2008

Histone Modifications in Trypanosoma Brucei

Veena Mandava

Histone Modifications in *Trypanosoma brucei*

A Thesis Presented to the Faculty of

The Rockefeller University

in Partial Fulfillment of the Requirements for

the degree of Doctor of Philosophy

by

Veena Mandava

June 2008

Histone Modifications in *Trypanosoma brucei*

Veena Mandava, Ph.D.

The Rockefeller University 2008

Trypanosoma brucei maintains an infection in its mammalian host by switching its surface antigen, thus evading host antibodies, in a process known as antigenic variation. Variant surface glycoproteins (VSG), the surface antigen, are expressed from genes located at as many as 20 Expression Sites (ES), which are present in subtelomeric regions of several chromosomes. How trypanosomes maintain monoallelic expression of *VSG* is a major question in trypanosome biology. Several theories have been proposed, including the presence of an Expression Site Body (a dedicated 'transcription factory'), regulation of RNA elongation, and some kind of transcriptionally restrictive chromatin structure at inactive ES.

In other organisms, the role of histone posttranslational modifications (PTMs) in influencing transcriptional states has been well-documented. To begin to understand the role of chromatin structure in the regulation of gene expression in trypanosomes, I developed a protocol for the purification of histones from mammalian-infective trypanosomes and identified PTMs using Edman degradation and mass spectrometry. I found that the N-termini of H4 and, possibly, H3 have a number of posttranslational modifications (PTMs), while H2A and H2B, in contrast, have relatively few. I also found a series of acetylated

lysines at the C-terminus of H2A. Interestingly, I observed that alanine 1 of H2A, H2B, and H4 is monomethylated, a novel modification found in trypanosomatids.

Next, I attempted to identify the targets of the putative histone deacetylase TbSIR2RP-1, a SIR2-related protein. Yeast SIR2 plays a key role in establishing heterochromatin at subtelomeric loci. TbSIR2RP-1 was chosen as a possible mediator of ES silencing, although this was later shown not to be the case. I showed that, unlike its yeast counterpart, TbSIR2RP-1 may not be an H4 deacetylase.

Finally, I began to characterize H3 K4 methylation, which, in other organisms, is traditionally associated with transcriptionally active chromatin. I showed that nucleosomes containing the histone variant H2BV are enriched for H3 K4 methylation. I initially hypothesized that this is due to ubiquitination at a conserved lysine at the H2BV C-terminus, but find that this is not the case. Instead, I propose that replacement of major histones H2A/H2B with H2AZ/H2BV acts as a trigger to stimulate H3 K4 methylation, which represents an alternative pathway leading to H3 methylation, compared to what is observed in yeast.

Acknowledgments

Graduate school was an intellectual and emotional challenge, and I am indebted to George Cross for navigating me through the past four years. I am especially grateful to have an advisor who buffered the inevitable disappointments that I experienced in my graduate school career by being accessible and gently nudging me onward. George is committed to educating his graduate students and pushing us to perform our work with rigor. Under his guidance, I have attained a level of scientific competence and an expectation of excellence that I will carry with me in my future research endeavors.

I would like to thank the members of my faculty advisory committee, Kirk Deitsch, David Allis, and Brian Chait, for their support and guidance over these past few years. Thanks especially to Dr. Chait for his tutorials on mass spectrometry. I am also grateful to C.C. Wang, my external advisor from U.C.S.F., for participating in my thesis defense.

I am grateful to my labmates, past and present, for their help in the lab: Joanna Lowell, Oliver Dreesen, Magda Kartvelishvili, Hee-Sook Kim, Bibo Li, Christian Janzen, Eiji Okubo, Jenny Li, Nicolai Siegel, and Luisa Figueiriedo. I would particularly like to thank Christian Janzen and Joanna Lowell, whose work provided the foundation for my project. Thanks also to Sandra Hake, a former postdoc in the Allis lab, for performing the HPLC necessary for the protein purification discussed here.

I would like to acknowledge the facilities available to us at Rockefeller University, especially the Bio-Imaging and Proteomics Resource Centers. Joseph Fernandez, Haiteng Deng, and Hongying Zhou performed the mass spectrometry that is discussed in this thesis. Thanks also to Alison North at the Bio-Imaging Facility for her expertise, particularly on the handling of trypanosomes for microscopy.

I would not have come this far in the MD/PhD program without the help of Olaf Andersen, Ruthie Gotian, Renee Horton, and Elaine Velez. I am grateful for their support and understanding as I struggle through this process.

Finally, I would like to acknowledge the support of my family, friends, my boyfriend Mohit, and my dog Ladu.

Table of Contents

List of Figures

Chapter 1. Introduction

Chapter 3. Purification of Histones from *T. brucei* Bloodstream Forms and Identification of Posttranslational Modifications

List of Tables

List of Abbreviations

BDF1	Bromodomain Factor 1
BF	Bloodstream Form
BRE1	Brefeldin A Sensitive
CHD1	Chromo-ATPase/helicase-DNA binding domain 1
CHP1	Heterochromatin-Associated Chromodomain Protein
CLR	Cryptic Loci Regulator
COMPASS	Complex of Proteins Associated with SET1
CTD	C-terminal domain
DOT1	Disruptor of Telomeric Silencing
ES	Expression Site
ESAG	Expression Site Associated Gene
ESB	Expression Site Body
FACT	Facilitates Chromatin Transcription
H2AZ, HTZ1	Histone H2A Variant
H2Bv	Histone H2B Variant
H3.3, H3V	Histone H3 Variants
HAT	Histone Acetyltransferase
HDAC	Histone Deacetylases
HIRA	Histone Regulation
HMT	Histone Methyltransferase
HP1	Heterochromatin Protein 1
HPLC	High Performance Liquid Chromatography
ISWI	Imitation SWI
JBP	J Binding Protein
Kac	Acetylated Lysine
Kme1, me2, me3	Mono-, Di-, or Tri-methyl Lysine
LC/ESI-MS/MS	Liquid Chromatography/Electrospray tandem MS
LSD1	Lysine Specific Demethylase
MALDI-TOF	Matrix Assisted Laser Desorption/Ionization – Time-of-Flight
MS	Mass Spectrometry
ORF	Open Reading Frame
PF	Procyclic Form
PA	Propionic Anhydride
PAF1	RNA polymerase II Associated Factor 1
PTM	Posttranslational Modification
QqTOF	Quadropole-quadropole Time-of-Flight
RAD6	Radiation Sensitive
RITS	RNA-induced initiation of transcriptional gene silencing
RNAi	RNA interference
SAGA	Spt-Ada-Gcn5 Acetyltransferase
SET	Su(var), E(z), Trx

SIR	Silent Information Regulator
SLIK	SAGA-like
SNF	Sucrose non-fermenting
Su(var)	Suppressor of Variegation
SWI	Switching mating type
SWR1	Swi2/Snf2-related ATPase
TAF	TBP Associated Factor
TAU	Triton Acetic Acid Urea
TbAUK	*T. brucei* Aurora kinase
TbSIR2RP	*T. brucei* SIR2-related protein
TcCRK1	*T. cruzi* Cdc2-related Kinase
UBP8	Ubiquitin-specific Processing Protease
UTR	Untranslated Region
VSG	Variant Surface Glycoprotein

Chapter 1. Introduction

Chromatin

In eukaryotic organisms, the genome is packaged with proteins into a condensed structure in the nucleus known as chromatin. The basic unit of chromatin is the nucleosome, formed by 146 bp of DNA wrapped around an octamer containing two copies of each of the core histones, H2A, H2B, H3, and H4. Nucleosomes form "beads on a string," and, with histone H1, form a 30 nm condensed fiber, which may be further condensed to form chromosomes. The structure of chromatin was once thought to be static and inert, forming only a physical framework for DNA to be packaged in the nucleus. Currently, the nucleosome is thought to be a dynamic structure that regulates the access of DNA to protein factors, influencing a variety of biological processes.

Chromatin may be divided into euchromatin or heterochromatin. Heterochromatin is transcriptionally inactive, microscopically condensed, has relatively few ORFs, is rich in repetitive DNA, and is replicated late in S phase; euchromatin is transcriptionally active, uncondensed, rich in ORFs, and replicates early in S phase. Understanding how these domains are formed and how the barriers between them are maintained are important fields of study in chromatin biology.

Histone Posttranslational Modifications

Core histones contain many posttranslational modifications (PTMs), including phosphorylation, acetylation, methylation, and ubiquitination. Histone PTMs influence a wide range of biological processes including transcriptional regulation and heterochromatin formation, nucleosome assembly, and DNA replication and repair [1]. These processes are mediated by the effector proteins that bind to modified histones [2]. Effector proteins with chromodomains bind methylated lysines [3-5], while proteins with bromodomains bind acetylated lysines [6-8].

Perhaps the most well-studied biological process affected by histone PTMs is gene expression. Histone acetylation, for example, is generally associated with transcriptionally permissive chromatin, or euchromatin [9]. In *Saccharomyces cerevisiae*, chromatin immunoprecipitation was used to show that nucleosomes associated with actively transcribed genes were enriched for acetylated lysines at the H3 and H4 N-termini [10]. Histone acetylation affects transcription both by neutralizing the positive charge of lysines and by acting as a binding platform for effector proteins. N-terminal acetylation is thought to disrupt interactions that stabilize the nucleosome, making the DNA more accessible to transcription-promoting complexes [11]. Also, acetylated lysines provide binding sites for a variety of effector proteins, including double bromodomain-containing human $TAF_{II}250$ [12]. $TAF_{II}250$, a subunit of the TFIID complex, binds

diacetylated H4, which allows for the recruitment of TFIID to sites of acetylation and subsequent assembly of the transcription initiation.

Histone methylation also affects transcriptional activity of chromatin, but these effects are dependent on the particular lysine that is methylated: H3 K9 and H4 K20 methylation are associated with silent chromatin, or heterochromatin, and H3 K4 and K79 methylation are associated with euchromatin [1]. The mechanisms by which methylated lysines promote transcription or induce heterochromatin formation are discussed.

Establishment of Heterochromatin

Two well-documented pathways illustrate the role of histone PTMs in heterochromatin formation. In the first, mutations in suppressor of variegation 3-9 (Su(var)3-9) and heterochromatin protein 1 (HP1) in *Drosophila* were shown to suppress position effect variegation, a phenomenon in which normally active genes may be silenced when placed in regions adjacent to heterochromatin, suggesting that the wild-type gene products were involved in heterochromatin formation [13,14]. In mammalian cells, Suv39h1, an Su(var)3-9 homologue, was shown be an H3 K9 histone methyltransferase (HMT) [15], which, in turn, creates a binding site for HP1 [3-5]. HP1 bound to methylated H3 K9 recruits a variety of factors including BRG1 [16], a chromatin remodeling factor that may generate heterochromatin, as well as Suv39h1 [17], thus propagating H3 K9 methylation

and heterochromatin assembly [16,18]. In addition to its role in establishing heterochromatin, HP1 is recruited to promoters in euchromatin for local gene silencing [19]. In *Schizosaccharomyces pombe*, H3 K9 methylation is dependent on Clr3, an H3 histone deacetylase (HDAC), demonstrating that multiple steps in the pathway leading to heterochromatin formation involve histone PTMs [20].

