0429</td>
<td>K</td>
</tr>
</tbody>
</table>

As mentioned above, PTMs are critical for gene transcription. Because of this, histone PTMs are implicated in many human disorders such as leukemia, epithelial cancers, Huntington’s disease, and fragile X syndrome. For example, the ubiquitylation of H2A is decreased significantly in prostate tumors \(^\text{38}\); aberrant acetylation of histone tails is strongly linked to carcinogenesis \(^\text{39}\), and detection of histone PTMs has been suggested for the diagnosis and treatment of lung cancer \(^\text{40}\). In addition, based on these findings drugs have been designed to target specific PTMs in the treatment of these disorders \(^\text{41}\). As shown in **Table 1.2**, histone deacetylase (HDAC) inhibitors that increase the level of histone acetylation are currently under investigation in ongoing clinical trials \(^\text{24}\).
**Table 1.2** HDAC inhibitors in clinical trials\(^2^4\),\(^4^2\). ALL-acute lymphoblastic leukemia; AML-acute myeloid leukemia; CLL-chronic lymphoblastic leukemia; EBV-Epstein-Barr virus; NHL-non-Hodgkin’s lymphoma; NSCLC-non-small-cell lung cancer; SCLC-small-cell lung carcinoma.

<table>
<thead>
<tr>
<th>HDAC inhibitor</th>
<th>Phase</th>
<th>Disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phenylbutyrate</td>
<td>I/II</td>
<td>Malignant glioma</td>
</tr>
<tr>
<td>Phenylbutyrate + tretinoin</td>
<td>I</td>
<td>Leukemia, myelodysplastic syndromes, myelodysplastic/myeloproliferative diseases</td>
</tr>
<tr>
<td>Phenylbutyrate + 5-azacytidine</td>
<td>I</td>
<td>Severe thalassemia</td>
</tr>
<tr>
<td>Phenylbutyrate + azacitidine</td>
<td>II</td>
<td>AML, NHL, multiple myeloma, NSCLC, prostate carcinoma</td>
</tr>
<tr>
<td>Phenylbutyrate + fluorouracil + indomethacin + interferon gamma</td>
<td>I/II</td>
<td>Advanced colorectal adenocarcinoma</td>
</tr>
<tr>
<td>Phenylbutyrate + dexamethasone + sargramostim</td>
<td>II</td>
<td>Refractory or relapsed AML</td>
</tr>
<tr>
<td>Arginine butyrate + gancyclovir</td>
<td>I</td>
<td>EBV-induced malignancies/lymphoproliferative disorders</td>
</tr>
<tr>
<td>Depsipeptide</td>
<td>I</td>
<td>CLL, AML, ALL</td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>Refractory/progressive SCLC, NSCLC</td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>Cutaneous T-cell lymphoma, relapsed peripheral T-cell lymphoma</td>
</tr>
<tr>
<td>SAHA</td>
<td>I</td>
<td>Advanced multiple myeloma</td>
</tr>
<tr>
<td>MS-275</td>
<td>I</td>
<td>Advanced solid tumors or lymphoma</td>
</tr>
<tr>
<td></td>
<td>I</td>
<td>Hematological malignancy</td>
</tr>
</tbody>
</table>
1.1.3 Chromatin Remodeling Enzymes

Several enzymes catalyze the transfer of modification groups to histones and their subsequent removal. Histone acetyl-transferases (HATs) transfer acetyl groups from acetyl-coenzyme A to histones while HDAC removes the acetyl group (Figure 1.2). These two reactions are at dynamic equilibrium and the net result determines the level of acetylation at particular sites. A large number of proteins have been recognized to have HAT activities. Most of the HATs function as transcriptional co-activators and prefer a specific substrate. Functioning in opposition to HATs, HDACs induce transcriptional repression usually together with proteins functioning in recruitment, corepression or chromatin remodeling. To date, at least 18 members have been identified as HDACs and categorized into three different classes. Most of them work through a Zn$^{2+}$-dependent mechanism. HDACs associate with a number of well characterized cellular oncogenes and tumor-suppressor genes. It is noteworthy that HDACs not only deacetylate histones but other proteins as well.

Similar to acetylation, histone phosphorylation is regulated by protein kinases and protein phosphatases, as shown in Figure 1.3. The phosphate donor can be a nucleotide triphosphate, such as ATP. Though protein kinases can phosphorylate histones in vitro at many residues, histones are phosphorylated in vivo only at serine, threonine, and tyrosine. The most extensively phosphorylated histone is H1.
Figure 1.2 The histone acetylation-deacetylation cycle. CoA: coenzyme A; TSA: trichostain A.

Figure 1.3 A protein phosphorylation-dephosphorylation cycle. ATP: adenosine triphosphate; ADP: adenosine diphosphate.

Similar to acetylation and phosphorylation, ubiquitination is reversible and the ubiquitination level is determined by the competing reactions of adding and removing ubiquitin moieties from histones. For ubiquitin to attach to other proteins, two or three
enzymes are required: a ubiquitin-activating enzyme E1, a ubiquitin-conjugating enzyme E2, and sometimes a ubiquitin-protein ligase E3. The ubiquitin moiety can be removed by deubiquitinylase enzymes.

Methylation was thought to be irreversible until 2004 when Shi discovered the Lysine Specific Demethylase 1 (LSD1), which reverses dimethylation. Since then, a number of proteins have been recognized having the ability to demethylate lysine and arginine through amine oxidation, hydroxylation or deimination. Figure 1.4 shows the proposed mechanisms of lysine demethylation by LSD1 and JmjC proteins. Out of 27 JmjC domain-containing proteins in the human genome, 15 have been reported to demethylate lysine or arginine. Studies suggested that histone demethylation is required for “fine-tuning” transcription. In addition, demethylases are found to correlate with a number of human diseases and possibly can be used as drugs to treat particular disorders. For example, both demethylases FBXL10 and JMJD5 were demonstrated to have potential in suppressing tumors.
Figure 1.4 Mechanisms of lysine demethylation by LSD1 and JmjC proteins. (Top) LSD1 demethylates H3K4Me2/Me1 via an amine oxidation reaction. (Bottom) JmjC proteins use αKG and iron as cofactors to hydroxylate the methylated histone substrate.

1.2 STRATEGIES TO INVESTIGATE HISTONE POSTTRANSLATIONAL MODIFICATIONS

1.2.1 Traditional Approaches

Traditionally, micro-sequencing and immunoassay are used to study histone PTMs. However, microsequencing uses Edman degradation, which involves tedious derivatization and in practice can only sequence peptides with no more than 30 amino acids. In addition, the proteins to be sequenced need to be well purified. Immunoassay includes three major approaches: immunoprecipitation, Western-blot, and immunofluorescence. Each of these requires the use of antibodies to target specific PTM sites. Therefore, problems with the availability, site specificity, and cross-reactivity of
antibodies arise when immunoassays are employed. Furthermore, immunoassay is not suitable to discover novel PTM sites. False positives occur when concomitant modifications are present. For example, conversion of serine 10 to glutamate or aspartate results in almost complete loss of H3 acetylation at K9.

Quantitative analysis is an important area in characterization of histones and their PTMs. Before MS became widely used, a number of methods were developed for protein/peptide quantification. Some of them are still used today to quantify histones and their PTMs. These methods can be categorized in four groups.

1. Polyacrylamide gel-based methods, such as autoradiography, fluorography, densitometry, and fluorescence labeling. Autoradiography and fluorography require samples to be radioactively labeled with isotopes such as $^3$H and $^{125}$I. After labeling, the sample is separated using SDS-PAGE, acetic acid-urea PAGE (AU-PAGE), Triton AU-PAGE (AUT-PAGE), or 2D SDS-PAGE and then the film is exposed to beta particles or gamma rays. The obtained autoradiograph and fluorograph are used to determine the relative abundance of histones of interest and their PTMs. Densitometry uses an optical digital densitometer measuring the light absorption of Coomassie stained gel bands. Fluorescence labeling methods here refer to difference gel electrophoresis and multiplexed proteomics. These two methods fluorescently label 2D gels and the fluorescent signals are utilized to determine the change of proteins and PTMs. However, polyacrylamide gels have limited separation capability. It is not uncommon for a single gel band or spot to contain more than one protein making it difficult to quantify specific proteins. In addition, the dynamic range of gel-based quantification is limited.
For example, the linear dynamic range of response from silver stains is typically 8-10 fold. Furthermore, different modifications have different responses in fluorography.

2. Chromatography-based methods. When proteins/peptides are separated by chromatography, estimation of the content of an eluted peak can be measured by its absorption of UV light. However, chromatography has limited resolution. In many cases, several proteins/peptides coelute into one peak. In addition, the molar absorbance of UV light varies depending on the composition of the proteins/peptides.

3. Protein array-based methods, such as chromatin immunoprecipitation and antibody microarrays. By using these methods, acetylation and phosphorylation have been quantified. However, these methods require antibodies to target specific modifications. Therefore, they are limited by availability, cross-reactivity, and site specificity of antibodies. Furthermore, these methods cannot be used to discover novel PTMs.

4. Sequence analysis-based methods, such as Edman degradation, which is the most widely used. Sequencing methods are very time consuming and require relatively pure proteins/peptides. With the development of MS, the frequency of Edman degradation used for protein analysis has decreased.

1.2.2 Mass Spectrometry-Based Approach

Mass spectrometry (MS) is being used widely to study histones and their PTMs. MS is an analytical technique used to measure the mass of gas phase ions. Compared
with the traditional approaches mentioned above, MS is relatively fast and able to detect all of the histones and PTM sites simultaneously. Most importantly, MS is excellent for the discovery of novel PTMs. The typical experimental design of MS is shown in Figure 1.5. Histones are first extracted from cells or tissues, and then subject to separation by chromatography or gel electrophoresis. Proteins are either analyzed intact by LC-MS or digested into peptides for analysis by LC-MS/MS and/or MALDI-TOF MS. When histones are separated by gel electrophoresis, the proteins are in-gel digested and analyzed by LC-MS/MS and/or MALDI-TOF MS.

\[ \text{Figure 1.5 Typical proteomic workflow for histone characterization.} \]
1.2.2.1 Ionization Sources of Mass Spectrometers

Typically a mass spectrometer consists of a source, mass analyzer, and detector. The source is used to vaporize and ionize the samples. The most widely used ionization methods are electrospray ionization (ESI) and matrix assisted laser desorption/ionization (MALDI). Both of these are soft ionization techniques capable of producing intact macromolecular ions with little fragmentation and in some cases even noncovalent complexes. Thus, ESI and MALDI are especially suited for the analysis of nonvolatile biopolymers.

Since the development by Dole 71, 72, the ESI source has undergone continuous improvement although the general arrangement remains basically the same as Figure 1.6. The sample is introduced to the source either by a syringe pump or as an eluent from liquid chromatography. As the sample solution reaches the end of the electrospray needle which has a high potential, a Taylor cone is formed. Currently the “Charged Residue Model” (Figure 1.7A) and “Ion Desorption Model” (Figure 1.7B) are popular mechanisms explaining electrospray. The Charged Residue Model 72 suggests that the droplet undergoes even or uneven fission and charges on the daughter droplets are proportional to their relative size. The Ion Desorption Model 73 suggests that columbic repulsion results in breaking off of the droplet and that compared to the original droplet the daughter droplets have negligible loss in droplet mass but a significant drop in charge. Either a curtain of heated nitrogen gas (Figure 1.6A) or a heated capillary (Figure 1.6B) is used for desolvation. The skimmer is used for ion focusing.
Figure 1.6 A schematic of an ESI source.

Figure 1.7A Charge Residue Model.  

Figure 1.7B Ion Desorption Model.
Developed in the 1980s, MALDI has a comparatively simple ion source setup as shown in Figure 1.8. In MALDI, a mixture of the sample and a matrix, which has a strong absorption band in UV spectrum, is spotted on a target and irradiated by UV light. The rapid heating of the mixture due to UV absorption by the matrix results in the sublimation of the matrix and expansion of the matrix into the gas phase with intact analyte molecules. Ionization may occur at any time during the expansion and acceleration, but the origin of the ion formation is not clear yet. Among many different lasers that have been used, UV lasers are most commonly used due to their easy operation and low price. A number of chemicals can be used as the matrix. A good matrix should have strong absorption at the laser wavelength, have a smaller MW than the analyte, be sublimable, demonstrate vacuum stability, and lack chemical reactivity. Based on these criteria, α-cyano-4-hydroxycinnamic acid is widely used for peptide and protein analysis. Unlike ESI, which creates singly and multiply charged ions, MALDI generates singly charged ions only.

Figure 1.8 A schematic of a MALDI ion source.
1.2.2.2 Liquid Chromatography-Mass Spectrometry (LC-MS)

Sample preparation, especially sample purification, is crucial in MS. By purifying crude histones, chemical background due to other proteins in the sample can be reduced. Several different chromatographic techniques have been employed to purify histones: size exclusion chromatography, ion-exchange chromatography, hydrophilic-interaction liquid chromatography, and RP-HPLC. Among the various liquid chromatographic techniques, RP-HPLC is the most widely used. For example, histones can be effectively separated by use of a reversed-phase C$_{18}$ column combined with water (solvent A) and acetonitrile (solvent B) with the ion pairing agent trifluoroacetic acid (TFA) as the mobile phases. In reversed-phase separations, the proteins and peptides are eluted in order of increasing hydrophobicities. The eluent can be directed to a mass spectrometer for online analysis or collected as fractions for offline MS analysis. By MS, the molecular weights of the separated proteins are determined. Specific shifts in the proteins’ masses are evidence of PTMs. Table 1.1 lists several common PTMs along with their associated mass shift. While this approach is very powerful for detecting PTMs, it does not provide sufficient information to determine the PTM sites. Nor can it differentiate proteins or PTMs that are isobaric (same nominal mass) such as trimethylation and acetylation. Additional LC-MS/MS analysis is often needed to corroborate the results obtained by LC-MS.
1.2.2.3 Tandem Mass Spectrometry

Tandem mass spectrometry, also known as MS/MS or MS², combines two or more mass analyzers of the same or different types. The first mass analyzer isolates an ion(s) of interest at a given m/z (precursor ion). The precursor ion(s) is then fragmented between the first and second mass analyzer typically via collisionally activated (or collisionally induced) dissociation. The second mass analyzer is used to detect and measure the m/z of the fragment ions (product ions). Four main scan modes are used in MS/MS.

1. Product ion scan. The precursor ions or ions at a given m/z are selected in MS1 and the resulting product ions are determined in MS2.

2. Precursor ion scan. Product ions are selected in MS1 and the associated precursor ions are then determined by scanning MS1.

3. Neutral loss scan. A neutral fragment is selected by locking MS1 and MS2 to scan a constant mass difference. The resulting data provide the precursor ions that give rise to the neutral loss.

4. Selected reaction monitoring (SRM). MS1 is set to pass a specific precursor ion and MS2 is set to pass a specific product ion. SRM measurements are the most sensitive because MS1 and MS2 are always locked on a single ion of interest. (Note: multiple reaction monitoring involves changing MS1 and MS2 very rapidly to monitor multiple species but at the cost of sensitivity.)

A number of fragmentation techniques have been developed, including collision-induced or activated dissociation (CID/CAD, or CXD), post source decay (PSD), infrared
multi-proton dissociation (IMPD), electron transfer dissociation (ETD), and electron capture dissociation (ECD). CXD is currently the most popular fragmentation method. The schematic process for CID is shown in Figure 1.9. The precursor ions collide with the collision gas and build up internal energy until the fragmentation threshold is reached, causing the product ions to form. When the internal energy is close to but lower than the fragmentation threshold, usually ions due to neutral losses (H₂O, NH₃, MeOH, CO, CO₂ etc.) are generated. At higher internal energy, peptide backbones are cleaved, resulting in b ions and y ions as shown in Figure 1.10. However, CXD cannot generate all of the desired product ions to obtain complete information about the primary structure of a polypeptide 77, 78.

ECD produces odd-electron, free-radical driven fragmentation that is often complementary to that obtained from CXD 79. ECD cleaves amide backbones and results in c ions and z ions as shown in Figure 1.10. Compared to CXD, ECD has a narrower dynamic range of fragmentation, which makes it particularly suitable for fragmentation of larger polypeptides, including intact proteins 80-82. Unlike CID, ECD can be used to fragment larger peptides produced by enzymes such as Lys-C or Arg-C 83-85. Though ECD has advantages over CXD, it is not practical to use ECD-only approaches to replace CXD completely. First, ECD has much poorer fragmentation efficiency than CXD. Thus, low abundant species are not easily characterized. Second, the instrument complexity and cost increase significantly with ECD. Currently ECD is exclusively used in FT-ICR mass spectrometer whereas CXD is a standard option on nearly all mass spectrometers. Despite
the differences between these approaches, they provide complementary information for determining the primary structure of a polypeptide. 

Figure 1.9 CID fragmentation.

Figure 1.10 Peptide fragmentation nomenclature.
1.2.2.4 Peptide Mass Fingerprinting (PMF)

PMF, also known as peptide mass mapping, is a method used to identify proteins by comparing the observed $m/z$ to the theoretical $m/z$ of the digested protein. Usually the peptide’s amino acid compositions and PTMs can be determined with relative certainty based on high-mass accuracy $m/z$ measurement alone. For this reason, PMF is very efficient in protein identification. It has been widely used in identifying and mapping PTMs on histones from a number of organisms. However, peptides that differ in amino acid composition or have isobaric masses can confound identification when only mass is considered. Thus, MS/MS is now commonly used in addition to PMF to corroborate the peptide assignments.

1.2.2.5 Acetic Acid-Urea Polyacrylamide Gel Electrophoresis (AU-PAGE) and Acetic Acid Triton-Polyacrylamide Gel Electrophoresis (AUT-PAGE)

Proteins can be separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) based on their MWs. However, SDS-PAGE has limited resolution for the separation of proteins of very similar MWs or isoforms that differ in mass by PTMs such as acetylation and phosphorylation. One approach to enhance the separation of modified proteins is to separate them on the basis of MW and charge. Described by Chalkley in 1969, AU-PAGE has become an important technique to prepare and separate small amounts of histone isoforms. In AU-PAGE proteins are separated at low pH (pH = 3). In acidic medium, histones are positively charged on the N-terminal and basic side chains. Acetylation removes one positive charge while
phosphorylation introduces a negative charge. Therefore, the overall charge state of histones with different levels of acetylation and phosphorylation varies and in turn the histones demonstrate different ion mobilities in AU-PAGE. The unmodified histones have a higher mobility relative to acetylated forms. The higher the acetylation level, the smaller the mobility. Generally the MW decreases from H1, H3, H2A, H2B, to H4. Therefore, based on size and charge, histones are separated as shown in Figure 1.11. From the top to the bottom are H1, H3, H2A, H2B, and H4. For the same histone, the unmodified isoform (Ac0) is at the bottom and the acetylated isoforms are at the top. As seen in Figure 1.11, isoforms of H3 and H4 are well separated. However, those of H2A and H2B overlap. To resolve H2A and H2B, the addition of Triton X-100 to the AU-PAGE (AUT-PAGE) is commonly performed. 

Core histones but not linker histones bind to nonionic detergent Triton X-100 with an affinity of H2A>H3, H4>H2B. When bound to Triton X-100, the effective mass of a protein increases within the gel without affecting its charge, and thus its mobility is significantly reduced during electrophoresis. By altering the histone mobility, isoforms are now well separated with AUT-PAGE. The addition of Triton X-100 does not affect the separation of H1. Histone H4 is still at the bottom of the gel. However, H2A moves up due to its strong binding with Triton X-100. In addition, due to the binding with Triton X-100 H2A is better resolved from H3 and H2B. By using 30-cm-long AUT PAGE gels, yeast histones were successfully separated based on their acetylation levels. H3 histones were separated into six bands, H4 five bands, H2A three bands, and H2B five bands. Theoretically speaking, separation between all core histones of diverse species can
always be achieved by adjusting the concentration of Triton X-100. An optimal concentration of Triton X-100 can be determined by producing a Triton X-100 gradient in the gel. As the concentration of Triton X-100 changes in the gel, the distribution pattern of histones changes but always with H4 at the bottom of the gel. Furthermore, changing the concentration of Triton X-100 can further improve the separation of specific histone isoforms.

Figure 1.11 AU-PAGE separation of yeast histones.
1.2.2.6 Quantitative Analysis of Histones and Their Posttranslational Modifications

With the development of MALDI and ESI, MS has been used more widely in quantitative analysis of histones and their PTMs. Due to variability in ionization with MALDI and ESI and the influence of other ions, the absolute signal abundance of a given peptide ion does not always reflect its amount. A number of quantitative approaches have been developed and are summarized in Figure 1.12. These approaches are separated into 2D gel electrophoresis, stable isotope labeling, and abundance-based approaches. Due to the concern of variations in protein and peptide ionization, all of the MS based quantitative strategies rely on the use of an internal or reference standard to account for the variations resulting from ionization differences and the sample matrix. An ideal internal/reference standard is chemically and physically identical to the analyte. These internal or reference standards are often generated by labeling the analyte to alter its mass (stable isotope labeling) and spectroscopy (difference gel electrophoresis), or are used for label-free abundance based comparisons.

Stable isotope labeling has been used to characterize and quantitate histone modifications. Stable isotope labeling approaches are separated into two categories based on the point in sample preparation when the label is added. In vivo labeling methods involve incorporating stable isotopes into cells grown in a medium containing labeled amino acids or glucose. Those isotopes are incorporated into the cells at the early stage of sample preparation and thus reduce variations between samples and yield highly accurate quantification. However, these methods rely on metabolic labeling and are not amenable to samples such as body fluid and tissues, which do not grow in cell culture.
vivo labeling can also require a relatively long time to fully incorporate the isotopes\(^9\). In contrast, in vitro labeling methods involve incorporation of stable isotopic tags onto certain amino acids via an in vitro chemical reaction. This approach can be applied not only to cultured cells but also to tissues and body fluids. In addition, more options are available for in vitro labeling. Unfortunately, these methods do not avoid the challenges faced by in vivo labeling, such as the high expense, long incorporation time, and chromatography shifts resulted from the incorporated isotopes\(^1\). In order to circumvent the limitations faced by in vivo and in vitro labeling, label-free abundance-based peptide/protein quantitation are of great interest for quantitative analysis of histones and their PTMs.

**Figure 1.12** Strategies for quantitative proteomic profiling. 2DE - two-dimensional gel electrophoresis, SILAC - stable isotope labeling with amino acids in cell culture, iTRAQ - isobaric tags for relative and absolute quantitation, ICAT - isotope-coded affinity tagging, MCAT - mass-coded abundance tagging, NIT - N-terminal isotope-encoded tagging\(^9\).
Label-free abundance-based quantitation avoids many of the complications associated with in vivo and in vitro labeling. Label-free approaches can be carried out with or without the use of internal standards. When internal standards are used, the standards are usually unmodified peptides either from the sample itself (endogenous standard) or other peptides spiked into the sample (exogenous standard). Kirschner developed a phosphorylation quantitative method that mixes synthetic, isotopically labeled peptides with studied peptides at a constant ratio. The spiked peptide has the same sequence as the studied peptide but a different mass thus making it an ideal internal standard. This approach requires pre-knowledge of the identity and quantity of the studied peptide. When an isotopically labeled peptide is not available, peptides with similar characteristics to the studied peptides are usually used as internal standards. For example, Laganà synthesized a peptide having the same sequence except for one amino acid with the studied peptide and successfully used it as an internal standard.

One key concern with exogenous standards is at what point in the sample preparation the standard is added. Ideally the standard should be added before any sample manipulation in order to properly account for sample losses that occur during sample preparation (Figure 1.13). Therefore, a better option of internal standards would be an endogenous peptide. Kinter developed a “Native Reference Peptide” method using an unmodified peptide as the internal standard to track fluctuations in instrumental response and sample preparation. Similarly, Zhang, Hansen, and Tabata also developed quantitative analysis methods using endogenous standards. However, peptides
meeting the criteria for endogenous standards are not always available and quantitative methods without the need for internal standards are desired.

One approach that does not rely on internal standards is the Accurate Mass and Time (AMT) tag method. AMT is a high-throughput method based on high-resolution / high-mass accuracy mass spectrometry combined with high efficiency LC separation to identify and quantify peptide ions. Relative abundances of peptides are compared with predetermined normalized peak abundances in order to determine changes in protein/peptide expression. Since accurate peak abundances are crucial for this approach, proper control of sample processing, and instrumental performance and data analysis are critical to obtain reliable quantitative results. For optimized AMT methods, ESI-MS results in a linear response for peptide abundance vs. peptide concentration in the sample. Thus AMT can be used to carry out relative quantitative proteomics by simply comparing peak abundances between MS runs. Statistical tools such as Student’s t-test can be used to evaluate the significance of peptide abundance variation when comparing multiple LC-MS runs. However, this approach requires at least three MS runs to obtain a meaningful comparison. Therefore, it can be time consuming and often requires more samples than label-based methods.
1.3 OBJECTIVE

As mentioned above, histones and their PTMs play crucial roles in gene transcription. Deciphering these roles facilitates our understanding of a variety of biological processes. Traditionally, micro-sequencing and immunoassays are used to study histones and their PTMs. However, these methods are not adept at discovery of novel PTMs. MS-based approaches have been applied to characterize histones and their
PTMs. The objective of my doctoral research is to develop novel methods for qualitative and quantitative analysis of histones and their PTMs using MS-based approaches.

Chapter 2 describes the development of methods for yeast histone preparation, separation, and LC-MS characterization. Due to the low number of histone genes in yeast, the histones are at low concentrations. Efforts have been made to extract histones from the whole cells and nuclei. Using a number of published protocols, I obtained yeast histones that were of low purity and poor yield. To improve the quantity and purity of histones obtained from yeast, I optimized the protocol through extensive washing of nuclei prior to lysis and acid extraction. Using this yeast histone extraction protocol, the separation of histones by SDS-PAGE was optimized. Traditionally, a Laemmli system is used for this purpose. However, this system has limited resolution for small proteins like yeast histones. Therefore, I evaluated the separations of tricine SDS-PAGE over a range of acrylamide cross-linking percentages. In addition to SDS-PAGE, the LC-MS profiling method was furthered refined to improve the separation of yeast histones. LC-MS is our most powerful approach for characterization of histones. It is able to detect all of the core histone isoforms simultaneously. In another words, it provides a bird’s-eye view of the global changes occurring on histones. Improving the separation of yeast histones allowed for the detection of more isoforms.

Chapter 3 describes the development of an approach to distinguishing between the isobaric histone modifications, acetylation and trimethylation. Acetylation and trimethylation are two common histone PTMs. These two modifications both have a mass of 42 Da and are not easily distinguished by mass spectrometry. They can be studied and
distinguished by use of antibodies. However, antibodies are limited by their availability, site specificity, and cross-reactivity. Some MS-based approaches have also been developed to study these two PTMs. However, those approaches require the use of high-resolution / high-mass accuracy mass spectrometers that are not available to many laboratories. Therefore, a method using low-mass accuracy mass spectrometers to distinguish acetylation and trimethylation was developed.

