



**Development of New Mass Spectrometry-based Methods for the
Analysis of Posttranslational Modifications**

**Entwicklung neuer massenspektrometrischer Methoden für die
Analyse posttranslationaler Proteinmodifikationen**

Doctoral thesis for a doctoral degree
at the Graduate School of Life Sciences,
Julius-Maximilians-Universität Würzburg,
Section Biomedicine

submitted by

Rasha ElBashir
From **Khartoum, Sudan**

Würzburg 2017



Submitted on:

Office stamp

Members of the *Promotionskomitee*:

Chairperson: Prof. Dr. Alexander Buchberger

Primary Supervisor: Prof. Dr. Andreas Schlosser

Supervisor (Second): Prof. Dr. Utz Fischer

Supervisor (Third): Prof. Dr. Hermann Schindelin

Supervisor (Fourth):

(If applicable)

Date of Public Defence:

Date of Receipt of Certificates:

Affidavit

I hereby confirm that my thesis entitled “*Development of New Mass Spectrometry-based Methods for the Analysis of Posttranslational Modifications*” is the result of my own work. I did not receive any help or support from commercial consultants. All sources and/or materials applied are listed and specified in the thesis.

Furthermore, I confirm that this thesis has not yet been submitted as part of another examination process neither in identical nor in similar form.

Würzburg, 05.07.2017

Place, Date

Rasha ElBashir

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt, die Dissertation „*Entwicklung neuer massenspektrometrischer Methoden für die Analyse posttranslatiionaler Proteinmodifikationen*“ eigenständig, d.h. insbesondere selbstständig und ohne Hilfe eines kommerziellen Promotionsberaters, angefertigt und keine anderen als die von mir angegebenen Quellen und Hilfsmittel verwendet zu haben.

Ich erkläre außerdem, dass die Dissertation weder in gleicher noch in ähnlicher Form bereits in einem anderen Prüfungsverfahren vorgelegen hat.

Würzburg, 05.07.2017

Ort, Datum

Rasha ElBashir

Abstract

Posttranslational modifications (PTMs) play a crucial role in many cellular processes. They are reversible, dynamic, and highly regulated events that alter the properties of proteins and increase their functional diversity. The identification and quantification of PTMs are critical for deciphering the molecular mechanisms of PTMs-related biological processes and disease treatment and prevention. Two of the most common and important PTMs that regulate many protein functions are acetylation and phosphorylation.

An important role of acetylation is the regulation of DNA/RNA-protein interactions. A prominent example for this are histones, whose tail regions are lysine-rich and can be highly acetylated at their N-terminal domain. In spite of the utmost importance of this PTM, methods that allow the accurate measuring the site-specific acetylation degree are missing. One of the challenges in quantifying the acetylation degree at an individual lysine residue of the histones N-termini is the occurrence of multiple lysines in close proximity. Herein, we describe the development of the "Fragment Ion Patchwork Quantification", a new mass spectrometry-based approach for the highly accurate quantification of site-specific acetylation degrees. This method combines $^{13}\text{C}_1$ -acetyl derivatization on the protein level, proteolysis by low-specificity proteases and quantification on the fragment ion level. Acetylation degrees are determined from the isotope patterns of acetylated b and y ions. We have shown that this approach allows determining the site-specific acetylation degrees of all lysine residues for all core histones of *Trypanosoma brucei*. In addition, we demonstrate the use of this approach to identify the substrate sites of histone acetyltransferases and to monitor the changes in acetylation of the histones of canonical nucleosome and transcription start site nucleosomes.

Phosphorylation is one of the most common and most important PTMs. The analysis of the human genome showed that there are about 518 kinases and more than 500,000 phosphorylation sites are believed to exist in the cellular proteome. Protein phosphorylation plays a crucial role in signaling many different cell processes, such as intercellular communication, cell growth, differentiation of proliferation and apoptosis. Whereas MS-based identification and relative quantification of singly phosphorylated peptides have been greatly improved during the last decade, and large-scale analysis of thousands of phosphopeptides can now be performed on a routine-base, the analysis of multi-phosphorylated peptides is still lagging vastly behind. The low pKa value of phosphate group and the asso-

ciated negative charge are considered the major source of the problems with the analysis of multi-phosphorylated peptides. These problems include the formation of phosphopeptide-metal complexes during liquid chromatography (e.g. Fe 3+), which leads to a drastic deterioration of the chromatographic properties of these peptides (peak tailing), the decreased ionization efficiencies of phosphorylated peptides compared to their unphosphorylated counterparts, the labile nature of phosphate during CID/HCD fragmentation, and the unsuitability of low-charged phosphopeptides for ETD fragmentation are the most important factors that hinder phosphorylation analysis by LC-MS/MS. Here we aimed to develop a method for improving the identification of multi-phosphorylated peptides as well as the localization of phosphorylation sites by charge-reversal derivatization of the phosphate groups. This method employs a carbodiimide-mediated phosphoramidation to convert the phosphates to stable aromatic phosphoramidates. This chemical modification of phosphite(s) reversed the negative charge of the phosphate group(s) and increased the number of the positive charges within the phosphopeptide. This modification prevented the formation of phosphopeptide-metal ion complexes that dramatically decreases or completely diminishes the signal intensity of protonated phosphopeptides, specifically multi-phosphorylated peptides. Furthermore, the increased net charge the (phospho-)peptides made them suitable for ETD fragmentation, which generated a high number of fragment ions with high intensities that led to a better phosphopeptide identification and localization of phosphite(s) with high confidence.

Zusammenfassung

Posttranslationale Modifikationen (PTMs) spielen eine entscheidende Rolle in vielen zellulären Prozessen. Sie sind reversible, dynamische und hochregulierte Ereignisse, die die Proteineigenschaften verändern und ihre funktionale Diversität erhöhen. Die Identifizierung und Quantifizierung von PTMs sind wesentlich für die Entschlüsselung der molekularen Mechanismen von PTM-regulierten biologischen Prozessen und für ein besseres Verständnis der Rolle posttranslativ Modifikationen bei einer Vielzahl von Krankheiten. Zwei der bedeutendsten PTMs, welche die Funktion unzähliger Proteine regulieren sind die Acetylierung an Lysin-Resten und die Phosphorylierung an Serin-, Threonin- und Tyrosinresten.

Im Rahmen dieser Arbeit wurden eine neue Methode zur Bestimmung des positionsspezifischen Acetylierungsgrades, sowie verbesserte Methoden für die Analyse der Phosphorylierung mittels Flüssigchromatographie-gekoppelter Tandem-Massenspektrometrie entwickelt. Wir haben eine neue MS-basierte Methode ("Fragment Ion Patchwork Quantification") entwickelt, welche es erlaubt die Acetylierungsgrade an individuellen Positionen mit hoher Genauigkeit zu messen. Diese Methode kombiniert die $^{13}\text{C}_1$ - Acetylderivatisierung von intakte Proteine, die Proteolyse durch Proteasen mit niedriger Spezifität, und die Quantifizierung auf dem MS2-Level. Die Acetylierungsgrade werden aus den Isotopenmustern von acetylierten b- und y-Ionen bestimmt. Obwohl unsere Methode zur Quantifizierung der positionsspezifischen Acetylierungsgrade auf jedes beliebige Protein angewandt werden kann, stand bei der Methodenentwicklung die Analyse der Histonacetylierung aufgrund ihrer herausragenden Bedeutung bei der Regulation der Genexpression im Vordergrund. Wir haben gezeigt, dass mit dieser Methode die Bestimmung der positionsspezifischen Acetylierungsgrade an allen Lysin-Resten aller Core-Histone von Nukleosomhistone von *Trypanosoma brucei* möglich ist. Darüber hinaus haben wir diese Methode angewandt, um die Substrat-Positionen von Histon-Acetyltransferasen zu identifizieren und um quantitative Veränderungen der Acetylierung an Histonen aus kanonischen Nukleosomen sowie Nukleosomen an Transkriptionsstartstellen zu analysieren.

Phosphorylierung ist eine der häufigsten und wichtigsten posttranslational Proteinmodifikationen. Im Verlauf des Sequenzierung des humanen Genoms wurden 518 Gene für Proteinkinasen entdeckt und es wird angenommen, dass im zellulären Proteom mehr als 500 000 Phosphorylierungsstellen existieren. Die Proteinphosphorylierung spielt eine

entscheidende Rolle in der Signalisierung vieler verschiedener Zellprozesse wie zum Beispiel der interzellulären Kommunikation, dem Zellwachstum, der Differenzierung der Proliferation und der Apoptose. Während bei der massenspektrometrie-basierte Identifizierung und relativen Quantifizierung von einfach phosphorylierten Peptiden in den letzten große Fortschritte erzielt wurden, und die Analyse tausender Phosphopeptide mittlerweile häufig routinemäßig durchgeführt werden kann, bereitet die massenspektrometrische Analyse mehrfach phosphorylierter Peptide nach wie vor große Probleme. Der niedrige pKa-Wert der Phosphatgruppe, und die damit einhergehende negative Ladung ist die Hauptursache für die Probleme bei der Analyse mehrfach phosphorylierter Peptide. Die mehrfache negative Ladung dieser Peptide führt zu einer ausgeprägten Neigung zur Komplexbildung mit mehrwertigen Metallionen (wie z.B. Fe^{3+}), welche zu einer dramatischen Verschlechterung der chromatographischen Eigenschaften dieser Peptide führt (Peak-Tailing), zu einer Verschlechterung der Ionisierungseffizienz, und zu einem ungewöhnlich niedrigen Protonierungsgrad im Positivionen-Modus, welcher diese Peptide für eine Fragmentierung mittels ETD ungeeignet macht.

Im Rahmen dieser Arbeit wurde mittels chemischer Modifikation der Phosphatgruppe versucht sowohl die Detektion von mehrfach-phosphorylierten Peptiden, als auch die Lokalisierung von Phosphorylierungsstellen zu verbessern. Hierfür wurden die Phosphatgruppen unter Verwendung des Aktivierungsreagenzes EDC in hydrolysestabile, aromatische Phosphoramidate überführt. Die durch diese Modifikation erzielte Ladungsumkehr führt wie erwartet zu einer verbesserten Signalintensität bei den entsprechend modifizierten Phosphopeptiden, sowie zu einem verbesserten Fragmentierungsverhalten bei ETD, und somit letztlich zu einer verbesserten Lokalisierbarkeit der Phosphatgruppe innerhalb des Peptids.

Acknowledgment

Contents

Abstract	i
Zusammenfassung	i
Acknowledgment	v
1. General Introduction	1
1.1. Proteomics Background	1
1.2. Protein Analysis by Mass Spectrometry	2
1.2.1. Ionization Techniques	3
1.2.1.1. Matrix-Assisted Laser Desorption/Ionization (MALDI)	4
1.2.1.2. Electrospray Ionization (ESI)	5
1.2.2. Mass Analyzers	6
1.2.2.1. Quadrupole (Q) Mass Analyzer	7
1.2.2.2. Time of Flight (TOF) Mass Analyzer	7
1.2.2.3. Orbitrap Mass Analyzer	10
1.2.3. Detector	10
1.3. MS/MS Fragmentation Methods	11
1.3.1. Collision Induced Dissociation (CID) and High-Energy Dissociation (HCD)	12
1.3.2. Electron Transfer Dissociation (ETD)	13
1.3.3. Electron Transfer and Higher-energy Dissociation	13
1.4. Liquid Chromatography Tandem Mass Spectrometry	14
1.5. Posttranslational Modifications (PTMs)	14
1.6. Proteins and PTMs Quantitation Methods	16
1.6.1. Label Free Quantitation (LFQ)	17
1.6.2. Stable Isotope Labeling-based Quantitation	18
1.6.2.1. Relative Quantitation	18
1.6.2.2. Absolute Quantitation	20
2. Thesis Objectives	23
3. Site-specific Acetylation Degree Quantitation	25
3.1. Introduction	25

Contents

3.2. Histones	26
3.2.1. Structural Chromatin Changes Induced by the Posttranslational Modifications of the Histones	26
3.2.2. Structural Chromatin Changes Induced by the Replacement of Canon- ical Histones with Histone Variants	27
3.3. Histone Acetylation	27
3.4. Quantitation of Lysine Acetylation Degree	29
4. Material and Methods	35
4.1. Reagents	35
4.2. Proteases	35
4.3. Protein and Peptides	35
4.4. Sample Preparation	36
4.4.1. Peptides	36
4.4.2. Proteins	36
4.4.2.1. Recombinant Histone H3	36
4.4.2.2. Preparation of Histone-enriched Fractions	37
4.5. Gel Electrophoresis	37
4.6. In-gel Chemical Acetylation and In-gel Digestion	37
4.7. NanoLC-MS/MS Analysis	39
4.8. Data Analysis	40
4.8.1. MS Raw Data Processing	40
4.8.2. Database Searching	40
4.8.2.1. PEAKS Database Search	40
4.8.2.2. Mascot Database Search	40
4.8.3. Fragment Ion Patchwork Quantification	41
5. Results and Discussion	45
5.1. Method Optimization	46
5.2. Method Validation	46
5.3. Proof of Principle	49
5.4. Identification of Substrate Sites of Histone Acetyltransferases 3 (HAT3) in <i>T. Brucei</i>	53
5.5. Differences in Acetylation Degrees Between Canonical and TSS-Nucleosomes in <i>T. Brucei</i>	63
5.6. Conclusion	67
6. Carbodiimide-mediated 4-ABA Labeling of Phosphosites	69
6.1. Introduction	69
6.2. Qualitative Analysis of Phosphorylation	70

6.3. Challenges In LC-MS/MS Phosphorylation Identification and Quantification	74
7. Material and Methods	79
7.1. Reagents	79
7.2. Proteases	79
7.3. Protein and Peptides	80
7.4. Sample Preparation	80
7.4.1. Peptides	80
7.4.2. Proteins	81
7.4.2.1. 4-ABA In-gel Chemical Modification and In-gel Digestion	81
7.4.2.2. 4-ABA In-solution Chemical Modification and Digestion .	82
7.5. NanoLC-MS/MS Analysis	84
7.6. Data Analysis	84
8. Results and Discussion	85
8.1. Proof of Principle	87
8.2. Method Validation	89
8.2.1. Labeling on the Peptide Level	89
8.2.2. Labeling on the Protein Level	91
8.2.2.1. Reaction Completion and Sequence Coverage	91
8.2.2.2. The Influence of 4-ABA Modification on the Overall Intensity	96
8.2.2.3. Fragmentation and Localization of Phosphorylation Site . .	99
8.3. Conclusion	100
Bibliography	103
Appendix A. Abbreviations	119
Appendix B. Amino Acids and their Masses	123
Appendix A. Peptides and Proteins Sequences	125
Appendix B. Site-specific Acetylation Degree Quantitation	129
Appendix C. Carbodiimide-mediated 4-ABA Labeling of Phosphosites	139
Publications Conferences	143
Curriculum Vitae	145

List of Figures

1.1. From genome to proteome	2
1.2. Types of proteomics.	2
1.3. Electrospray ionization (ESI).	6
1.4. Quadrupole with hyperbolic rods and applied potentials and quadrupole stability diagram.	8
1.5. Schematic description of the principle of linear TOF (LTOF) instrument tuned to analyze positive ions produced by MALDI.	8
1.6. Schematic description of the principle of a TOF instrument equipped with a reflectron.	9
1.7. Three-dimensional schematic of an Orbitrap cell.	11
1.8. Simplified nomenclature for peptide fragments	12
1.9. Illustration of Bottom-up, middle-down, and top-down strategies.	15
1.10. Common quantitative mass spectrometry (MS)-based proteomics workflows.	17
1.11. Principle of relative quantitation PTMs.	19
1.12. Comparison between relative and absolute quantitation of PTMs.	21
3.1. Acetylation enzymatic reaction.	25
3.2. Chromatin structure.	26
3.3. Histone acetylation alters chromatin structure.	28
3.4. General Challenges of PTMs Absolute Quantitation.	30
3.5. Isotopic pattern and MS/MS spectrum of the tryptic peptide TKKTIT-SKKSKKASKGSDAASGVKTAQR of D ₃ -derivatized recombinant histone H3	32
3.6. Reaction scheme of acetic anhydride-1,1\prime- ¹³ C ₂	34
4.1. Workflow for the determination of site-specific acetylation degree	38
4.2. Fragment ion patchwork quantification approach for measuring site-specific acetylation degrees.	43
5.1. General workflow for the fragment ion patchwork quantification.	45
5.2. Retention time of the base peaks of pre-acetylated peptide and the chemical modified isoform.	46
5.3. Recombinant histone H3 <i>T. brucei</i> expressed in <i>E. coli</i>	47

List of Figures

5.4. Reaction completion efficiency of light acetylated Histone H3.	48
5.5. Reaction completion efficiency of 10% acetylated Histone H3.	50
5.6. Coverage map for acetylated recombinant histone H3 from <i>T. brucei</i>	51
5.7. Acetylation degree of recombinant <i>T. brucei</i> H3 resembles 10%, 50% and 90% acetylation at all sites.	52
5.8. Workflow for the identification of substrate site of histone acetyltransferases (HATs).	53
5.9. Acetylation degrees of histone H4 form wild-type and knockout strains. . .	55
5.10. Acetylation degrees of histone H3 form wild-type and knockout strains. . .	56
5.11. Acetylation degrees of histone H2A form wild-type and knockout strains. . .	57
5.12. Acetylation degrees of histone H2B form wild-type and knockout strains. . .	58
5.13. Acetylation degrees of histone H2B.V form wild-type and knockout strains.	59
5.14. Acetylation degrees of histone H2A.Z form wild-type and knockout strains.	60
5.15. Acetylation degrees of histone H4V form wild-type and knockout strains. . .	61
5.16. Acetylation degrees of histone H3V form wild-type and knockout strains. . .	62
5.17. Sequence alignment of the canonical H4 and its variant H4V.	63
5.18. Workflow for comparing the acetylation patterns of canonical and TSS nu- cleosomes.	64
5.19. Acetylation degree of canonical and TSS H4.	65
5.20. Acetylation degree of canonical and TSS H3.	66
6.1. Mechanism of protein phosphorylation and related cellular processes.	69
6.2. Chemical structures of phosphoserine, phosphothreonine and phosphothre- onine peptides.	70
6.3. CID fragmentation scheme of phosphorylated peptide	72
6.4. ETD fragmentation scheme of phosphorylated peptide	73
6.5. Retention time difference caused by aniline labeling of carboxyl and phos- phate groups of PKA.	77
6.6. Reverse of the phosphate negative charge.	77
7.1. Workflow for the 4-ABA in-gel chemical modification	82
7.2. Workflow for the 4-ABA in-solution chemical modification	83
8.1. Challenges in LC-MS analysis of phosphorylation.	86
8.2. Principle of the carbodiimide-mediated 4-ABA labeling.	87
8.3. Experimental MALDI-TOF results of TWpYV before and after 4-ABA modification.	88
8.4. Different approaches for the carbodiimide-mediated 4-ABA labeling.	89
8.5. Reaction completion of singly and doubly phosphorylated peptides modified with 4-ABA.	90

8.6. Sequence coverage of PKA chemically modified in-gel with 4-ABA	92
8.7. Reaction completion of in-gel 4-ABA chemical modification.	93
8.8. Sequence coverage of β -casein chemically modified in-gel with 4-ABA . . .	93
8.9. SDS-PAGE gel of β -casein before and after 4-ABA treatment.	94
8.10. Sequence coverage of β - casein chemically modified in-solution with 4-ABA and digested in-gel.	95
8.11. Reaction completion of in-solution 4-ABA chemical modification.	96
8.12. Comparison of peak intensities of in-gel 4-ABA modified sample and un- modified sample	97
8.13. Comparison of peak intensities of in-solution 4-ABA modified sample and unmodified sample	98
8.14. MS/MS spectra of IGRFpSEPHAR from PKA before and after treatment. .	99
8.15. MS/MS spectra of phosphorylated HKEpSGNHYAM from PKA before and after treatment.	100

List of Tables

- 1.1. List of common matrices used for MALDI MS analysis. 5
- 1.2. Some of the common PTMs, their masses and chemical composition. . . . 16

1. General Introduction

1.1. Proteomics Background

Proteins are macromolecules that are essential parts of a living organism and facilitate a number of processes within the cell. They provide a wide range of important functions, such as catalysis of biochemical reactions, formation of scaffolds that maintains cell shape, cell signaling, immune responses, cell adhesion, nutrient storage, and transportation of molecules (including other proteins) between subcellular organelles and across the cell membrane [1]. Proteomics is the large-scale study and characterization of a complete set of proteins present in a cell, organ, or organism at a given time [2]. Proteomics is more challenging than genomics; a single gene can give rise to a different number of proteins through various processes (Fig. 1.1). For example, it has been reported that the human genome has 21,000 genes that give rise to approximately 1.8 million of protein species have been predicted from the reported 21,000 genes, thus far annotated in the human genome[3]. The dynamic nature of the proteome increases its complexity; unlike the immutable genome, it changes frequently in response to tens of thousands of intra- and extracellular environmental signals[4, 5].

Many cellular processes such as cellular signaling, cellular differentiation, and regulation of gene expression are regulated by proteins. Cellular functions of proteins are regulated elegantly at several distinct levels including transcriptional regulation of gene expression activity, posttranscriptional mRNA regulation by mechanisms such as alternative splicing and micro-RNA, post-translational regulation by covalently bonded modifications, and protein degradation or cleavage. The ability of detecting significant differences between two cellular states is a universal approach to unraveling the cellular and molecular mechanisms involved in a process with an ultimate goal of discovering new markers, diagnostics and indirectly to track new therapeutic routes [6]. Although protein folding and refolding play a crucial role in protein functionality and thus cellular functionality, posttranslational modifications (PTMs) contributes significantly to the structural and functional diversity of the proteins.

In the past few decades mass spectrometry-based proteomics has emerged as a key tool in studying biological systems. The application of techniques based on mass spectrometry for

1. General Introduction

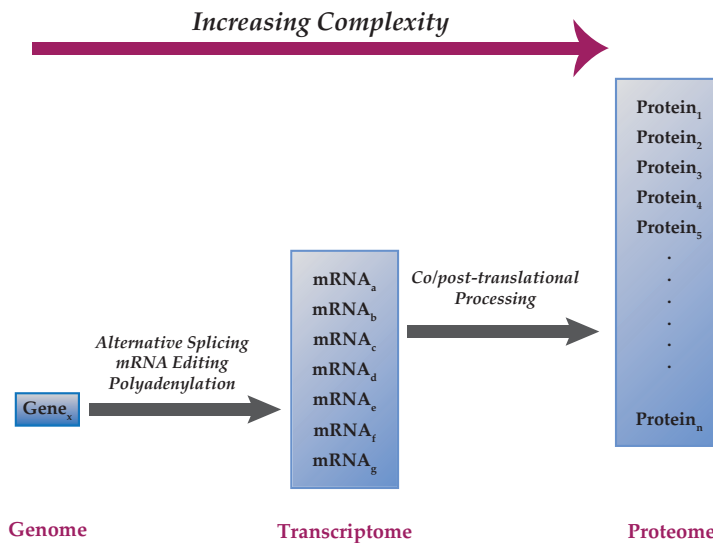


Figure 1.1.: From genome to proteome: a major increase in complexity and dynamics. A single gene can give rise to multiple gene products. Multiple protein isoforms can be generated by RNA processing when RNA is alternatively spliced or edited to form mature mRNA. mRNA, in turn, can be regulated by stability and efficiency of translation. Proteins can be regulated by additional mechanisms, including posttranslational modification, proteolysis, or compartmentalization [3].

the qualitative and quantitative comprehensive proteomic analysis derived from complex mixtures had an essential role in our understanding of cellular processes (Fig. 1.2).

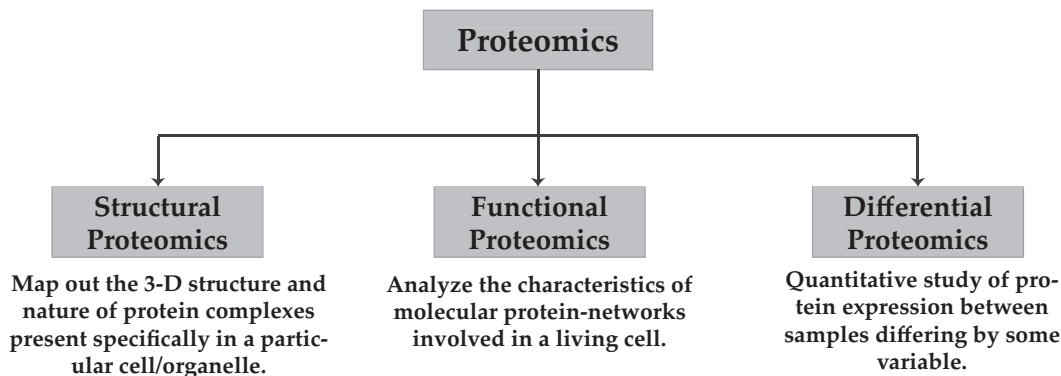


Figure 1.2.: Types of proteomics.

1.2. Protein Analysis by Mass Spectrometry

Mass spectrometry has become an indispensable tool in proteomics, not only because it can be used to identify proteins confidently, but it also helps to quantify their abundances accurately in a sample. MS-based proteomics approaches help to provide valuable information about proteome profile, protein-protein interactions, and the identification and localization of posttranslational modifications (PTMs).

A mass spectrometer is composed of three indispensable parts: a) the ionization source which converts ions for analytes into the gas phase and delivers them into the mass spectrometer, b) the mass analyzer which measures the mass-to-charge ratio (m/z) of ions, and c) the detector which records the measured m/z value of ions. There are many different types of mass spectrometers that use different ways of ionization, analysis, and detection of analytes. However, when choosing an appropriate instrumentation for a proteomic measurement, the following analytical figures of merit should be considered [7]:

1. *Mass Resolution*: the ability to distinguish different ions with similar m/z . Resolution is equal to the full width at half-height of a peak.
2. *Sensitivity*: the amount of signal gain with increasing concentration of analyte.
3. *Dynamic Range*: the ratio of the largest to the smallest detectable signal.
4. *Speed*: the amount of time that is taken to collect a spectrum.
5. *Precision*: the variation in ion abundance when the same analyte is measured multiple times.
6. *Mass Accuracy*: the difference between the measured mass and calculated mass.
7. *Mass Range*: the m/z range that can be measured by a mass analyzer.
8. *Detection Limit*: the lowest quantity of an analyte needed to generate a signal larger than background noise level.

For the work presented in this dissertation, tandem mass spectrometers Orbitrap Fusion TribridTM, LTQ-Orbitrap Velos ProTM were used to analyze proteins and peptides in addition to LTQ and MALDI TOF-TOF mass spectrometers for single peptides analysis. The characteristics of these different mass spectrometers are based on the mass analyzer used to separate the ions according to their individual m/z values and the ionization methodology used to generate these charged ions.

Tandem mass spectrometry (MS/MS) combines two stages of mass analysis with an intermediate fragmentation step to selectively examine a particular ion in a mixture of ions. Most tandem mass spectrometers are hybrid instruments, i.e. different types of mass analyzers, which are described in subsection 1.2.2, are used for the first and second stage of mass analysis. MS/MS permits very rapid, sensitive, accurate measurements of many different types of samples. The common components of a mass spectrometer are described next.

1.2.1. Ionization Techniques

The first step in the mass spectrometric analysis of molecules is to generate ions from analytes. The classic ionization methods like fast atom bombardment ionization (FAB),

1. General Introduction

field ionization (FI)/field desorption (FD), and chemical ionization (CI) are not used much in modern mass spectrometry and proteomics. Converting the polar, nonvolatile and thermally labile biomolecules such as peptides and proteins into the gas phase without extensive degradation and losses was the primary concern faced chemists when analyzing biomolecules. The development of the two soft ionization techniques, matrix-assisted laser desorption/ionization (MALDI) and electrospray ionization (ESI), brought a revolution to the analysis of biomolecules by MS. A salient feature of soft ionization, its capability to transfer a wide range of biomolecules into the gas phase intact, thus making them amenable to mass spectrometric analysis.

1.2.1.1. Matrix-Assisted Laser Desorption/Ionization (MALDI)

Matrix-assisted laser desorption/ionization (MALDI) plays an outstanding role in the analysis of biomolecules such as proteins and peptides. The sample for the MALDI is uniformly mixed and co-crystallized with a large quantity of matrix which provides protection for the molecule from decomposition due to excessive energy. The matrix absorbs the ultraviolet light (nitrogen laser light, wavelength 337 nm) and converts it to heat energy. A small part of the matrix heats rapidly and vaporize, together with the sample. The matrix not only plays a key role in absorbing the laser light energy and causing the analyte to vaporize while protecting it from degradation, but it also serves as a proton donor and receptor, acting to ionize the analyte in both positive and negative ionization modes, respectively [8]. Selecting an appropriate matrix solution is pivotal for achieving proper ionization efficiency, table 1.1 shows some of the common matrices used for protein and peptide analysis. It is a prerequisite that the MALDI matrix must meet a number of criteria simultaneously to achieve mass spectra with high resolution, high sensitivity, and high S/N ratios:

- ▶ It should be vacuum stable.
- ▶ It must promote analyte ionization.
- ▶ It should have a strong absorbance at the wavelength with which the laser irradiates the sample.
- ▶ It must have sufficient acidic/basic properties to enable the generation of quasi-molecular ions from the analyte.
- ▶ Matrix and analyte should be soluble in the same solvent to avoid an even co-crystallization between the matrix and the analyte in order to enhance reproducibility.

MALDI mass spectrum predominantly consists of singly-charged ions, $[M + H]^+$ and $[M + Na]^+$, where M is the molecular mass. There are two main theories for why the

MALDI spectrum consists mainly of singly charged ions. The first, the *Lucky Survivor Model*, is based on the assumption that during ablation, ions recombine with counterions, but a few “lucky” ones are not neutralized by absorption of photoelectrons or electrons from the metallic target, and they can be detected as so-called Lucky Survivor [9]. The second, the *Coupled Photophysical and Chemical Dynamics (CPCD) model*, is based on the assumption that the laser irradiation interacts with the matrix and produces primary ions and radicals (first step) which then, while the plume is still dense, collide with analyte molecule producing the ions we see (second step) [10].

One drawback of MALDI is that the solid matrix sample preparation cannot provide separation of the biological samples, and hence, LC-MALDI was developed. LC-MALDI methods utilize discontinuous separation on LC and drop the fraction of mobile phase onto the MALDI multiple-well plate followed by crystallization and ionization of matrix. Thus, LC-MALDI MS methods have some issues for high-throughput and complex sample analysis.

Chemical Name	Common Abbr.	Sample Compatibility
α -Cyano-4-hydroxycinnamic acid	CHCA	Peptides, small proteins
Sinapic acid	SA	High molecular weight proteins
2,5-dihydroxybenzoic acid	DHB	Peptides, phosphopeptides, glycoproteins
2, 6-dihydroxyacetophenone / diammonium hydrogen citrate	DHAP/DAHC	Phosphopeptides

Table 1.1.: List of common matrices used for MALDI MS analysis.

1.2.1.2. Electrospray Ionization (ESI)

ESI is a gentle method of ionization that converts ions into gaseous phase with the help of strong electrical field. Unlike MALDI, ESI produces vaporized ions continuously from HPLC elution. During ESI process, a continuous stream of biomolecules eluted with the mobile phase containing volatile solvent and precharged analytes is sprayed from a fine tip capillary into a fine aerosol, which is driven by a maintained high voltage (2 - 6 kV) applied between the emitter and the inlet of a mass spectrometer [11]. As extensive solvent evaporation occurs, the size of droplets shrinks until they reach the Rayleigh limit where the Coulombic explosion occurs, and droplets fall apart, producing smaller droplets that repeat this process and ions that enter mass spectrometer for mass analysis [12]. ESI is especially suitable for analyzing large biomolecules, such as proteins and peptides because it produces multiply charged ions with relatively small m/z that falls within the optimal mass range of some mass analyzers. The m/z is calculated according to the formula:

1. General Introduction

$$m/z = [M + nH]^{n+}$$

Under standard conditions, an electrolyte-containing analyte solution is pumped at low flow rates through a metal capillary needle biased to a high potential relative to a counter electrode. The strong electric field present at the capillary tip, the liquid is dispersed into an aerosol of fine charged droplets. Through an orifice in the counter electrode and a differential pumping system, these droplets are transferred from atmospheric pressure into the vacuum of the mass analyzer, and free ions are released from the shrinking droplets. In this work, we used ESI-nanospray ionization that operates with a flow rate of 200 η L/min and 400 η L/min. The use of such low flow rates produce smaller droplets and improve ionization efficiency and sensitivity.

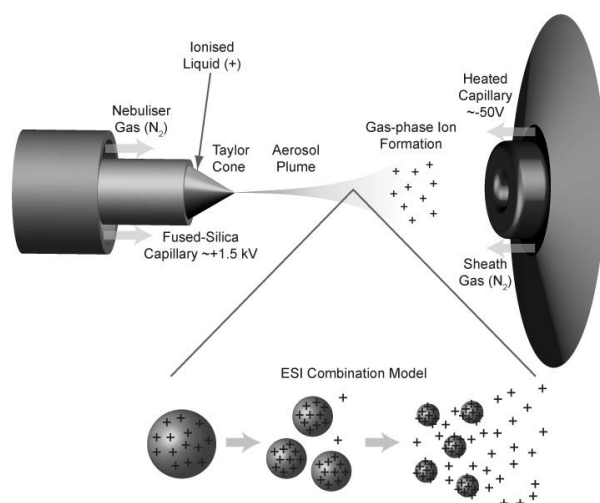


Figure 1.3.: Electrospray ionization (ESI) (*Source: www.LAMONDLAB.COM*)

1.2.2. Mass Analyzers

The mass analyzer is an essential component of the mass spectrometer that separates and measures the ionized analytes with or without fragmentation in the gas phase based on their m/z . There are different types of mass analyzers with different ion separation mechanisms widely used in the detection of ions; all have their own advantages and limitations. Quadrupole (Q), Linear Trap Quadrupole (LTQ), Time of Flight (TOF), Fourier Transform Ion Cyclotron Resonance (FT-ICR), and Orbitrap are the common mass analyzers used but LTQ and Orbitrap have emerged as the most popular ones in proteomic measurement [13]. At their most basic level, mass analyzers can be differentiated based on whether trapping or storage of ions is required to enable mass separation and analysis. Furthermore, tandem mass spectrometry has been developed by combining different

analyzers together to meet the requirements for different applications.

1.2.2.1. Quadrupole (Q) Mass Analyzer

The quadrupole mass analyzer, as the name implies, consists of four parallel hyperbolic metal rods with a rectangular array (Fig. 1.4). Each pair is applied with constant positive DC (direct current) and negative DC to oscillate the electric fields that selectively stabilize or destabilize the paths of ions passing through a radio frequency (RF) quadrupole field. DC and RF potential can affect the trajectories of a beam of ions in the X-Z planes and Y-Z planes separately [14]. For a given DC and RF combination, only a specific predefined m/z range will be allowed to pass through the quadrupole and reach the detector; other ions will be filtered onto the walls by collisions because of their unstable trajectories. This feature of the quadrupole analyzer allows mass filtering of an ion with a particular m/z or scanning of an m/z range by continuously altering the combination of DC/RF voltage [15].

Quadrupoles offer two main advantages; tolerance to relatively high pressures and a significant mass range with the capability of analyzing up to an m/z 4000, which is useful because ESI of proteins and other biomolecules commonly produce charge distributions from m/z 1000 to 3500 [16]. Additionally, quadrupoles have good reproducibility, and they are a relatively low-cost instrument. However, it has a low resolution, and it is not compatible with pulsed ionization methods such as MALDI. Additionally, the single quadrupole by itself is not suitable for tandem mass analysis. In order to perform tandem MS analysis with a quadrupole instrument, some mass spectrometers are designed to combine it with a TOF to establish a QTOF [17] or to place three quadrupoles in series to establish a QqQ [16], these combinations make mass analyzers complementary to each other. In QqQ each quadrupole has a separate function; the first (Q1) serves as a mass filter for the precursor, the second (Q2) is an RF-only collision cell where precursors are fragmented, and the third (Q3) serves as an analyzer of the fragment ions generated in the collision cell (Q2) [16]. Triple-quadrupole mass spectrometers have higher selectivity than single-quadrupole mass spectrometers, which results in less interference of co-eluting compounds and matrix [18]. Furthermore, QqQs have better Signal-to-Noise (S/N), have wider linear range, and they are tandem-in-space mass spectrometers, where ions are scanned in real time rather than accumulated, which allow QqQs to be used in quantitation experiments [18]. Moreover, QqQ have better accuracy and reproducibility, especially at low concentrations.

1.2.2.2. Time of Flight (TOF) Mass Analyzer

The TOF analyzer, the most straightforward and a widely used mass spectrometer, was developed by Stephens in the middle of the 20th century [19]. TOF is the fastest MS

1. General Introduction

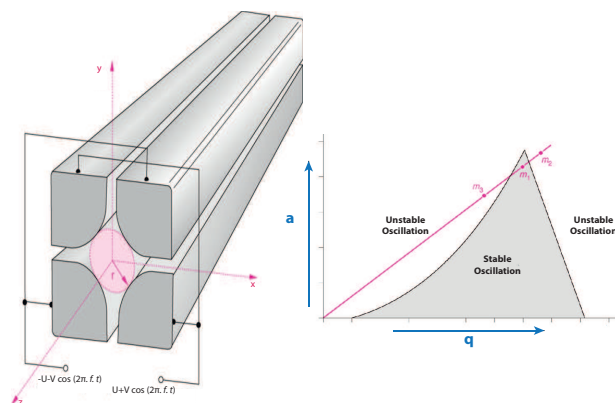


Figure 1.4.: Quadrupole with hyperbolic rods and applied potentials and quadrupole stability diagram.

analyzers, it has the highest mass range, and it is also the choice for the majority of MALDI-based mass spectrometers. It measures the time the ions take to move from the ion source to the detector with the same amount of kinetic energy; it uses an electric field to accelerate the ions through the same potential and then measures the time they take to reach the detector according to the equation [19]:

$$t = L\sqrt{\frac{m}{z} \frac{1}{2U}}$$

Where, t is the time an ion takes to move from the ion source to the detector, L is the distance over which the ion travels, U is the electric potential difference in the acceleration region, m is the mass of the ion, and z is the charge state of the ion. L and U and are constants in an instrument and in MALDI $z = 1$; therefore, the time is only determined by the mass of the ion, lighter ions reaches the detector faster.

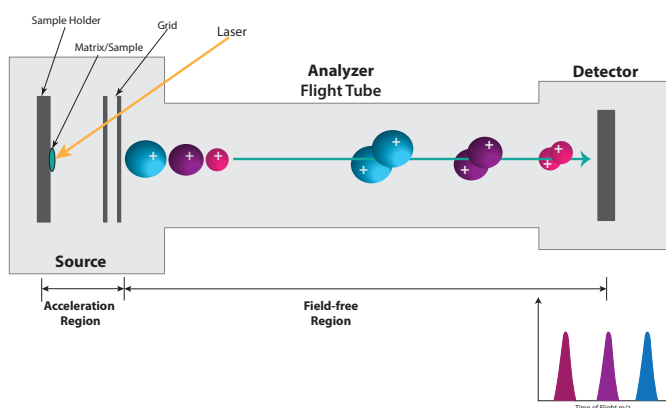


Figure 1.5.: Schematic description of the principle of linear TOF (LTOF) instrument tuned to analyze positive ions produced by MALDI. After their formation during a laser pulse, ions are subject to the applied electric field. Ions are continuously accelerated and drift in a free-field region. They travel through this region with a velocity that depends on their m/z ratios. Ions are thus dispersed in time [20].

Poor mass resolution is considered a critical drawback of LTOF since it's proportional to the flight time and the flight path. There are two possibilities by which the LTOF resolution can be increased [20]:

1. Lengthen the flight tube, which will decrease the performance of the TOF analyzer, due to the loss of ions by scattering after the collision with the gas molecules.
2. Increase the flight time by lowering the acceleration voltage but this will lead to reduced sensitivity.

Another way to enhance LTOF mass resolution is to use a reflectron (RTOF) consists of a series of equally spaced grid electrodes connected through a resistive network of equal-value resistors. It acts as a mirror and uses a constant electrostatic field to deflect the ion beam toward the detector. The reflectron is located behind the field-free region facing the ion source. Unlike LTOF, the detector for RTOF is positioned adjacent to the ion source to capture the arrival of ions after they are reflected. The reflectron corrects the kinetic energy dispersion of the ions leaving the source with the same m/z ratio, as shown in (Fig. 1.6). Ions with higher kinetic energy and hence, with higher velocity will infiltrate the reflectron deeper than ions with lower kinetic energy. Consequently, faster ions will spend more time in the reflectron and will reach the detector at the same time as the slower ions with the same m/z . Although the reflectron increases the flight path without increasing the dimensions of the mass spectrometer, this positive effect on the resolution is of lower interest than its capability to correct the initial kinetic energy dispersion. However, the reflectron increases the mass resolution at the expense of sensitivity and introduces a mass range limitation.

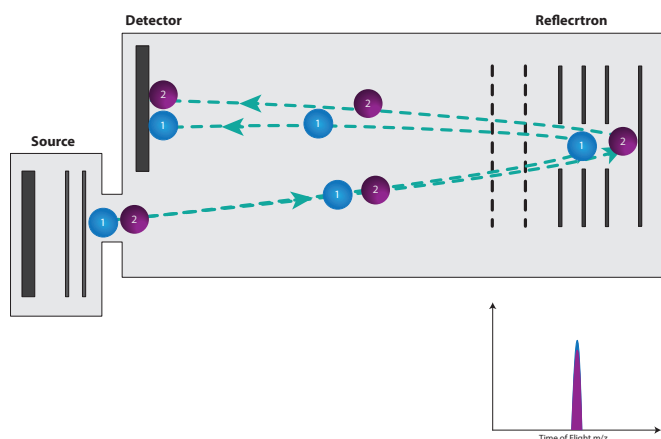


Figure 1.6.: Schematic description of the principle of a TOF instrument equipped with a reflectron. Blue is ion of a given mass with high kinetic energy; purple is ion of the same mass but with a kinetic energy that is too low. The latter reach the reflectron later, with properly chosen voltages, path lengths and fields, both kinds of ions reach the detector simultaneously [20].

1. General Introduction

1.2.2.3. Orbitrap Mass Analyzer

The Orbitrap mass spectrometer is the latest development in trapping devices used as an m/z analyzer; it is famous for its ultimate and superior mass resolution and accuracy. The Orbitrap is an electrostatic ion trap that uses the Fourier transform to obtain mass spectra, it has drawn attention due to its analytical performance in terms of resolution, mass accuracy, space charge capacity and linear dynamic range, relatively small size and cost. The analyzer consists of a “spindle-like central electrode” and two outer cylinder electrodes facing each other (Fig. 1.7) [21]. When a voltage is applied in the chamber, ions are trapped because their movements towards the central electrode by the radial electric field are balanced by an opposing centrifugal force created by a tangential velocity [13]. With a correct potential of the electric field, ions oscillate around the central electrode, while an axial electric field excites ions into a larger trajectory. The axial oscillation of ions is harmonic, and the frequency of ion axial oscillation is simply related to the m/z . The induced ion current is detected by outer electrodes, and then converted into frequency by Fourier transform, finally output into a mass spectrum.

An LTQ-Orbitrap has emerged as the most popular mass spectrometer in proteomic measurements. It possesses a high resolving power analyzers (ion trap, Orbitrap) combined with multiple fragmentation techniques e.g. collision induced dissociation (CID), high-energy collision dissociation (HCD), and electron transfer dissociation (ETD) and) that make it a powerful tool for protein and peptide identification. It has a dual-pressure linear ion trap, the first ion trap operates at high pressure, which allows more efficient ion trapping, isolation, fragmentation and the second one operates at low pressure, which enables faster spectra acquisition. Precursor ions are then axially ejected from the LTQ and trapped in the C-trap. The ions are then squeezed into a small cloud and injected to the orbitrap, where they are electrostatically trapped while rotating around the central electrode and perform axial oscillation. The oscillating ions induce an image current into the outer halves of the orbitrap, which can be detected using a differential amplifier. Ions of only one mass generates a sine wave signal.

1.2.3. Detector

The final element of the mass spectrometer is the detector that records either the charge induced or the current produced when an ion passes by or hits a surface. The choice of detector is based on the required detection sensitivity, speed, stability needed, and application-specific requirements, such as the thermal and chemical stability. Detectors must meet a number of desired properties to achieve high efficiency and sensitivity such as: high amplification, fast time response, high collection efficiency, low-cost, large dynamic range, long-term stability and long life [22]. Different types of detectors are used mass

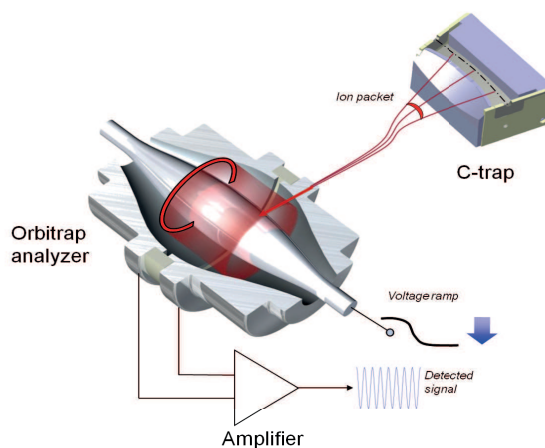


Figure 1.7.: Cross-section of the C-trap and Orbitrap analyzer. Ion packet enters the analyzer during the voltage ramp and form rings that induce current detected by the amplifier. (*Copyright: Thermo Fischer Scientific*)

spectrometry; they can be divided into two classes [22]; 1) Detectors to detect ions of a single mass that arrive sequentially at one point. 2) Detectors that have the ability to count multiple masses and detect the arrival of all ion simultaneously along the plane (array collectors). The detection of ions is always based on their charge, their mass or their velocity. If a detector must ideally be free of any discrimination effect, it's efficiency generally decreases when the mass of the ion increases. This induces limitations for the detection of high mass ions and can compromise a quantity quantitative analysis from these data because the signal decreases exponentially with increases increasing mass.

1.3. MS/MS Fragmentation Methods

Peptide sequencing is the key step in bottom-up mass spectrometry-based proteomics. The definitive protein identification requires fragmentation of the sample inside the mass spectrometer and the analysis of the products generated. In tandem mass spectrometry, soft ionization techniques like electrospray ionization generate intact peptides and proteins in the gas phase without fragmentation or degradation of the biomolecule that enters the first analyzer, which isolates the precursor ion. This precursor ion can then undergo spontaneously or by some activation a fragmentation step to yield product ions and neutral fragments, which are then analyzed by a second mass analyzer that records the m/z values of fragmented ions.

There are different fragmentation methods in proteomics, including collision-induced dissociation (CID), high-energy collision dissociation (HCD), electron transfer/capture dissociation (ETD/ECD), and electron transfer and higher-energy dissociation (EThcD), all of them can provide structural information about peptides and consequently the proteins

1. General Introduction

(Fig. 1.8). The types of fragment ions observed in an MS/MS spectrum depend on many factors including primary sequence, charge state, the amount of internal energy, and how the energy is introduced [23].

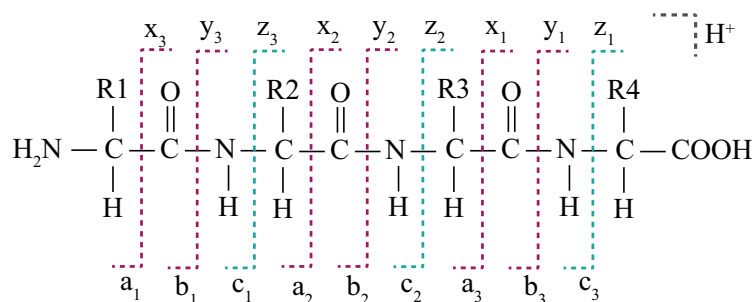


Figure 1.8.: Simplified nomenclature for peptide fragments. Fragments will only be detected if they carry at least one charge. If this charge is retained on the N-terminal fragment, the ion is classed as either a, b or c. If the charge is retained on the C-terminal, the ion type is either x, y or z. A subscript indicates the number of residues in the fragment [24].

1.3.1. Collision Induced Dissociation (CID) and High-Energy Dissociation (HCD)

CID is widely used in peptide sequencing by MS/MS. During CID, protonated peptides are accelerated by an electric potential in the vacuum of the mass spectrometer, internally heated, and collided in the collision cell with an inert gas such as helium or nitrogen [17]. The multiple collisions of the precursor peptide with the inert gas induce a vibrational energy which redistributes along the backbone of the peptide ion [25]. The vibrational energy rapidly increases the internal energy of the peptide ion, which subsequently results in dissociation of the amide bonds along the peptide backbone.

CID fragmentation primarily generates *b*- and *y*- fragments ions from the precursor peptide, which can be used to infer the amino acid sequence. Occasionally, the *b*- fragment ions are subject to loss of a carbonyl group (C=O, ≈ 28 Da), thereby generating *a*- fragment ions [26]. Due to the slow heating, energetic feature associated with this method, internal fragmentation and neutral losses of H₂O, NH₃, and labile PTMs are common [27].

One very common alternative fragmentation technique is the HCD, which is basically a high energy CID [26]. The fragmentation pattern of HCD is featured with higher activation energy and shorter activation time comparing the traditional ion trap CID. HCD also generates *b*- and *y*-type fragment ions and the high energy of HCD leads to a predominance of *y*-ions; *b*-ions can be further fragmented to *a*-ions or smaller species [28]. The precursor ions are isolated in the ion trap and then injected, via the C-trap, to an octopole collision cell where precursor ions are collided with nitrogen gas to generate fragment ions. The fragment ions are then send to Orbitrap via the C-trap for mass measurement. Because Orbitrap measures HCD generated fragment ions with high resolution, high-quality MS/MS

spectra, and high mass accuracy. HCD spectra contain richer fragment ions, which provides more information content for peptide identification, post-translational modification identification and also quantitative proteomics analysis based on iTRAQ [29] and TMT [30] isobaric tags.

Despite the fact that HCD provides more information than CID, both are not suitable for fragmentation of intact proteins, and peptides with labile post-translational modifications, such as phosphorylation [31].

1.3.2. Electron Transfer Dissociation (ETD)

Electron transfer dissociation (ETD) is different from the CID-based methods in many ways; it is a radical induced fragmentation method at low energy [31]. In ETD fragments, the initial radical anion fluoranthene transfers an electron to a protonated peptide by electron transfer reaction. The odd-electron peptide undergoes rearrangement and induces the fragmentation of the peptide backbone, resulting in a *c*-*z* type cleavage of the C α -N bond [32]. This process generates predominantly *c*- and *z*-type ions rather than the typical *b* and *y* ions generated in CID [31]. Due to the radical-directed cleavage of the peptide backbone, ETD is particularly suited for the structural analysis of modified peptides with labile side chains, since these stay intact during backbone cleavage. This is attractive for the analysis of peptides with labile modifications, which are partly or completely lost in CID/HCD, such as phosphorylation. Another unique feature of ETD is that the electron transfer reaction can induce the breakage of disulfide bond, which is not available in other fragmentations, making ETD a powerful tool to determine the disulfide linkage in protein structure.

ETD perform better than CID or HCD on higher charge states, and thus, it is suitable for fragmenting peptides with a 3+ charge or higher. However, it may yield a low number of total identifications due to its slow scan rate .

1.3.3. Electron Transfer and Higher-energy Dissociation

Electron-transfer and higher-energy collision dissociation EThcD is a 2-fold fragmentation strategy combining ETD and HCD fragmentation [33]. In the first step of EThcD, an ETD fragmentation of the peptide precursor generates *c*- and *z*-ions in addition to the undesired unreacted precursor ion. In the second step, all generated ions including the unreacted precursor ions are subjected to HCD, which yields *b*/*y*- and *c*/*z*-type fragment ions in a single spectrum. The *b*/*y*-ions are generated from the unreacted precursor and the yield of *c*/*z*-ions increases by the fragmentation of the charge reduced precursor.

EThcD is suitable for the analysis of phosphorylation, and it can be applied to doubly charged peptides as well as to the highly charged peptides.

1.4. Liquid Chromatography Tandem Mass Spectrometry

Liquid chromatography (LC) is a powerful and robust tool for discovery and validation studies in LC-MS proteomics. It provides a very high sensitivity and wide dynamic range, which are essential to confidently analyze complex biological samples. The analysis of proteins and their PTMs by LC-MS can in principle be achieved by three different strategies; Top-Down, Middle-Down and Bottom-Up mass spectrometry (Fig. 1.9). In Top-Down approach, intact protein ions are introduced into the gas phase by ESI and are subsequently fragmented in the mass spectrometer, yielding the molecular masses of both the protein and the fragment ions. If a sufficient number of informative fragment ions are observed [34], this analysis can provide a complete description of the primary structure of the protein and reveal all of its modifications, as well as any correlations that exist between these modifications [35]. Middle-Down approach is emerging as an attractive alternative to Top-Down analysis. In this approach proteins are proteolytically digested with enzymes that target less abundant amino acid residues than trypsin, such as GluC or AspN to generate peptides with masses greater than 3 kDa [36]. MS fragmentation methods such as electron capture dissociation (ETD), used alongside the middle-down approach, allow sequencing of large polypeptides with highly diverse kinds posttranslational modifications (PTMs) of modified and by increasing the coverage of peptide sequences and the retention of labile PTMs [37]. In Bottom-up proteomics, whereby enzymatically digested peptides from complex samples/mixtures are used to identify proteins and their post-translational modifications (PTMs), is the mainstream proteomics approach [38]. This is largely due to the ease and efficiency of detecting enzymatically generated peptides by liquid chromatography electrospray ionization mass spectrometry (LC-ESI-MS) versus top-down MS analysis of intact proteins [38] that might precipitate under the LC conditions.