The constitutive heterochromatin assembly pathway involving H3 K9 methylation is mediated by RNAi machinery [21,22]. In *S. pombe*, deletion of components of the RNAi machinery resulted in loss of H3 K9 methylation and de-repression of reporter genes integrated into the centromere, indicating a role for RNAi in transcriptional repression by H3 K9 methylation [23]. Centromeric repeats are sufficient to nucleate heterochromatin formation [22]. Small RNAs are generated at centromeric repeats by RNAi machinery, which, in turn, are used to target the RITS (RNA-induced initiation of transcriptional gene silencing) complex to the centromere [21]. The RITS complex is thought to recruit Clr4, an H3 K9 HMT, as well as HDACs and other silencing factors. Recuitment of Clr4 results in the propagation of H3 K9 methylation and binding of Swi6, an HP1 homologue. The RITS complex also depends on H3 K9 methylation for binding through one of its components, chromodomain-containing Chp1. The main regions of heterochromatin in *S. pombe* – the centromere, the mating-type loci, and the telomere – all produce small RNAs from their repetitive sequences, which are presumably used to nucleate heterochromatin assembly [24].

In the second pathway, heterochromatin is established with the help of SIR (silent information regulator) proteins, which were first studied in *S. cerevisiae*. The SIR2/3/4 complex is recruited to the telomere through the interaction of SIR4 with RAP1, a telomere binding protein [25-28]. SIR2 is an HDAC for K9 and K14 on H3 and K16 on H4 [29]. SIR2 plays a valuable role in the establishment of heterochromatin since SIR2 deletion mutants demonstrate defects in silencing at the telomere [30], the rDNA locus [31,32], and the mating loci [33]. SIR2 creates a binding platform for SIR3, which interacts with deacetylated H4 K16 [34,35], and recruits additional SIR4 and SIR2 to propagate heterochromatin assembly [36]. In this way, heterochromatin is nucleated at the telomere and propagates through the interaction of SIR proteins with the H3 and H4 N-termini [37,38]. In *S. cerevisiae* and other organisms, histone PTMs clearly play a role in determining whether chromatin is in a transcriptionally permissive or restrictive configuration.

Chromatin Remodeling during Transcription Elongation

In addition to transcription initiation, it is important to review transcription elongation by RNA pol II in yeast because a number of histone PTMs are involved in this process (reviewed in [39]). The C-terminal domain (CTD) of RNA pol II consists of a heptapeptide repeat with serines at positions 2 and 5, the phosphorylation of which is required for the transition from transcription initiation

to elongation [40]. Phosphorylation of serine 5 occurs at the 5´ end of the ORF, while phosphorylation of serine 2 occurs at a later phase of elongation [41]. The phosphorylated CTD creates a binding platform for factors at different stages of the elongation process [42]. An RNA pol II-associated factor (Paf1) associates with the serine 5-phosphorylated CTD [43-45] and participates in the recruitment of HMT Set1 (discussed later) [46-48]. Paf1 also recruits FACT [45], a histone chaperone complex, which destabilizes nucleosomes by removing H2A/H2B dimers out of the path of RNA pol II [49]. FACT also deposits H2A/H2B dimers to facilitate nucleosome re-assembly following transcription. The factors that are recruited to the elongating RNA pol II illustrate that histone modification and chromatin remodeling are integral processes to transcription elongation.

Methylation of Histone H3 Lysine 4

H3 K4 methylation has been implicated in transcriptional activation and elongation [39,50]. In *S. cerevisiae*, chromatin immunoprecipitation with specific antibodies against mono-, di-, and tri-methyl H3 K4 coupled to DNA microarrays showed the methylation status of H3 K4 across an average gene [10]. Specifically, trimethylation of K4 was enriched at the 5´ end of transcribed genes, dimethylation in the middle, and monomethylation at the 3´ end of genes. The same study showed that nucleosome occupancy is low at the promoters of active genes. The same methods were used in higher eukaryotes to show that

trimethylation of H3 K4 was enriched at the 5´ end of actively transcribed genes, although patterns of dimethylation appear to differ [51-53]. Inactive genes were hypomethylated at H3 K4, demonstrating that H3 K4 methylation is correlated with transcriptional activity [51].

The enzymes that methylate and demethylate H3 K4 have been identified, and deletion mutants of these enzymes solidify the relationship between H3 K4 methylation and transcriptional activity. In *S. cerevisiae*, H3 K4 is methylated by Set1, named for the presence of a SET domain, which is common to many HMTs [54]. Deletion of Set1 resulted in an expected loss of H3 K4 methylation as well as decreased transcription of 80% of yeast ORFs [55]. H3 K4 is demethylated by LSD1, a known member of a number of corepressor complexes [56]. Consistent with this, RNAi inhibition of LSD1 showed transcriptional de-repression in HeLa cells.

Like other histone PTMs, H3 K4 methylation mediates its effects by acting as a platform for the binding of effector proteins. For example, BPTF, a member of the NURF ATP-dependent chromatin remodeling complex, preferentially binds trimethylated H3 K4, thereby recruiting NURF to nucleosomes enriched for H3 K4 trimethylation [57,58]. Also, chromodomain-containing Chd1, a chromatin remodeling enzyme, binds methylated H3 K4 [59]. A component of SAGA and SLIK HAT complexes, Chd1 recruits these multisubunit coactivator complexes to sites of H3 K4 methylation. These studies and others show that H3 K4

methylation promotes transcription by recruiting chromatin remodeling enzymes and HAT coactivator complexes to its target genes.

Finally, H3 K4 methylation was implicated in transcriptional elongation when it was shown that elongating RNA pol II interacts with a Complex of Proteins Associated with Set1 (COMPASS) [46-48]. COMPASS binds the CTD when the CTD is phosphorylated at serine 5, but not serine 2. These studies show that COMPASS is associated with a specific form of the RNA pol II elongation complex, one that is present at the promoter region and 5´ end of the ORF, which corresponds with the localization of trimethyl H3 K4. Some speculate that H3 K4 trimethylation by Set1 is required for the regulation of transcription initiation and promoter clearance, possibly through the recruitment of additional transcription factors [60]. The results also imply, as the authors suggest, that methylation may serve as a memory of those genes that have been recently transcribed, a mechanism that allows for the propagation of the active state through cell division.

H2B K123 Ubiquitination is Required for H3 K4 and K79 Methylation in Yeast

Histone PTMs have been shown to facilitate or impede the placement of additional PTMs on the same or neighboring histones, in a phenomenon dubbed cross-talk [61]. One well-documented example of cross-talk is the relationship between H2B K123 ubiquitination and H3 K4 and K79 methylation in *S.*

cerevisiae [62-64]. Mutation of H2B lysine 123 to arginine abolishes H3 K4 and K79 methylation. However, mutation of H3 lysine 4 to arginine does not affect H2B ubiquitination, indicating that H2B K123 ubiquitination is required for H3 K4 and K79 methylation in a unidirectional pathway. Only 5% of H2B is ubiquitinated, while a far higher percentage of H3 is methylated at K4, indicating that the effect of ubiquitination extends to neighboring nucleosomes. The mechanism by which H2B ubiquitination signals for H3 K4 and K79 methylation is currently unknown (reviewed in [65]).

Monoubiquination of H2B is achieved through the actions of Rad6, an E2 ubiquitin-conjugating enzyme, and Bre1, an E3 ligase [63,66]. Loss of either of these proteins or the interactions between them result in loss of ubiquitination and, subsequently, H3 K4 methylation. The Paf1 complex is also required for H2B ubiquitination [47,67]. Paf1 may mediate H2B ubiquitination by facilitating an interaction between Rad6 and phosphorylated RNA pol II CTD, since mutation of the CTD or absence of serine 5 phosphorylation resulted in reduced levels of ubiquitination [68]. The interaction between Rad6 and the elongating form of RNA pol II suggests that H2B ubiquitination is involved in transcriptional elongation.

Removal of ubiquitin from H2B is achieved by deubiquitinating enzyme Ubp8, a member of the SAGA and SLIK HAT coactivator complexes [69,70]. Ubp8 was shown to be associated with upstream regulatory elements of a SAGA/SLIK-dependent gene during activating conditions, while Rad6 was

present during non-activating conditions. Also, deletion of Ubp8 results in decreased H3 K4 trimethylation at active genes and decreased expression of a SAGA/SLIK-dependent gene. One possible theory for the role of Udp8 is that it recruits SAGA to sites of H2B ubiquitination, and deubiquitination is required for subsequent transcriptional activation [65].

One of the aims of this thesis is to determine whether the relationship between H3 K4 and K79 methylation and H2B K123 ubiquitination is conserved in evolutionarily divergent organisms.

Histone Variants

There are several ways in which chromatin structure may be altered [71]. The first is by histone PTMs by pathways that have been discussed in detail. The second is by chromatin remodeling complexes, which disrupt histone-DNA interactions, allowing for transcription factor binding to DNA and subsequent gene activation [72]. The third way in which chromatin structure may be altered is by the replacement of major histones with histone variants.

Incorporation of histone variants into the nucleosome has been implicated in a number of processes, including gene activation and silencing, chromosome segregation, X chromosome inactivation, and DNA repair [71]. I will focus on H2A.Z and H3.3, the incorporation of both of which is associated with specific transcriptional states. In *S. cerevisiae*, the H2A.Z homologue, Htz1, localizes to

regions of euchromatin that are adjacent to heterochromatin, suggesting that Htz1 may act as a boundary against the propagation of heterochromatin [73]. Accordingly, deletion of Htz1 resulted in the spread of Sir2 and Sir3 into euchromatic regions and reduced expression of associated genes. Htz1 is also enriched at inactive promoters that are poised for activation [74-76]. Upon activation, H2A.Z is evicted from the nucleosome, enabling active transcription complexes to be assembled at the promoter [76,77]. However, in mammalian cells, H2A.Z appears to be associated with heterochromatin [78]. In developing mouse embryos, H2A.Z localizes to pericentric heterochromatin and interacts with a pericentric heterochromatin binding protein. Also, H2A.Z may be required for binding of HP1 to chromatin, since depletion of H2A.Z disrupted HP1-chromatin interactions [79]. These studies suggest that histone variant H2A.Z is involved in the maintenance of heterochromatin and euchromatin, although its exact role is organism-specific.