Chapter 4 describes the discovery of JmjC domain-containing demethylases. Before 2004, methylation was believed to be an irreversible modification. In 2004 the first demethylase, LSD1, was reported. LSD1 reverses dimethylation but not trimethylation. Shengjiang Tu proposed that trimethylation should also be reversible by JmjC domain-containing proteins. To test the hypothesis, sensitive MS-based methods were developed to assess the changes of histone methylation in yeast. Using the optimized protocols developed in Chapter 2 several JmjC domain-containing proteins were either knocked out or overexpressed in yeast to screen for demethylase activity. The MS-based screening indicated that several of the candidates possessed demethylase activity. Next, specific trimethylation sites on yeast histone H3 were identified as potential substrates for these enzymes.

Chapter 5 examines the potential of histone cross-talk between acetylation and methylation on yeast histones due to alterations in demethylase activities. It has been proposed that methylation and acetylation are synergistic. The work presented in Chapter 4 provides some clues that methylation and acetylation are closely linked. To obtain a better understanding of the relationship between methylation and acetylation, the JmjC
proteins were altered to probe synergistic changes in acetylation in response to changes in methylation. Our results show that H3 acetylation changes significantly in all strains except for ecm5Δ. LC-MS/MS results indicate that Gis1 and Jhd1 facilitate K56Ac, while Rph1 decreases the level of K56Ac. Taking into consideration that JmjC proteins are demethylases, this study implies that H3 methylation is correlated with acetylation on H2B1, H2B2, and H3.

Chapter 6 contains a summary of the research along with a discussion of potential future work.
CHAPTER 2

SEPARATION AND CHARACTERIZATION OF HISTONES FROM
SACCHROMYCES CEREVISIAE

2.1 INTRODUCTION

Yeast are an ideal model organism for the characterization of histones. Yeast have many beneficial properties, such as rapid growth, dispersed cells, and being amenable to genetic manipulation, that make it particularly suitable for biological studies. Of the various strains of yeast, budding yeast Sacchromyces cerevisiae (S. cerevisiae) and fission yeast Schizosacchromyces prombe are most widely studied. S. cerevisiae is the first eukaryote whose genome was completely sequenced. In addition to these advantages, yeast have the lowest copy number of histone genes of all eukaryotes. Among the 16 chromosomes in S. cerevisiae, six contain histone genes but only three encode the core histones. The relatively low number of histone genes allows for experiments that involve genetic manipulation of all histones.

As a consequence of the low copy number of histones and the small size of yeast genome, yeast histones are expressed at very low concentrations. Therefore, extraction procedures must be optimized to obtain a high yield of histones. Even with special care and optimized extraction procedures, yeast histone extracts often contain a high
background\textsuperscript{119, 120}. Therefore, sample clean-up and separation is critical for subsequent MS analysis. Gel electrophoresis and liquid chromatography are the most commonly used approaches to separating histones.

SDS-PAGE is widely used for protein separation. The Tris-glycine-SDS (Laemmli) system is the most commonly used. However, this Laemmli system is not effective for the separation of peptides and small proteins. Proteins and polypeptides below 20 kDa run with the SDS front and are only partly separated. Efforts have been made to modify the Laemmli system to extend the separation to proteins of low MWs. The resolving power of a polyacrylamide gel is primarily determined by the pore size of the gel; however, other factors, such as buffer composition, pH, and gel additives, also have an effect. By increasing the pH of a stacking gel from 8.9 to 9.2, the linear calibration range of a 10\% gel was extended to proteins of 10-20 kDa\textsuperscript{121}. By using tricine, instead of glycine, Schagger successfully separated proteins from 1 to 100 kDa\textsuperscript{122}. Glycine (pK\textsubscript{a} 9.6) migrates very slowly in the acidic stacking gel and it usually stacks proteins of large MWs. Tricine allows the stacking and destacking of small proteins to be achieved at the same pH as in the separating gel and at a low concentration of acrylamide. At the usual pH values of SDS gels (pH 6.8-8.8), tricine (pK\textsubscript{a} 8.15) has lower negative charge than glycine, allowing it to migrate faster than glycine in a stacking gel. As a result, small SDS-coated proteins are separated from SDS micelles and the stacking limit is shifted to proteins of low MWs. With tricine as the tailing ion, resolution of proteins of 5-20 kDa is increased greatly\textsuperscript{122}. Tris-tricine gel has proven useful not only for small molecules but also for mixtures of lipopolysaccharides and lipooligosaccharides\textsuperscript{123}. 

32
Liquid chromatography (LC) and LC coupled with mass spectrometry (LC-MS) have recently been implemented for effective separations of histones \textsuperscript{124}. LC separations of histones are performed using hydroxylapatite, size exclusion, hydrophilic interaction, or reversed-phase-based stationary phases \textsuperscript{106, 125-133}. Reversed-phase liquid chromatography (RPLC) is most applicable for coupling with ESI-MS because the mobile phase is compatible with ESI. By adjusting the LC gradient and flow rate, not only can impurities be removed but proteins of interest can be well separated. For LC, the most commonly used mobile phase is water and acetonitrile with trifluoroacetic acid (TFA) as the ion-pairing agent. One disadvantage of TFA is that it binds strongly with ionized amino acid side chains so that the ion pairs with analytes can not be broken apart at normal ESI conditions. The strongly bound adducts lead to signal suppression as follows: 1) reduced overall ionization efficiency of the TFA-analyte ion and 2) analyte signal is distributed across the multiple TFA adducts in the observed MS spectra \textsuperscript{134}. Thus, optimization has to be done to reduce the effect of TFA. A variety of weaker ion-pairing additives, such as formic and acetic acid, have been investigated \textsuperscript{135-138}. These reagents often lead to higher sensitivity but poor chromatographic performance. For example, the replacement of TFA with formic or acetic acid resulted in increased peak width at half height by 32-104% \textsuperscript{139, 140}. Because high performance separations are required for histones, TFA is still the preferred ion-pairing agent for RPLC ESI-MS \textsuperscript{134, 141}.

This chapter describes the development and optimization of methods for yeast histone preparation, separation, and characterization. To improve the quantity and purity
of histones obtained from yeast, the protocols for purification were modified to include extensive washing of nuclei prior to lysis and acid extraction. Because histones are low MWs (<25 kDa) and have similar chemical properties, their separation by SDS-PAGE and RPLC was also optimized. We determined that a constant voltage of 117 V provided the optimal separation of yeast histone H2A, H2B, H3, and H4 on 16.5% Tris-tricine-SDS gels. Using optimized mobile phase conditions, histones were also well separated on a reversed-phase C\textsubscript{18} column with minimal interference from mobile phase adducts.

2.2 EXPERIMENTAL

2.2.1 Bovine Thymus Histone Extraction

Bovine calf thymus was purchased from Worthington Biochemical Corp (Lakewood, NJ, USA). Histones were purified using a standard acid-extraction procedure\textsuperscript{142} (Appendix C.8). In brief, frozen calf thymus was thawed in buffer C (0.25 M sucrose, 10 mM MgCl\textsubscript{2}, 25 mM KCl, 50 mM Tris-HCl, pH 7.5, and 50 mM NaHSO\textsubscript{3}) and homogenized by using a Plytron homogenizer. The homogenized tissue was then centrifuged at 4500 rpm for 5 min. The nuclear pellet was resuspended twice in 10 volumes of buffer C, followed by five washings with 0.14 M NaCl containing 50 mM NaHSO\textsubscript{3}. Nuclei were suspended in 0.25 M HCl and shaken for 2 h at 4 °C. After centrifugation at 13400 rpm for 10 min and dialysis of the supernatant against 300 volumes of 0.025 M HCl for 12 h, histones were precipitated with 10 volume of acetone at -20 °C overnight. The resulting bovine histones were speed-vacuum dried and dissolved in HPLC water at a concentration of 1 µg/µl.
2.2.2 Yeast Histone Extraction

Histones were purified as reported previously\textsuperscript{143} (Appendix C.7). In brief, cells were purified by sequential resuspension in buffer A (50 mM Tris-HCl, pH 7.5, 30 mM dithiothreitol), buffer S (20 mM HEPES, pH 7.4, 1.2 M Sorbitol), and buffer S with 10 mg/ml Zymolyase, followed by a 60-min incubation at 30°C. Ice-cold buffer B (20 mM PIPES, pH 6.8, 1.2 M Sorbitol, 1 mM MgCl\textsubscript{2}) was then added and the cells were collected by centrifuging at 5500 g for 10 min. The cells were further purified by three successive resuspensions in ice-cold freshly made NIB buffer (0.25 M sucrose, 60 mM KCl, 15 mM NaCl, 5 mM MgCl\textsubscript{2}, 1 mM CaCl\textsubscript{2}, 15 mM MES, pH 6.6, 0.8% Triton X-100, 1 mM PMSF, 1 mM NaF), three successive resuspensions in wash buffer A (10 mM Tris, pH 8.0, 0.5% NP-40, 75 mM NaCl, 30 mM sodium butyrate, 1 mM PMSF, 1 mM NaF) and two successive resuspension in wash buffer B (10 mM Tris, pH 8.0, 0.4 M NaCl, 30 mM sodium butyrate, 1 mM PMSF, 1 mM NaF). The pellet was resuspended in 0.2 M H\textsubscript{2}SO\textsubscript{4} and spun down at 5500 g for 10 min. 100% TCA was added to the supernatant to a final concentration of 20%. After overnight precipitation, histones were collected by centrifuging at 5500 g for 30 min. The pellet was washed in 500 μl of ice-cold acetone with 0.1% HCl and then in 500 μl of pure acetone. The obtained histones were air dried and stored at -20 °C.

2.2.3 Liquid Chromatography-Mass Spectrometry

Histones were characterized by LC-MS performed by use of a Micromass LCT (Micromass, Wythenshawe, UK) coupled with a Waters 2690 LC system (Waters,
Milford, MA). The LCT instrument was used in positive ion mode at the following optimal conditions: capillary voltage = 2500 V, source temperature = 100 ºC, resolution = 4000, and sample cone voltage = 40 V. Spectra were collected in continuum mode on LCT at a rate of 1.00 scan/sec. The separation was performed at a flow rate of 50 µl/min using the gradient provided in Table 2.1. 0.1% TFA in water and acetonitrile were used as mobile phases A and B, respectively.

LC-MS experiments were also performed using a Micromass Q-TOF mass spectrometer (Micromass Q-Tof-2, Micromass, Wythenshawe, UK) coupled with a Dionex Ultimate 3000 LC system (Dionex, Sunnyvale, CA). The Q-TOF was also operated in positive ion mode at the following conditions: source temperature = 100 ºC, resolution = 8000 and cone voltage = 21-50 V. Spectra were collected in continuum mode at a rate of 2.0 scan/sec. The separation was performed at a flow rate of 25 µl/min using the gradient provided in Table 2.2. 0.05% TFA in water and acetonitrile were used as mobile phases A and B, respectively.

NaI was used for external mass calibration over the mass range of 500-2500 m/z for LCT and Q-TOF. Spectra obtained on both instruments were analyzed by use of MassLynx 4.0 data analysis software. 10 µg of samples were injected on both instrument and separated on a Discovery Bio wide pore C\textsubscript{18} column (150 × 1 mm, 5 µm, 300 Å, Supelco, Park Bellefonte, PA). TFA was obtained from Sigma Aldrich (St. Louis, MO, USA). All solvents used were HPLC grade.
Table 2.1 Reversed-phase gradient used for LC-MS (LCT).

<table>
<thead>
<tr>
<th>Time/min</th>
<th>0</th>
<th>2</th>
<th>22</th>
<th>26</th>
<th>28</th>
<th>30</th>
<th>31</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>%B</td>
<td>30</td>
<td>45</td>
<td>60</td>
<td>60</td>
<td>95</td>
<td>95</td>
<td>30</td>
<td>30</td>
</tr>
</tbody>
</table>

Table 2.2 Reversed-phase gradient used for LC-MS (Q-TOF).

<table>
<thead>
<tr>
<th>Time/min</th>
<th>0</th>
<th>2</th>
<th>10</th>
<th>30</th>
<th>35</th>
<th>36</th>
<th>40</th>
<th>42</th>
<th>80</th>
</tr>
</thead>
<tbody>
<tr>
<td>%B</td>
<td>20</td>
<td>30</td>
<td>35</td>
<td>50</td>
<td>60</td>
<td>95</td>
<td>95</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>

2.2.4 *Sodium Dodecyl Sulfate–Polyacrylamide Gel Electrophoresis (SDS-PAGE)*

*Separation of Histones*

SDS-PAGE separation of histones was performed by use of precast 16.5% Tris-tricine gels (3.3% C, Bio-Rad, Hercules, CA) with Tris/Tricine/SDS (Bio-Rad, Hercules, CA) running buffer. The yeast histones were separated following the detailed protocol provided in Appendix C.4. In brief, 980 µl of sample buffer were mixed with 20 µl of β-mercaptoethanol (βME). The resulting sample buffer was mixed with yeast histones at a ratio of 4:1. The sample mixture was incubated in boiling water for 10 min and then spun down at 14 krpm for 15 min. After the loading wells were thoroughly washed with the running buffer, histone samples were loaded on the gel. 14 µl of Precision Plus Protein Standards (Bio-Rad, Hercules, CA) were used as the MW marker. Gels were run at a constant voltage of 100-250 V for 3-4.5 h. The optimal voltage for histone separations was determined to be 117 V.
After the separation was complete, SDS was thoroughly washed from the gel twice by microwaving the gel in 100 ml of distilled water for 90. The gel was then shaken in 100 ml of fresh distilled water on an orbital shaker for 5 min. To stain the histones, the gel was microwaved in 50 ml of Bio-Safe Coomassie (Bio-Rad, Hercules, CA) for 90 seconds and then shaken for 15 min or until gel bands were visible. The gel was destained by incubating in 100 ml of distilled water with constant shaking.

2.2.5 Liquid Chromatography-Tandem Mass Spectrometry

After the yeast histones were separated by SDS-PAGE, the gel bands were in-gel digested with trypsin as described in Chapter 3. The tryptic digest was subject to LC-MS/MS for identification by use of LTQ FT-ICR mass spectrometer (Thermo Fisher, San Jose, CA) and LCQ Deca XP+ ion trap mass spectrometer (Thermo Fisher). The working conditions are as described in Chapter 3.

2.3 RESULTS AND DISCUSSION

2.3.1 Yeast Histone Extraction

As mentioned previously, special care needs to be taken during yeast histone extraction to obtain a high recovery and relatively pure proteins. A number of protocols have been developed to extract yeast histones from whole cells and nuclei \(^{118, 120, 144-146}\). Histones extracted from whole cells typically contain more non-histone proteins than when extracted from nuclei \(^{118}\). Even histones extracted from yeast nuclei contain larger amounts of impurities than other eukaryotes \(^{118}\). The protocol used in this thesis was
modified from a bulk histone purification protocol developed by Lo \(^{143}\). As seen in **Figure 2.1**, the yeast histone extract contained a high abundance of histones \(^{118,120}\). The modified bulk histone purification protocol yielded about 80 times higher purity histones than was obtained with chaotropic guanidine-based protocols \(^{147}\). The primary difference in the two approaches was that the bulk histone extraction involves creation of spheroplasts (removal of cell wall) prior to cell lysis. The bulk histone purification is believed to not only yield higher amounts of histones but also preserves histone acetylation during sample preparation \(^{118}\). Our modified bulk histone extraction procedure was used for the majority of histone preparations due to its high yield.

During the course of optimizing the histone purifications, several observations were made regarding the experimental factors that affect data quality.

(1) Histone H3 is labile to rapid proteolysis. Spontaneous breakdown of H3 is commonly seen in the histone preparation from a number of organisms, including yeast \(^{148-151}\). We observed this degradation in some of our yeast samples. It results from lack of HDAC inhibitor in the NIB buffer and/or a low concentration of butyrate in subsequent buffers. Protease inhibitors slow the degradation but have less impact than HDAC inhibitors.

(2) The cell concentration at time of harvest must not exceed \(OD_{600} = 1.3\). If the cells grow to a higher concentration, a large amount of impurities will be extracted along with the histones (**Figure A.1**).
Figure 2.1 An LC trace of yeast histones obtained on the Q-TOF mass spectrometer. The LC gradient is provided in Table 2.2.

2.3.2 Yeast Histone Variants

During embryogenesis or the maturation of specialized cells \(^{152}\), histone variants with small amino acid sequence differences develop. Since the first revelation of histone variants in 1969 \(^{153}\), their roles in gene expression, DNA repair, and meiotic events have received great interest \(^{154-156}\). A full listing of known histones for different organisms can be found in the NCBI histone database \(^{157}\). Human histones are known to include 14 variants of H1, 36 of H2A, 30 of H2B, 20 of H3, and 6 of H4 \(^{158}\). In contrast, yeast histones have 2 variants of H2A, 2 of H2B, and one of H4 and H3 (Table 2.3). It is
noteworthy to point out that histone H1 is not listed in the Table 2.3 because its presence in yeast is debated. Some reports claim that a yeast H1 protein exists and others refer to so-called H1 variants in yeast as H1-like proteins. The notion that yeast lack a true H1 is rationalized due to an unusually short nucleosomal repeat length.

Furthermore, no unanimity has been reached on the nomenclature of yeast histone variants. The names vary across many sources. Histone H2A coded by the \( HTZ1 \) gene is referred to as H2AZ, H2A.Z, H2AFZ, and Htz1. Histone H3 coded by the \( HHT1/HHT2 \) genes resemble H3.3 in mammalian cells and is commonly referred to as H3.3. This thesis will adopt the use of common names when referring to yeast histones.

Yeast possess fewer number of histone variants than any other organism. The low number of histone genes is ideal for performing genetic manipulation of histones that would be difficult with other eukaryotes. Histone variants are usually highly conserved and differ by only a few amino acids. Often these sequence variations confound the characterization of histones. For example, yeast H2B has two variants, H2B1 and H2B2 that differ by 4 residues. These amino acid difference lead to slight difference in hydrophobicity and pI. The H2A variants, H2A.1 and H2A.2, differ by two amino acids. The amino acids at 124 and 125 are transposed yielding proteins with the same molecular weight and the same physical properties. H2A-1 shares the same sequence as H2A.1 fragment 15-132. H3 has three histone variants. Two of the variants are coded by the same gene \( HHT2 \). The H3.3 variants are translated from genes \( HHT1 \) and \( HHT2 \) with identical sequences and share the same sequence as the other H3 variant coded by \( HHT2 \).
except for two additional amino acids (135 and 136). Only the H3.3 variant has been observed and studied by MS. Histone H4 is unique in yeast in that it has no variants. It is coded by two genes, *HHF1* and *HHF2*, which code for the same protein sequence.

Histone variants are often expressed at different levels *in vivo*. MS approaches do not observe all the different variants due to limitations in the mass spectrometer’s dynamic range of detection. For example, H2B1 and H2B2 are both easily observed with H2B2 at higher abundance (Figure 2.1). The much lower abundant H2AZ variant is extremely challenging to detect. Typically 1 mg of histones is extracted from 1 liter of yeast cells whereas 12 liters of yeast cells are needed to affinity purify just 5-10 µg of FLAG tagged H2AZ\(^{147}\). Even though H2A is easily observed by LC-MS, the lower abundant H2AZ and H2A-1 fall below the range of detection\(^{173,174}\).
Table 2.3 Yeast histone variants. (a - data not available; RH - relative hydrophobicity, calculated by using software SSRCalc, version 3.2, choosing the condition of 300 Å column with TFA; SL – sequence length measured by the number of amino acids.)

<table>
<thead>
<tr>
<th>Histone</th>
<th>Variant</th>
<th>Accession #</th>
<th>Gene code</th>
<th>MW</th>
<th>RH</th>
<th>pI</th>
<th>SL</th>
</tr>
</thead>
<tbody>
<tr>
<td>H2B</td>
<td>H2B1</td>
<td>NP_010510</td>
<td>HTB1</td>
<td>14121</td>
<td>52.94</td>
<td>10.10</td>
<td>131</td>
</tr>
<tr>
<td></td>
<td>H2B2</td>
<td>NP_009553</td>
<td>HTB2</td>
<td>14106</td>
<td>53.45</td>
<td>10.07</td>
<td>131</td>
</tr>
<tr>
<td>H4</td>
<td>H4</td>
<td>NP_009563</td>
<td>HHF1</td>
<td>11237</td>
<td>60.43</td>
<td>11.36</td>
<td>103</td>
</tr>
<tr>
<td></td>
<td>H4</td>
<td>NP_014368</td>
<td>HHF2</td>
<td>11237</td>
<td>60.43</td>
<td>11.36</td>
<td>103</td>
</tr>
<tr>
<td>H2A</td>
<td>H2AZ</td>
<td>NP_014631</td>
<td>HTZ1</td>
<td>14152</td>
<td>65.43</td>
<td>10.65</td>
<td>134</td>
</tr>
<tr>
<td></td>
<td>H2A-1</td>
<td>AAA66318</td>
<td>--</td>
<td>12801</td>
<td>72.03</td>
<td>10.57</td>
<td>118</td>
</tr>
<tr>
<td></td>
<td>H2A.1</td>
<td>NP_010511</td>
<td>HTA1</td>
<td>13858</td>
<td>73.70</td>
<td>10.67</td>
<td>132</td>
</tr>
<tr>
<td></td>
<td>H2A.2</td>
<td>NP_009552</td>
<td>HTA2</td>
<td>13858</td>
<td>73.70</td>
<td>10.67</td>
<td>132</td>
</tr>
<tr>
<td></td>
<td>--a</td>
<td>AAS64349</td>
<td>HHT2</td>
<td>14982</td>
<td>54.60</td>
<td>11.33</td>
<td>134</td>
</tr>
<tr>
<td>H3</td>
<td>H3.3</td>
<td>NP_009564</td>
<td>HHT1</td>
<td>15225</td>
<td>56.51</td>
<td>11.43</td>
<td>136</td>
</tr>
<tr>
<td></td>
<td>H3.3</td>
<td>NP_014367</td>
<td>HHT2</td>
<td>15225</td>
<td>56.51</td>
<td>11.43</td>
<td>136</td>
</tr>
<tr>
<td></td>
<td>Cse4p</td>
<td>AAB60309</td>
<td>CSE4</td>
<td>26710</td>
<td>75.14</td>
<td>9.30</td>
<td>229</td>
</tr>
</tbody>
</table>

Table 2.4 Sequence variations in yeast histone H2B1 and H2B2.

<table>
<thead>
<tr>
<th>Residue</th>
<th>2</th>
<th>3</th>
<th>27</th>
<th>35</th>
</tr>
</thead>
<tbody>
<tr>
<td>H2B1</td>
<td>A</td>
<td>K</td>
<td>T</td>
<td>A</td>
</tr>
<tr>
<td>H2B2</td>
<td>S</td>
<td>A</td>
<td>V</td>
<td>V</td>
</tr>
</tbody>
</table>

As stated previously, yeast histones have proved to be an excellent model organism for the study of histones. In particular, the low number of sequence variants greatly simplifies the characterization of histone PTMs such as acetylation. All of the core histones, except for H2A, have a large number of sites that are reported to be
acetylated (Table 2.5). Studies show that 63% of yeast H4 exists in di-, tri-, and tetra-acetylated forms and that the monoacetylated form of H4, H2A, and H2B is exclusively N-terminal. The acetylation of these histones is not random but highly specific. Following acetylation at the N-terminus, H2A is next acetylated on K7 followed by K4. As to the acetylation of H2B, it was observed that monoacetylation (Ac1) of H2B occurred exclusively at the N-terminus and that the second acetylation was split 3:2 between K22 and K16. Upon complete acetylation of K16, the third acetylation was split 3:2 between K22 and K16 and the fourth acetylation resulted in 100% modification of K22 and K16. The fifth and sixth acetylation sites have not been reported yet due to the low abundance of penta- and hexa-acetylation.

The directionality of yeast H4 acetylation is still debated. Some researchers believe that acetylation of yeast histone H4 occurs in a stepwise fashion referred to as the zipper model. In this model, acetylation proceeds in the following order: K16 (first), K12, K8, and K5 (last). However, studies have shown that yeast H4 acetylation does not follow the C-terminal to N-terminal zipper strictly. It has been shown that monoacetylated H4 (H4Ac1) occurred predominantly on K16 but also contained minor forms of acetylated K12. The second and third acetylations were split between K12 and K8 and the fourth acetylation occurred on K5. As to the order of sequential acetylation on yeast H3, no results have been reported, though studies on human H3 have been reported.

Acetylation of yeast histones is not only selective but occurs at different rates for given sites on each histone. It was determined that the average acetylation half-lives were
approximately 15 min for bulk H4, 10 min for H3, 4 min for H2B, and 5 min for H2A. In addition, acetylation half-lives are affected by growth media. For example, the half-life of H4 acetylation is 20.6 min in the complete medium (a blend of yeast extract, peptone, and dextrose - YPD), and 13.6 min in synthetic defined medium. The turnover rates for acetylation of H3 are the same across the acetylated forms. In contrast, the mono-, and di-acetylated H4 and H2B have lower turnover rates than multi-acetylated forms of these histones.

**Table 2.5** Yeast histone acetylation sites. Ac - acetylation.

<table>
<thead>
<tr>
<th>Histone</th>
<th>Ac state</th>
<th>Ac site</th>
</tr>
</thead>
<tbody>
<tr>
<td>H4</td>
<td>0, 1, 2, 3, 4, 5</td>
<td>N-terminal, K5, K8, K12, K16</td>
</tr>
<tr>
<td>H3</td>
<td>0, 1, 2, 3, 4, 5, 6</td>
<td>K9, K14, K18, K23, K27, K56</td>
</tr>
<tr>
<td>H2A</td>
<td>0, 1, 2, 3</td>
<td>N-terminal, K4, K7</td>
</tr>
<tr>
<td>H2B</td>
<td>0, 1, 2, 3, 4, 5, 6, 7</td>
<td>N-terminal, K6, K11, K16, K17, K21, K22</td>
</tr>
</tbody>
</table>

### 2.3.3 Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

**Separation of Yeast Histones**

SDS-PAGE is a powerful tool for separation and characterization of proteins. With the ready availability of precast gels one can rapidly characterize protein mixtures. SDS binds proteins at a ratio of ~1.4 g SDS per gram of proteins and neutralizes the protein’s charge distribution. This nearly constant ratio of SDS to protein provides a uniform mass-to-charge ratio for most proteins. Under conditions of uniform charge,
SDS-PAGE separates proteins according to their molecular size. The smaller the protein, the faster the protein runs in the gel. The total amount of acrylamide (%T) and cross-linker (%C) determines the pore size of a gel and thus determines the separation capacity of the gel. The calculation of %T and %C is as follows:

\[
%T = \frac{\text{Acrylamide (grams)} + \text{Bisacrylamide (grams)}}{100 \text{ ml}} \times 100 \quad \text{eqn 2.1}
\]

\[
%C = \frac{\text{Bisacrylamide (grams)}}{\text{Acrylamide (grams)} + \text{Bisacrylamide (grams)}} \times 100 \quad \text{eqn 2.2}
\]

The pore size is most commonly controlled by altering %T and holding %C constant. Higher %T results in smaller pores and better resolution but significantly longer running time. SDS gels usually contain a combination of a stacking gel and a resolving gel to improve protein separation. The stacking gel is a large pore polyacrylamide gel used to concentrate proteins in a tight band onto the resolving gel. The proteins are then separated as they migrate through the resolving gel according to their MWs. By varying %T of the resolving gel, proteins with different ranges of MW can be separated. For example, a 5% Tris-HCl gel is good for the separation of proteins with MW of 75–250 kDa while an 18% gel is ideal for proteins of 10-250 kDa.