1.5. Posttranslational Modifications (PTMs)

Posttranslational modifications (PTMs) are covalent processing events that change the properties of a protein by proteolytic cleavage or by addition of a modifying group to one or more amino acids [39]. PTMs occur on most amino acid residues in proteins except the side chains of leucine, isoleucine, valine, alanine and phenylalanine [40]. In addition, there is one other type of PTM generated by protein backbone cleavage, termed proteolysis. Multiple different PTMs that are present on the same protein can interact to alter the protein function. It is also possible that individual PTMs on different proteins interacts in a series to form a signaling cascade and regulate various cellular processes.

"Writers" and "erasers" are specific classes of enzymes that modulate the status of PTMs on proteins; "writers" are responsible for adding the modification on the protein (e.g.

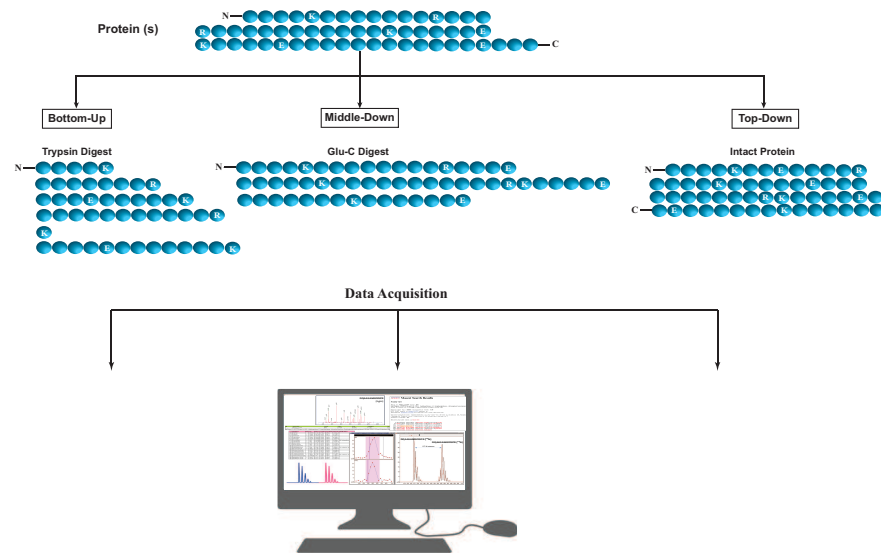


Figure 1.9.: Illustration of Bottom-up, middle-down, and top-down strategies. In the bottom-up approach (left), proteins are digested into short peptides. The middle-down approach (middle) uses a limited digest (e.g., Glu-C or Asp-N), while the top-down approach (right) uses intact proteins without any digestions. All peptides and proteins are acquired under approach-specific conditions and analyzed by bioinformatics tools.

acetyl transferases and kinases) while the "erasers" are responsible for the removal of these PTMs (e.g. deacetylases and phosphatases) [41]. In addition to these enzymes, "readers", which are specific protein domains that can recognize specific PTM (e.g. bromodomain) can be found [41].

Traditional methods of PTM identification have relied upon gel stains specific to particular modifications, and immunochemistry [42]. However, these methods are typically limited by the requirement of purified proteins in quantities that are often not available, the limited specificity and availability of antibodies, and/or the capability to identify only a single targeted modification per experiment [39]. Bottom-up proteomics with the help of high-resolution mass spectrometry enables the detection, site-specific localization, and quantitation of a vast number of PTMs [43]. However, for the unambiguous assignment of a given modification to a specific site, it is necessary to show that the mass shift detected in the precursor ion is also observed in the fragment ions carrying the modified amino acid residue. To preserve this information, it is important to choose the suitable fragmentation method since the presence of PTMs on a peptide may alter the dissociation pathway, which may lead to the loss of the information.

So far, over 300 different PTMs have been identified (see the database of protein PTMs at <http://abrf.org/index.cfm/dm.home>), of these, phosphorylation, acetylation, glycosylation, methylation, ubiquitination, and sumoylation are among the most commonly studied and well-characterized and the number of known PTMs is expanding rapidly due to the

PTM Type	Modified Residue(s)	Δ Monoisotopic Mass
Acetylation	[X]@N-term, K, S, C	42.01057
Amidation	[X]@C-term	-0.98402
Carbamidomethyl	C	57.02146
Carboxymethyl	C	58.00548
Carboxylation	M]@N-term, W, K, D, E	43.98983
Citrullination	R	0.98402
Deamidation	Q, N	0.98402
Dimethylation	[X]@N-term, K, R, N, P	28.03130
Disulfide Bond Formation	C	-2.01565
Dehydration	[N, Q]@C-term, [C]@N-term, S, T, Y	-18.01057
Formylation	[X]@N-term, K, S, T	27.99492
Glycosylation	N, S, T	≥ 200
Monomethylation	[X]@N-term, K, R, S, E, D, H, [X]@C-term	14.01565
Nitration	Y, W	44.98508
Oxidation	M, C, H, W	15.99492
Phosphorylation	S, T, Y	79.96633
Pyroglumatic acid from Gln	[Q]@N-term	-17.02655
Pyroglumatic acid from Glu	[E]@N-term	-18.01057
S-Nitrosylation	C	28.99016
Sulifation	Y	79.95682
Trimethylation	[A]@N-term, K, R	42.04695
Ubiquitination	K	383.22810

Table 1.2.: Some of the common PTMs, their masses and chemical composition.

rapid development in proteomics and mass spectrometry in the past few years. Table 1.2 lists some of the common PTMs, in this thesis acetylation and phosphorylation are highlighted and discussed in greater details in following chapters.

1.6. Proteins and PTMs Quantitation Methods

Molecular characterization of novel PTMs by LC MS/MS involves 4 essential aspects: i) confident identification of the peptide sequences containing PTMs, ii) precise localization of PTM sites, iii) accurate determination of PTM mass shifts, and iv) accurate quantitation of the occupancy. Although the identification of many PTMs is straightforward, the determination of the exact site of the PTM is not always feasible, and despite the great importance of PTMs quantitation for biological function, their quantitation has been hampered by a lack of suitable methods. The majority of methods available so far

for protein quantitation, which are essential for gaining insights into protein functions and dynamics in biological systems are label free quantitation methods and stable isotope-based quantitation methods, each have specific strengths and weaknesses. The choice of the quantitation method is generally based on the biological question and nature of the samples and affected by the available mass spectrometer and the experience of the user.

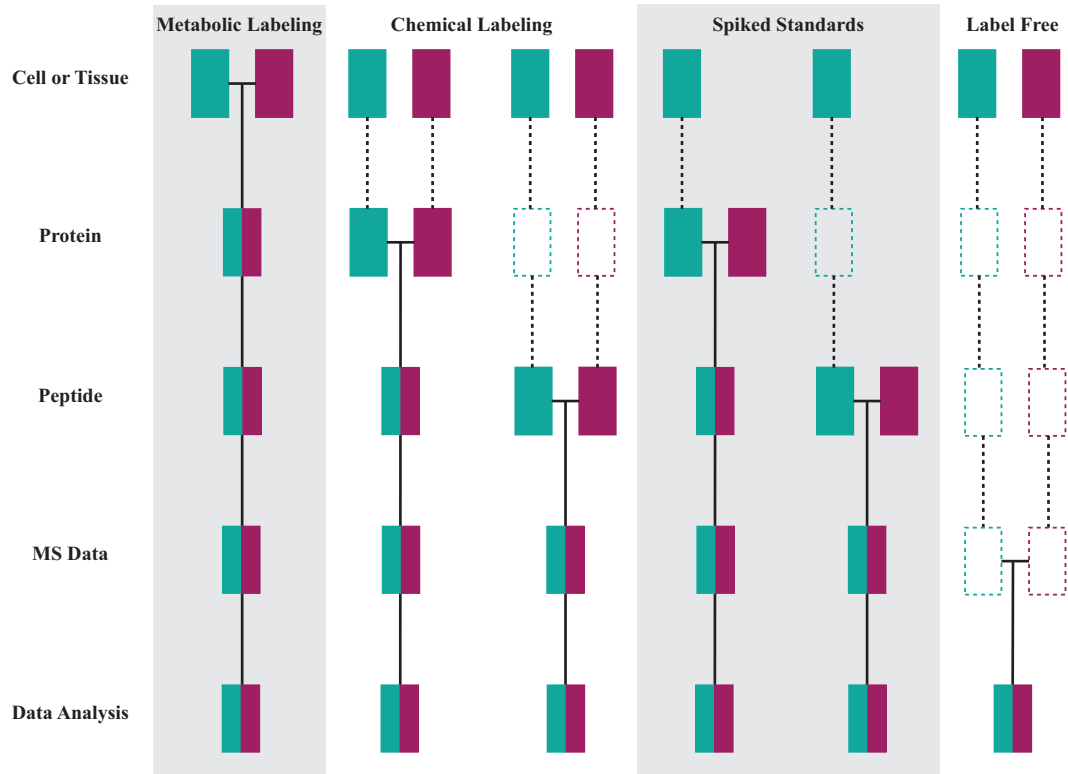


Figure 1.10.: Common quantitative mass spectrometry (MS)-based proteomics workflows. Purple and green boxes represent two experimental conditions. Horizontal lines indicate when samples are combined. Dashed lines indicate the points at which experimental variation and thus quantification errors can occur. (Source: Quantitative mass spectrometry in proteomics: a critical review [44])

1.6.1. Label Free Quantitation (LFQ)

Label-free quantitation is widely used because it's applicable to samples from any source and it allows simultaneous identification and quantification of proteins without a laborious and costly process of introducing stable isotopes into samples, and thus reduce sample loss and the overall assay cost. There are several methods commonly used for label-free quantitation that can employ either relative or absolute quantitation, such as selected reaction monitoring, ion intensity, and spectrum counting. Since label-free methods do not require additional sample preparation for isotope labeling, they reduce sample loss and the overall assay cost. However, because samples to be quantified are prepared and measured separately, label-free approaches have limited quantification performance in terms of accuracy,

1. General Introduction

precision, and reproducibility [45]. To overcome these limitations, a strategy involving the combination of selected reaction monitoring on a QqQ has been developed for relative and absolute quantitation [46]. When an internal standard is used in selected reaction monitoring, a calibration curve can be established and absolute quantitation can be achieved. This strategy combines the liquid chromatography retention time and the m/z of precursor and product ions to generate time-based MS/MS transitions, thus eliminating false positive results and increasing sensitivity.

1.6.2. Stable Isotope Labeling-based Quantitation

Stable isotope labeling quantitation methods are widely used in quantitative proteomics, they employ the incorporation of heavy nonradioactive stable isotopes - in general, ^{13}C , ^{15}N , ^{18}O , and ^2H that change the mass of the protein or peptide without affecting their analytical or biochemical properties. The physiochemical properties of these heavy isotopes are identical to their light counterparts (except ^1H and ^2H exhibit slight differences), which ensures the same behavior of the sample during preparation, LC elution, and ionization. There are two main quantitation methods available in isotope labeling quantitative proteomics, *Relative* and *Absolute* quantitation.

1.6.2.1. Relative Quantitation

In this method, two samples in different states are compared; control is light, and the sample of interest is heavy labeled. These two samples are then mixed in 1:1 ratio before they are subjected to MS analysis. Relative quantitation is applied to compare PTMs among samples and only shows if the protein is up- or down-regulated by measuring the fold change between light and heavy labeled proteins without knowing the exact degree of modification of an individual protein (Fig. 1.11). The advantage of this method is that both modified and unmodified peptides co-elute from the LC at the same retention time, but they can be distinguished by the mass shift introduced by the heavy stable isotope.

The stable isotope labeling methods for relative quantitation are subdivided into two classes: metabolic labeling methods that utilize biological incorporation of stable isotope labels into living cells and chemical methods that utilize a reagent to introduce the mass tag.

1. **Metabolic Labeling:** In metabolic labeling strategies, two populations of cells are cultivated in cell culture. The two culture media are similar except one is containing a particular natural (light) amino acid and the other contains the heavy version of the amino acid [47]. After a number of cell divisions, each occurrence of this particular amino acid will be replaced by its isotope-labeled version and which will

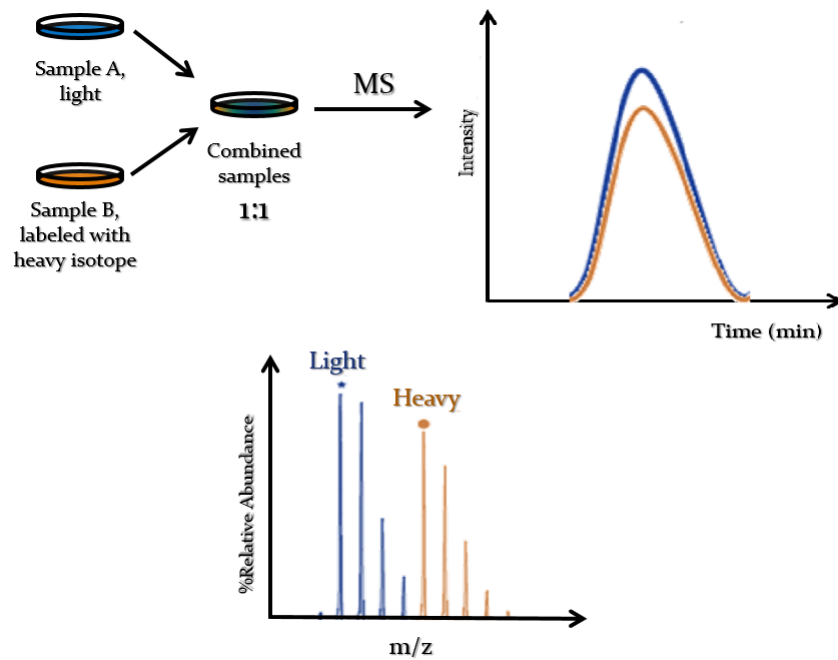


Figure 1.11.: Principle of relative quantitation of PTMs. Two identical samples, one is light and the other is labeled with heavy isotope, are mixed in 1:1 ratio and subjected to MS analysis. Light and heavy peptides co-elute from the LC-system. The MS spectra of the light and heavy peptides can easily be distinguished by the mass shift introduced by the heavy isotope. By comparing the peak intensities of light and heavy peptides, the fold change between the samples can be determined.

be incorporated into all newly synthesized proteins, which generates a labeled standard for every protein in a sample of interest. Since there is hardly any chemical difference between the labeled amino acid and the natural amino acid isotopes, the cells behave exactly like the control cell population grown in the presence of natural amino acid. Proteins of each culture are then combined in 1:1 ratio prior digestion, resulting in peptides in light and heavy form of the same sequence. These peptide mixtures are then analyzed by LC-MS/MS, and they can be distinguished by the mass difference of between heavy and light peptides and relatively compared by the ratio of peak intensities. The typical two approaches of metabolic labeling are ^{15}N and SILAC (Stable isotope labeling by amino acids in cell culture), the advantage of these methods is that the isotopic incorporation rate is relatively high, often more than 98% [44]. ^{15}N labeling can only be used for micro-organisms as well as a few number of higher (multi-cellular) organisms, while SILAC was originally described for mammalian cell culture [47], but it has been adapted to label yeast, bacteria, and plants as well. Therefore, SILAC is gaining popularity because of its ease of implementation, the high quality of quantitative data obtained, minimized error, robustness, reproducibility, and compatibility with existing experimental workflows. However, the drawback of SILAC and any other metabolic labeling methods is that

1. General Introduction

they depend on endogenous labeling of cell lines, so they are not suitable for use with primary tissue such as patient samples and are restricted to the examination of biological systems which can be maintained in controlled conditions. In addition, the method may be hampered by in-vitro alteration of labeled arginine to proline which needs a computational approach to be corrected.

2. **Chemical Labeling:** In chemical protein and peptide labeling methods, any reactive group may be modified using stable-isotope labels. It includes chemical labeling using isotope coded affinity tags (ICAT), isobaric tags (iTRAQ), tandem mass tags (TMT), dimethylation labeling. ICAT is only suitable for cysteine-rich proteins. It labels cysteines as heavy and light versions before they are combined, digested and analyzed by LC-MS [48]. TMT and iTRAQ are applied after digestion; they target the primary amine groups of the resultant peptides. They consist of three parts: the peptide reactive group (amine specific), the reporter mass group and the mass balance group [49]. They are built such that peptides from differentially labeled samples have identical mass (isobaric) but can be distinguished following fragmentation inside the mass spectrometer by the differentially isotope encoded reporter ions in the lower mass range region of tandem MS (MS/MS) spectra. The intensities of the reporter ions form the basis for quantification [49]. Dimethylation labeling also targets the primary amines, it can be applied in both the protein level and peptide level.

1.6.2.2. Absolute Quantitation

The determination of the absolute concentration of protein (or proteins) present in a complex sample is invaluable for the determination of the stoichiometry of proteins within a sample that facilitate the understanding the complicated molecular mechanism for cellular responses. These responses are often controlled through a coordinated direct and indirect interactions between proteins present within the cell, that allows the cell to communicate a response across the cellular compartments to recruit and produce other critical proteins needed [50].

Absolute quantitation of proteins requires the addition of stable-isotope-labeled reference standard synthetic peptides of known absolute quantity to the sample [51]. These synthetic peptides correspond to particular native peptides created during digestion of the sample under investigation. Because the amount of labeled peptide spiked into the sample prior to digest is known, the amount of unlabeled peptide, and thus protein, can then be determined. The major roles of these standards are to correct variations in experiments and establish standard curves, based on comparing the MS signal intensities of the unknown sample with those of the standard [52]. Several methods utilizing such reference peptides and proteins exist, including AQUA (absolute quantification), iTRAQ (isobaric

tags), QconCAT, PSAQ (protein standard absolute quantification), absolute SILAC, and FlexiQuant [51]. The selected reaction monitoring (SRM) or multiple reactions monitoring (MRM) scanning mode is usually employed to quantitate in the mass spectrometer, such as triple quadrupole. The parent ions of interest are selected to fragment in CID, and then specific product ions are scanned and recorded. The peak areas of product ions are used to quantitate the amounts of the analytes. This method has advantages when analyzing complicated samples, such as proteins from the serum, since it can exclude interferences of co-eluted analytes [44].

The absolute quantitation of PTMs at a specific residue is a major hurdle. To address abundance and stoichiometry, specific enrichment is generally required to be able the detection of modified peptides from the high background of unmodified peptides. Additionally, modified and unmodified peptides are expected to have significantly different ionization efficiencies due to their different physicochemical properties. Furthermore, the protease cleavage is influenced by the presence of PTMs at or close to the protease cleavage site, which generates peptides with different sequences that can not be compared.

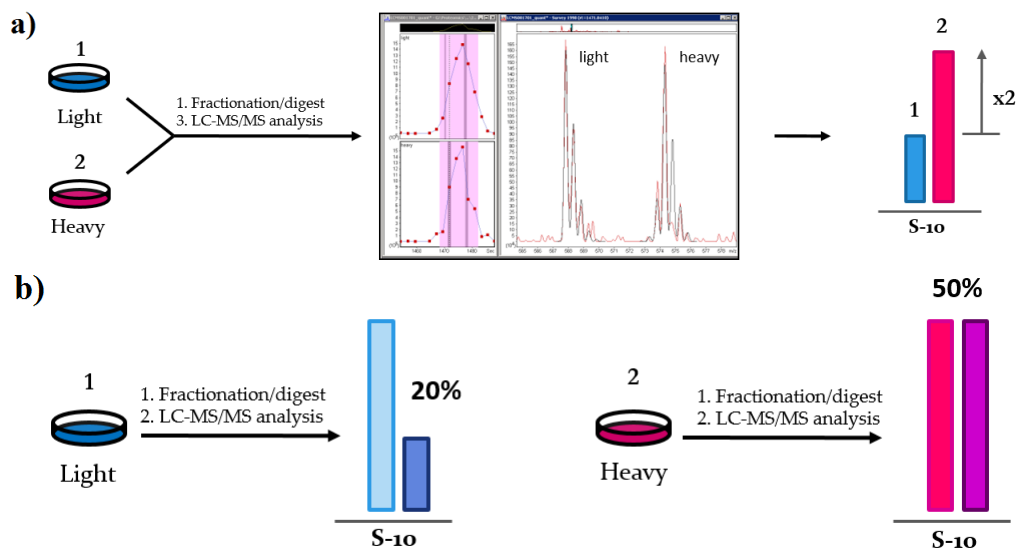


Figure 1.12.: Comparison between relative and absolute quantitation of PTMs, e.g. phosphorylation of S10 of any given protein. **a):** In relative quantitation of PTMs, only the fold change between the light and heavy is known but not the exact degree of PTM at a specific-residue. **b):** In absolute quantitation of PTMs, the exact degree of PTM at a specific residue is measured for each sample separately.

2. Thesis Objectives

Posttranslational modifications (PTMs) play a key role in many cellular processes. In the past decade, a number of mass spectrometry-based methods have been developed to identify, characterize and quantify various types of PTMs. Despite the success of these methods in deciphering the biological functions of PTMs, there is still room for further enhancements. The identification of many PTMs is straightforward; however, the determination of the exact site and the degree of the PTM are not always feasible. Many PTMs impair the mass spectrometric analysis of the corresponding peptides. The addition of a modifying group changes the physicochemical properties of the modified peptides and influences the proteolytic cleavage sites, which generates different peptides that hamper the quantitative analysis. In general, the quantitative analysis of PTMs mostly relies on relative quantification, which only allows the measurement of fold changes between different conditions without knowing the degree of modification of individual site. MS-based methods for measuring the site-specific modification degree are vastly missing.

Acetylation of histone tails serves as an important regulator of gene expression. There are many well-established MS-based methods for the relative and absolute quantitation of lysine acetylation. However, the determination of site-specific acetylation degrees for all lysine residues within histone tails has not been possible with currently existing methods.

Phosphorylation is one of the most important and well-studied PTMs; it affects most of the cellular activities in eukaryotes. MS-based methods for the identification and relative quantification of singly phosphorylated peptides have been greatly improved and well established. However, the identification efficiency for multiply phosphorylated peptides is still insufficient compared to that for singly phosphorylated ones. Due to analytical limitations, detection of multi-phosphorylated peptides and correct localization of phosphosite(s) are formidable challenges.

The aim of this thesis was to develop improved MS-based methods that allow:

1. Accurate quantitation of the site-specific acetylation degrees, specifically in histones,
2. Detection of multi-phosphorylated peptides, and
3. Improved localization of phosphorylation site(s).

3. Site-specific Acetylation Degree Quantitation

3.1. Introduction

Acetylation is one of the major and oldest discovered protein modification in eukaryotes. It is catalyzed by a wide range of acetyltransferases, where an acetyl group from the acetyl coenzyme A (Ac-CoA) is transferred and attached covalently to various distinct positions within the protein either cotranslationally or posttranslationally at the α -amino group of the protein N-terminus, or at the ϵ -amino group of lysine residue respectively [53].

Protein N $^{\alpha}$ -terminal acetylation is an irreversible modification catalyzed by N-terminal acetyltransferases (NATs) (Fig. 3.1). Studies showed that N $^{\alpha}$ -terminal acetylation occurs in approximately 80% of eukaryotic proteins, while in prokaryotes this modification is considered to be a rare event [54].

The acetylation of the lysine ϵ -amino group is a reversible, dynamic and tightly regulated PTM catalyzed by lysine acetyltransferases (KATs), which transfer the acetyl group of acetyl-CoA (Ac-CoA) to the ϵ -amino group of a lysine residue [53]. This acetyl group can be removed by a family of hydrolytic enzymes called lysine deacetylases (KDAC) (Fig. 3.1).

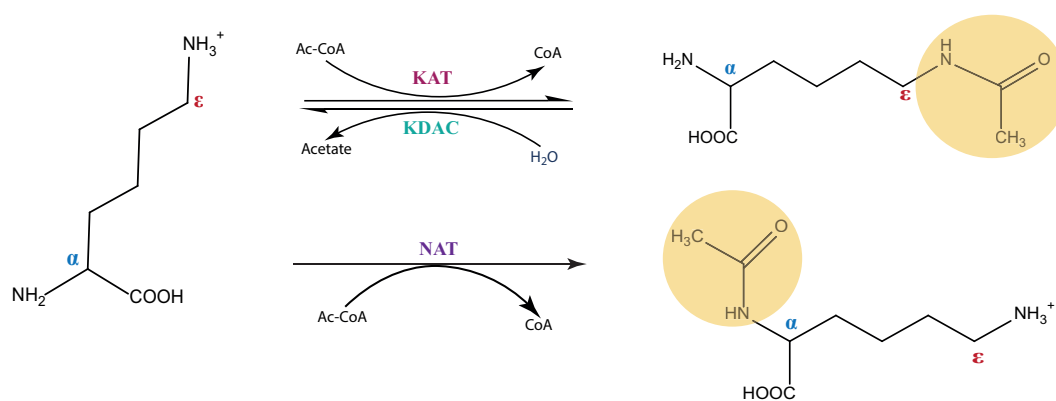


Figure 3.1.: Acetylation enzymatic reaction of lysine ϵ -amino group (top) and N $^{\alpha}$ -terminal of protein (bottom).

Acetylation, in general, is crucial in many cellular processes such as DNA recognition, chro-

3. Site-specific Acetylation Degree Quantitation

matin remodeling, cell cycle regulation, RNA splicing, RNA transcription, metabolism, and cytoskeleton reorganization [55, 56]. Lysine acetylation is often used to regulate DNA or RNA-binding properties of proteins. A prominent example for this are histones, whose tail regions are lysine-rich and often extensively modified by acetylation.

3.2. Histones

Histones are a key architectural and regulatory component of chromatin. The basic repeating unit of chromatin is the nucleosome (Fig. 3.2); 147 bp of DNA wrapped 1.7 times around an octamer of two copies of each of the canonical histones: H2A, H2B, H3, and H4 [57]. Chromatin is a dynamic structure that not only helps to package the entire eukaryotic genome into the confines of the nucleus but also regulates the accessibility of DNA for transcription, recombination, DNA repair, and replication [58]. Structural chromatin changes can be induced by the posttranslational modifications of the histones or by the replacement of canonical histones with histone variants. Chromatin has been categorized into heterochromatin, which is condensed in interphase nuclei and reflects a functionally inactive state of the genome, and euchromatin, the condensed form, which is a marker for active transcription.

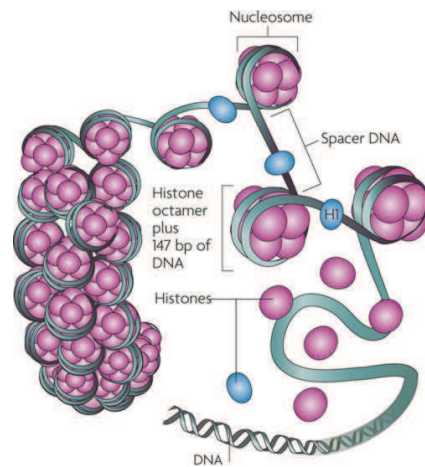


Figure 3.2.: Chromatin structure. Nucleosomes are composed of 147 bp DNA wrapped around an octamer of two H2A/H2B dimers and one H3/H4 tetramer. Histone H1 (light blue) interacts with the linker DNA. (Source: Figueiredo et al., 2009.) [59]

3.2.1. Structural Chromatin Changes Induced by the Posttranslational Modifications of the Histones

A striking feature of histone proteins is that they are subject to a number of diverse reversible (PTMs) such as acetylation [60]. To date, 19 PTMs have been reported to

occur at histones in addition to acetylation [61]. These modifications include mono-, di- and trimethylation, ubiquitination, phosphorylation, sumoylation, propionylation, formylation, citrullination, and ADP-ribosylation [62]. The large number and diversity of histone modifications give cells access to an enormous combinatorial potential to help regulate gene expression or to store information that may be passed on through cell division [41]. These PTMs are proposed to contribute the histone code that dictates the functions of the proteins in gene expression and chromatin dynamics. Many scientists believe that lysine acetylation is the central component of the histone code that controls whether the gene is turned on or off, because acetylation occurs on very specific residues, usually lysine.

3.2.2. Structural Chromatin Changes Induced by the Replacement of Canonical Histones with Histone Variants

Histone variants are non-canonical (non-allelic) variants of histones that can have significant differences in primary sequence. Each canonical histone protein has its own repertoire of variants that differ mostly in their amino acid sequence in the N-terminal region [63]. Histone variants are not restricted in their expression to the S phase but are expressed throughout the cell cycle [64]. Some variants exchange with the preexisting canonical histones during development and differentiation. This replacement often results in the variants becoming the predominant species in the differentiated cell, which lead to the suggestion that the histone variants have specialized functions altering and regulating chromatin dynamics. Like in canonical histones, variants are subject to a number of PTMs [65]. The resulting replacement of modified histone variants might modulate barriers to transcription and perpetuate active chromatin states [64].

3.3. Histone Acetylation

When the chromatin is tightly packed, the histones prevent the access of transcription machinery to the DNA and thus, the transcription activities are suppressed. When histone tails are acetylated by histone acetyltransferases (HATs), the positively charged lysine residues that lie on the outside of the nucleosome are neutralized, which affect the interaction between the lysine residues and DNA, leading to weaker contact with the negatively charged DNA backbone [66]. This modification consequently leads to a loosened chromatin structure, known as euchromatin, which makes the DNA accessible and facilitates the interaction with transcription machinery. When the acetyl groups are removed by histone deacetylases (HDACs), the euchromatin returns to its condensed, transcriptionally repressive heterochromatin conformation (Fig. 3.3). Histone acetylation can also alter chromatin structure by providing a docking site for histone binding proteins. The acetylated lysine residues are recognized by bromodomains, which mediate the specific binding

3. Site-specific Acetylation Degree Quantitation

of these proteins to a distinct acetylated lysine residue or recruit machinery required to stabilize or remodel specific chromatin states.

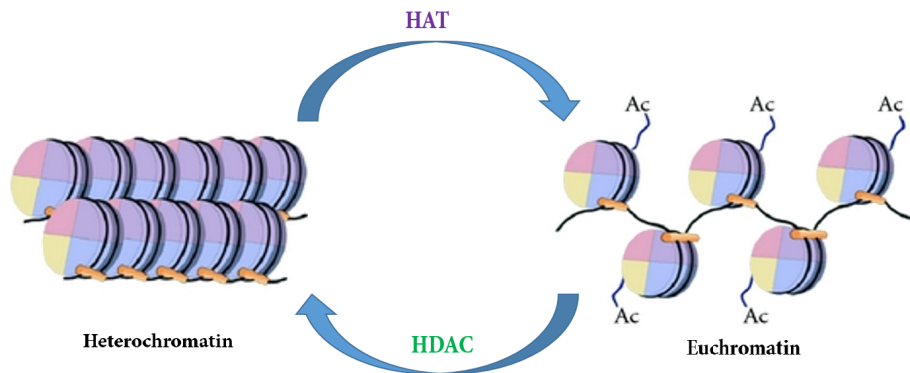


Figure 3.3.: Histone acetylation alters chromatin structure. Shown in this illustration, the dynamic state of histone acetylation/deacetylation regulated by HAT and HDAC enzymes. Acetylation of histones alters accessibility of chromatin and allows DNA binding proteins to interact with exposed sites to activate gene transcription and downstream cellular functions. *Figure was adapted with permission from Rodd et. al 2012 [67]*

The nature of the lysine-rich histone tails, in both canonical and variants, makes acetylation the predominant PTM, that is facilitated by a variety of histone acetyltransferase complexes. Two main superfamilies of HATs have been identified based on their catalytic domains and sequence homology: i) the Gcn5-related N-acetyltransferases (GNATs), named after the founding member Gcn5; and ii) the MYST family HATs, named for the founding members of this family: MOZ, Ybf2, Sas2, and Tip60 [68]. In general, GNATs are involved in cellular growth while MYST HATs are involved in the control of transcription and cell growth and survival [69]. To date, twenty-two lysine acetyltransferases have been found in human, implying a significant role of this modification in a wide range of cellular processes, but many of their specific targets are not fully deciphered yet [70]. In contrast, only six HATs, HAT1, HAT2, HAT3, ELP3a, ELP3b and PHD, are encoded within the genome of *Trypanosoma brucei* [71], which is a unicellular eukaryotic parasite that causes sleeping sickness in humans and livestock in sub-Saharan Africa and is considered as the most primitive eukaryotes. Their impact as pathogens and their simplicity comparing to other eukaryotes create great interest to investigate them in order to understand their cell biology and to find more efficient treatment of the diseases. Given this relatively small number of HATs, *T. brucei* represents an excellent system to investigate the specific target of individual HATs.

To identify the specific enzyme responsible for the acetylation of a specific lysine residue of the histones tails and to quantify the acetylation degree of that specific lysine residue we collaborated with the group of Dr. Nicolai T. Siegel (Institute of Infectious Diseases at University of Würzburg, Germany). Previous studies have shown that the histone H4 N-terminus of *T. brucei* is heavily modified, while, in contrast to other organisms, the histone

H2A and H2B N-termini have relatively few modifications [72]. However, the degrees of these modification has not been reported. Siegel demonstrated that in *T. brucei* HAT2 and HAT3 are responsible for the acetylation of H4K10 and H4K4, respectively [73, 74]. Siegel reported that in *T. brucei* HAT2 and HAT3 are responsible for the acetylation of H4K10 and H4K4, respectively. Using specific antibodies, the acetylation degrees of H4K10 and H4K4 were estimated to be 7% and 73%, respectively [74]. However, neither the other targets for these two HATs nor the acetylation degrees of the other acetylated site were deciphered due to the lack of specific antibodies; the high number of acetylated sites at histone tails made the use of specific antibodies impractical.

3.4. Quantitation of Lysine Acetylation Degree

Localization and the absolute quantitation of the acetylation degree of an individual lysine residue are initial and crucial steps in deciphering the molecular mechanisms of acetylation-related biological processes in all organisms. Despite the great importance of acetylation quantitation for biological function, its quantitative measurements have been hampered by a lack of suitable methods.

In the past, the antibody-based approach was the favored tool for quantifying histones acetylation and other PTMs through various assays [75]. These include immunofluorescence (IF), western blot (WB), and chromatin immunoprecipitation (ChIP), and ChIP-DNA microarray (ChIP-on-chip) [75]. Despite the powerfulness of the the antibody-based approach, the use of antibodies encounters limitations in both identification and quantitation of PTMs. They exhibit promiscuity when two or more histone PTMs are placed too close together [76], many antibodies lack the desired specificity to distinguish subtle differences in the chemical composition of PTMs [77], and they are limited in their ability for a direct determination of acetylation stoichiometries at specific lysine residues within the histones tails [77].

Currently, MS has become the leading method for PTMs quantitation. The majority of mass spectrometry-based methods available so far for acetylation quantitation as well as for other known PTMs are the relative quantitation methods. These methods provide information on how the levels of acetylation at individual sites differ between two samples, without providing information about the absolute acetylation degree on the individual site [44]. SILAC [78], isobaric mass tag labeling [79], and enzymatic ^{18}O labeling [80] have been combined with antibody-based enrichment of peptides containing acetylated lysine residues to analyze the acetylation quantitatively.

For better understanding of the role of acetylation, it is important to know what fraction of each acetylation site is actually modified. Using the absolute quantification (AQUA) method, the amount of acetylated peptide can be compared to that of a known and labeled

3. Site-specific Acetylation Degree Quantitation

peptide standard. However, AQUA standards are expensive and have not been used for proteome-wide quantification of acetylation [81]. Thus, cost-effective absolute quantitation methods were sought to know what fraction of each acetylation site is actually modified.

In general, three main problems hamper the absolute quantitation of the acetylation degree and other PTMs degree; a) Different retention time of modified and unmodified peptides, b) Different ionization efficiencies of modified and unmodified peptides, and c) The influence of the presence of PTMs on the protease cleavage sites, these three issues lead to difficulties in using mass spectrometry-based methods to quantify them (Fig. 3.4).

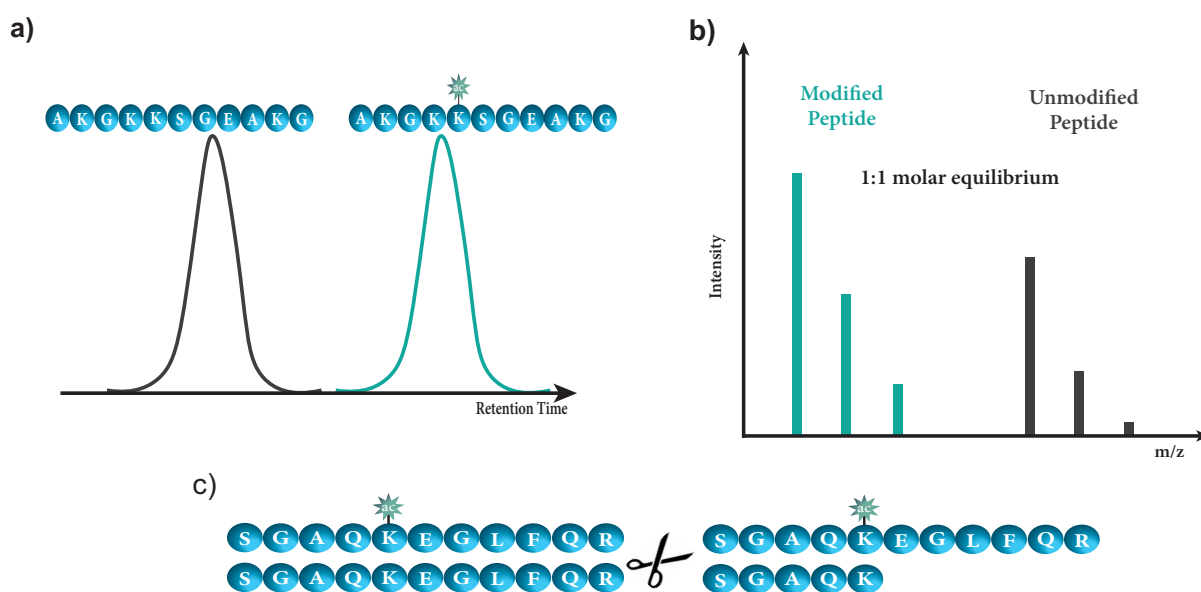


Figure 3.4.: Illustration of the general challenges facing PTMs quantitation **a)** Different retention time of modified and unmodified peptides caused by the different elution of hydrophobic and hydrophilic peptides. **b)** Different ionization efficiencies of modified and unmodified peptide, which are in a molar equilibrium, resulted in different intensities. **c)** Partially acetylated protein digested with the common protease trypsin, which cleaves predominantly the amide peptide bond on the carboxyl side Lys and Arg except before proline. Modified lysine sites generally become “missed cleavages”, producing different peptide segments for the modified and unmodified forms that are then difficult to compare quantitatively.

To circumvent these issues, proteins can be modified *in-vitro* to convert the amino acids to the same chemical species. This *in-vitro* modification eliminates the influence of modified sites on the proteolytic cleavage and generates peptides having the same retention time and similar ionization efficiencies. Subsequently, the monoisotopic peaks intensities of the endogenously modified and the chemically modified peptides can be compared in a quantitative manner. Unfortunately, converting all peptides to the same chemical species causes them to be contained in a single monoisotopic peak, from which we cannot distinguish the original state of each individual peptide. This problem can be solved by using heavy isotopically labeled reagents, which allows the modified peptides to have an identical retention time and ionization efficiencies but their monoisotopic peaks can be differentiated

by the mass shift introduced by the heavy stable isotope.

Despite finding solutions for the general problems that hamper the acetylation quantitation, the accurate quantitation of the site-specific acetylation degree at the N-terminal tail of histones remains challenging due to the occurrence of lysine residues in close proximities.

Several groups have developed different bottom-up approaches for the site-specific acetylation degree. Yuan and coworkers developed a fluorescence assay based on the specific recognition of H3K14ac by the PB1(2) bromodomain in human, to detect changes in H3K14 acetylation degrees [82]. Despite the high sensitivity of this assay to H3K14, it is inapplicable to determine the acetylation degrees of other sites. Quantification of the site-specific acetylation degree using *in-vitro* derivatization techniques with stable isotope-labeled reagents has been proven to be the most efficient and direct approach. Smith and coworkers [83, 84] reported the first method in 2003, where they used a mixture of deuterated (d_6) acetic anhydride and deuterated (d_4) acetic acid as the derivatization reagent to convert all endogenously unmodified lysine residues at N-terminal tail of yeast histone H4 to deuterioacetylated species. D_3 -acetyl derivatization causes a mass shift of $\Delta 3$ Da between protiated acetyl and deuterated acetyl groups. Smith and coworkers succeeded to determine the fraction of *in-vivo* acetylation at residues K5, K8, K12, and K16 of histone H4 using targeted LC-MS/MS methods. With this targeted method, an identification scan was run first, and then the m/z of the peptides of interest containing the lysine residues to be quantified were subjected for further fragmentation. They reported that this method would not be successful if CID of the peptide does not give high yields of fragment ions required for the analysis [83]. Later, Cotter group [85, 86] used this approach to investigate the interplay between methyl and acetyl marks on neighboring acetyl residues within a single tryptic peptide of the histone H3 tail.

Recently, Feller et al. [87] applied D_3 -acetyl derivatization in combination with a targeted MS-based approach to study specific alterations of acetylation signatures in histone tails after the ablation of specific HATs and HDACs. D'Urzo et al. [88] developed a method for the site-specific quantification of lysine acetylation in the N-terminal tail of histone H4 using a double-labeling, targeted MS/MS approach combined with trypsin and chymotrypsin digestion. ϵ -amino groups were first modified with propionic acid anhydride prior digestion. The newly formed peptide N-termini were subjected to a second derivatization step with d_6 - (heavy) or d_0 - (light) acetic acid anhydride. Subsequently, samples were mixed at different ratios and peptides monitored by multiple reaction monitoring (MRM) LC-MS/MS. Bennet and coworkers [89] have also reported a similar method that uses *N*-acetoxy- d_3 -succinimide (D_3 -NAS) instead of d_6 acetic anhydride for chemical acetylation. Where they coupled the chemical derivatization to filter-aided sample preparation (FASP) and subsequently, used targeted MS method similar to Smith et al. Baeza *et al.* [90] developed a method to quantify the stoichiometry of site-specific acetylation at the

3. Site-specific Acetylation Degree Quantitation

proteome-wide scale, without immunoenrichment. In this method chemical derivatization of ϵ amino groups at the protein level was achieved by using $^{13}\text{C}_4$, d_6 acetic anhydride prior performing an in-silico enrichment of the acetyl peptides.

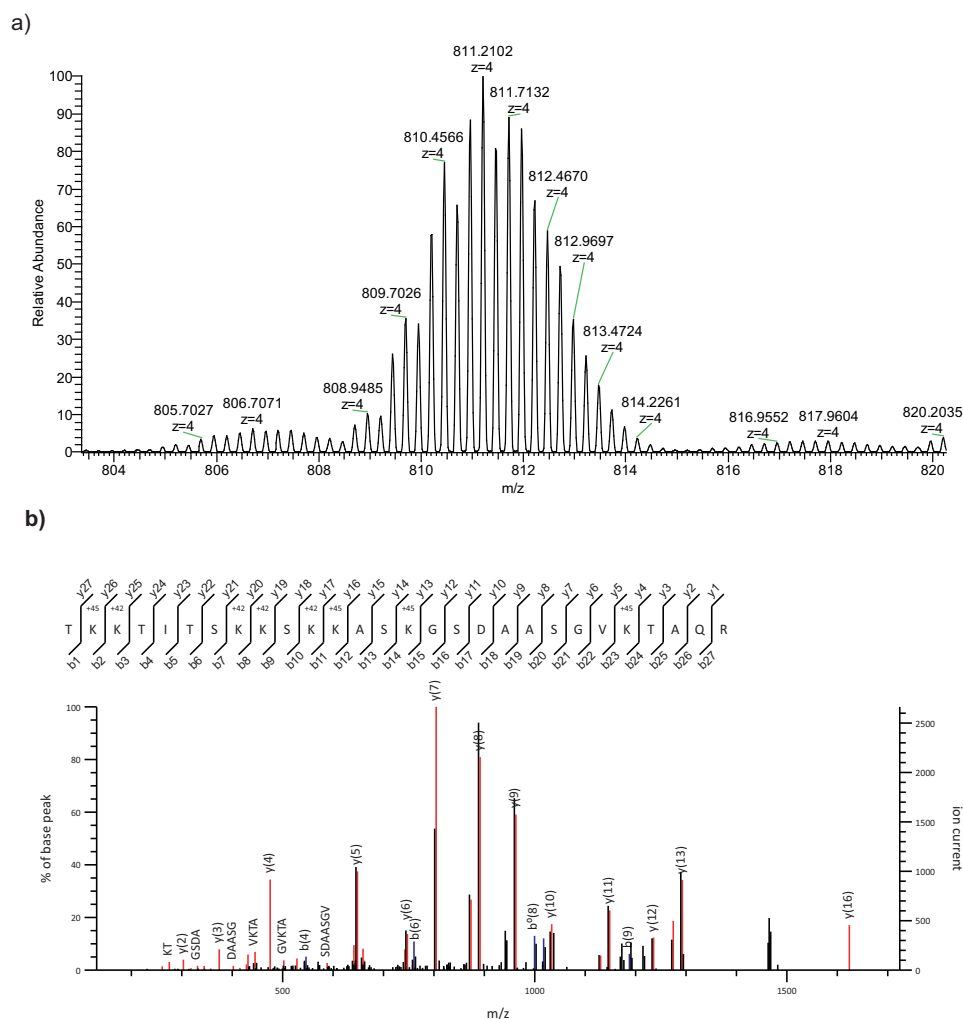


Figure 3.5: a) Isotopic pattern and b) MS/MS fragment ions of D_3 -derivatized recombinant histone H3 tryptic peptide TTKKTITSKKS KSKKASKGSDAASGVKTAQR, 2+ ($m/z = 811.2102$) containing eight acetylated lysine residues. Modifying the eight lysine residues with either light (D_0) or heavy (D_3) acetyl results in a total of nine isotopologues ($0 \text{ D}_3 - 8 \text{ D}_3$). All isotopologues except the isotopologues 0 D_3 and 8 D_3 consist of a multiple isotopomers (e.g. 56 for 3 D_3 and 5 D_3). The intensities of b, y and the internal ion fragments can be used for the determination of the acetylation degree. Only the acetylation degree of the terminal lysine residues K11, K12, K23, and K32 can be quantified, the acetylation degrees of K17, K18, K20, and K21 can not be determined.

Although Smith's method is the basis for the quantitation of site-specific acetylation degree, the method is accompanied with some limitation that makes the quantitation inaccurate. First, the use of deuterium causes a shift of few seconds in the LC retention time between protiated and deuterated peptides, thereby impairs quantification on the fragment ion level. Second, the use of trypsin with chemically acetylated histones generates

relatively long peptides with a high number of lysine residues. The method allows the acetylation quantitation of the terminal lysine residues but not the ones in the middle. A good example is the tryptic peptide of *T. brucei* H3(10-37), which has eight lysine residues (Fig. 3.5), only the acetylation of K11, K12, K23, and K32 can be determined, and the information about the acetylation degree of the other 4 residues cannot be obtained. Third, D₆-acetic anhydride induces a 3 Da mass shift per acetyl group, for multi-acetylated peptides, this causes a lag in the transmission of the isotopic ion clusters of endogenously acetylated peptides and their chemically acetylated forms to the collision cell for simultaneous fragmentation.

Middle-down and top-down approaches have emerged recently as attractive alternatives to quantify the degree of acetylation at the histones tails, albeit they are encountering certain limitations and present significant challenges. The poor fragmentation efficiencies of the high molecular weight ions and the distribution of large precursor ions across many charge states reduce the overall sensitivity and thus, hamper the identification and quantification of the site-specific acetylation degree.

To overcome these limitations, we developed the “fragment ion patchwork quantification”; a new MS-based approach that allows the determination of the site-specific acetylation degrees from fragment ion spectra with high accuracy, even within clusters of lysine residues. In this method, we used a mixture of ¹³C₁- acetic anhydride and ¹³C- sodium acetate as derivatization reagent on the protein level, followed by proteolysis by low-specificity proteases and quantification on the fragment ion level [91]. Acetic anhydride is known to modify α -amino groups, tyrosine, histidine, cysteine in addition to the desired ϵ -amino group. Performing the modification reaction in conditions of saturated sodium acetate (pH 7) instead of acetic acid results in an increased specificity. Modifying the protein with ¹³C₂-acetic anhydride acetylates all free amino groups and monomethylated amino groups (Fig. 3.6), but will not react with the dimethylated, trimethylated, or acylated amino groups (e.g., acetylated, formylated, or ubiquitinated amino groups). As a consequence, all the lysine residues in the protein become acetylated either endogenously or chemically.

The use of ¹³C₂ acetic anhydride causes the labeled and unlabeled isoforms to coelute from the LC C₁₈ column, eliminating the shift in the LC retention time, and thus, makes the quantification straightforward. Chemical acetylation with ¹³C₁- acetic anhydride results in smaller mass shift, 1 Da per acetyl group, compared to the 3 Da induced by the D₃-derivatization. The smaller mass shift makes it easier to transmit the isotopic ion clusters of endogenously acetylated peptides and their chemically acetylated forms at the same efficiency to the collision cell for simultaneous fragmentation. Unlike the common used sequence-specific protease trypsin, low-specificity proteases as elastase, thermolysin, and papain eliminates the influence of the presence of acetylation on the cleavage site. Using

3. Site-specific Acetylation Degree Quantitation

these different low-specificity proteases generates a high number of overlapping peptides leading to a high number of fragment ions for quantification. Compared to sequence specific proteases (e.g. trypsin, Arg-C or Glu-C), these low-specificity proteases yield better sequence coverage and allow a more accurate quantification of histones acetylation.

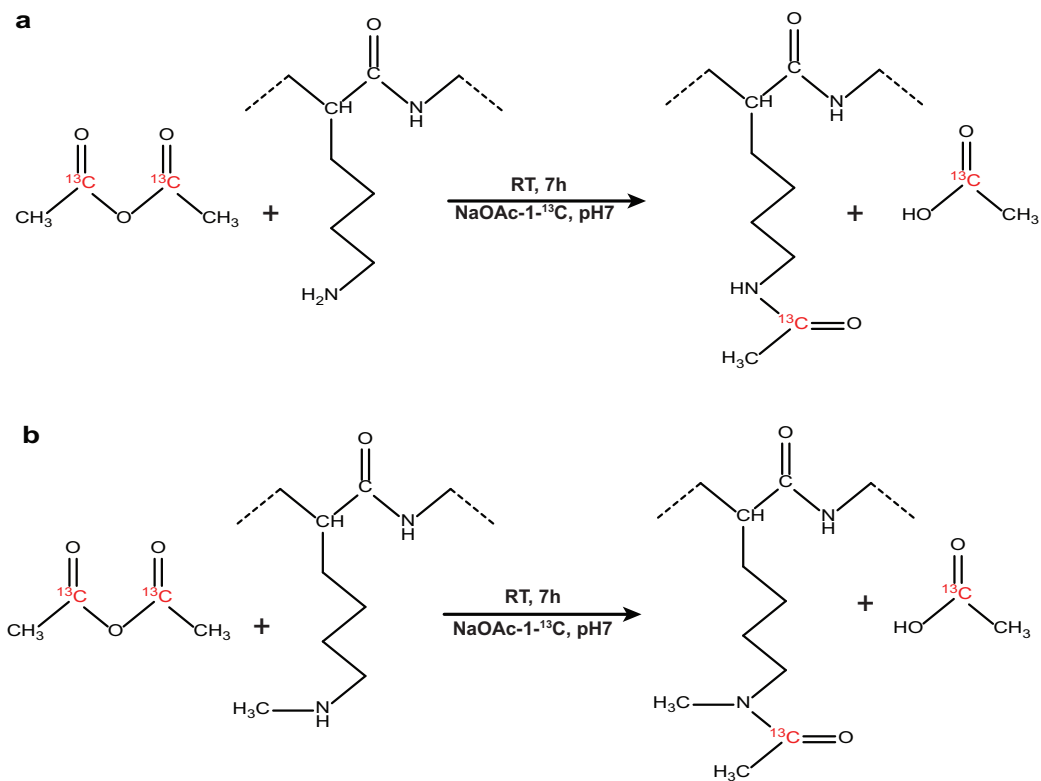


Figure 3.6.: Reaction scheme of acetic anhydride- $1,1\text{'-}^{13}\text{C}_2$ with a) unmodified lysine residue and b) monomethylated lysine residue. Incubating the samples under the given conditions guarantees nearly a complete modification (~98%) at all sites regardless the number of lysine residues within a protein.