While major histones are incorporated into a nucleosome during S phase, variant histones may be deposited during S phase or in a replication-independent manner. In *S. cerevisiae*, the protein Swr1, a Swi2/Snf2-related chromatin remodeling ATPase, was shown to catalyze the exchange of H2A/H2B dimers for Htz1/H2B dimers [80-82]. Swr1 and Htz1 co-immunoprecipitated with Bdf1, a protein containing two bromodomains, which may be involved in targeting the complex to euchromatin [81]. These studies identify a chromatin remodeling

mechanism by which histone variants replace canonical histones outside of S phase.

Like H2A.Z, deposition of histone variant H3.3 takes place outside of S phase [83]. In *Drosophila*, H3.3 differs from major H3 by only four amino acids; site-directed mutagenesis of these residues in major H3 allowed for some replication-independent deposition of major H3. The four distinctive H3.3 amino acids may mediate its interaction with the histone chaperone HIRA, which is required for replication-independent nucleosome deposition of H3.3 [84]. H3.3 deposition also depends on chromodomain protein Chd1, a chromatin remodeling enzyme that interacts with HIRA [85]. Like H2A.Z, replication-independent deposition of H3.3 employs chromatin-remodeling enzymes to replace major H3 with an H3 variant within the nucleosome.

The consequence of replacing major H3 with H3.3 is not completely understood, but appears to involve transcriptional activation. Chromatin immunoprecipitation showed that H3.3 is found at the promoters of transcriptionally active genes [86]. Correspondingly, H3.3 is enriched for PTMs associated with transcriptional activity [87]. Also, H3.3 may prevent the spread of heterochromatin since it is incorporated at a demonstrated boundary element [88]. Some propose that deposition of H3.3 alleviates repression of major H3 resulting from PTM-mediated silencing [83]. They suggest a mechanism by which transcriptional activation can be achieved through histone variant incorporation into the nucleosome.

Histone variants are present in a number of divergent organisms, including *Trypanosoma brucei* and other kinetoplastids. The role of histone variants in these organisms is currently being investigated.

Trypanosoma brucei and African Sleeping Sickness

Trypanosoma brucei is a flagellated protozoan parasite responsible for African Sleeping Sickness in humans and is one of the species responsible for Nagana in livestock. The prevalence of these diseases in sub-Saharan Africa is due to the geographic distribution of *Glossina* sspp. (tsetse), the insect vector responsible for the transmission of trypanosomes. The World Health Organization estimates that 70 million people live in areas where disease transmission could take place, and between 300,000 and 500,000 people are currently infected. Livestock infections have contributed to widespread economic deprivation and also perpetuate the human infection by acting as reservoirs for the parasite. African Sleeping Sickness presents with central nervous system symptoms, including confusion, poor motor and sensory system functioning, and sleep cycle disturbances; trypanosomiasis results in coma and death, if left untreated. Current drug treatments, which are not specific for trypanosomes, are toxic and have limited efficacy, so new drugs and vaccines must be developed to specifically target the parasite.

Variant Surface Glycoproteins

T. brucei successfully combats mammalian host defenses to establish and maintain an infection. Early observations showed that relapsing fever, which was characteristic of trypanosome infections, was concurrent with waves of parasitemia [89]. These authors and others speculated that, as an antibody response was mounted against each wave of parasites, some trypanosomes evade the host immune response by spontaneously switching their surface antigen, in a process known as antigenic variation. The surviving cells then initiate a new wave of parasitemia.

The dense coat comprised of surface antigen was first visualized by electron microscopy [90]. Purification and characterization of the surface coat antigen led to the identification of Variant Surface Glycoproteins (VSG), approximately 10 million identical copies of which are found on the plasma membrane [91]. Due to the dense packing of VSGs, only VSG epitopes are available to the host antibody, while invariant surface proteins are protected [92]. It was recognized, then, that each wave of parasitemia in an infected patient was due to a different VSG expressed as the surface coat [93]. Recent sequencing and bioinformatic analysis of the *T. brucei* genome shows that there are approximately 1600 *VSG* genes and pseudogenes [94,95]. Sequencing of additional strains shows that >50% of a megabase chromosome can be comprised of *VSG* [96].

The *T. brucei* genome consists of 11 diploid megabase chromosomes (1-5.2 Mb), an indeterminate number of intermediated chromosomes (200-700 kB), and approximately 100 minichromosomes (25-150 kB) [97,98]. Megabase chromosomes contain housekeeping genes and *VSG* ES, intermediate chromosomes also contain ES, and minichromosomes act as reservoirs for arrays of silent *VSG*. Although not all of the *VSG* are intact, the repertoire is sufficient to maintain an infection for the lifespan of the host.

Lifecycle Stages

Trypanosomes are transmitted to mammalian hosts by tsetse. Upon infection, the parasites differentiate into proliferating bloodstream forms (BF), which are characterized morphologically as long slender forms (Fig. 1.1, [99]) [100]. As parasitemia increases, trypanosomes differentiate to non-proliferating stumpy forms, which presumably prolongs the host's survival by limiting parasite numbers. Both long slender and short stumpy BF express VSG. During a bloodmeal, tsetse injests the parasites, which differentiate to the proliferating procyclic forms (PF) in the insect midgut. PF express a small family of invariant proteins called procyclins to form a protective surface coat [101]. As the trypanosomes pass to the salivary glands, they differentiate through a proliferating epimastigote form to a non-dividing terminally differentiated and infective metacyclic form. Metacyclic forms express a limited repertoire of VSGs

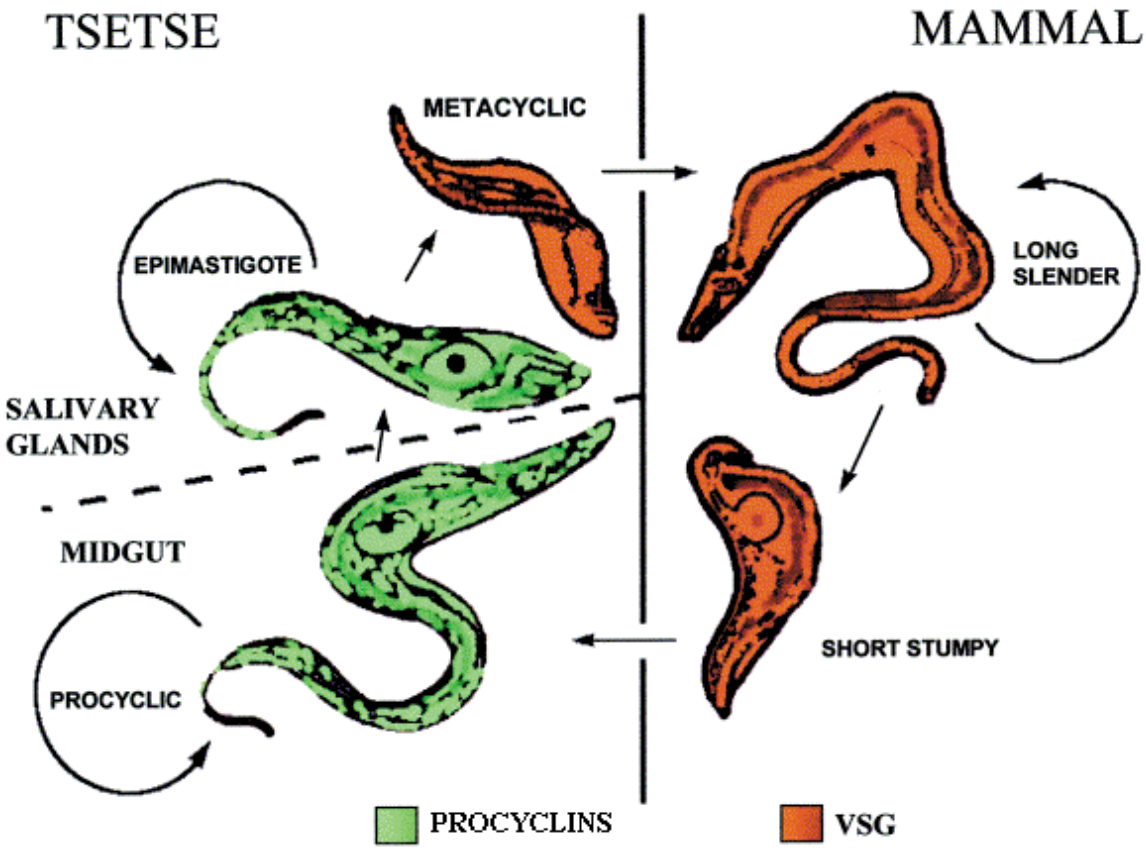


Figure 1.1. *Trypanosoma brucei* life cycle. During a bloodmeal, the tsetse injests bloodstream form trypanosomes. In the midgut, trypanosomes differentiate into procyclic forms, which express a small family of invariant surface proteins called procyclins. Trypanosomes migrate to the salivary glands of the tsetse and differentiate into metacyclic forms, which express a limited repertoire of VSGs in preparation for infection in a mammalian host. Following infection, parasites again differentiate into bloodstream forms. Trypanosomes expressing VSGs or procyclins on their surface coat are shown in red and green, respectively. Reproduced from [9].

in preparation for their adaptation to the mammalian host [102]. Once the parasites have infected humans, they differentiate again into proliferating bloodstream forms.

Differentiation from BF to PF is accompanied by a number of physiological changes that enable the trypanosome to adapt to its new environment [103]. Lifecycle differentiation from BF to PF can be achieved in culture, so that these processes may be studied [104]. Upon differentiation to PF, trypanosomes no longer require a defense against a host antibody response, so the VSG coat is shed and all ES become transcriptionally inactive [101,105]. Cells that were in a non-proliferative state in the mammalian host, re-enter the cell cycle and begin to undergo DNA synthesis [106,107]. Metabolic changes are observed upon differentiation to PF [108], since BF survive primarily on its host's glucose stores. Changes in the cellular architecture and morphology are also observed, including repositioning of the kinetoplast [109]. These adaptations require coordinated changes in gene expression in response to environmental stimuli.

Expression Sites

*VSG*s are expressed in BF from polycistronic transcription units known as Expression Sites (ES) that, importantly, are transcribed by RNA polymerase I (RNA pol I) [110]. There are approximately 20 ES in the genome of the *T. brucei* Lister 427 strain [111,112], each of which contain a single *VSG*. The remaining

hundreds of silent *VSGs* are found at non-transcribed telomeric loci and in transcriptionally silent subtelomeric arrays [94-96].

ES are located at the telomeres of some of the 11 diploid megabase chromosomes and variable numbers of smaller 'intermediate' chromosomes (Fig. 1.2, [113]). Transcription of the ES is initiated at the ES promoter [114]. ES are comprised of a ~50 kb array of genes, termed Expression Site Associated Genes (*ESAG*s), terminating at the subtelomeric *VSG* [115-117]. Eleven families of *ESAG*s have been identified, some of which have been characterized in detail [94]. *ESAG6* and *ESAG7*, for example, encode the two subunits of a heterodimeric transferrin receptor that localizes to the flagellar pocket [118,119]; sequence variation of *ESAG6* and *ESAG7* between ES led to the idea that ES switching optimizes transferrin uptake from the host serum [120], although this remains a controversial hypothesis [121]. However, despite their close physical association with the *VSG*, none of the *ESAG*s has yet been shown to be involved in regulating *VSG* expression. In addition to the *ESAG*s, a number of repetitive DNA sequences are associated with the ES; these include 50 bp repeats upstream of the ES promoter, 70 bp repeats upstream of the *VSG*, and the telomeric repeats.