Yeast histones are low MW proteins (10-26 kDa) with variants that have similar MWs. Therefore, it is crucial to optimize the SDS-PAGE conditions to achieve the best separation. As stated previously, Tris-glycine-SDS gels have limited resolving power for
proteins below 20 kDa. Tris-tricine-SDS gels are significantly better at resolving proteins below 20 kDa. In this study, both Tris-glycine and Tris-tricine gels at %Ts ranging from 12-18% were evaluated. It was determined that 16.5%T Tris-tricine gels (8.6 × 6.8 cm, 4% stacking gel) were optimal for yeast histone separation. Interestingly, the Tris-tricine gel itself does not contain SDS, which could impact the gel’s resolving power. However, SDS was used during sample preparation and also present in the running buffer. Separation was further improved by the addition of beta mercaptoethanol (βME) to reduce the proteins and by denaturing the proteins with heat prior to loading the gel. It is important to note why higher %T was not used in this study. When we performed separations on higher %T gels better resolution of the histone isoforms were achieved. However, %T > 17% resulted in a poor yield after in-gel trypsin digestion. The rationale for the low yield is that the small pore size prevents the migration of trypsin into the gel. 16.5% was determined to be the ideal compromise between high resolution and compatibility with in-gel digestion.

The optimal yeast histone separation on 16.5% Tris-tricine-SDS gel is shown in Figure 2.2. The MWs of yeast histones decrease from H3 (MW = 15225 Da), H2B (MW_{H2B1} = 14121 Da, and MW_{H2B2} = 14106 Da), H2A (MW = 13885 Da), to H4 (MW = 11237 Da). Thus, the four distinct bands in Figure 2.2 are expected to be H3, H2B, H2A, and H4 from the top to the bottom. The identities of the bands were confirmed by LC-MS/MS (Table 2.6). Due to the limited resolving power of the gel and the small MW difference between H3 and H2B, a minor amount of H2B migrated into the band of H3. It is interesting to notice that the 15 kDa protein standard migrated to about the same level.
as H2A. This observation agrees with the report that the mobilities of small proteins in SDS-PAGE may no longer be proportional to their MWs.\textsuperscript{141}

![Image of yeast histone separation on a 16.5% Tris-tricine-SDS gel](image)

**Figure 2.2** Yeast histone separation on a 16.5% Tris-tricine-SDS gel.

### 2.3.4 Identification of Yeast Histone Separation Pattern from Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

To confirm our prediction of the gel band identities in Figure 2.2, each band was excised and tryptic digested individually. The digest was subject to LC-MS/MS by use of both LTQ FT-ICR and LCQ Deca XP+ ion trap mass spectrometers. The identity of each gel band obtained from both mass spectrometers is consistent. Since LTQ FT-ICR is a high resolution / high mass accuracy mass spectrometer, it provided better sequence
coverage for each protein than the ion trap mass spectrometer. Table 2.6 provides the results obtained from the LTQ FT-ICR mass spectrometer.

Table 2.6 LC-MS/MS identification of yeast histone separation pattern on a SDS-PAGE gel. SC – sequence coverage; PS - protein score.

<table>
<thead>
<tr>
<th>Histone</th>
<th>SC</th>
<th>PS</th>
<th>Peptide detected</th>
<th>Hit</th>
</tr>
</thead>
<tbody>
<tr>
<td>H4</td>
<td>87</td>
<td>741</td>
<td>6GGKGLGK\textsubscript{Ac}GGAK\textsubscript{Ac}CR\textsuperscript{17}</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>20KLRDIQGITKPAIR\textsuperscript{35}</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>21ILRDIQGITKPAIR\textsuperscript{35}</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>24DIQGITKPAIR\textsuperscript{35}</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>24DIQGITKPAIR\textsuperscript{36}</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>46ISGLIYEEVR\textsuperscript{55}</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>60SFLESVIR\textsuperscript{67}</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>68DSVTYTEHAK\textsuperscript{77}</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>79KVTSDLVYALK\textsuperscript{91}</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>79KVTSDLVYALKR\textsuperscript{92}</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>80TVTSMDVYALK\textsuperscript{91}</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>80TVTSMDVYALKR\textsuperscript{92}</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>93QGRTLYGFGG\textsuperscript{102}</td>
<td>1</td>
</tr>
<tr>
<td>H2A</td>
<td>62</td>
<td>672</td>
<td>8AGSAAKASQRSAKAGLTFPVGR\textsuperscript{30}</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>19SAKAGLTFPVGR\textsuperscript{30}</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>32AGLTFPVGR\textsuperscript{30}</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>77TRIIAR\textsuperscript{82}</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>83HLQAI\textsuperscript{89}</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>83HLQAIRNDDELNK\textsuperscript{96}</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>90NDDELNK\textsuperscript{96}</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>90NDDELNKLLGNVTIAQGGVLPNIHQNNLLPK\textsuperscript{119}</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>97LLGNVTIAQGGVLPNIHQNNLLPK\textsuperscript{119}</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>97LLGNVTIAQGGVLPNIHQNNLLPKK\textsuperscript{120}</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>124ATK\textsubscript{Ac}ASQEL\textsuperscript{131}</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 2.6 Continued.

<table>
<thead>
<tr>
<th>Histone</th>
<th>SC</th>
<th>PS</th>
<th>Peptide detected</th>
<th>Hit</th>
</tr>
</thead>
<tbody>
<tr>
<td>H2B</td>
<td>100</td>
<td>2076</td>
<td>$^1$SAKA$\text{K}<em>{\text{Ac}}$KPASK$</em>{\text{Ac}}$APAEK$^{16}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$^4$AEKKPASK$^{11}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$^4$AEK$_{\text{Ac}}$KPASK$^{11}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$^4$AEK$<em>{\text{Ac}}$KPASK$</em>{\text{Ac}}$APAEK$^{16}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$^{12}$APAEEK$\text{K}_{\text{Ac}}$PAAK$^{21}$</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$^{12}$APAEEK$<em>{\text{Ac}}$K$</em>{\text{Ac}}$APAEK$^{21}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$^{12}$APAEEK$\text{K}<em>{\text{Ac}}$K$</em>{\text{Ac}}$K$^{22}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$^{12}$APAEEK$\text{K}<em>{\text{Ac}}$K$</em>{\text{Ac}}$TSTSTDGK$^{30}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$^{17}$K$_{\text{Ac}}$PAAK$^{21}$</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$^{22}$K$_{\text{Ac}}$TSTSTDGK$^{30}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$^{22}$KTSTSTDGK$^{31}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$^{22}$KTSTSTDGKK$^{32}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$^{23}$TSTSTDGK$^{30}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$^{35}$ARKETYSSYIYK$^{46}$</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$^{35}$AR$_{\text{Me}_2}$KETYSSYIYK$^{46}$</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$^{37}$KETYSSYIYK$^{46}$</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$^{37}$KETYSSYIYKVKK$^{49}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$^{37}$KETYSSYIYKVKKQTHPDTGISQK$^{60}$</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$^{38}$ETYSSYIYK$^{46}$</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$^{38}$ETYSSYIYKVKK$^{49}$</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$^{47}$VLKQTHPDTGISQK$^{60}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$^{50}$QTHPDTGISQK$^{60}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$^{61}$SMSILNSFVNDIFER$^{75}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$^{76}$IATEASK$^{82}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$^{76}$IATEASKLAAYNK$^{88}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$^{76}$IATEASKLAAYNK$^{89}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$^{76}$IATEASKLAAYNKSTISAR$^{95}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$^{83}$LAAYNK$^{88}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$^{83}$LAAYNK$^{89}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$^{96}$STISAREIQTAVR$^{102}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$^{96}$EIQTAVRLILPGELAK$^{111}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$^{103}$LILPGELAK$^{111}$</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$^{103}$LILPGELAKHAVSEGTR$^{119}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$^{112}$HAVSEGTRA$^{130}$KTYSSSTQA</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$^{120}$AVTKYSSSTQA</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 2.6 Continued.

<table>
<thead>
<tr>
<th>Histone</th>
<th>SC</th>
<th>PS</th>
<th>Peptide detected</th>
<th>Hit</th>
</tr>
</thead>
<tbody>
<tr>
<td>H3</td>
<td>76</td>
<td>1850</td>
<td>1ARTK$_{Me}$QTAR$^8$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2TKQTAR$^8$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3TK$_{Me}$QTAR$^8$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3TK$_{Me2}$QTAR$^8$</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3TK$_{Me3}$QTAR$^8$</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>9KSTGGK$^14$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>9KSTGGK$_{Ac}$APR$^{17}$</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>9KSTGGK$_{Ac}$APR$^{17}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>9K$<em>{Ac}$STGGK$</em>{Ac}$APR$^{17}$</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10STGGKAPR$^{17}$</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10STGGK$_{Ac}$APR$^{17}$</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>18K$<em>{Ac}$QLASK$</em>{Ac}$AAR$^{26}$</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>19QLASK$^{23}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>27KSAPSTGGVK$^{36}$</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>27KSAPSTGGVK$_{Me}$KPHR$^{40}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>27KSAPSTGGVK$_{Me2}$KPHR$^{40}$</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>27KSAPSTGGVK$_{Me3}$KPHR$^{40}$</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>27K$<em>{Ac}$SAPSTGGVK$</em>{Me}$KPHR$^{40}$</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>27KSAPSTGGVK$_{Me2}$KPHR$^{40}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>27KSAPSTGGVK$_{Me3}$KPHR$^{40}$</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>28SAPSTGGVK$_{Me2}$KPHR$^{40}$</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>28SAPSTGGVK$_{Me3}$KPHR$^{40}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>41YKPGTVALR$^{49}$</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>41YKPGTVALREIR$^{52}$</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>41YKPGTVALREIRR$^{53}$</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>53RFQKSTELLIR$^{63}$</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>53RFQK$_{Ac}$STELLIR$^{63}$</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>57STELLIR$^{63}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>57STELIRK$^{64}$</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>57STELIRKLPFQR$^{69}$</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>64KLPFQR$^{69}$</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>64KLPFQLVRR$^{72}$</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>65LPFQR$^{69}$</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>70LVREIAQDFKTDLR$^{83}$</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>73EIAQDFKTDLR$^{83}$</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>73EIAQDFK$_{Me}$TDLR$^{83}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>73EIAQDFK$_{Me2}$TDLR$^{83}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>73EIAQDFK$_{Me3}$TDLR$^{83}$</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>116RVTIQKK$^{122}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>116RVTIQK$^{121}$</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>117VTIQKK$^{122}$</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>125DIKLAR$^{128}$</td>
<td>2</td>
</tr>
</tbody>
</table>
2.3.5 Optimization of Liquid Chromatography-Mass Spectrometry Histone Profiling

The Micromass LCT and Q-TOF are both time-of-flight mass spectrometers with electrospray ionization sources made by the same manufacturer. The Q-TOF provides higher resolving power than the LCT and is ideal for separation and characterization of complex mixtures. Therefore, we performed optimization of LC-MS separations of standard bovine histone mixtures on the Q-TOF. To find the ideal conditions we optimized TFA concentration, flow rate, and cone voltage.

Being an anionic ion-pairing agent, in acidic mobile phases TFA binds cationized amino acid side chains. These adducts result in poor sensitivity and confound data analysis. We noted that TFA has different binding affinities with different histone variants. H4 and H2A appear to contain a larger number of TFA adducts than the other histones (Figure 2.3). To minimize TFA adducts, we optimized the concentration of TFA in mobile phases and the ESI cone voltage. My studies showed that at a flow rate of 25 µl/min a TFA concentration no lower than 0.05% was required for a good LC separation. Once the minimum TFA concentration was determined the cone voltage was varied from 21 V to 50 V. The cone voltage effects the translational and internal energy imparted upon the ions during desolvation. Higher cone voltages result in high-energy collisions with background neutrals. These collisions drive off TFA adducts from the gas-phase ions resulting in improved sensitivity. However, higher cone voltages also result in increased internal energy of the ion and eventually lead to ion fragmentation. It was observed that histones were easily fragmented when the cone voltage reached 30 V.
Figure 2.3 LC-MS spectra of bovine histones from a Q-TOF mass spectrometer with a cone voltage of 21 V.
Higher cone voltages resulted in nearly complete fragmentation of the histones (Figure A.3). Persistent TFA adducts were even observed at cone voltages as high as 50 V (Figure A.4). A flow rate of 25 µl/min, 0.05% TFA, and a cone voltage of 25 V provided the histone separation (Figure 2.4) with a minimum of TFA adducts (Figure A.5).

Figure 2.4 An LC trace of bovine histones from a Q-TOF mass spectrometer.

At the optimized conditions, namely 0.05% TFA and a cone voltage of 25 V, 10 µg of yeast histones were analyzed by LC-MS on the Q-TOF. The separation profile for
yeast histones is shown in Figure 2.1. The relative hydrophobicity (See Table 2.3) of H2B1, H2B2, H4, and H2A correlated with the observed order of elution with the exception of H3. The relative hydrophobicity of H3 falls between H2B and H4, so it was expected to elute in the third peak. However, it eluted last. This unexpected behavior may be due to the highly modified nature of histone H3. The effects of histone PTMs on relative retention have not been predicted and thus could account for H3’s uncharacteristic elution\textsuperscript{178}.

The MS spectra of each eluted histone in Figure 2.1 are shown in Figure 2.5. The predominant species in H2B1, H2B2, H2A, and H4 are the N-terminal acetylated form. Their polyacetylated forms were at much lower abundances. Due to the low abundances, some of the polyacetylated forms were not detected on the Q-TOF. Of the four core histones, H3 is the only one with a free N-terminal. The abundance of the H3 acetylated forms is in contrast to the other histones. Instead of the mono-acetylated, the di-acetylated is the most abundant form. The polyacetylated forms were at lower abundances. With the degradation of H3, the residue H3 (fragment 1) showed a somewhat different pattern from the intact H3. In fragment 1, unmodified H3 was barely seen; the monoacetylated form is dominant; the abundances of the polyacetylated forms decrease sequentially. The identity of H3 fragment 2 has not been determined yet. Across all of the H3 related spectra it is seen that the peaks are 14 Da apart, which may have resulted from H3 methylation\textsuperscript{179}.  

55
Figure 2.5 LC-MS spectra of wide type yeast histones from a Q-TOF mass spectrometer.
2.3.6 Optimization of Liquid Chromatography-Mass Spectrometry for Separation of Yeast Histones Isoforms

As stated previously, the Micromass Q-TOF mass spectrometer provides higher resolution than the Micromass LCT instrument. However, the former possesses lower sensitivity than the latter due to the additional collision cell. For this reason, the Q-TOF is ideal for the analysis of complex samples, such as crude histone extracts. For simple samples that require high sensitivity, the LCT provides better results. Thus, we performed optimization of LC-MS separations of recombinant yeast histones on the LCT. To find the ideal conditions the cone voltage and LC gradient were optimized. Like on the Q-TOF, adducts were observed when TFA was used as the ion-pairing reagent. We optimized our LC-MS conditions to minimize the effect of TFA. The optimal conditions for separation were determined to be a TFA concentration = 0.1% and a flow rate = 50 µl/min. We also observed that a cone voltage = 40 V was optimal for reducing TFA adducts without inducing protein fragmentation (Figure 2.6 and 2.7).

HPLC separation of histones is critical to obtain the best sensitivity and detection limit possible on a mass spectrometer. Often times, a high-resolution histone separation is hard to achieve due to the sample complexity and the similar physical properties of histone isoforms. Therefore, the working conditions were optimized to maximize the chromatographic separation for the specific histone isoforms of interest. For this section the LC separation was optimized to fully resolve recombinant yeast histone H3 and its K56 acetylated isoform. First, the yeast recombinant histones were dialyzed to remove any impurities before separation. Second, the column was thoroughly washed and
equilibrated at the initial gradient (30% B) for at least 30 min. Third, the LC gradient was optimized to achieve the best separation of the two isoforms of interest. As shown in Figure 2.8 the yeast histones H3 and H4 were fully resolved into two sharp peaks with flat baselines. More importantly, the unmodified yeast histone H3 (MW = 15225 Da) was separated from its K56 acetylated form (MW = 15225 + 42 = 15276 Da). Reversed-phase LC separation of unmodified and acetylated histone isoform is a remarkable accomplishment and has not been previously reported.

The high-resolution separation of histone acetylation isoforms was used to demonstrate that H3 K56 is a substrate for a novel histone acetyltransferase, Rtt109. Rtt109 and Asf1 were incubated with recombinant (H3/H4)$_2$ tetramers. The H3 acetylated isoform is clearly visible in the sample containing acetyl coenzyme A. To verify the site of acetylation, the samples were in-solution digested with trypsin and analyzed by LC-MS/MS. The LC-MS/MS data definitively showed acetylation occurring only on K56. No other modifications were observed for H3 or H4 (Figure 2.9 and Table 2.7).
Figure 2.6 LC-MS spectra of bovine histones from an LCT mass spectrometer.
Figure 2.7 Deconvoluted LC-MS spectra of bovine histones from an LCT mass spectrometer.
Figure 2.8 LC-MS of yeast histones from an LCT mass spectrometer. a: negative control of yeast histones; b: yeast histones with acetyl coenzyme A.
Precursor ion: 639.44$^{2+}$

*: -NH$_3$

': -H$_2$O

Figure 2.9 LC-MS/MS spectra of precursor ion at 639.44$^{2+}$. 
Table 2.7 LC-MS/MS verification of K56 acetylation in the recombinant yeast histones. SC – sequence coverage; PS – protein score.

<table>
<thead>
<tr>
<th>Histone</th>
<th>SC</th>
<th>PS</th>
<th>Peptide detected</th>
<th>Hit</th>
</tr>
</thead>
<tbody>
<tr>
<td>H3</td>
<td>74</td>
<td>606</td>
<td>\textbf{FQK$_{Ac}$STELLIR}$^{53}$</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>\textbf{KSAPSTGGVK}$^{36}$</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>\textbf{YKPGTVALK}$^{49}$</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>\textbf{STELLIR}$^{63}$</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>\textbf{EIAQDFK}$^{79}$</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>\textbf{EIAQDFKTDLR}$^{83}$</td>
<td>3</td>
</tr>
<tr>
<td>H4</td>
<td>96</td>
<td>854</td>
<td>\textbf{ILRDNIQGITKPAIR}$^{35}$</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>\textbf{ILRDNIQGITKPAIRR}$^{36}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>\textbf{DNIQGITKPAIR}$^{35}$</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>\textbf{DNIQGITKPAIRR}$^{36}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>\textbf{RISGLIYEEVR}$^{55}$</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>\textbf{ISGLIYEEVR}$^{55}$</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>\textbf{SFLESVIR}$^{67}$</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>\textbf{SFLESVIRDSVTYTEHAK}$^{77}$</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>\textbf{DSVTYTEHAK}$^{77}$</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>\textbf{DSVTYTEHAKR}$^{78}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>\textbf{KTVTSLDVYYALK}$^{91}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>\textbf{KTVTSLDVYYALKR}$^{92}$</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>\textbf{TVTSLDVYYALK}$^{91}$</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>\textbf{TVTSLDVYYALKR}$^{92}$</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>\textbf{TYGFGG}$^{102}$</td>
<td>44</td>
</tr>
</tbody>
</table>

2.4 CONCLUSIONS

The development of methods for the characterization of yeast histones and their PTMs were discussed in this chapter. We adapted the bulk histone purification protocol by extensive washing of nuclei. The resulting procedure gave a high yield of yeast histones with good preservation of acetylated isoforms. We next optimized our SDS-PAGE and LC-MS separations. SDS-PAGE is a powerful technique for protein separation. However, the traditional Laemmli system poorly resolves histones. We
observed that replacement of glycine with tricine gave better resolution. We also optimized the performance of our LC-MS experiments. The reversed-phase LC is an efficient means for purification of histone. When combined with mass spectrometry the LC-MS technique can be used to profile intact proteins. TFA is the most popular mobile phase additive used in LC-MS. Although it forms adducts with the analytes, at optimal conditions the effects are minimal and yeast core histones can be detected simultaneously on both LCT and Q-TOF. We optimized the concentration of TFA, the LC gradient, and the mass spectrometer cone voltage. Using optimized LC-MS methods we demonstrated the power of LC-MS by separating isoforms of histone H3.
CHAPTER 3

UNAMBIGUOUS DETERMINATION OF ISOBARIC HISTONE MODIFICATIONS BY LC-MS AND HIGH-MASS ACCURACY

3.1 INTRODUCTION

Methylation and acetylation of lysines in histones are two crucial posttranslational modifications that regulate gene transcription\textsuperscript{180} and have been shown to be misregulated in many forms of cancers\textsuperscript{181, 182}. In histone H3, K4, K9, K27, K36, and K79 can be methylated which is related to diverse transcription states\textsuperscript{183}. Acetylation of histone H3 K9, K14, K18, K23, and K27 is generally associated with gene activation\textsuperscript{184, 185}. It is noteworthy that histone H3 K9 methylation (H3K9Me) participates in heterochromatin formation and gene silencing, whereas acetylation of histone H3 K9 (H3K9Ac) has been reported to be associated with gene activation\textsuperscript{186}. The identification and quantitation of acetylation and methylation is therefore of high significance for understanding the role of histone modifications in gene regulation. The differentiation between trimethylation (ΔM = 42.0470 Da) and acetylation (ΔM = 42.0106 Da) is challenging since they are isobaric modifications that differ by 0.0364 Da and requires a mass resolving power of 27, 472 for a 1000 Da peptide.
Traditionally immunoassay methods, such as western blot and immunoprecipitation, are widely used to characterize histone modifications due to their high sensitivity. However, site specificity of antibodies is affected by adjacent amino acid modifications leading to poor specificity for epitopes with concomitant modifications. In addition, these approaches are also limited by the availability of site specific antibodies for known modifications and are not readily applicable for discovery and rapid characterization of novel modifications.

In recent years, mass spectrometry has been demonstrated to be very successful in the study of histone posttranslational modifications. The high-mass resolving power and high-mass accuracy of FT-ICR mass spectrometry enable it successfully distinguish between acetylation and trimethylation by peptide mass fingerprinting. Tandem mass spectrometry is also a powerful tool for distinguishing between these two modifications. By collision induced dissociation, acetylated peptides usually generate an ammonium ion at 126 Da while trimethylated peptides generate a neutral loss of trimethyl amine (N(CH$_3$)$_3$). Therefore, acetylation and trimethylation can be distinguished by the presence of these diagnostic peaks. However, these peaks are not always observed, especially for low abundant species. Another approach to distinguish trimethylation from acetylation is to introduce heavy isotopes to the acetyl or methyl groups so that trimethylation and acetylation will no longer be isobaric and then can be easily distinguished by mass alone. However, the labeling method is time consuming.
Relative retention in reversed-phase liquid chromatography (RPLC) has long been used as a supplement to MS for peptide identification. A number of models have been proposed to calculate peptide hydrophobicity although most endogenous posttranslational modifications do not significantly alter peptide hydrophobicity. Phosphorylation was believed to increase peptide hydrophobicity due to the addition of anionic/acidic phosphate groups, resulting in reduced retention. However, recent work by Steen and coworkers showed that this effect was not always born out. In some cases the increased hydrophobicity was compensated for by charge neutralization when the peptide contains basic amino acids that are positively charged under standard LC-MS conditions. In their study, the phosphorylated peptides eluted about the same time or later than the unmodified due to the presence of basic residues. Only peptides with a larger number of phospho moieties than basic residues showed shorter elution times than their unmodified forms. Their work demonstrated that modifications that result in reduction of net charges would reduce the overall hydrophobicity and increase the retention time on RPLC. For modifications with no introduction of charges such as mono-, di-, and trimethylation, we would expect no significant change of retention time. However, modifications that alter the number of basic sites would see a change in relative retention times. N-terminal acetylation of the alpha-amino group results in a substantial change in hydrophobicity and retention time. Guo and Mant reported that N-terminal acetylated peptides eluted 1.5-3.5 min later than the isoforms of the same peptides with a free N-terminus. Likewise, acetylation of lysine side-chains would neutralize the positive charge and lead to an increase in peptide hydrophobicity. Thus acetylated peptides in
general would be expected to elute later in reversed-phase chromatography than their unmodified isoforms. For example, Hunt showed that peptides $^9\text{KSTGGK}_{\text{Ac}}\text{APR}^{17}$ and $^9\text{K}_{\text{Me3}}\text{STGGKAPR}^{17}$ were separated by 17 min $^{192}$. Hunt and coworkers have also reported an increase of histone peptide hydrophobicity by incubating histones with propionic anhydride before in-solution trypsin digestion. The formation of a propionyl amide effectively neutralizes the charges of unmodified or endogenously monomethylated lysine $^{214}$. Propionic anhydride derivatization has also been applied to in-gel trypsin digestion of histones $^{215}$. Acetic anhydride is another derivatization reagent of lysine. It adds one acetyl group to the unmodified lysine residue resulting in an increased hydrophobicity of histone peptides $^{216}$.

Herein we report the general observation that acetylation of epsilon-amino groups of lysine results in shifted retention time whereas methylation of these same groups does not. The combination of relative retention shifts with tandem mass spectrometry allows for the unequivocal determination of trimethylation vs. acetylation.

3.2 EXPERIMENTAL

3.2.1 Yeast Histone Extraction

$S.\ cerevisiae$ strain BY4743 was obtained from Open Biosystems. Cells were grown at 30 °C and 200 rpm for 9 h. Growth was continued until cultures reached the late logarithmic/stationary phase and then harvested for histone extraction by centrifuging at 5500 $\times$ g for 10 min. Histones were purified as described in Chapter 2.
3.2.2 Yeast Histone H3 In-Gel Digestion

Histone H3 of *S. cerevisiae* was separated from other histones by use of SDS-PAGE with pre-cast 16.5% Tris-tricine gels (Bio-Rad Laboratories, Hercules, CA). H3 gel bands were in-gel digested with trypsin as previously described 204 (Appendix C.1). In brief, the H3 gel bands were excised and cut into small pieces and washed twice (one hour each) with freshly made 50% methanol/5% acetic acid solution. The gel pieces were then dehydrated in 200 µl of acetonitrile for 5 min followed by 5-min rehydration in 200 µl of 100 mM NH₄CO₃. This dehydration-rehydration procedure was repeated once, followed by another 5-min rehydration in acetonitrile. 30 µl of freshly prepared trypsin (20 ng/µl in 25 mM NH₄HCO₃) were added and rehydrated on ice for 10 min, then digested at 37 °C for one hour. Tryptic digested peptides were extracted with 50% acetonitrile/5% formic acid three times and dried to about 10 µl in a vacuum concentrator.