4. Material and Methods

4.1. Reagents

All chemicals were purchased from Sigma-Aldrich Chemical Co. unless stated otherwise.

NuPAGE[®] LDS Sample Buffer (4X), NuPAGE[®] Sample Reducing Agent containing 500mM DTT (Dithiotreitol), NuPAGE[®] MOPS SDS Running Buffer, SimplyBlue[™] SafeStain, iodoacetamide (BioUltra), LC-MS grade ammonium bicarbonate (ABC) (Eluent additive for LC-MS), LC-MS grade acetonitrile ACN (LC-MS CHROMASOLV[®]), formic acid FA (for mass spectrometry ~98%), trifluoroacetic acid TFA (CHROMASOLV[®], for HPLC, ≥99.0%), sodium acetate-1-¹³C (99 atom % ¹³C), sodium acetate (ACS reagent, ≥99.0%), acetic anhydride, formic acid (for MS), acetone (for HPLC, ≥99.9%), L-Cysteine (TLC, ≥99.7%), acetic anhydride-1,1'-¹³C₂ (99 atom % ¹³C) was purchased from Cambridge Isotope Laboratories, and α-Cyano-4-hydroxycinnamic acid)

4.2. Proteases

For protein digestion, the low specificity proteases elastase from Promega (#V1891), thermolysin from R&D Systems (#3097-ZN-020) and papain (from papaya latex; #P4762) from Sigma-Aldrich were used. For peptide digestion, Glu-C, Sequencing Grade from Promega (#V1651) was used.

4.3. Protein and Peptides

The standard peptides used were [Glu1]-Fibrinopeptide B (Glu-Fib, EGVNDNEEGFF-SAR) from Sigma-Aldrich (St. Louis, USA) and a customized synthetic peptide LGG-Lys(ac)-GELGGKGE from JPT Peptide Technologies. Recombinant Histone H3 expressed in *E.coli* was used as standard protein for acetylation experiments. Histone fractions from *wild-type* HAT3 knockout strains were derived from Lister 427 bloodstream. (Sequences in Appendix A).

4.4. Sample Preparation

4.4.1. Peptides

In the early tests, peptides were modified using ^{12}C acetic anhydride and ^{12}C sodium acetate to test the completion of the reaction and the generated byproducts.

1pmol/ μl aliquotes of GluFib and customized synthetic peptide LGG-Lys(ac)-GELGGKGE were desalted with C-18 STAGE tips to remove any salts that might interfere with the reaction. Each C-18 STAGE tip was prewashed with 1x 20 μl MeOH and centrifuged for 2 min. at 2000-4000 g. 1x 20 μl of the elution buffer (buffer (B)) containing 80% ACN 0.5% FA was added to each C-18 STAGE tip and centrifuged for 2 min. at 2000-4000 g. Each C-18 STAGE tip was then conditioned with 2x 25 μl of the washing buffer (buffer (A*)) containing 2% ACN 0.3% TFA and centrifuged for 2 min. at 2000-4000 g. Samples were acidified with 0.1% FA before they were loaded into the washed and conditioned C-18 STAGE tips and centrifuged for 10 min. at 2000 g. Tips were then washed 3 times with 25 μl buffer (A*) and centrifuged for 2 min. at 2000-4000 g to assure that all chemical and salts were eliminated from the sample. In the final step, the samples were collected in clean eppendorf tube by eluting them from the C-18 STAGE tips with 2x 20 μl buffer (B) using a syringe to push the sample through the C-18 material and subsequently dried down in SpeedVac. 60 μl of 1M NaAcO was added to the control samples, for test samples 10 μl of acetic anhydride was added first followed by 50 μl of 1M NaAcO. All samples were incubated in a thermomixer at 25°C for 2 hours. Incubated samples were desalted using C-18 STAGE tips, dried down in SpeedVac and redissolved in 10 μl of 1% FA prior to the MS analysis.

Dried synthetic peptide LGG-Lys(ac)-GELGGKGE aliquotes were then modified by suspending them in a mixture of 10 μl acetic anhydride-1,1'- $^{13}\text{C}_2$ and 50 μl 1M sodium acetate-1- ^{13}C (pH 6.8) at room temperature to introduce the heavy stable isotope to the unmodified lysine residue following the same protocol.

4.4.2. Proteins

4.4.2.1. Recombinant Histone H3

Recombinant histone H3 from *T.brucei* (TriTryp database [92] ID Tb927.1.2510) was obtained from AG Janzen, Biocenter University of Würzburg. Recombinant histone H3 from textit*T.brucei* was expressed in *E.coli* BL21 and was purified from inclusion bodies by cation exchange chromatography according to a protocol published by Luger *et al.* [93]. The histone was dialyzed three times against distilled water containing 2 mM 2-

mercaptoethanol for 1h at 4°C, followed by one dialysis step overnight. The intact protein was analyzed by static nanoESI.

4.4.2.2. Preparation of Histone-enriched Fractions

***T. brucei* Cell Lines and Culture Conditions:** Both, wild type and HAT3 knockout strains were derived from Lister 427 bloodstream-form MITat 1.2 (clone 221) and were cultured at 37°C in HMI-11 up to a cell density of 1.5×10^6 cells/ml. Generation and characterization of HAT3 knockout strain has been described previously [74].

For each replicate, mononucleosomes were isolated from 3×10^8 bloodstream-form *T. brucei* cells using protocols similar to those published previously [94]. Histones were extracted from the mononucleosome-containing supernatant applying a standard acid extraction protocol [95] and were concentrated using StrataClean™ resin (Agilent Technologies, #400714) as described earlier [96]. Samples were obtained from AG Siegel, Institute of Infectious Diseases, University of Würzburg.

4.5. Gel Electrophoresis

NuPAGE® Novex® Pre-Cast 4-12% Bis-Tris gels 1.0 mm x 10 Well from Invitrogen were used. The sample preparation was carried out according to the workflow proposed by Invitrogen. The histone samples were resuspended in LDS sample buffer (10% glycerol, 141 mM Tris Base, 106 mM Tris HCl, 2% LDS, 0.51 mM EDTA, 0.22 mM SERVA® Blue G250, 0.175 mM Phenol Red, pH 8.5). Subsequently, reduced with 50 mM DTT at 70 °C for 10 minutes. Samples were cooled at room temperature before they were alkylated with 120 mM iodoacetamide and incubated in the dark for 20 minutes at room temperature. A sample volume of ~23 μ L was added per Well. NuPAGE® MOPS SDS running buffer (50mM MOPS, 50mM Tris, 1 mM EDTA, 0.1% SDS, pH 7.7) was used to separate the proteins. For recombinant histone H3 proteins were separated by SDS-PAGE for 40 min. at 200V while for acid extracted histones proteins were separated by SDS-PAGE at 200V for 25 min so that all histones are contained in one band at the level of 20 kDa marker.

Gels were washed three times for 5 minutes with ultrapure water before they are stained for 45 minutes with Simply Blue™ Safe Stain (Life Technologies). The gels then were washed with ultrapure water for approximately 2 hours.

4.6. In-gel Chemical Acetylation and In-gel Digestion

Gel regions containing histones were excised from the gel. Gel slices were chopped and destained with 70 % acetonitrile in 100 mM NH_4HCO_3 (pH 8), shrunk with 100 % ace-

4. Material and Methods

tonitrile and dried in a vacuum concentrator (Concentrator 5301, Eppendorf) separately.

Recombinant histone H3 dried gel bands were chemically modified in-gel with mixtures of 10 μ l of 90:10, 50:50 and 10:90 ^{12}C and ^{13}C acetic anhydride in corresponding mixtures of 50 μ l of 1M ^{12}C and ^{13}C sodium acetate buffer (pH 6.8) to mimic 10 %, 50 %, and 90 % acetylation degrees, respectively. For the acid extracted histones, dried gel pieces were suspended in a mixture of 10 μ l acetic anhydride-1,1'- $^{13}\text{C}_2$ and 50 μ l 1M sodium acetate-1- ^{13}C (pH 6.8). Gel pieces were incubated for 3.5 hours at 25 °C on a shaker. Solutions were replaced after 3.5 hours with fresh reaction mixtures, and the incubations were continued for further 3.5 hours (Fig. 4.1).

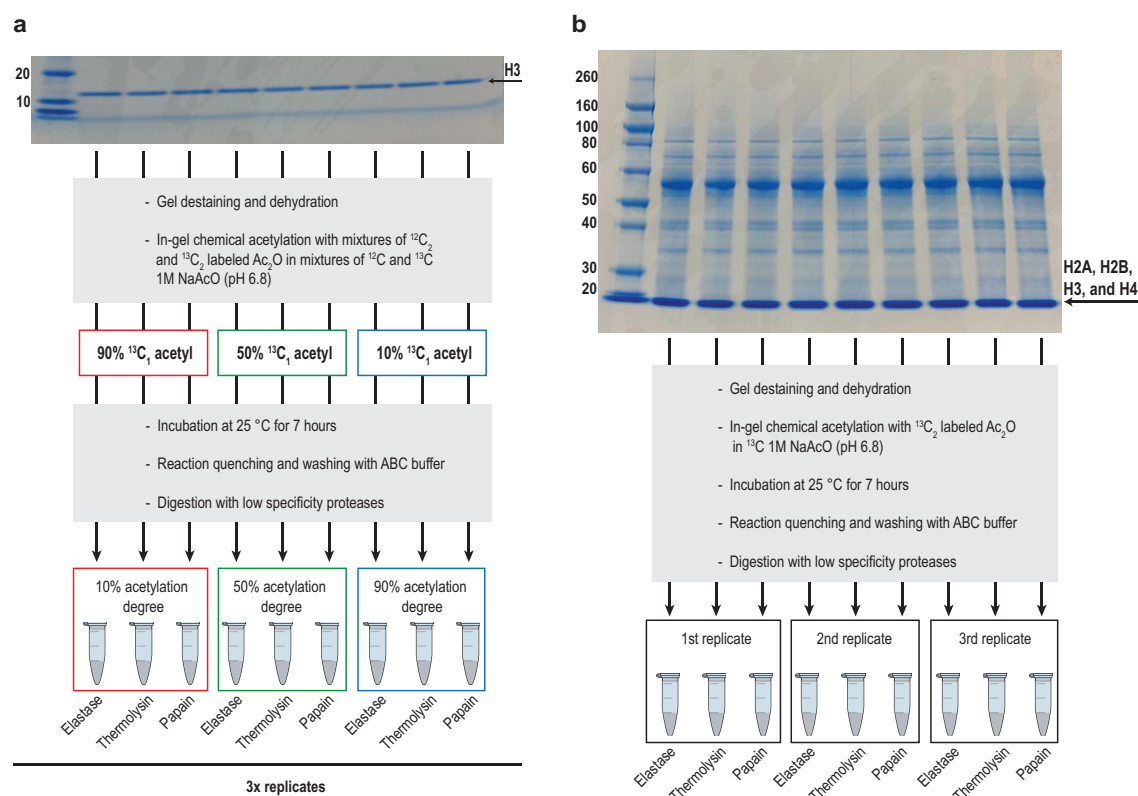


Figure 4.1.: a) Workflow for the determination of site-specific acetylation of recombinant histone H3 from *T. brucei*. b) Workflow for the determination of site-specific acetylation of acid extracted histones. For both workflows, treated gel bands were incubated at 25°C for 7 hours in total. Reaction was quenched with ABC buffer and gel bands were washed with 500 mM, 250 mM and 100 mM ABC buffer. ACN was used to shrink the gels before drying them down in SpeedVac. Proteins were proteolytically digested in gel by low specificity proteases elastase, thermolysin and papain, respectively[97].

The reactions were quenched by adding 200 μ l of 500 mM ABC buffer to each sample and shook vigorously for 15 min.. After removing the supernatant, 200 μ l of 100% ACN was added, samples shook vigorously for 5 min, supernatants were removed and the gel pieces were dried in SpeedVac. The samples were washed with 250 mM and 100 mM ABC buffer, respectively, in the same manner. Gel pieces were dried in a vacuum concentrator prior

in-gel digestion.

Dried gel pieces were suspended in 100 mM NH_4HCO_3 (pH 8) containing 0.1 μg of a protease. Digests with elastase, thermolysin and papain were performed overnight at 37 °C. 5 mM L-Cysteine was added to papain digests in order to activate the protease. Peptides were extracted from the gel slices with 5 % formic acid and transferred to LC vials.

4.7. NanoLC-MS/MS Analysis

For the peptide, unmodified recombinant histone H3 and $^{12}\text{C}_1$ -acetylated histone H3 nanoLC-MS/MS analyses were performed on LTQ Orbitrap Velos ProTM equipped with NSI source and coupled to EASY-nLC 1000 (Thermo Scientific). Peptides were loaded on a trapping column (2 cm x 75 μm ID, PepMap C-18 3 μm particles, 100 Å pore size) and separated on an EASY-SprayTM column (25 cm x 75 μm ID, PepMap RSLC C18 2 μm , 100 Å pore size) with 30 minute linear gradient from 3 % to 30 % acetonitrile and 0.1 % formic acid. The global MS analyses were performed in a data-dependent acquisition mode. The top 5 most intense ions were selected for higher-energy collision-induced dissociation (HCD) and collision-induced dissociation (CID), respectively. In both cases, normalized collision energy was 30%. Dynamic exclusion of selected ions for MS² fragmentation was 30 s and a single lock mass at m/z 371.101241 for the polysiloxane was used for internal mass calibration. Minimum signal threshold for precursor selection was set to 50 000, and the isolation width of the precursor selection window was set to 2.5. Overfilling of the C-trap was prevented by automatic gain control (AGC) and set to 10⁶ ions for a full FTMS scan and 5 × 10⁵ ions for MSⁿ mode. Intact peptides were detected in the Orbitrap mass analyzer at a resolution of 60 000.

NanoLC-MS/MS analyses for $^{13}\text{C}_1$ - acetylated recombinant histone H3 and acid extracted histones proteins were performed on an Orbitrap FusionTM (Thermo Scientific) equipped with an EASY-SprayTM Ion Source and coupled to an EASY-nLC 1000 (Thermo Scientific). Peptides were loaded on a trapping column (2 cm x 75 μm ID, PepMap C-18 3 μm particles, 100 Å pore size) and separated on an EASY-SprayTM column (25 cm x 75 μm ID, PepMap C-18 2 μm particles, 100 Å pore size) with either a 45 minute (for recombinant histone H3) or a 120 minute (for histone fractions obtained by acid extraction) linear gradient from 3 % to 30 % acetonitrile and 0.1 % formic acid. Both MS and MS/MS scans were acquired in the Orbitrap analyzer with a resolution of 60 000. HCD fragmentation with 35 % normalized collision energy was applied. A Top Speed data-dependent MS/MS method with a fixed cycle time of 3 seconds was used. Dynamic exclusion was applied with a repeat count of 3 and an exclusion duration of 45 seconds. Minimum signal threshold for precursor selection was set to 50 000, and the width of the precursor selection window was

4. Material and Methods

set to 6 Da. Predictive AGC was used with AGC a target value of 2×10^5 for MS scans and 5×10^4 MS/MS scans. EASY-IC was used for internal calibration.

4.8. Data Analysis

4.8.1. MS Raw Data Processing

Each raw data file was processed with Mascot Distiller 2.4 (Matrix Science) using two different types of parameter settings. One parameter set (ID) was used for generating peak lists for database searching with Mascot, the other parameter set (Quant) was used for generating peak lists for quantitative analysis.

4.8.2. Database Searching

4.8.2.1. PEAKS Database Search

The MS data of the unmodified intact histone H3 and $^{12}\text{C}_1$ -acetylated H3 were analyzed with PEAKS 8 and searched against a custom database containing the proteins of interest. For data refinement, no merge option was selected, precursor options corrected and no correction of charge options and no filter was selected. Carboxyamidomethyl (C) was set as fixed modification. Acetylation (Protein N-term), acetylation (K), methylation (K), dimethylation (K), trimethylation (K) and oxidation (O) were set as variable modification. Database searches were performed without protease specificity (enzyme: none), 8 ppm mass tolerance for the precursor, 0.02 Da mass tolerance for fragment ions, 8 max variable PTM per peptide and FDR estimation enabled. De novo ALC score was set to $\geq 50\%$ and peptides with score < 15 were filtered out.

4.8.2.2. Mascot Database Search

ID peak lists were searched with Mascot Server 2.4 against a custom database compiled from TriTrypDB [74]. In addition to all histone sequences, the custom database contained the top 300 of proteins identified in our histone samples. Since acetylation with $^{13}\text{C}_1$ -acetyl is a proxy for in-vivo unmodified lysine residues, we replaced all K for O in all protein sequences of the database fasta file. O was defined as a new amino acid with the chemical composition $^{13}\text{CC}_7\text{H}_{14}\text{N}_2\text{O}_2$ (monoisotopic mass 171.108883 Da). In order to take N-terminal acetylation into account we defined J as a new amino acid with the chemical composition of $^{13}\text{C}_1$ -acetyl ($^{13}\text{CCH}_2\text{O}$, monoisotopic mass 43.01392 Da). J was added to the N-terminus of all protein sequences of the database fasta file. The fasta file contained each protein sequence in two versions, one starting with J and the other starting

with JM. All variable protein N-terminal and lysine modifications were defined relative to $^{13}\text{C}_1$ -acetyl.

In addition to carbamidomethyl at C ($\text{C}_2\text{H}_3\text{NO}$) as fixed modification, database searching was performed with the following variable modifications: oxidation at M (O), light acetyl at O ($^{13}\text{C}_1\text{C}$), light acetyl at protein N-terminus ($^{13}\text{C}_1\text{C}$), mono-methyl at O (CH_2), mono-methyl at protein N-terminus (CH_2), di-methyl at O ($^{13}\text{C}_1\text{CH}_2\text{O}_1$), di-methyl at protein N-terminus ($^{13}\text{C}_1\text{CH}_2\text{O}_1$), tri-methyl at O ($^{13}\text{C}_1\text{C}_2\text{H}_4\text{O}_1$), and tri-methyl at protein N-terminus ($^{13}\text{C}_1\text{C}_2\text{H}_4\text{O}_1$). These definitions take into account that mono-methylated amino groups react with acetic anhydride, whereas di- and tri-methylated amino groups do not react. Database searches were performed without protease specificity (enzyme: none), 7 ppm mass tolerance for the precursor, and 0.02 Da mass tolerance for fragment ions. Mascot search results were exported as XML file. In addition to standard settings, start and end positions of the peptides were included in the export file, peptides with score <15 were removed.

4.8.3. Fragment Ion Patchwork Quantification

For MS^2 -isotope pattern based quantification, scripts programmed in R version 3.1.2 using parallel processing when possible (doParallel) were executed on a Microsoft Windows Server 2008 R2 operated computer (20 cores, 128 GB RAM). XML files containing Mascot identifications and peak list files with complete isotope patterns of fragment ions (exported as mzData-files from Mascot Distiller after processing with the Quant options, see above), were parsed and combined. For each peptide, fragment ion masses were calculated with OrgMassSpecR, and fragment ions with one, two or three acetylation sites (incl. N-terminal) were kept for further analysis. Data belonging to the same experiment, e.g. digests with different enzymes or technical replicates, were combined.

The unambiguous identification of the correct positioning of (multiple) modifications is not always possible. We considered the site information from the Mascot search as reliable, if the delta score [98] (i.e. the difference of lower ranking peptide isoform scores to the rank 1 score) is higher than 10. From these unambiguously identified peptides, a list of reliable modification sites was generated. These were compared to the set of modifications of each isoform of ambiguous peptides (i.e. with delta scores <10). If the number of ambiguous peptide isoforms could be reduced to one by considering only reliable modifications, all other ambiguous isoforms were eliminated and the peptide was considered as non-ambiguous. In case that none or more than one isoform could be explained by reliable modifications alone, all of these were used for further analysis, but marked in the resulting table as ambiguous.

In the first quantitation round, theoretical isotope patterns for fragment ions with a single

4. Material and Methods

acetylation site were calculated with Rdisop, one heavy ($^{13}\text{C}_1$) acetylated and one light (natural) acetylated isotope pattern for each fragment. The isotope purity of the heavy acetic anhydride (99.5 % according to certificate of analysis) was considered for the calculation of heavy isotope patterns. For each fragment ion, calculated theoretical isotope patterns were combined in heavy to light proportions from 0 to 100 % in 0.01 % steps. Resulting mixed isotope patterns (i.e. the monoisotopic, +1 and +2 isotope peak) were compared with the corresponding experimental isotope pattern. For each mixing ratio the deviation from the experimental pattern was calculated, the mixing ratio with the minimum deviation from the experimental pattern equals the acetylation degree of this site. To remove calculated isotope patterns, which are inferred by other fragment ions or are only partial due to a low intensity, the following quality criteria were applied using additionally the -1 isotope peak of the isotope patterns: Pearson correlation r between experimental and theoretical isotope pattern > 0.99 , maximum difference in the relative abundance of a single isotope peak < 20 .

The acetylation degree of particular sites could be influenced by other modifications in the same peptide. Therefore, we calculated the acetylation degree for each fragment ion covering a particular site and used the peptide context information about other modifications (e.g. methylations) for grouping of acetylation degrees during outlier removal and averaging.

Outliers (outside 1.5x interquartile range, IQR) from the distribution of acetylation degrees were removed. Subsequently, the median acetylation degrees for each site (and modification context) were used for calculating the acetylation degree of fragment ions containing two acetylation sites. Here, heavy and light isotope patterns were calculated with one of the acetylation sites treated as acetylated with the median acetylation degree from the first quantification round. For comparing the mixed theoretical isotope patterns with the experimental patterns, all peaks from the monoisotopic to the +4 isotope peak were used to identify the best matching mixing ratio and the acetylation degree of this particular site. The same quality criteria as from the first round were applied, but using the -1 to the +4 peaks. Outliers were removed as above from the combined acetylation degrees from the first and second round. It should be noted, that for each fragment ion containing two acetylation sites, an acetylation degree of both sites is calculated when possible, i.e. if a median value from the first round was available for both sites. In addition to obtaining values for closely acetylation sites at all, this strategy generally increases the number of quantitation events per site.

The analysis of fragment ions with three acetylation sites was performed correspondingly, except that theoretical isotope patterns were calculated using for two of the three acetylation sites the median acetylation degrees from the second round. The difference between the experimentally observed and the theoretical acetylation degree was less than 3% in all

cases.

The acetylation degree of a site that is additionally modified by other type of modification cannot be determined with high accuracy since the corresponding peptides are chemically distinct and have different physicochemical properties. In such a case, we use spectral counting to estimate the ratio between acetylation and other types of modifications.

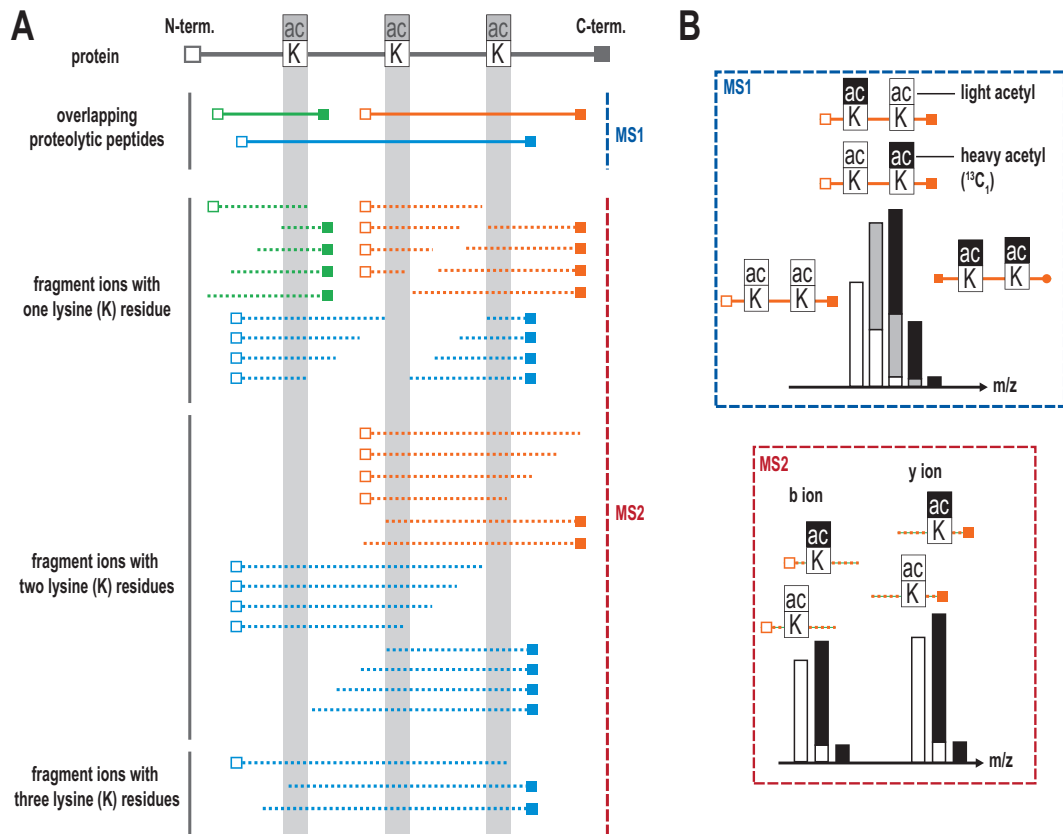


Figure 4.2.: Fragment ion patchwork quantification approach for measuring site-specific acetylation degrees. (A) All unmodified amino groups were chemically $^{13}\text{C}_1$ -acetylated on the protein level. Proteins were digested with a set of different low-specificity proteases in order to obtain high sequence coverage. All identified fragment ions with up to three acetyl groups were used to calculate the site-specific acetylation degrees. (B) MS1: The isotope patterns of acetylated peptides are mixtures from isotope patterns of different isotopologues. Full isotope patterns were selected with a widened precursor selection window for HCD fragmentation. MS2: Acetylation degrees were calculated by determining the fraction of $^{13}\text{C}_1$ -acetyl in acetylated b and y ions from their isotope patterns [97].

5. Results and Discussion

A general experimental workflow for the fragment ion patchwork quantification approach is shown in Fig. 5.1. First, a protein is treated with $^{13}\text{C}_2$ -acetic anhydride, to convert all free amino groups to the acetylated form. Endogenously acetylated lysine residues carry $^{12}\text{C}_1$ -acetyl group, while the chemically acetylated lysine residues carry an isotopically labeled $^{13}\text{C}_1$ -acetyl group. This eliminates the influence of acetylated sites on proteolytic cleavage and on the ionization efficiency of the generated peptides. Once the complete chemical acetylation is achieved, the protein is digested by a low-specificity protease and analyzed by LC-MS/MS. The endogenously acetylated peptide and the same peptide that was chemically acetylated have the same chemical properties but are distinguishable by mass spectrometry.

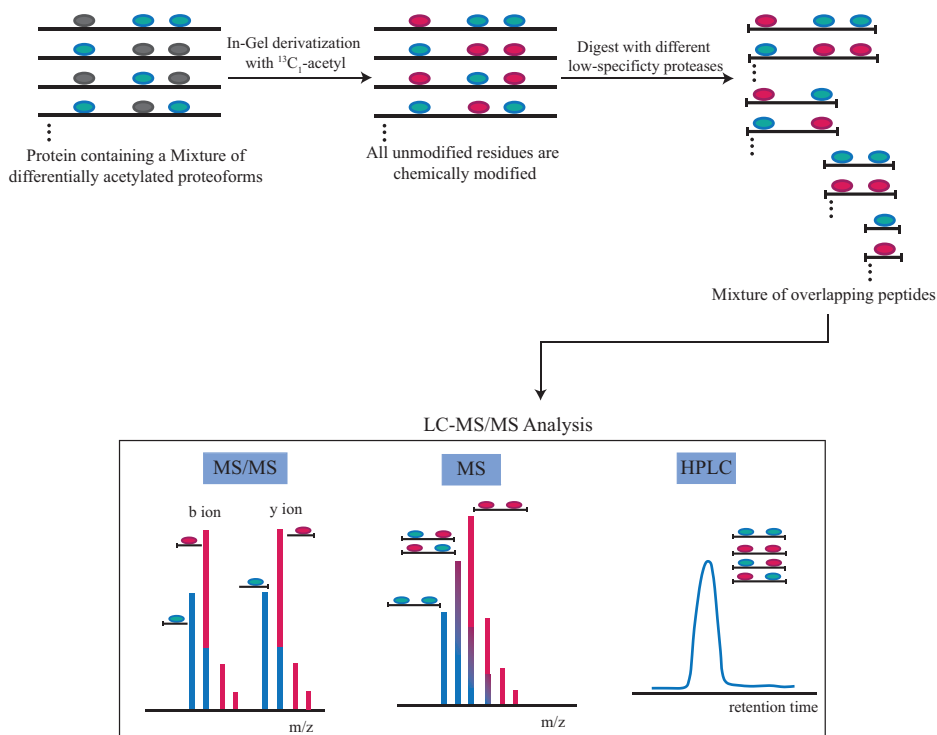


Figure 5.1.: General workflow for the fragment ion patchwork fragment ion patchwork quantification approach for measuring site-specific acetylation degrees. A protein, which is partially endogenously acetylated, is chemically modified with $^{13}\text{C}_2$ -acetic anhydride and digested with a set of different low-specificity proteases. The site-specific acetylation degrees was calculated using the intensities of b and y.

5.1. Method Optimization

In the early stages of method development, we used the synthetic peptide LGG-Lys(ac)-GELGGKGE, which has one acetylated lysine residue and an unmodified lysine residue, to test the reaction completion, the generation of byproducts and the elution pattern of the pre-acetylated and the chemically acetylated peptides (Fig. 5.2). We tested different mixtures of reagents to achieve complete acetylation with minimal generation of byproducts, and to overcome the limitation of deuterium (appendix B). The best one we found that meets all the requirements was a mixture of $^{13}\text{C}_2$ -acetic anhydride and 1M $^{13}\text{C}_2$ -sodium acetate (pH 6.8) [99]. Acetic anhydride is a known and widely used acetylation reagent. It acetylates ϵ -amino groups, α -amino groups, tyrosine, histidine, cysteine. The use of acetic anhydride in combination of saturated sodium acetate instead of acetic acid increases its specificity. In contrast to D_3 -acetyl derivatization, $^{13}\text{C}_1$ -acetyl causes no retention time shift between isotopologue peptides in C_{18} reversed-phase chromatography that allows the endogenously acetylated and the chemically acetylated peptides to coelute, which is the prerequisite for an accurate quantitation on the fragment ion level.

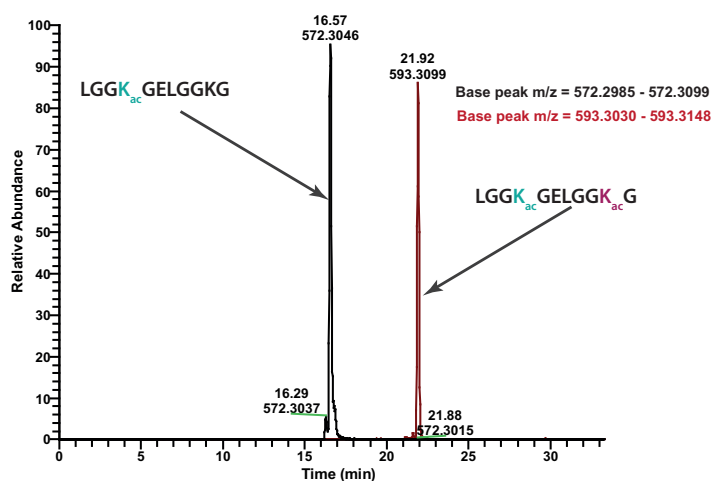


Figure 5.2.: Retention time of the base peaks of pre-acetylated peptide LGG-Lys(ac)-GELGGKGE, 2+ ($m/z = 572.3042$) *black* and the chemical modified isoform LGG-Lys(ac)-GELGG-Lys(ac)-GE, 2+ ($m/z = 593.3089$) *red*.

5.2. Method Validation

To validate the robustness of our new approach, we used recombinant histone H3 from *Trypanosoma brucei*, a protozoan parasite causing African trypanosomiasis, expressed in *E. coli*.

In the first step, we performed an LC-MS/MS analysis of the intact recombinant histone H3 to ensure that it is completely unmodified (Fig. 5.3). We then modified the recombi-

nant histone H3 in-gel prior to proteolysis using a mixture of $^{12}\text{C}_2$ -acetic anhydride and $^{12}\text{C}_2$ -sodium acetate to test the reaction completion efficiency, since a complete chemical acetylation reaction is essential for accurate quantification.

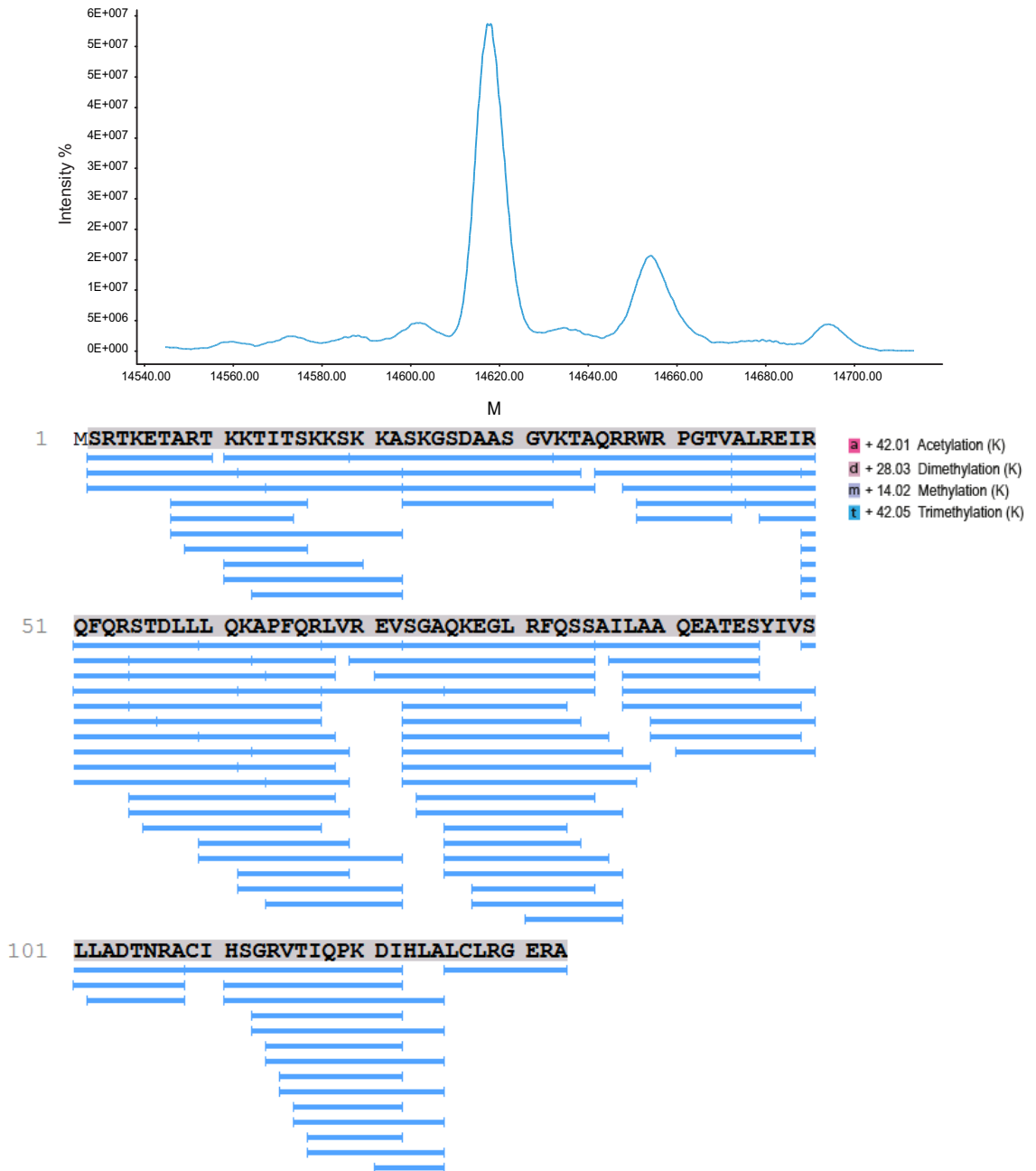


Figure 5.3.: Recombinant histone H3 *T. brucei* expressed in *E. coli*. *Upper panel:* NanoESI spectrum of the intact recombinant histone H3. The experimental mass of the largest peak (14618 Da) is in accordance with the theoretical mass of 14619 Da. *Lower panel:* Sequence coverage and the generated peptides of an elastase digest of unmodified recombinant histone H3. MS analysis of the intact protein confirmed that it is completely unmodified.

The digests were analyzed by data-dependent nanoLC-MS/MS on a Velos Orbitrap in-

5. Results and Discussion

strument using HCD fragmentation and detection of the fragment ions in the Orbitrap mass analyzer with high resolution and high mass accuracy.

PEAKS database search showed that a complete sequence coverage was achieved when digesting the modified protein with the low-specificity protease elastase, and shorter peptides were generated comparing to those from a tryptic digest.

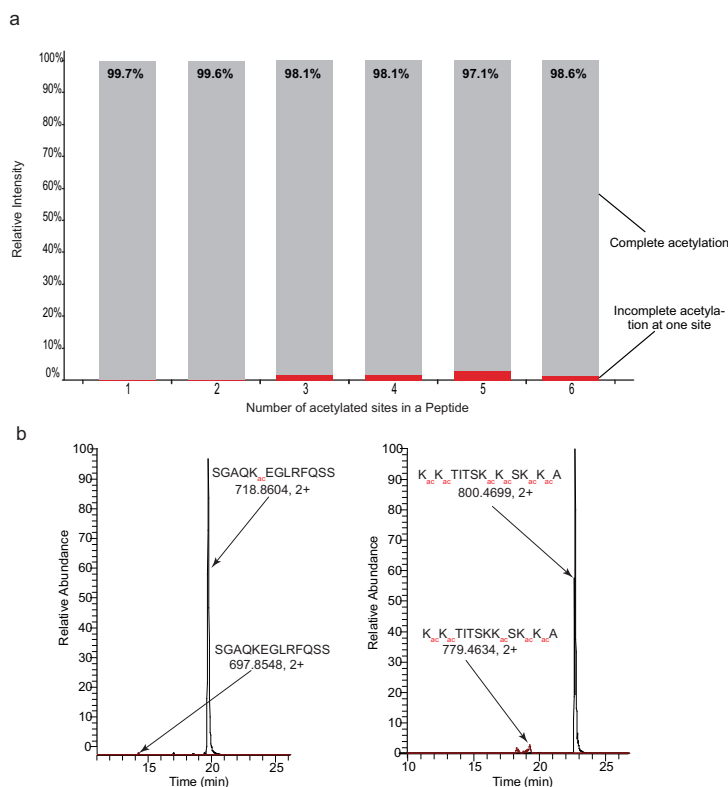


Figure 5.4.: Reaction completion efficiency of light acetylated Histone H3: **a)** Summed relative intensities of acetylated peptides with different numbers of acetylation sites. The precursor intensities of all identified peptides in an LC-MS run of the recombinant histone H3 chemically derivatization with $^{12}\text{C}_1$ -acetyl was summed up for fully acetylated peptides (gray bars) or incompletely acetylated peptides (red bars), binned by their number of acetylation sites (protein N-terminus plus Lysine residues). The percentage of the summed intensities for completely modified peptides is noted on top of the gray bars. **b)** Typical extracted ion chromatograms (XICs) of chemical acetyl-derivatized peptides. Shown are XICs of the peptides SGAQKEGLRFQSS, 2+ and KKTITSKSKSKKA, 2+ from recombinant histone H3, both fully acetylated (black) and incompletely acetylated on one site (red).

The completeness of the chemical acetylation reaction was verified by comparing the summed peptide intensities of completely and incompletely modified peptides, which was 99.7% for singly acetylated peptides and at least 97.1% for peptides with a higher number of acetylated sites, indicating that the applied conditions resulted in virtually complete chemical acetylation (Fig. 5.4). Only peptides with complete or acetylation missing on one site have been identified and for most of the fully acetylated peptides no incompletely labeled isoforms were identified. Precursor intensities were derived from processing the

raw-data file with MaxQuant, searched against a custom database with about 300 entries plus common contaminants and filtered for 1 % FDR on protein and peptide level. Searches were performed with no enzyme specificity, allowed peptide lengths were min. 6 and max. 25. In addition to the fixed modification carbamidomethylation (C), acetylation (protein N-term and K), Gln- to pyroGlu (Q) and oxidation (M) were allowed as variable modifications. For calculating the precursor intensity sums, intensity values from “ModificationSpecificPeptides.txt”-file were used.

5.3. Proof of Principle

To validate the accuracy and efficiency of our method we used the recombinant histone H3 to mimic an endogenously 10%, 50%, and 90% acetylation degree at each individual lysine residue and on all the protein N-terminus. This was achieved by modifying the protein in-gel using a mixture of $^{12}\text{C}_2$ and $^{13}\text{C}_2$ - acetic anhydride in $^{12}\text{C}_2$ and $^{13}\text{C}_2$ -sodium acetate, respectively, in 1:9, 1:1 and 9:1 ratios.

Nine aliquots of histone H3 were separated by SDS-PAGE, three samples each were chemically modified in-gel to mimic 10%, 50%, and 90% (Fig 4.1 a). One sample of each triplicate identical in the acetylation degree was digested in-gel with either elastase, thermolysin, or papain, which yielded complete sequence coverage and many overlapping peptides, resulting in a high number of fragment ions for quantification. The digests were analyzed by data-dependent nanoLC-MS/MS on an Orbitrap Fusion instrument using HCD fragmentation and detection of the fragment ions in the Orbitrap mass analyzer with high resolution and accurate mass. This experiment was repeated three times with three acetic anhydride mixtures and the three proteases elastase, thermolysin, and papain. Completeness of the chemical acetylation was 99.9% for singly acetylated peptides and at least 97.7% for peptides with higher number of acetylated sites (Fig. 5.5).

Mascot database search performed showed that a complete sequence coverage for all nine samples was successfully achieved. More importantly, all lysine residues, as well as the protein N-terminus, were covered by a large number of overlapping peptides (Fig. 5.6). The use of different low-specificity proteases generated a high number of overlapping peptides, and subsequently, a high number of b and y fragment ions, between 30 to 600, for each site. All fragment ions from all identified peptides generated from the multiprotease experiments were used for quantification. The high number of fragment ions generated was more than high enough to calculate the acetylation degrees of all lysine residues of H3 with high accuracy. In a cluster of lysine residues stretching from position 10 to 23, the acetylation degrees of the terminal sites K10, K11, and K23 was calculated from 200 to 600 fragment ions, while for those in the middle of the lysine cluster K16, K17, and K19 were calculated from 30 to 150 b and y fragment ions. Although the number of usable

5. Results and Discussion

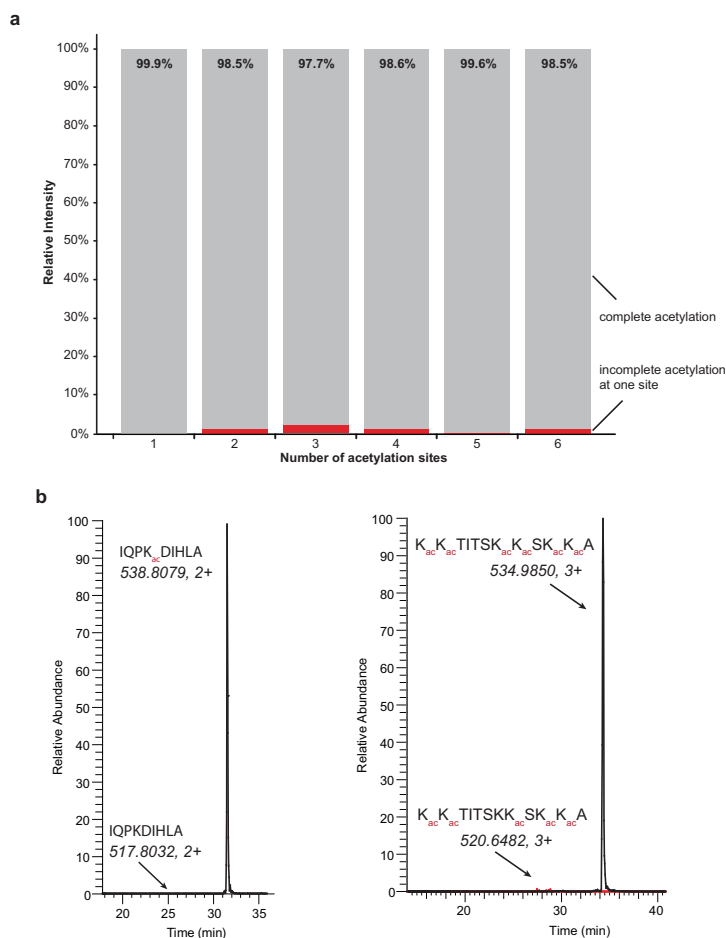


Figure 5.5.: Reaction completion efficiency of 10% acetylated Histone H3: **a)** Summed relative intensities of acetylated peptides with different numbers of acetylation sites. The precursor intensities of all identified peptides in an LC-MS run of the recombinant histone H3 chemically derivatization with $^{13}\text{C}_1$ -acetyl was summed up using the same criteria in method validation. **b)** XICs of the peptides QPKDIHLA, 2+ and KKTITSKKSCKKA, 2+ from recombinant histone H3, both fully acetylated (black) and incompletely acetylated on one site (red) [97].

fragment ions for sites in the middle of a cluster of lysine residues in a close proximity to each other significantly drops, the accurate calculation of the acetylation degrees for these individual sites remains achievable, which was not possible with Smith's method [83].

Site-specific acetylation degrees were calculated using an in-house algorithm developed by Dr. Jens T. Vanselow. The software calculates theoretical isotopic patterns of acetylated fragment ions with either one light or one heavy acetyl residue. These are mixed in different ratios and matched to the experimentally observed isotopic pattern. The fraction of the light acetyl isotopic pattern in the best fitting theoretical isotopic pattern directly corresponds to the acetylation degree. This procedure is performed in the first instance with all singly acetylated fragment ions only, and preliminary acetylation degrees were calculated for all sites covered with singly acetylated fragments. Acetylation sites within the middle of a lysine cluster are typically not covered by singly acetylated fragment ions.

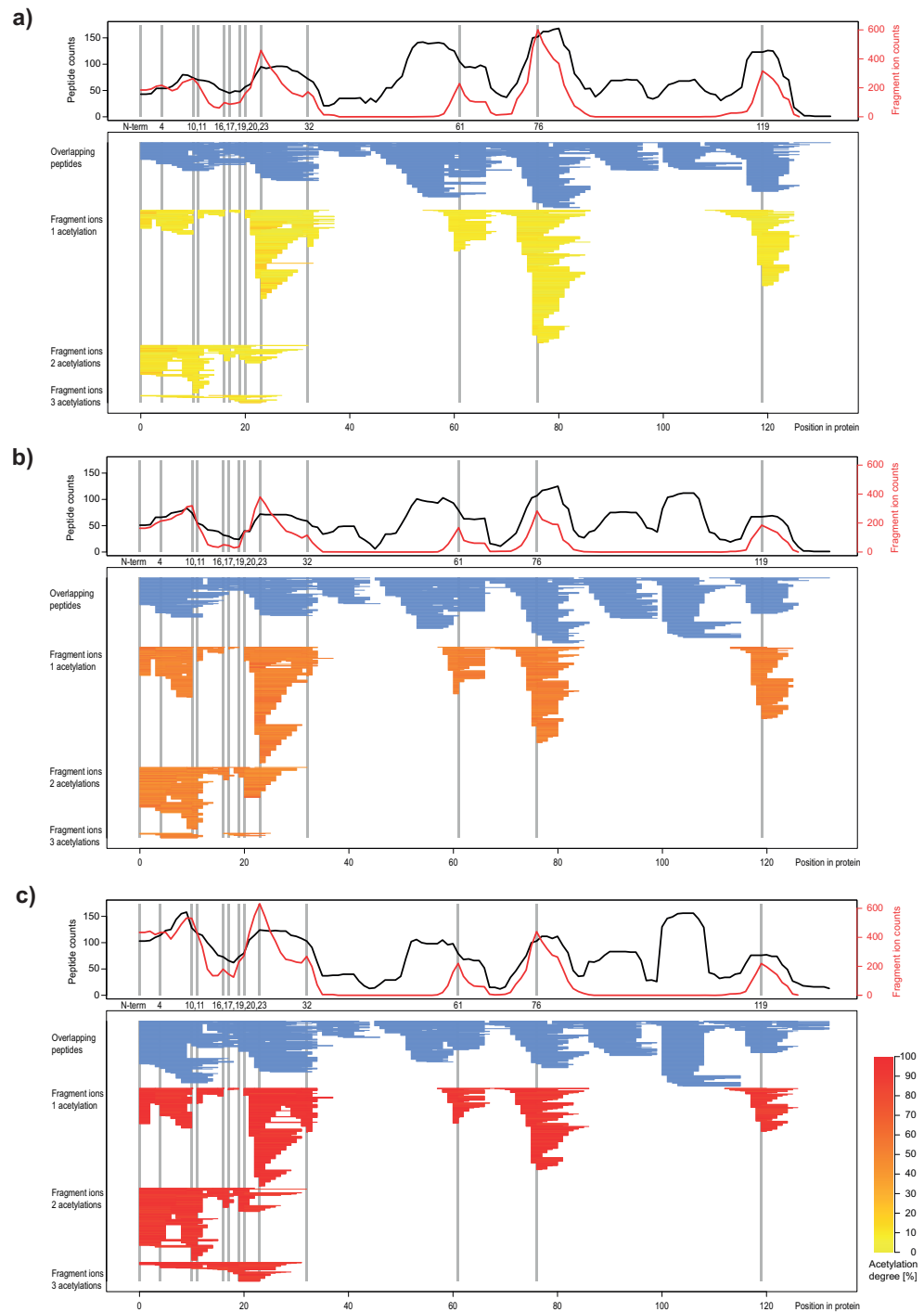


Figure 5.6.: Coverage map for acetylated recombinant histone H3 from *T. brucei*, with a) 10%, b) 50% and c) 90% acetylation degree on all sites. All identified peptides and fragment ions from digests with elastase, thermolysin, and papain are displayed. Upper panel: number of peptides (black) and fragment ions (red) used for calculation of the acetylation degree per amino acid position. Patchwork panel: overlapping peptides (blue lines) and fragment ions used for quantitation (10% yellow, 50% orange and 90% red lines). Acetylation sites are highlighted (gray vertical bars) and positions are indicated. The coverage maps of all other experiments with recombinant histone H3 are shown in (B).