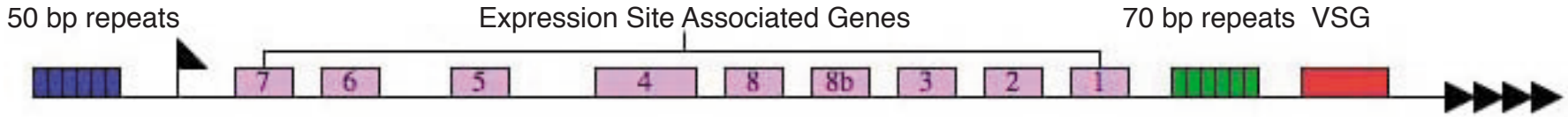


Figure 1.2. Architecture of the *VSG* Expression Site (ES). ES are 50 kb polycistronic transcription units that terminate at the subtelomeric *VSG*. Up to eleven families of Expression Site-Associated Genes (ESAG 1-11) are located in the ES. The ES is flanked by 50 bp repeats and telomeric repeats, and there are 70 bp repeats upstream of the *VSG*. Reproduced from [23].

Mechanisms of VSG Switching

There are three major mechanisms by which *VSG* expression can switch [122]. The most common [123], gene conversion, occurs when the *VSG* in the active ES is replaced by another gene, and the original gene is not retained in the genome [124,125]. The new *VSG* can either be from a telomere (telomere conversion) or a chromosome-internal site (duplicative transposition). The second mechanism is known as reciprocal recombination, and occurs when the active ES *VSG* is exchanged with a new telomeric *VSG*, and no genetic material is lost [126-128]. *In situ* switching, the third mechanism, occurs when transcription at the active ES is switched off and concurrently switched on at another ES [128]. The switching frequency for VSGs varies from 10^{-3}–10^{-5} switches/cell/generation [129]. The mechanisms that regulate *VSG* switching are far from understood and are discussed in some detail below.

Chromatin Structure in Trypanosoma brucei

Trypanosomes are highly divergent from other model organisms (Fig. 1.3, [130]), and chromatin structure has been shown to differ. Unlike other organisms, condensed chromosomes are not visible during nuclear division in trypanosomes. Chromatin from BF, PF, and rat liver were examined by electron microscopy at increasing ionic strength [131]. At high salt concentrations,

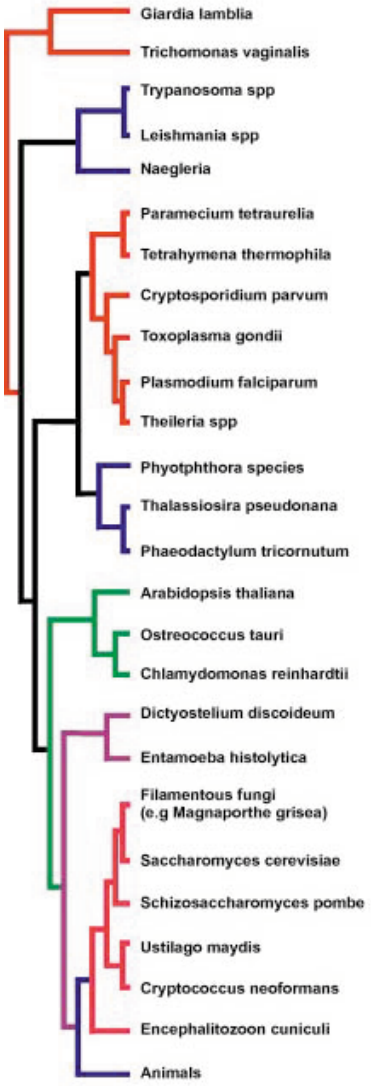


Figure 1.3. Evolutionary tree. An evolutionary tree shows the divergence of trypanosomes from other model organisms. It also shows the divergence of other parasitic organisms, including *Giardia lamblia*, *Toxoplasma gondii*, and *Plasmodium falciparum*. Reproduced from [130].

condensation was achieved in BF and, to a lesser extent, in PF, but trypanosome chromatin was not compacted to the same degree as rat liver chromatin. Comparison of PF and BF histones on a TAU gel demonstrated different band patterns. Also, incubation of BF and PF chromatin with micrococcal nuclease showed different digestion patterns; PF chromatin was digested more readily than BF, suggesting that the DNA-histone interactions in PF are weaker than in BF. Both BF and PF chromatin are more readily digested than rat liver chromatin. The authors of this study suggest that there are structural differences between BF and PF chromatin, both of which have weaker DNA-histone interactions than other organisms.

An additional study demonstrated that trypanosome chromatin is more accessible to T7 RNA polymerase than chromatin from other organisms [132]. The T7 polymerase is a single DNA-binding protein that can access the T7 promoter without the help of any other protein factors [133,134]. When the T7 system was introduced into the mouse germ line, only short promoter-proximal transcripts could be detected, indicating that the chromatin structure of higher eukaryotes provided minimal accessibility to T7 polymerase [133]. In contrast, T7 transcription in trypanosomes is very processive, producing transcripts ~3 kb from the transcription start site [132]. These studies confirm that trypanosome chromatin is in a more accessible confirmation than chromatin from other organisms.

Divergent Histone Sequences in Trypanosomes and Other Organisms

The above studies demonstrate that histones and chromatin structure are different in *T. brucei* compared to other model organisms, and histone sequences differ considerably (Fig. 3.28). Comparison of the yeast N-terminal histone sequence to *T. brucei* and other divergent organisms, including *Plasmodium falciparum* and *Giardia lamblia*, demonstrate the sequence divergence (Fig. 1.4). Interestingly, *G. lamblia* and *Entamoeba histolytica* have N-terminal extensions of H3 and H4, respectively. Some well-studied lysines are conserved among the divergent organisms, such as H3 K4 and H4 K12, K16, and K20. The conservation of these lysines through millions of years of evolution suggests that they play important roles. *Leishmania major*, *Trypanosoma cruzi*, and *T. brucei* histone sequences are well-conserved. Therefore, studying PTMs in any one of these organisms will most likely reveal common principles in all of them.

Sequence alignments of the histone N-termini of these organisms demonstrate a few important points. First, although histone sequences are highly conserved among most model organisms, they do differ among evolutionarily divergent organisms. Therefore, PTMs cannot be divined simply by looking at sequence alignments, so it is necessary to perform an unbiased survey of histone PTMs from each of these organisms. Also, commercially available antibodies are not specific for PTMs from divergent organisms, so, in order to study them, specific antibodies must be raised. Second, conservation of amino

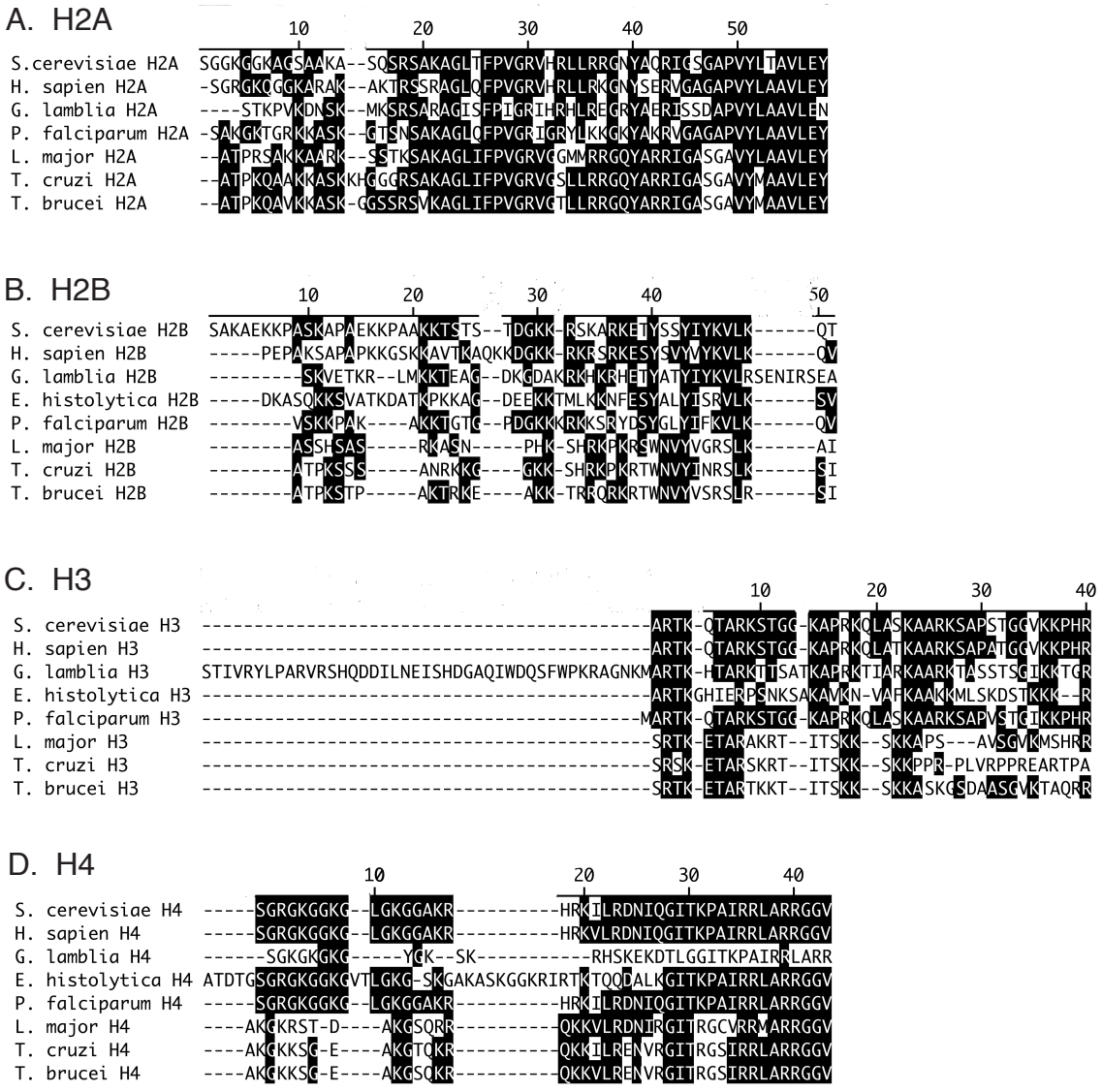


Figure 1.4. Sequence alignments of the N-termini of each core histone from evolutionarily divergent organisms. Identical residues are shaded. The sequence ruler is numbered according to the *S. cerevisiae* sequence. (A) H2A. (B) H2B. (C) H3. (D) H4.

acids and PTMs may demonstrate important sequences that are essential to all organisms, thereby illustrating the fundamental properties of chromatin biology. Third, functional studies of individual PTMs may uncover new pathways in which they are involved and new relationships between PTMs, revealing the full scope of chromatin biology. Therefore, it is important to study histone PTMs in divergent organisms both for their own sake and to enrich our understanding of chromatin in more frequently studied model organisms.