3.2.3 Nano-Liquid Chromatography-Tandem Mass Spectrometry (Nano-LC-MS/MS)

The digested peptides were subject to nano-LC-MS/MS analysis by use of either an LTQ FT-ICR mass spectrometer (Thermo Fisher, San Jose, CA) or an LCQ Deca XP+ ion trap mass spectrometer (Thermo Fisher) coupled with a Shimadzu LC 10ADvp capillary system (Columbia, MD, USA). Peptide separations were carried out with a commercial C₁₈ column (5 cm, 5 µm, I.D. 75 µm, New Objective, MA) using a gradient and working conditions as previously described 126. The peptides were separated using a 120 min gradient of mobile phase A (0.1% formic acid in water) and mobile phase B
(0.1% formic acid in acetonitrile). Mobile phase B was increased linearly from 5 to 60% in 80 min, held at 60% for 5 min, then increased to 95% in 5 min, held for 5 min and then returned to 5% to equilibrate the column for 15 minutes. The column was washed between each run to minimize carryover. One micro liter of the digest was injected onto the column. The electrospray voltage was maintained at 1.3 kV and capillary temperature was set at 200 °C. The mass spectrometric detection range was 200-2000 (m/z). Three-five data-dependent MS/MS scans with dynamic exclusion were carried out between each full MS scan. The product ion mass spectra were analyzed by use of in-house developed software, MassMatrix. The search parameters included the following variable modifications: acetylation of lysine and N-terminal, and methylation of lysine and arginine. Each of the tandem MS spectra matched by the database search was manually validated.

3.3 RESULTS AND DISCUSSION

3.3.1 Posttranslational Modification Patterns of Yeast Histone H3

Acetylation and trimethylation are isobaric posttranslational modifications that differ in mass by 0.0364 Da. High-resolution / high-mass accuracy mass spectrometers, such as FT-ICR and Orbitraps, can readily distinguish these modifications by mass alone. All initial experiments were carried out on a hybrid LTQ-FT mass spectrometer in order to establish with high confidence the pattern of histone modifications for yeast histone H3. The full MS spectra were collected via FT-ICR MS and product ion mass spectra were obtained in the LTQ. The high-mass accuracy precursor ion spectra were used to
establish the presence of trimethylation vs. acetylation as described previously. All peptide assignments were supported by MS/MS spectra with manually validated database search matches. The peptides identified with the LTQ-FT are listed in Table 3.1.

High-mass accuracy is a proven approach to discriminate between isobaric peptides. For example, a precursor ion was observed at $m/z = 358.7179$. Based on nominal mass the peptide could be either the acetylated or trimethylated peptide $^{18}\text{KQLASK}^{23}$ or peptide $^{117}\text{VTIQQK}^{122}$. Based on accurate mass the trimethylated $^{18}\text{KQLASK}^{23}$ or $^{117}\text{VTIQQK}^{122}$ would have mass errors as high as 50 ppm. Such mass errors are highly improbable given a properly calibrated FT-ICR mass spectrometer. The more likely assignment is the acetylated peptide $^{18}\text{KQLASK}^{23}$ with a mass error of -4.04 ppm. The assignment of the backbone peptide sequence was corroborated by MS/MS. In this manner we confirmed that K4, K36, and K79 were (tri)methylated and K9, K14, K18, K23, K27, and K56 were acetylated on yeast histone H3. As indicated in Table 3.1, each modification had a resulting error less than 6 ppm. These observed modification patterns determined by mass spectrometry are consistent with those obtained from other techniques $^{185, 218-221}$.

In addition to the absolute mass, peptide mass shifts between different isoforms can also be used to corroborate modification identifications $^{222}$. As shown in Figure 3.1, mass shifts of 42.0104 and 42.0096 Da between isoforms of peptide $^{9}\text{KSTGGKAPR}^{17}$ suggested K9 and K14 in the peptide were both acetylated rather than trimethylated. The tandem MS spectra shown in Figure 3.2 establish the predominant acetylation first occurring at K14 followed by K9. In the case of peptide $^{73}\text{EIAQDFKTDLR}^{83}$, a mass
shift of 42.0450 Da (shown in Figure 3.3) between the isoforms suggested that the peptide was trimethylated. More importantly, the presence of isoforms at $m/z = 675.3560$, 682.3629, and 689.3711 correspond in mass to the mono-, di-, and trimethylated species. The tandem MS spectra shown in Figure 3.4 indicate K79 as the site of mono-, di-, and trimethylation. In the figure for trimethylation of K79, tandem mass spectra of the triply charged precursor ion at $m/z = 460.2520$ is shown instead of the doubly charged precursor ion at $m/z = 689.3711$ because the former showed better fragmentation.

In summary, the data from accurate mass, relative mass shifts, and tandem MS spectra support the conclusion that yeast histone H3 K4, K36, and K79 were methylated while K9, K14, K18, K23, K27, and K56 were acetylated. These well validated results, as summarized in Table 3.2, serve as the standards that will be used in our examination of relative retention time shifts between acetylation and trimethylation.
Table 3.1 Peptides of yeast H3 detected by an LTQ FT-ICR mass spectrometer. Note: False assignments are italicized.

<table>
<thead>
<tr>
<th>Measured $(m/z)$</th>
<th>Calculated $(m/z)$</th>
<th>RT (min)</th>
<th>Sequence</th>
<th>Error (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>337.7123$^{2+}$</td>
<td>337.7141$^{2+}$</td>
<td>8.16</td>
<td>$^{18}$KQLASK$^{23}$</td>
<td>-5.33</td>
</tr>
<tr>
<td>358.7179$^{2+}$</td>
<td>358.7192$^{2+}$</td>
<td>16.15</td>
<td>$^{18}$KAcQLASK$^{23}$</td>
<td>-4.18</td>
</tr>
<tr>
<td></td>
<td>358.7375$^{4+}$</td>
<td></td>
<td>$^{18}$KMe$_3$QLASK$^{23}$</td>
<td>-54.64</td>
</tr>
<tr>
<td></td>
<td>358.7375$^{2+}$</td>
<td></td>
<td>$^{17}$VTIQKK$^{12}$</td>
<td>-54.64</td>
</tr>
<tr>
<td>352.7056$^{2+}$</td>
<td>352.7068$^{2+}$</td>
<td>6.10</td>
<td>$^{3}$TKQTAR$^{8}$</td>
<td>-3.40</td>
</tr>
<tr>
<td>359.7131$^{2+}$</td>
<td>359.7146$^{2+}$</td>
<td>6.91</td>
<td>$^{3}$TKMe$_3$QTAR$^{8}$</td>
<td>-4.17</td>
</tr>
<tr>
<td>366.7208$^{2+}$</td>
<td>366.7224$^{2+}$</td>
<td>7.13</td>
<td>$^{3}$TKMe$_2$QTAR$^{8}$</td>
<td>-4.36</td>
</tr>
<tr>
<td>373.7287$^{2+}$</td>
<td>373.7303$^{2+}$</td>
<td>6.99</td>
<td>$^{3}$TKMe$_3$QTAR$^{8}$</td>
<td>-4.28</td>
</tr>
<tr>
<td></td>
<td>373.7127$^{2+}$</td>
<td></td>
<td>$^{3}$TK$_{2+}$QTAR$^{8}$</td>
<td>44.42</td>
</tr>
<tr>
<td>387.2156$^{2+}$</td>
<td>387.2175$^{2+}$</td>
<td>8.20</td>
<td>$^{10}$STGGKAPR$^{17}$</td>
<td>-4.91</td>
</tr>
<tr>
<td>408.2213$^{2+}$</td>
<td>408.2228$^{2+}$</td>
<td>15.48</td>
<td>$^{10}$STGGK$_{Ac}$APR$^{17}$</td>
<td>-3.67</td>
</tr>
<tr>
<td></td>
<td>408.2410$^{2+}$</td>
<td></td>
<td>$^{10}$STGGK$_{Me3}$APR$^{17}$</td>
<td>-48.26</td>
</tr>
<tr>
<td>451.2635$^{2+}$</td>
<td>451.2650$^{2+}$</td>
<td>7.43</td>
<td>$^{9}$KSTGGKAPR$^{17}$</td>
<td>-3.32</td>
</tr>
<tr>
<td>472.2687$^{2+}$</td>
<td>472.2703$^{2+}$</td>
<td>10.47</td>
<td>$^{9}$KSTGGK$_{Ac}$APR$^{17}$</td>
<td>-3.38</td>
</tr>
<tr>
<td></td>
<td>472.2885$^{2+}$</td>
<td></td>
<td>$^{9}$KSTGGK$_{Me3}$APR$^{17}$</td>
<td>-41.92</td>
</tr>
<tr>
<td>493.2735$^{2+}$</td>
<td>493.2756$^{2+}$</td>
<td>16.49</td>
<td>$^{9}$K$<em>{Ac}$STGGK$</em>{Ac}$APR$^{17}$</td>
<td>-4.26</td>
</tr>
<tr>
<td></td>
<td>493.2938$^{2+}$</td>
<td></td>
<td>$^{9}$K$<em>{Me3}$STGGK$</em>{Ac}$APR$^{17}$</td>
<td>-41.15</td>
</tr>
<tr>
<td></td>
<td>493.3119$^{2+}$</td>
<td></td>
<td>$^{9}$K$<em>{Me3}$STGGK$</em>{Me3}$APR$^{17}$</td>
<td>-77.84</td>
</tr>
<tr>
<td>528.8105$^{2+}$</td>
<td>528.8123$^{2+}$</td>
<td>20.80</td>
<td>$^{18}$K$<em>{Ac}$QLASK$</em>{Ac}$AAR$^{26}$</td>
<td>-3.40</td>
</tr>
<tr>
<td></td>
<td>528.8303$^{2+}$</td>
<td></td>
<td>$^{18}$K$<em>{Ac}$QLASK$</em>{Me3}$AAR$^{26}$</td>
<td>-37.82</td>
</tr>
<tr>
<td></td>
<td>528.8487$^{2+}$</td>
<td></td>
<td>$^{18}$K$<em>{Me3}$QLASK$</em>{Me3}$AAR$^{26}$</td>
<td>-72.23</td>
</tr>
<tr>
<td>366.7161$^{4+}$</td>
<td>366.7172$^{4+}$</td>
<td>13.89</td>
<td>$^{2}$K$<em>{Ac}$SAPSTGGVK$</em>{Me}$KPHR$^{40}$</td>
<td>-3.00</td>
</tr>
<tr>
<td>370.2193$^{4+}$</td>
<td>370.2211$^{4+}$</td>
<td>13.89</td>
<td>$^{2}$K$<em>{Ac}$SAPSTGGVK$</em>{Me3}$KPHR$^{40}$</td>
<td>-3.24</td>
</tr>
<tr>
<td>373.7237$^{4+}$</td>
<td>373.7250$^{4+}$</td>
<td>13.89</td>
<td>$^{2}$K$<em>{Ac}$SAPSTGGVK$</em>{Me3}$KPHR$^{40}$</td>
<td>-3.48</td>
</tr>
<tr>
<td></td>
<td>373.7159$^{4+}$</td>
<td></td>
<td>$^{27}$K$<em>{Ac}$SAPSTGGVK$</em>{Ac}$KPHR$^{40}$</td>
<td>20.87</td>
</tr>
<tr>
<td>373.7141$^{4+}$</td>
<td>373.7159$^{4+}$</td>
<td>15.63</td>
<td>$^{27}$K$<em>{Ac}$SAPSTGGVK$</em>{Ac}$KPHR$^{40}$</td>
<td>-4.82</td>
</tr>
<tr>
<td></td>
<td>373.7250$^{4+}$</td>
<td></td>
<td>$^{27}$K$<em>{Me3}$SAPSTGGVK$</em>{Me}$KPHR$^{40}$</td>
<td>-29.17</td>
</tr>
<tr>
<td>377.2179$^{4+}$</td>
<td>377.2198$^{4+}$</td>
<td>15.63</td>
<td>$^{27}$K$<em>{Ac}$SAPSTGGVK$</em>{Me3}$KPHR$^{40}$</td>
<td>-5.04</td>
</tr>
<tr>
<td>380.7218$^{4+}$</td>
<td>380.7237$^{4+}$</td>
<td>15.63</td>
<td>$^{27}$K$<em>{Ac}$SAPSTGGVK$</em>{Me3}$KPHR$^{40}$</td>
<td>-4.99</td>
</tr>
<tr>
<td>384.2256$^{4+}$</td>
<td>384.2276$^{4+}$</td>
<td>15.63</td>
<td>$^{27}$K$<em>{Ac}$SAPSTGGVK$</em>{Me3}$KPHR$^{40}$</td>
<td>-5.21</td>
</tr>
<tr>
<td></td>
<td>384.1936$^{4+}$</td>
<td></td>
<td>$^{27}$K$<em>{Ac}$SAPSTGGVK$</em>{Ac}$KPHR$^{40}$</td>
<td>83.29</td>
</tr>
<tr>
<td>497.9493$^{3+}$</td>
<td>497.9519$^{3+}$</td>
<td>15.63</td>
<td>$^{27}$K$<em>{Ac}$SAPSTGGVK$</em>{Ac}$KPHR$^{40}$</td>
<td>-5.22</td>
</tr>
<tr>
<td></td>
<td>497.9855$^{3+}$</td>
<td></td>
<td>$^{27}$K$<em>{Me3}$SAPSTGGVK$</em>{Me3}$KPHR$^{40}$</td>
<td>-73.50</td>
</tr>
<tr>
<td>502.6213$^{3+}$</td>
<td>502.6237$^{3+}$</td>
<td>15.63</td>
<td>$^{2}$K$<em>{Ac}$SAPSTGGVK$</em>{Me}$KPHR$^{40}$</td>
<td>-4.77</td>
</tr>
<tr>
<td>507.2933$^{3+}$</td>
<td>507.2956$^{3+}$</td>
<td>15.73</td>
<td>$^{2}$K$<em>{Ac}$SAPSTGGVK$</em>{Me3}$KPHR$^{40}$</td>
<td>-4.53</td>
</tr>
<tr>
<td>511.9651$^{3+}$</td>
<td>511.9675$^{3+}$</td>
<td>15.63</td>
<td>$^{2}$K$<em>{Ac}$SAPSTGGVK$</em>{Me}$KPHR$^{40}$</td>
<td>-4.69</td>
</tr>
<tr>
<td>695.9089$^{2+}$</td>
<td>695.9126$^{2+}$</td>
<td>27.47</td>
<td>$^{53}$RFQKSTEILLIR$^{63}$</td>
<td>-5.32</td>
</tr>
<tr>
<td>716.9148$^{2+}$</td>
<td>716.9178$^{2+}$</td>
<td>29.35</td>
<td>$^{53}$RFQK$_{Ac}$STEILLIR$^{63}$</td>
<td>-4.18</td>
</tr>
<tr>
<td></td>
<td>716.9360$^{2+}$</td>
<td></td>
<td>$^{53}$RFQK$_{Me3}$STEILLIR$^{63}$</td>
<td>-29.57</td>
</tr>
<tr>
<td>638.8659$^{2+}$</td>
<td>638.8671$^{2+}$</td>
<td>32.58</td>
<td>$^{54}$FQK$_{Ac}$STEILLIR$^{63}$</td>
<td>-1.88</td>
</tr>
</tbody>
</table>

Note: [pm]
Table 3.1 Continued.

<table>
<thead>
<tr>
<th>Measured $(m/z)$</th>
<th>Calculated $(m/z)$</th>
<th>RT (min)</th>
<th>Sequence</th>
<th>Error (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>582.6666$^3+$</td>
<td>582.6689$^3+$</td>
<td>29.92</td>
<td>$^{70}$LVREIAQDFK$^{Me3}$TDLR$^{83}$</td>
<td>-3.95</td>
</tr>
<tr>
<td>582.6349$^3+$</td>
<td></td>
<td></td>
<td>$^{70}$LVREIAQDFK$^{Ac}$TDLR$^{83}$</td>
<td>54.41</td>
</tr>
<tr>
<td>668.3486$^2+$</td>
<td>668.3495$^2+$</td>
<td>26.85</td>
<td>$^{73}$EIAQDFKTDLR$^{83}$</td>
<td>-1.35</td>
</tr>
<tr>
<td>668.3789$^2+$</td>
<td></td>
<td></td>
<td>$^{28}$SAPSTGGVK$^{Me}$KPHR$^{40}$</td>
<td>-45.33</td>
</tr>
<tr>
<td>675.3560$^2+$</td>
<td>675.3573$^2+$</td>
<td>27.25</td>
<td>$^{73}$EIAQDFK$^{Me}$TDLR$^{83}$</td>
<td>-1.92</td>
</tr>
<tr>
<td>675.3867$^2+$</td>
<td></td>
<td></td>
<td>$^{28}$SAPSTGGVK$^{Me2}$KPHR$^{40}$</td>
<td>-45.45</td>
</tr>
<tr>
<td>682.3629$^2+$</td>
<td>682.3651$^2+$</td>
<td>27.35</td>
<td>$^{73}$EIAQDFK$^{Me2}$TDLR$^{83}$</td>
<td>-3.22</td>
</tr>
<tr>
<td>682.3763$^2+$</td>
<td>682.3945$^2+$</td>
<td>27.92</td>
<td>$^{28}$SAPSTGGVK$^{Ac}$KPHR$^{40}$</td>
<td>-19.64</td>
</tr>
<tr>
<td>689.3711$^2+$</td>
<td>689.3729$^2+$</td>
<td>26.92</td>
<td>$^{73}$EIAQDFK$^{Me3}$TDLR$^{83}$</td>
<td>-2.61</td>
</tr>
<tr>
<td>689.3947$^2+$</td>
<td></td>
<td></td>
<td>$^{73}$EIAQDFK$^{Ac}$TDLR$^{83}$</td>
<td>23.79</td>
</tr>
<tr>
<td>689.3842$^2+$</td>
<td></td>
<td></td>
<td>$^{28}$SAPSTGGVK$^{Ac}$K$^{Me}$PHR$^{40}$</td>
<td>-19.00</td>
</tr>
<tr>
<td>689.4024$^2+$</td>
<td></td>
<td></td>
<td>$^{28}$SAPSTGGVK$^{Me3}$K$^{Me}$PHR$^{40}$</td>
<td>-45.40</td>
</tr>
<tr>
<td>335.5339$^3+$</td>
<td>335.5351$^3+$</td>
<td>20.96</td>
<td>$^{41}$YKPGTVALR$^{48}$</td>
<td>-3.58</td>
</tr>
<tr>
<td>534.3211$^3+$</td>
<td>534.3233$^3+$</td>
<td>31.44</td>
<td>$^{57}$STELLIRKLPFQR$^{69}$</td>
<td>-4.12</td>
</tr>
</tbody>
</table>

Table 3.2 Posttranslational modification sites on histone H3 of *S. cerevisiae*.

<table>
<thead>
<tr>
<th>Residue</th>
<th>Posttranslational modification</th>
</tr>
</thead>
<tbody>
<tr>
<td>K4</td>
<td>Me1, Me2, Me3</td>
</tr>
<tr>
<td>K9</td>
<td>Ac</td>
</tr>
<tr>
<td>K14</td>
<td>Ac</td>
</tr>
<tr>
<td>K18</td>
<td>Ac</td>
</tr>
<tr>
<td>K23</td>
<td>Ac</td>
</tr>
<tr>
<td>K27</td>
<td>Ac</td>
</tr>
<tr>
<td>K36</td>
<td>Me1, Me2, Me3</td>
</tr>
<tr>
<td>K56</td>
<td>Ac</td>
</tr>
<tr>
<td>K79</td>
<td>Me1, Me2, Me3</td>
</tr>
</tbody>
</table>
Figure 3.1 Full MS results of peptide $K_{STGGKAPR}^{17}$ from an LTQ FT-ICR mass spectrometer. The unmodified and modified isoforms are doubly charged.
Figure 3.2 Tandem MS results and LC traces of peptide $^9$KSTGGKAPR$^{17}$ from an LTQ FT-ICR mass spectrometer.
Figure 3.3 Full MS results of peptide $^{73}$EIAQDFKTLR$^{83}$ from an LTQ FT-ICR mass spectrometer. The unmodified and modified isoforms are doubly charged.
Figure 3.4 Tandem MS results and LC traces of peptide $^{73}$EIAQDFKTDLR$^{83}$ from an LTQ FT-ICR mass spectrometer.
3.3.2 Retention Time Shifts on an FT-ICR Mass Spectrometer

As demonstrated above, accurate absolute mass and accurate relative mass shifts are powerful approaches to distinguish between trimethylation and acetylation. For data obtained on low-mass accuracy instruments, additional corroborative data such as the presence of supporting MS/MS reporter ions are required. As previously reported, relative retention time can also be used to distinguish between sequence assignments for isobaric peptides; $^{73}$EIAQDFKTDLR$^{83}$ vs. $^{28}$SAPSTGGVK$_{Me}$KPHR$^{40}$, $^{73}$EIAQDFK$_{Me}$TDLR$^{83}$ vs. $^{28}$SAPSTGGVK$_{Me2}$KPHR$^{40}$, and $^{73}$EIAQDFK$_{Me2}$TDLR$^{83}$ vs. $^{28}$SAPSTGGVK$_{Me3}$KPHR$^{40}$. This approach can also be extended to effectively distinguish between trimethylation and acetylation.

Acetylation of lysine effectively neutralizes the lysine’s positive charge under acidic reversed-phase separation conditions. The change in charge results in increased retention time relative to the unmodified peptide isoform. This effect was observed for all acetylated peptide isoforms of yeast histone H3. As shown in Figure 3.2, unmodified fragment $^{9}$KSTGGKAPR$^{17}$ eluted at 7.4 min, whereas the K14 acetylated isoform eluted at 10.5 min and the K9 and K14 acetylated isoform eluted at 16.5 min. The effect of acetylation on retention time was also observed for peptides $^{10}$STGGKAPR$^{17}$, $^{18}$KQLASK$^{23}$, and $^{53}$RFQKSTELLIR$^{63}$ (the results were shown in Table 3.1 and Figure 3.5-3.7). From these peptides we conclude that the increase in retention time is cumulative and of similar magnitude as N-terminal acetylation.

Unlike the addition of the acetyl group, the addition of methyl groups to lysine does not neutralize the lysine’s charge. Therefore, it was expected that methylation would
not dramatically alter retention time. As shown in Figure 3.4, the unmodified form of peptide $^{73}\text{EIAQDFKTDLR}^{83}$ eluted at 26.9 min whereas its mono-, di-, and trimethylated isoforms eluted at 27.2, 27.4, and 26.9 min, respectively. The similarity in retention time for the unmodified and methylated forms indicates that the addition of methyl groups to lysine has only a small effect on peptide retention time. The same effect was observed for peptides $^{70}\text{LVREIAQDFKTDLR}^{83}$, whose results are shown in Table 3.1 and Figure 3.8 and Figure 3.9.

Based on the above observation that acetylation has a significant effect on retention time while methylation has little, we hypothesize that a peptide subject to both acetylation and trimethylation would have its acetylated isoforms eluted several minutes later than the unmodified/methylated isoforms. Given this assertion, we propose that trimethylation and acetylation are easily distinguished based on relative retention times and supporting MS/MS. This hypothesis is supported by observations for the peptide $^{27}\text{KSAPSTGGVKKPHR}^{40}$ in which K27 and K36 are subject to concomitant acetylation and methylation, respectively. As seen in Figure 3.10, the methylated K36 isoforms eluted around 13.8 min whereas the isoforms in which K27 is acetylated (along with K36 methylations) eluted around 15.7 min. This assignment is further supported by the peptide assignments based on accurate mass in Table 3.1 and the MS/MS spectra shown in Figure 3.11 and 3.12. Thus the group of peptides eluting earlier are methylated isoforms of K36 and the later group is the same isoforms but with K27 acetylated. Figure 3.10 clearly shows the effect of acetylation vs. methylation of lysine on retention time. Given the supporting data of MS/MS and retention time it should be possible to
unambiguously distinguish between these isobaric modifications without the need for
accurate mass. To test this assertion the same analysis of acetylation and trimethylation
was carried out on a common ion trap mass spectrometer.
Figure 3.5 LC traces (left) and spectra (right) of peptide $^{10}$STGGKAPR$^{17}$ from an LTQ FT-ICR mass spectrometer.
Figure 3.6 LC traces (left) and spectra (right) of peptide $^{18}$KQLASK$^{23}$ from an LTQ FT-ICR mass spectrometer.
Figure 3.7 LC traces (left) and spectra (right) of peptide $^{53}$RFQKSTELLIR$^{63}$ from an LTQ FT-ICR mass spectrometer.
Figure 3.8 LC traces (A) and spectra (B) of peptide $^{3}\text{TKQTR}^8$ from an LTQ FT-ICR mass spectrometer.
Figure 3.9 LC traces (A) and spectra (B) of peptide $^{70}$LVREIAQDFKTDLR$^{83}$ from an LTQ FT-ICR mass spectrometer.
Figure 3.10 LC traces of peptide $^{27}$KSAPSTGGVKKPHR$^{40}$ from an LTQ FT-ICR mass spectrometer.
Figure 3.11 Full MS results of peptide $^{27}$KSAPSTGGVKKPHR$^{40}$ from an LTQ FT-ICR mass spectrometer. All isoforms are triply charged.
Figure 3.12 Tandem MS results of peptide $^{27}$KSAPSTGGVKKPHR$^{40}$ from an LTQ FT-ICR mass spectrometer.
3.3.3 Unambiguous Determination of Acetylation and Trimethylation at Low-Mass Accuracy

Results from the FT-ICR mass spectrometer demonstrated convincingly that acetylated isoforms elute several minutes later than the corresponding unmodified and methylated isoforms. This observation suggests that retention time can be used effectively to support MS/MS data for determining the presence of either acetylation or trimethylation. The nano-LC-MS/MS experiments were then performed on an LCQ Deca XP+ ion trap mass spectrometer.

To illustrate the power of the combined approach, peptide $^{27}$KSAPSTGGVK$^{36}$, in which K27 and K36 are subject to acetylation or trimethylation, was examined. With the low-mass accuracy of the ion trap mass spectrometer it was impossible to conclude whether K27/K36 was either acetylated or trimethylated. However, based on the relative retention time shift we can confidently conclude that the isoform eluting at 10.8 min was acetylated and not trimethylated because it eluted 4 min later than its unmodified form (Figure 3.13). Using retention time, we also identified the modification patterns of K18 and K16 in peptide $^{18}$KQLASKAAR$^{26}$. As shown in Table 3.3, the isoform of peptide $^{18}$KQLASKAAR$^{26}$ eluting at 25.5 min was believed to be the isoform with K18 acetylated because it eluted 5.9 min later than the other isoform. This effect was also observed on peptide $^{1}$SGRGKGGKGLGKGGAKR$^{17}$ (Figure 3.14). Methylated peptides also showed similar retention times as expected (Figure 3.15). Based on our observations, we conclude that relative retention time from RPLC plays a powerful
supporting role in distinguishing between the isobaric modifications, acetylation and trimethylation.