5. Results and Discussion

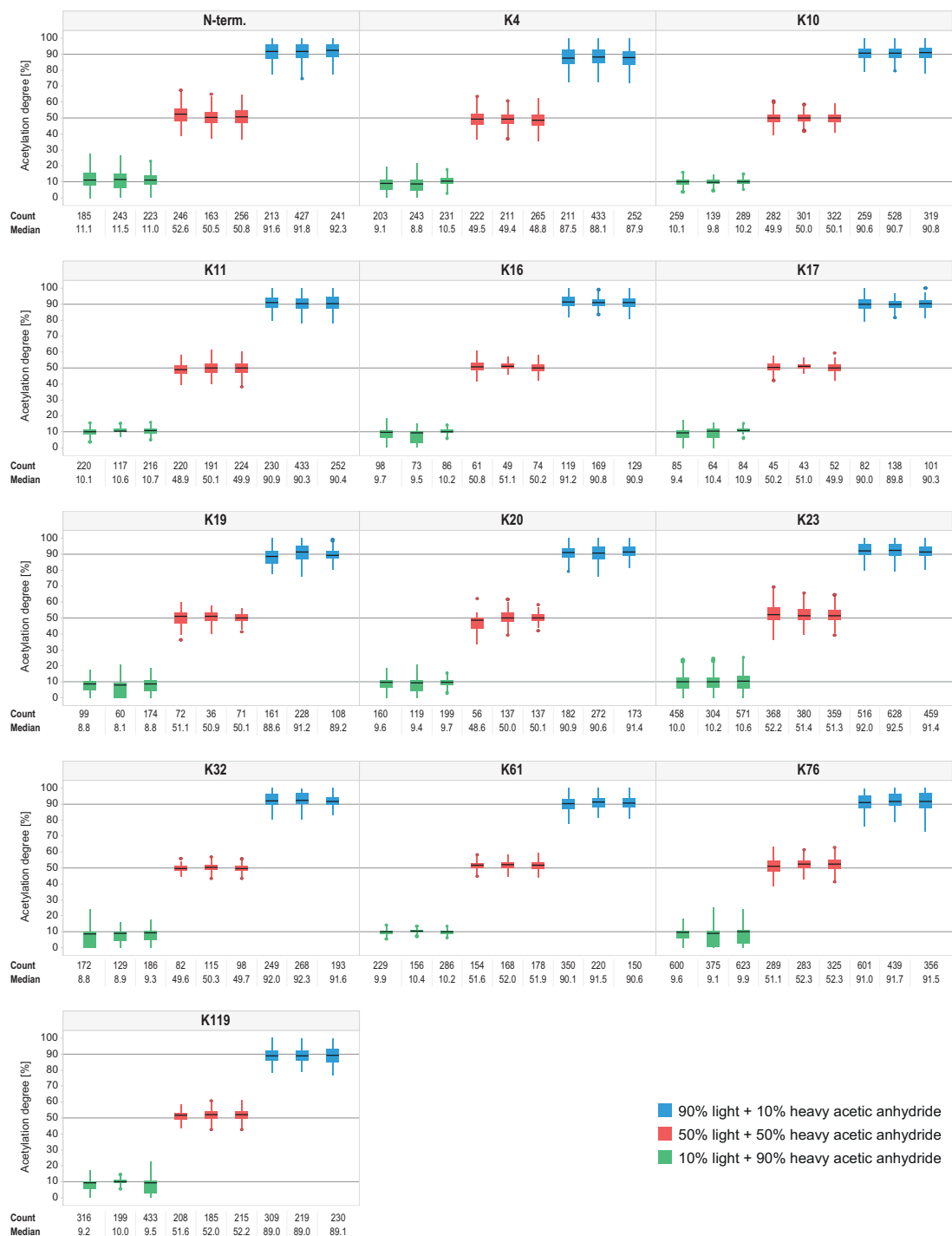


Figure 5.7.: Recombinant histone H3 from *T. brucei* was chemically acetylated with 10:90, 50:50 and 90:10 mixtures of light and heavy acetic anhydride. Boxplots show the distribution of determined acetylation degrees for all 13 acetylation sites of H3 (protein N-terminus and 12 lysine residues) from three replicates (box: 50 % of the data between 25 and 75 %; black line: median; whiskers: 1.5x interquartile range, IQR; outliers: outside 1.5x IQR). The tables below the boxplots contain the number of fragment ions used for quantitation and the median acetylation degree [97].

5.4. Identification of Substrate Sites of Histone Acetyltransferases 3 (HAT3) in *T. Brucei*

Therefore, in a second and third round, all isotopic patterns of all b and y ions with two and three acetyl residues, respectively, are analyzed in the same way as in the first round, but taking the preliminary acetylation degrees of all except one other site into account. Finally, all calculated acetylation degrees from all detectable b and y ions with up to three acetylated residues are used to calculate site-specific acetylation degrees for all sites. Experimentally determined acetylation degrees for all sites were in excellent agreement with the theoretical values of 10, 50, and 90% (Fig. 5.7), respectively, and the quantification strategy is highly reproducible. These results demonstrate the high accuracy of the fragment ion patchwork quantification approach.

5.4. Identification of Substrate Sites of Histone Acetyltransferases 3 (HAT3) in *T. Brucei*

Allfrey and colleagues were the first to discover histone acetylation and the involvement of particular enzymes (KATs) [100]. So far, twenty-two lysine acetyltransferases have been identified in human; their specific targets are not fully deciphered yet [70]. We choose *T. brucei*, which has only six HATs [71] as a model organism to identify the substrate sites of histone acetyltransferases 3 (HAT3).

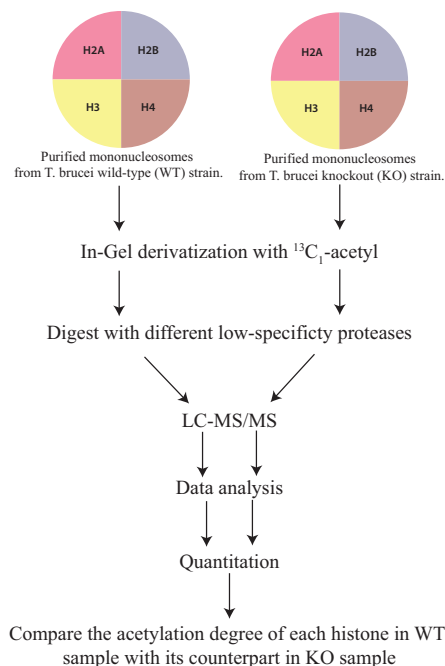


Figure 5.8.: Workflow for the identification of substrate site of histone acetyltransferases (HATs).

For this, we applied the fragment ion patchwork quantification to identify the target lysine(s) of histone acetyltransferase HAT3 in *T. brucei*. HAT3 is known to acetylate H4K4

5. Results and Discussion

[73], but no other targets for HAT3 has been identified. For unbiased target site identification, site-specific histone acetylation degrees for wild type and HAT3 knockout strains were measured in biological triplicates. Histones from the two different strains were enriched by acid extraction and separated by SDS-PAGE. Regions containing histones were excised from the corresponding gel lanes, and proteins were chemically $^{13}\text{C}_1$ -acetylated in-gel (Fig. 5.8). Proteins of each sample were digested in-gel with elastase, thermolysin, and papain separately, the digests were then analyzed by nanoLC-MS/MS. Peptides and proteins were identified using Mascot, and acetylation degrees were calculated from acetylated b and y ions of the identified peptides. We were able to measure site-specific acetylation degrees of almost all lysine residues for all core canonical histones H2A, H2B, H3, and H4 as well as variant histones H2A.Z, H2B.V, H3.V and H4.V, for the wild type as well as for the HAT3 knockout strain.

For histone H4 from wild type cells, we measured a high acetylation degree of 80% on K4, an acetylation degree of 6% for K10, and a low acetylation degree of 4% for K2. The acetylation degree for K14 through K93 was 0%. By comparing the acetylation of H4 lysines from wild type and knockout cells, the acetylation degree of H4K4 decreased from 80% in the wild type to 10% in the HAT3 knockout (Fig. 5.9), whereas the acetylation degrees of all other sites were not influenced by the knockout. These findings confirm the previous report that HAT3 is responsible for in vivo acetylation of H4K4 in *T. brucei* [74]. The N-terminal acetylation degree was quantified in only one (WT) and two (KO) experiments. Based on spectral counting, we found strong monomethylation on the N-terminus and substantial mono-, di-, and trimethylation on K17 and K18 of histone H4 (Fig. 5.9), these sites are not additionally acetylated (acetylation degree: 0%).

For histone H3 we measured a low acetylation degree (6%) on N-terminal, and (0%) acetylation at all other identified sites from wild type cells, the knockout of HAT3 has no influence on the acetylation of this site (Fig. 5.10). Based on spectral counting, we found a mixture of mono-, di-, and trimethylation on the N-terminus of histone H3, in addition to the low acetylation degree. Therefore, the degree of acetylation at N-terminus of H3 cannot be determined with high accuracy due to the presence of other modification at the site.

These findings are also in good agreement with the previous reported data suggesting that N-terminal of histone H3 appears to have a number of modifications [72]. We also detected a mixture of mono- and trimethylation on H3K4. Moreover, our results showed that H3K76 is trimethylated to a great extent, with a presence of di- and monomethylation. These results are consistent with the previously reported data [101]. It is believed that the trimethylation of H3K4 and H3K76 promote the replacement of H2B by H2BV at the nucleosome transcription start site of *T. brucei* [102].

5.4. Identification of Substrate Sites of Histone Acetyltransferases 3 (HAT3) in *T. Brucei*

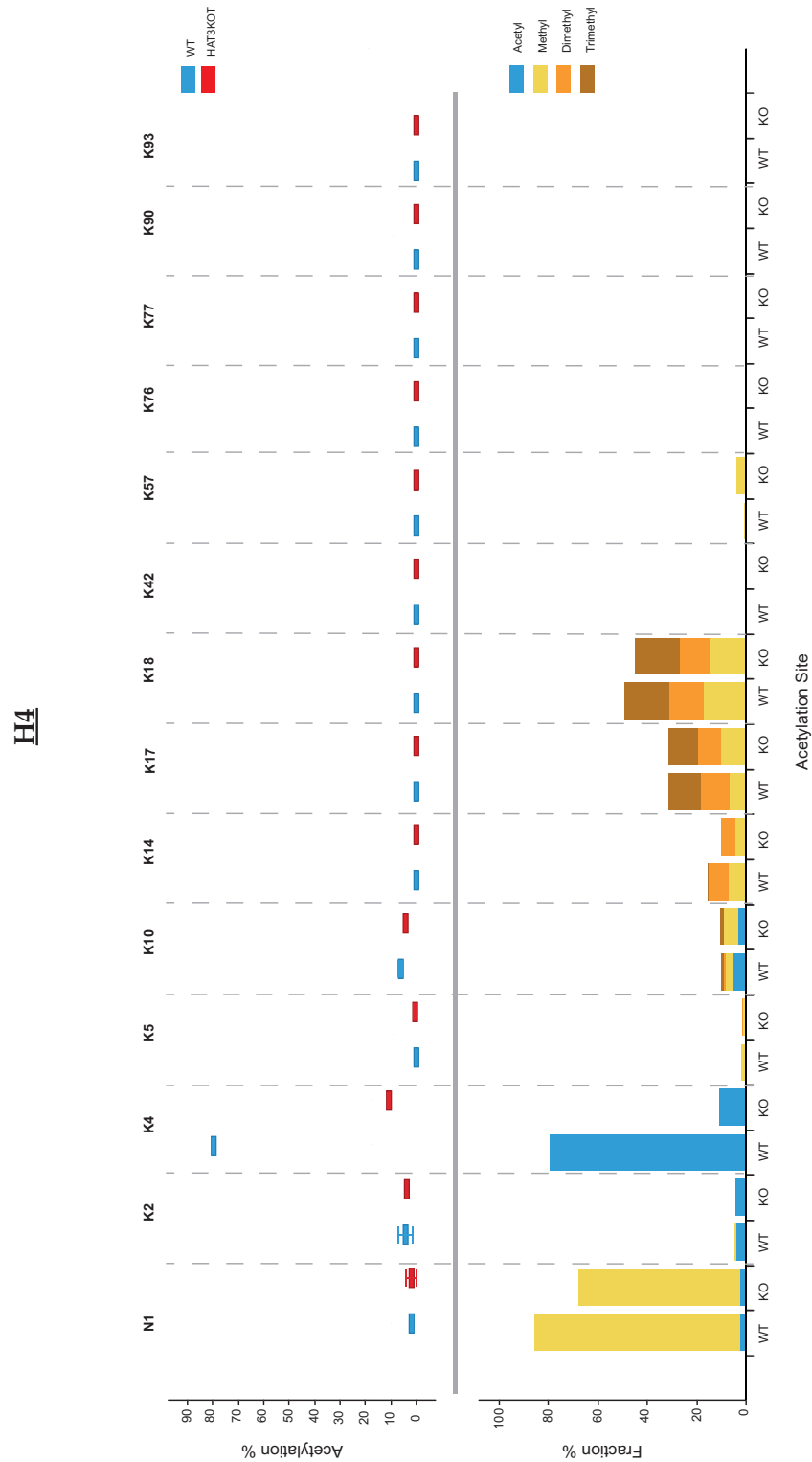


Figure 5.9.: Mononucleosomes from *T. brucei* wild-type (WT) and HAT3 knockout (KO) strains were purified and analyzed by fragment ion patchwork quantification. Upper panel: Acetylation degrees of all sites of histone H4. Shown is the median of three independent replicates; error bars show the median average deviation. Lower panel: Occurrence of other modifications; stacked bars represent the fraction of identified mono-, di-, and trimethylated peptides; the acetylation fraction was calculated from the corresponding acetylation degree [97].

5. Results and Discussion

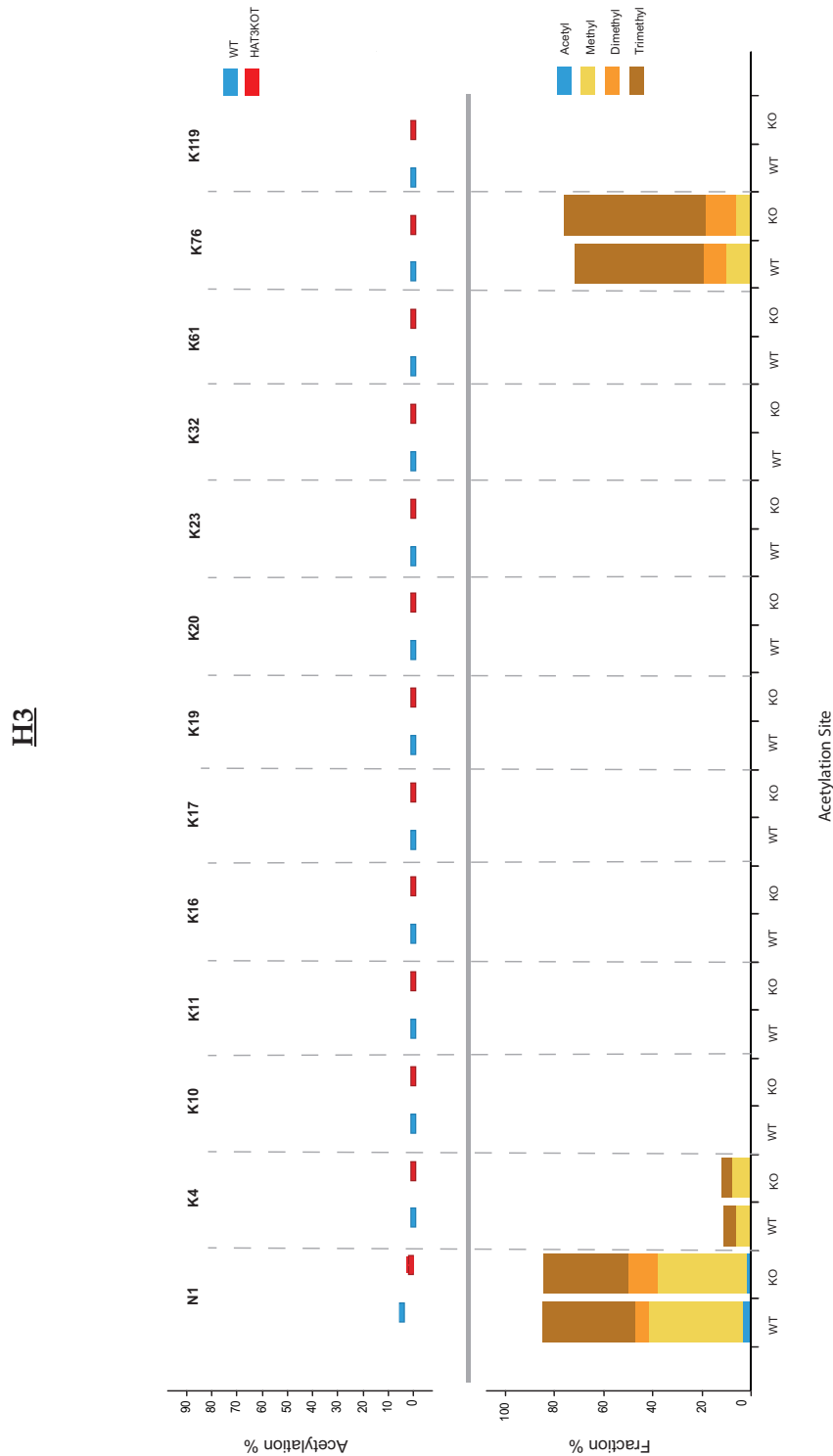


Figure 5.10.: Mononucleosomes from *T. brucei* wild-type (WT) and HAT3 knockout (KO) strains were purified and analyzed by fragment ion patchwork quantification. Upper panel: Acetylation degrees of all sites of histone H3. Shown is the median of three independent replicates; error bars show the median average deviation. The N-terminal acetylation degree was quantified in all 3 replicates of (WT) and (KO) experiments. Lower panel: Occurrence of other modifications; stacked bars represent the fraction of identified mono-, di-, and trimethylated peptides; the acetylation fraction was calculated from the corresponding acetylation degree.

5.4. Identification of Substrate Sites of Histone Acetyltransferases 3 (HAT3) in *T. Brucei*

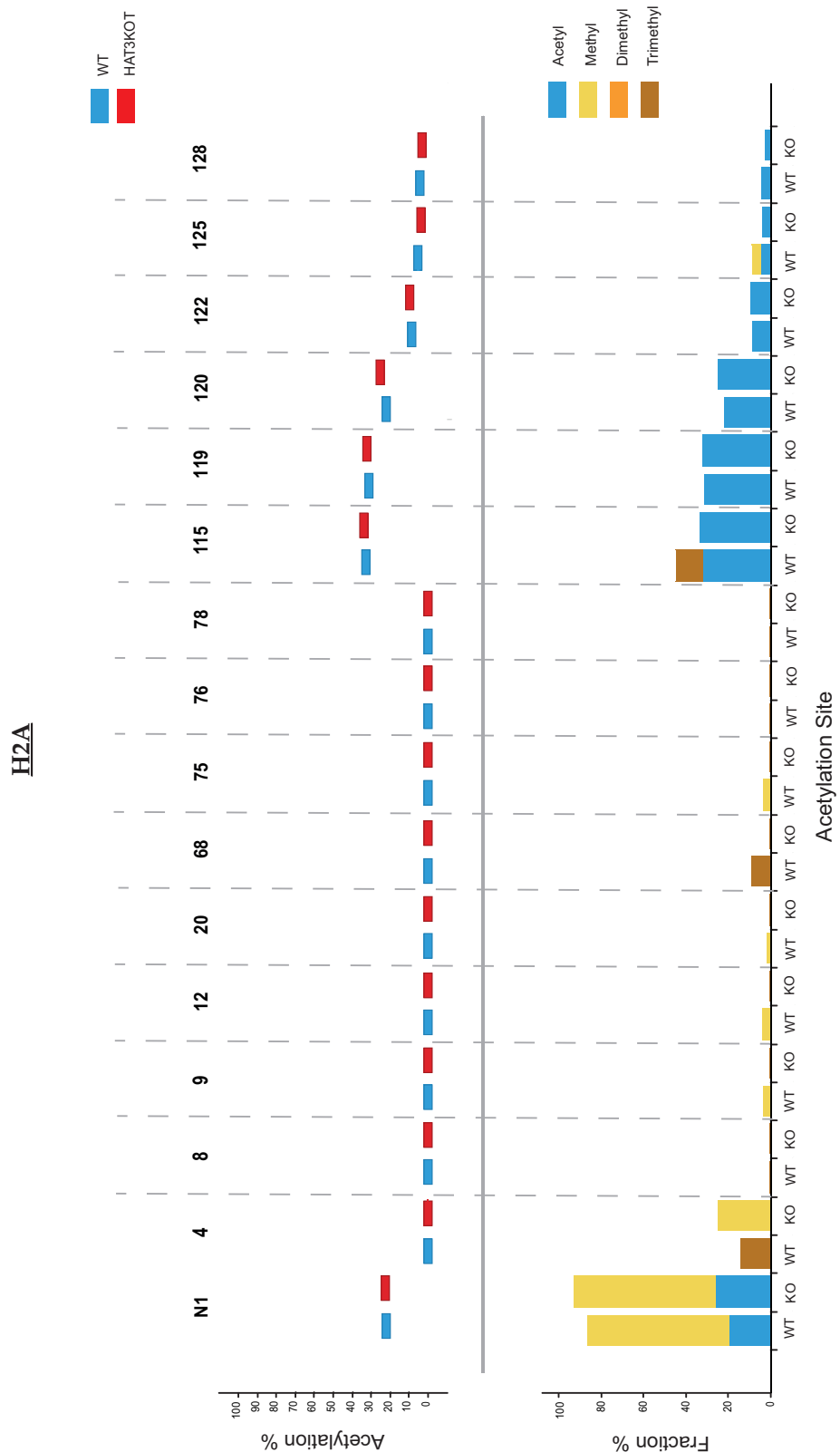


Figure 5.11.: Mononucleosomes from *T. brucei* wild-type (WT) and HAT3 knockout (KO) strains were purified and analyzed by fragment ion patchwork quantification. Upper panel: Acetylation degrees of all sites of histone H2A. Shown is the median of three independent replicates; error bars show the median average deviation. Lower panel: Occurrence of other modifications; stacked bars represent the fraction of identified mono-, di-, and trimethylated peptides; the acetylation fraction was calculated from the corresponding acetylation degree.

5. Results and Discussion

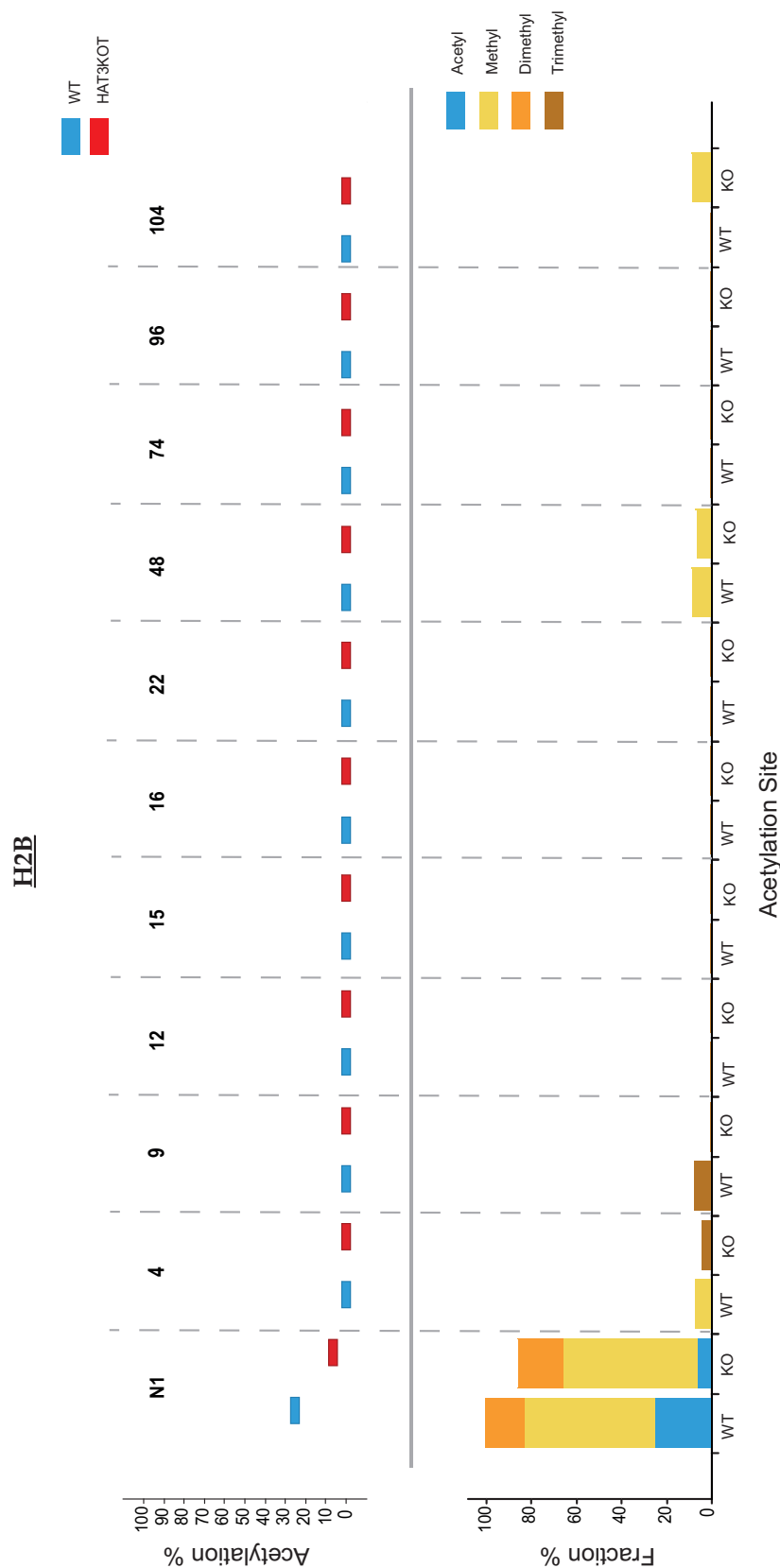


Figure 5.12.: Mononucleosomes from *T. brucei* wild-type (WT) and HAT3 knockout (KO) strains were purified and analyzed by fragment ion patchwork quantification. Upper panel: Acetylation degrees of all sites of histone H2B. Shown is the median of three independent replicates; error bars show the median average deviation. Lower panel: Occurrence of other modifications; stacked bars represent the fraction of identified mono-, di-, and trimethylated peptides; the acetylation fraction was calculated from the corresponding acetylation degree.

5.4. Identification of Substrate Sites of Histone Acetyltransferases 3 (HAT3) in *T. Brucei*

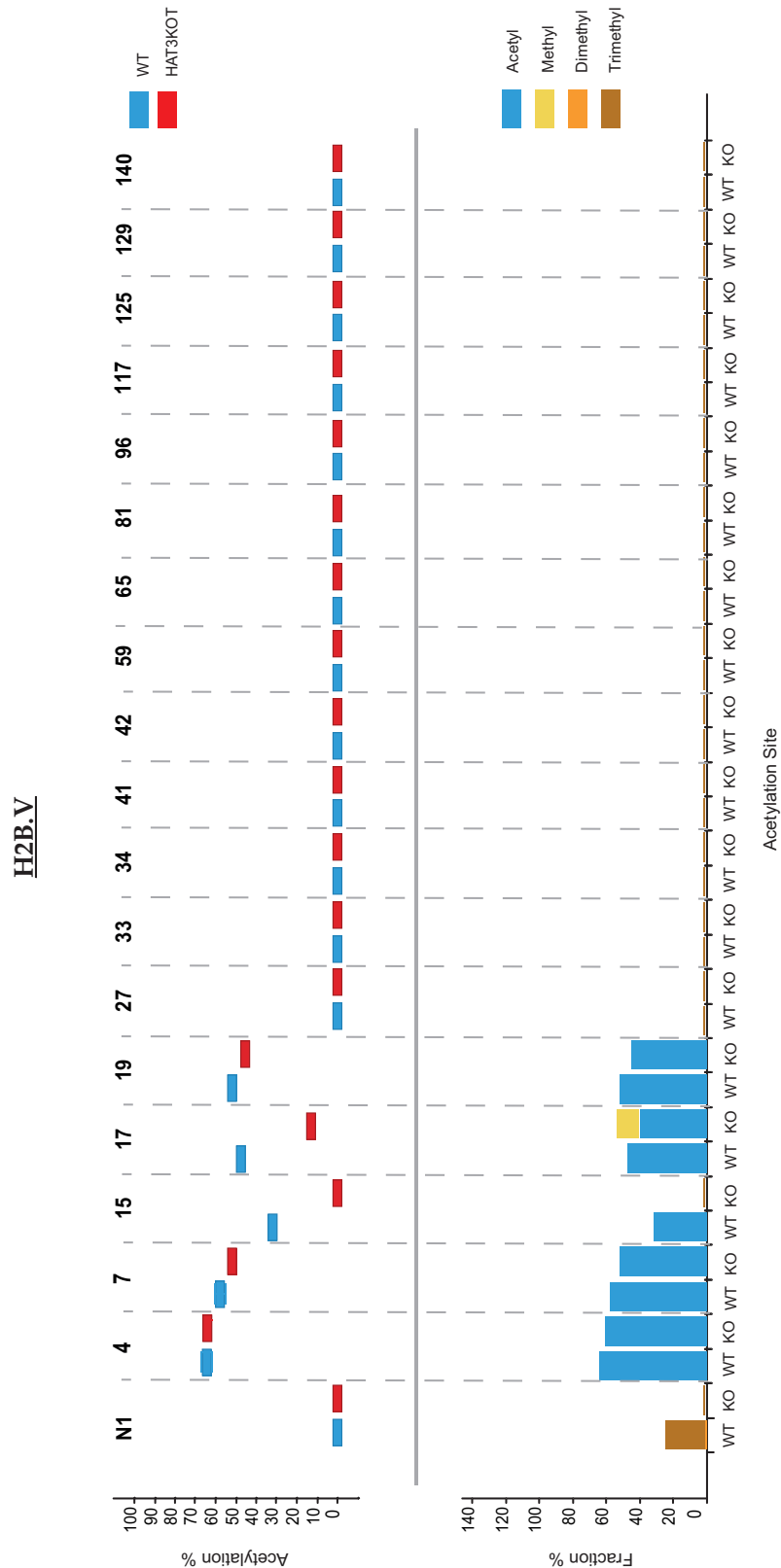


Figure 5.13.: Mononucleosomes from *T. brucei* wild-type (WT) and HAT3 knockout (KO) strains were purified and analyzed by fragment ion patchwork quantification. Upper panel: Acetylation degrees of all sites of histone H2B.V. Shown is the median of three independent replicates; error bars show the median average deviation. Lower panel: Occurrence of other modifications; stacked bars represent the fraction of identified mono-, di-, and trimethylated peptides; the acetylation fraction was calculated from the corresponding acetylation degree.

5. Results and Discussion

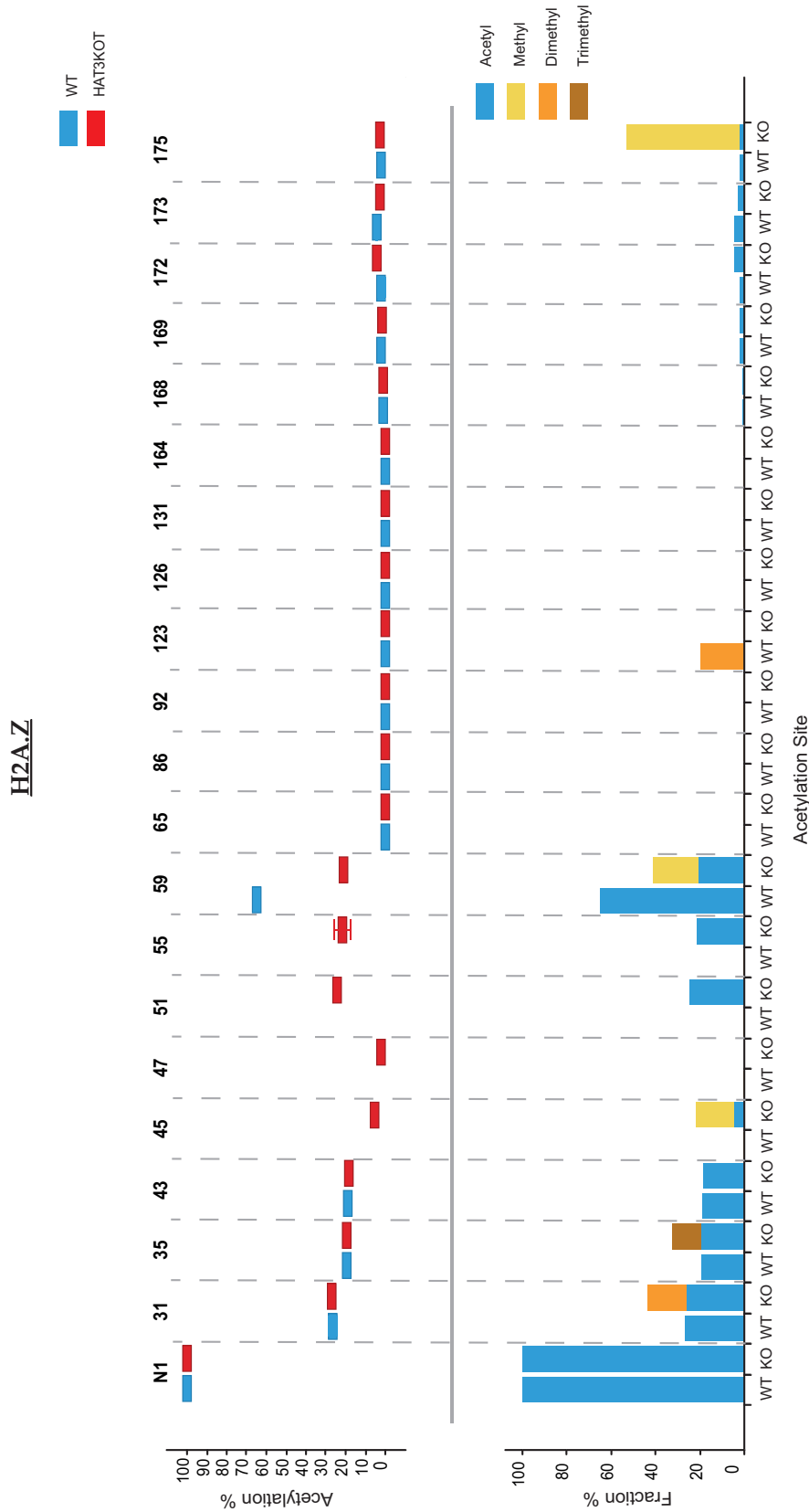


Figure 5.14.: Mononucleosomes from *T. brucei* wild-type (WT) and HAT3 knockout (KO) strains were purified and analyzed by fragment ion patchwork quantification. Upper panel: Acetylation degrees of all sites of histone H2A.Z. Shown is the median of three independent replicates; error bars show the median average deviation. Lower panel: Occurrence of other modifications; stacked bars represent the fraction of identified mono-, di-, and trimethylated peptides; the acetylation fraction was calculated from the corresponding acetylation degree.

5.4. Identification of Substrate Sites of Histone Acetyltransferases 3 (HAT3) in *T. Brucei*

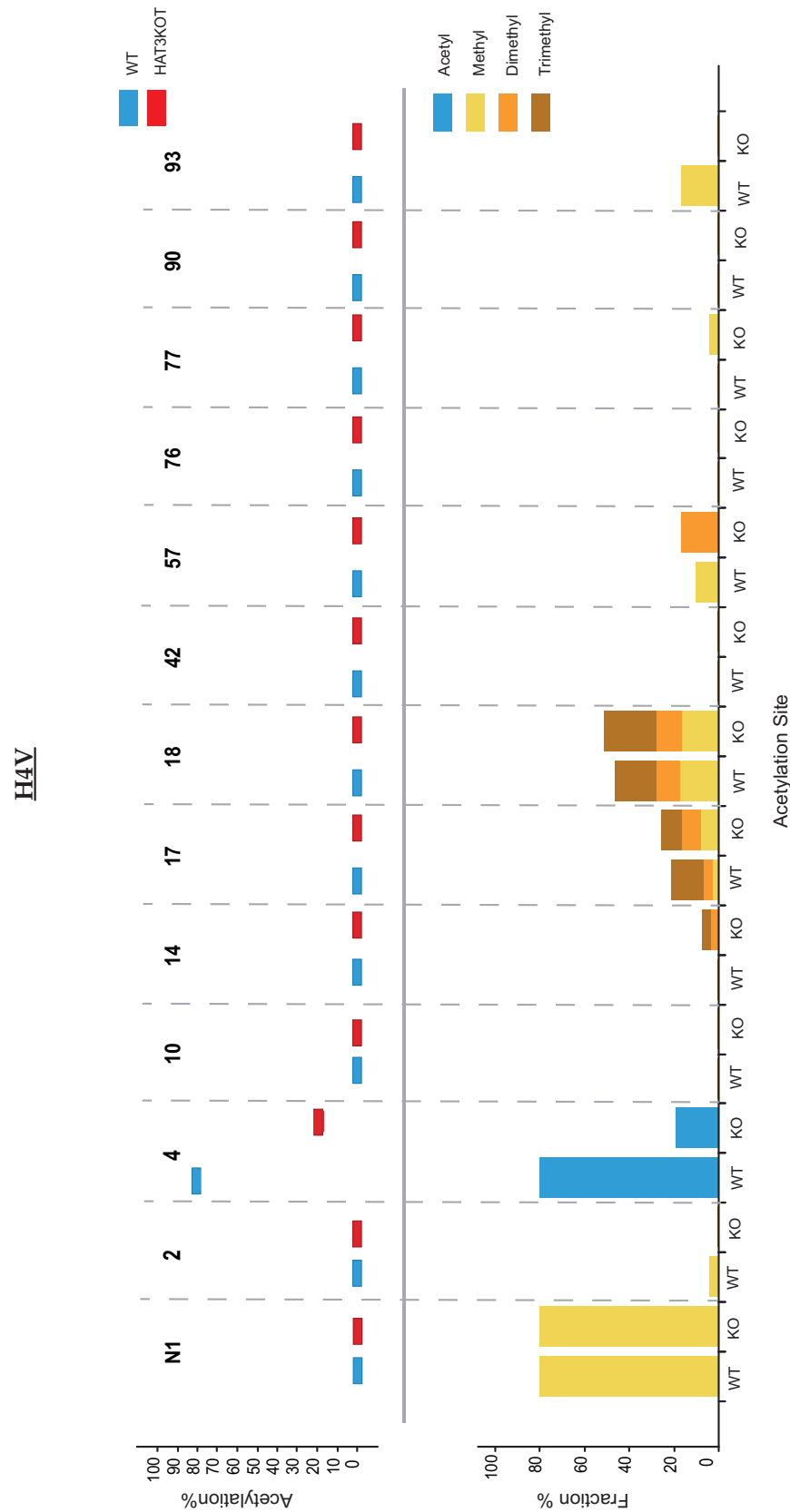


Figure 5.15.: Mononucleosomes from *T. brucei* wild-type (WT) and HAT3 knockout (KO) strains were purified and analyzed by fragment ion patchwork quantification. Upper panel: Acetylation degrees of all sites of histone H4V. Shown is the median of three independent replicates; error bars show the median average deviation. Lower panel: Occurrence of other modifications; stacked bars represent the fraction of identified mono-, di-, and trimethylated peptides; the acetylation fraction was calculated from the corresponding acetylation degree.

5. Results and Discussion

For H2A, acetylation degrees of 20 % on the N-terminal, 33% on K115, 31% on K119, 22% on K120, 8% on K122, 5% on K125, 4% on K128, and 0% on all other identified lysine residues were measured in the wild-type samples (Fig. 5.11). The acetylation degrees for these sites from the knockout samples did not change, which implies that HAT3 is not the acetyltransferase responsible for acetylating H2A. In H2B, the N-terminal is the only acetylated site (Fig. 5.12). On this site we measured an acetylation degree of 25% in wild type samples and 6% in knockout samples. This implies that HAT3 is not the only acetyltransferase responsible for acetylating N-terminal of H2B.

Like canonical histones, histones variants have lysine rich tails that undergo a number of various PTMs. For H2B.V, acetylation degrees of 64% on K4, 58% on K7, 32% on K15, 47% on K17, 52% on K19, and 0% on all other identified residues were measured in wild type samples. The knockout of HAT3 has not influenced the acetylation degree of K4, which implies that HAT3 is not responsible for the acetylation of this residue. The acetylation degree of K7, K17, and K19 have been dropped to 50%, 13%, and 44%, respectively. We believe that HAT3 is not the only HAT responsible for the acetylation of these sites and in the absence of the non-essential HAT3 another HAT can acetylate them. Our results also showed that the acetylation degree of K15 dropped from 32% in the wild-type to 0% in the HAT3 knockout. (Fig. 5.13).

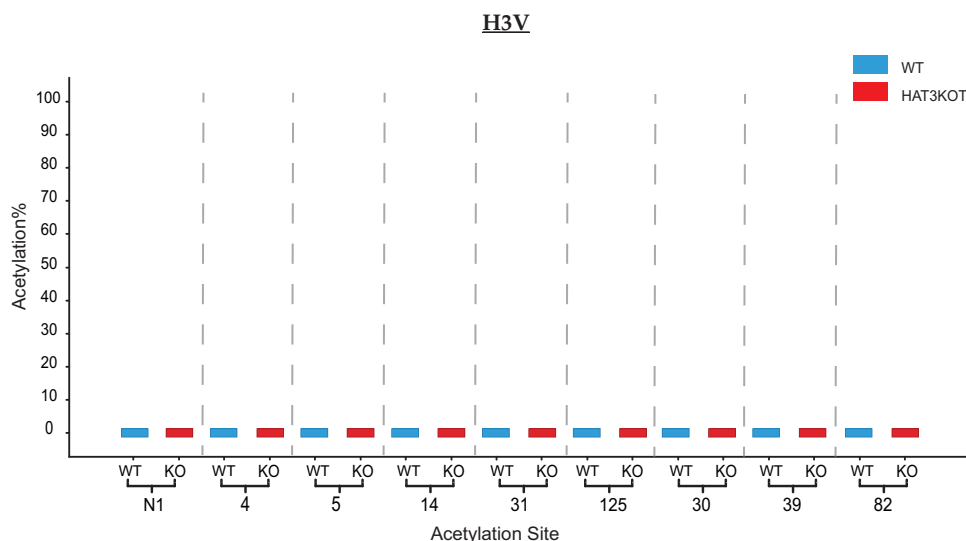


Figure 5.16.: Mononucleosomes from *T. brucei* wild-type (WT) and HAT3 knockout (KO) strains were purified and analyzed by fragment ion patchwork quantification. Upper panel: Acetylation degrees of all sites of histone H3V. Shown is the median of three independent replicates; error bars show the median average deviation. Lower panel: Occurrence of other modifications; stacked bars represent the fraction of identified mono-, di-, and trimethylated peptides; the acetylation fraction was calculated from the corresponding acetylation degree.

The N-terminal of H2A.Z is highly acetylated, where we measured an acetylation degree of 100% in both wild type and knockout samples. We also measured acetylation degrees of

5. Results and Discussion

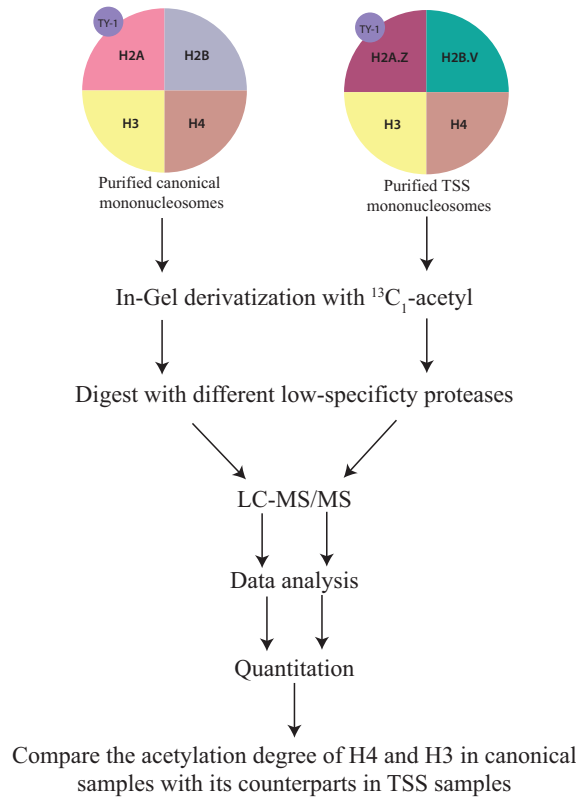


Figure 5.18.: Workflow for comparing the acetylation patterns of canonical and TSS nucleosomes.

We obtained the samples from Dr. Siegel group (Institute for Molecular Infection Biology (IMIB), Würzburg), where they prepared the mononucleosomes from two cell lines expressing Ty1-tagged H2A or H2A.Z. The tagged histones and histone variants were then immunoprecipitated, histones were extracted from the mononucleosome-containing supernatant by applying a standard acid extraction protocol, and concentrated before SDS-PAGE gel electrophoresis. Siegel *et al.* reported that histone variants H2A.Z and H2B.V mark the transcription start site, while the variants H3V and H4V mark the transcription termination site of the polycistronic transcription unit in *T. brucei*. The presence of the variants at the TSS alongside with the presence of PTMs at histones tails, make the DNA accessible to the transcription machinery.

For H4, we measured an acetylation degree of 95% at TSS H4K10 comparing to 6% acetylation degree of the canonical H4K10 (Fig. 5.19). It is suggested that high acetylation of K4K10 promote the replacement of H2A by H2A.Z at TSS. H2A.Z dimerizes with H2BV and both are essential for viability of *T. brucei*, thus, all *T. brucei* nucleosomes that contain H2A.Z also contain H2BV. ChIP-seq data that was published by Siegel *et al.* [104] showed that H4K10, which is acetylated by the essential acetyltransferase HAT2, is enriched at the transcription start site (TSS) in *T. brucei* as well as the trimethylation of H3K4. We also measured higher acetylation degrees at the N-terminal (64%), K2 (65%), and K5 (43%)

5.5. Differences in Acetylation Degrees Between Canonical and TSS-Nucleosomes in *T. Brucei*

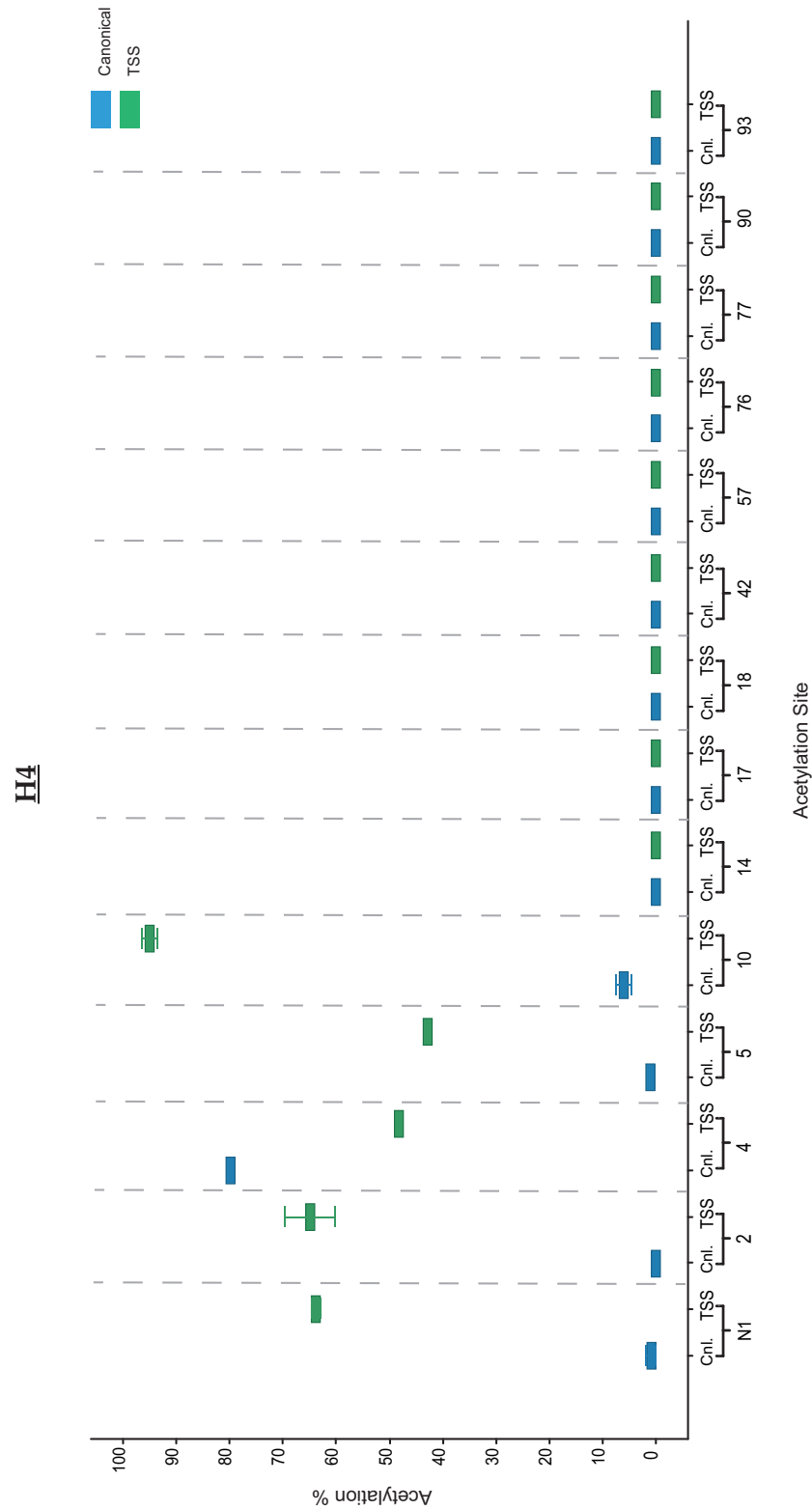


Figure 5.19.: Acetylation degree of canonical and TSS H4. Mononucleosomes from *T. brucei* wild-type (WT) and TSS strains were purified and analyzed by fragment ion patchwork quantification. H4K10 is enriched at TSS leading to an open chromatin structure and thus, promote the accessibility of DNA to transcription machinery. The acetylation degrees of all individual sites were measured.

5. Results and Discussion

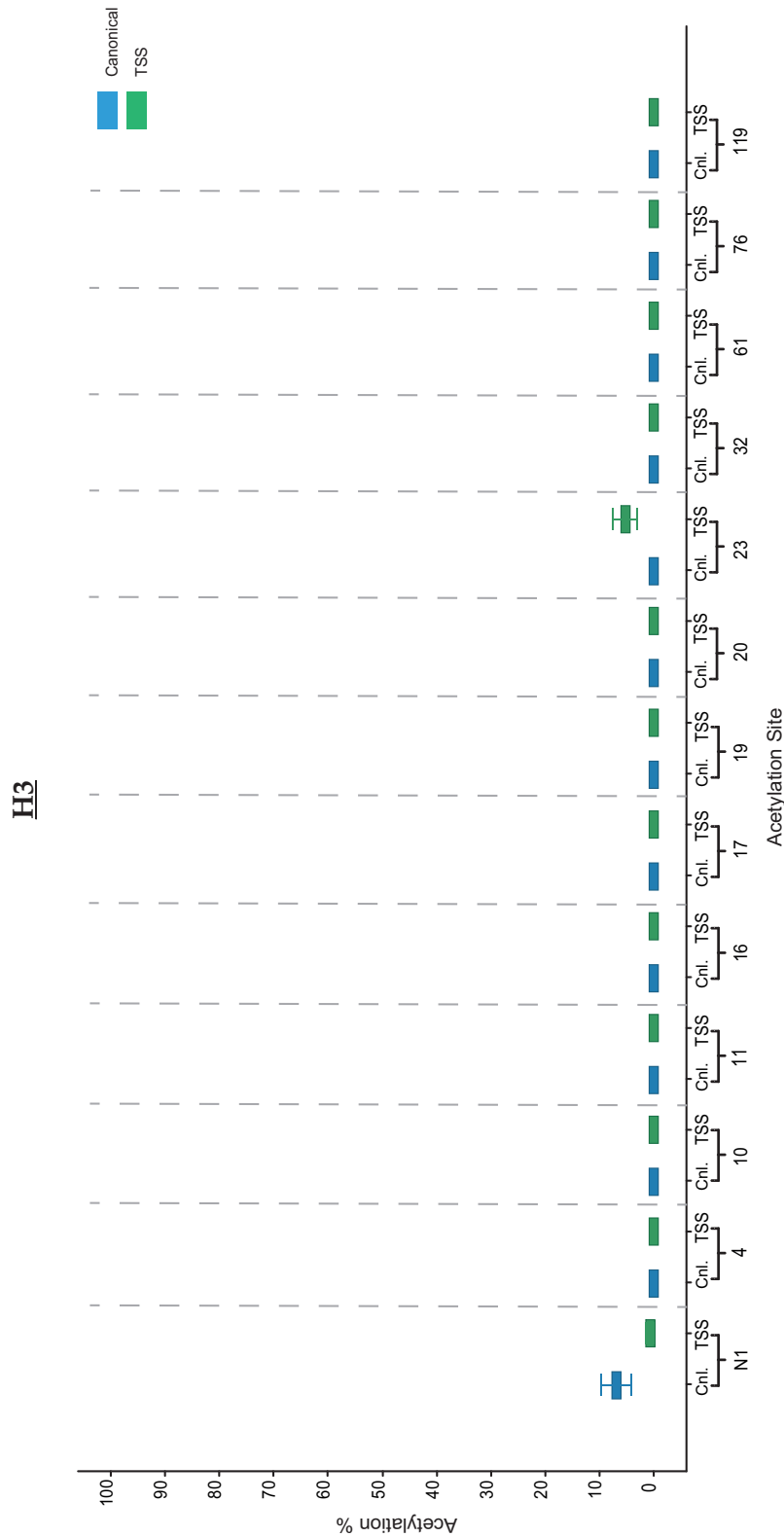


Figure 5.20.: Acetylation degree of canonical and TSS H3. Mononucleosomes from *T. brucei* wild-type (WT) and TSS strains were purified and analyzed by fragment ion patchwork quantification. H4K10 is enriched at TSS leading to an open chromatin structure and thus, promote the accessibility of DNA to transcription machinery. The acetylation degrees of all individual sites were measured.

of the TSS nucleosome comparing to 0% at these three sites in the canonical nucleosome. The acetylation degree measured for H4K4 was 80% and 48% for canonical and TSS nucleosomes, respectively. For all other identified sites, the acetylation degree was 0% in both samples.

For H3, we measured a low acetylation degree of 5% at H3K23 in TSS sample comparing to 0% at the same site in canonical samples (Fig. 5.20). The acetylation degree at the N-term was 6% and 0% for canonical and TSS samples, respectively. All other identified sites of H3 are not acetylated in both samples.