Transcription in Trypanosoma brucei

Transcription in trypanosomes differs from most other organisms in a few important ways [135-137]. First, genes are arranged in polycistronic transcription units. Genes are grouped randomly, and not as functional units. Second, there is apparently unregulated RNA pol II transcription of protein-coding genes. No RNA pol II promoter has been identified upstream of protein-coding genes. It is hypothesized that RNA pol II "touches down" on a chromosome at sites that are relatively free of ORFs and initiates transcription bi-directionally; genes on either side of this site are arranged in the same orientation [138,139]. Since transcription is apparently unregulated, mRNA transcripts have sequences within the 3' UTR that direct its degradation so that levels of transcripts may be controlled [140].

The third way in which transcription is different is that trypanosome mRNAs undergo trans-splicing. RNA pol II transcribes spliced leader RNA (SLRNA), which is capped at the 5´ end [141-143]. SLRNA is trans-spliced to the 5´ end of each ORF. Trans-splicing at the 5´ end of the subsequent ORF releases the 3´ end of the transcript, which then undergoes polyadenylation [144,145]. Trans-splicing uncouples the processes of transcription and mRNA capping. In other organisms, mRNA is capped at the 5´ end co-transcriptionally by capping machinery that associates with the elongating RNA pol II [42], which necessitates that RNA pol II transcribes protein-coding genes.

Trans-splicing alleviates the need for RNA pol II to transcribe all mRNA, so, in *T. brucei*, RNA pol I transcribes VSGs and procyclin [110]. This may allow for independent regulation of these genes. In keeping with the uncoupling of transcription and mRNA capping, the serine-rich trypanosome RNA pol II CTD lacks heptapeptide repeats [146,147]; therefore, it is not clear whether the CTD will act as a binding platform for elongation and RNA processing factors as it does in other organisms [42].

The apparently unregulated transcription of most protein-coding genes is consistent with the observation that trypanosome chromatin is in an open conformation. For this reason, it seems likely that most trypanosome chromatin is euchromatin, with the possible exception of subtelomeric regions where *VSG* are located.

Regulation of Monoallelic VSG Expression

The mechanism regulating whether an ES is transcriptionally silent or active has not been elucidated. Some theories are discussed below.

Transcription Initiation. Initiation of transcription does not appear to be a regulated process in trypanosomes. The absence of identifiable homologues of common transcription factors and the organization of genes into polycistronic units suggest that transcription itself may not be highly controlled [148]; trypanosomes appear to preferentially utilize other means of regulating mRNA abundance, especially by RNA processing and degradation [135-137].

It is not surprising, then, that ES promoters are nearly identical in sequence, and the deletion of sequences surrounding them does not affect the transcription levels of ES-associated genes [149]. Also, replacement of the ES promoter with an rRNA promoter allows for effective silencing and activation in inactive and active ES, respectively, but not after differentiation to PF [150]. These studies demonstrate that the ES promoter and surrounding regions are not critical for ES activation. Finally, transcription is initiated at many, if not all, ES, all of which argues against a strong role for regulation at the level of transcription initiation [111].

Transcription Elongation and RNA Processing. Although transcription is initiated at a number of 'silent' ES, transcription is rapidly attenuated such that only a single *VSG* is transcribed, suggesting that ES are regulated at the level of

transcription elongation [111]. Using RT-PCR, transcripts were analyzed using four sets of primers that annealed at different sites throughout the ES. PCR products were then sequenced to determine the number of ES represented. Immediately downstream of the promoter (1-200 bp), transcripts from nearly all ES are present; 1000-1200 bp from the promoter, heterogeneity of transcripts was reduced, but multiple ES were still represented; at the *VSG*, only transcripts from the active ES were identified. Transcripts were then fractionated by oligo(dT) chromatography into polyA+ and polyA- groups, which were analyzed by the same methods. The authors showed that only transcripts from active ES have a polyA tail, demonstrating that preferential RNA processing contributes to the dominance of active ES transcripts.

Expression Site Body. Monoallelic expression of *VSGs* may also be mediated by a specialized structure, the expression site body (ESB), which is not present in the PF nucleus [151]. An antibody was developed against a subunit of RNA pol I, and immunofluorescence was used to show a transcriptionally active RNA pol I-containing structure in the nucleus distinct from the nucleolus. The active expression site was tagged and shown to co-localize with the RNA-pol I containing body, but inactive ES do not co-localize. Also, the ESB is still present following DNA digestion by DNAse I, demonstrating that the ESB is a physical structure and not simply an accumulation of RNA pol I on an active ES.

Chromatin Structure. A number of studies suggest that chromatin structure at the ES may influence its transcriptional activity. The active ES

displays greater sensitivity to digestion with DNase I and single-strand-specific endonucleases than inactive ESs, which suggested a more open chromatin conformation at the active ES [152,153]. Micrococcal nuclease digestion at the promoter region of an active and inactive ES gave similar digestion patterns, suggesting that there are no differences in the chromatin conformation at the ES promoter [149]. Therefore, any differences in chromatin structure between active and inactive ES would be found closer to the 3´ end of the ES, where the *VSG* is located.

Chromatin structure was further implicated in transcriptional silencing when it was shown that transcription was repressed from both ES and rRNA promoters following their insertion into an inactive ES [154]. Upon differentiation to PF, the rRNA promoter was de-repressed, while the ES promoter remained repressed, suggesting that there are two separate mechanisms controlling ES silencing in BF and PF. The authors concluded that silencing in BF was due to a position effect caused by restrictive chromatin structure, while silencing in PF may be, at least partially, due to lack of BF ES promoter-specific factors. Other studies demonstrate that transcription levels decreased as the promoters were placed closer to a telomere [155]. These studies suggest that ES silencing in BF is due to a repressive chromatin conformation, but that ES silencing requires mechanisms in addition to telomere-mediated silencing (see next section).

A similar study was performed using a T7 polymerase and promoter, which allowed for the accessibility of chromatin at ES to be examined directly, in

the absence of protein factor interactions that might confound results from the RNA pol I studies [156]. A T7 RNA polymerase-driven luciferase reporter cassette was integrated in an inactive ES as well as in non-transcribed, non-ES sites, including the rRNA spacer region and upstream of the chromosome-internal VSG118 gene. The T7 promoter was active in all locations in BF, indicating that these sites provided the minimum chromatin accessibility required for transcription by T7 polymerase. Therefore, repression of RNA pol I-mediated transcription might be due to the absence of protein factors required for transcription elongation.

These studies sparked an interest in the effects of chromatin remodeling and histone modifying enzymes on ES activity. In a screen for genes involved in ES silencing, a Swi2/Snf2-related chromatin remodeling enzyme, TbISWI, was identified in trypanosomes [157]. In both BF and PF, RNAi knockdown of TbISWI led to de-repression of a fluorescent marker placed downstream of an inactive ES promoter. However, no *VSG* transcripts from inactive ES were observed, indicating that TbISWI may be involved in silencing at the ES promoter, but not at the subtelomeric *VSG*. Deletion or knockdown of trypanosome histone modifying enzymes have similarly been tested for their role in ES silencing. The results of these studies are discussed later.

Telomeric Silencing. Insertion of two *VSG* genes into a single active ES resulted in the stable expression of two VSGs on the trypanosome surface coat [158]. Double expressors had similar growth rates and infectivity in rats as wild-

type cells. When compared to a study mentioned earlier in which double expressors were engineered by inducing rapid switching between ES [159], the results imply that regulation is based not on the expression of a single *VSG*, but on the expression of only one ES subtelomeric region.

To investigate telomeric silencing in *T. brucei*, one group seeded *de novo* telomeres with an expression cassette 2 kb and 5 kb downstream of an rRNA promoter [160]. Expression from the rRNA promoter was low when the telomere was 2 kb from the promoter and increased as the distance between the telomere and the promoter increased, demonstrating that RNA pol I transcription is repressed by telomeres. However, seeding of *de novo* telomeres 2 kb and 5 kb downstream of a silent *VSG* both resulted in strong repression, indicating that another mechanism is involved in silencing at BF ES. This was confirmed upon differentiation to PF, when *VSG* expression was de-repressed at 5 kb from the telomere, but not at 2 kb.

Additional evidence that telomeric silencing was not the sole mechanism governing *VSG* silencing was provided by the same group [161]. The construct used to seed *de novo* telomeres was introduced into telomerase deficient cells with a unique restriction site upstream of the telomeric repeats. Cleavage with a conditional endonuclease resulted in loss of the telomere, which, in turn, resulted in de-repression of reporter constructs at non-ES telomeres, but not at ES telomeres. The authors provide direct evidence for the existence of telomeric silencing at non-ES loci, but they show that telomeres are not an absolute

requirement for ES silencing. However, they cannot exclude the possibility that telomere binding proteins may recruit silencing proteins or induce DNA or chromatin modifications; telomere loss would result in loss of binding proteins, but would not reverse existing modifications. Therefore, restrictive chromatin structure formed as a result of these modifications may still be a factor in ES silencing. To conclude with more certainty that silencing is not nucleated at the telomere, the authors could cleave an active site telomere and induce an *in situ* switch. They could then determine if a formerly expressed *VSG* is effectively silenced in the absence of a telomere.

Although telomeres have a limited role in *VSG* silencing, there is some evidence to suggest that they influence VSG switching. Deletion of telomerase resulted in the progressive shortening of telomeres [162]. The problem of critically short telomeres at the active ES was solved by recombination with a telomere from a silent ES, resulting in the replacement of the active site *VSG* with the *VSG* associated with the donor telomere (telomere conversion). In wild-type cells, the active ES is subject to breakage at the telomere, possibly due to destabilization of telomere structure that allows RNA pol I to transcribe to the 3´ end of the ES [163,164]. The authors propose that telomere breakage at the active site is the impetus for VSG switching, which occurs at random at a frequency of 10^{-3}–10^{-5} switches/cell/generation [129].

Conclusion. There is some evidence to suggest that chromatin structure plays a role in the regulation of *VSG* expression. Although the chromatin

structure in trypanosomes has been shown to be in a relatively open conformation, perhaps heterochromatin forms at the 3' end of ES, thus restricting transcription of all but one *VSG*. Whether heterochromatin forms at the ES telomere, as it does in other organisms, is debatable. Once thought to mediate ES silencing, telomeres were shown to not be absolute requirements for silencing, but may play a role in VSG switching. Characterization of chromatin structure and the factors that influence it, the most important of which are histone modifications, is necessary for understanding the mechanisms regulating *VSG* expression.