3.4 CONCLUSIONS

The present study was undertaken to explore the feasibility of distinguishing acetylation from trimethylation of lysine by using retention time on RPLC. To address the question, histone H3 of *S. cerevisiae* was characterized by nano-LC-MS/MS on FT-ICR and ion trap mass spectrometers. The data convincingly demonstrate that acetylation resulted in substantial changes in retention time, whereas methylation did not. Thus relative retention plays a powerful supporting role in distinguishing between these two isobaric modifications.
Figure 3.13 Tandem MS results and LC traces of peptide $^{27}\text{KSAPSTGGVK}^{36}$ from an LCQ Deca XP+ ion trap mass spectrometer.
Figure 3.14 LC traces (top) and spectra of peptide $^{1}_{\text{Ac}}$SGRGKGGKGLGKGGAK$^{17}$ on an LCQ Deca XP+ ion trap mass spectrometer.
Figure 3.15 Tandem MS results and LC traces of peptide \(^{73}\text{EIAQDFKTDLR}^{83}\) from an LCQ Deca XP+ ion trap mass spectrometer.
Table 3.3 Peptides detected on an LCQ Deca XP+ ion trap mass spectrometer.

<table>
<thead>
<tr>
<th>Measured ($m/z$)</th>
<th>Calculated ($m/z$)</th>
<th>RT (min)</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>446.31$^{3+}$</td>
<td>445.90$^{3+}$</td>
<td>32.40</td>
<td>$^{73}$EIAQDFKTDLR$^{83}$</td>
</tr>
<tr>
<td>675.88$^{2+}$</td>
<td>675.36$^{2+}$</td>
<td>32.74</td>
<td>$^{73}$EIAQDFK$_{Me}$TDLR$^{83}$</td>
</tr>
<tr>
<td>455.76$^{3+}$</td>
<td>455.25$^{3+}$</td>
<td>32.89</td>
<td>$^{73}$EIAQDFK$_{Me2}$TDLR$^{83}$</td>
</tr>
<tr>
<td>460.56$^{3+}$</td>
<td>459.92$^{3+}$</td>
<td>32.54</td>
<td>$^{73}$EIAQDFK$_{Me3}$TDLR$^{83}$</td>
</tr>
<tr>
<td>669.07$^{2+}$</td>
<td>668.38$^{2+}$</td>
<td>15.38</td>
<td>$^{28}$SAPSTGGVK$_{Me}$KPHR$^{40}$</td>
</tr>
<tr>
<td>675.76$^{2+}$</td>
<td>675.39$^{2+}$</td>
<td>16.09</td>
<td>$^{28}$SAPSTGGVK$_{Me2}$KPHR$^{40}$</td>
</tr>
<tr>
<td>683.04$^{2+}$</td>
<td>682.39$^{2+}$</td>
<td>15.68</td>
<td>$^{28}$SAPSTGGVK$_{Me3}$KPHR$^{40}$</td>
</tr>
<tr>
<td>466.68$^{2+}$</td>
<td>466.26$^{2+}$</td>
<td>6.52</td>
<td>$^{27}$KSAPSTGGVK$^{36}$</td>
</tr>
<tr>
<td>487.51$^{2+}$</td>
<td>487.27$^{2+}$</td>
<td>10.83</td>
<td>$^{27}$K$_{Ac}$SAPSTGGVK$^{36}$</td>
</tr>
<tr>
<td>508.15$^{2+}$</td>
<td>507.81$^{2+}$</td>
<td>16.73</td>
<td>$^{18}$KQLASK$_{Ac}$AAR$^{26}$</td>
</tr>
<tr>
<td>529.18$^{2+}$</td>
<td>528.81$^{2+}$</td>
<td>25.50</td>
<td>$^{18}$K$<em>{Ac}$QLASK$</em>{Ac}$AAR$^{26}$</td>
</tr>
<tr>
<td>538.59$^{3+}$</td>
<td>538.32$^{3+}$</td>
<td>79.04</td>
<td>$^{1}_{Ac}$SGRGKGGKGLGKGGAKR$^{17}$</td>
</tr>
<tr>
<td>828.09$^{2+}$</td>
<td>827.98$^{2+}$</td>
<td>82.17</td>
<td>$^{1}<em>{Ac}$SGRGKGGKGLGKGGAK$</em>{Ac}$R$^{17}$</td>
</tr>
</tbody>
</table>
4.1 INTRODUCTION

The covalent modification of histones plays pivotal roles in the epigenetic control of gene expression. Histone methylation, which occurs on the side chains of lysine and arginine and is most prominent in histone H3 and H4, is linked to transcriptional activation, differentiation, imprinting, and X inactivation. In general, histone methylation at H3 K4, K36, and K79 is associated with euchromatin and gene activation, whereas histone methylation at H3 K9 and K27, and H4 K20 is associated with heterochromatin and repressed genes. Lysine may be mono-, di-, or trimethylated – where the degree of modification is likely related to different biological events. In the budding yeast *S. cerevisiae*, histone H3 K4, K36, and K79 are methylated to different degrees by histone methyltransferases Set1, Set2, and Dot1, respectively. While methylation of these sites has been suggested to be a marker of active transcription, it may also play critical roles in gene silencing and DNA damage responses.

The above site-specific histone methyltransferases have been studied for a number of years. It was believed that histone methylation is stable and irreversible. However,
emerging evidence shows that histone methylation is dynamic \(^{230, 231}\), which suggests the existence of histone demethylases. Shi’s 2004 landmark discovery of human amine oxidase LSD1, the first unequivocal histone demethylase, demonstrated that histone methylation is indeed reversible \(^{51}\). LSD1 demethylates mono- and dimethylated K4 or K9 on H3 \(^{51, 232, 233}\), and its homologues are found in organisms ranging from \textit{S. pombe} to mammals. However, due to the nature of the amino oxidation demethylation mechanism (which requires a free lone pair of electrons on the lysine ε-nitrogen), LSD1 is unable to demethylate trimethylated lysine. Thus, it was clear that an additional class of histone lysine demethylases would be requisite. Shengjiang Tu posited, as described in the following paragraph, that Jumonji C (JmjC) domain-containing proteins might be new histone demethylases. This hypothesis, apparently shared by other research groups, has been empirically verified recently with the discovery of novel human JmjC domain-containing histone demethylases \(^{234-239}\).

Histone demethylation in budding yeast is still poorly understood. In particular, the trimethylated lysine demethylases remain to be determined. The fact that histone lysine methylation is prevalent in \textit{S. cerevisiae} and that it does not encode an LSD1 homologue \(^{51}\) suggests that it is imperative to identify yeast histone demethylases. In our search for new yeast demethylases, we postulated that candidate histone demethylases would satisfy four specific criteria. Proteins containing the JmjC domain satisfy each of these criteria as follows. \(i\) Candidates must be able to cleave the stable N-CH\(_3\) bond. We surmised that an Fe\(^{2+}\)/α-ketoglutarate-dependent dioxygenase could potentially catalyze histone demethylation via an oxygen radical mechanism in which an unstable carbinolamine
intermediate is generated and followed by spontaneous elimination of formaldehyde. Several groups have independently suggested a similar mechanism. Furthermore, an Fe\(^{2+}\)/α-ketoglutarate family protein, AlkB, was recently shown to catalyze the hydroxylation and subsequent demethylation of the DNA lesions 1-methyladenine and 3-methylcytosine. The fact that the JmjC domain displays structural similarity to Fe\(^{2+}\)/α-ketoglutarate enzymes suggests that it might be capable of catalyzing demethylation.

(ii) Candidates must be known to influence histone modifications. In satisfaction of this criterion, two *S. pombe* JmjC proteins, Epe1 and Msc1, have been reported to be associated with histone modifications.

(iii) Candidates must be involved in transcription regulation, such as gene silencing. Consistent with this, several JmjC-containing proteins interact with either the tumor suppressor retinoblastoma protein (Rb) or histone deacetylases (HDACs) and regulate transcription. For example, RBP2, RBP1, and Jumonji interact with Rb; JMJD2A form complexes with HDACs and Rb. The JmjC proteins SMCX and SMCY are known to be involved in X-inactivation. Moreover, the JmjC domain is found in conjunction with zinc finger, PHD, ARID, or TUDOR domains – all of which are protein-protein interaction or DNA binding domains involved in transcription. Additionally, PHD and TUDOR domains have been shown to be methylated histone binding domains.

(iv) Given the conservation of histone methylation, candidate histone demethylases should be conserved in all eukaryotic organisms, from *S. cerevisiae* to human, which is true for JmjC domain proteins.
Herein, we increase the repertoire of JmjC demethylases by demonstrating that four out of the five JmjC-containing proteins encoded by *S. cerevisiae* influence histone methylation levels *in vivo*. The five yeast proteins are Ecm5 (YMR176W), Gis1 (YDR096W), Rph1 (YER169W), Jhd1 (YER051W), and Jhd2 (YJR119C) (Figure B.1A). YJR119C gene product is named Jhd2 (*JmjC* domain-containing histone demethylase 2 by following the naming of Jhd1). Here, we utilized proteomic MS techniques which have been successfully used in studying individual yeast histone methyltransferase before. We systematically analyzed global (overall pattern) and site-specific histone H3 methylation in wild type (WT) and five yeast histone demethylase candidate knockout strains. The changes in methylation patterns led us to predict that Rph1, Jhd1, Gis1, and Jhd2 are histone demethylases. Subsequent overexpression data are complementary to the deletion data. Particularly, overexpression of *RPH1* and *JHD2* significantly decrease histone H3 K36Me3 and K4Me3 levels, respectively. *In vitro* enzymatic assays demonstrate that Rph1 demethylates histone H3 K36Me3 primarily and K36Me2 secondarily. The *RPH1* overexpression strain is sensitive to UV irradiation damage. In addition, its histone H3K36Me3 demethylase activity is responsible for the phenotype. Rph1 and Jhd1 might be involved in transcription elongation.
4.2 EXPERIMENTAL

4.2.1 Yeast Strains and Cell Culture

The WT yeast, knockout strains (ecm5Δ, jhd1Δ, gis1Δ, rph1Δ, and jhd2Δ), and overexpression strains (JHD1, RPH1, and JHD2) were graciously supplied by the laboratory of Dr. Ming-Daw Tsai (OSU). Cells were cultured by Shengjiang Tu and Ester Bulloch as described in Chapter 5.

4.2.2 Histone Extraction and Mass Spectrometry Analysis

Histones were purified and tryptic in-gel digested as described in Chapter 3 (Appendix C.1 and C.7). Histone H3 was separated from other histones on a 16.5% Tris-tricine SDS-PAGE gel as described in Chapter 2 (Appendix C.4). LC-MS was performed on Micromass LCT (Micromass, Wythenshawe, UK) as described in Chapter 2. Nano LC-MS/MS was performed on an LCQ Deca XP+ ion trap mass spectrometer (Thermo Finnigan, San Jose, CA) as described in Chapter 3. For MADLI-TOF MS, a Bruker Reflex III was operated in linear, positive ion mode at a constant accelerating voltage of 28 kV. The N₂ laser power was minimized to the level just necessary to obtain signal. Based on the samples of interest, protein standards were used to calibrate the instrument over the mass range of 500-2500 m/z. Saturated α-cyano-4-hydroxy-cinnamic acid (HCCA) solution in 0.1% TFA/acetonitrile was mixed with the sample at a ratio of 5:1. Eight-tenths µl of the resulting mixture was spotted on a ground steel MTP384 target (Bruker Atonics, Billerica, MA) and then air dried. For samples that contained a large
amount of unwanted salt, the dried sample on the MTP384 target was washed with 0.1% TFA.

High resolution LC-MS/MS experiments were performed on an LTQ-FTICR mass spectrometer (Thermo Electron). Quantification of K4 isoforms was obtained by calculating the peak areas in the MALDI-TOF MS spectra, whereas quantification of K36 and K79 isoforms was performed by analyzing the precursor ions in the nano-LC-MS/MS spectra. The reported values represent an average from four trials.

4.2.3 Assays of In Vitro Activities

The in vitro enzyme assays were performed in the Laboratory of Dr. Tsai (OSU) by Shengjiang Tu. H3 K4, K36, and K79 peptide sequences for the assays are ARTK_{Me1-3}QTARKST, STGGVK_{Me1-3}KPHRY, and IAQDFK_{Me1-3}TDLRFQ. After the reactions for the assays were complete, the samples were spotted for MADLI-TOF MS analysis as described above.

4.2.4 Growth Phenotype Assays

Cells were grown to mid-log phase (OD_{600} = 0.5-1.0). Five microliters of five-fold or ten-fold serial dilutions of each culture were plated. For UV irradiation assays, YP plates were subjected to UV irradiation (Stratalinker UV Crosslinker, Stratagene). For 6-AU sensitivity assays, SC-Uracil plates were supplemented with 30 µg/ml 6-azauracil (6-AU). Plates were photographed after cells were grown for 2-5 days at 30 ºC. A pipetman was used to dilute the cells.
4.3 RESULTS

4.3.1 Global Changes of Histone H3 Modifications in JmjC Knockout Strains

We first studied the overall modifications of histone H3 in WT and the knockout strains by LC-MS. Figure 4.1 shows the MS spectra of the samples. The peak at 15225 Da corresponds to unmodified histone H3; the subsequent 17 modified species are each separated by an interval of 14 Da. Although this series of peaks could represent incremental methylation, a shift of 42 Da (3 × 14 Da) can also arise via acetylation. Before the ambiguity between acetylation and trimethylation is clarified, “methylation equivalents” will be used to describe the mass differences between these peaks.

At first glance, the spectra of the knockout strains seem to have no difference from the WT. To address the H3 modification changes in the five mutants, the relative change of each peak was analyzed. First, the percentage of each peak in a given spectrum was calculated to reflect the abundance of this species in the sample. This percentage was then compared with its counterpart in the WT strain to generate the relative percentage change. For a given peak (i), the following equation was used to calculate the “relative percentage change”:

\[ 100 \times \frac{\text{percentage (mutant, } i\text{) - percentage (WT, } i\text{)}}{\text{percentage (WT, } i\text{)}} \]

As shown in Figure 4.2, a consistent shift in H3 modifications was observed. Relative to WT, except for ecm5Δ, each knockout strain displayed a discernable and reproducible decrease in histone H3 species containing less than 6 methylation equivalents. In contrast,
each knockout strain showed an increase in histone H3 species with 6-17 methylation equivalents. These results suggest that highly modified histones accumulate in the knockout strains.

When analyzed carefully, the pattern of changes in Figure 4.2 provided clues about the substrate specificities of the JmjC proteins. The magnitude of changes followed the order $rph1\Delta > jhd1\Delta \geq jhd2\Delta > gis1\Delta$. Furthermore, the shift of H3 modification patterns in the $rph1\Delta$ and $jhd2\Delta$ strains appeared to be similar; those of $jhd1\Delta$ and $gis1\Delta$ were also similar. These differences were interesting and potentially significant. Focusing on the data of 7-15 methylation equivalents, the $rph1\Delta$ and $jhd2\Delta$ strains showed significant accumulation of 9, 12, and 15 methylation equivalents, whereas the $jhd1\Delta$ and $gis1\Delta$ strains did not show such a pattern. If trimethylation is the predominant state of methylation (which is in fact the case, as described below), then a reasonable interpretation is that Rph1 and Jhd2 are trimethyl lysine demethylases, while Gis1 and Jhd1 are likely mono- or dimethyl lysine demethylases. This was our working hypothesis from this point forward.

To test the above hypothesis, the following experiments were performed: (i) tryptic digestion of histone H3 from WT and the five knockout mutants; (ii) analysis of the digested samples using high-resolution LTQ FT-ICR to identify fragments containing methylated lysine and differentiate the trimethylated fragments from the corresponding acetylated fragments; and (iii) analysis of the relative abundance of unmodified, mono-, di-, and trimethylated species of each fragment.
4.3.2 High-Resolution Mass Spectrometry Analysis of S. cerevisiae Histone Methylation Patterns

To decipher the identities of the H3 modifications that previously described as “methylation equivalents”, high-resolution mass spectrometry of the digested peptides was used. Based on the fact that trimethylation and acetylation have different mass shift and different retention times on RP HPLC, H3 peptides with trimethylation and acetylation were unambiguously assigned. The data demonstrate that H3 K4, K36, and K79 are solely methylated. For the details of acetylation and methylation identification, refer to Chapter 3.
Figure 4.1 LC-MS spectra of H3.
Figure 4.2 Global changes of histone H3 methylation in JmjC knockout strains.
4.3.3 Methylation Changes at Specific Sites in the Knockout Strains

After unambiguous assignment of methylation, quantitative analyses of histone H3 methylation were conducted to compare the abundance of each methylated fragment between WT and the five mutant strains by use of LC-MS/MS and/or MALDI-TOF MS. The tryptic H3 peptide that contained K4 has the sequence, $^3$TKQTAR$^8$. It is a small peptide with a monoisotopic MW of 703.40 Da. Because of the poor fragmentation of singly charged ions and the limited mass range of the LCQ ion trap mass spectrometer (400-2000 m/z), the peptide $^3$TKQTAR$^8$ could not be examined by use of the LCQ ion trap mass spectrometer. Therefore, K4 methylation was studied by MALDI-TOF MS. The MALDI-TOF spectrum of $^3$TKQTAR$^8$ is shown in Figure 4.3. The relative abundance of each isoform of $^3$TKQTAR$^8$ was calculated by taking the peak area of each isoform and dividing it by the sum of peak areas of all K4 isoforms.

In a similar fashion we attempted to use MALDI-TOF MS to characterize the methylation of K36 and K79. However, the peptides of interest, monomethylated $^{28}$SAPSTGGVKKPHR$^{40}$ (MW = 1334.74 Da) and unmodified $^{73}$EIAQDFKTLDR$^{83}$ (MW = 1334.68 Da), are isobaric and interfere with the determination of relative peptide abundances. Therefore, it was not possible to distinguish these two peptides with the MALDI-TOF. As discussed in Chapter 3, peptides with different hydrophobicities have different retention times on RP HPLC. The relative hydrophobicities of $^{28}$SAPSTGGVKKPHR$^{40}$ and $^{73}$EIAQDFKTLDR$^{83}$ are 6.00 and 24.30, respectively. Therefore, the former peptide would elute earlier when separated by RP HPLC. In this manner, these two peptides should be easily distinguished by LC-MS/MS. As expected,
the isoforms of peptide $^{28}\text{SAPSTGGVKKPHR}^{40}$ eluted at 14.90-16.93 min, while the isoforms of peptide $^{73}\text{EIAQDFKTDLR}^{83}$ eluted at 64.66-69.08 min (Figure 4.4). The LC-MS/MS spectra of peptides $^{28}\text{SAPSTGGVKKPHR}^{40}$ and $^{73}\text{EIAQDFKTDLR}^{83}$ are shown in Figure B.2-B.7.

As seen in Figure 4.3 and 4.4, H3 K4, K36, and K79 were predominantly trimethylated in all samples. The unmodified, mono-, and dimethylated K4 and K79 were observed simultaneously with the trimethylated form in all samples. However, due to the low abundance of unmodified K36, the unmodified isoform of $^{28}\text{SAPSTGGVKKPHR}^{40}$ could not be detected in all samples. Quantitative analyses of methylation at K4, K36, and K79 are summarized in Figure 4.5. For a given isoform of peptides $^{28}\text{SAPSTGGVKKPHR}^{40}$ and $^{73}\text{EIAQDFKTDLR}^{83}$, the “percentage of each isoform” is calculated in the same way as for K4 methylation by MALDI-TOF MS. Statistical analysis was performed by using Students’ t test. Although the overall changes relative to WT were small, the major changes were reproducible and consistent. Detailed analyses of the five knockout strains are presented below.

(a) $rph1\Delta$ and $jhd2\Delta$—The $jhd2\Delta$ strain showed an increased level of K4Me3 and decreased levels of K4Me2 and K4Me. Similarly, the $rph1\Delta$ strain showed a higher level of K36Me3 and reduced levels of K36Me2 and K36Me. The data support the prediction from the global analysis that Rph1 and Jhd2 are histone trimethyl lysine demethylases, and further suggest that they target histone H3 K36 and K4, respectively.

(b) $gis1\Delta$ and $jhd1\Delta$—These strains showed modest changes in methylation of K36 only (Figure 4.5B). In contrast to the $rph1\Delta$ strain, the $gis1\Delta$ and $jhd1\Delta$ strains
displayed decreased level of K36Me3 but increased levels of K36Me and K36Me2. Therefore, we predict that Gis1 and Jhd1 are H3 K36Me or K36Me2 demethylases.

(c) ecm5Δ—In agreement with the LC-MS results, the methylation patterns in the ecm5Δ strain did not display significant differences relative to the WT strain. We did not have enough experimental evidence to conclude whether Ecm5 is a histone demethylase.

In summary, we are proposing that four of the five yeast JmjC proteins are histone H3 K4 or K36 demethylases.

Figure 4.3 MALDI-TOF MS spectrum of $^3$TKQTAR$^8$. 
Figure 4.4 LC-MS/MS spectra of peptides $^{28}$SAPSTGGVKKPHR$^{40}$ and $^{73}$EIAQDFKTDLR$^{83}$. RT – retention time.
Figure 4.5 Identification of demethylation sites in knockout strains by MS. *, p < 0.05; #, p < 0.1.
4.3.4 Methylation Changes in RPH1, JHD1, and JHD2 Overexpression Strains

A potential argument about the knockout strains is that the full impact on methylation levels could be masked due to redundant or dominating effects of histone methyltransferases. If an effect complementary to that observed in a knockout strain is observed in the corresponding overexpression strain, a stronger link between a given enzyme and its proposed site of demethylation is established. Therefore, we examined the consequence of individually overexpressing RPH1, JHD1, and JHD2. JmjC protein overexpression was verified by Western blot (Figure B.8). As shown in Figure 4.6, global analysis of histone H3 from the overexpression strains demonstrated that the relative percentages of the lower-methylated species increased, while those of the higher-methylated species decreased. Compared with the patterns observed in the knockout strains (Figure 4.3), the patterns in Figure 4.6 were in the opposite direction and of higher magnitudes, strongly suggesting that these JmjC proteins are directly involved in histone H3 methylation. Specifically, as Figure 4.7A indicates, in the JHD2 overexpression strain, but not in the JHD1 or RPH1 overexpression strains, the K4Me3 level decreased but K4Me level increased significantly. This effect was complementary to that observed in the jhd2Δ strain shown in Figure 4.5A, which supports our previous assertion that Jhd2 catalyzes demethylation of H3K4Me3. Similarly, in Figure 4.7B, histone H3 K36Me3 level in the RPH1 overexpression strain reduced by half, which was consistent with our predication that Rph1 is an H3 K36Me3 demethylase. It is interesting to notice that the H3 K36Me3 level also modestly reduced in the JHD2 overexpression strain (Figure 4.7B). It is possible that overexpression of JHD2 mildly affects histone H3
methylation by a mechanism yet to be established. In Figure 4.7C, similar to Figure 4.5C, the methylation levels at K79 changed very mildly in all strains. Thus no histone H3 K79 demethylases were found here.

As reported previously, the first 21 amino acids of yeast histone H3 were sometimes lost during sample preparation. Within the 21 amino acids, the only possible methylation site is K4. In Figure 4.8, if we take a look at the peaks between 13072 and 13114 Da (heavily modified populations), truncated H3 (residues 22-135) showed similar methylation patterns and intensities in both the WT and JHD2 overexpression strains, indicating that (i) JHD2 does not affect K36 or K79 methylation levels, and (ii) the JHD2-induced difference in the region of 15309-15351 Da in Figure 4.9 is solely the consequence of K4 demethylation, which is consistent with the prediction from the jhd2 knockout and overexpression results shown in Figure 4.6A and Figure 4.7A. Taken together, the data indicate that Jhd2 is an H3K4Me3 demethylase.
Figure 4.6 Global changes of histone H3 methylation in JmjC overexpression strains.
Figure 4.7 Identification of demethylation sites in overexpression stains by MS. *, p<0.05; #, p<0.
Figure 4.8 LC-MS spectra of full length (left) and truncated (right) histone H3 from WT, RPH1 and JHD2. The average MW of unmodified fragment 22-135 of H3 is 12988 Da.
4.3.5 *In Vitro Enzyme Assays of Rph1 Demethylase Activity*

To confirm that it is in fact the JmjC domains that catalyze histone demethylation, *in vitro* assays were performed with recombinant JmjC proteins. To obtain active enzymes, the full-length proteins (from *E. coli* and yeast) and truncated JmjC proteins were purified and tested. When full-length recombinant Rph1, Jhd2, and Jhd1 were overexpressed in *E. coli*, most had poor solubility, so no further assays could be performed. When the full-length proteins were overexpressed and purified from yeast, only Jhd2 showed slight enzyme activities toward K4 peptides. When the truncated recombinant Rph1, Gis1, Jhd2, and Ecm5 purified from *E. coli* were tested, only Rph1 showed robust enzyme activity. For this reason, only the results of Rph1 enzyme assays are reported below.

The truncated Rph1 construct, Rph1-(1-373), which contained the JmjC domain, was overexpressed as a Glutathione-S-transferase (GST) fusion protein, purified and then assayed for demethylase activity. Nine synthetic peptides containing mono-, di-, and trimethylated lysine at K4, K36, and K79 were tested. The abundance of the methylated isoforms was determined by MALDI-TOF MS as described above. The results showed that high concentrations of salt in the sample solutions resulted in ununiform droplets on the MALDI-TOF target and in turn produced poor and un reproducible signals. Therefore, a minimal amount of salt was added to the samples. To further reduce the effect of salt, the sample was washed with 0.1% TFA after the sample was air dried on the MALDI-TOF target.
As predicted, the K36Me3 was readily demethylated to K36Me2, which was further converted to K36Me but with a lower efficiency (Figure 4.9). No activity toward methylated K4 or K79 peptides was observed from the Rph1 enzyme assays. More importantly, the activity of Rph1 toward K36Me3 and K36Me2 was ablated by omission of Fe$^{2+}$ and α-ketoglutarate, addition of EDTA, and replacement of the WT enzyme with the H235A mutant (Figure 4.10). Although the dependence on α-ketoglutarate confirms α-ketoglutarate as a co-substrate, the residual activity observed in the absence of Fe$^{2+}$ could be explained by the presence of trace Fe$^{2+}$ in the purified protein sample. That a metal cofactor is indeed essential for catalysis is suggested by the complete loss of activity in the presence of EDTA. The fact that mutation of residue 235, which based on sequence alignment (Figure B.1B) is predicted to bind Fe$^{2+}$, inactivated Rph1 reinforced the conclusion that it is the JmjC domain of Rph1 that is responsible for catalysis. Collectively, the above data unequivocally demonstrate that Rph1 is an Fe$^{2+}$/α-ketoglutarate-dependent histone lysine demethylase with specificity for H3 K36Me3.
Figure 4.9 *In vitro* enzyme assays of GST-Rph1(1-373). (Left) MALDI-TOF MS data on demethylation of the K36Me2 peptide. (Right) MALDI-TOF MS data on demethylation of the K36Me3 peptide.
Figure 4.10 Control experiments for *in vitro* enzyme assays of GST-Rph1(1-373). Dotted and shaded bars represent the percentages of conversion from K36Me3 and K36Me2 peptides, respectively.
4.3.6 Overexpression of RPH1 Causes a Growth Defect Upon UV Irradiation

As a first step to elucidate the specific biological contexts in which the five *S. cerevisiae* JmjC proteins are requisite, phenotype screening was performed. Under normal conditions, no discernible growth phenotype was observed in any of the individual knockout strains. However, when galactose was used as the carbon source in synthetic medium (SC-Uracil), the growth rate of the *RPH1* overexpression strain dropped significantly (Figure 4.11A, right panel). By contrast, no difference was displayed among these strains in SC-Uracil medium with dextrose (glucose) (Figure 4.11A, left panel). Interestingly, in the rich media (YP glucose or YP galactose), the growth defect was not observed. The results suggested that overexpression of *RPH1* resulted in a growth defect in synthetic media but not in rich media.