The acetylation degrees of all identified sites of H2A, H2B, H2A.Z, and H2B.V were measured, and they were similar to the acetylation degrees measured for wild-type samples. However, comparing the acetylation degrees of the canonical H2A and H2B with their variants H2A.Z and H2B.V is irrelevant for the purposes of this study.

5.6. Conclusion

Fragment ion patchwork quantification is a smart and powerful approach for measuring site-specific acetylation degrees. The robustness and straightforwardness of this approach allowed us to quantify the acetylation degree of all lysines for all core histones from *Trypanosoma brucei* within a single analysis. It is thereby remarkable that this was achieved from a crude histone-enriched fraction that contained several hundred additional proteins. Full coverage of acetylation sites, which is a formidable challenge with other existing methods, was achieved by the use of low-specificity proteases even for large lysine clusters in histone tails. The use of different low-specificity proteases generated a high number of overlapping peptides, and subsequently a high number of b and y fragment ions for each site. Using these fragment ions for the quantitation allowed us to determine the site-specific acetylation degree with high confidence.

Using our approach, we were able to circumvent the problem of different elution times caused by the deuterioacetylation. The absence of any retention time shift between isotopologues is prerequisite when quantification is performed on the fragment ion level, which is the only way to reveal site-specific acetylation degrees for peptides with more than one lysine. In contrast to targeted methods, our approach allows quantification without prior knowledge of peptide features, such as sequence, retention time, charge state, modifications, and is compatible with the application of low-specificity proteases.

With this method a variety of crucial biological and pharmacological questions can be addressed, for example, the identification of acetyltransferase and deacetylase target sites or the effect of HDAC inhibitors on cancer treatments. Due to the high accuracy of this method, it should be possible to measure even subtle changes in the acetylation degree.

5. Results and Discussion

The outlined concept of fragment ion patchwork quantification is even more general and could be adopted for measuring the site-specific modification degree of other types of lysine modifications, such as formylation, propionylation, myristoylation, or ubiquitination.

Some of the results of this chapter have been published in the *Journal of Analytical Chemistry* in 2015.

6. Carbodiimide-mediated 4-ABA Labeling of Phosphosites

6.1. Introduction

Phosphorylation is the most pervasive and pleiotropic modification in biology that mainly occurs on serine (Ser), threonine (Thr) and tyrosine (Tyr) [105]. It is a reversible PTM that is tightly regulated by the dynamic interplay between protein kinases and phosphatases on their target proteins. Phosphorylation occurs when kinases transfer a phosphate group from adenosine triphosphate (ATP) to the target protein, while dephosphorylation is catalyzed by phosphatases releasing inorganic phosphate (Fig. 6.1). It is a key mechanism for regulating proteins activities and intimately involved in almost every cellular process including metabolism, cell signaling, protein-protein interactions, cytoskeleton remodeling, PTMs cross-talk and cell growth and development [106].

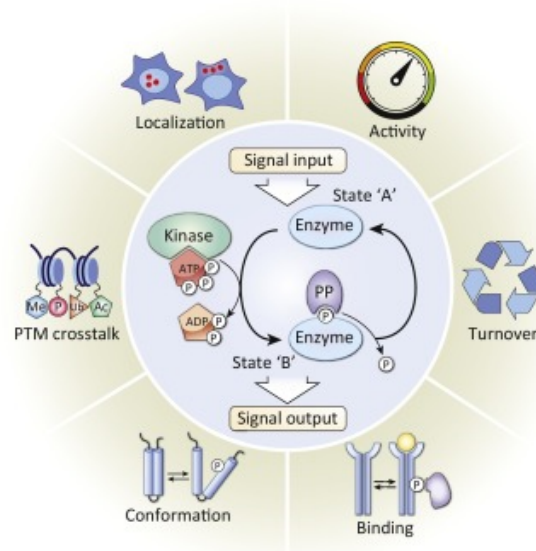


Figure 6.1.: Mechanism of protein phosphorylation and related cellular processes. **Center circle** shows the mechanism of the reversible protein phosphorylation that occurs by the addition of phosphate donated from ATP and the removal of phosphate from a phosphorylated protein substrate, catalyzed by protein kinase and phosphatase enzymes, respectively. **Outer circle** shows the cellular processes regulated by the phosphorylation and dephosphorylation of Ser, Thr and Tyr [107].

6. Carbodiimide-mediated 4-ABA Labeling of Phosphosites

In eukaryotes, ~30% of proteins are phosphorylated at any given time [108]. Analysis of the human genome showed that there are about 518 kinases and 214 phosphatases, about 80% of the kinases are serine/threonine kinases whereas tyrosine kinases represents 20% [109, 110]. The phosphoamino acid content ratio (pSer:pThr:pTyr) of human cell is 90:10:1 [111]; thus, the majority of protein phosphorylation happens on serine residues, followed by threonine and then tyrosine.

To date, four types of protein phosphorylation are known, depending on the phosphorylated amino acid [112]: (i) *O*-phosphomonoesters (*O*-phosphates), resulting from the phosphorylation of the hydroxyl group of serine, threonine, and tyrosine residues Fig. 6.2, (ii) *N*-phosphoramidates (*N*-phosphates), resulting from the phosphorylation of the amino groups of arginine, lysine and histidine residues, (iii) *S*-phosphothioesters (*S*-phosphates) resulting from the phosphorylation of cysteine, and (iv) *O*-acylphosphates resulting from the phosphorylation of aspartic acid and glutamic acid. These different phosphorylation species differ significantly in their chemical stability [113]. Under acidic conditions, all *O*-phosphates are highly stable, phosphocysteine is moderately stable, whereas *N*-phosphates and *O*-acylphosphates are unstable [113]. Therefore, *O*-phosphates are suitable for the LC-MS/MS analysis.

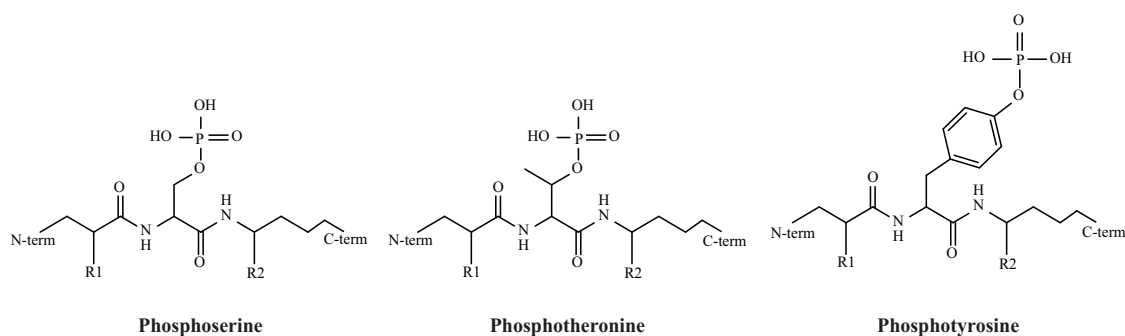


Figure 6.2.: Chemical structures of phosphoserine, phosphothreonine and phosphotyrosine peptides.

Protein phosphorylation events are dynamic and reflect the current physiological states and needs of the cell. Therefore, understanding these events on a molecular level is foremost important toward understanding the difference between “healthy” and “abnormal” cellular states. It also provides new approaches for therapy of disease linked to the imbalanced kinases and phosphatases activities like cancer [114], diabetes [115], and Alzheimer’s disease [116].

6.2. Qualitative Analysis of Phosphorylation

Several analytical methods have been developed for the analysis of *O*-phosphorylation. Radioactive labeling of the protein with ^{32}P to monitor phosphorylation and Edman

degradation to localize the phosphosites are considered the traditional methods used for the analysis of phosphorylation [117]. In these methods, ^{32}P is incorporated into cellular proteins via treatment with radiolabeled ATP, to add a phosphorus label to the phosphosites. The radioactive proteins are detectable during fractionation procedures e.g. two-dimensional gel electrophoresis or HPLC and the radioactive peptide fraction can then be identified. Phosphorylation site(s) can be localized by proteolytic digestion of the radiolabeled protein, separation, and detection of phosphorylated peptides (e.g. by two-dimensional peptide mapping), followed by peptide sequencing by Edman degradation [118]. Despite the successes reported of these methods, they have the disadvantage of being laborious, requiring large amounts of purified phosphoproteins, and requiring the use of considerable amount of radioactive material. The analysis of phosphoproteins can be carried out with the aid of phosphospecific antibodies in combination with Western blot. These antibodies can be used to immunoprecipitate, and therefore to enrich, phosphorylated proteins from complex mixtures of proteins. However, these antibodies are highly selective to phosphotyrosine, and currently, there are no antibodies that are suitable for the enrichment of phosphoserine and phosphothreonine [119]. A major drawback of these methods is that phosphorylation can not be directly localized.

Phosphoprotein Specific Staining: Direct staining of phosphoproteins with a phosphospecific dye after an SDS-PAGE gel is the easiest approach to detect phosphoproteins in a sample [120]. The Pro-Q Diamond fluorescent dye selectively binds to phosphoproteins; the sensitivity of the dye depends on the number of phosphosites in the protein [121]. The later dyeing of the total proteins on the same gel with SYPRO Ruby dye allows distinguishing the low represented but highly phosphorylated proteins from highly represented but poorly phosphorylated ones. This method is very useful for a preliminary screening, but still not sufficient for a comprehensive analysis of the phosphoproteome. Therefore, LC-MS/MS has become the favored and most efficient method for the analysis of phosphoproteins.

Mass Spectrometry-based Methods: Tandem mass spectrometry has an enormous impact on the phosphorylation analysis. Recent methodological advances in MS combined with the continuous improvements in MS instrumentation, protocols, and data processing procedures facilitate the accurate determination of the molecular weight of intact phosphoproteins and the identification of phosphosites without the need of using radioactive labels or phosphospecific antibodies. Comparing the molecular weight of the phosphoprotein with its unphosphorylated counterpart gives a rough estimation of the average number of phosphosites [122]. The high resolution of current mass spectrometers allows the determination of the distribution of these phosphosites. However, despite the constant

6. Carbodiimide-mediated 4-ABA Labeling of Phosphosites

improvements, MS analysis of phosphorylation still facing some limitations that make the identification and quantification of phosphopeptides a formidable challenge.

Phosphopeptides Fragmentation: The MS analysis of phosphopeptides has some peculiarities compared to the analysis of their unmodified counterparts. For example, CID spectra of pSer/pThr-phosphopeptides exhibit a predominant peak for the neutral loss of phosphoric acid H_3PO_4 (97.9768 Da) accompanied by a few sequence-specific fragment ions of moderate to low abundance (Fig. 6.3). This indicates that the peptide is phosphorylated but the sequence information is very poor and the correct localization of the phosphorylation site in the peptide is impaired. In the case of multi-phosphorylated peptides, this effect is more pronounced due to the occurrence of several neutral losses. The loss of H_3PO_4 from pSer/pThr fragment ion is indistinguishable from loss of 18 Da (H_2O) from a non-phosphorylated fragment ion, which complicates the interpretation of a CID spectrum. The intensity of the neutral loss depends on the nature of the mass analyzer, the peptide sequence, and the peptide charge. Thus, doubly charged peptide ions often exhibit a stronger neutral loss than triple-charged peptide ions [123].

Phosphotyrosine exhibits a neutral loss of the phosphate group HPO_3 (79.9663 Da). MS/MS spectra of tyrosine phosphorylated peptides tend to resemble those of their unphosphorylated counterparts, but they can be distinguished by the characteristic immonium ion of $m/z = 216.0426$ of phosphotyrosine. This fragment ion is specific for phosphotyrosine; therefore, precursor ion scanning coupled with subsequent MS/MS analysis of the selected precursor ions has been used to identify these pTyr-containing peptides from complex mixtures [124]. The neutral loss of the phosphate group can also be used for the search for phosphopeptides in complex mixtures due to its typical mass [125]. A second step of activation can be used to obtain more sequence information from peptides that show an extensive neutral loss in their CID spectra. Performing an MS^3 or multistage activation (MSA) generate more sequence-informative fragments from the ion produced by the neutral loss and may overcome the lack of backbone fragmentation in phosphopeptides [126].

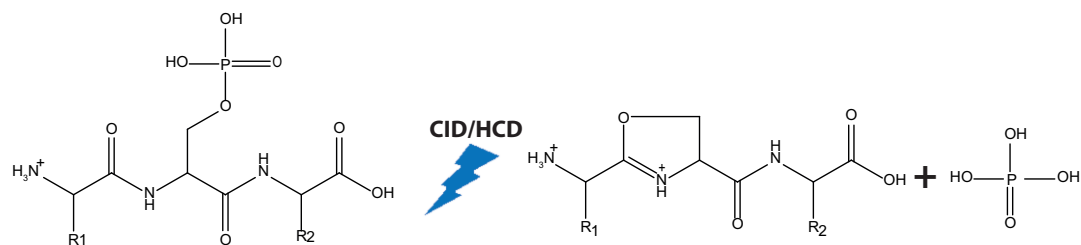


Figure 6.3.: Fragmentation scheme for loss of phosphoric acid from a multiply protonated phosphopeptide by CID.

Although the neutral loss exhibited by CID/HCD facilitates phosphopeptide identifica-

tion, it hampers backbone fragmentation for peptide sequencing and assignment of the phosphorylation site. Therefore, two alternative fragmentation methods, electron capture dissociation (ECD) [127] and electron transfer dissociation (ETD) [128] that show tremendous promise in fragmentation of large peptides and peptides with PTMs have been proposed (Fig 6.4). These two nonergodic fragmentation variants do not lead to a neutral loss on the phosphate group since cleavage of the peptide backbone takes place between the nitrogen atom and the C α atom of the peptide bond, leading to the formation of *c*- and *z*-ions. Due to this advantage, the labile phosphate groups remain on the resulting *c* and *z* fragment ions, enabling phosphorylation site assignment. However, ECD and ETD lead to very good fragment spectra for highly charged peptides (charge of 3+ or higher), while fragmentation for doubly charged ions is relatively poor [129]. Comparing the performance of CID with ECD or ETD, CID appears to be better suited for phosphopeptide identification, whereas ECD and ETD are better suited for the localization of the phosphorylation site [130, 131]. The recently introduced EThcD fragmentation combines ETD and HCD to generate both *b*/*y* and *c*/*z* ions that provide data-rich MS/MS spectra, yield higher peptide sequence coverage and more confident localization of phosphorylation sites [132].

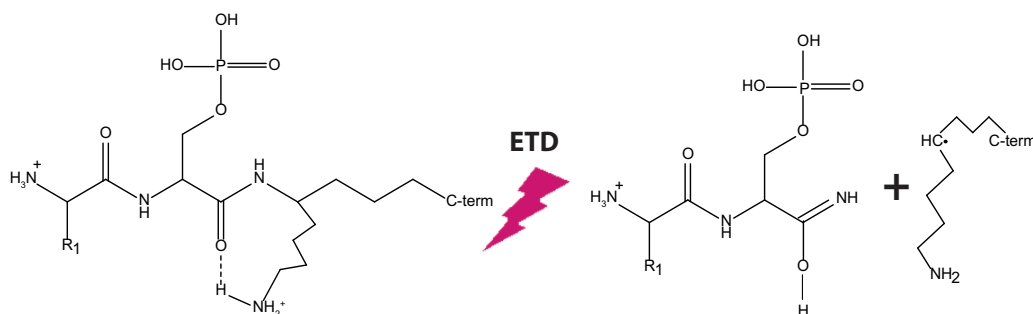


Figure 6.4.: ETD fragmentation scheme of phosphorylated peptidied.

Phosphopeptides Enrichment: Although protein phosphorylation is a very frequent event in cells, often only a small fraction of the protein in the phosphorylated form is found due to the large excess of unphosphorylated protein. Therefore, it is often necessary to perform enrichment of phosphopeptides prior the LC-MS/MS analysis to reduce the suppression of the signal by the highly abundant unphosphorylated peptides [133]. Furthermore, in the case of the data-dependent acquisition, the instrument preferably fragment and collect spectra of high abundant peptides [133]. In addition, the higher the complexity of a sample the higher the chance of isobaric overlap, which might affect the data analysis [134]. It is, thus, useful for qualitative phosphorylation analysis, if possible, only phosphorylated peptides are analyzed. Therefore, various methods have been developed for phosphoproteins and phosphopeptides enrichment. For example, phosphoprotein enrichment by anti-phosphoantibodies and immunoprecipitation can be achieved. This

6. Carbodiimide-mediated 4-ABA Labeling of Phosphosites

method is used mainly to enrich proteins with phosphorylated tyrosine since phosphoantibodies have higher specificity for phosphotyrosine than phosphoserine and phosphothreonine [119].

In addition, various methods for phosphopeptides enrichment according to their physicochemical properties exist. The phosphate group of serine, threonine, and tyrosine has distinct chemical properties compared to unmodified and differently modified peptides, which allows the enrichment of the phosphate group. For example, the low acidity constant (pK_a value, $pK_a \approx 2$) of the phosphate group causes a shift in the isoelectric point (pI value) of the phosphopeptide. This may be used for the precipitation of the phosphopeptides [135], for the partial enrichment of the phosphopeptides in isoelectric focusing [136] or in ion exchange chromatography. Cation exchange chromatography [137], and anion exchange chromatography [138] can also be used for phosphopeptide enrichment as well as hydrophilic interaction liquid chromatography (HILIC) that can be used for the prefractionation and enrichment of phosphopeptides [139]. Electrostatic Repulsion-Hydrophilic Interaction Chromatography (ERLIC), which is a subset of HILIC separations that employs hydrophilic interaction and electrostatic forces [140] and permits the selective isolation of phosphopeptides.

Phosphate group has two nearby oxygen atoms, that can form strong bidentate interactions with metal ions. The enrichment of phosphopeptides by means of immobilized metal affinity chromatography (IMAC) and metal oxide affinity chromatography (MOAC) can also be exploited by the addition of such as TiO_2 [141], $Al(OH)_3$ [142], ZrO_2 [143], Nb_2O_5 [144], SnO_2 [145], or HfO_2 [146]. MOAC shows more specificity towards mono-phosphorylated peptides, while IMAC is more specific for multi-phosphorylated peptides, but has a low capacity and selectivity when used with highly complex samples [147]. A promising approach for phosphopeptide enrichment is the sequential elution from IMAC (SIMAC), where both mono- and multi-phosphorylated peptides can be enriched using IMAC and TiO_2 [148].

6.3. Challenges In LC-MS/MS Phosphorylation Identification and Quantification

Phosphorylation is the best studied posttranslational modification, and despite continuous improvements, phosphoproteomics still faces many challenges that are often neglected. These challenges can be biological or MS related challenges. Among those challenges:

1. Biological challenges:
 - a) Phosphorylation is a transient modification with a short lifetime, specially phosphotyrosine due to its very active phosphatase [149], which makes capturing the

6.3. Challenges In LC-MS/MS Phosphorylation Identification and Quantification

phosphorylated site for analysis difficult.[150].

- b) The low stoichiometry of phosphorylation, only a small fraction of the population of a particular protein may be phosphorylated at any given time [110].
- c) Different phosphorylated isoforms that have different phosphorylated sites within a protein may exist simultaneously in this substoichiometric population [150].

2. MS related challenges:

- a) The presence of PTMs impairs the proteolytic digestion efficiency [151].
- b) The loss of phosphopeptides during sample preparation and chromatography [152].
- c) Different ionization efficiencies of phosphorylated and unphosphorylated peptides [153].
- d) Weak ionization efficiency of multi-phosphorylated peptides in +ve mode [118].
- e) Poor MS/MS spectra generated by CID or HCD fragmentation due to the peculiar behavior of the labile phosphate group upon fragmentation [123] impairs both:
 - i. The correct localization of phosphorylation site(s).
 - ii. The identification of peptide/protein sequence.

Herein, we tried to develop a method to overcome most of the MS-related challenges. In order to preserve the labile phosphate group during fragmentation, it's better to use ETD instead of CID/HCD fragmentation. However, ETD is more efficient to fragment peptides with charges higher than 3+. This fragmentation method might be effective for the analysis of singly phosphorylated peptides but will not be suitable for the analysis of multi-phosphorylated sites since the negative phosphate groups will decrease the overall phosphopeptide's net charge. Therefore, reversing and increasing the charge state of the multi-phosphorylated peptides will have a huge impact in improving the ETD fragmentation of these peptides and thus, pinpointing the correct sites of phosphorylation, which will lead to better sequence identification.

A number of methods have been developed to increase the charge-state of peptides for ETD fragmentation during LC separation. Kjeldsen et al. [154] have found that adding *m*-nitrobenzyl alcohol to both mobile phases of LC separation increases the abundance of the 3+ charged tryptic peptide, and thus, improve ETD performance for peptide sequencing and identification. However, the analysis of phosphorylation in complex samples is hampered by the non-specific binding of nonphosphorylated species containing multiple acidic residues. The addition of DHB to loading and washing buffers [141] or blocking the acidic residues with *O*-methyl esterification [155] can increase the specificity to phosphopeptides binding. Li and Cole demonstrated increased charging for polypeptides with

6. Carbodiimide-mediated 4-ABA Labeling of Phosphosites

decreasing ESI tip orifice diameter [156]. Iavarone and Williams reported that applying high surface tension, low relative volatility solvents could enhance ESI charging, or “supercharge” [157].

Some other methods have been developed to derivatize carboxyl groups on peptides to increase charge states during LC-MS. Lu and co-workers used EDC/HOAt/DMF chemistry to attach pyrimidyl piperazine groups carboxyl groups to enhance ionization efficiencies of standard phosphopeptides by MALDI and ESI and showed that the increase from 1+ to 2+ precursors allowed fragmentation by ETD [158]. However, fragmentation was still not efficient due to the low charge-state. Ko and Brodbelt derivatized carboxyl groups by attaching basic or fixed-charge groups to their side chains [159]. They thoroughly analyzed the charge states and ETD fragmentation efficiencies, both of which increased substantially for the modified peptides. Despite efforts to optimize the EDC/HOAt/DMF chemistry, they reported that labeling was incomplete and dependent on the peptide sequence; they estimated the reaction efficiency as $\geq 70\%$ of the C-terminal carboxyls, but much lower and less consistent for Asp and Glu side-chains. Smith and coworkers have developed a derivatization method that yields nearly complete derivatization of all carboxyl groups (Asp, Glu, and C-termini) on small proteins or peptides [160]. They used 2-steps reaction where they first protected the native amino groups by reductive methylation. In the second step the performed amidation reaction, whereby the peptide carboxylic acids are converted to tertiary or quaternary amine groups leading to substantially higher charge states after derivatization [161]. By applying the concept of phosphoramidate chemistry (PAC) introduced by Zhou [162], which entails reaction of the phosphate group in a phosphopeptide with primary amines to form a phosphoramidate. Panchaud et al. [163] introduced the ANIBAL approach, where they used aniline and benzoic acid in their ^{12}C and ^{13}C forms to modify carboxylic and amino groups at the protein level and subsequently use the information for the quantitative analysis. Although this method is also applicable to label phosphosites, a major draw back is the increased hydrophobicity of the labeled peptides, which results in prolonged retention time, and thus the loss of many of these peptides during liquid chromatography. These losses decrease the number of identified peptides and consequently decreases the protein sequence coverage. When we tested this approach, we measured a ~ 6 min retention time shift/labeled site between labeled and unlabeled peptides. This retention time shift increases depending on the peptide sequence, peptide length, and the number of labeled sites.

Shi et al. [164] introduced a method to modify the phosphate groups on phosphoserine peptides to the corresponding phosphoramidates, using 2-aminobenzylamine (2-ABA). Upon CID fragmentation, the modified peptides release a positively charged phosphoramidate that forms cyclophosphoramidate (CyPAA) ion via gas-phase intramolecular elimination that is stable in solution. This ion is specific for detecting phosphoserine peptides and has sufficient mass defect that allows separating the ion from isobaric interfering ions and

6.3. Challenges In LC-MS/MS Phosphorylation Identification and Quantification

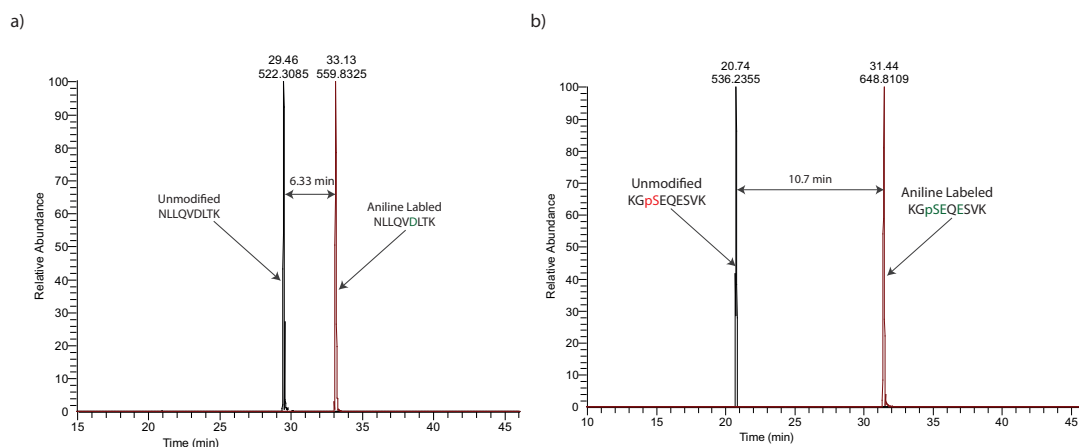


Figure 6.5.: Retention time difference caused by aniline labeling of carboxyl and phosphate groups of PKA. **a)** Base peak of peptide NLLQVDLTK, 2+ ($m/z = 522.3085$) with a retention time 29.28 min. Labeling of the Asp residue with aniline increases the hydrophobicity of the peptide and causes ~ 6 min increase in the retention time. **b)** Base peak of peptide KGpSEQESVK, 2+ ($m/z = 536.2355$) with a retention time 20.44 min. Labeling of the two Glu residues and the phosphate group of pSer with aniline causes ~ 11 min increase in the retention time.

allows fast analysis of phosphoserine-containing peptides.

Here we introduce a method for carbodiimide-mediated labeling the phosphate groups using 4-aminobenzylamine (4-ABA) to reverse the charge of the phosphate(s) and increase the net charge of the modified phosphopeptide. 4-ABA has two amino groups, both of which could potentially be reactive to the carbodiimide-activated peptidyl phosphates. When the carbodiimide-mediated phosphoramidation reaction is performed under slightly acidic conditions (pH 5.5), the aromatic amino group of the 4-ABA forms a stable amide bond with the phosphate. We found that 4-ABA is more stable than 2-ABA, and the peptides labeled with 4-ABA are less hydrophobic than those labeled with aniline. Unlike the labeling approaches with aniline and 2-ABA that were performed solely at the peptide level, our approach is applicable to both peptide and protein levels. Performing the labeling at the protein level reduces sample variations during the sample preparation.

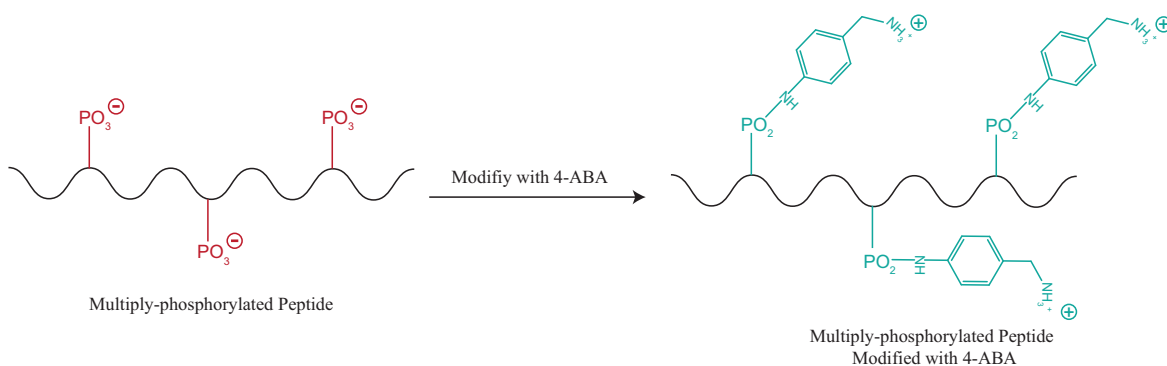


Figure 6.6.: Reverse of the phosphate negative charge.

7. Material and Methods

7.1. Reagents

All chemicals were purchased from Sigma-Aldrich Chemical Co. unless stated otherwise. Formic acid FA (for mass spectrometry ~98%), trifluoroacetic acid TFA (CHROMASOLV[®], for HPLC, ≥99.0%), trifluoromethanesulfonic acid TFMS (ReagentPlus[®], ≥99%), heptafluorobutyric acid HFBA (for ion chromatography, ≥99.5%), trichloroacetic acid (BioUltra, ≥99.5%), glycine hydrochloride (99% HPLC), Phosphoric acid (for HPLC, 85-90%), *N*-(3-Dimethylaminopropyl)-*N'*-ethylcarbodiimide hydrochloride EDC (≥99.0%), *N*-Hydroxysulfosuccinimide sodium salt Sulfo-NHS (≥98% HPLC), hydrochloric acid HCl (ACS reagent, 37%), chloroform (ReagentPlus[®], ≥99.8%, contains 0.5-1.0% ethanol as stabilizer), methanol (CHROMASOLV[®], for HPLC, ≥99.9%), MES hydrate (≥99.5%), sodium dodecyl sulfate solution SDS (BioUltra, for molecular biology, 10% in H₂O), sodium deoxycholate SDC (BioXtra, ≥98.0%), 5-Cyclohexylpentyl β-D-maltoside CYMAL 5 (≥98.0%), urea (ACS reagent, 99.0-100.5%), 2,5-Dihydroxybenzoic acid DHB, LC-MS grade acetonitrile (ACN) (LC-MS CHROMASOLV[®]), LC-MS grade ammonium bicarbonate (ABC) (Eluent additive for LC-MS), iodacetamide (BioUltra), NuPAGE[®] LDS Sample Buffer (4X), NuPAGE[®] Sample Reducing Agent containing 500mM DTT (Dithiothreitol), NuPAGE[®] MOPS SDS Running Buffer, SimplyBlue[™] SafeStain, iodoacetamide (BioUltra), LC-MS grade ammonium bicarbonate (ABC) (Eluent additive for LC-MS), LC-MS grade acetonitrile ACN (LC-MS CHROMASOLV[®]), formic acid FA (for mass spectrometry ~98%), trifluoroacetic acid TFA (CHROMASOLV[®], for HPLC, ≥99.0%), acetone (for HPLC, ≥99.9%), ¹⁸O Water (99 atom % ¹⁸O) was purchased from Cambridge Isotope Laboratories, and α-Cyano-4-hydroxycinnamic acid)

7.2. Proteases

Trypsin Gold mass spectrometry grade (#V5280) elastase (#V1891) from Promega, thermolysin from R&D Systems (#3097-ZN-020) and were used.

7.3. Protein and Peptides

The standard peptides used were [Glu1]-Fibrinopeptide B (Glu-Fib, EGVNDNEEGFF-SAR) from Sigma-Aldrich (St. Louis, USA), synthetic peptide TWpYV, synthetic peptide LQKpSPGPQRER, synthetic peptide FQSRGpSpSPLQLN and a customized synthetic peptide SIINFEKL from JPT Peptide Technologies. PKA and B-Casein from Sigma-Aldrich (St Louis, USA) (Sequences in Appendix A).

7.4. Sample Preparation

7.4.1. Peptides

In the early tests, peptides were modified using 4-ABA to test the completion of the reaction and the generated byproducts.

1pmol/ μ l aliquotes of the synthetic peptides SIINFEKL, TWpYV, LQKpSPGPQRER and synthetic FQSRGpSpSPLQLN were desalted with C-18 STAGE tips to remove any salts that might interfere with the reaction. Each C-18 STAGE tip was prewashed with 1x 20 μ l MeOH and centrifuged for 2 min. at 2000-4000 g. 1x 20 μ l of the elution buffer (buffer (B)) containing 80% ACN 0.5% FA was added to each C-18 STAGE tip and centrifuged for 2 min. at 2000-4000 g. Each C-18 STAGE tip was then conditioned with 2x 25 μ l of the washing buffer (buffer (A*)) containing 2% ACN 0.3% TFA and centrifuged for 2 min. at 2000-4000 g. Samples were acidified with 0.1% FA before they were loaded into the washed and conditioned C-18 STAGE tips and centrifuged for 10 min. at 2000 g. Tips were then washed 3 times with 25 μ l buffer (A*) and centrifuged for 2 min. at 2000-4000 g to assure that all chemical and salts were eliminated from the sample. In the final step, the samples were collected in clean eppendorf tube by eluting them from the C-18 STAGE tips with 2x 20 μ l buffer (B) using a syringe to push the sample through the C-18 material and subsequently dried down in SpeedVac.

To each dried aliquot, 15 μ l of 1.25 M EDC of the freshly prepared EDC in 1 M MES buffer (pH 6) was added followed by 25 μ l of 3 M 4-ABA in 1M MES buffer and 6 μ l concentrated HCl. The final pH of the sample should be between 5 and 5.5 to assure that the aromatic amine of the 4-ABA is reacting with the activated carboxyl and phosphate groups. Samples are incubated at 25°C for 4 hours while shaking at 550 rpm on a thermomixer. Labeled samples are then cleaned with C₁₈ STAGE Tips as described above. The dried samples are redissolved in 10 μ l 1% FA prior LC-MS/MS analysis. For MALDI analysis, acidified SIINFEKL is mixed 1:1 with CHCA matrix and the acidified phosphopeptide TWpYV is mixed 1:1 with DHB/PA matrix.

7.4.2. Proteins

7.4.2.1. 4-ABA In-gel Chemical Modification and In-gel Digestion

Gel Electrophoresis: NuPAGE® Novex® Pre-Cast 4-12% Bis-Tris gels 1.0 mm x 10 Well from Invitrogen were used. The sample preparation was carried out according to the workflow proposed by Invitrogen. The PKA and β -Casein samples were resuspended in LDS sample buffer (10% glycerol, 141 mM Tris Base, 106 mM Tris HCl, 2% LDS, 0.51 mM EDTA, 0.22 mM SERVA® Blue G250, 0.175 mM Phenol Red, pH 8.5). Subsequently, reduced with 50 mM DTT at 70 °C for 10 minutes. Samples were cooled at room temperature before they were alkylated with 120 mM iodoacetamide and incubated in the dark for 20 minutes at room temperature. A sample volume of ~23 μ l (12 pmol) was added per Well. NuPAGE® MOPS SDS running buffer (50mM MOPS, 50mM Tris, 1 mM EDTA, 0.1% SDS, pH 7.7) was used to separate the proteins. PKA and β -Casein proteins were separated by SDS-PAGE for 40 min. at 200V. Gels were washed three times for 5 minutes with ultrapure water before they are stained for 45 minutes with Simply Blue™ Safe Stain (Life Technologies). The gels then were washed with ultrapure water for approximately 2 hours.

4-ABA In-gel Chemical Modification: Gel regions containing PKA or β -Casein were excised from the gel. Gel slices were chopped and destained with 70 % acetonitrile in 100 mM NH_4HCO_3 (pH 8), shrunk with 100 % acetonitrile and dried in a vacuum concentrator (Concentrator 5301, Eppendorf) separately.

A mixture of 3 M 4-ABA in 1 M MES buffer (pH 6) and concentrated HCl was prepared; the final pH of this mixture is ranging between 5.3 and 5.8. A mixture of 1.25 M EDC and 20 mM Sulfo-NHS, both were prepared with 1 M MES buffer (pH 6). To each dried gel band, 25 μ l of the EDC/Sulfo-NHS was added, once the gel is swollen and this mixture diffused in the gel 50 μ l of 4-ABA/MES/HCl mixture is added, the final pH should be checked and adjusted if it's outside the 5-5.5 range. Modified samples and control samples, to which 40 μ l MES buffer were added, were incubated at 25°C for 3 hours while shaking at 550 rpm on a thermomixer. Every 1 hour, the reagents were removed from the gel and replaced with freshly prepared reagents.

At the end of the incubation period, 200 μ l of 500 mM ABC buffer was added to each sample and shook vigorously for 30 min.. After removing the supernatant, 200 μ l of 100% ACN was added, samples shook vigorously for 5 min, supernatants were removed and the gel pieces were dried in SpeedVac. The samples were washed with 250 mM and 100 mM ABC buffer, respectively, in the same manner. Gel pieces were dried in a vacuum concentrator prior in-gel digestion.

7. Material and Methods

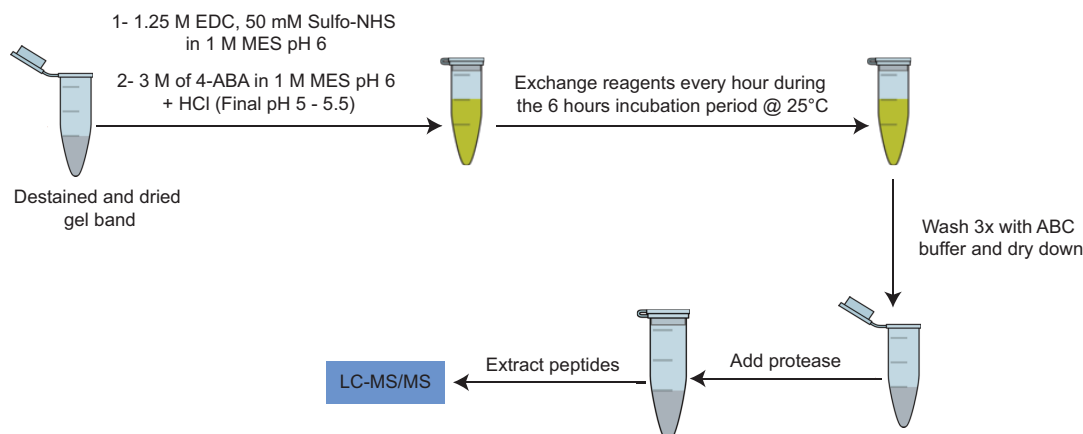


Figure 7.1.: Workflow for the 4-ABA in-gel chemical modification.

7.4.2.2. 4-ABA In-solution Chemical Modification and Digestion

Acetone Precipitation: Four volumes of -20°C acetone were added to one volume of the β -Casein samples. Samples were vortexed for 1 minute and incubated overnight at -20°C . Samples were centrifuged for 15 minutes at 15,000 g at 4°C . The supernatant was carefully removed and discarded without poking the pellet. The residual acetone was evaporated from the uncapped tube containing the pellet at room temperature for 30 minutes.

4-ABA In-solution Chemical Modification: A mixture of 3 M 4-ABA in 1 M MES buffer (pH 6) and concentrated HCl was prepared; the final pH of this mixture is ranging between 5.3 and 5.8. A mixture of 125 mM EDC and 50 mM Sulfo-NHS, both were prepared with 1 M MES buffer (pH 6). 15 μl of the EDC/Sulfo-NHS was added followed by 25 μl of 4-ABA/MES/HCl mixture is added, the final pH should be checked and adjusted if it's outside the 5-5.5 range. Modified samples and control samples, to which 40 μl MES buffer were added, were incubated at 25°C for 6 hours while shaking at 550 rpm on a thermomixer. During this incubation period, 10 μl of freshly prepared EDC/Sulfo-NHS mixture was added every 30 min, and 10 μl of freshly prepared 4-ABA/MES/HCl was added every 2 hours and the pH was checked and adjusted. For the control sample, same volumes of MES buffer were added at each step. At the end of the incubation period, the sample was precipitated using chloroform/methanol precipitation.

Chloroform-Methanol Precipitation: Four volumes of -20°C methanol were added to one volume of each protein sample, and the mixture was vortexed for 1 minute. One volume of chloroform was then added, and the mixture was vortexed for 1 minute. The sample was centrifuged at 14,000 g for 5 min at 4°C . The aqueous methanol layer was removed from the top of the sample without disrupting the protein layer found at the

phase boundary between the aqueous methanol layer and the chloroform layer. Four volumes of methanol were added, and the mixture was vortexed 1 minute. The sample was spun at 14,000 g for 15 min at 4°C. The supernatant was removed without disturbing the pellet, and the pellet was air dried before running the gels.

Gel Electrophoresis: NuPAGE® Novex® Pre-Cast 4-12% Bis-Tris gels 1.0 mm x 10 Well from Invitrogen were used. The sample preparation was carried out according to the workflow proposed by Invitrogen. β -Casein samples were resuspended in LDS sample buffer (10% glycerol, 141 mM Tris Base, 106 mM Tris HCl, 2% LDS, 0.51 mM EDTA, 0.22 mM SERVA® Blue G250, 0.175 mM Phenol Red, pH 8.5). Subsequently, reduced with 50 mM DTT at 70 °C for 10 minutes. Samples were cooled at room temperature before they were alkylated with 120 mM iodoacetamide and incubated in the dark for 20 minutes at room temperature.

A sample volume of ~23 μ l (12 pmol) was added per Well. NuPAGE® MOPS SDS running buffer (50mM MOPS, 50mM Tris, 1 mM EDTA, 0.1% SDS, pH 7.7) was used to separate the proteins. β -Casein proteins were separated by SDS-PAGE for 40 min. at 200V. Gels were washed three times for 5 minutes with ultrapure water before they are stained for 45 minutes with Simply Blue™ Safe Stain (Life Technologies). The gels then were washed with ultrapure water for approximately 2 hours.

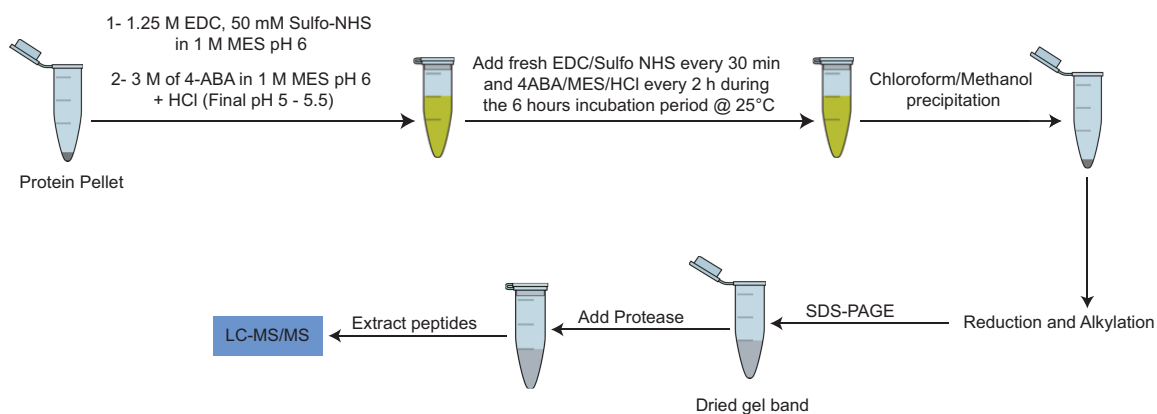


Figure 7.2.: Workflow for the 4-ABA in-solution chemical modification followed by in-gel digestion.

In-gel Digestion: Dried gel pieces of the in-gel modified samples and the in-solution modified samples were suspended in 100 mM NH_4HCO_3 (pH 8) containing 0.1 μ g of a protease. Digests with trypsin for PKA samples and elastase or thermolysin for β -casein samples were performed overnight at 37 °C. Peptides were extracted from the gel slices with 5 % formic acid and transferred to LC vials.

7.5. NanoLC-MS/MS Analysis

NanoLC-MS/MS analyses for unmodified proteins, in-gel modified and in-solution modified proteins were performed on an Orbitrap Fusion[™] (Thermo Scientific) equipped with an EASY-Spray[™] Ion Source and coupled to an EASY-nLC 1000 (Thermo Scientific). Peptides were loaded on a trapping column (2 cm x 75 μ m ID, PepMap C-18 3 μ m particles, 100 Å pore size) and separated on an EASY-Spray[™] column (25 cm x 75 μ m ID, PepMap C-18 2 μ m particles, 100 Å pore size) with either a 45 minute linear gradient from 3 % to 30 % acetonitrile and 0.1 % formic acid. Both MS and MS/MS scans were acquired in the Orbitrap analyzer with a mass range of 350 -1550 m/z and a resolution of 60 000 and 15 000, respectively using HCD fragmentation. HCD fragmentation with 35 % normalized collision energy with was applied. The ETD reaction time was set to 50 ms. A Top Speed data-dependent MS/MS method with a fixed cycle time of 3 seconds was used. Dynamic exclusion was applied with duration of 10 seconds. Minimum signal threshold for precursor selection was set to 50 000, and the width of the precursor selection window was set to 2.5 Da. Predictive AGC was used with AGC a target value of 2×10^5 for MS scans and 5×10^4 MS/MS scans. EASY-IC was used for internal calibration.

7.6. Data Analysis

The MS data of the unmodified and modified SIINFEKL, TWpYV, LQKpSPGPQRER, FQSRGpSpSPLQLN, PKA, β -Casein were analyzed with PEAKS 8 and searched against a custom database containing the peptides and proteins of interest. For data refinement, no merge option was selected, precursor options corrected and no correction of charge options and no filter was selected. For peptides searches, deamidation (NQ), dehydration, phosphorylation (ST), 4-ABA at (DE) $C_7H_8N_2O_{-1}$, 4-ABA (C-term) $C_7H_8N_2O_{-1}$, phospho 4-ABA (STY) $C_7H_9N_2O_{-2}P$. Database searches were performed without protease specificity (enzyme: none), 8 ppm mass tolerance for the precursor, 0.02 Da mass tolerance for fragment ions, 9 max variable PTM per peptide and FDR estimation enabled.

For PKA and β -casein, carboxyamidomethyl (C) was set as a fixed modification. In addition to the variable modifications set for peptide searches, acetylation (Protein N-term) and oxidation (M) were added. Database searches for β -casein were performed without protease specificity (enzyme: none), for PKA without protease specificity (enzyme: none) and with protease specificity (enzyme: trypsin) with 3 maximum missed cleavages allowed, 8 ppm mass tolerance for the precursor, 0.02 Da mass tolerance for fragment ions, 8 max variable PTM per peptide and FDR estimation enabled.

8. Results and Discussion

Protein phosphorylation is critical for many cellular processes. Therefore, the development of mass spectrometry-based methods to identify phosphosites and to quantify phosphorylation provide a valuable tool to decipher these processes. Despite the continuous development of advanced MS methods, LC-MS/MS analysis of multi-phosphorylated peptides in the positive ion mode still faces several challenges. Some of these challenges have been partially addressed and solved in the recent years, but they still represent common drawback that affects the correct localization of phosphorylation site(s), phosphopeptide identification, and characterization, which are critically dependent on the fragmentation behavior of phosphopeptide ions and the properties of the precursor ion.

All LC-MS-related challenges originate mainly from the negative charge of the phosphate. Phosphate increases the acidity of phosphopeptides compared to their unphosphorylated cognates, and hence decreases the protonated molecules under the conditions typically used for the positive ion mode LC-MS/MS. Therefore, the ionization efficiency, which depends largely on the peptides basicity, of phosphorylated peptides is lower compared their unphosphorylated cognates [165]. Additionally, phosphopeptides form complexes with metals such as Fe (III) on the C₁₈ column, which causes a strong retention of phosphopeptides on the C₁₈ column. Thus, the amount of phosphopeptides eluted from the LC system may be less than the amount injected, especially for multi-phosphorylated peptides [165]. The low ionization efficiencies of multi-phosphorylated peptides in positive mode and the labile nature of phosphorylation during CID/HCD fragmentation lead to poor quality of CID/HCD spectra [166] and consequently, incorrect localization of phosphorylation site(s) and impaired identification of protein/peptide sequences (Fig. 8.1). ETD fragmentation can identify the CID/HCD-labile phosphorylation; however, ETD is efficient for higher-charged peptides, typically $\geq 3+$ [129]. The presence of multi-phosphorylated sites within a peptide decreases its net charge and thus makes it unsuitable for ETD fragmentation. Furthermore, during ETD salt bridges might be formed between the phosphate group and the basic amino acid side chains, which prevents the separation of fragments following backbone cleavage [167].

All the above mentioned issues make the analysis of multi-phosphorylated peptides complicated and troublesome. The analysis of multi-phosphorylated peptides with clusters of phosphosite(s) and the localization of these phosphosites are crucial to deciphering the

8. Results and Discussion

biological functions of the sophisticated phosphoproteome dynamics. Good examples for proteins with such clusters are the PER2, which is key regulator involved in the resetting of the circadian clock [168] and B-RAF, which plays an important role in various developmental processes [169].

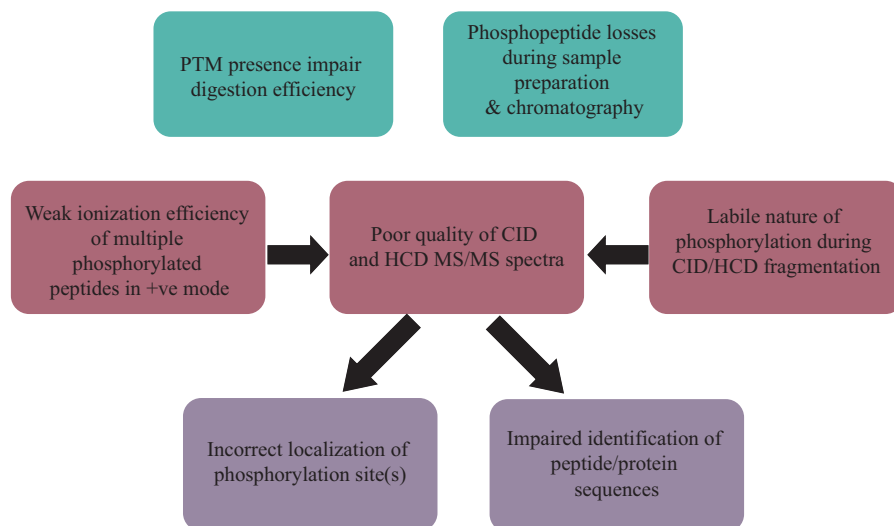


Figure 8.1.: Challenges in LC-MS analysis of phosphorylation.

We aimed to develop a method that overcomes the MS-related challenges in the multi-phosphorylation analysis, where we used the 4-aminobenzelamine (4-ABA) to reverse the charge of the phosphate(s) and increase the net charge of the modified phosphopeptide. 4-ABA has two amino groups, both of which could potentially be reactive to the carbodiimide-activated peptidyl phosphates. We performed the carbodiimide-mediated phosphoramidation reaction under slightly acidic conditions (pH 5.5). At this pH, the aromatic amine (pKa \sim 5) is largely unprotonated and, thus, is reactive as a nucleophile and forms a stable amide bond with the peptidyl phosphates. In contrast, the aliphatic amine (pKa \sim 10) is protonated at the same pH, preventing the amine group from being a nucleophile.

Using 4-ABA to modify the peptidyl phosphates under the conditions used generates aromatic phosphoramidates that are highly stable against hydrolysis. High hydrolytic stability of aromatic phosphoramidates is essential to preserve the derivatization through the sample preparation and separation of the modified phosphopeptides before LC-MS/MS analysis. Furthermore, 4-ABA labeling of the negatively charged phosphates reverse the charge and increase the net charge of the phosphorylated peptides. This modification makes the phosphopeptides suitable for ETD fragmentation and facilitates the correct localization of phosphosite(s). Moreover, 4-ABA balances the hydrophobicity of the multi-phosphorylated peptides due to the attachment of additional positive charge state.

The general concept of the 4-ABA labeling of phosphate and carboxyl groups, which is

based on the principle of phosphoramidate chemistry (PAC), is shown in Fig. 8.2. First, the residues of interest are activated by the water-soluble carbodiimide EDC forming an unstable active-ester intermediate. This highly reactive species can either react under slightly acidic conditions with the aromatic primary amine of the 4-ABA forming a stable amide bond, or it can be hydrolyzed by water. The use of EDC/Sulfo-NHS mixture instead of using only EDC accelerates the reaction of the O-acylisourea intermediate with the amine and, thus lowers the probability of hydrolysis. As a result of this reaction, the activated residues and the amine groups form a stable amide bond.