Trypanosome Chromatin Biology

***T. brucei* genome.** Sequencing of the *T. brucei* genome identified several putative histone-modifying enzymes: 5 HAT, 7 HDAC (including 3 SIR2-related proteins), 2 DOT1-like HMT, 3 SET-domain-containing HMT, and 4 arginine HMT [148]. The genome sequencing also revealed that genes encoding for major histones are present in tandem arrays, which presumably allows for increased expression of these genes. There are 13 copies of *H2A* on chromosome 7, 14 copies of *H2B* on chromosome 10, 7 copies of *H3* on chromosome 1, and 10 copies of *H4* on chromosome 5. Trypanosome histone sequences, which are very divergent compared to other organisms, are

discussed in Chapter 3. Finally, one copy each of genes encoding variants of the four major histones were identified.

PTMs on trypanosome histones. Two studies have been published identifying PTMs on *T. brucei* PF histones and *T. cruzi* H4 [165,166]. The findings are discussed in Chapter 3.

Histone modifying enzymes. A number of studies have been published recently identifying histone modifying enzymes in trypanosomes, including DOT1-like HMTs, HATs, and HDACs. Two DOT1-like HMTs, DOT1A and DOT1B, are responsible for the dimethylation and trimethylation, respectively, of *T. brucei* H3 K76, the *S. cerevisiae* H3 K79 homologue [166]. Dimethylation of H3 K76 was shown to be cell cycle-regulated; K76 dimethylation was only detectable during mitosis, while trimethylation was present throughout the cell cycle. Disruption of DOT1A (lethal) and DOT1B (non-essential) impaired life cycle differentiation and progression through the cell cycle. Specifically, DOT1A knockdown cells were able to undergo mitosis before completing DNA replication, resulting in the presence of apparently haploid cells. The authors propose that DOT1A and DOT1B play a regulatory role at the mitotic checkpoint through methylation of H3 K76.

Deletion of yeast DOT1 disrupts telomeric silencing [167]. Similarly, recent experiments show that deletion of DOT1B results in de-repression at inactive ES (L. Figueiredo, unpublished data). Specific antibodies were used to show that two VSGs are expressed simultaneously in mutant cells in the absence

of drug selection. Therefore, deletion of DOT1B appears to disrupt telomeric silencing, a role that is homologous to its yeast counterpart, and may be a regulator of ES silencing.

Of the 5 putative HATs in *T. brucei*, HAT3, which is non-essential, has been shown to function *in vivo* as an H4 K4 HAT [168]. An HDAC was not identified for acetylated H4 K4. Exposure of cells to cycloheximide, a protein synthesis inhibitor, led to an immediate loss of unmodified H4 K4. The authors suggest that newly synthesized histones are unmodified at H4 K4 in the cytosol and are acetylated inside the nucleus, possibly irreversibly. The biological outcome of H4 K4 acetylation is currently unclear, especially since it seems not to be essential.

Two studies have provided an overview of the 7 putative HDACs in *T. brucei* [169,170]. Fluorescence microscopy studies of 3 *T. brucei* SIR2-related proteins, TbSIR2RP-1-3, show that TbSIR2RP-1 localizes to the nucleus, while TbSIR2RP-2 and TbSIR2RP-3 are found in the mitochondria, suggesting that perhaps only TbSIR2RP-1 will have a role in transcriptional silencing [170]. Two studies showed that TbSIR2RP-1 does not affect ES silencing [170,171]. However, deletion of *TbSIR2RP-1* results in modest de-repression of reporter expression at non-ES subtelomeric loci, suggesting that TbSIR2RP-1 plays a role in establishing telomeric silencing [170]. Finally, TbSIR2RP-1 has *in vitro* deacetylase and ADP-ribosyltransferase activity against histones, but its *in vivo*

target has yet to be identified [172]. These studies are discussed in more detail in Chapter 4.

Of the remaining 4 HDACs, DAC1-4, two were encoded by essential genes (*DAC1* and *DAC3*) and two were not (*DAC2* and *DAC4*) [169]. Interestingly, *Δdac4* cells demonstrate cell cycle defects, which were illustrated by growth delays and an accumulation of cells in late G2 phase. However, neither the *Δdac2* nor *Δdac4* cells demonstrated changes in VSG expression or an inability to undergo lifecycle differentiation.

An overview of the histone modifying enzymes and corresponding PTMs shows that they affect several biological processes in trypanosomes, including progression through the cell cycle, life cycle differentiation, and gene expression. De-repression of inactive ES following disruption of DOT1B provides preliminary evidence for a role for histone PTMs in the maintenance of ES silencing.

Histone variants. Two studies have been published characterizing histone variants in *T. brucei* [173,174]. One study identified H2A and H2B variants [174], H2AZ and H2BV, which are discussed in more detail in Chapter 5. Briefly, both variants are encoded by essential genes. The authors show, by co-immunoprecipitation, that H2AZ and H2BV heterodimerize in the same nucleosome. By chromatin immunoprecipitation, it was shown that H2AZ and H2BV were enriched at repetitive DNA. H2AZ and H2BV were not detected at inactive ES, although their presence at active ES could not be determined. Finally, it was shown that H2AZ does not co-localize with sites of BrUTP

incorporation, indicating that the histone variants are not found at sites of active transcription.

H3 variant H3V is encoded by a non-essential gene [173]. Immunofluorescence shows that the localization of H3V shifts throughout the cell cycle and co-localizes with the mitotic spindle and telomeres. Chromatin immunoprecipitation confirmed that H3V was enriched at telomeric DNA, although subtelomeric VSG transcription was unaffected by the loss of H3V. The authors conclude that H3V may have a telomere-specific function and suggest a role for the variant in ES silencing, which may have been masked if another histone played a functionally redundant role to H3V.

Thesis Research

Packaging of DNA into nucleosomes acts to repress transcription by RNA pol II. In yeast and other organisms, elongating RNA pol II recruits chromatin remodeling and histone modifying enzymes that presumably enable it to overcome this transcriptional repression. In trypanosomes, although DNA-histone interactions are weaker than in higher eukaryotes, it seems likely that heterochomatin assembly at the 3´ end of inactive ES inhibits RNA pol I processivity. It is hypothesized that the ESB possesses elongation factors that enable RNA pol I to overcome nucleosomal repression, and it is likely that some of these factors will be chromatin remodeling and histone modifying enzymes.

In trypanosomes, we are far from understanding the processes of heterochromatin assembly and RNA elongation, both of which are informed by histone PTMs in other organisms. As a first step toward studying chromatin structure in trypanosomes, I performed a detailed survey, using Edman degradation and mass spectrometry (MS), to identify as many as possible of the PTMs present on histones H2A, H2B, H3, and H4, which I purified from the BF of *T. brucei* strain Lister 427. Although histones are highly conserved in most organisms, the amino acid sequences of *T. brucei* histones are very different, and we cannot infer what modifications will be present, based on sequence alignments (Fig. 3.28). Similar studies have been performed in *T. cruzi* and *T. brucei* PF. Structural differences between BF and PF chromatin have already been identified [131]; I compared our findings from the two lifecycle stages and attempt to identify stage-specific histone PTMs. Also, I compared our findings to marks found in yeast and other organisms to make predictions about the role of trypanosome PTMs.

Next, I began to characterize a putative histone deacetylase, TbSIR2RP-1. TbSIR2RP-1 was chosen for the role that yeast SIR2 plays in heterochromatin formation. I investigated the target of TbSIR2RP-1 and find that, unlike its yeast counterpart, TbSIR2RP-1 is not an H4 deacetylase. I suggest that TbSIR2RP-1 may be an H2A deacetylase, but further work will be required to justify this claim.

Finally, I began to look at marks particularly associated with transcriptional elongation in other organisms. Based on our MS data and a sequence

comparison between yeast and trypanosome H3, I infer that H3 lysine 4 is trimethylated. Commercial antibodies that recognize modifications in other organisms cannot be expected to specifically cross-react with trypanosome PTMs, and the most likely ones do not (unpublished data). Therefore, a specific antibody was raised against trimethyl H3 K4 to begin characterizing this mark.

I also generated a hypothesis regarding the relationship between PTMs on different histones based on sequence comparisons. I found that histone variant H2Bv lysine 129 is homologous to *S. cerevisiae* H2B lysine 123. However, unlike yeast, I showed that trypansome H2Bv K129 ubiquitination is not required for downstream H3 K4 and K76 methylation. Interestingly, I found that H3 K4 and K76 methylation is enriched in nucleosomes containing H2Bv. I propose that H2Bv deposition in the nucleosome is sufficient to act as a trigger for downstream methylation of H3 K4 and K76.

Chapter 2. Materials and Methods

Amplification of BF 221wt Trypanosomes in Rats

Approximately $5x10^5$ BF 221wt were used to infect 3 CD-1 female swiss white mice. The infection took place for 3 days after which point the mice were assessed for parasitemia. Whole blood from the animals that maintained the infection was used to infect 3-10 Sprague Dawley male rats ($1\text{-}2x10^7$ tryps were needed to infect each rat). Following a 3 day infection period, rats were exsanguinated from the aorta, and the blood was centrifuged at 1,000g for 10 min at 0 ºC in a Sorvall RT6000B centrifuge. Three layers were observed following centrifugation, and the top layer (plasma) was discarded. Trypanosomes were resuspended in BBSG (0.050 M Bicine; 0.077 M glucose; 0.050 M NaCl; 0.005 M KCl; pH 8.0) and centrifuged at 1,000g for 10 min at 0 ºC in a Sorvall RT6000B centrifuge. The top layer was discarded, and the bottom layer was resuspended in BBSG.

Prior to this stage, DEAE-52 beads were poured to a final volume of 75 ml and washed with 150-200 ml BBSG. The column was washed with cold BBSG immediately before application of the cells. Trypanosomes were added to the column under reduced pressure until the column almost ran dry. One column volume (75 ml) of BBSG was added to the column, and the trypanosomes were collected. This protocol yields on the order of 10^{10}-10^{11} cells.