To avoid the growth defect observed in synthetic media, the following assays were performed on YP plates. As previously reported, Rph1 is a repressor of PHR1, which encodes for a photolyase in response to UV damage. To determine whether the histone demethylase activities are required in the UV damage responses, UV sensitivity assay was performed. Deletion of individual JmjC genes had no discernible effect on the viability of cells in response to UV irradiation (Figure B.9A). However, overexpression of *RPH1*, but not *JHD1* or *JHD2*, resulted in a growth defect in response to UV irradiation (Figure 4.11B). The phenotype was observed in the range of 5-20 mJ/cm² UV irradiation. To test whether the trimethylated H3K36 demethylase activity is responsible for the phenotype, a demethylase activity-defective mutant, *rph1*<sup>H235A</sup>, was tested in the UV sensitivity assay. Under the same UV irradiation treatment, the *rph1*<sup>H235A</sup> mutant
abolished the phenotype (Figure 4.11C). The results indicated that the demethylase activity is required for the *RPH1* dependent phenotype. Furthermore, MS data (Figure 4.5B) demonstrated that overexpression of *RPH1*, not *JHD1*, reduced the H3K36Me3 level by 50%. Collectively, the data indicate that Rph1 is involved in cell survival in response to DNA damage, likely by changing the level of histone H3K36Me3.

### 4.3.7 *RPH1* and *JHD1* Overexpression Strains Are Slightly Resistant to 6-Azauracil

The methylation status of H3K36 is associated with transcriptional elongation. Previous studies showed that deletion of Set2, the histone H3K36 methyltransferase, results in a modest resistance to 6-azauracil (6-AU) treatment. 6-AU, a commonly used drug for testing transcription elongation defects, inhibits nucleotide metabolism and leads to depletion of cellular GTP and UTP. Nucleotide depletion affects elongation efficiency. To test whether the JmjC proteins are involved in transcription elongation, we performed 6-AU sensitivity assays. Preliminary studies showed that overexpression of *RPH1* or *JHD1* resulted in slight resistance to 6-AU (Figure B.9B). The result was only observed in the strains overexpressing H3K36 demethylase (Rph1 or Jhd1), but not in the strain overexpressing H3K4 demethylase (Jhd2). The data are consistent with previous observations that deletion of H3 K36 methyltransferase gene *SET2* resulted in a modest resistance to 6-AU. Our data support the idea that demethylation of H3K36 by Rph1 and Jhd1 acts to decrease the methylation level of H3K36 and to interfere with transcription.
Figure 4.11 Growth phenotypes of the JmjC overexpression strains. A, Cell were diluted tenfold and spotted on SC-Uracil plates with 2% dextrose (left) or galactose (right). B, Cells were grown in rich medium (YP) with 2% glucose (DEX) or galactose (GAL) and exposed to UV irradiation (10 mJ/cm²). C, The rph1 deletion strain (rph1Δ) was transformed with plasmids containing no insert (Vector), or GAL1 promoter driven wild type (RPH1), and activity-deficient mutant (rph1H235A), respectively. UV sensitivity assays were performed as described in B.
4.4 DISCUSSION

4.4.1 JmjC Domain Encodes Histone Demethylase Activities

While our work was in progress, YER051W, named as JHD1 later, was shown to encode a yeast H3 K36 demethylase \(^{234}\) and was the only yeast JmjC-containing histone demethylase reported before this study was published. It is not clear whether Jhd1 can demethylate H3K36Me3. Our \textit{in vivo} MS data predict that Jhd1 is an H3K36Me2 or H3K36Me demethylase. Our data also suggest that Rph1, Jhd2, and Gis1 are demethylases. Among them, Jhd2 and Rph1 are trimethyl lysine demethylases with specificity on H3 K4 and K36, respectively; Gis1 is predicted to be similar to Jhd1, acting as an H3 K36Me or K36Me2 demethylase. The specific demethylase activities of Rph1 towards histone H3K36 have been verified by \textit{in vitro} enzyme assays. It is the JmjC domain that is responsible for Rph1 demethylase activity. We also observed very weak but reproducible demethylase activities of Jhd2 and Gis1 with specificity on K4 and K36 peptides, respectively. In summary, four out of the five yeast JmjC proteins (Jhd1, Rph1, Jhd2, and Gis1) are histone demethylases.

As for human enzymes, several groups independently reported six JmjC proteins to be H3 demethylases. While JHDM1 and JHDM2A demethylate mono- and dimethyl H3 K36 and K9 respectively \(^{234, 236}\), JMJD2 family histone demethylases (JMJD2A, JMJD2B, JMJD2C, and JMJD2D) demethylate trimethyl H3 K9/K36 \(^{235, 237-239}\). It is interesting to note that each of these six JmjC proteins has a zinc finger, PHD, or TUDOR domain. This matches one of our searching criteria and implies a role for each of these proteins in transcription.
In summary, by using different approaches, we and other groups have reached the same conclusion that JmjC domain encodes histone demethylase activity. Our study is the first systematic MS proteomic report on all of JmjC proteins in one single organism. Our \textit{in vivo} MS analysis results matched \textit{in vitro} enzyme assay results. In the knockout MS data (Figure 4.5), the changes were relatively small in comparison with those in the overexpression data (Figure 4.7). Phenotype screening assays demonstrated that overexpression of \textit{RPH1} results in phenotypes which were not observed in the deletion strains. The differences suggest that deletion strains have a smaller impact on histone H3 methylation. Given that all three sites (H3 K4, K36, and K79) are heavily methylated (more than 60%), we believe that in yeast, histone methyltransferases are dominating. Furthermore, localized effects of the demethylases may not contribute to changes in global MS data. In spite of these effects, small but consistent changes were observed. The demonstrated agreement between \textit{in vitro} activity assays and \textit{in vivo} MS data for Rph1 provides solid support for the reliability of our approach. This approach could be applied in other organisms to identify histone demethylases.

\textbf{4.4.2 Rph1 is an H3K36Me3 Demethylase Involved in DNA Damage Response and Transcription Elongation}

In this study, \textit{in vivo} MS data indicated that Rph1 is an H3K36Me3 demethylase, which was verified by \textit{in vitro} assays. Based on the \textit{in vivo} and \textit{in vitro} analysis, it is safe to conclude that it is the JmjC domain in Rph1 that demethylates H3K36Me3 and H3K36Me2. The \textit{in vitro} assays did not show any detectable activity of Rph1 towards the
H3 K4 and K79 peptides. To the best of my knowledge, Rph1 is the first yeast H3K36Me3 demethylase that has been reported. The phenotype screening experiments demonstrated that RPH1 overexpression strain is sensitive to UV irradiation and the catalytic activity of Rph1 JmjC domain is responsible for the defect. The phenotype indicated that Rph1, an H3K36Me3 demethylase, is involved in DNA damage responses. Consistently, it has been reported that H3K36 methyltransferase Set2 is involved in DNA damage response \(^{229}\). Furthermore, the UV irradiation defect caused by RPH1 overexpression strain was not observed in JHD1 overexpression strain. The difference can be explained by our prediction that Jhd1 is not an H3K36Me3 demethylase while Rph1 is. These results indicate that UV damage responses are correlated to H3K36 methylation levels, particularly on the level of H3K36Me3.

As shown in Figure 4.7B, the K36Me3 level decreased sharply in RPH1 overexpression strain. The decrease in the global level (Figure 4.6) suggests that overexpression of RPH1 has global effects. Since the link between H3K36 methylation and transcription in yeast is well established and H3K36 methyltransferase Set2 has been shown to interact with RNA polymerase II and to play important roles in transcription elongation \(^{220, 256, 257, 259, 260}\), H3K36 demethylase Rph1 might affect transcription elongation. Indeed, when RPH1 was overexpressed, there was a slight 6-AU phenotype, suggesting its participation in transcription elongation. Furthermore, another H3K36 demethylase Jhd1 showed the same phenotype as Rph1. The H3K36 demethylases Rph1 and Jhd1 might shift H3K36 to less methylated species and result in lower transcription activity. This could be the reason that they are mildly resistant to 6-AU.
4.4.3 Jhd2 is an H3K4 Demethylase

No JmjC domain containing H3K4 demethylase was reported prior to this study. The in vivo MS analysis presented here suggests that Jhd2 is an H3K4 demethylase. A weak H3K4 demethylase activity was observed in in vitro assays. Jhd2 is a nuclear protein with unknown function. A recent large scale yeast protein-protein interaction study identified three Jhd2 interacting proteins Fun30, Sas10, and YIL091C. Sas10 overexpression results in derepression of mating type genes at HML and HMR. Fun30 is a putative ATP-helicase and homolog of Snf2. Overexpression of Fun30 affects chromosome stability and integrity. Interestingly, the Snf2 family yeast protein Isw1, an ATP-dependent chromatin-remodeling factor, has been shown to be recruited to active genes by H3K4 methylation. These studies suggest that Jhd2 could be a component of a histone modification and chromatin remodeling complex related to silencing. This study demonstrates that Jhd2 is a novel histone demethylase acting on H3K4.

Methylation on H3K4 has many interesting consequences. On one hand, K4 trimethylation occurs in the promoter region of genes, and is involved in transcription activation. On the other hand, K4 methylation mediates rDNA and telomere silencing. The H3K4 methyltransferase Set1 is required for silencing at the rDNA region, telomeres, and HML mating type loci. Therefore, Jhd2 might negatively regulate transcription, or repress silencing on the rDNA region, telomeres, or the HML mating type loci.
4.4.4 Possible H3K79 Demethylases

Of the three known lysine methylation sites on histone H3 (K4, K36, and K79), only K4 and K36 have been shown to be affected by demethylases. Since H3K79 is highly (90%) methylated in yeast, small changes in the demethylase knockout mutants could be masked by the relatively high histone methyltransferase activities. Therefore, it can not be ruled out that Ecm5 or a non-JmjC protein may be an H3K79 demethylase. Further investigation is needed to resolve the question.

4.4.5 Yeast JmjC Proteins in Transcription

Of the five JmjC proteins in *S. cerevisiae*, Jhd1, Jhd2, and Ecm5 each has a PHD domain while Rph1 and Gis1 have zinc finger domains. These domains might be associated with transcription regulations related to histone demethylation. Rph1 and Gis1 have been reported to act as repressors. Rph1 binds to a promoter element, URS_{PHR1}, likely via the zinc finger domains. It is speculated that the PHD domains in Jhd1 or Jhd2 may bind to chromatin via methylated histone or DNA. Genome-wide location analysis by use of CHIP-chip of the JmjC proteins will provide details regarding their roles in transcription.

4.5 CONCLUSIONS

Based on the prediction that histone lysine demethylases may contain the JmjC domain, we examined the methylation patterns of five knockout strains (*ecm5Δ, gis1Δ, rph1Δ, jhd1Δ*, and *jhd2Δ (yjr119cΔ)*) and three overexpression strains (*RPH1, JHD1, and*
JHD2) of *S. cerevisiae*. Mass spectrometry analysis of histone H3 showed increased modifications in all mutants except for *ecm5Δ*. The relative abundance of specific fragments indicated that histone K36Me3 and K4Me3 accumulated in *rph1Δ* and *jhd2Δ* strains, respectively, while both histone K36Me2 and K36Me accumulated in *gis1Δ* and *jhd1Δ* strains. The methylation patterns in overexpression strains were the reverse of those obtained in the complementary knockout strains. *In vitro* enzymatic assays confirmed that the JmjC domain of Rph1 primarily demethylates K36Me3 and K36Me2 secondarily. Overexpression of *RPH1* generated a growth defect in response to UV irradiation. Collectively, in addition to Jhd1, our results identified three novel JmjC domain containing histone demethylases (Gis1, Rph1, and Jhd2) and their sites of action in *S. cerevisiae*. 
CHAPTER 5

LABEL-FREE MASS SPECTROMETRY DETERMINATION OF HISTONE CROSS-TALK IN SACCHAROMYCES CEREVISIAE

5.1 INTRODUCTION

Histones are the chief proteins in chromatin. The histone tails that extend from nucleosomes are subject to a variety of posttranslational modifications (PTMs), such as acetylation, methylation, phosphorylation, and ubiquitination. Different combinations of PTMs are believed by some to represent a form of histone code. These PTMs play important roles in chromatin remodeling and transcriptional regulation. They also influence histone and DNA interactions as well as histone / histone, and histone / non-histone interactions. Studies have showed that some PTMs are functionally inter-related. Modifications to one site may reduce or enhance the modification at another site. Such correlation (or cross-talk) between distinct modifications has been shown to occur on the same or different histone tail(s).

Acetylation and methylation are the most widely studied histone PTMs. They play indispensable but distinct roles in gene activities. For example, methylation of H3K9 facilitates heterochromatin-associated protein binding and results in gene silencing while acetylation of H3K9 results in gene activation. Attempts have been
made to reveal the correlation between methylation and acetylation. Studies showed that methylation of H3K4 is associated with hyperacetylation of H3\(^{106, 185, 220}\), while methylation of H3K27 and K9 is associated with hypoacetylated H3\(^{284}\).

Unlike the other sites of acetylation in yeast, K56 is not within the N-terminal tail of H3 and is the last residue on the α-helix \(^{285}\). K56 acetylation (K56Ac) occurs in S phase and largely disappears in G2/M phase in the cell cycle\(^{286-289}\). It is estimated that 20% of H3 could be acetylated at K56 in asynchronous cells\(^{285}\). The modification has been implicated in regulating replication\(^{221, 290}\) and the alteration in H3K56 acetylation displays increased sensitivity toward certain DNA-damaging agents \(^{26, 291-297}\). Though recent studies have offered a great amount of insights into the function of K56Ac, few attempts have been made to reveal the cross-talk between K56Ac and other PTMs such as methylation of lysine.

Previous studies showed that four of the five JmjC domain-containing proteins are histone H3 demethylases specific for certain residues: Rph1 for di- and trimethylation of both K9 and K36\(^{179, 298}\), Jhd1 for K36 dimethylation\(^{179, 234}\), Jhd2 for K4 trimethylation\(^{179, 299}\), and Gis1 for mono- or dimethylation of K36\(^{179}\). These enzymes not only demethylate H3 but play important roles in biological processes\(^{300-303}\). For example, Rph1 and Gis1 are DNA damage-responsive repressors of PHR1\(^{255}\), and Jhd1 is involved in methylated histone residue binding\(^{304}\). The molecular function of Ecm5 remains unknown and it is desirable to study its effects on methylation levels together with the other four JmjC domain-containing proteins. To the best of our knowledge, no studies
have been published on the change of H3K56Ac in response to alteration of the JmjC proteins.

With the development of ESI and MALDI, MS has been widely used in the characterization of histones and their PTMs. It bypasses the limitations of antibodies such as availability, site specificity, and cross reactivity. Furthermore, MS offers an accurate and efficient means of discovering novel PTMs. A large number of mass spectrometry-based approaches have been developed for histone PTM quantification. Due to the simplicity, efficiency, and low cost, label-free quantitative approaches have been widely used. Internal standard-based label-free quantitation employs an endogenous or exogenous standard monitoring the fluctuation of instrument conditions and variances during sample preparation. Usually these approaches require pre-knowledge of the identity and/or quantity of the studied peptides. Even so, sometimes it is difficult to find peptides that serve as the standard. Therefore, label-free and internal standard-free quantitation is desirable. In using these approaches, sample processing, instrument performance, and data analysis need to be under strict control. At least three MS runs are carried out for each sample to obtain meaningful statistics. Recent research demonstrated that mass spectral peak areas of peptide ions correlated well with protein abundances. Therefore, direct quantitation from LC-MS raw chromatograms enables fast access to multi-replicate analysis. Studies also suggested that spectral counts relate well to protein abundance. Each time when a peptide from a particular protein is fragmented and identified, it is termed a “hit”. The spectral count is the sum of all of these hits including redundant spectra with different charge states. This approach expands
the dynamic range over isotopic labeling approaches \textsuperscript{311,315} and offers a simple means to measure the relative abundance of proteins in a mixture.

In this chapter, mass spectrometry was employed to determine histone cross-talk among H2A, H2B, H3, and H4 in yeast resulting from knocking out and overexpressing the JmjC domain-containing proteins. We observed distinct changes in acetylation of H3K56 as measured by spectral counting and spectral peak abundance approaches.

5.2 EXPERIMENTAL

5.2.1 Yeast Cell Culture and Media

All \textit{S. cerevisiae} strains - wild type (WT), knockout strains (\textit{gis1}\Delta, \textit{ecm5}\Delta, \textit{rph1}\Delta, \textit{jhd1}\Delta, and \textit{jhd2}\Delta), and overexpression strains (\textit{RPH1}, \textit{JHD1}, and \textit{JHD2}) - were provided by Dr. Tsai (\textit{Chapter 4}). The cells were grown as described in \textit{Appendix C.6}. In brief, for the knockout strains 10 ml of YPD media was inoculated with a freshly streaked single colony and grown at 30 °C and 200 rpm till OD\textsubscript{600} = 1.7–1.9. The culture was then diluted into 200 ml of YPD media and incubated at 30 °C and 200 rpm till OD\textsubscript{600} = 1.1–1.3. Cells were harvested for histone extraction by centrifugation at 6000 g and 4 °C for 10 min.

For overexpression strains, 10 ml of SC-Uracil media containing 2\% glucose (SD media) was inoculated with a freshly streaked colony line and incubated at 30 °C and 200 rpm till OD\textsubscript{600} ≥ 1.1. The culture was then diluted into 15 ml of SD media and grown till OD\textsubscript{600} ≥ 1.3. At this point, aliquots from the 15 ml starter cultures were diluted into 200 ml of SC-Uracil media with 2\% sucrose to give an initial OD\textsubscript{600} of 0.02. Incubation was
continued until an OD$_{600}$ of 1.0–1.2 was reached. Cells were pelleted by centrifugation at 7000 g and 4 °C for 10 min and then washed with 40 ml of ice-cold sterile water to remove the sucrose. The cells were then resuspended in 200 ml of SC-Uracil media with 2% galactose, incubated at 30 °C with agitation at 200 rpm for 4 h, and finally harvested. Efforts were made to overexpress the five JmjC proteins individually. However, overexpression of Ecm5 and Gis1 was not successful due to hindered growth. Thus, only overexpression strains $RPH1$, $JHD1$, and $JHD2$ were studied.

5.2.2 Histone Preparation

Yeast histones were extracted as described in Chapter 3 and separated on 16.5% Tris-Tricine-SDS gels (Bio-Rad, Hercules, CA) as described in Chapter 2. Gel bands representative of H3 were excised and trypsin digested as described in Chapter 3. To remove impurities and eliminate potential variances in sample preparation, in-gel digestion was conducted by use of a 96-well C$_{18}$ ZipPlates (Millipore, Bedford, MA).

5.2.3 Liquid Chromatography-Mass Spectrometry (LC-MS) Analysis

The LC-MS experiment was conducted on a Q-TOF instrument (Micromass, Wythenshawe, UK) coupled with an Ultimate 3000 HPLC system (Dionex, Sunnyvale, CA). The Q-TOF was used in positive ion mode at conditions: source temperature = 100 °C, resolution = 8000, and cone voltage = 25 V. NaI was used for external mass calibration over the mass range of 500-2500 m/z. The spectra were collected in continuum mode at a rate of 2 scan/sec and analyzed by using MassLynx 4.0 software.
Histones were separated on a Discovery Bio wide pore C\textsubscript{18} column (150 × 1 mm, 5 µm, 300 Å, Supelco, Park Bellefonte, PA) at a flow rate of 25 µl/min using a gradient as shown in Table 2.2. 0.05% TFA in water and acetonitrile were used as mobile phases A and B, respectively. 10 µg of histone mixture were injected. The column was washed thoroughly after each sample and then equilibrated at 20% B for at least 30 min. Three LC-MS runs were carried out for each knockout and overexpression strain.

5.2.4 Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS) Analysis

The tryptic digest was subject to nano-LC-MS/MS analysis by use of an LTQ-Orbitrap mass spectrometer (Thermo Finnigan, San Jose, CA) coupled with an Ultimate 3000 HPLC system (Dionex, Sunnyvale, CA). Three LC-MS/MS runs were carried out for each knockout and overexpression strain. The electrospray was facilitated by use of an Advance-LTQ/Orbitrap/FT MS spray source (Michrom Bioresources, Auburn, CA) and spray tip (I.D.40 µm, Michrom). 0.1% formic acid in water and acetonitrile were used as mobile phases A and B, respectively. Peptides were preconcentrated on a trap column (300 µm × 5 mm, 100 Å, Dionex) for 10 min with an isocratic loading buffer (0.03% TFA and 2% acetonitrile in water) at a flow rate of 20 µl/min and then online separated on a C\textsubscript{18} column (150 × 0.1 mm, 3 µm, 200 Å, Michrom) at a flow rate of 2 µl/min. A 70-min gradient was used: mobile phase B was increased linearly from 2 to 55% in 40 min, then increased to 90% in 2 min, held at 90% for 2 min, then returned to 2% to equilibrate the column for 20 minutes. The column was washed thoroughly.
between each run to minimize carryover. One microliter of the digest was injected onto the column.

The mass spectrometer was operated in positive ion mode at the following conditions: resolution = 30,000, normalized collision energy = 35 V, activation Q = 0.25, activation time = 30 ms, and isolation width = 1.0 m/z. The electrospray voltage was maintained at 2.0 kV and capillary temperature was 175 °C. Mass calibration was performed immediately prior to running the samples over the mass range 150-2000 (m/z) by using a standard peptide solution containing Caffeine, Bradykinin fragment 1-5, Angiotensin I, Melittin, and Glu-Fibrinopeptide (Sigma-Aldrich, St. Louis, MO). The average mass accuracy for the calibration solution was < 3 ppm. Ten data-dependent MS/MS scans with dynamic exclusion were carried out between each full MS scan. The full MS scan spectra were collected in profile mode using FTMS as the analyzer. The dependent MS scan spectra were collected in centroid mode using ion trap as the analyzer. Peptides generated from knockout and overexpression strains were run in a batch along with the WT and standard bovine histone digests. The spectra obtained were searched against in-house developed software, MassMatrix 217. The search parameters included the following variable modifications: acetylation of lysine and N-terminal, methylation of lysine and arginine, ±10.00 ppm mass tolerance for precursor ions, ±1.00 Da mass tolerance for product ions, pp value of output ≥ 6.0, pp tag of output ≥ 2.0. Each of the tandem MS spectra matched by the database search was manually validated.
5.3 RESULTS AND DISCUSSION

5.3.1 Global Changes of Histone H3 Acetylation

Our previous study of yeast histones indicated that the five JmjC domain-containing proteins affect lysine methylation on H3. When we take a closer look at Figure 4.2 and 4.6, there is evidence that acetylation levels on H3 are also affected. As seen in Figure 4.3, compared to WT, the acetylation levels on H3 appear to change significantly in knockout stains \textit{gis1}\Delta, \textit{rph1}\Delta, \textit{jhd1}\Delta, and \textit{jhd2}\Delta. The effect of \textit{ecm5}\Delta on acetylation is trivial. \textit{gis1}\Delta and \textit{jhd1}\Delta have approximately the same magnitude of changes in acetylation; \textit{rph1}\Delta has the most significant changes; the changes of \textit{jhd2}\Delta are between those of \textit{rph1}\Delta and \textit{gis1}\Delta. In addition, in all of the strains compared to WT, the level of monoacetylation decreased while the levels of tri-, tetra-, and penta-acetylation increase. It is interesting to note that these proteins affect the acetylation isoforms to different degrees as listed in Table 5.1. As seen in Figure 4.6 the acetylation pattern for the overexpression strains \textit{RPH1}, \textit{JHD1}, and \textit{JHD2} are opposite to those from the knockout strains. This observation suggests that \textit{RPH1}, \textit{JHD1}, and \textit{JHD2} affect acetylation on H3. For some reason, the magnitude of acetylation changes in overexpression strain \textit{JHD2} is at much higher levels than those in \textit{RPH1} and \textit{JHD1}.

Yeast histone H3 is known to be acetylated on K9, K14, K18, K23, K27, and K56. It is expected that these sites are acetylated nonrandomly. However, no research on the acetylation sequence of yeast histone H3 has been reported, though studies on human H3 have been undertaken\textsuperscript{177}. From the LC-MS data it is hard to determine which peak in
Figure 4.2 and 4.6 reflects a change for any given site of acetylation. However, the results suggest that on yeast histone H3 methylation does correlate with acetylation.

**Table 5.1** Changes of H3 acetylation levels in JmjC knockout strains.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Change of acetylation</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>gis1Δ</em></td>
<td>Ac2 &lt; Ac1 ≈ Ac3 ≈ Ac4 &lt; Ac5</td>
</tr>
<tr>
<td><em>rph1Δ</em></td>
<td>Ac2 &lt; Ac1 &lt; Ac4 &lt; Ac3 &lt; Ac5</td>
</tr>
<tr>
<td><em>jhd1Δ</em></td>
<td>Ac2 &lt; Ac1 &lt; Ac4 &lt; Ac3 &lt; Ac5</td>
</tr>
<tr>
<td><em>jhd2Δ</em></td>
<td>Ac2 &lt; Ac1 &lt; Ac4 &lt; Ac5 &lt; Ac3</td>
</tr>
</tbody>
</table>

5.3.2 *Enzyme Digestion of Histones*

A variety of enzymes have been used to generate peptides from proteins. Arg-C is commonly used to cleave at the carboxyl side of arginine residues and leaves both the unmodified and modified lysines intact in the peptide. Unfortunately, compared to the most commonly used enzyme trypsin, Arg-C is much more expensive. Therefore, chemical derivatization is performed to modify lysine so that trypsin would have limited cleavage at lysine and arginine. Acetic anhydride adds one acetyl moiety to the unmodified lysine so that trypsin would not be able to cleave at the C-terminal of the unmodified lysine. Propionic anhydride adds a propinyl group to the unmodified and monomethylated lysine so that trypsin does not cleave at the C-terminal of the unmodified lysine nor the monomethylated lysine. Therefore, through
these derivatizations desired peptides can be generated by using the cheaper enzyme trypsin instead of expensive Arg-C. However, these derivatizations limit the cleavage of trypsin only to a certain degree and the digestion does not resemble Arg-C completely. It is believed by some researchers that propionic derivatization restricts trypsin digestion to the C-terminal side of arginine residues \(^{214, 320}\). However, studies showed that dimethylation of lysine does not prevent tryptic cleavage at the C-terminal of lysine \(^{107, 321, 322}\).

The digestion completeness is always a concern in peptide quantitation, especially for quantification of lysines that can be posttranslationally modified. The ideal condition for quantification of a specific residue is that the residue is contained in only one peptide. However, usually enzyme digestion, even after overnight incubation, is not complete so that the studied residue is contained in more than one peptide \(^{107, 189, 195, 216, 323}\). Thus, quantification of such a residue is challenging. Quantification of lysine residue is even more difficult when peptides are generated by trypsin or Arg-C. Unfortunately, no reports have addressed this concern.