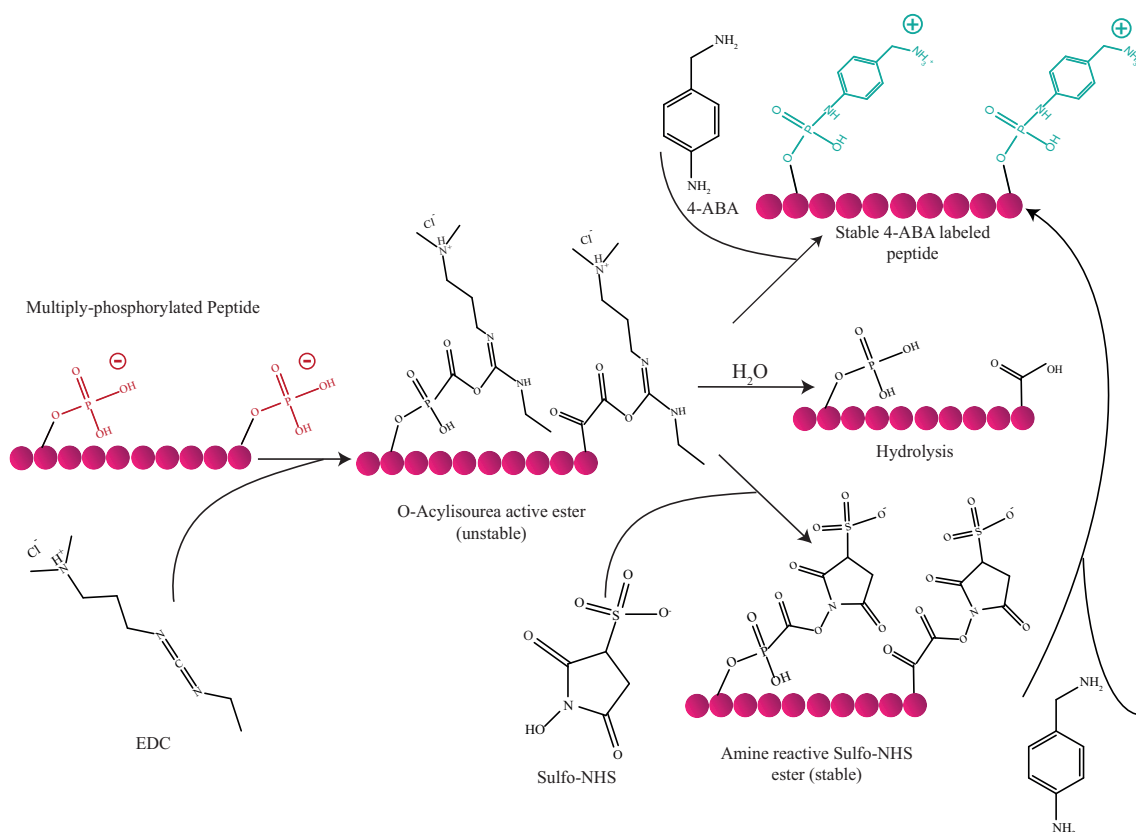
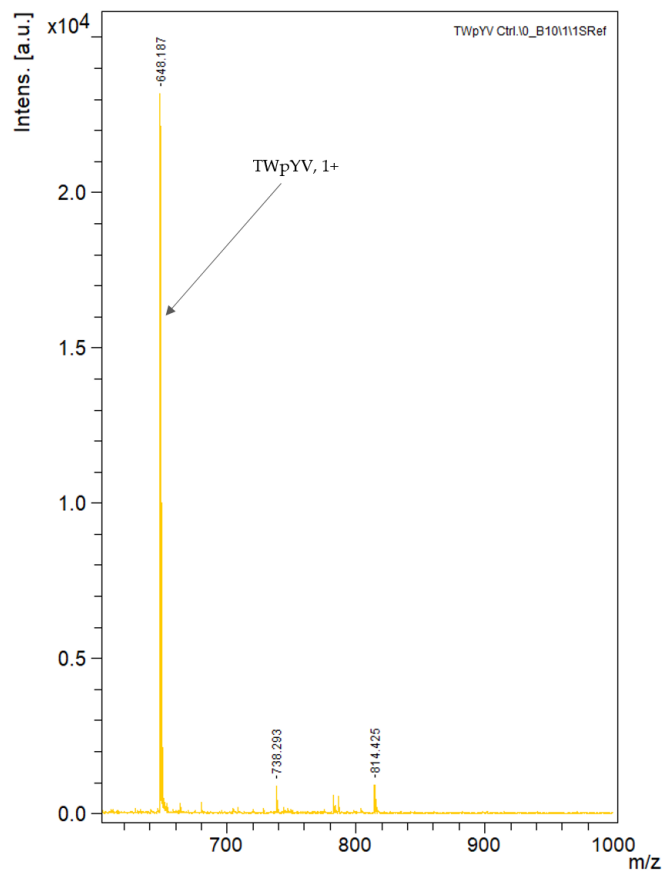


Figure 8.2.: Principle of the carbodiimide-mediated 4-ABA labeling of phosphate groups and carboxyl groups at peptide or protein level. 4-ABA has two amino groups, when the carbodiimide-mediated phosphoramidation reaction is performed under slightly acidic conditions (pH 5.5), the aromatic amino group of the 4-ABA is largely unprotonated and, thus, is reactive as a nucleophile and forms a stable amide bond with the phosphate. At the same pH, the aliphatic amine is protonated, which prevents the amine group from being a nucleophile. Figure adapted from [170]

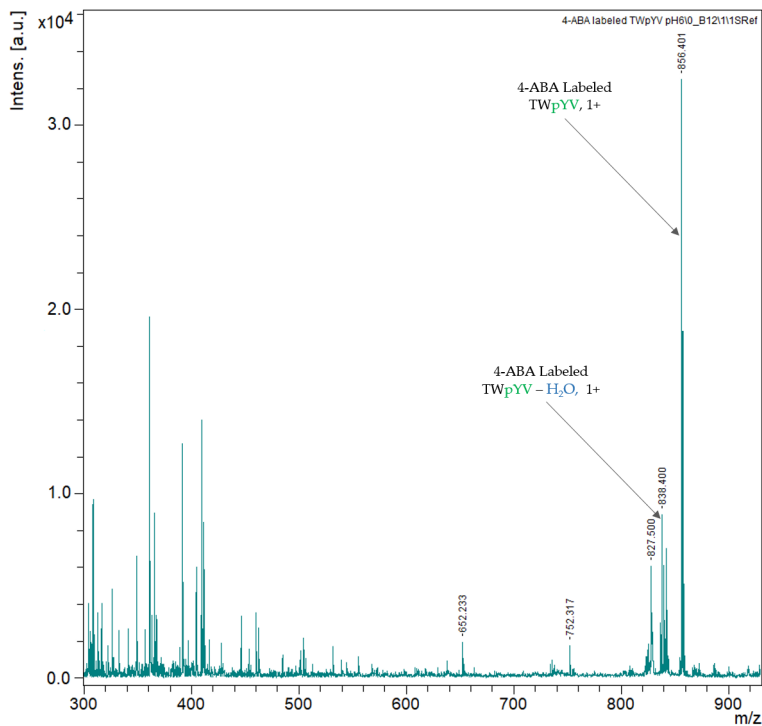
8.1. Proof of Principle

In the early stages of method development, we used the synthetic peptides TWpYV ($m/z = 648.2429$) to test our hypothesis and the reaction completion. By testing different mixtures of reagents, we found the best reaction conditions with which we can achieve

8. Results and Discussion



(a) TWpYV Control.



(b) 4-ABA Labeled TWpYV

Figure 8.3.: Experimental MALDI results of TWpYV before and after 4-ABA modification. **a)** Control TWpYV, 1+ ($m/z = 648.187$). **b)** Completely 4-ABA labeled TWpYV, 1+ ($m/z = 856.401$) and completely 4-ABA labeled TWpYV in addition to water loss at $m/z = 838.400$.

a complete labeling is activating the residues of interest with 1.25 M EDC in 1 M MES buffer (pH 6). A mixture of 3 M 4-ABA in 1 M MES buffer (pH 6) and concentrated HCl was added directly after the EDC was added to the dry peptide to be tested. Samples were incubated at 25 °C while shaking for 4 hours (Fig. 8.3).

8.2. Method Validation

Carbodiimide-mediated 4-ABA labeling of phosphosites can be performed at the peptide level as well as at the intact protein level depending on the application.

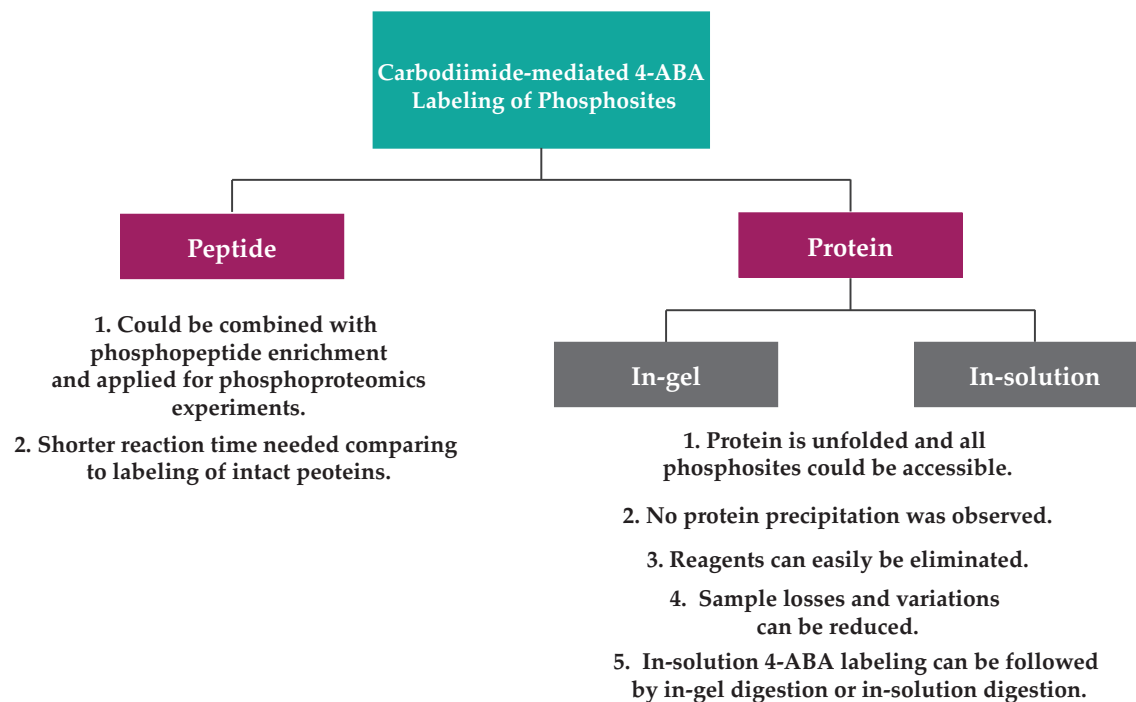


Figure 8.4.: Different approaches for the carbodiimide-mediated 4-ABA labeling.

8.2.1. Labeling on the Peptide Level

We used the synthetic peptides LQKpSPGPQRER and FQSRGpSpSPLQLN to test the reaction completion, the generation of side products (Fig 8.5), and to compare the ionization efficiencies of modified and unmodified peptide. We tested different mixtures of reagents to achieve complete labeling with minimal generation of byproducts (appendix). The best reaction condition we found that meets all the requirements was activating the residues of interest with a mixture of 1.25 M EDC and 50 mM Sulfo-NHS both in 1 M MES buffer (pH 6). A mixture of 3 M 4-ABA in 1 M MES buffer (pH) and concentrated HCl was added directly after the EDC/Sulfo-NHS mixture was added to the dry peptide

8. Results and Discussion

to be tested. Samples were incubated at 25 °C while shaking for 4 hours. No peptide precipitation was observed by using these concentrations of EDC and Sulfo-NHS.

Modifying the phosphopeptides with 4-ABA not only modifies the phosphate groups but also modifies the carboxyl groups. This modification increases the hydrophilicity of the peptide and causes a mass shift of $184.0402 n + 104.0738 m$; where n is the number of phosphosites and m is the number of carboxyl groups within a peptide. No major side products were detected; the only side reaction found was the deamidation of Asn and Gln (Fig. 8.5). The peptide charge states observed increases substantially upon derivatizing the peptides with 4-ABA at their phosphate groups and carboxyl groups, which makes them suitable for ETD fragmentation. We did not detect any noticeable difference in the signal intensity between the 4-ABA modified samples and unmodified samples.

The peptide derivatization reaction employed proceeds nearly to completion, thereby minimizing any increase in sample complexity that would result from multiple incomplete reaction products for each peptide. The results show that the high reaction efficiency for the singly phosphorylated peptide and the doubly phosphorylated peptide appears to be independent of the number of the sites to be modified, and their position within the peptide.

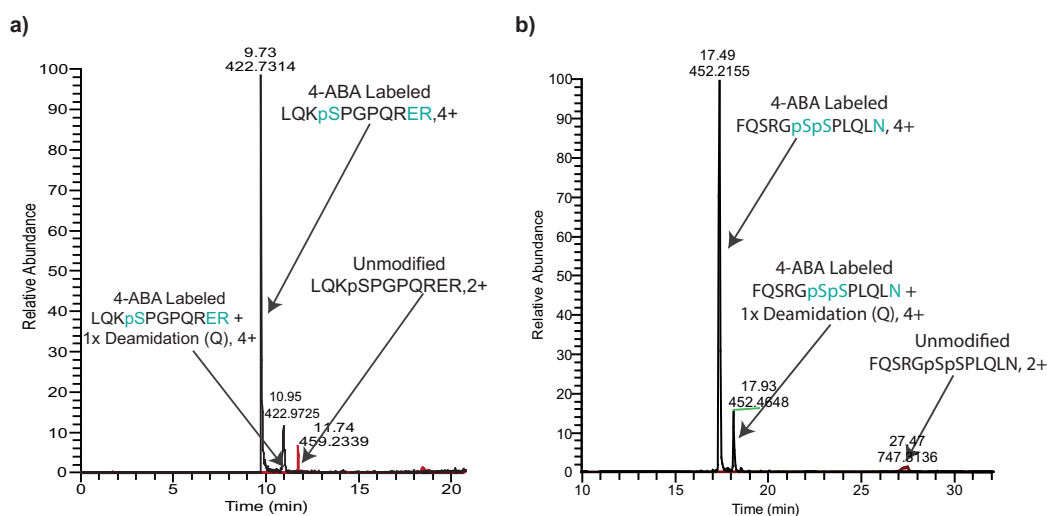


Figure 8.5.: Reaction completion of singly and doubly phosphorylated peptides modified with 4-ABA. **a)** XICs of 4-ABA modified LQKpSPGPQRER, *black*: completely modified LQKpSPGPQRER, 4+ ($m/z = 422.7314$, RT = 9.73) and deamidated 4-ABA modified LQKpSPGPQRER, 4+ ($m/z = 422.9725$, RT = 10.95), *red*: unmodified LQKpSPGPQRER, 2+ ($m/z = 459.2339$, RT = 11.74). **b)** XICs of 4-ABA modified FQSRGpSpSPLQLN, *black*: completely modified FQSRGpSpSPLQLN, 4+ ($m/z = 452.2155$, RT = 17.49) and deamidated 4-ABA modified FQSRGpSpSPLQLN, 4+ ($m/z = 452.4648$, RT = 17.93), *red*: unmodified FQSRGpSpSPLQLN, 2+ ($m/z = 747.3136$, RT = 27.47). Both modified and unmodified peptides were detected within the same sample.

8.2.2. Labeling on the Protein Level

To validate the robustness of our new approach and its applicability on the intact protein we needed to ensure a complete modification of pSer, pThr, pTyr, Glu, Asp, and the protein C-terminal have been achieved. The completeness of the labeling reaction is a prerequisite for pinpointing the correct site(s) of phosphorylation. The presence of unmodified or partially modified peptides within a sample leads to the existence of chemical entities with different retention times and masses than the fully-modified peptide, which will complicate the data analysis.

To validate the applicability of the 4-ABA modification on the protein level, four main factors have been taken into account: i) reaction completion, ii) sequence coverage, iii) fragmentation, and iv) overall intensity. For this, we used the commercially available, recombinant protein kinase A (PKA) catalytic subunit as a model phosphoprotein. The recombinant version is known to have four main phosphorylation sites [171]. As a test case for a protein with a cluster of phosphorylation sites, we used the commercially available β -casein. To avoid any variations that may occur when performing the modification after the enzymatic digestion, we modified the intact PKA in-gel as described in the method subsection 7.4.2.1 and later the intact β -casein in-solution as described in the method subsection 7.4.2.2. These variations are usually caused by losing some peptides during sample preparation, specifically hydrophobic peptides, due to their adsorption on the surfaces of pipette tips and Eppendorf tubes. Modifying the intact protein in-solution using 1.25 M EDC to activate the residues to be modified caused the protein to cross-link, which made the modified protein stuck in the well of the SDS-PAGE gel and do not migrate during the gel electrophoresis. Therefore, performing the modification in-gel represents an alternative to the in-solution modification.

8.2.2.1. Reaction Completion and Sequence Coverage

The preliminary results of applying the 4-ABA modification on the intact protein in-gel were promising. However, a complete labeling of all sites that could be modified and a good sequence coverage of the PKA have not been achieved (Fig. 8.6). To evaluate the reaction completion when modifying the intact protein in-gel, we compared the peak intensities of modified, partially modified, and unmodified peptides within the same sample (Fig. 8.7).

To assure that the reaction went to completion, we increased the reaction time to 8 hours instead of 6 hours while adding freshly prepared reagents as mentioned in subsection 7.4.2.1. Despite increasing the reaction time, some residues remained completely unmodified. This probably caused by "oily-like" nature of the 4-ABA, which hinders the diffusion of 4-ABA into the gel pieces and make it difficult to reach some residues that are difficult to be accessible (Fig. 8.8).

8. Results and Discussion

To circumvent the issues with the in-gel chemical modification, we tried again to modify the intact protein in-solution. In order to avoid or to reduce the chances of cross-linking when modifying the protein in-solution, we lowered the concentration of EDC. We tested different concentrations of EDC to activate the desired residues while using the same concentrations of all other reagents (see section 7.4.2.2). We found that using 125 mM instead of 1.25 M do not cause precipitation of protein during the incubation and do not cause the protein to aggregate in the well of SDS-PAGE gel. After performing two-steps

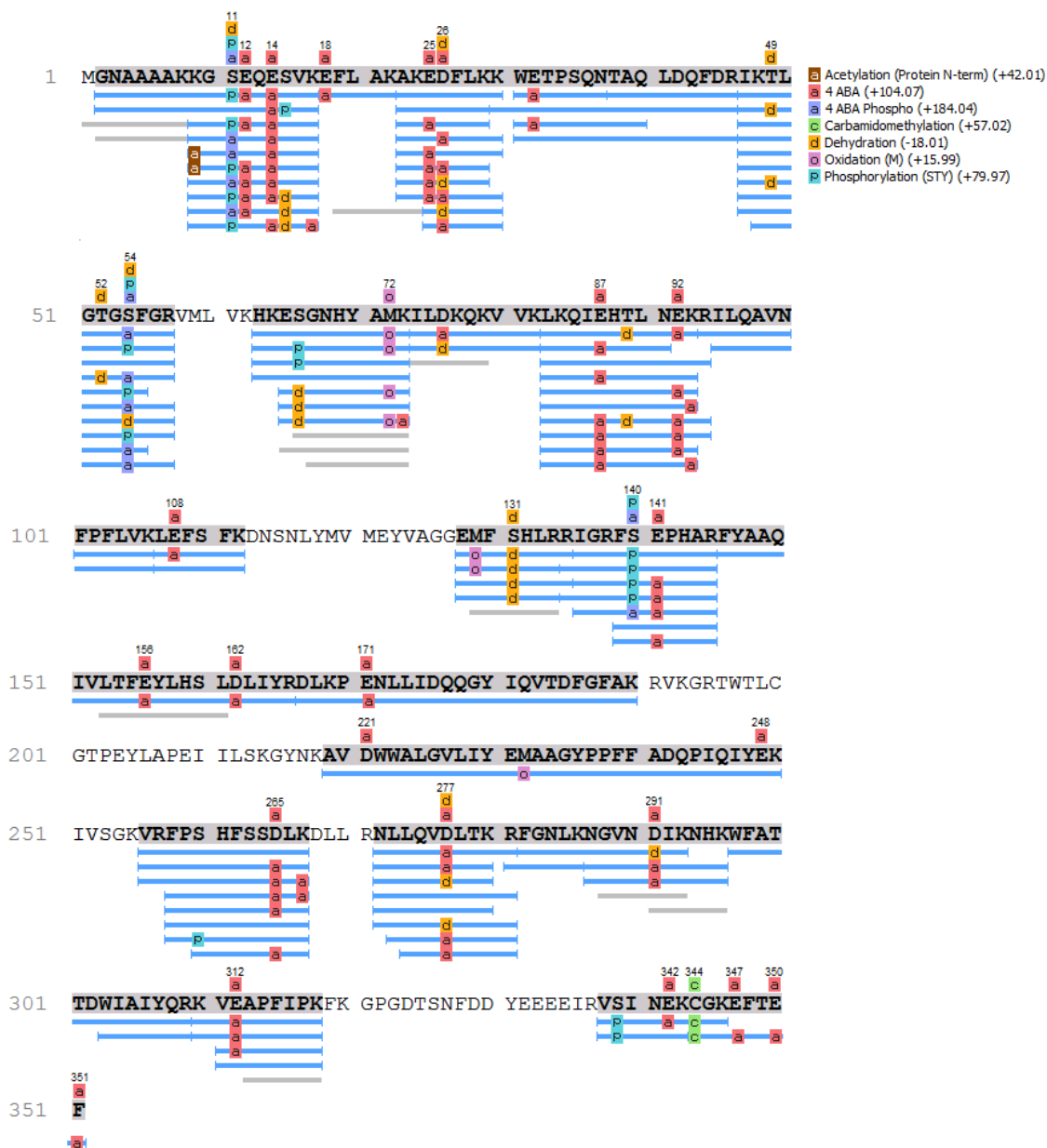


Figure 8.6.: Sequence coverage of tryptic digest of PKA chemically modified in-gel with 4-ABA. The type of modification is shown on the top of the modified residue. Performing the modification in-gel resulted in incomplete labeling of many residues.

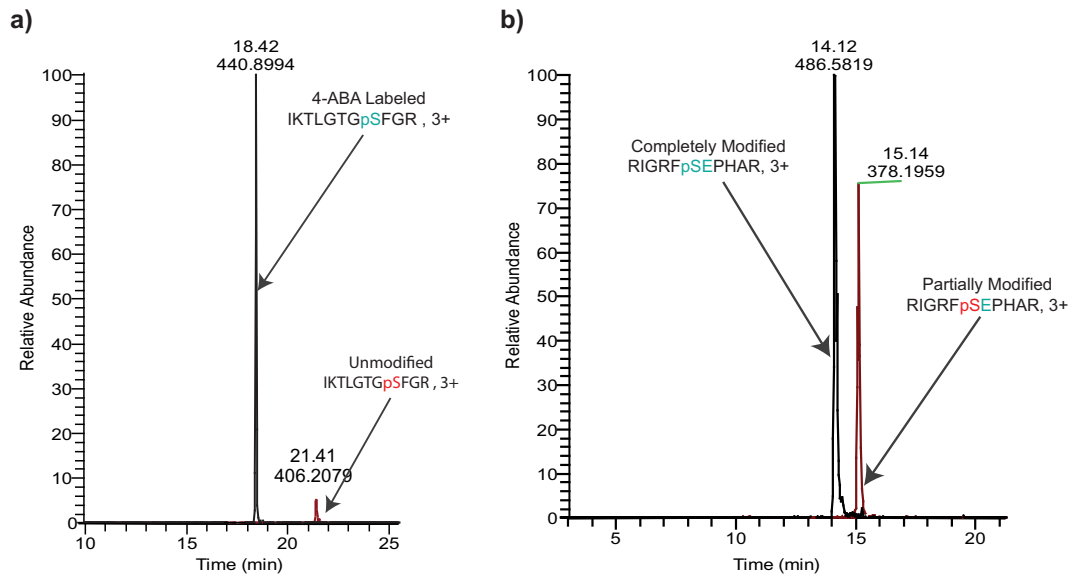


Figure 8.7.: Reaction completion of tryptic digest of PKA modified in-gel with 4-ABA: **a)** XIC of IKTLGTGpSFGR, 3+ (black) completely modified peptide $m/z = 440.8995$, RT = 18.42 and (red) unmodified peptide $m/z = 406.2075$, RT = 21.41 within the same sample **b)** XIC of RIGRFpSEPHAR, 3+ (black) completely modified peptide $m/z = 486.5819$, RT = 14.12 and (red) partially modified peptide $m/z = 378.1959$, RT = 15.14 within the same sample.

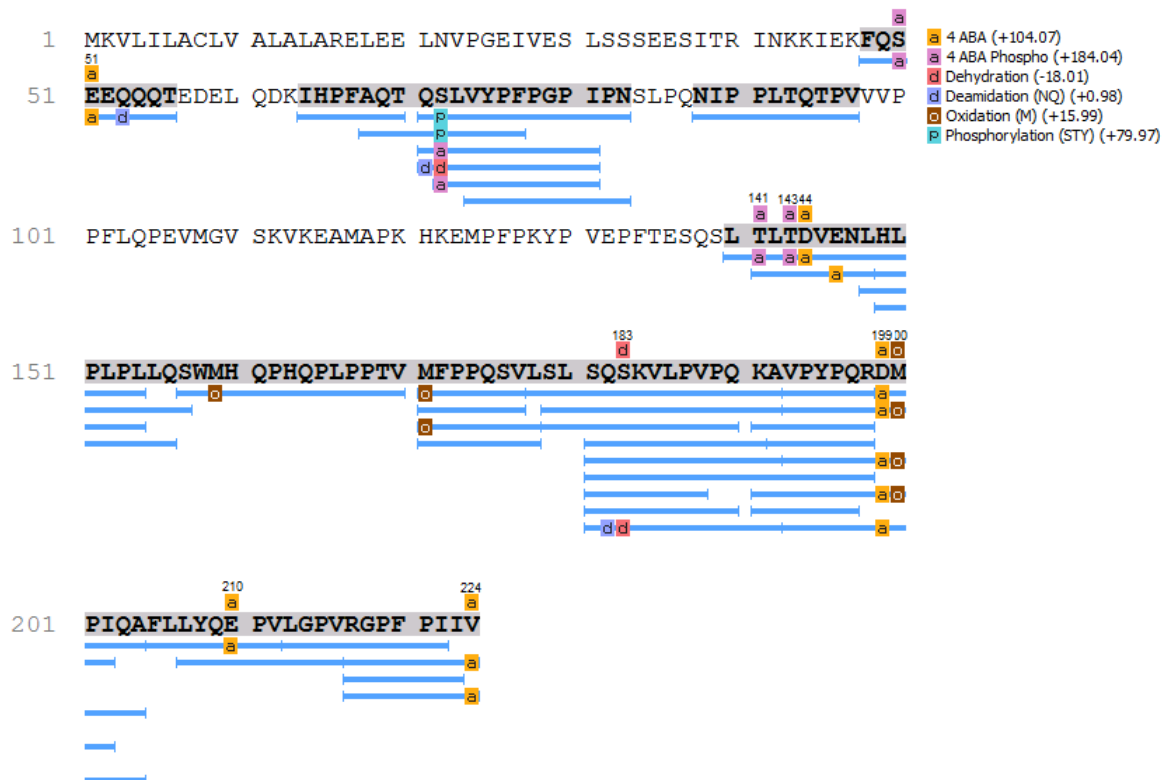


Figure 8.8.: Sequence coverage of thermolysin digest of β -casein chemically modified in-gel with 4-ABA. The type of modification is shown on the top of the modified residue.

8. Results and Discussion

protein precipitation at the end of the reaction to ensure complete removal of the reagents, we performed an in-solution proteolytic digestion. The in-solution digestion was hampered by the residual 4-ABA that was still present after the chloroform-methanol precipitation followed by 2x acetone precipitation. Therefore, in-gel digestion was a better choice than in-solution digestion since the residual 4-ABA can be removed during the gel electrophoresis (C)

The staining of the SDS-PAGE gel showed a noticeable reduction in the protein concentration (Fig. 8.9), although both, control and modified samples, had the same initial concentration and incubated under the same conditions as the modified samples. Performing the 4-ABA modification in the presence of SDS resulted in a further reduction in the protein intensity. This reduction is caused by the cross-linking promoted by the presence of SDS.

Results showed that applying the modification in-solution under the conditions mentioned in section 7.4.2.2, resulted in more labeled residues and better sequence coverage than the in-gel chemical modification (Fig. 8.8 and 8.10).

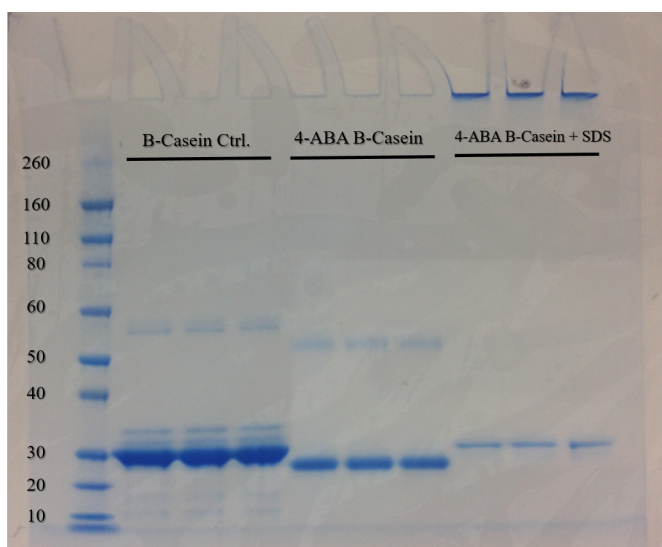


Figure 8.9.: SDS-PAGE gel of β -casein before and after 4-ABA treatment. Picture shows the difference intensities after acetone precipitation of control samples and 4-ABA modified samples. For the three gel bands on the right, the protein pellet was resuspended in 0.1 %SDS before the 4-ABA treatment.

β -Casein has four phosphoserines in close proximity. The identification and the pinpointing of phosphorylation of these sites are challenging in the common positive ion mode. The presence of multi-phosphorylation sites within a peptide decreases the peptide's protonation, which results in a low ionization efficiencies in the positive ion mode [166]. A complete labeling of these four phosphoserines has not been achieved under the conditions applied. We are not certain if this incomplete labeling is due to an incomplete reaction

or due to the structure of the β -casein that makes these sites inaccessible to 4-ABA. However, the increased positive charge resulting from the carboxyl residues labeling of the tetraphosphorylated peptides compensate for the net charge. This improves the ionization efficiency of the tetraphosphorylated peptide comparing to its counterpart from the control sample.

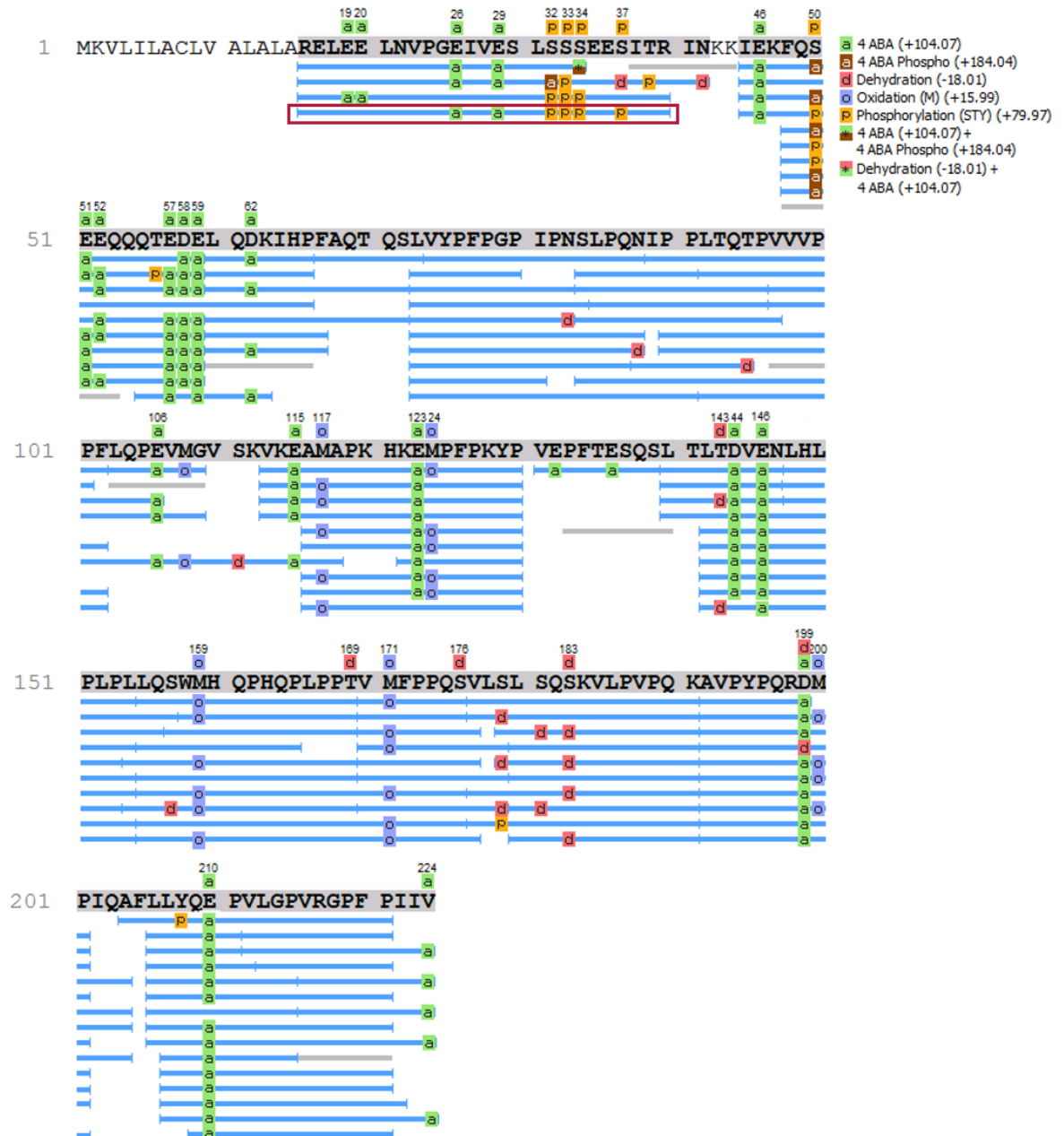


Figure 8.10.: Sequence coverage of thermolysin digest of β -casein chemically modified in-solution with 4-ABA and digested in-gel. Labeling the protein in-solution resulted in better labeling and better sequence coverage than the in-gel labeling showed in Fig 8.8. The type of modification is shown on the top of the modified residue.

8. Results and Discussion

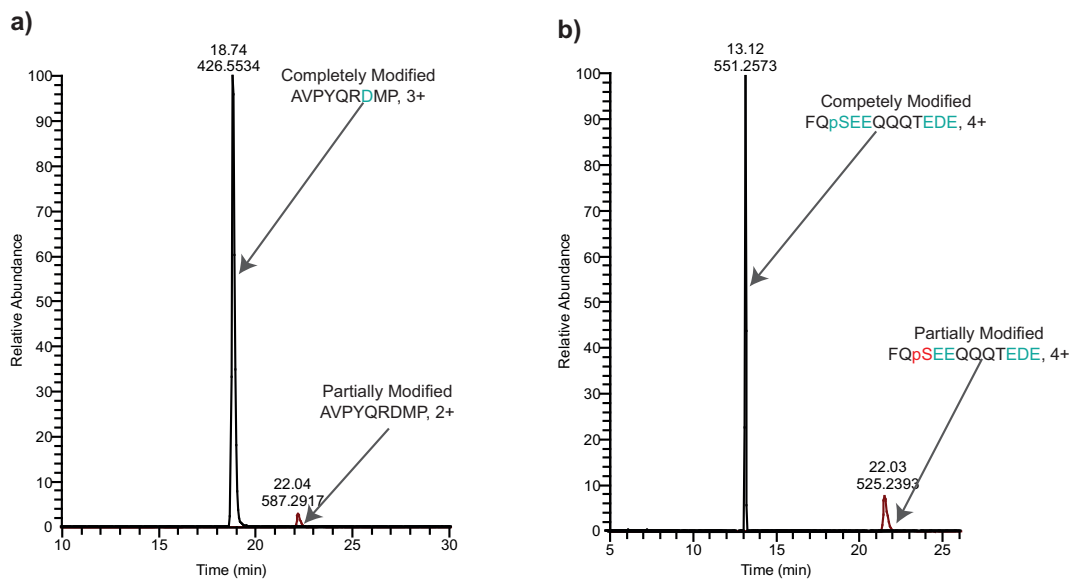


Figure 8.11.: Reaction completion of thermolysin digest of β -casein modified in-solution with 4-ABA: **a)** XIC of AVPYQRDMP (black) completely modified AVPYQRDMP, 3+ ($m/z = 426.5533$, RT = 18.72) and (red) unmodified AVPYQRDMP, 2+ ($m/z = 406.2075$, RT = 21.41) **b)** XIC of FQpSEEQQQTEDE, 4+ (black) completely modified peptide $m/z = 551.2573$, RT = 13.12 and (red) partially modified peptide $m/z = 525.2329$, RT = 22.03.

8.2.2.2. The Influence of 4-ABA Modification on the Overall Intensity

We wanted to test if modifying the intact protein with 4-ABA has an influence in changing the overall peptide intensity. For this, we compared the total ion chromatogram (TIC) of a control sample and a 4-ABA modified sample; both samples have the same initial protein concentration and volume. We hypothesized that modifying the protein with 4-ABA will increase the overall intensity.

For the in-gel modified sample, the overall intensity of the modified sample is reduced by approximately the half (Fig. 8.12 a), comparing to the control sample, which contradicts our proposed hypothesis. However, comparing the peaks intensities of completely modified peptides and their unmodified counterparts in both samples showed that the intensities of modified peptides are higher than unmodified peptides (Fig. 8.12 b and 8.12 c), which support our initial assumption. For peptides that do not contain any site to be modified (acidic residue or phosphosite), the peptide intensity did not change. (Fig. 8.12 d). The observed differences in the TIC intensities of the unmodified and in-gel modified samples is probably due to the incomplete labeling.

The TIC intensity of in-solution 4-ABA modified β -casein has not been affected by the modification (Fig. 8.13 a). However, comparing the peaks intensities of completely modified peptides and their unmodified counterparts in both samples showed that the intensities of modified peptides are substantially higher than unmodified peptides (Fig. 8.13 b).

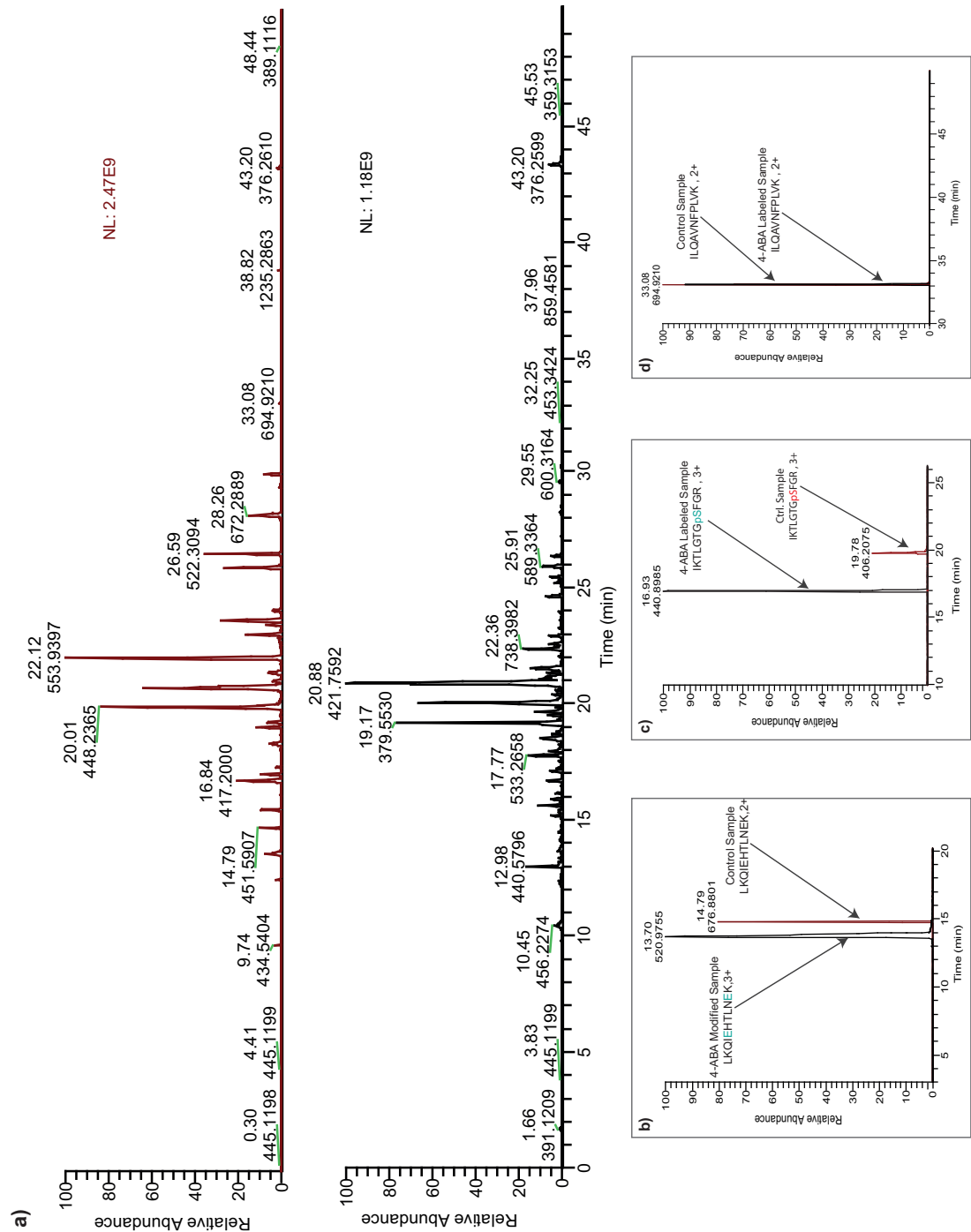


Figure 8.12.: Comparison between the peak intensities of a PKA tryptic digests of an in-gel 4-ABA modified sample and a control sample. **a)** TIC of tryptic digests of (red) unmodified PKA and (black) 4-ABA in-gel modified PKA. **b)** XIC of LKQIEHTLNEK (black) 4-ABA modified sample LKQIEHTLNEK, 3+ ($m/z = 520.9717$) and (red) control sample LKQIEHTLNEK, 2+ ($m/z = 676.8801$). **c)** XIC of IKTLGTGpSFGR, 3+ (black) 4-ABA modified sample $m/z = 440.8995$ and (red) control sample $m/z = 406.2075$. **d)** XIC of ILQAVNFPFLVK, 2+ ($m/z = 694.9175$) (black) 4-ABA modified sample and (red) control sample.

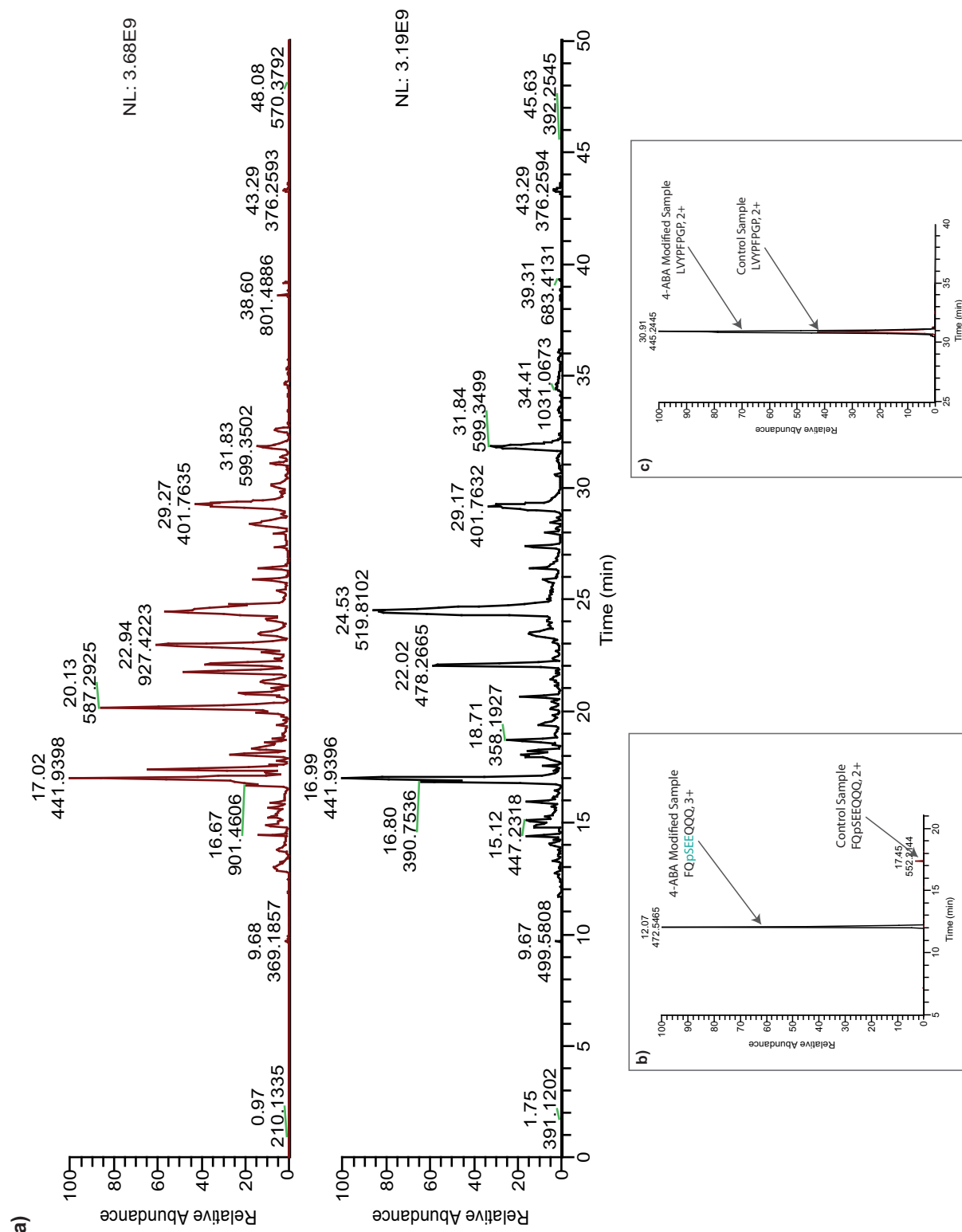


Figure 8.13.: Comparison between the peak intensities of a β -casein thermolysin digests of an in-solution 4-ABA modified sample and a control sample. **a)** TIC of thermolysin digests of (red) unmodified β -casein and (black) 4-ABA in-solution modified β -casein. **b)** XIC of FQpSEEQQQ, 3+ (black) 4-ABA modified sample FQpSEEQQQ, 3+ ($m/z = 472.5465$) and (red) control sample FQpSEEQQQ, 2+ ($m/z = 512.2225$). **c)** XIC of LVYPPFGP, 2+ (black) 4-ABA modified sample and (red) control sample.

By comparing the intensities of peptides that do not contain any phosphate group or carboxyl, the intensity of the peptide from the modified sample was roughly double the intensity of the one from the unmodified sample (Fig. 8.13 c). We could not verify the reason for the decrease in the intensity of these peptides.

8.2.2.3. Fragmentation and Localization of Phosphorylation Site

Localization of phosphorylation sites is crucial to understanding the dynamics of the cellular events. The determination of the exact site of phosphorylation is more challenging than the identification of the phosphopeptide. Site-determining fragment ions are required to localize the site(s) of phosphorylation with high confidence. During CID/HCD, phosphopeptides exhibit neutral losses that can be used to identify the phosphorylation

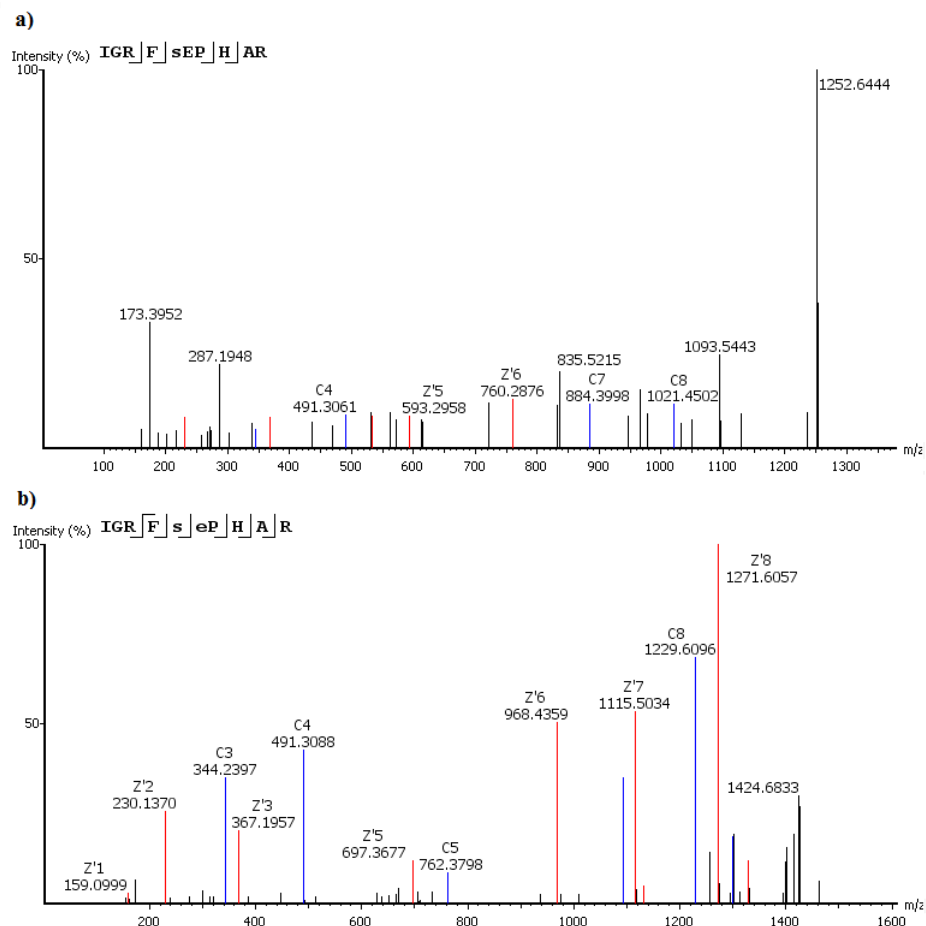


Figure 8.14.: MS/MS spectra of phosphorylated IGRFpSEPHAR from PKA before and after 4-ABA modification. **a)** ETD fragmentation of unmodified IGRFpSEPHAR, 3+ ($m/z = 417.1985$) **b)** ETD fragmentation of 4-ABA modified IGRFpSEPHAR, 3+ ($m/z = 486.5813$).

site. However, the neutral loss fragment ions mass is the same as the mass of the water loss from unmodified residue, which leads to an ambiguous phosphosite localization [172].

8. Results and Discussion

4-ABA labeling improves the ionization efficiency and increases the net charge of (multi-) phosphorylated peptides, which makes them suitable for ETD fragmentation (Fig. 8.14). This facilitates the unambiguous pinpointing of the correct phosphorylation site(s). ETD fragmentation of the 4-ABA modified phosphopeptides generated an almost complete series of c and z ions. The more site-determining ions are detected, the higher the confidence in phosphosite localization. Compared to HCD of the unlabeled phosphopeptides, we found a substantial increase in the peptide ETD backbone fragmentation of the 4-ABA labeled phosphopeptides (Fig. 8.15).

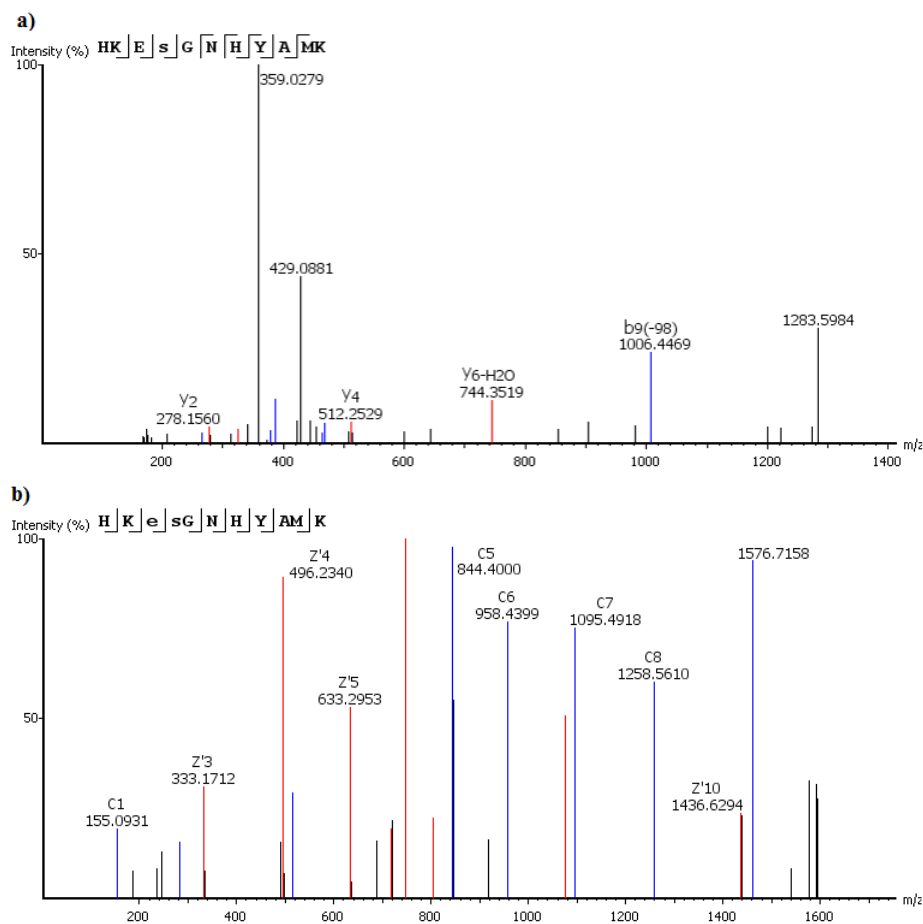


Figure 8.15.: MS/MS spectra of phosphorylated HKEpSGNHYAM from PKA before and after 4-ABA modification. **a)** HCD fragmentation of unmodified HKEpSGNHYAM, 3+ ($m/z = 461.1951$) **b)** ETD fragmentation of 4-ABA modified HKEpSGNHYAM, 3+ ($m/z = 530.5779$).