Purification of Histones from BF 221wt cells

For purification of histones, $1x10^{10}$ BF 221wt cells were centrifuged at 1,800g for 10 min at 4 ºC in Sorvall SS-34 rotor. Cell pellet was washed with 10 ml Trypanosome Dilution Buffer (0.005 M KCl, 0.080 M NaCl, 0.001 M $MgSO_4.7H_2O$, 0.020 M Na_2HPO_4, 0.002 M $NaH_2PO_4.2H_2O$, 0.020 M glucose). Cells were resuspended in freshly made 20 ml 8% Polyvinyl pyrrolidone-40 + 0.05% octyl glucoside and immediately processed with pre-cooled Polytron (PTA head 10, setting 6) with six 1 minute bursts at 4 ºC [175]. Samples were checked under microscope to ensure that cells were lysed before centrifuging at 16,000g for 5 min at 4 ºC in Sorvall SS-34 rotor. The pellet was washed in 10 ml Buffer A (10 mM Tris, pH 8.0; 75 mM NaCl; 0.05% octyl glucoside). The pellet was resuspended in 10 ml Buffer B (10 mM Tris, pH 8.0; 2 M NaCl; 1% octyl glucoside). Octyl glucoside can be substituted with NP-40 for a better yield of histones, but NP-40 detergent contamination interfered with MS analysis in the final sample. Nuclei were lysed with a dounce homogenizer (50 strokes, which was probably excessive) and centrifuged at 16,000g for 20 min at 4 ºC in Sorvall SS-34 rotor. The supernatant was collected and divided into 0.9 ml aliquots. For acid extraction, 0.1 ml 4 M H_2SO_4 was added to each aliquot and rotated overnight at 4 ºC. Samples were centrifuged at 16,000g for 10 min at 4 ºC in a microcentrifuge. The supernatant was collected and neutralized by adding 1 M

Tris, pH 9.0. The sample was dialyzed overnight against Concanavalin A Equilibration Buffer (20 mM Tris, pH 7.4; 0.5 M NaCl) at 4 ºC. All of the solutions used up to this point were treated with 1 mM DTT, 50 mM sodium butyrate, 0.5 mg/ml TLCK, 0.174 mg/ml PMSF, and a protease inhibitor cocktail (Sigma).

The sample was concentrated in a 15 ml Vivaspin concentrator by centrifugation at 2,000g for 3 hours at 4 ºC in a Sorvall RT6000B centrifuge. To remove remaining VSG, which was otherwise the main contaminant in the purified histone preparation, the sample was applied to a 1 ml Concanavalin A column (Sigma) at 4 ºC. The column was washed with 20 ml Wash Buffer (1 M NaCl; 5 mM $MgCl_2$; 5 mM $MnCl_2$; 5 mM $CaCl_2$) and 2 ml Equilibration Buffer. The sample was applied to column, followed by 5 ml equilibration buffer. Flow through was collected immediately after sample was added to column. Flow through was frozen in liquid Nitrogen and stored overnight at –70 ºC prior to lyophilization.

The sample was fractionated by RP-HPLC on a C8 column using an acetonitrile gradient (0 to 60%) in 0.1% trifluoroacetic acid. Fractions were collected every minute for approximately 70 min, although later it was shown that histones eluted between 50 and 65 min. Protein content was monitored by absorbance at 214 nm. Fractions were lyophilized and resuspended in 50 μl ddH_2O. The composition of each fraction was determined by SDS-PAGE and MALDI of the intact proteins.

Analysis of Histones by Edman Degradation and Mass Spectrometry

Edman degradation and mass spectrometry experiments were conducted by Joseph Fernandez and Haiteng Deng in the Proteomics Resource Center at The Rockefeller University. Edman was performed using a Procise 494-HT with a PTH C18 HPLC column (Applied Biosystems). Modified PTH standards (acetyl or methyl lysines) were made by reacting free amino acids (Sigma) with PITC to create PTC-amino acids. These were then converted to PTH-amino acids by incubation with 25% trifluoroacetic acid.

Histones were digested with endoproteinases Glu-C, trypsin, or Asp-N, according to the manufacturer's instructions (Roche). Some histone preparations were also modified with propionic anhydride prior to trypsin digestion, to block lysines from cleavage. Histones were dissolved in 15-30 μl 1.0 M Tris, pH 8.0, and an equal volume of 1% propionic anhydride in acetonitrile (Aldrich)[176]. The reaction was incubated 30-120 min (longer incubation times resulted in non-specific modification at serines or threonines), dried, and digested with trypsin. Some digests were fractionated by HPLC in order to obtain information about minor species. Peptides were isolated on a Vydac C18 column (1.0x150 mm) using a Helwett-Packard/Agilent 1100 HPLC system.

Digests were first analyzed by MALDI-TOF, which was conducted using a DE-STR mass spectrometer operating in the delayed extraction and reflector modes (Applied Biosystems). Samples were then analyzed by ESI-LC-MS/MS

connected to a PE Sciex Q-STAR (Quadropole-Time of Flight mass analyzer). Nanoflow, an electrospray ionization source, was performed using a 75 mm ID C18 column.

Creating a ΔTbsir2rp-1 Cell Line

In a BF 221wt background, the *Tbsir2rp-1* single knockout was created by transfection of 10 μg pMH166 fragment (created by PvuII/HindIII digest) in $1.5x10^7$ cells and selection in HMI-9 media + 5.0 μg/ml hygromycin. The *Tbsir2rp-1* double knockout was created by transfection of 10 μg pMH167 fragment (created by AlwNI/ApaLI digest) in $3.0x10^7$ *Tbsir2rp-1* single knockout cells and selection in HMI-9 media + 1.5 μg/ml neomycin. pMH166 and pMH167 were provided by M. Hoek [171]. Cell lines were verified by Southern blot.

Triton-Acetic-Acid-Urea (TAU) Gel Electrophoresis

Purified histones from BF 221wt and *ΔTbsir2rp-1* cells were dried and resuspended in 5 μl freshly prepared loading buffer (0.36 g/ml Urea; 0.02% pyronin Y; 5% acetic acid; 25 mg/ml protamine sulfate). Samples were loaded on 15% or 20% polyacrylamide (60% acrylamide: 0.4% bisacrylamide), 6 M Urea, 0.74% Triton X-100, 5% acetic acid gel and electrophoresed at 300 V in 5% acetic acid. TAU gels allow for the separation of histones based on their size

and charge, which enables us to visualize different modified forms of these proteins [177]. Histones were visualized by silver staining.

Purification of Histone H3 Lysine 4 Trimethyl and Unmethylated Antibodies

Polyclonal antibodies were raised by immunizing rabbits (Sigma) with KLH-conjugated peptide RTKme0ETARTC (H3K4me0) or RTKme3ETARTC (H3K4me3). The anti-H3K4me3 antibody was purified by first depleting the antiserum from rabbit 14,995 bleeds #1-3 on Sulfolink coupling gel (Pierce) linked to peptide VSGAQKme3EGLRFC (H3K76me3). Each 1 ml of H3K4me3 antiserum was incubated on an H3K76me3-conjugated column for 30 min at 4 ^{o}C. Flowthrough and column washes were collected and concentrated in Vivaspin concentrators. Following elution of the bound material from the H3K76me3 column, the depleted antiserum was incubated on the column a second time. Flowthrough and column washes were collected. Twice depleted H3K4me3 antiserum was affinity purified on an H3K4me3-conjugated column as described [178]. H3K4me0 antisera from rabbit 14,993 bleed #1 was affinity purified on an H3K4me0-conjugated column.

The following peptides (10 ng/ml) were used to test antibody specificity: RTKme0ETARTC (H3K4me0), RTKme3ETARTC (H3K4me3), VSGAQKme0EGLRFC (H3K76me0), and VSGAQKme3EGLRFC (H3K76me3). Antibodies were incubated with peptide for 1 hr at RT prior to incubation of the Western blot, which contained 5 lanes of cell lysate from $2.0x10^{6}$ PF 427wt cells.

Fluorescence Microscopy

Both BF 221wt and PF 427wt cells were stained with anti-H3K4me0 and H3K4me3 antibodies. Approximately $5.0x10^6$ cells were collected by centrifugation at 1,800g for 10 min. The cell pellet was resuspended in 1 ml medium and fixed in 2% formaldehyde for 5 min at RT. Cells were collected by centrifugation at 800g for 3 min and washed twice in PBS for 5 min with rotation. Cells were resuspended in 1 ml PBS, and 100 μl ($5.0x10^5$ cells) were dropped on silanized coverslips. Cells were allowed to adhere to the coverslips for 20 min. Excess liquid was aspirated, and the coverslips were washed with PBS. Cells were permeabilized in 0.1% NP-40 (0.2% for BF cells) for 5 min at RT and were then washed twice with PBS. Cells were blocked twice with PBG for 10 min at RT with shaking, then incubated with 250 μl anti-H3K4me0 (diluted 1:100 for PF and 1:50 for BF) or anti-H3K4me3 (dil 1:2000 for PF and 1:750 for BF) antibodies diluted in PBG for 1 hr at RT with shaking. Cells were washed 3x in PBS for 5 min with shaking and once in PBG for 5 min. Cells were incubated with 50 μl Alexa anti-rabbit secondary antibody (diluted 1:200) for 1 hr at RT with shaking. Coverslips were kept covered from this point forward to protect them from light. Cells were incubated in 0.5 mg/ml DAPI diluted in PBS for 10 min at RT with shaking, then washed 3x with PBS for 10 min with shaking. Coverslips were mounted on slides with embedding media (20 mg p-phenylene diamene dissolved in 20 ml PBS diluted in glycerol). For BF cells, trypanosome dilution

buffer was substituted for PBS. Cells were examined using DeltaVision deconvolution microscopy (Applied Precision).

H2Bv-FLAG cell lines

Cell lines were derived from the single marker (sm) cell line, which constitutively expresses T7 polymerase and the *tet* repressor, allowing for inducible expression of ectopic genes that are downstream of the T7 promoter [179]. Plasmid pJEL69, which is used to introduce an inducible, ectopic copy of H2Bv-FLAG, was made by PCR amplifying the *H2Bv* ORF to include the FLAG epitope (DYKDDDDK) at the N-terminus (J. Lowell, unpublished data). pJEL69 was linearized with NotI and transfected in the single marker cell line. Selection took place in HMI-9 with 2.5 μg/ml phleomycin. The endogenous copies of *H2Bv* were replaced with puromycin and hygromycin drug resistance markers using deletion plasmids pJEL74 and pJEL75, respectively, to create BFJEL18 cells [174]. pJEL plasmids were provided by J. Lowell.

Plasmids used to introduce mutations in H2Bv-FLAG were derived from pJEL69. Lysine 129 (AAA) was mutated to alanine (GCA) or arginine (AGA) using the QuikChange Site Directed Mutagenesis kit (Strategene). The *H2BvΔCTD* mutation was made by PCR amplifying *H2Bv* 1-127 with a premature stop codon from pJEL69 and using it to replace the wildtype copy of *H2Bv* in the

same plasmid. FLAG-tagged *H2Bv* mutants were used to replace endogenous *H2Bv* alleles in single marker cells as described above.

A cell line (BFJEL8) containing an inducible, ectopic copy of H2B-FLAG was also used [174].