In my studies, LC-MS/MS spectra of both knockout and overexpression strains were analyzed to evaluate the change levels of H3K56Ac in response to the JmjC proteins. Yeast H3K56 was found to be represented by several tryptic peptides as shown in Table 5.2. As seen in the table, complete digestion of yeast H3 gave peptide \(^{54}\text{FQK}_{\text{Ac}}\text{STELLIR}^{63}\) or \(^{57}\text{STELLIR}^{63}\); incomplete digestion could result in several other peptides: \(^{41}\text{YKPGTVALREIRRFQ}^{56}\), \(^{50}\text{EIRRFQKSTELLIR}^{63}\), \(^{53}\text{RFQKSTELLIR}^{63}\), \(^{53}\text{RFQKSTELLIRK}^{64}\), \(^{54}\text{FQKSTELLIRK}^{64}\), and \(^{54}\text{FQKSTELLIRKLPFQR}^{69}\). In Table
5.2, peptides not containing K56 (#8-11) due to its C-Terminal cleavage are also listed. The rational for their inclusion is that when K56 is not modified, trypsin cleavage at the C-terminal of K56 would yield peptides $5^4FQK^{56}$ or $5^3RFQK^{56}$ (For the complete sequence of yeast H3, see Figure S5.1). These two peptides are hydrophilic and too small to be detected reliably. Therefore, peptides $5^7STELLIR^{63}$, $5^7STELLIRK^{64}$, $5^7STELLIRKLPFQR^{69}$, and $5^7STELLIRKLPFQRLVR^{72}$ are indicators of unmodified K56. To truly reflect the changes of H3K56Ac all of the peptides listed in Table 5.2 are to be used for quantitation.

<table>
<thead>
<tr>
<th>Peptide #</th>
<th>Peptide Sequence</th>
<th>K56Ac State</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$5^0EIRRFQKSTELLIR^{63}$</td>
<td>0, 1</td>
</tr>
<tr>
<td>2</td>
<td>$5^3RFQKSTELLIR^{63}$</td>
<td>0, 1</td>
</tr>
<tr>
<td>3</td>
<td>$5^3RFQKSTELLIRK^{64}$</td>
<td>0, 1</td>
</tr>
<tr>
<td>4</td>
<td>$5^4FQKSTELLIR^{63}$</td>
<td>0, 1</td>
</tr>
<tr>
<td>5</td>
<td>$5^4FQKSTELLIRK^{64}$</td>
<td>0, 1</td>
</tr>
<tr>
<td>6</td>
<td>$5^4FQKSTELLIRKLPFQR^{69}$</td>
<td>0, 1</td>
</tr>
<tr>
<td>7</td>
<td>$4^1YKPGTVALREIRRFQK^{56}$</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>$5^7STELLIR^{63}$</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>$5^7STELLIRK^{64}$</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>$5^7STELLIRKLPFQR^{69}$</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>$5^7STELLIRKLPFQRLVR^{72}$</td>
<td>0</td>
</tr>
</tbody>
</table>
5.3.3 **Quantification of H3K56Ac by Spectral Counting**

Based on the notion that spectral counts correlate with protein abundances, it is safe to rationalize that spectral counts of peptides listed in Table 5.2 reflect the abundance of K56 isoforms. To investigate the relative change of K56Ac in response to the JmjC proteins, MS/MS spectra for the peptides representative of unmodified (Ac0) and acetylated (Ac1) K56 were counted for each knockout and overexpression strain. The percentage of K56Ac1 (P) was calculated by dividing the spectral counts of K56Ac1 by the sum of spectral counts of K56Ac0 and K56Ac1. The relative change of K56Ac in the knockout and overexpression strains was compared to WT following the equation 5.1. The spectral counting results are shown in **Figure 5.1**.

\[
\text{Relative change} = \frac{P(\text{knockout/overexpression}) - P(\text{WT})}{P(\text{WT})} \times 100 \quad \text{eqn 5.1}
\]

As seen in Figure 5.1, *ecm5Δ* had little effect on K56Ac as Figure 4.2 suggested. In contrast, the other strains affect K56Ac to different degrees. Ideally, if the JmjC domains affect K56Ac, knockout and overexpression strains would demonstrate opposite effects as seen from Jhd1. However, due to the variability, results for the knockout and overexpression strains of Rph1 and Jhd2 do not corroborate this prediction. The sums of the spectral counts for three runs of each strain are listed in **Table 5.3**. As seen, spectral counts for peptides representative of K56 are small numbers even for three runs. In this manner, this approach can only provide an estimation of K56Ac changes and could not give an accurate ratio as other researchers have experienced \(^{309, 311, 312, 315}\). Therefore, by
this approach it was not possible to conclude exclusively the effects of the JmjC proteins on K56Ac. To give an accurate change of K56Ac, peak area abundance was employed and the results are presented below.

Table 5.3 Spectral counts of peptides representative of unmodified (Ac0) and acetylated (Ac1) K56 in knockout and overexpression strains. WT1 - WT cultured for gis1Δ, rph1Δ, and jhd1Δ; WT2 - WT cultured for ecm5Δ and jhd2Δ; WT3 - WT cultured for RPH1, JHD1, and JHD2.

<table>
<thead>
<tr>
<th></th>
<th>WT1</th>
<th>gis1Δ</th>
<th>rph1Δ</th>
<th>jhd1Δ</th>
<th>WT2</th>
<th>ecm5Δ</th>
<th>jhd2Δ</th>
<th>WT3</th>
<th>RPH1</th>
<th>JHD1</th>
<th>JHD2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ac0</td>
<td>37</td>
<td>27</td>
<td>35</td>
<td>31</td>
<td>67</td>
<td>58</td>
<td>40</td>
<td>62</td>
<td>50</td>
<td>58</td>
<td></td>
</tr>
<tr>
<td>Ac1</td>
<td>22</td>
<td>11</td>
<td>18</td>
<td>14</td>
<td>24</td>
<td>20</td>
<td>15</td>
<td>26</td>
<td>23</td>
<td>14</td>
<td>27</td>
</tr>
</tbody>
</table>

5.3.4 Quantification of H3K56Ac by Peak Area Abundance

As stated above, trypsin digestion of yeast histone H3 resulted in eleven peptides representative of K56. All of these peptides were taken into consideration to quantify the change of K56Ac. The percentage of K56Ac1 was calculated (eqn 5.2) by dividing the peak abundance of K56Ac1 over the sum of abundances for K56Ac0 and K56Ac1 from the eleven peptides. The relative change of K56Ac in the knockout and overexpression strains was compared to the WT following the eqn 5.1. The results are shown in Figure 5.2. As seen in Figures 5.1 and 5.2, the K56Ac change in response to Gis1, Ecm5, and Jhd1 strains displays the same trend by measurement of peak abundance and spectral counting with the exception of rph1Δ. Interestingly, it seems that Jhd2 does not significantly affect K56Ac because the relative change of K56Ac1 is positive in both
knockout and overexpression strains. Taking together the results obtained from spectral counting and peak area abundance, we concluded that K56Ac1 level decreases upon knocking out Gis1 while K56Ac1 level change minimal upon knocking out protein Ecm5. Upon overexpression of Rph1 K56Ac1 decreased while overexpression of Jhd1 caused an increase in K56Ac1. Jhd2 showed little to no effect on K56Ac. Given that Gis1 is a demethylase specific for K36, Rph1 for K9 and K36, Jhd1 for K36, and Jhd2 for K4, the results shown here imply that methylation (Me) of K9 and K36 correlate with K56Ac while K4Me has no significant correlation with K56Ac.

\[
P = \frac{\sum \text{Peak abundance (Ac1)}}{\sum \text{Peak abundance (Ac0)} + \sum \text{Peak abundance (Ac1)}} \times 100 \quad \text{eqn 5.2}
\]

One of the challenges in peptide quantitation arise when multiple peptides representative of the same residue are present in the sample. Incomplete enzymatic digestion is not uncommon. Ideally, all of those peptides must be taken into consideration when quantifying changes at a specific residue. In several reports on peptide quantitation, only one peptide representative of the residue is used for quantitation \textsuperscript{179, 215, 320, 323}. No discussions as to the effect of incomplete digestion are made. Other reports have been made on the use of selected peptides to quantitate K56Ac using only fragment \textsuperscript{53}RFQKSTELLIR\textsuperscript{63} \textsuperscript{320}. To examine the effect of using selected peptides we determined K56Ac changes by use of peak abundances of peptides \textsuperscript{53}RFQKSTELLIR\textsuperscript{63} and \textsuperscript{54}FQKSTELLIR\textsuperscript{63} individually. Both of these two peptides are observed in tryptic and Arg-C digest \textsuperscript{320, 323} even after overnight digestion at 37 °C. The results are shown in
**Figure 5.3** and **5.4.** As seen from Figure 5.3 and 5.4, except for strains *RPH1* and *JHD2*, quantification of K56Ac relative changes by using peptides $^{53}$RFQKSTELLIR$^{63}$ and $^{54}$FQKSTELLIR$^{63}$ alone show opposite trends. For example, in strain *jhd1Δ* K56Ac1 level decreases when peptide $^{54}$FQKSTELLIR$^{63}$ was used for quantification while increases when peptide $^{53}$RFQKSTELLIR$^{63}$ was used. Thus, it is dangerous to use just a single peptide for quantitation when multiple peptides represent the specific residue, especially when the study is carried out with mutant strains alone$^{215, 320, 323}$.

$^{53}$RFQKSTELLIR$^{63}$ and $^{54}$FQKSTELLIR$^{63}$ are the most frequently observed and reported peptides when enzymes Arg-C and trypsin are used for digestion. In most cases, these two peptides are the dominant fragments among the K56 related peptides. In this study, we were also attempted to quantify K56Ac by using these two peptides together out of the eleven. The percentage of K56Ac1 was calculated by peak abundance of K56Ac1 over the sum of peak abundances of K56Ac0 and K56Ac1 from these two peptides by following equation 5.2. The relative change of K56Ac in the knockout and overexpression strains is compared to the WT following equation 5.1. The results are shown in **Figure 5.5.** As seen in the figure, when both of the peptides $^{53}$RFQKSTELLIR$^{63}$ and $^{54}$FQKSTELLIR$^{63}$ are used for quantification, the results between the three runs of each strain have a better precision. Compared to the results shown in Figure 5.2, it can be seen that the trend of K56Ac upon alteration of the JmjC proteins are the same. This observation suggests that $^{53}$RFQKSTELLIR$^{63}$ and $^{54}$FQKSTELLIR$^{63}$ are better indicators for K56Ac than either alone. However, to truly reflect the changes of K56Ac all peptides representative of K56 should be considered.
Figure 5.1 H3K56Ac changes in knockout (top) and overexpression (bottom) strains in response to JmjC proteins quantified by spectral counting. * - only two trials of data and error bars are not available.
Figure 5.2 H3K56Ac changes in knockout (top) and overexpression (bottom) strains in response to JmjC proteins quantified by peak abundance of all K56 related peptides. * - only two trials of data and error bars are not available.
Figure 5.3 H3K56Ac changes in knockout (top) and overexpression (bottom) strains in response to JmjC proteins quantified by peak abundance of peptide $^{54}$FQKSTELLIR$^{63}$. * only two trials of data and error bars are not available.
Figure 5.4 H3K56Ac changes in knockout (top) and overexpression (bottom) strains in response to JmjC proteins quantified by peak abundance of peptide 53RFQKSTELLIR63. * - only two trials of data and error bars are not available.
Figure 5.5 H3K56Ac changes in knockout (top) and overexpression (bottom) strains in response to JmjC proteins quantified by peak abundance of peptides $^{53}$RFQKSTELLIR$^{63}$ and $^{54}$FQKSTELLIR$^{63}$. * - only two trials of data and error bars are not available.
5.3.5 Global Changes of Histone H4 Acetylation

Cross-talk is believed to occur either within the same or different histone tail(s). To assess acetylation changes on H4 in response to the JmjC proteins, LC-MS of yeast histone mixture was performed. The histone separation is shown in Figure 2.1. As seen in the figure, histone H4 eluted into a single peak. Its spectrum is shown in Figure 5.6. As seen, histone H4 is highly acetylated. N-terminal acetylation (i.e. monoacetylation, Ac1, peak 1) and diacetylation (Ac2, peak 2) are the dominant forms; tri-, tetra-, and penta-acetylation are in much lower abundances. To find out the relative change of each acetylation isoform in response to the JmjC proteins, the relative abundances of peaks 1-7 (as denoted in Figure 5.6) were calculated by using equation 5.3. The relative peak abundance \( R(i) \) of a knockout and overexpression strain was then compared to the corresponding relative peak abundance of the WT. By this means, the relative change (%) of each acetylation isoform was calculated by use of equation 5.4. The results are summarized in Figure 5.7.

\[
R(i) = \frac{\text{Peak area}(i)}{\sum_{i=1}^{n=7} \text{Peak area}(i)} \quad \text{eqn 5.3}
\]

\[
\text{Relative change}(i) = \frac{R(i, \text{knockout/overexpression}) - R(i, \text{WT})}{R(i, \text{WT})} \times 100 \quad \text{eqn 5.4}
\]

As seen in Figure 5.7, in the knockout strains acetylation levels of Ac1, Ac2, Ac3, and Ac4 change less than 15% compared to the WT. In the overexpression strains, except...
for Ac3 in *JHD2*, the changes of acetylation Ac1-Ac4 are less than 10%. Due to the low abundance of Ac5, its relative abundance may be masked by the baseline peaks. It is noteworthy that the acetylation changes of all isoforms in *rph1Δ* and *RPH1* are less than 5%. Taken all of the H4 data together, it is safe to conclude that alteration of the JmjC proteins have minimal effects on H4 Ac1-Ac4 acetylation levels.

![Figure 5.6 LC-MS spectrum of yeast histone H4.](image)

**Figure 5.6** LC-MS spectrum of yeast histone H4.
Figure 5.7 Relative changes of H4 acetylation in knockout (top) and overexpression (bottom) strains in response to JmjC domains.
5.3.6 Global Changes of Histone H2B Acetylation

H2B acetylation was evaluated in response to the JmjC proteins. Since some isoforms of H2B1 and H2B2 coelute (See Figure 2.5), peaks for H2B1 and H2B2 in Figure 2.1 were averaged. The resulting spectrum of H2B1 and H2B2 are shown in Figure 5.8. Due to the low abundances of hexa- (Ac6) and hepta-acetylation (Ac7), these two H2B isoforms could not be reliably detected in some strains. For this reason, Ac6 and Ac7 were not quantified in this study. The relative peak abundances of peaks 1-17 was calculated following equation 5.5 for the quantification of Ac0-Ac5. The relative change of acetylation was calculated following equation 5.4. The results are shown in Figure 5.9 and 5.10.

\[
R(i) = \frac{\text{Peak area}(i)}{\sum_{i=1}^{n=17} \text{Peak area}(i)} \quad \text{eqn 5.5}
\]

![Yeast H2B1 and H2B2 spectrum.](image)

Figure 5.8 Yeast H2B1 and H2B2 spectrum.
Figure 5.9 H2B1 acetylation changes in knockout (top) and overexpression (bottom) strains in response to the JmjC proteins.
Figure 5.10 H2B2 acetylation changes in knockout (top) and overexpression (bottom) strains in response to the JmjC proteins.
The following observations were made from Figure 5.9. (1) In response to the JmjC proteins, the relative change of Ac0-Ac3 levels in the knockout strains are less than 10% except for Ac3 in $jhd1\Delta$ and Ac0 in $jhd2\Delta$. Because the relative abundances of Ac4 and Ac5 are less than 3% across all strains, the relatively large change of acetylation levels in these two isoforms may not signal significant changes of Ac4 and Ac5. Based on this observation, we conclude that knocking out Gis1 and Ecm5 proteins has little effect on H2B1 acetylation. (2) When the JmjC proteins are overexpressed, the relative change of acetylation levels in some isoforms, such as Ac0 in $RPH1$ and $JHD1$ and Ac5 in $JHD1$, increases greatly. Compared to the knockout strains, we can see that except for Ac0 in $JHD1$ the acetylation change patterns in all Ac isoforms from $JHD1$ and $JHD2$ are the same in the overexpression strains. In contrast, except for Ac0 the acetylation change patterns in all of the isoforms from $RPH1$ are opposite in the knockout and overexpression strains. Theoretically speaking, if one JmjC protein affects H2B acetylation, the acetylation change patterns in its knockout and overexpression strains are to be opposite. Thus, here it is safe to conclude that Rph1, demethylase of K9 and K36, correlates with H2B1 acetylation, while Jhd1 and Jhd2 have no significant correlation to H2B1 acetylation.

Compared to Figure 5.9, Figure 5.10 demonstrates the similar acetylation change patterns in H2B2 to H2B1. We conclude the following: (1) Knocking out Gis1 and Ecm5 proteins has little effect on H2B2 acetylation. (2) Rph1 correlates H2B2 acetylation, while Jhd1 and Jhd2 have no significant correlation to H2B2 acetylation.
5.3.7 *Global Changes of Histone H2A Acetylation*

In response to alteration of the JmjC proteins, H2A acetylation was also evaluated. As seen in Figure 5.11, H2A is predominantly N-terminal acetylated across all knockout and overexpression strains. Di- and triacetylation are at very low abundance. No significant differences among the strains were observed from the LC-MS spectra. To confirm our visual observation, the relative abundance of each acetylation isoform and the relative change of each isoform compared to WT were calculated by following equations 5.1 and 5.2, respectively. Peaks representative of each acetylation isoform are as denoted in Figure 5.11. The results are shown in Figure 5.12. As seen in Figure 5.12, in response to the JmjC proteins the overall acetylation changes in both knockout and overexpression strains are less than 10%. Given the low abundance of Ac2 and Ac3, we conclude that the JmjC proteins have little effect on H2A acetylation.
Figure 5.11 Yeast H2A spectrum.
Figure 5.12 H2A acetylation changes in knockout (top) and overexpression (bottom) strains in response to the JmjC proteins.
5.4 CONCLUSIONS

In this study, histone acetylation on H3, H4, H2A, and H2B were assessed in alteration of JmjC proteins Gis1, Ecm5, Rph1, Jhd1, and Jhd2. Results show that these JmjC proteins have minimal effects on acetylation of H4 and H2A. Gis1, Ecm5, Jhd1, and Jhd2 have minimal effects on acetylation of H2B1 and H2B2, while Rph1 has significant effects on acetylation of H2B1 and H2B2. With the alteration of the JmjC proteins, H3 acetylation changes significantly in all strains expect for ecm5Δ. LC-MS/MS results indicate that Gis1 and Jhd1 facilitate K56Ac, while Rph1 decreases K56Ac level. Taking into consideration that JmjC proteins are demethylases, this study implies that H3 methylation is correlated with acetylation on H2B1, H2B2, and H3.

\[
\begin{array}{cccccc}
1 & 11 & 21 & 31 & 41 \\
ARTKQTARKS & TGGKAPRKQL & ASKAARKSAP & STGGVKKPHR & YKPGTVALRE \\
51 & 61 & 71 & 81 & 91 \\
IRRFQKSTEL & LIRKLPFQRL & VREIAQDFKT & DLRFQSSAIG & ALQESVEAYL \\
101 & 111 & 121 & 131 \\
VSLFEDTNLA & AIHAKRVTIQ & KKDIKLARRL & RGERS \\
\end{array}
\]

**Figure S5.1** Sequence of yeast histone H3.
CHAPTER 6

SUMMARY

This dissertation focuses on characterization and quantitation of histones and their posttranslational modifications (PTMs) by mass spectrometry-based approaches. The goal was to understand the inherent biological significance of histones and their PTMs. To this end, a pipeline of mass spectrometry-based methodologies were developed and applied to characterization and quantitation of histones and their PTMs.

Chapter 2 describes the development of methods for yeast histone preparation, separation, and LC-MS characterization. Due to the low number of histone genes in yeast, the histones are at low concentrations. Efforts have been made to extract histones from the whole cells and nuclei. Using a number of published protocols, I obtained yeast histones that were of low purity and poor yield. To improve the quantity and purity of histones obtained from yeast, I optimized the protocol through extensive washing of nuclei prior to lysis and acid extraction. Using this yeast histone extraction protocol, the separation of histones by SDS-PAGE was optimized. Traditionally, a Laemmli system is used for this purpose. However, this system has limited resolution for small proteins like yeast histones. Therefore, I evaluated the separations of tricine SDS-PAGE over a range of acrylamide cross-linking percentages and found that 16.5% tricine gels gave the best
separation for yeast histones. In addition to SDS-PAGE, the LC-MS profiling method was further refined to improve the separation of yeast histones. At the optimal working conditions, all of the core histone isoforms were detected simultaneously with minimal TFA adducts.

Chapter 3 describes the development of an approach to distinguishing between the isobaric histone modifications, acetylation and trimethylation. Acetylation and trimethylation are two common histone PTMs. These two modifications both have a mass of 42 Da and are not easily distinguished by mass spectrometry. They can be studied and distinguished by use of antibodies. However, antibodies are limited by their availability, site specificity, and cross reactivity. Some MS-based approaches have also been developed to study these two PTMs. However, those approaches require the use of high-resolution / high-mass accuracy mass spectrometers that are not available to many laboratories. Our studies showed that on reversed-phase HPLC acetylated peptide isoform have longer retention time and this effect of acetylation on retention time is accumulative. In contrast, mono-, di-, and trimethylation have little effect on retention time. In this manner, acetylation and trimethylation can be unambiguously distinguished even on a low-mass accuracy mass spectrometer.

Chapter 4 describes the discovery of JmjC domain-containing demethylases. Before 2004, methylation was believed to be an irreversible modification. In 2004 the first demethylase, LSD1, was reported. LSD1 reverses dimethylation but not trimethylation. Shengjiang Tu proposed that trimethylation should also be reversible by JmjC domain-containing proteins. To test the hypothesis, sensitive MS-based methods
were developed to assess the changes of histone methylation in yeast. Using the optimized protocols developed in Chapter 2 several JmjC domain-containing proteins were either knocked out or overexpressed in yeast to screen for demethylase activities. The MS-based screening indicated that four of the five JmjC proteins are histone H3 demethylases: Gis1 and Jhd1 are demethylases for K36 mono- or dimethylation, Rph1 for K36 di- and trimethylation, and Jhd2 for K4 trimethylation.

Chapter 5 examines the potential of histone cross-talk between acetylation and methylation on yeast histones due to alterations in demethylase activities. It has been proposed that methylation and acetylation are synergistic. The work presented in Chapter 4 provided evidence that methylation and acetylation are closely correlated. To obtain a better understanding of the relationship between methylation and acetylation, the JmjC proteins were altered to probe synergistic changes in acetylation due to histone demethylation. Our results show that with the alteration of the JmjC proteins, H3 acetylation changes significantly in all strains except for $ecm5^\Delta$. LC-MS histone profiles showed a correlation between JmjC mutant strains and acetylation on H2B1, H2B2, and H3. Further, LC-MS/MS analysis indicated that Gis1 and Jhd1 facilitate K56Ac, while Rph1 decreases K56Ac level.
LIST OF REFERENCES


35. Jason, L. J.; Finn, R. M.; Lindsey, G.; Ausio, J., Histone H2A ubiquitination does not preclude histone H1 binding, but it facilitates its association with the nucleosome. *J Biol Chem* 2005, 280, (6), 4975-82.


187


197


Figure A.1 An LC trace of yeast histones from an LTQ-Orbitrap mass spectrometer. The histones were extracted from cells with OD$_{600} = 1.55$. 
Figure A.2 LC-MS spectra of bovine H2B from a Q-TOF mass spectrometer with a cone voltage of 30 V.
Figure A.3 LC-MS spectra of bovine histone H2B from a Q-TOF mass spectrometer with a cone voltage of (a) 21 V, (b) 25 V, (c) 35 V, and (d) 50 V.
Figure A.4 LC-MS spectra of bovine histones from a Q-TOF mass spectrometer with a cone voltage of 50 V.
Figure A.5 LC-MS spectra of bovine histones from a Q-TOF mass spectrometer with a cone voltage of 25 V.
APPENDIX B: SUPPLEMENTARY INFORMATION FOR CHAPTER 4
Figure B.1 The five JmjC domain-containing proteins in *S. cerevisiae*. A, Domain structures. The diagrams were adapted from SMART database. Each protein name is followed by the open reading frame name of the gene. ZF represents a zinc finger domain. B, Core β sheet sequence alignment of the 5 yeast JmjC proteins. Predicted β sheet sequences are indicated with “b” and highlighted in bold and italicized letters; the predicted cofactor Fe$^{2+}$ binding sites are in red; the predicted α-ketoglutarate binding sites are in blue.
Figure B.2 LC-MS/MS spectra of $^{73}$EIAQDFKTLR$^{83}$. 

Precursor ion: 669.04$^{2+}$

*: -NH$_3$

': -H$_2$O
**Figure B.3** LC-MS/MS spectra of $^{73}$EIAQDFK$_{Me}$TDLR$^{83}$. 

- ***:** \(-\text{NH}_3\)
- **':** \(-\text{H}_2\text{O}\)

Precursor ion: 676.18$^{2+}$
Figure B.4 LC-MS/MS spectra of \(^{73}\text{EIAQDFK}_{\text{Me}_2}\text{TDLR}^{83}\).
Figure B.5 LC-MS/MS spectra of $^{73}$EIAQDFK$_{Me3}$TDLR$^{83}$. 

Precursor ion: 690.54$^{2+}$

': -H$_2$O
Figure B.6 LC-MS/MS spectra of $^{28}$SAPSTGGVK$_{Me2}$KPHR$^{40}$. 

Precursor ion: 676.17$^{2+}$

$^*: -NH_3$

$''$: +H$_2$O
Figure B.7 LC-MS/MS spectra of $^{28}$SAPSTGGVK$_{Me3}$KPHR$^{40}$. 
Figure B.8 Western blot of overexpressed JmjC proteins. In the controls (-), no galactose was introduced. Protein overexpression (+) was induced by addition of 2% galactose to SC-Uracil or YP medium. After 4-hour induction, cells were harvested and the whole cell lysate were applied for immunoblot. The JmjC proteins with C-terminal His tag were examined with anti-His antibody.