8.3. Conclusion

The qualitative and quantitative analysis of phosphoproteins by means of LC-MS remains a challenging task in the field of proteomics. As it has been reported by others, the analysis of multi-phosphorylated peptides is a major hurdle since the presence of multiple negative

phosphate groups decreases the peptide's net charge, increases its hydrophilicity, and causes poor ionization efficiency [165]. Moreover, the use of CID or HCD fragmentation of phosphopeptides cleaves the liable phosphate group, which makes the confident localization of phosphosite and the identification of the corresponding phosphopeptide a major hurdle [166]. ETD fragmentation preserves the liable phosphate group, albeit it is not suitable for the doubly charged peptide [129]. Therefore, we aimed to develop a method that applies a modification of the phosphate groups with 4-aminobenzylamine (4-ABA) combined with ETD fragmentation to overcome these limitations.

The 4-aminobenzylamine (4-ABA) carbodiimide-mediated labeling is applicable to intact proteins as well as to peptides and modifies the carboxyl groups as well as the phosphate groups. Applying the 4-ABA labeling to the intact protein reduces the sample variations that may occur during the sample preparation. 4-ABA labeling reverses the negative charge of the phosphates, increases the net charge, improves the ionization efficiency, balances the hydrophobicity, facilitates correct phosphosite(s) pinpointing, and causes a mass shift of $184.0402 n + 104.0738 m$, where n is the number of phosphosites and m is the number of carboxyl groups.

The localization of phosphosite(s) is often arbitrary and unreliable. 4-ABA labeling of phosphoprotein, and consequently, phosphopeptide is a promising method for the analysis of phosphorylation, specifically multi-phosphorylated peptides, without the need for enrichment. The increased number of positive charges at the 4-ABA modified peptides makes them suitable for ETD fragmentation. ETD fragmentation of the 4-ABA modified (phospho)peptides generated a high number of site-determining c- and z ions with high signal intensities, which facilitates the correct localization of phosphosite(s) with high confidence. The increased number of fragment ions having high signal intensity leads to additional peptide identifications during database searching of the mass spectra, and consequently more protein identifications and higher sequence coverage.

4-ABA labeling can also be advantageous for the analysis proteins and other types of HCD-liable modification using ETD. The derivatization of proteins with 4-ABA not only labels the phosphate groups but additionally it labels all the carboxyl groups-Glu, Asp, and C-terminus. The increased number of ETD fragment ions and the high signal intensities of detected c- and z ions improves the mass spectra and hence peptides identification and the sequence coverage during the database searching. The identification of a high number of (unique) peptides will enable the detection of protein variations in biological samples with high confidence.

Despite the promising preliminary results we obtained by labeling the phosphoproteins with 4-ABA, a complete 4-ABA labeling of all residues has not been achieved under the conditions we applied. Upon further method optimization, this method can be used to pinpoint the correct phosphorylation site(s) in proteins with clusters of phosphosite(s) e.g.

8. *Results and Discussion*

PER2 [168] and B-RAF [169]. By applying few adjustment, 4-ABA labeling of proteins can be used in a quantitative manner. But first, a complete labeling of all residues that can be modified has to be achieved, since the completeness of the labeling reaction is a prerequisite for the quantitative analysis.

Bibliography

- [1] P. N. Campbell. Principles of Biochemistry, volume 21-2, page 114. Worth Publishers New York NY, second edition edition, 1993.
- [2] M. R. Wilkins, J. C. Sanchez, A. A. Gooley, R. D. Appel, I. Humphery-Smith, D. F. Hochstrasser, and K. L. Williams. Progress with Proteome Projects: Why All Proteins Expressed by a Genome Should Be Identified and How to Do It. *Biotechnology and Genetic Engineering Reviews*, 13(1):19–50, 1996. PMID: 8948108.
- [3] O. N. Jensen. Modification-specific Proteomics: Characterization of Post-translational Modifications by Mass Spectrometry. *Current Opinion in Chemical Biology*, 8(1):33–41, 2004.
- [4] M. T. A. Sharabiani, M. Siermala, T. O. Lehtinen, and M. Vihinen. Dynamic Covariation Between Gene Expression and Proteome Characteristics. *BMC Bioinformatics*, 6(1):215, 2005.
- [5] K. Chandramouli and P.-Y. Qian. Proteomics: Challenges, Techniques and Possibilities to Overcome Biological Sample Complexity. *Human Genomics and Proteomics : HGP*, 2009:239204, 2009.
- [6] B. Deracinois, C. Flahaut, S. Duban-Deweer, and Y. Karamanos. Comparative and Quantitative Global Proteomics Approaches: an Overview. *Proteomes*, 1(3):180–218, 2013.
- [7] Scott A. McLuckey, , and J. Mitchell Wells. Mass Analysis at the Advent of the 21st Century. *Chem. Rev.*, 101(2):571–06, 2001. PMID: 11712257.
- [8] F. Hillenkamp, M. Karas, D. Holtkamp, and P. Klüsener. Energy Deposition in Ultraviolet Laser Desorption Mass Spectrometry of Biomolecules. *Int. J. Mass Spectrom. Ion Processes*, 69(3):265–76, 1986.
- [9] T. W. Jaskolla and M. Karas. Compelling Evidence for Lucky Survivor and Gas Phase Protonation: the Unified MALDI Analyte Protonation Mechanism. *Journal of The American Society for Mass Spectrometry*, 22(6):976–88, 2011.
- [10] R. Knochenmuss. MALDI Ionization Mechanisms: the Coupled Photophysical and Chemical Dynamics Model Correctly Predicts 'temperature'-selected Spectra. *Journal of Mass Spectrometry*, 48(9):998–1004, 2013.

Bibliography

- [11] C. S. Ho, C. W. Lam, M. H. Chan, R. C. Cheung, L. K. Law, L. C. Lit, K. F. Ng, M. W. Suen, and H. L. Tai. Electrospray Ionisation Mass Spectrometry: Principles and Clinical Applications. *Clin Biochem Rev*, 24(1):3–12, 2003.
- [12] P. Kebarle and U. H. Verkerk. Electrospray: from Ions in Solution to Ions in the Gas Phase, What We Know Now. *Mass Spectrom. Rev.*, 28(6):898–17, 2009.
- [13] A. Makarov, E. Denisov, A. Kholomeev, W. Balschun, O. Lange, K. Strupat, , and S. Horning. Performance Evaluation of a Hybrid Linear Ion Trap/Orbitrap Mass Spectrometer. *Analytical Chemistry*, 78(7):2113–20, 2006. PMID: 16579588.
- [14] P. E. Miller and M. B. Denton. The Quadrupole Mass Filter: Basic Operating Concepts. *J. Chem. Educ.*, 63(7):617, 1986.
- [15] G. Brucker and K. Van Antwerp. Comparison of Ion Trap Mass Spectrometer and Quadrupole Mass Spectrometer. *Society of Vacuum Coaters*, page 122–29, 2009.
- [16] J. Throck Watson and O. David Sparkman. Mass Spectrometry/Mass Spectrometry, in *Introduction to Mass Spectrometry: Instrumentation, Applications and Strategies for Data Interpretation*, pages 173–228. John Wiley & Sons, Ltd, fourth edition, 2008.
- [17] O. Nunez, E. Moyano, and M. T. Galceran. High Mass Accuracy In-source Collision-induced Dissociation Tandem Mass Spectrometry and Multi-Step Mass Spectrometry as Complementary Tools for Fragmentation Studies of Quaternary Ammonium Herbicides. *J. Mass Spectrom.*, 39(8):873–83, 2004.
- [18] G. Hopfgartner, E. Varesio, V. Tschappat, C. Grivet, E. Bourgoigne, and L. A. Leuthold. Triple Quadrupole Linear Ion Trap Mass Spectrometer for the Analysis of Small Molecules and Macromolecules. *Journal of Mass Spectrometry*, 39(8):845–55, 2004.
- [19] W. E. Stephens. A Pulsed Mass Spectrometer with Time Dispersion. *Phys. Rev.*, 69:674–674, Jun 1946.
- [20] E. d. Hoffmann and V. Stroobant. *Mass Spectrometry Principles and Applications*, pages 128–33. John Wiley & Sons Ltd, 3rd edition, 2007.
- [21] A. Makarov. Electrostatic Axially Harmonic Orbital Trapping: a High-performance Technique of Mass Analysis. *Anal. Chem.*, 72(6):1156–62, 2000. PMID: 10740853.
- [22] D. W. Koppenaal, C. J. Barinaga, M. B. Denton, R. P. Sperline, G. M. Hieftje, G. D. Schilling, F. J. Andrade, and Barnes J. H. Mass Spectrometry Detectors. *Anal. Chem.*, 77(21):418–27, 2005.
- [23] B. Paizs and S. Suhai. Fragmentation Pathways of Protonated Peptides. *Mass Spectrometry Reviews*, 24(4):508–548, 2005.

- [24] K. Biemann. Appendix 5. Nomenclature for Peptide Fragment Ions (positive Ions). In *Mass Spectrometry*, volume 193 of *Methods in Enzymology*, pages 886–87. Academic Press, 1990.
- [25] X. J. Tang, P. Thibault, and R. K. Boyd. Fragmentation Reactions of Multiply-protonated Peptides and Implications for Sequencing by Tandem Mass Spectrometry with Low-energy Collision-induced Dissociation. *Analytical Chemistry*, 65(20):2824–2834, 1993. PMID: 7504416.
- [26] H. Steen and M. Mann. The ABC’s (and XYZ’s) of Peptide Sequencing. *Nature Reviews. Molecular Cell Biology*, 5(9):699–711, 2004.
- [27] X. Han, A. Aslanian, and J. R. Yates, 3rd. Mass Spectrometry for Proteomics. *Curr Opin Chem Biol*, 12(5):483–90, 2008.
- [28] J. V. Olsen, B. Macek, O. Lange, A. Makarov, S. Horning, and M. Mann. Higher-energy C-trap Dissociation for Peptide Modification Analysis. *Nat M*, 4(9):709–12, 2007.
- [29] E. S. Boja and H. Rodriguez. Mass Spectrometry-based Targeted Quantitative Proteomics: Achieving Sensitive and Reproducible Detection of Proteins. *Proteomics*, 12(8):1093–10, 2012.
- [30] P. Pichler, T. Köcher, J. Holzmann, T. Möhring, G. Ammerer, and K Mechtler. Improved Precision of iTRAQ and TMT Quantification by an Axial Extraction Field in an Orbitrap HCDCell. *Analytical Chemistry*, 83(4):1469–74, 2011. PMID: 21275378.
- [31] L. M. Mikesch, B. Ueberheide, A. Chi, J. J. Coon, J. E. P. Syka, J. Shabanowitz, and D. F. Hunt. The Utility of ETD Mass Spectrometry in Proteomic Analysis. *Biochimica et Biophysica Acta*, 1764(12):1811–22, 2006.
- [32] E. Elviri. in *Tandem Mass Spectrometry - Applications and Principles*, pages 163–78. InTech, 2012.
- [33] C. K. Frese, A. F. M. Altelaar, H. van den Toorn, D. Nolting, J. Griep-Raming, A. J. R. Heck, and S. Mohammed. Toward Full Peptide Sequence Coverage by Dual Fragmentation Combining Electron-transfer and Higher-energy Collision Dissociation Tandem Mass Spectrometry. *Analytical Chemistry*, 84(22):9668–73, 2012. PMID: 23106539.
- [34] C. Denekamp and E. Rabkin. Radical Induced Fragmentation of Amino Acid Esters Using Triphenylcorrole (CuIII) Complexes. *J. Am. Soc. Mass. Spectrom.*, 18(4):791–01, 2007.
- [35] B. T. Chait. Chemistry. Mass Spectrometry: Bottom-up or Top-down? *Science*, 314(5796):65–6, 2006.

Bibliography

- [36] B. A. Garcia. What Does the Future Hold for Top-Down Mass Spectrometry? *J Am Soc Mass Spectrom*, 21(2):193–02, 2010.
- [37] A. Kalli and K. Hakansson. Electron Capture Dissociation of Highly Charged Proteolytic Peptides from Lys-N, Lys-C and Glu-C Digestion. *Mol Biosyst*, 6(9):1668–81, 2010.
- [38] J. Mayne, Z. Ning, X. Zhang, A. E. Starr, R. Chen, S. Deeke, C. K. Chiang, B. Xu, M. Wen, K. Cheng, D. Seebun, A. Star, J. I. Moore, and D. Figeys. Bottom-Up Proteomics (2013-2015): Keeping Up in the Era of Systems Biology. *Anal Chem*, 88(1):95–121, 2016.
- [39] M. Mann and O. N. Jensen. Proteomic Analysis of Post-translational Modifications. *Nat Biotechnol*, 21(3):255–61, 2003.
- [40] C. T. Walsh, S. Garneau-Tsodikova, and Jr. Gatto, G. J. Protein Posttranslational Modifications: the Chemistry of Proteome Diversifications. *Angew Chem Int Ed Engl*, 44(45):7342–72, 2005.
- [41] M. S. Cosgrove. Writers and Readers: Deconvoluting the Harmonic Complexity of the Histone Code. *Nat Struct Mol Biol*, 19(8):739–40, 2012.
- [42] T. H. Steinberg, T. K. Pretty On, K. N. Berggren, C. Kemper, L. Jones, Z. Diwu, R. P. Haugland, and W. F. Patton. Rapid and Simple Single Nanogram Detection of Glycoproteins in Polyacrylamide Gels and on Electroblobs. *Proteomics*, 1(7):841–55, 2001.
- [43] J. Cox and M. Mann. Quantitative, High-resolution Proteomics for Data-driven Systems Biology. *Annu. Rev. Biochem.*, 80(1):273–99, 2011. PMID: 21548781.
- [44] M. Bantscheff, M. Schirle, G. Sweetman, J. Rick, and B. Kuster. Quantitative Mass Spectrometry in Proteomics: a Critical Review. *Analytical and Bioanalytical Chemistry*, 389(4):1017–31, 2007.
- [45] Z. Li, R. M. Adams, K. Chourey, G. B. Hurst, R. L. Hettich, and C. Pan. Systematic Comparison of Label-free, Metabolic Labeling, and Isobaric Chemical Labeling for Quantitative Proteomics on LTQ Orbitrap Velos. *Journal of Proteome Research*, 11(3):1582–90, 2012. PMID: 22188275.
- [46] V. Lange, P. Picotti, B. Domon, and R. Aebersold. Selected Reaction Monitoring for Quantitative Proteomics: a Tutorial. *Mol. Syst. Biol.*, 4:222–222, 2008.
- [47] S. E. Ong, B. Blagoev, I. Kratchmarova, D. B. Kristensen, H. Steen, A. Pandey, and M. Mann. Stable Isotope Labeling by Amino Acids in Cell Culture, SILAC, as a Simple and Accurate Approach to Expression Proteomics. *Molecular & Cellular Proteomics*, 1(5):376–86, 2002.

- [48] S. P. Gygi, B. Rist, S. A. Gerber, F. Turecek, M. H. Gelb, and R. Aebersold. Quantitative Analysis of Complex Protein Mixtures Using Isotope-Coded Affinity Tags. *Nat Biotech*, 17(10):994–99, 1999.
- [49] M. H. Elliott, D. S. Smith, C. E. Parker, and C. Borchers. Current Trends in Quantitative Proteomics. *Journal of Mass Spectrometry*, 44(12):1637–60, 2009.
- [50] J. C. Silva, M. V. Gorenstein, G. Z. Li, J. P. C. Vissers, and S. J. Geromanos. Absolute Quantification of Proteins by LCMSE : a Virtue of Parallel MS Acquisition. *Molecular & Cellular Proteomics*, 5(1):144–56, 2006.
- [51] M. Bantscheff, S. Lemeer, M. M. Savitski, and B. Kuster. Quantitative Mass Spectrometry in Proteomics: Critical Review Update from 2007 to the Present. *Anal Bioanal Chem*, 404(4):939–65, 2012.
- [52] O. Chahrour, D. Cobice, and J. Malone. Stable Isotope Labelling Methods in Mass Spectrometry-based Quantitative Proteomics. *J Pharm Biomed Anal*, 113:2–20, 2015.
- [53] B. Plevoda and F. Sherman. The Diversity of Acetylated Proteins. *Genome Biol*, 3(5):reviews0006, April 2002.
- [54] T. Arnesen, P. Van Damme, B. Plevoda, K. Helsens, R. Evjenth, N. Colaert, J. E. Varhaug, J. Vandekerckhove, J. R. Lillehaug, F. Sherman, and K. Gevaert. Proteomics Analyses Reveal the Evolutionary Conservation and Divergence of N-terminal Acetyltransferases from Yeast and Humans. *Proceedings of the National Academy of Sciences*, 106(20):8157–62, 2009.
- [55] T. Kouzarides. Acetylation: a Regulatory Modification to Rival Phosphorylation? *The EMBO Journal*, 19(6):1176–79, 2000.
- [56] A. Drazic, L. M. Myklebust, R. Ree, and T. Arnesen. The World of Protein Acetylation. *Biochim Biophys Acta*, 1864(10):1372–401, 2016.
- [57] R. D. Kornberg. Chromatin Structure: A Repeating Unit of Histones and DNA. *Science*, 184(4139):868–71, 1974.
- [58] K. Luger, M. L. Dechassa, and D. J. Tremethick. New Insights into Nucleosome and Chromatin Structure: an Ordered State or a Disordered Affair? *Nature Reviews. Molecular Cell Biology*, 13:436–47, June 2012.
- [59] Luisa M. Figueiredo, George A. M. Cross, and Christian J. Janzen. Epigenetic Regulation in African Trypanosomes: a New Kid on the Block. *Nature Reviews Microbiology*, 7:504–13, July 2009.
- [60] T. Kouzarides. Chromatin Modifications and their Function. *Cell*, 128(4):693–05, 2007.

Bibliography

- [61] H. Huang, S. Lin, B. A. Garcia, and Y. Zhao. Quantitative Proteomic Analysis of Histone Modifications. *Chem Rev*, 115(6):2376–18, 2015.
- [62] B. D. Strahl and C. D. Allis. The Language of Covalent Histone Modifications. *Nature*, 403:41–5, January 2000.
- [63] R. T. Kamakaka and S. Biggins. Histone Variants: Deviants? *Genes & Development*, 19(3):295–16, 2005.
- [64] S. Henikoff, T. Furuyama, and K. Ahmad. Histone Variants, Nucleosome Assembly and Epigenetic Inheritance. *Trends in Genetics*, 20(7):320–6, 2004.
- [65] M. L. Hennrich, P. J. Boersema, H. van den Toorn, N. Mischerikow, A. J. Heck, and S. Mohammed. Effect of Chemical Modifications on Peptide Fragmentation Behavior Upon Electron Transfer Induced Dissociation. *Anal Chem*, 81(18):7814–22, 2009.
- [66] R. E. Kingston and G. J. Narlikar. ATP-dependent Remodeling and Acetylation as Regulators of Chromatin Fluidity. *Genes & Development*, 13(18):2339–52, 1999.
- [67] A. L. Rodd, K. Ververis, and T. C. Karagiannis. Current and Emerging Therapeutics for Cutaneous T-cell Lymphoma: Histone Deacetylase Inhibitors. *Lymphoma*, 2012:1–10, 2012.
- [68] C. E. Brown, T. Lechner, A. K. Howe, and J. L. Workman. The Many HATs of Transcription Coactivators. *Trends in Biochemical Sciences*, 25(1):15 –19, 2000.
- [69] W. Zhang, J. R. Bone, D. G. Edmondson, B. M. Turner, and S. Y. Roth. Essential and Redundant Functions of Histone Acetylation Revealed by Mutation of Target Lysines and Loss of the Gcn5p Acetyltransferase. *The EMBO Journal*, 17(11):3155–67, 1998.
- [70] C. Choudhary, B. T. Weinert, Y. Nishida, E. Verdin, and M. Mann. The Growing Landscape of Lysine Acetylation Links Metabolism and Cell Signalling. *Nat Rev Mol Cell Biol*, 15(8):536–50, 2014.
- [71] V. L. Alonso and E. C. Serra. Lysine Acetylation: Elucidating the Components of an Emerging Global Signaling Pathway in Trypanosomes. *J Biomed Biotechnol*, 2012:452–67, 2012.
- [72] V. Mandava, J. P. Fernandez, H. Deng, C. J. Janzen, S. B. Hake, and G. A. Cross. Histone Modifications in *Trypanosoma Brucei*. *Mol Biochem Parasitol*, 156(1):41–50, 2007.
- [73] T. N. Siegel, T. Kawahara, J. A. Degrasse, C. J. Janzen, D. Horn, and G. A. Cross. Acetylation of Histone H4K4 Is Cell Cycle Regulated and Mediated by HAT3 in *Trypanosoma Brucei*. *Mol Microbiol*, 67(4):762–71, 2008.

- [74] T. Kawahara, T. N. Siegel, A. K. Ingram, S. Alsford, G. A. M. Cross, and D. Horn. Two Essential MYST-family Proteins Display Distinct Roles in Histone H4K10 Acetylation and Telomeric Silencing in Trypanosomes. *Molecular Microbiology*, 69(4):1054–68, 2008.
- [75] Y. Katan-Khaykovich and K. Struhl. Dynamics of Global Histone Acetylation and Deacetylation in Vivo: Rapid Restoration of Normal Histone Acetylation Status Upon Removal of Activators and Repressors. *Genes Dev.*, 16:743–52, March 2002.
- [76] S. M. Fuchs, K. Krajewski, V. L. Baker, R. W. and Miller, and B. D. Strahl. Influence of Combinatorial Histone Modifications on Antibody and Effector Protein Recognition. *Current Biology*, 21(1):53–8, 2011.
- [77] C. M. Smith, Z. W. Haimberger, C. O. Johnson, A. J. Wolf, P. R. Gafken, Z. Zhang, M. R. Parthun, and D. E. Gottschling. Heritable Chromatin Structure: Mapping "memory" in Histones H3 and H4. *Proc. Natl. Acad. Sci. U.S.A.*, 99 Suppl 4:16454–61, December 2002.
- [78] M. Downey, J. R. Johnson, N. E. Davey, B. W. Newton, T. L. Johnson, S. Galaang, C. A. Seller, N. Krogan, and D. P. Toczyski. Acetylome Profiling Reveals Overlap in the Regulation of Diverse Processes by Sirtuins, Gcn5, and Esa1. *Molecular & Cellular Proteomics*, 14(1):162–76, 2015.
- [79] T. Svinkina, H. Gu, J. C. Silva, P. Mertins, J. Qiao, S. Fereshetian, J. D. Jaffe, E. Kuhn, N. D. Udeshi, and S. A. Carr. Deep, Quantitative Coverage of the Lysine Acetylome Using Novel Anti-acetyl-lysine Antibodies and an Optimized Proteomic Workflow. *Mol Cell Proteomics*, 14(9):2429–40, 2015.
- [80] E. J. Vazquez, J. M. Berthiaume, V. Kamath, O. Achike, E. Buchanan, M. M. Montano, M. P. Chandler, M. Miyagi, and M. G. Rosca. Mitochondrial complex I defect and increased fatty acid oxidation enhance protein lysine acetylation in the diabetic heart. *Cardiovascular Research*, 107(4):453, 2015.
- [81] S. A. Gerber, J. Rush, O. Stemman, M. W. Kirschner, and S. P. Gygi. Absolute Quantification of Proteins and Phosphoproteins from Cell Lysates by Tandem MS. *Proceedings of the National academy of Sciences of the United States of America*, 100:6940–5, June 2003.
- [82] O. F. Sanchez, D. Williamson, L. Cai, and C. Yuan. A Sensitive Protein-based Sensor for Quantifying Histone Acetylation Levels. *Talanta*, 140:212–8, 2015.
- [83] C. M. Smith, P. R. Gafken, Z. Zhang, D. E. Gottschling, J. B. Smith, and D. L. Smith. Mass Spectrometric Quantification of Acetylation at Specific Lysines Within the Amino-terminal Tail of Histone H4. *Anal Biochem*, 316(1):23–33, 2003.
- [84] C. M. Smith. Quantification of Acetylation at Proximal Lysine Residues Using Isotopic Labeling and Tandem Mass Spectrometry. *Methods*, 36(4):395–03, 2005.

Bibliography

- [85] E. Hersman, D. M. Nelson, W. P. Griffith, C. Jelinek, and R. J. Cotter. Analysis of Histone Modifications from Tryptic Peptides of Deuteroacetylated Isoforms. *Int. J. Mass Spectrom.*, 312:5–16, February 2012.
- [86] K. L. Fiedler, P. Bheda, J. Dai, J. D. Boeke, C. Wolberger, and R. J. Cotter. A Quantitative Analysis of Histone Methylation and Acetylation Isoforms from their Deuteroacetylated Derivatives: Application to a Series of Knockout. *J. Mass Spectrom.*, 48:608–15, May 2013.
- [87] C. Feller, I. Forné, A. Imhof, and P. Becker. Global and Specific Responses of the Histone Acetylome to Systematic Perturbation. *Molecular Cell*, 57(3):559–71, 2015.
- [88] A. D’Urzo, A. P. Boichenko, T. van den Bosch, J. Hermans, F. Dekker, V. Andrisano, and R. Bischoff. Site-specific Quantification of Lysine Acetylation in the N-terminal Tail of Histone H4 Using a Double-labelling, Targeted UHPLC MS/MS Approach. *Anal Bioanal Chem*, 408(13):3547–53, 2016.
- [89] D. Vitko, P. Majek, E. Schirghuber, S. Kubicek, and K. L. Bennett. FASIL-MS: an Integrated Proteomic and Bioinformatic Workflow to Universally Quantitate in Vivo-acetylated Positional Isomers. *J Proteome Res*, 15(8):2579–94, 2016.
- [90] J. Baeza, J. A. Dowell, M. J. Smallegan, J. Fan, D. Amador-Noguez, Z. Khan, and J. M. Denu. Stoichiometry of Site-specific Lysine Acetylation in an Entire Proteome. *J Biol Chem*, 289(31):21326–38, 2014.
- [91] J.F. Riordan and B.L. Vallee. [66] Acetylation. In *Enzyme Structure*, volume 11 of *Methods in Enzymology*, pages 565–70. Academic Press, 1967.
- [92] M. Aslett, C. Aurrecochea, M. Berriman, J. Brestelli, B. P. Brunk, M. Carrington, D. P. Depledge, S. Fischer, B. Gajria, X. Gao, M. J. Gardner, A. Gingle, G. Grant, O. S. Harb, M. Heiges, C. Hertz-Fowler, R. Houston, F. Innamorato, J. Iodice, J. C. Kissinger, E. Kraemer, W. Li, F. J. Logan, J. A. Miller, S. Mitra, P. J. Myler, V. Nayak, C. Pennington, I. Phan, D. F. Pinney, G. Ramasamy, M. B. Rogers, D. S. Roos, C. Ross, D. Sivam, D. F. Smith, G. Srinivasamoorthy, Jr. Stoeckert, C. J., S. Subramanian, R. Thibodeau, A. Tivey, C. Treatman, G. Velarde, and H. Wang. TriTrypDB: a Functional Genomic Resource for the Trypanosomatidae. *Nucleic Acids Res.*, 38, January 2010.
- [93] K. Luger, T. J. Rechsteiner, and T. J. Richmond. Expression and Purification of Recombinant Histones and Nucleosome Reconstitution, pages 1–16. Humana Press, Totowa, NJ, 1999.
- [94] J. E. Lowell, F. Kaiser, C. J. Janzen, and G. A. M. Cross. Histone H2AZ Dimerizes with a Novel Variant H2B and is Enriched at Repetitive DNA in *Trypanosoma Brucei*. *Journal of Cell Science*, 118(24):5721–30, 2005.

- [95] D. Shechter, H. L. Dormann, C. D. Allis, and S. B. Hake. Extraction, Purification and Analysis of Histones. *Nature Protocols*, 2:1445–57, 2007.
- [96] A. Günzl and B. Schimanski. *Tandem Affinity Purification of Proteins*, pages 1919–55. John Wiley & Sons, Inc., 2001.
- [97] R. ElBashir, J. T. Vanselow, A. Kraus, C. J. Janzen, T. N. Siegel, and A. Schlosser. Fragment Ion Patchwork Quantification for Measuring Site-specific Acetylation Degrees. *Anal Chem*, 87(19):9939–45, 2015.
- [98] M. M. Savitski, S. Lemeer, M. Boesche, M. Lang, T. Mathieson, M. Bantscheff, and B. Kuster. Confident Phosphorylation Site Localization Using the Mascot Delta Score. *Molecular & Cellular Proteomics*, 10(2), 2011.
- [99] R. L. Lundblad. *Chemical Reagents for Protein Modification*, Third Edition, chapter The Modification of Amino Groups, pages 31–66. CRC Press, 2004.
- [100] V. G. Allfrey, R. Faulkner, and A. E. Mirsky. Acetylation and Methylation of Histones and Their Possible Role in the Regulation of RNA Synthesis. *Proceedings of the National Academy of Sciences*, 51(5):786–94, 1964.
- [101] A. Gassen, D. Brechtefeld, N. Schandry, J. M. Arteaga-Salas, L. Israel, A. Imhof, and C. J. Janzen. DOT1A-dependent H3K76 Methylation Is Required for Replication Regulation in *Trypanosoma Brucei*. *Nucleic Acids Research*, 40(20):10302, 2012.
- [102] V. Mandava, C. J. Janzen, and G. A. Cross. Trypanosome H2B.V Replaces H2B in Nucleosomes Enriched for H3 K4 and K76 Trimethylation. *Biochem Biophys Res Commun*, 368(4):846–51, 2008.
- [103] C. Bönisch and S. B. Hake. Histone H2A Variants in Nucleosomes and Chromatin: More or Less Stable? *Nucleic Acids Research*, 40(21):107–19, 2012.
- [104] T. N. Siegel, D. R. Hekstra, L. E. Kemp, L. M. Figueiredo, J. E. Lowell, D. Fenyo, X. Wang, S. Dewell, and G. A. Cross. Four Histone Variants Mark the Boundaries of Polycistronic Transcription Units in *Trypanosoma Brucei*. *Genes Dev*, 23(9):1063–76, 2009.
- [105] J. D. Nardoizzi, K. Lott, and G. Cingolani. Phosphorylation Meets Nuclear Import: a Review. *Cell Commun Signal*, 8:32, 2010.
- [106] P. Cohen. The Origins of Protein Phosphorylation. *Nat Cell Biol*, 4(5):E127–E130, 2002.
- [107] S. J. Humphrey, D. E. James, and M. Mann. Protein Phosphorylation: A Major Switch Mechanism for Metabolic Regulation. *Trends Endocrinol Metab*, 26(12):676–87, 2015.

Bibliography

- [108] P. Cohen. The Role of Protein Phosphorylation in Human Health and Disease. *Eur. J. Biochem.*, 268(19):5001–10, 2001.
- [109] G. Manning, D. B. Whyte, R. Martinez, T. Hunter, and S. Sudarsanam. The Protein Kinase Complement of the Human Genome. *Science*, 298(5600):1912–34, 2002.
- [110] M. Mann, S. E. Ong, M. Gronborg, H. Steen, O. N. Jensen, and A. Pandey. Analysis of Protein Phosphorylation Using Mass Spectrometry: Deciphering the Phosphoproteome. *Trends Biotechnol*, 20(6):261–8, 2002.
- [111] J. V. Olsen, B. Blagoev, F. Gnad, B. Macek, C. Kumar, P. Mortensen, and M. Mann. Global, In-Vivo, and Site-specific Phosphorylation Dynamics in Signaling Networks. *Cell*, 127:635–48, November 2006.
- [112] A. Sickmann and H. E. Meyer. Phosphoamino Acid Analysis. *PROTEOMICS*, 1(2):200–6, 2001.
- [113] T. M. Martensen. Chemical-properties, Isolation, and Analysis of O-phosphates in Proteins. *Methods in Enzymology*, 107:3–23, 1984.
- [114] P.-K. Chong, H. Lee, J. W.-F. Kong, M. C.-S. Loh, C.-H. Wong, and Y.-P. Lim. Phosphoproteomics, Oncogenic Signaling and Cancer Research. *PROTEOMICS*, 8(21):4370–82, 2008.
- [115] L. K. Iwai, C. Benoist, D. Mathis, and F. M. White. Quantitative Phosphoproteomic Analysis of T Cell Receptor Signaling in Diabetes Prone and Resistant Mice. *Journal of proteome research*, 9:3135–45, 2010.
- [116] P. Rudrabhatla, P. Grant, H. Jaffe, M. J. Strong, and H. C. Pant. Quantitative Phosphoproteomic Analysis of Neuronal Intermediate Filament Proteins (NF-M/H) in Alzheimer’s Disease by iTRAQ. *The FASEB Journal*, 24:4396–07, 2010.
- [117] J. X. Yan, N. H. Packer, A. A. Gooley, and K. L. Williams. Protein Phosphorylation: Technologies for the Identification of Phosphoamino Acids. *Journal of Chromatography A*, 808:23–41, May 1998.
- [118] D. T. McLachlin and B. T. McLachlin. Analysis of Phosphorylated Proteins and Peptides by Mass Spectrometry. *Current Opinion in Chemical Biology*, 5(5):591–02, 2001.
- [119] A. Pandey, A. V. Podtelejnikov, B. Blagoev, X. R. Bustelo, M. Mann, and Lodish H. F. Analysis of Receptor Signalling Pathways by Mass Spectrometry: Identification of Vav-2 as a Substrate of the Epidermal and Platelet-derived Growth Factor Receptors. *Proc. Natl. Acad. Sci*, 97:179–84, Jan 2000.
- [120] T. H. Steinberg, B. J. Agnew, K. R. Gee, W.-Y. Leung, T. Goodman, B. Schulenberg, J. Hendrickson, J. M. Beechem, R. P. Haugland, and W. F. Patton. Global

- Quantitative Phosphoprotein Analysis Using Multiplexed Proteomics Technology. *PROTEOMICS*, 3(7):1128–44, 2003.
- [121] B. Schulenberg, R. Aggeler, J. M. Beechem, R. A. Capaldi, and W. F. Patton. Analysis of Steady-state Protein Phosphorylation in Mitochondria Using a Novel Fluorescent Phosphosensor Dye. *Journal of Biological Chemistry*, 278(29):27251–5, 2003.
- [122] G. Carmel, B. Leichus, X. Cheng, S. D. Patterson, U. Mirza, B. T. Chait, and J. Kuret. Expression, Purification, Crystallization, and Preliminary X-ray Analysis of Casein Kinase-1 from *Schizosaccharomyces*. *J. Biol. Chem.*, 269:7304–09, 1994.
- [123] J. P. DeGnore and J. Qin. Fragmentation of Phosphopeptides in an Ion Trap Mass Spectrometer. *Journal of the American Society for Mass Spectrometry*, 9(11):1175–88, 1998.
- [124] A. Nita-Lazar, H. Saito-Benz, and F. M. White. Quantitative Phosphoproteomics by Mass Spectrometry: Past, Present, and Future. *Proteomics*, 8(21):4433–43, 2008.
- [125] A. Schlosser, R. Pipkorn, D. Bossemeyer, and W. D. Lehmann. Analysis of Protein Phosphorylation by a Combination of Elastase Digestion and Neutral Loss Tandem Mass Spectrometry. *Analytical Chemistry*, 73(2):170–76, 2001. PMID: 11199962.
- [126] M. J. Schroeder, J. Shabanowitz, J. C. Schwartz, D. F. Hunt, and J. J. Coon. A Neutral Loss Activation Method for Improved Phosphopeptide Sequence Analysis by Quadrupole Ion Trap Mass Spectrometry. *Analytical Chemistry*, 76(13):3590–8, 2004. PMID: 15228329.
- [127] R. A. Zubarev, D. M. Horn, E. K. Fridriksson, N. L. Kelleher, N. A. Kruger, M. A. Lewis, B. K. Carpenter, and F. W. McLafferty. Electron Capture Dissociation for Structural Characterization of Multiply Charged Protein Cations. *Analytical Chemistry*, 72(3):563–73, 2000. PMID: 10695143.
- [128] J. E. Syka, J. J. Coon, M. J. Schroeder, J. Shabanowitz, and D. F. Hunt. Peptide and Protein Sequence Analysis by Electron Transfer Dissociation Mass Spectrometry. *Proc Natl Acad Sci U S A*, 101(26):9528–33, 2004.
- [129] D. M. Good, M. Wirtala, G. C. McAlister, and J. J. Coon. Performance Characteristics of Electron Transfer Dissociation Mass Spectrometry. *Mol Cell Proteomics*, 6(11):1942–51, 2007.
- [130] C. Stingl, C. Ihling, G. Ammerer, A. Sinz, and K. Mechtler. Application of Different Fragmentation Techniques for the Analysis of Phosphopeptides Using a Hybrid Linear Ion Trap-FTICR Mass Spectrometer. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1764(12):1842–52, 2006. Posttranslational Modifications in Proteomics.

Bibliography

- [131] S. M. Sweet, C. M. Bailey, D. L. Cunningham, J. K. Heath, and H. J. Cooper. Large Scale Localization of Protein Phosphorylation by Use of Electron Capture Dissociation Mass Spectrometry. *Molecular & Cellular Proteomics*, 8(5):904–12, 2009.
- [132] C. K. Frese, H. Zhou, T. Taus, A. F. M. Altelaar, K. Mechtler, A. J. Heck, and S. Mohammed. Unambiguous Phosphosite Localization Using Electron-transfer/higher-energy Collision Dissociation (EThcD). *Journal of Proteome Research*, 12(3):1520–25, 2013. PMID: 23347405.
- [133] B. Bodenmiller, L. N. Mueller, M. Mueller, B. Domon, and R. Aebersold. Reproducible Isolation of Distinct, Overlapping Segments of the Phosphoproteome. *Nature Methods*, 4:231–37, 2007.
- [134] B. Bodenmiller, L. N. Mueller, P. G. Pedrioli, D. Pflieger, M. A. Junger, J. K. Eng, R. Aebersold, and W. A. Tao. An Integrated Chemical, Mass Spectrometric and Computational Strategy for (quantitative) Phosphoproteomics: Application to *Drosophila Melanogaster* Kc167 Cells. *Mol Biosyst*, 3:275–86, 2007.
- [135] E. C. Reynolds, P.F. Riley, and N.J. Adamson. A Selective Precipitation Purification Procedure for Multiple Phosphoserine-containing Peptides and Methods for Their Identification. *Analytical Biochemistry*, 217(2):277–84, 1994.
- [136] G. Maccarrone, N. Kolb, L. Teplytska, I. Birg, R. Zollinger, F. Holsboer, and C. W. Turck. Phosphopeptide Enrichment by IEF. *ELECTROPHORESIS*, 27(22):4585–95, 2006.
- [137] S. A. Beausoleil, M. Jedrychowski, D. Schwartz, J. E. Elias, J. Villén, J. Li, M. A. Cohn, L. C. Cantley, and S. P. Gygi. Large-scale Characterization of HeLa Cell Nuclear Phosphoproteins. *Proceedings of the National Academy of Sciences of the United States of America*, 101(33):12130–35, 2004.
- [138] G. Han, M. Ye, H. Zhou, X. Jiang, S. Feng, X. Jiang, R. Tian, D. Wan, H. Zou, and J. Gu. Large-scale Phosphoproteome Analysis of Human Liver Tissue by Enrichment and Fractionation of Phosphopeptides with Strong Anion Exchange Chromatography. *PROTEOMICS*, 8(7):1346–61, 2008.
- [139] D. E. McNulty and R. S. Annan. Hydrophilic Interaction Chromatography Reduces the Complexity of the Phosphoproteome and Improves Global Phosphopeptide Isolation and Detection. *Molecular & Cellular Proteomics*, 7(5):971–80, 2008.
- [140] A. J. Alpert. Electrostatic Repulsion Hydrophilic Interaction Chromatography for Isocratic Separation of Charged Solutes and Selective Isolation of Phosphopeptides. *Analytical Chemistry*, 80(1):62–76, 2008. PMID: 18027909.

- [141] M. R. Larsen, T. E. Thingholm, O. N. Jensen, P. Roepstorff, and T. J. D. Jorgensen. Highly Selective Enrichment of Phosphorylated Peptides from Peptide Mixtures Using Titanium Dioxide Microcolumns. *Molecular & Cellular Proteomics*, 4(7):873–86, 2005.
- [142] F. Wolschin, S. Wienkoop, and W. Weckwerth. Enrichment of Phosphorylated Proteins and Peptides from Complex Mixtures Using Metal Oxide/hydroxide Affinity Chromatography (MOAC). *PROTEOMICS*, 5(17):4389–97, 2005.
- [143] H. K. Kweon and K. Hakansson. Selective Zirconium Dioxide-based Enrichment of Phosphorylated Peptides for Mass Spectrometric Analysis. *Analytical Chemistry*, 78:1743–9, 2006.
- [144] S. B. Ficarro, J. R. Parikh, N. C. Blank, and J. A. Marto. Niobium(v) Oxide (Nb₂O₅) Application to Phosphoproteomics. *Analytical Chemistry*, 80(12):4606–13, 2008. PMID: 18491922.
- [145] M. Sturm, A. Leitner, J. K. Smatt, M. Linden, and W. Lindner. Tin Dioxide Microspheres as a Promising Material for Phosphopeptide Enrichment Prior to Liquid Chromatography-(tandem) Mass Spectrometry Analysis. *Advanced Functional Materials*, 18(16):2381–9, 2008.
- [146] J. G. Rivera, Y. S. Choi, S. Vujcic, T. D. Wood, and L. A. Colon. Enrichment/Isolation of Phosphorylated Peptides on Hafnium Oxide Prior to Mass Spectrometric Analysis. *Analyst*, 134:31–33, 2009.
- [147] L. Negroni, S. Claverol, J. Rosenbaum, E. Chevet, M. Bonneu, and J.-M. Schmitter. Comparison of {IMAC} and {MOAC} for Phosphopeptide Enrichment by Column Chromatography. *Journal of Chromatography B*, 891–892:109–12, 2012.
- [148] T. E. Thingholm and M. R. Larsen. Sequential Elution from IMAC (SIMAC): an Efficient Method for Enrichment and Separation of Mono- and Multi-phosphorylated Peptides. *Methods Mol. Biol.*, 1355:147–60, 2016.
- [149] J. Kremerskothen and A. Barnekow. Determination of Phosphotyrosine Phosphatase (ptpase) Activity in Normal and Transformed Cells. *Molecular and Cellular Biochemistry*, 125(1):1–9, 1993.
- [150] A. M. Palumbo, S. A. Smith, C. L. Kalcic, M. Dantus, P. M. Stemmer, and G. E. Reid. Tandem Mass Spectrometry Strategies for Phosphoproteome Analysis. *Mass Spectrometry Reviews*, 30(4):600–25, 2011.
- [151] C. Dickhut, I. Feldmann, J. Lambert, and R. P. Zahedi. Impact of Digestion Conditions on Phosphoproteomics. *J Proteome Res*, 13(6):2761–70, 2014.
- [152] D. Winter, J. Seidler, Y. Ziv, Y. Shiloh, and W. D. Lehmann. Citrate Boosts the

Bibliography

- Performance of Phosphopeptide Analysis by UPLC-ESI-MS/MS. *J Proteome Res*, 8(1):418–24, 2009.
- [153] Y.i Ishihama, F.-Y. Wei, K. Aoshima, T. Sato, J. Kuromitsu, and Y. Oda. Enhancement of the Efficiency of Phosphoproteomic Identification by Removing Phosphates After Phosphopeptide Enrichment. *Journal of Proteome Research*, 6(3):1139–44, 2007. PMID: 17330947.
- [154] F. Kjeldsen, A. M. B. Giessing, C. R. Ingrell, and O. N. Jensen. Peptide Sequencing and Characterization of Post-translational Modifications by Enhanced Ion-charging and Liquid Chromatography Electron-transfer Dissociation Tandem Mass Spectrometry. *Analytical Chemistry*, 79(24):9243–52, 2007. PMID: 18020370.
- [155] S. B. Ficarro, M. L. McClelland, P. T. Stukenberg, D. J. Burke, M. M. Ross, J. Shabanowitz, D. F. Hunt, and F. M. White. Phosphoproteome Analysis by Mass Spectrometry and Its Application to *Saccharomyces Cerevisiae*. *Nature Biotechnology*, 20:301–5, 2002.
- [156] Y. Li and R. B. Cole. Shifts in Peptide and Protein Charge State Distributions with Varying Spray Tip Orifice Diameter in Nanoelectrospray Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Analytical Chemistry*, 75(21):5739–46, 2003. PMID: 14588013.
- [157] A. T. Iavarone and E. R. Williams. Supercharging in Electrospray Ionization: Effects on Signal and Charge. *International Journal of Mass Spectrometry*, 219(1):63–72, 2002. *Structure of Biological Molecules*.
- [158] L. Zhang, Y. Xu, H. Lu, and P. Yang. Carboxy Group Derivatization for Enhanced Electron-transfer Dissociation Mass Spectrometric Analysis of Phosphopeptides. *Proteomics*, 9(16):4093–7, 2009.
- [159] B. J. Ko and J. S. Brodbelt. Enhanced Electron Transfer Dissociation of Peptides Modified at C-terminus with Fixed Charges. *J Am Soc Mass Spectrom*, 23(11):1991–2000, 2012.
- [160] C. J. Krusemark, B. L. Frey, P. J. Belshaw, and L. M. Smith. Modifying the Charge State Distribution of Proteins in Electrospray Ionization Mass Spectrometry by Chemical Derivatization. *J Am Soc Mass Spectrom*, 20(9):1617–25, 2009.
- [161] B. L. Frey, D. T. Ladrer, S. B. Sondalle, C. J. Krusemark, A. L. Jue, , J. J. Coon, and L. M. Smith. Chemical Derivatization of Peptide Carboxyl Groups for Highly Efficient Electron Transfer Dissociation. *Journal of the American Society for Mass Spectrometry*, 24:1710–21, 2013.
- [162] H. Zhou, J. D. Watts, and R. Aebersold. A Systematic Approach to the Analysis of Protein Phosphorylation. *Nature Biotechnology*, 19:375–78, 2001.

- [163] A. Panchaud, J. Hansson, M. Affolter, R. Bel Rhlid, S. Piu, P. Moreillon, and M. Kussmann. ANIBAL, Stable Isotope-based Quantitative Proteomics by Aniline and Benzoic Acid Labeling of Amino and Carboxylic Groups. *Mol Cell Proteomics*, 7(4):800–12, 2008.
- [164] Y. Shi, B. Bairami, M. Morton, and X. D. Yao. Cyclophosphoramidate for Ion as Mass Defect Marker for Efficient Detection of Protein Serine Phosphorylation. *Analytical Chemistry*, 80(19):7614–23, 2008.
- [165] S. Liu, C. Zhang, J. L. Campbell, H. Zhang, V. K. Yeung, K. K.-and Han, and G. A. Lajoie. Formation of Phosphopeptide-metal Ion Complexes in Liquid Chromatography/electrospray Mass Spectrometry and Their Influence on Phosphopeptide Detection. *Rapid Communications in Mass Spectrometry*, 19(19):2747–56, 5 2005.
- [166] B. A. Garcia, J. Shabanowitz, and D. F. Hunt. Analysis of Protein Phosphorylation by Mass Spectrometry. *Methods*, 35(3):256–64, 2005.
- [167] A. F. Lopez-Clavijo, C. A. Duque-Daza, A. J. Creeser, and H. J. Cooper. Electron Capture Dissociation Mass Spectrometry of Phosphopeptides: Arginine and Phosphoserine. *International Journal of Mass Spectrometry*, 390:63 – 70, 2015. Special Issue: Biological Cation Radicals, Ion-Electron and Ion-Ion Interactions.
- [168] K. Vanselow, J. T. Vanselow, P. O. Westermarck, S. Reischl, B. Maier, T. Korte, A. Herrmann, H. Herzel, A. Schlosser, and A. Kramer. Differential Effects of PER2 Phosphorylation: Molecular Basis for the Human Familial Advanced Sleep Phase Syndrome (FASPS). *Genes Dev*, 20(19):2660–72, 2006.
- [169] A. E. Eisenhardt, A. Sprenger, M. Röring, R. Herr, F. Weinberg, M. Köhler, S. Braun, J. Orth, B. Diedrich, U. Lanner, N. Tscherswinski, S. Schuster, N. Dumaz, E. Schmidt, R. Baumeister, A. Schlosser, J. Dengjel, , and T. Brummer. Phosphoproteomic Analyses of B-raf Protein Complexes Reveal New Regulatory Principles. *Oncotarget*, 7(18):26628–52, 2016.
- [170] G. T. Hermanson. *Bioconjugate Techniques*, chapter 4, pages 259–73. Academic Press, Boston, third edition edition, August 2013.
- [171] T. Langer, S. Sreeramulu, M. Vogtherr, B. Elshorst, M. Betz, U. Schieberr, K. Saxena, and H. Schwalbe. Folding and Activity of cAMP-dependent Protein Kinase Mutants. *FEBS Lett.*, 579(19):4049–54, 2005.
- [172] A. Stensballe, O. N. Jensen, J. V. Olsen, K. F. Haselmann, and R. A. Zubarev. Electron Capture Dissociation of Singly and Multiply Phosphorylated Peptides. *Rapid Communications in Mass Spectrometry*, 14(19):1793–1800, 2000.

A. Abbreviations

2-ABA	2-aminobenzylamine
4-ABA	4-aminobenzylamine
Å	Angstrom
ABC	Ammonium Bicarbonate
Ac	Acetylation
AC	Alternating Current
AC ₂ O	Acetic Anhydride
Ac-COA	Acetyl-Coenzyme A
ACN	Acetonitrile
ADP	Adenosine Diphosphate
AGC	Automatic Gain Control
ANIBAL	Aniline And Benzoic Acid Labeling
AQUA	Absolute Quantiation
Arg-C	Endoproteinase Arg-C
Asp-N	Endoproteinase Asp-N
ATP	Adenosine Triphosphate
Bis-Tris	2-[Bis(2-Hydroxyethyl)Amino]-2-(Hydroxymethyl)-1,3-Propandiol
°C	Degree Celsius
CHCA	α -Cyano-4-Hydroxycinnamic Acid
ChIP	Chromatin Immunoprecipitation
CI	Chemical Ionization
CID	Collision Induced Dissociation
CyPAA	Cyclophosphoramidate Ion
Da	Dalton
DAHC	Diammonium Hydrogen Citrate
DC	Direct Current
DHAP	2, 6-Dihydroxyacetophenone
DHB	2,5-Dihydroxybenzoic Acid
DMF	Dimethylformamide
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic Acid
DTT	Dithiothreitol

A. Abbreviations

<i>E. Coli</i>	<i>Escherichia Coli</i>
ECD	Electron Capture Dissociation
EDC	<i>N</i> -(3-Dimethylaminopropyl)- <i>N'</i> -ethylcarbodiimide Hydrochloride
EDTA	Ethylenediaminetetraacetic Acid
ERLIC	Electrostatic Repulsion Hydrophilic Interaction Chromatography
ESI	Electrospray Ionization
ETD	Electron Transfer Dissociation
EThcD	Electron Transfer and Higher-energy Dissociation
FA	Formic Acid
FAB	Fast Atom Bombardment Ionization
FASP	Filter Aided Sample Preparation
FD	Field Desorption
FDR	False Discovery Rate
FI	Field Ionization
FT-ICR	Fourier Transform Ion Cyclotron Resonance
FTMS	Fourier Transform Mass Spectrometer
GC-MS	Gas Chromatography Mass Spectrometry
Glu-C	Endoproteinase Glu-C
HAT	Histone Acetyltransferase
HCD	High-energy Collision Dissociation
HCl	Hydrochloric Acid
HDAC	Histone Deacetylase
HILIC	Hydrophilic Interaction Chromatography
HPLC	High Performance Liquid Chromatography
IAA	Iodoacetamide
ICAT	Isotope-coded Affinity Tag
IMAC	Immobilized Metal Affinity Chromatography
iTRAQ	Isobaric Tags For Absolute Quantitation
KAT	Lysine Acetyltransferase
KDAC	Lysine Deacetylase
KO	Knock Out
LC	Liquid Chromatography
LC-ESI-MS	Liquid Chromatography Electrospray Ionization Mass Spectrometry
LC-MS	Liquid Chromatography Mass Spectrometry
LC-MS/MS	Liquid Chromatography Tandem Mass Spectrometry
LDS	Lithiumdodecylsulfat
LFQ	Label Free Quantitation
LTOF	Linear Time-of-Flight
LTQ	Linear Ion Trap

Lys-C	Endoproteinase Lys-C
M	Molar
μ	Micro
m	Milli
m/z	Mass to Charge Ratio
MALDI	Matrix-assisted Laser Desorption Ionization
MeOH	Methanol
MES	2-(<i>N</i> -morpholino)ethanesulfonic Acid
ml	Milliliter
mM	Millimolar
MOAC	Metal Oxide Affinity Chromatography
MRM	Multiple Reaction Monitoring
mRNA	Messenger Ribonucleic Acid
MS	Mass Spectrometry
MS/MS	Tandem Mass Spectrometry
MSA	Multistage Activation
MS ⁿ	Multiple Stage Mass Spectrometry
η/n	Nano
NaAcO	Sodium Acetate
NAS	<i>N</i> -Acetoxysuccinimide
NAT	<i>N</i> -Terminal Acetyltransferases
η LC	Nano Liquid Chromatography
p	Pico
PA	Phosphoric Acid
PAC	Phosphoramidate Chemistry
PAGE	Polyacrylamide Gel Electrophoresis
PKA	CAMP-Dependent Protein Kinase Catalytic Subunit Alpha
ppm	Part-per-Million
PSAQ	Protein Standard Absolute Quantification
PTM	Post-translational Modification
Q	Quadrupole
QqQ	Triple Quadrupole
Q-TOF	Quadrupole Time-of-Flight
Quant	Quantitation
RF	Radio Frequency
RNA	Ribonucleic Acid
RTOF	Reflectron Time-of-Flight
S	Synthesis Phase
S/N	Signal to Noise Ratio

A. Abbreviations

SA	Sinapic Acid
SDS	Sodium Dodecyl Sulfate
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
SILAC	Stable Isotope Labeling By Amino Acid In Cell Culture
SIMAC	Sequential Elution From IMAC
SRM	Selected Reaction Monitoring
Sulfo-NHS	<i>N</i> -Hydroxysulfosuccinimide
<i>T. brucei</i>	<i>Trypanosoma Brucei</i>
TFA	Trifluoroacetic Acid
TIC	Total Ion Chromatogram
TMT	Tandem Mass Tag
TOF	Time-of-Flight
TSS	Transcription Start Site
TTS	Transcription Termination Site
<i>v/v</i>	Volume Per Volume
<i>w/v</i>	Weight per Volume
WB	Western Blot
WT	Wild Type
XIC	Extracted Ion Chromatogram
<i>z</i>	Charge

B. Amino Acids and their Masses

Name	3-letter Code	1-letter Code	Monoisotopic Mass (Da)
Alanine	Ala	A	71.037114
Arginine	Arg	R	156.10111
Asparagine	Asn	N	114.04293
Aspartic Acid	Asp	D	115.02694
Cysteine	Cys	C	103.00919
Glutamic Acid	Glu	E	129.04259
Glycine	Gly	G	57.021464
Histidine	His	H	137.05891
Isoleucine	Ile	I	113.08406
Leucine	Leu	L	113.08406
Lysine	Lys	K	128.09496
Methionine	Met	M	131.04048
Phenylalanine	Phe	F	147.06841
Proline	Pro	P	97.052764
Serine	Ser	S	87.032029
Threonine	Thr	T	101.04768
Tryptophan	Trp	W	186.07931
Tyrosine	Tyr	Y	163.06333
Valine	Val	V	99.068414

A. Peptides and Proteins Sequences

1. Peptide Sequences

[Glu1]-Fibrinopeptide B EGVNDNEEGFFSAR

LGG-Lys(ac)-GELGGKGE

EAI^{pT}AAPFAK-NH₂

TW^{pYV}

LQK^{pSP}GPQRRER

FQSRG^{pSpS}PLQLN

SIINFEKL

2. Proteins Sequences

H2A_Tb427.07.2940: (*Trypanosoma Brucei*)

```
1  MATPKQAVKK  ASKGGSSRSV  KAGLIFPVGR  VGTLRRGQY  ARRIGASGAV
51  YMAAVLEYLT  AELLELSVKA  AAQQTKKTKR  LTPRTVTLAV  RHDDDLGALL
101 RNVTMSRGGV  MPSTLNKALAK  KQKSGKHAKA  TPSV
```

H2B_Tb427.10.10460: (*Trypanosoma Brucei*)

```
1  MATPKSTPAK  TRKEAKKTRR  QRKRTWNVYV  SRSLRSINSQ  MSMTSRTMKI
51  VNSFVNDLFE  RIAAEAATIV  RVNRKRTLGA  RELQTAVRLV  LPADLAKHAM
101 AEGTKAVSHA  SS
```

A. Peptides and Proteins Sequences

H3_Tb427.01.2430: (*Trypanosoma Brucei*)

1 MSRTKETART KKTITSKKSK KASKGSDAAS GVKTAQRRWR PGTVALREIR
51 QFQRSTDLLL QKAPFQRLVR EVSGAQKEGL RFQSSAILAA QEATESYIVS
101 LLADTNRACI HSGRVTIQPK DIHLALCLRG ERA

H4_Tb427.05.4170: (*Trypanosoma Brucei*)

1 MAKGKKSGEA KGSQKRQKKV LRENVRGITR GSIRRLARRG GVKRISGVII
51 DEVRGVLKSF VEGVVRDATA YTEYSRKKTV TAVDVVNALR KRGKILYGYA

H2A.Z_Tb427.07.6360: (*Trypanosoma Brucei*)

1 MSLTGDDAVP QAPLVGGVAM SPEQASALTG GKLGGKAVGP AHGKKGKGGK
51 GKRGGKTGGK AGRRDKMTRA ARADLNFPVG RIHSRLKDGL NRKQRCGASA
101 AIYCAALLEY LTSEVIELAG AAKAQKTER IKPRHLLLA I RGDEELNQIV
151 NATIARGGVV PFVHKSLEK I IKKSKRGS

H2B.V: (*Trypanosoma Brucei*)

1 MPPTKGGKRP LPLGGKGGK RPPGQTTKSS SSRKKSARR GKKQQRWDLY
51 IHRTLRLQVYK RGTLKAAVR VLSSFIEDMY GKIQAQAVHV ACINNVKTLT
101 AREIQTSARL LLPPELAKHA MSEGTKAVAK YNASREEAYS KVL

H3V: (*Trypanosoma Brucei*)

1 MAQMKKITPR PVRPKSVASR PIQAVARAPV KKVENTPPQK RHRWRPGTV
51 ALREIRRLQS STDFLIQRAP FRRFLREVVS NLKDSYRMSA ACVDIQEAX
101 ETYITSVFMD ANLCTLHANR VTLPFKDIQL ALKLRGERN

H4V_Tb427.02.2670 (*Trypanosoma Brucei*)

1 MAKGKRVGES KGAQKRQKKV LRDNVRGITR GSIRRLARRG GVKRISGVII
51 DEVRGVLKSF VEGVVRDATA YTEYSRKKTV TAAHVVFALR KRGKVLYGYD

TY1-H2A_Tb427.07.2820: (*Trypanosoma Brucei*)

1 MEVHTNQDPL DATPKQAVKK ASKGGSSRSV KAGLIFPVGR VGTLRRGQY
51 ARRIGASGAV YMAAVLEYLT AELLELSVKA AAQQTKKTKR LTPRTVTLAV
101 RHDDDLGALL RNVTMSRGGV MPSTLNKALAK KQKSGKHAKA TPSV

TY1-H2A_Tb427.07.2820: (*Trypanosoma Brucei*)

1 MEVHTNQDPL DSLTGDDAVP QAPLVGGVAM SPEQASALTG GKLGGKAVGP
51 AHGKGGKGGK GKRGGKTGGK AGRRDKMTRA ARADLNFPVG RIHSRLKDGL
101 NRKQRCGASA AIYCAALLEY LTSEVIELAG AAAKAQKTER IKPRHLLLAI
151 RGDEELNQIV NATIARGGVV PFVHKSLEKK IIKKSKRGS

P05132: cAMP-dependent protein kinase catalytic subunit alpha (Mouse)

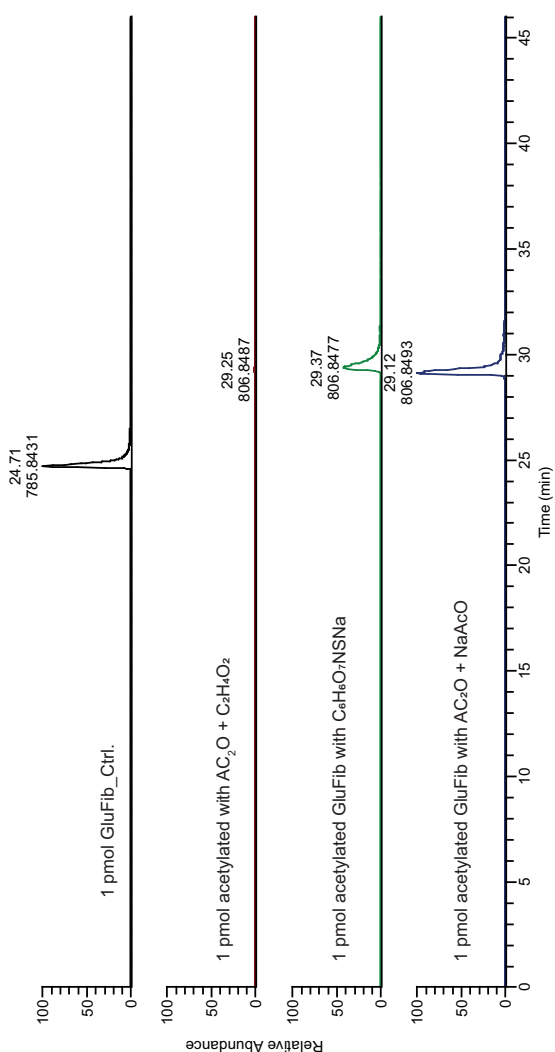
1 MGNAAAARKG SEQESVKEFL AKAKEDFLKK WETPSQNTAQ LDQFDRIKTL
51 GTGSFGRVML VKHKESGNHY AMKILDKQKV VKLKQIEHTL NEKRILQAVN
101 FPFLVKLEFS FKDNSNLYMV MEYVAGGEMF SHLRRIGRFS EPHARFYAAQ
151 IVLTFEYLHS LDLIYRDLKP ENLLIDQQGY IQVTDFGFAK RVKGRWTWTLK
201 GTPEYLAPEI ILSKGYNKAV DWWALGVLIY EMAAGYPPFF ADQPIQIYEK
251 IVSGKVRFPS HFSSDLKDLL RNLLQVDLTK RFGNLKNGVN DIKNHKWFAT
301 TDWIAIYQRK VEAPFIPKFK GPGDTSNFDD YEEEEIRVSI NEKCGKEFTE
351 F

P02666 : Beta-casein Bos taurus (Bovine)

1 MKVLILACLV ALALARELEE LNVPGEIVES LSSSEESITR INKKIEKFQS
51 EEQQQTEDEL QDKIHPFAQT QSLVYFPFGP IPNSLPQNIPLTQTPVVVP
101 PFLQPEVMGV SKVKEAMAPK HKEMFPKYP VEPFTESQSL TLTDVENLHL
151 PLPLLQSWMH QPHQPLPPTV MFPPQSVLSL SQSKVLPVPQ KAVPYPQRDM
201 PIQAFLLYQE PVLGPVRGPF PIIIV

B. Site-specific Acetylation Degree Quantitation

Comparison Between Using Different Reagents for Acetylation

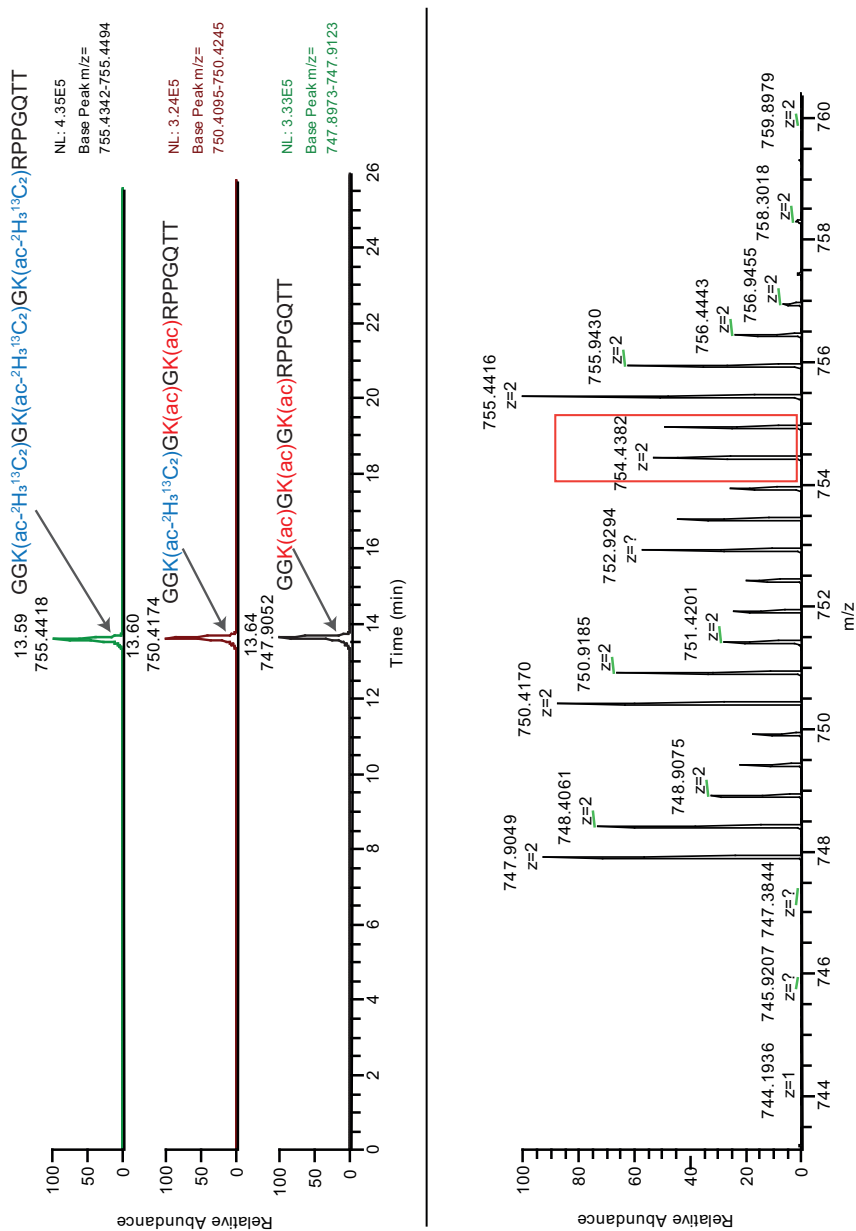


Acetylation of 1 pmol GluFib using different reagents. (*Black*) GluFib control sample EGVNDNEEGFFSAR, 2+ ($m/z = 785.8423$), (*red*) GluFib treated with 50 μ l acetic acid ($C_2H_4O_2$) and 10 μ l acetic anhydride (AC_2O), (*green*) GluFib treated with 20 μ l Sulfo-NHS ($C_6H_6O_7NSNa$) prepared with 100 mM sodium acetate ($NaAcO$) pH 6, and (*green*) GluFib treated with 10 μ l acetic anhydride (AC_2O) prepared with 1M sodium acetate ($NaAcO$) pH

B. Site-specific Acetylation Degree Quantitation

6.8. All samples were incubated at 25°C for 2 hours while shaking causing an acetylation of the N-terminal of GluFib EGVNDNEEGFFSAR, 2+ ($m/z = 806.8475$)

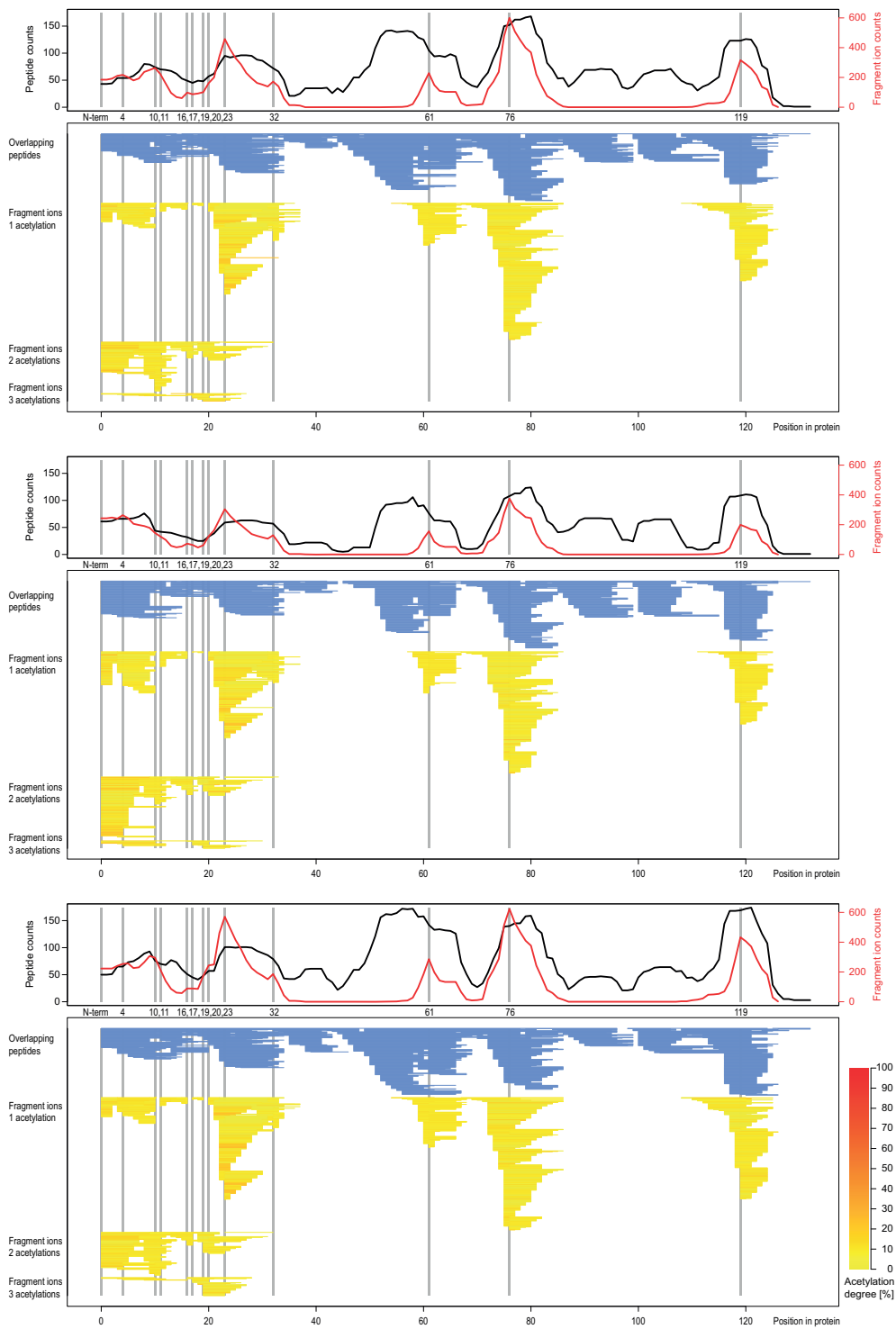
Acetylation Using $^{13}\text{C}_4\text{d}_6$ Labeled Acetic Anhydride and Acetic Acetate- d_3



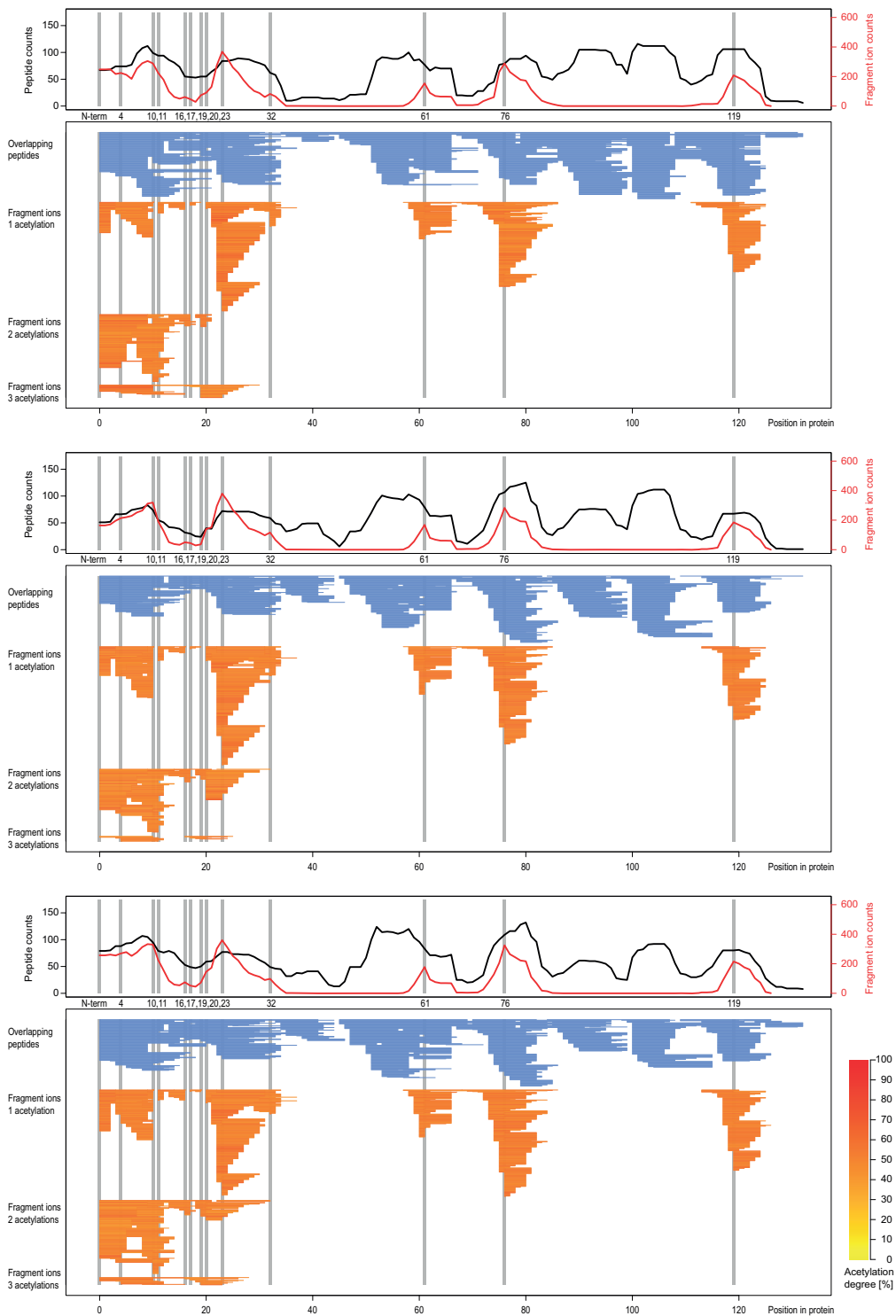
Upper panel: XICs of GGKKGKGRPPGQTT (G14-T26) from in-gel acetylated H2B.V from *T. brucei* using acetic anhydride- $^{13}\text{C}_4\text{d}_6$. (Black) $\text{GGK}_{\text{ac}}\text{GK}_{\text{ac}}\text{GK}_{\text{ac}}\text{RPPGQTT}$, 2+ ($m/z = 747.9052$), (red) $\text{GGK}_{\text{ac-}^2\text{H}_3^{13}\text{C}_2}\text{GK}_{\text{ac}}\text{GK}_{\text{ac}}\text{RPPGQTT}$, 2+ ($m/z = 750.4174$), and (green) $\text{GGK}_{\text{ac-}^2\text{H}_3^{13}\text{C}_2}\text{GK}_{\text{ac-}^2\text{H}_3^{13}\text{C}_2}\text{GK}_{\text{ac-}^2\text{H}_3^{13}\text{C}_2}\text{RPPGQTT}$, 2+ ($m/z = 755.4416$). Lower panel: Experimental isotopic pattern of GGKKGKGRPPGQTT, 2+. The use of $^{13}\text{C}_4\text{d}_6$ labeled acetic anhydride cause a mass shift of 5 Da between the light and heavy peptide. Despite the good separation between the two isoforms, the use of this reagent makes the data analysis complicated and unreliable.

Quantification Patchworks for Recombinant *T. Brucei* H3 Histone

a)

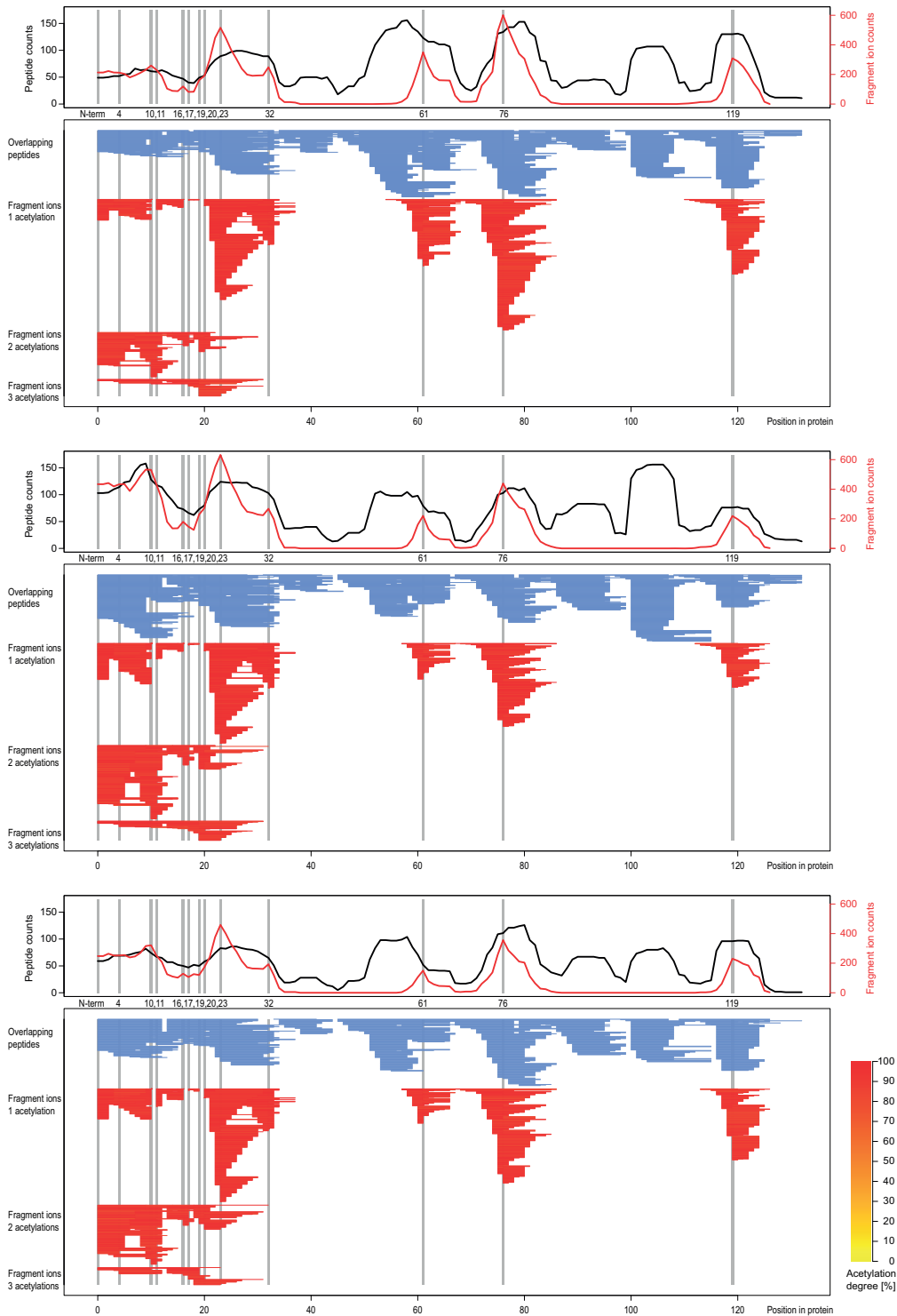


b)



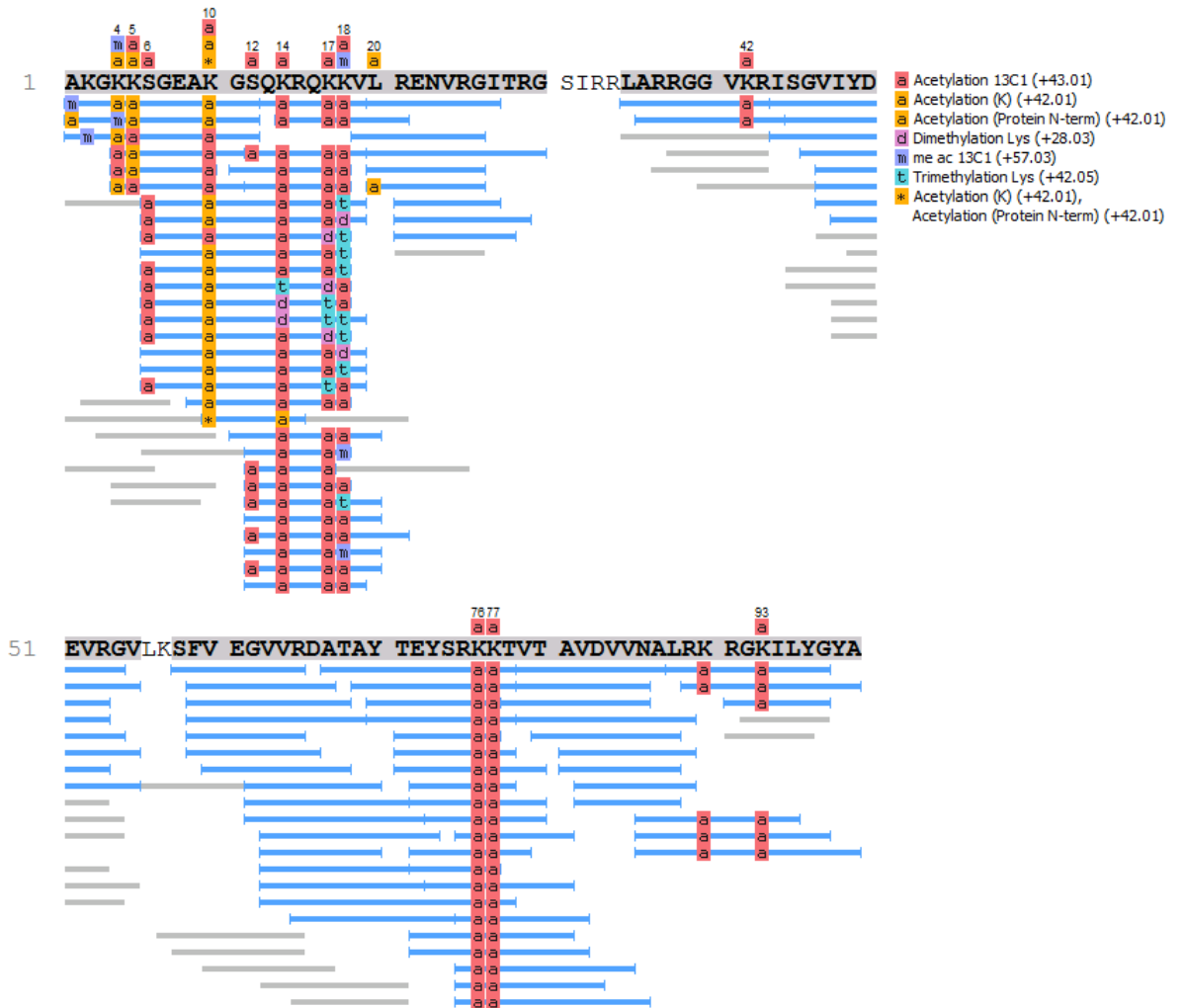
B. Site-specific Acetylation Degree Quantitation

c)



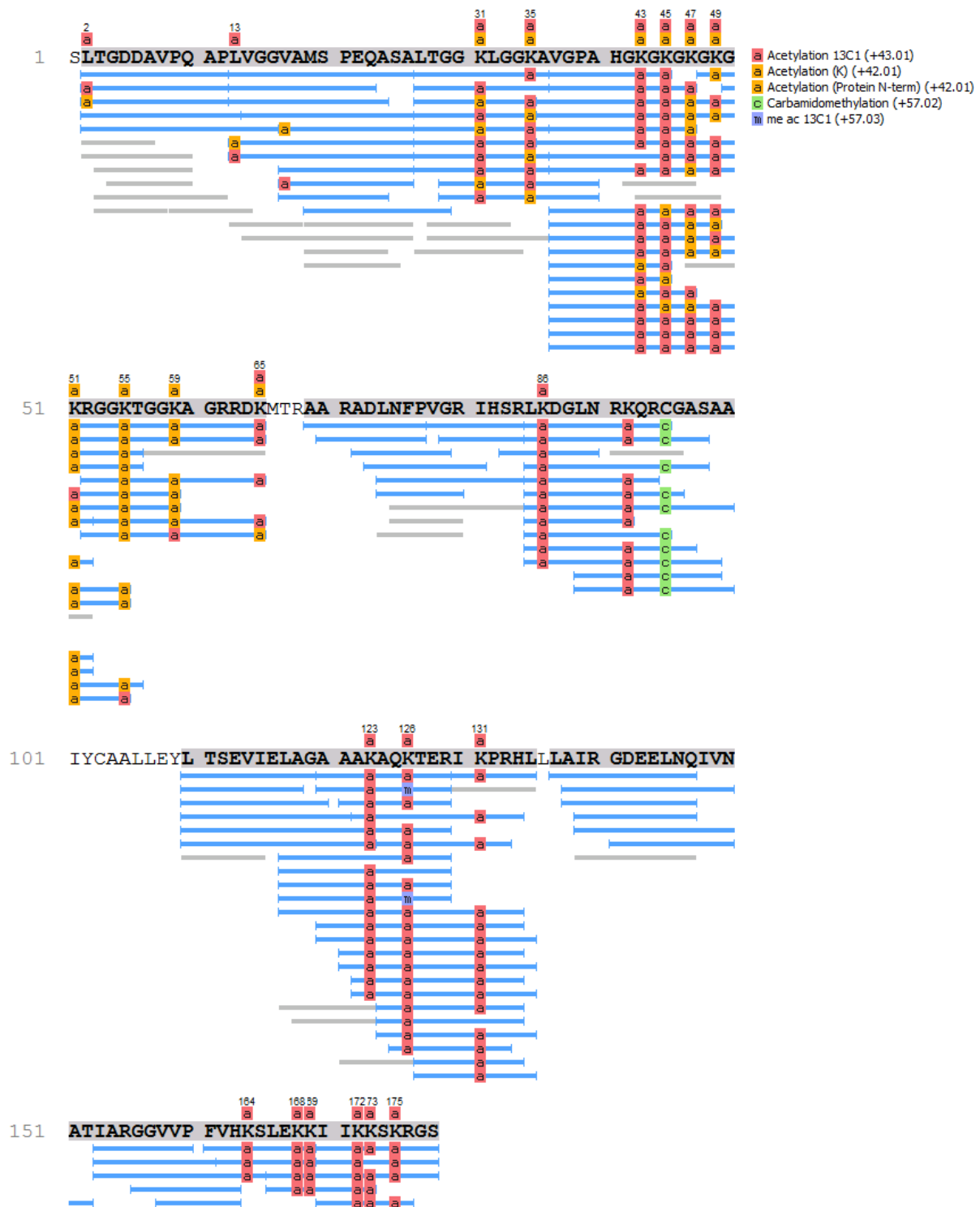
(a) 10% light, 90% heavy, (b) 50% light, 50% heavy, (c) 90% light, 10% heavy. **Upper panels:** number of peptides (black) and fragment ions used for calculation of the acetylation degree (red) per amino acid position. **Patchwork panels:** overlapping peptides (blue lines) and fragment ions used for quantitation (yellow to red lines, acetylation degree is color coded). Acetylation sites are highlighted (gray vertical bars). Fragment ions are only counted and shown once (relevant only for fragments with more than one acetylation site).

Sequence Coverage of $^{13}\text{C}_1$ Chemically Modified Histones from *T. Brucei*

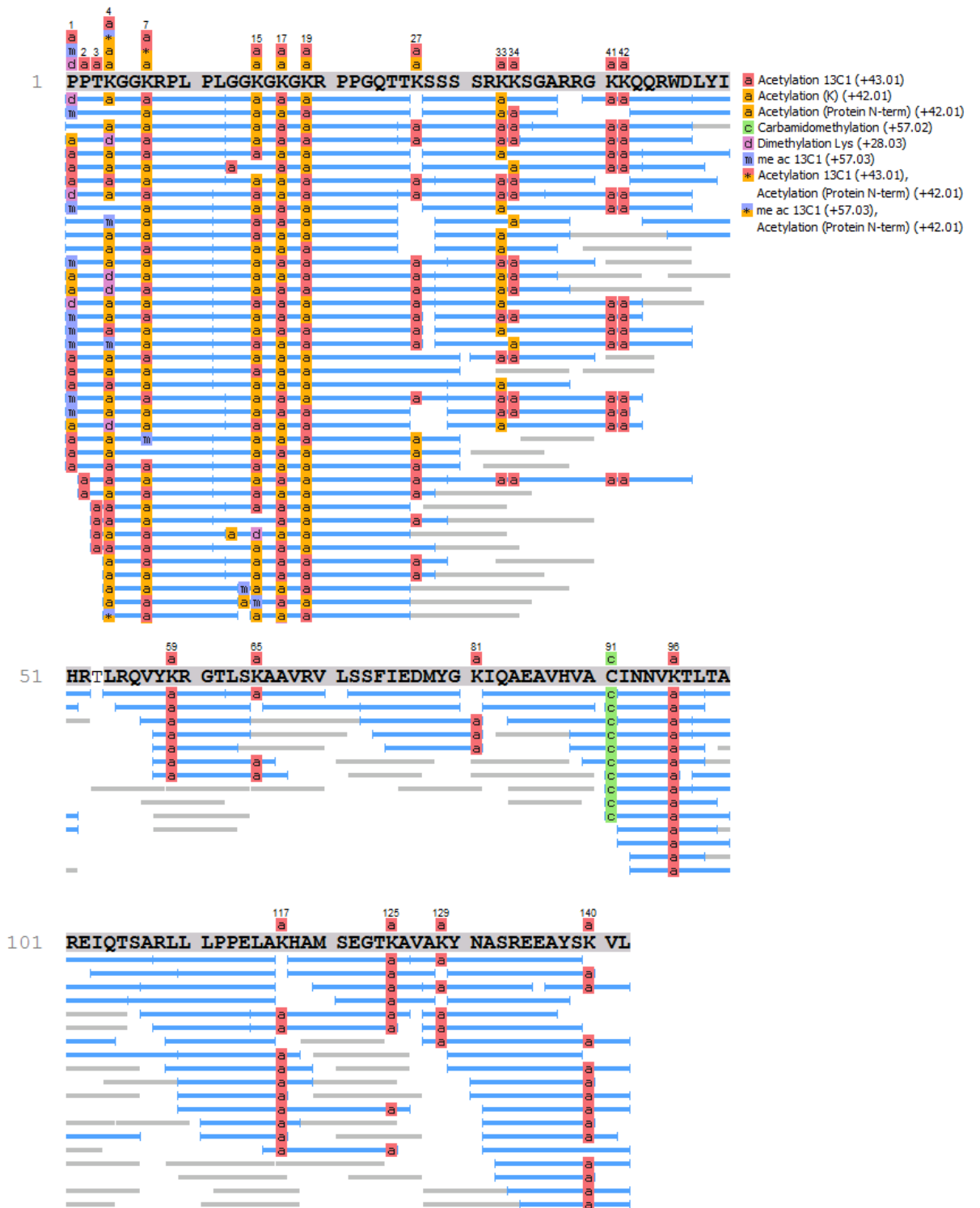


In-gel acetylated histone H4 with $^{13}\text{C}_1$ -acetic anhydride.

B. Site-specific Acetylation Degree Quantitation

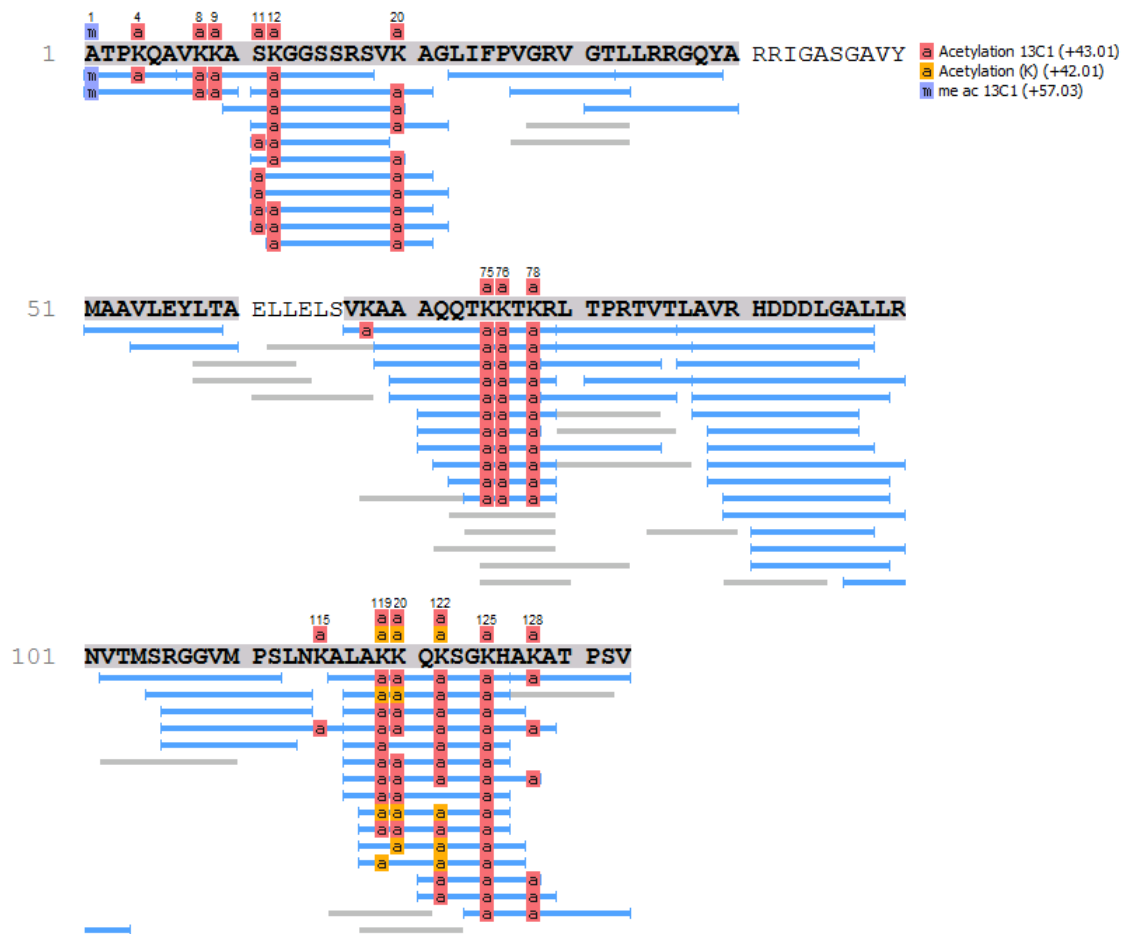


In-gel acetylated histone H2A.Z with $^{13}\text{C}_1$ -acetic anhydride.



In-gel acetylated histone H2B.V with $^{13}\text{C}_1$ -acetic anhydride.

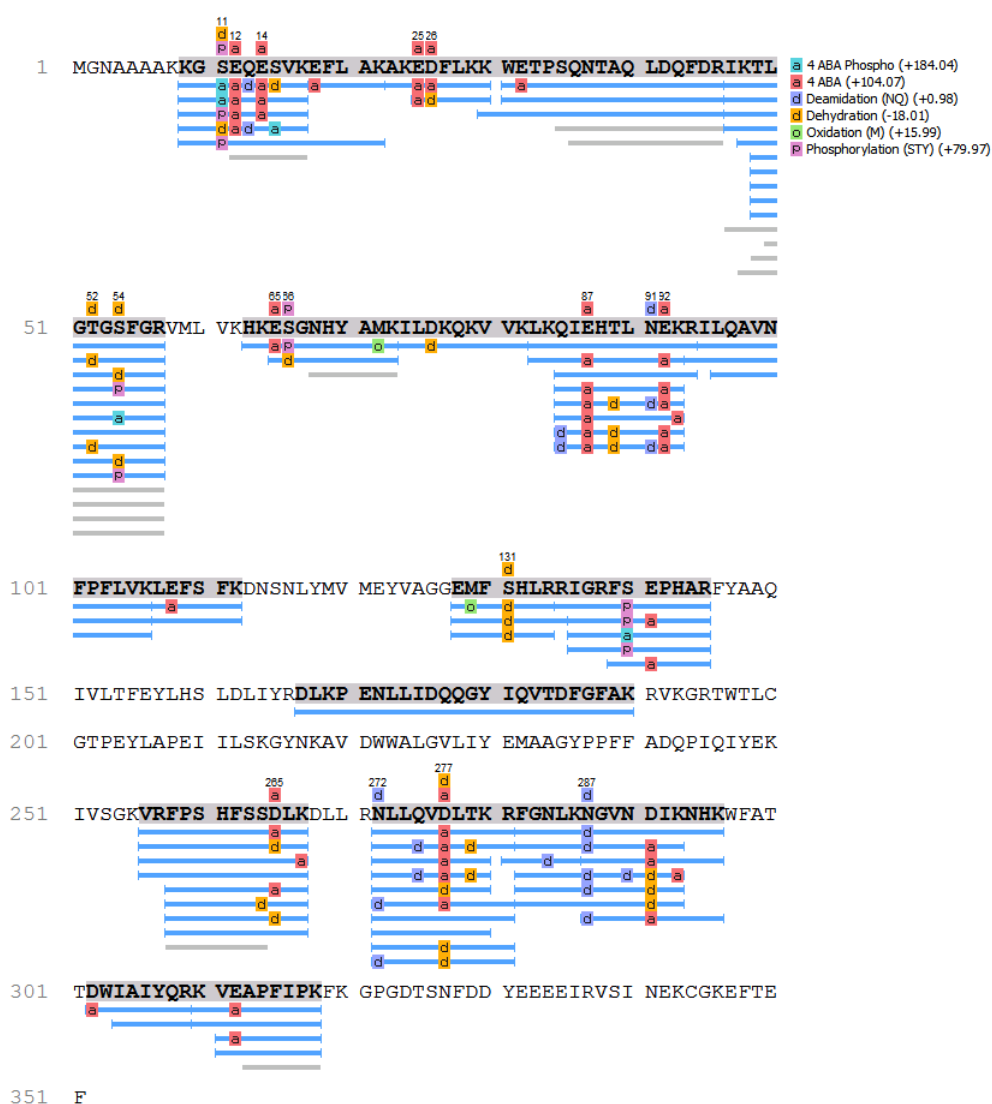
B. Site-specific Acetylation Degree Quantitation



In-gel acetylated histone H2A with $^{13}\text{C}_1$ -acetic anhydride.

C. Carbodiimide-mediated 4-ABA Labeling of Phosphosites

2- In-gel 4-ABA Chemical Labeling



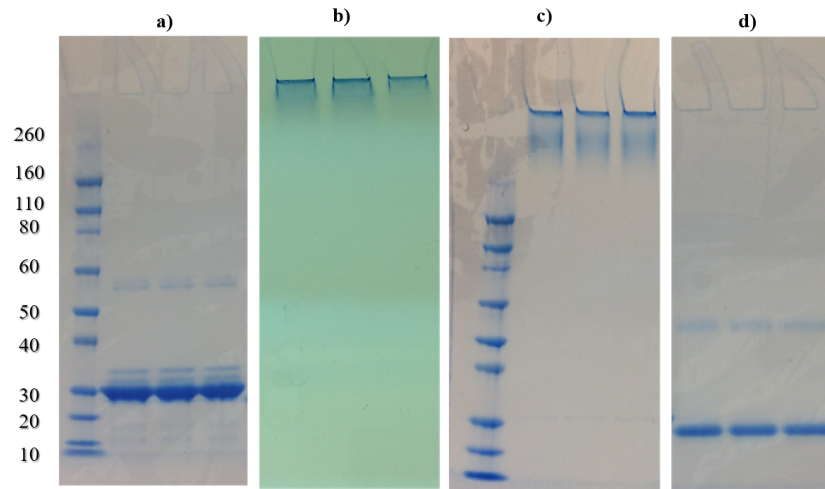
Sequence coverage of tryptic digest of PKA chemically modified in-gel with 4-ABA. Sample was incubated for 4 hours at 25°C while shaking. Incubating the sample for this period resulted in a) many sites have not been modified and b) an incomplete reaction in many of the modified sites.

C. Carbodiimide-mediated 4-ABA Labeling of Phosphosites



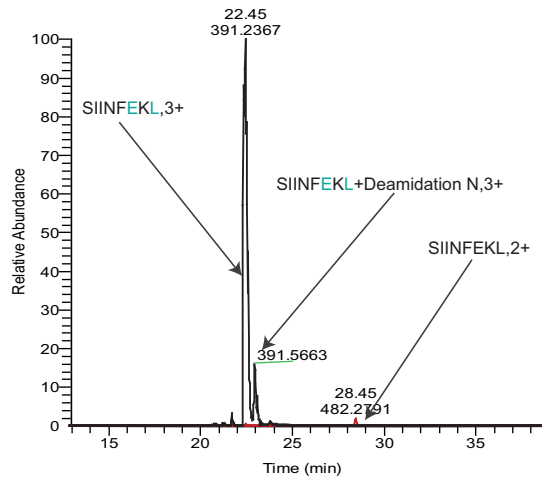
Sequence coverage of tryptic digest of PKA chemically modified in-gel with 4-ABA. Sample was incubated for 5 hours at 25°C while shaking. Incubating the sample for this period resulted in a) many sites have not been modified and b) an incomplete reaction in many of the modified sites.

3- In-solution 4-ABA Chemical Labeling



Gel bands of β -casein before and after 4-ABA treatment, from different gels, using different EDC concentrations to activate the desired residues; **a)** β -casein control sample, **b)** 4-ABA modified β -casein using 250 mM EDC, **c)** 4-ABA modified β -casein using 175 mM EDC, and **d)** 4-ABA modified β -casein using 125 mM EDC. Modifying the protein in gel using EDC concentrations higher than 125 mM cause the protein to cross-link and aggregate in the SDS-PAGE gel well.

3- 4-ABA Labeling of Carboxyl Groups



Reaction completion of 4-ABA labeled SIINFEKL **a)** Red: completely modified SIINFEKL,3+ ($m/z = 391.2373$, RT = 22.45) and deamidated SIINFEKL,3+ ($m/z = 391.5663$, RT =), black: unmodified SIINFEKL,2+ ($m/z = 482.2789$, RT = 28.45).

Publications & Conferences

ElBashir, R., Vanselow, J. T., Kraus, A., Janzen, C. J., Siegel, T. N., and Schlosser, A.: Fragment Ion Patchwork Quantification for Measuring Site-Specific Acetylation Degrees. In: *Analytical Chemistry* 87(19), pp 9939-45, 2015

ElBashir, R., Vanselow, J. T., Kraus, A., Janzen, C. J., Siegel, T. N., and Schlosser, A.: Site-specific Quantitation of the Lysine Acetylation Degree: A Generic MS2-based Method Using $^{13}\text{C}_1$ -acetyl Labeling and Unspecific Proteases. In: *63rd Annual American Society for Mass Spectrometry, St. Louis, MO. USA. May 2015*

ElBashir, R., Vanselow, J. T., and Schlosser, A.: Determination of Site Specific Acetylation Degree by All-in-One Fragmentation of Deuteroacetylated Isotopologues. In: *47th Annual Meeting for the German Society of Mass Spectrometry. Frankfurt, Germany. March 2014*

ElBashir, R., Vanselow, J. T., and Schlosser, A.: Mass Spectrometry-Based Method for the Quantitative Analysis of Histone Acetylation. In: *8th International Symposium of Graduate School of Life Sciences University of Würzburg, Germany. October 2013*

ElBashir, R and Digel, I.: Effect of Nitric Oxide on Hydrogels. In: *International Congress of Students and Young Scientists. Almaty, Kazakhstan. March 2009*

Curriculum Vitae