Figure B.9 Growth phenotype assays of yeast cells. A, UV irradiation assays on the deletion strains. Wild type (WT) yeast, ecm5Δ, gis1Δ, jhd1Δ, jhd2Δ, and rph1Δ were grown to an OD_{600} of 1.00. Cell viability was assayed by placing serial dilutions of cells on YPD plates followed by exposure to UV-irradiation (10 mJ/cm²). B, Overexpression of H3K36 demethylases slightly influenced 6-AU resistance. Cells were grown in rich medium (YP) with 2% galactose for 4 hours and assayed on SC-Uracil plates in the absence (left panel) or presence of 30 µg/ml 6-azauracil (6-AU) (right panel).
APPENDIX C: PROTOCOLS

C.1 In-gel digestion

A. Reagents

Wash solution:

50% MeOH
5% HAc
45% H₂O

Dithiothreitol (DTT) solution:

5 mg/ml DTT in 100 mM NH₄HCO₃

Iodoacetamide (Iodo) solution:

15 mg/ml Iodo in 100 mM NH₄HCO₃

Trypsin solution:

20 µg of Promega sequencing grade modified trypsin (V511)
1.0 ml of 25 mM ABC

Extraction buffer:

50% CAN
5% formic acid
45% H₂O

Acetone (ACN)

100 mM ammonium bicarbonate (ABC)
B. Procedure

1. Cut gel bands as closely as possible and combine two bands of the same protein together. Cut them into small pieces.
2. Soak bands in 200 µl of wash solution for 1 h.
3. Remove wash solution and add 200 µl of fresh wash solution.
4. Remove wash solution and dry gel pieces in 200 µl of ACN for 5 min.
5. Remove ACN and dry gel pieces in speed-vac for 2-3 min.
6. Add 75 µl of DTT solution and incubate 30 min at room temperature to reduce the disulfide bonds.
7. Remove DTT and add 75 µl of Iodo solution and incubate 30 min in dark at room temperature to alkylate the SH-group. If disulfide bonds are not a concern, steps 7 and 8 can be skipped.
8. Remove Iodo solution and wash with 200 µl of 100 mM ABC for 5 min.
9. Remove ABC; dehydrate in 200 µl of ACN for 5 min.
10. Remove ACN and rehydrate in 200 µl of ABC for 5 min.
11. Remove ABC and dehydrate in 200 µl of ACN for 5 min.
12. Remove CAN and dry gel pieces in speed-vac for 2-3 min.
13. Add 30 µl of trypsin solution to each sample (or adjust the amount of trypsin according to the amount of gel pieces).
14. Rehydrate on ice for 10 min.
15. Remove excess trypsin (may not be necessary).
16. Add 20 µl of 25 mM ABC and digest at 37 °C for 2 hours or overnight depending on the specific enzymes and proteins.

17. Add 40 µl of extraction buffer, incubate for 10 min, and then collect solution.

18. Repeat extraction twice.

C.2 Preparation for MALDI-TOF MS experiment

A. Reagent

α-cyano-4-hydroxycinnamic acid (CHCA) matrix solution:

- 0.1% TFA
- 50% acetone
- 50% H₂O

B. Procedure

1. Clean MALDI plate: sonicate the plate in 10% formic acid in MeOH for 30 min, then sonicate in pure water for 30 min, and rinse it with pure MeOH till dry.

2. Mix sample and matrix at 1:5 ratio or other ratios and spot 1 µl onto MALDI plate. Air dry and spot wash with 0.1 % TFA to get rid of the salt (not necessary for 25 mM ABC).
C.3 AU-PAGE separation of histones

A. Reagents (All the concentrations indicated below are final concentrations)

Resolving Gel: takes about one hour or longer to polymerize

<table>
<thead>
<tr>
<th>Component</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 M Urea</td>
<td>18.0 g</td>
</tr>
<tr>
<td>15% Acrylamide</td>
<td>18.75 ml (40% stock 38.7:1.3)</td>
</tr>
<tr>
<td>5% Acetic Acid</td>
<td>2.50 ml</td>
</tr>
<tr>
<td>dd H₂O</td>
<td>~8.20 ml</td>
</tr>
<tr>
<td>TEMED</td>
<td>100 µl</td>
</tr>
<tr>
<td>APS</td>
<td>500 µl</td>
</tr>
</tbody>
</table>

Scavenger: 1 ml (runs over resolving gel)

<table>
<thead>
<tr>
<th>Component</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>8 M Urea</td>
<td>4.8 g</td>
</tr>
<tr>
<td>5% Acetic Acid</td>
<td>500 µl</td>
</tr>
<tr>
<td>0.6 M β-mercaptoethanol</td>
<td>470 µl</td>
</tr>
<tr>
<td>dd H₂O</td>
<td>up to 10 ml</td>
</tr>
</tbody>
</table>

Stacking Gel: set up best at warm temperature

<table>
<thead>
<tr>
<th>Component</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 M Urea</td>
<td>7.2 g</td>
</tr>
<tr>
<td>7.5% Acrylamide</td>
<td>3.75 ml (40% stock)</td>
</tr>
<tr>
<td>0.375 M KOAc</td>
<td>5.00 ml</td>
</tr>
<tr>
<td>dd H₂O</td>
<td>5.20 ml</td>
</tr>
<tr>
<td>TEMED</td>
<td>40 µl</td>
</tr>
<tr>
<td>APS</td>
<td>200 µl</td>
</tr>
</tbody>
</table>

2× Sample Buffer:

<table>
<thead>
<tr>
<th>Component</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>8 M Urea</td>
<td>4.8 g</td>
</tr>
<tr>
<td>8% β-mercaptoethanol</td>
<td>0.8 ml</td>
</tr>
<tr>
<td>10% Acetic Acid</td>
<td>1.0 ml</td>
</tr>
<tr>
<td>0.001% methylene blue</td>
<td>~10 µl (1% stock)</td>
</tr>
<tr>
<td>dd H₂O</td>
<td>up to 10 ml</td>
</tr>
</tbody>
</table>

Coomassie Stain:

<table>
<thead>
<tr>
<th>Component</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>25 % isopropanol</td>
<td></td>
</tr>
<tr>
<td>10% glacial acetic acid</td>
<td></td>
</tr>
<tr>
<td>0.25% Coomassie brilliant blue</td>
<td></td>
</tr>
</tbody>
</table>
Fixing Solution:

50% ethanol/10% acetic acid

Washing Solution:

50% methanol/5% acetic acid

Destaining Solution:

30% methanol/10% acetic acid

Storage Buffer:

5% acetic acid in water or pure water

B. Procedure

1. Clean two glass plates (18 cm x 16 cm) thoroughly and let them air dry. Add two side plastic spacers and one bottom plastic spacer. Then clap them tightly to make sure the plates are sealed and no acrylamide drips out.

2. Degas the resolving gel for 15 min, and then add 500 µl of APS (ammonium persulfate) first and 100 µl of TEMED second and mix them completely.

3. Pour about 30 ml of the resulting resolving gel slowly into the space between those two plates, which are set up in step 1, till about one centimeter below the comb. You must pour the gel immediately after adding APS and TEMED since these reagents catalyze the polymerization process.

4. Add some water saturated butanol to the above gel to get a straight line. Avoid introducing air bubbles. Wait for about 1 h to let it polymerize.

5. Take off the bottom spacer and transfer the plates into electrophoresis tank. Clap them tightly. Rinse off the butanol solution. Add 5% acetic acid running buffer to
the bottom chamber along the side wall to avoid any air bubbles and then add buffer to the top chamber. Check for leaking. If bubbles are trapped under the gel, use a syringe with a bent needle to aspirate them.

6. Plug wires into the electrodes (Connect the red wire to the cathode and the black one to the anode). Histones run toward the negative (cathode) end of the gel since they are positively charged. Pre-run the gel at a constant voltage of 300 for at least 6 h at 4 °C.

7. The next day turn off the voltage first, add 1 ml scavenger into the glass plates. Run at 300 V for 1 hour at 4°C.

8. Prepare fresh stacking gel solution. Put in 37 °C water bath to dissolve urea completely. Filter and degas it and then add 200 µl of APS and 40 µl of TEMED.

9. Take the gel out of the electrophoresis tank. Pour out the running buffer and rinse off excess resolving gel with distilled water. Put back the bottom spacer, clap the plates, and dry the interface between the two plates with Kimwipes.

10. Pour stacking gel solution and avoid any bubbles. Push comb in immediately and wait for 3 hour or till it polymerizes completely.

11. Prepare samples. Usually 20 -25 µg of samples are loaded. Add equal volume to each well (usually about 15 µl). Calculate the volume of sample needed. If not enough, add distilled water to make the volume the same for each sample and mix with the same volume of 2x loading buffer. Then heat the resulting mixture at ≥95°C for at least 2 minutes to denature the protein and expose the charges. Add 2
µl of 0.001% methylene blue, vortex, and then centrifuge for 30 seconds to mix them completely.

12. Take off the bottom spacer and put back the plates into electrophoresis tank. Clap it tightly. Add 5% acetic acid running buffer to the bottom and top chamber as before. Take out the comb. Make each well straight. Rinse each well with 5% acetic acid several times before loading samples with Multiflex pipette tips.

13. Plug in the wires as before and run the gel at 300 V at 4 °C for 18 hours (adjust the time based on the samples and running conditions) till the front of the dye is run out of the gel.

14. Stop the voltage. Take the plates out of the electrophoresis tank. Transfer the gel carefully to a container with fixing solution.

15. Fix the gel for 3-4 hours on a shaker.

16. Aspirate the fixing solution and wash the gel twice in washing solution (30 min each).

17. Cover the gel in Coomassie stain solution for 30 minutes.

18. Destain the gel in destain solution till the background is clear.
C.4 Histone separation by use of 16.5% Tris-Tricine-SDS gels

A. Materials and reagents

Equipment: Bio-Rad Mini Format Vertical Electrophoresis Apparatus

Gels: The gels are 16.5% Tris-tricine-SDS gels with 4% stacking gels (Bio-Rad with a catalog number of 161-1107).

Chemicals and reagents:

10 X Tris/Tricine/SDS buffer from Bio-Rad with a catalog number 161-0744.

Tricine sample buffer from Bio-Rad with a catalog number 161-0739.

MW Marker: Precision plus protein standards from Bio-Rad with a catalog number 161-0373.

B. Procedure

1. Prepare sample buffer by mixing 980 µl of tricine sample buffer with 20 µl of β-mercaptoethanol.

2. Mix samples with the resulting sample buffer at a ratio of 1:3.

3. Incubate the solution from step 2 in 95~100 ºC water bath for 10 min and spin down.

4. Prepare 1 × Tris/Tricine/SDS buffer (running buffer).

5. Set up the electrophoresis apparatus, install the gel, and fill the chambers with 1 × Tris/Tricine/SDS buffer.

6. Wash each well of the gel with the running buffer.

7. Load samples from step 3 to the desired wells.
8. Run the gel at a constant voltage of 120 V till 15 kDa band from the marker runs to the end of the gel (about 2cm above the bottom of the gel).

10. Stop the voltage and take the gel out of the device carefully.

11. Fix, wash, stain, and destain the gel as described for AU-PAGE.

12. The alternative approach to step 10 is as the following.

**Alternative microwave procedure for gels**

**A. Wash**

Place the gel into a microwavable tray containing 100 ml of distilled water, and microwave till boiling. Wash the gel with 100 ml of fresh distilled water, and microwave again till boiling. Wash the gel with another 100 ml of fresh distilled water and place on a shaker for 5 min.

**B. Stain**

Add commercial Coomassie stain solution to cover the gel completely and microwave the gel till boiling. Place the tray on a shaker for 5~20 min or till the gel bands are visible.

**C. Destain**

Discard staining solution and add 200 ml of distilled water. Incubate at room temperature on a shaker till the band intensity is satisfying.

Note: Frequently replacing water and washing for a longer time may increase band intensity.
C.5 Human histone extraction

A. Reagents

NP-40 Lysis Buffer (store at 4 ºC):

- 10 mM Tris-HCl (pH 7.4)
- 10 mM NaCl
- 3 mM MgCl$_2$
- 0.5% NP-40 (Nonidet P-40)
- 0.15 mM Spermine
- 0.5 mM Spermidine

To make 100 ml of this solution (take from the stock solution):

- 1.0 ml 1M Tris-HCl (pH 7.4)
- 0.2 ml 5M NaCl
- 0.3 ml 1M MgCl$_2$
- 0.5 ml 100% NP-40 (Nonidet P-40)
- 15 µl 1M Spermine
- 50 µl 1M Spermidine

NP-40 (Nonidet P-40) is very viscous. So add all of the others first and then add 0.5 ml of 100% NP-40 drop by drop in different places of the above solution. After mixing completely, store at 4 ºC. When ready to use, take the necessary volume of the above solution and add PMSF and Protease inhibitor. For 5 ml of the buffer, add 50 µl of 100 mM PMSF stock solution and 5 µl of protease inhibitor (x1000 stock or x100 stock).

The final concentration of PMSF is 1 mM and protease inhibitor is 1x.

B. Procedure

1. Pellet the cells down at 1500 rpm at 4 ºC for 10 minutes in a disposable 50-ml conical tube. Discard the supernatant.
2. Wash the cell pellet twice with 10 ml of ice cold 1XPBS. Pellet cells at 1500 rpm for 10 minutes at 4 °C.

3. Resuspend the pellets in 5 ml (for 100 million cells) of ice cold NP-40 lysis buffer and incubate on ice for 5 min.

4. Pellet the nuclei at 1000 rpm for 10 min at 4 °C and discard the supernatant slowly with a pipette since the pellet is very loose.

5. Wash the pellet with 5 ml of TBS and centrifuge at 1000 rpm for 5 min at 4 °C (to wash away the lysis buffer. It is not necessary if no further reaction is performed.) (TBS: Tris-HCl, 10 mM, pH 7.5; Saline NaCl 150 mM).

6. Discard the supernatant, add ~ 400 µl of 0.2 M sulfuric acid to the pellet, vortex and incubate on ice for 30 min (First add 150 µl to transfer the pellet to an eppendorf tube and then use another 200 µl to wash the tube).

7. Spin down at 14 krpm in cold room for 15 min. Transfer the supernatant to another clean 2-ml eppendorf tube.

8. Add acetone to a final concentration of 80% to the supernatant and precipitate at -20 °C for overnight (e.g. For 400 µl of the solution, add 1600 µl of ice cold pure acetone).

9. Spin down at 14 krpm for 15 min in cold room and wash it with ice cold acetone once. Let it air dry for one min and add HPLC water to dissolve the histones (500 µl of H2O for 10 million cells).

10. Use BCA analysis to determine the protein concentration and store in -80 °C.
C.6 Cell growth of *S. cerevisae*

I. Knockout strains

A. Media

YP medium:

- 10% yeast extract
- 20% Bacto peptone

Glucose solution:

- 20% in water

B. Procedure

1. Prepare and autoclave YP medium.
2. Prepare glucose solution and sterilize it by filtering.
3. Mix 1.0 ml of glucose solution with 9.0 ml of YP medium in a sterile flask, which results in YPD medium.
4. Streak a colony of a knockout strain to the YPD medium, which results in the start culture.
5. Incubate the start culture at 30 °C 200 rpm overnight till the cells grow to saturated phase (e.g. OD$_{600}$ ≥ 1.2).
6. Mix 180 ml of YP medium with 20 ml of glucose solution in a sterile flask.
7. Transfer appropriate amount of the saturated cell culture to the 200 ml YPD medium from step 6 and incubate the resulting culture at 30 °C 200 rpm overnight till OD$_{600}$ = 1.0 ~ 1.3. The double time of the cell is 1.5 h.
8. Spin down the cells in a JA14 rotor at 6000 rpm for 10 min and discard the medium.
II. Overexpression strains

A. Media

S-Ura medium:

6.7 % yeast nitrogen base w/o amino acids
0.077 % -Ura DO supplement

Galactose solution:

20% galactose in water

Glucose solution:

20% glucose in water

Sucrose solution:

20% sucrose in water

B. Procedure

1. Mix 1.0 ml of glucose solution with 9.0 ml of S-Ura medium, which results in SD-Ura medium.

2. Streak 10~20 colonies of an overexpressed strain to the SD-Ura medium and incubate the culture at 30 ºC 200 rpm overnight, which results in the start culture. The double time for overexpressed strains varies from 2 h to 3 h.

3. Add 2.0 ml of the start culture to 16.0 ml of fresh SD-Ura.

4. Mix 180 ml of S-Ura medium with 20 ml of sucrose solution, which results in SC-Ura medium.

5. Add appropriate amount of cell culture from step 3 to 200 ml of SC-Ura medium and incubate the culture at 30 ºC 200 rpm overnight till OD$_{600}$ = 1.1 ~ 1.3.
6. Spin down the cells at 7 krpm for 10 min and wash the cells with 30 ml of sterile water.

7. Spin down the cells again and discard the water.

8. Resuspend the cells in 20 ml of S-Ura medium and then transfer the cells to a flask containing 160 ml of S-Ura medium and 20 ml of 20% galactose.

9. Incubate the cells at 30 °C 200 rpm for 4 h and check OD$_{600}$. 

10. Spin down the cells as before and wash the cells with 30 ml of sterile water.
C.7 Yeast histone extraction

A. Reagents

Buffer A:

50 mM Tris-HCl, pH 7.5
30 mM DTT

Buffer S:

20 mM HEPES, pH 7.4
1.2 M Sorbitol

Buffer B:

20 mM PIPES, pH 6.8
1.2 M Sorbitol
1 mM MgCl₂

Buffer NIB:

0.25 M sucrose
60 mM KCl
15 mM NaCl
5 mM MgCl₂
1 mM CaCl₂
15 mM MES, pH 6.6
0.8% Triton X-100
1 mM PMSF (fresh)
1 mM NaF (fresh)

Wash Buffer A:

10 mM Tris, pH 8.0
0.5% NP-40
75 mM NaCl
30 mM sodium butyrate
1 mM PMSF (fresh)
1 mM NaF (fresh)

Wash Buffer B:

10 mM Tris, pH 8.0
0.4 M NaCl
30 mM sodium butyrate
1 mM PMSF (fresh)
1 mM NaF (fresh)

B. Procedure

1. Resuspend the cell pellet in 30 ml of sterile water and transfer to JA-20 tubes. Spin down at 6500 rpm for 10 min in a JA-20 rotor.

2. Resuspend the pellet in 5 ml of Buffer A. Incubate at 30 °C for 15 min with 100 rpm shaking.

3. Spin down the pellet as before and resuspended in 10 ml of Buffer S.

4. Spin down the pellet and resuspend in 5 ml of Buffer S. Add 0.2 ml of 10 mg/ml Zymolyase 100T or approximately 2 mg Zymolyase/g yeast cells. Incubate at 30 °C for 60 min with 100 rpm shaking.

All the following procedures are to be carried out on ice or at 4 °C and use pre-cooled solutions.

5. Spin down the cells and resuspend in 10 ml of ice-cold Buffer B and spin down at 6500 rpm for 10 min in a JA-20 rotor.

6. Three successive times: resuspend the pellet in 5 ml of ice-cold NIB buffer with p1000, and incubate on ice for 20 min. Spin down the pellet at 7500 rpm for 15 min.

7. Three successive times: resuspend the pellet in 5 ml of wash buffer A, and hold on ice water for 15 min. Spin down the pellet.

8. Two successive times: resuspend the pellet in 5 ml of wash buffer B and hold on ice water for 5 min. Spin down the pellet.
9. Extract histones by resuspending the pellet in 1.4 ml of 0.2 M H\textsubscript{2}SO\textsubscript{4} and transfer the pellets in H\textsubscript{2}SO\textsubscript{4} to a 2-ml eppendorf tube. Incubate on ice water for 1 hour with occasionally vortexing. Note: Concentrated H\textsubscript{2}SO\textsubscript{4} is 18 M.

10. Spin down at 14000 rpm for 15 min. Transfer the supernatant, which contained the extracted histones, to a clean 2-ml tube.

11. Add trichloroacetic acid solution (100 % w/v in water) to the supernatant to a final concentration of 20% (i.e. 250 µl per ml of supernatant) and precipitation of the extracted histones will be observed immediately. Keep the solution on ice water for 12-16 h. Spin down the precipitate at 14000 rpm for 30~40 min and discard the supernatant.

12. Wash the pellet with 500 µl of ice cold acidified acetone (0.1% HCl in acetone), vortex, and centrifuge for 20 min at 14000 rpm. Remove the acidified acetone carefully.

13. Wash the pellet with 500 µl of ice cold acetone and centrifuge for 20 min at 14000 rpm. Remove acetone carefully and air dry the pellet on ice and store at -80 ºC.
C.8 Histone extraction from bovine calf thymus

A. Reagents

Buffer C:

- 10 mM MgCl$_2$
- 25 mM KCl
- 50 mM Tris-HCl, pH 7.5
- 50 mM NaHSO$_3$
- 250 mM sucrose

Solution I:

- 1.4 M NaCl
- 0.5 M NaHSO$_3$

Solution II:

- 0.25 M HCl

Solution III:

- 0.025 M HCl

B. Procedure

1. Thaw the frozen tissue in Buffer C
2. Cut the tissue into small pieces and weigh them.
3. Put the tissue into homogenization tubes and add 10 volume of Buffer C to it (e.g. 10 ml of buffer C for 1 g of tissue).
4. Homogenize the tissue 2-3 times (3-5 second each time) using a homogenizer.
5. Spin down the homogenized tissue at 3000 rpm for 5 minutes.
6. Resuspend the tissue in 10 volume of Buffer C and spin down as before.
7. Repeat step 7 once.
8. Wash the pellet with 10 volume of Solution I five times.
10. Resuspend the pellet in 10 volume of solution II and shake it for 2 hours in cold room.

11. Spin down the sample at 16000 rpm for 10 minutes and dialyse the supernatant for \( \geq 12 \) hours using 300 volume of solution III in cold room.

12. Precipitate histones with 10 volume of acetone overnight at -20 °C.

13. Spin down the precipitate at 3000 rpm for 5 minutes and air dry it.

14. Store the sample at -80 °C.
C.9 Histone extraction from *Drosophila*

A. Reagents

D-22 medium without serum

Buffer A:

- 300 mM sucrose
- 2 mM MgOAc
- 3 mM CaCl\(_2\)
- 10 mM Tris, pH 8.0
- 0.5 mM 1, 4-dithiothreitol (DTT)
- 0.5 mM phenylmethysulfonyl fluoride (PMSF)
- 5 mM n-butyrate

Buffer B:

- 2 M sucrose
- 5 mM MgOAc
- 10 mM Tris, pH 8.0
- 0.5 mM DTT
- 0.5 mM PMSF
- 5 mM n-butyrate

Buffer C:

- 0.25% glycerol
- 5 mM MgOAc
- 5 mM Tris, pH 8.0
- 0.5 mM DTT
- 0.5 mM PMSF
- 5 mM n-butyrate
- 0.1 mM EDTA, pH 8.0
- 0.2% NP-40

B. Procedure

1. Grow the *Drosophila* Kc cell at 24 °C in 75-cm\(^2\) tissue culture flasks (Corning) using D-22 medium without serum for 6-8 h. All subsequent procedures were performed at 4 °C or on ice.
2. Cells were pelleted and resuspended at 2.5-5.0 \times 10^7 \text{ cell/ml} in Buffer A.

3. Homogenize the cells in a Dounce homogenizer with pestle A until >95% of the cells were lysed, and nuclei were liberated (ascertained by staining the homogenate with DAPI and checking with phase and UV optics (usually 20-30 strolls)).

4. The resulting homogenate was diluted with an equal amount of buffer B and layered over a 5-10-ml heavy sucrose cushion in Buffer B; the interface was stirred gently to form a crude gradient.

5. The above gradients were centrifuged at 15000 rpm for 30 min at 0 °C.

6. Aspirate the supernatant and resuspend the pellet in Buffer C.

7. Store the above solution at -80 °C.
C.10 Cell culture and histone extraction from *Tetrahymena*

**A. Reagents**

Medium:

- 0.25% proteose peptone
- 0.25% yeast extract
- 0.5% glucose
- 0.3 mM magnesium sulfate
- 18 µM calcium chloride
- 33 µM ferric chloride or ferric citrate

Nucleus wash buffer (pH 7.4):

- 0.25 M sucrose
- 10 mM Tris
- 3 mM CaCl$_2$
- 1 mM MgCl$_2$
- 1 mM iodoacetamide
- 10 mM butyrate
- 1 mM PMSF,
- 0.2 M H$_2$SO$_4$

Acidified acetone:

- 0.1% concentrated HCl in acetone

10 mM Tris-HCl, pH 7.4

Trichloroacetic acid (TCA) and/or perchloric acid (PCA)

**B. Cell growth**

**Note:** Speed and practice are very important in the following procedure. If you are thawing test vials, thaw the four tubules of the cryovial one by one and discard the cryovial.

1. Remove one cryovial from the cane and put the cane back under liquid N$_2$. 
2. Remove one tubule from the cryovial using sterile forceps (sterilized by dipping in ethanol, which is eliminated by flaming)

3. Drop the tubule into a 4-in test tube, containing 3 ml of PPY with penicillin and streptomycin, which has been prewarmed in a 35 °C water bath on a shaker

4. Turn the shaker on, and return the cryovial containing the remaining frozen tubules to its cane and the liquid N₂ tank.

5. After 1.5-2 min of shaking at 35 °C, pick up the test tube, load a sterile Pasteur pipette with warm medium from the tube, insert it into the tubule, and flush it. A good way to be sure that the tubule contents have been flushed out is to create bubbles at the bottom of the tubule.

6. Grow the cells in the medium in siliconized aspirator bottles, aerated by bubbling air through the medium, till the cell concentration is over 5 × 10⁵ cells/ml (one generation is about 2.5 h). Starving cells at concentrations greater than 10⁶ cells/ml greatly reduces the performance of subsequent mating mixtures.

7. To starve the cultures, collect and wash the cells once with 10 mM Tris-HCl (pH 7.4), then centrifuge at 100 g for 2 min in a clinical centrifuge at room temperature.

8. Resuspend the cells at a concentration of approximately 2 × 10⁵ cells/ml in an Erlenmeyer flask in a volume between one-tenth and one-fifth of the flask volume (Cultures were starved for at least 24 h).

9. Equal numbers of cells from two starved cultures of different mating types were mixed at 30 °C in an Erlenmeyer flask ten times of the volume of the mating
mixture. The concentration of the mating mixture is $1.5-3 \times 10^5$ cells/ml. More than 85-90% pairing will be observed in 3-4 h.

10. Keep the mated samples at -80 °C.

C. Histone extraction from the nuclei

1. Wash isolated nuclei with nucleus wash buffer to remove remaining protein and lipid contaminants from nucleus isolation.

2. Thoroughly resuspend the nuclei pellet in an appropriate volume, as determined by the following formula, of 0.2 M H$_2$SO$_4$.

\[
1-3 \times 10^6 \text{ macronuclei or } 2-6 \times 10^7 \text{ micronuclei per } 100 \mu l \text{ of } 0.2 \text{ M H}_2\text{SO}_4.
\]

**Note:** Certain histones may not be recovered well from more dilute samples because of poor precipitation or adsorptive losses. To minimize losses, use only polypropylene and/or siliconized tubes for this step.

3. Shake gently at 4 °C, 20 min or overnight. Note that histone may degrade with prolonged exposure to acid.

4. Centrifuge at moderate speed (5000~10000 rpm for 15 min) and collect the acid-soluble supernatant. Acid-insoluble pellets may be discarded because they contain little histone.

5. Precipitate proteins from the acid supernatant by adding 100% TCA to a final concentration of 20% (precipitate all histones, including histone H1) or PCA to 5.4% (precipitate all histones except H1). Incubate on ice for 15 min to 1 h. Extended incubations (overnight) may be necessary to precipitate small amounts of protein efficiently.
6. Collect precipitates by centrifugation at >10000 rpm for 20 min. Discard the supernatant.

7. Remove acid from the precipitate by washing sequentially with acidified acetone followed by two washes with straight acetone. Both acetone solutions should be cooled to -20 °C before use.

8. Allow precipitates to dry completely at room temperature and then resuspend in water.

**Note:** If resuspension of precipitate is difficult, warm (60 °C to boiling) for about 2 min, but only if samples will then be run on a denaturing gel.