Purification of H2Bv for MS and Western

Histones were purified from BFJEL18 as described above. In addition to protease inhibitors, solutions were treated with 10 mM iodoacetamide. Following acid extraction and dialysis in TBS, the crude histone preparation was incubated on anti-FLAG M2 affinity gel (Sigma) with rotation for 2 hr at 4 ºC. Bound material was washed three times with TBS and eluted by boiling the gel in 2x SDS loading buffer for 3 min. The eluate was analyzed by Coomassie stain. Eight gel slices were cut from the gel ranging from ~17 to 35 kDa and analyzed by tandem MS.

Eluate from $2.0x10^8$ cells was also analyzed by Western blot with an anti-ubiquitin antibody (P4D1, Santa Cruz Biotechnology). Immunoprecipitated H2Bv-FLAG was evaluated alongside 100 ng ubiquitin (Sigma) and cell lysate from $2.0x10^6$ BFJEL18 cells. The blot was incubated in anti-ubiquitin antibody (diluted 1:100) ON at 4 ºC.

Preparation of mononucleosomes and co-immunoprecipitation

Mononucleosomes were prepared from cell lines containing the following inducible fusion proteins: H2B-FLAG, H2Bv-FLAG, H2Bv K129A-FLAG, and H2Bv K129R-FLAG [174]. Approximately $1.5x10^8$ cells were collected by centrifugation at 1,800g for 10 min. Cells were then washed in trypanosome dilution buffer and resuspended in 1 ml permeabilization buffer (100 mM KCl; 10 mM Tris, pH 8.0; 25 mM EDTA). Digitonin was added to a final concentration of 40 μM and incubated at RT for 5 min. Permeabilized cells were collected by centrifugation at 800g for 3 min. To remove EDTA, cells were washed 3x with 1 ml cold isotonic buffer (100 mM KCl; 10 mM Tris, pH 8.0; 10 mM $CaCl_2$; 5% glycerol). The cell pellet was resuspended in 0.5 ml isotonic buffer and incubated with 2 U micrococcal nuclease for 5 min at 25 ºC in a water bath. The reaction was stopped with EDTA (10 mM final concentration). To solubilize chromatin, NaCl (200 mM final concentration) and NP-40 (0.05% final concentration) were added, and the reaction was incubated on ice for 5 min. Soluble mononucleosomes were collected in the supernatant following centrifugation at 16,000g for 10 min. 1.5% mononucleosomes were set aside for Western blot analysis. Isotonic buffer was treated with 0.5 mg/ml TLCK, 0.174 mg/ml PMSF, and a protease inhibitor cocktail (Sigma).

For the co-immunoprecipitation, mononucleosomes were incubated on anti-FLAG M2 affinity gel (Sigma) with rotation for 2 hours at 4 ºC. Bound

material was washed once with isotonic buffer with 200 mM NaCl for 10 min with rotation. Bound material was then washed twice with isotonic buffer for 10 min with rotation. Immunoprecipitate was eluted by boiling the gel in 2x SDS loading buffer for 3 min. Approximately 33% of the H2Bv immunoprecipitate and 11% of the H2B immunoprecipitate were analyzed by Western blot. The 3:1 ratio of H2Bv to H2B IP resulted in approximately equal loading of total H3, which was demonstrated with a general anti-H3 antibody (Abcam 1791).

Chapter 3. Purification of Histones from *T. brucei* Bloodstream Forms and Identification of Posttranslational Modifications

Introduction

As described in chapter 1, previous studies of *T. brucei* ES have suggested a role for chromatin structure in the regulation of *VSG* expression. Whether chromatin is in a permissive or restrictive state for access by transcription machinery is influenced in part by which histone modifications are present. To begin to understand the relationship between chromatin structure and gene expression in trypanosomes, I proposed to first identify the histone modifications present in BF cells. Typically, histones are highly conserved proteins, so modifications can be identified by homology using well-studied model systems such as yeast as a guide. However, trypanosome histones diverge significantly from this consensus, so I expected to uncover unique modifications that have not been previously described, although some of the marks that are identified will be conserved in other organisms as well. I did not expect that trypanosome histones would contain many modifications because sequencing of the *T. brucei* genome showed that there are few putative histone modifying enzymes [148], and only a few of these have been shown to function in their predicted roles [169,172].

Some histone modifications in PF trypanosomes have been identified previously by our lab [165]. Here, a more comprehensive survey of histone PTMs is performed in BF trypanosomes. Since VSG expression is limited to bloodstream and metacyclic forms, our preliminary survey may uncover lifecycle-specific candidate modifications.

To begin this analysis, I developed a purification strategy for histones from BF trypanosomes. The strategy used to purify PF histones was not successful in BF because it required the purification of nuclei. BF nuclei are more fragile than in PF and lysed during purification [175]. Therefore, I avoided the purification of nuclei, but maintained the same basic outline as the PF protocol [165]. The protocol includes the following steps: cell lysis and collection of the insoluble material, including the nuclei and plasma membrane, lysis of the nuclei, collection of the soluble nuclear contents, extraction of basic proteins with acid, removal of contaminating VSG using a Concanavalin A column, and separation of histones by HPLC.

Analysis of purified histones was performed using Edman degradation and mass spectrometry. Edman degradation is a quantitative method that can be used to uncover modifications in the N-terminus by comparison to amino acid standards with known elution patterns. Quantitation is possible by calculating the area under the curve representing the modified amino acid as a fraction of the total amino acid. Edman degradation is limited by the inability to sequence proteins that are blocked at the N-terminus by modification of the primary amine,

the requirement for a significant amount of starting material, the ability to provide accurate sequence information for only the first ~20 amino acids, and the lack of sensitivity for the identification of minor modifications. Mass spectrometric (MS) methods have been developed more recently to identify histone modifications and complement the information provided by Edman degradation. To be analyzed by MS, histones must first be digested with an endopeptidase, and resulting peptides are separated according to their mass by MALDI-TOF. Sequence information for individual peptides is provided by MS/MS, which fragments amide bonds to produce a series of peaks with differences in mass that correspond to the mass of each amino acid and its potential modification. The drawbacks of MS are that the protein of interest must be amenable to digestion by the available endoproteases and the peptides produced may be too large or small or may have other characteristics that prevent them from registering on a MALDI-TOF spectrum. Also, tandem MS (MS/MS) data may be incomplete, which complicates the assignment of modifications to a specific lysine, especially when multiple lysines are present on a single peptide. Finally, the presence of multiple species of a modified peptide on a single spectrum further complicates the assignment of peaks to an individual species. With these drawbacks in mind, I proceeded with a global survey of BF histones.

Results

Purification of Histones from Bloodstream-Form Trypanosomes

To develop a protocol to purify histones from BF trypanosomes, I took advantage of their two essential properties: they are found in the nucleus and are basic proteins. Histones were previously purified from PF by isolating nuclei, then using acid extraction and RP-HPLC purification [165,175]. This method was attempted in BF trypanosomes, but the yield was extremely low (<1%) (Fig. 3.1). The greatest loss of material occurred during nuclei isolation, which I attributed to the increased fragility of the BF nucleus. Consequently, the protocol I developed did not involve the purification of nuclei as an intermediate step. The cells were lysed with a Polytron and the nuclei and plasma membrane were separated from the cytosolic fraction by centrifugation, resulting in a crude nuclear preparation. Several methods were employed for disruption of the plasma membrane, including osmotic lysis and lysis with RIPA buffer (137 mM NaCl; 20 mM Tris-Cl, pH 7.4; 10% (v/v) glycerol; 1% (v/v) Triton X-100; 0.5% (w/v) sodium deoxycholate; 0.1% (w/v) SDS; 2.0 mM EDTA; 1.0 mM PMSF; 20 mM leupeptin) containing 1% NP-40, but mechanical disruption using a Polytron appeared to be the most effective.

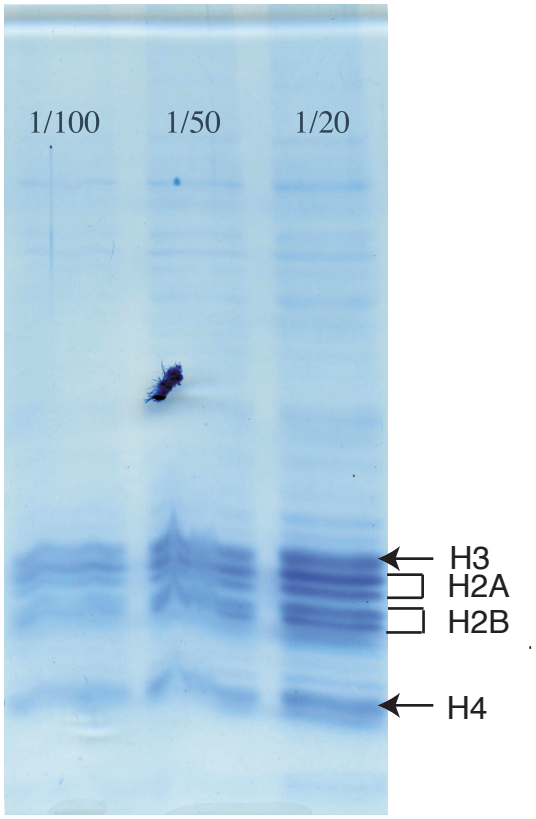


Figure 3.1. Histones from purified nuclei. SDS-PAGE of acid-extracted histones following purification of nuclei by Rout and Field protocol (178). Lanes 1-3 contain 1/100, 1/50, and 1/20 of total sample. Histones were identified by MALDI-TOF MS.

Following cell lysis and isolation of the crude nuclear fraction by centrifugation, the nuclei were washed in a low-salt buffer (10 mM Tris, pH 8.0; 75 mM NaCl; 0.05% NP-40) and resuspended in a high-salt buffer (10 mM Tris, pH 8.0; 2 M NaCl; 1% NP-40). The nuclei were mechanically disrupted using a dounce homogenizer, and the soluble fraction containing histones was separated from the insoluble material, including the nuclear and plasma membranes by centrifugation. Initially, histones were retained in the insoluble fraction, until the salt concentration in the high-salt buffer was increased from 0.4 M to 2.0 M NaCl (Fig. 3.2). Another method of liberating histones from the insoluble fraction involved permeabilizing the nuclei with digitonin and treating the chromatin with micrococcal nuclease. However, because increasing the salt concentration effectively released most of the histones into the supernatant, it was unnecessary to treat the nuclei with micrococcal nuclease.

The supernatant was then treated with 0.4 M H_2SO_4, resulting in the solubilization of basic proteins, which were separated from insoluble proteins by centrifugation. At this stage, histones were heavily contaminated with VSG, which was not efficiently removed in later stages of our purification strategy, so VSG was depleted by passing the sample over a ConA column. To remove detergent and acid from the basic proteins before introducing the sample onto the column, the acid was neutralized with 1 M Tris, pH 9.0, and the sample was dialysed against ConA equilibration buffer (20 mM Tris, pH 7.4; 0.5 M NaCl).

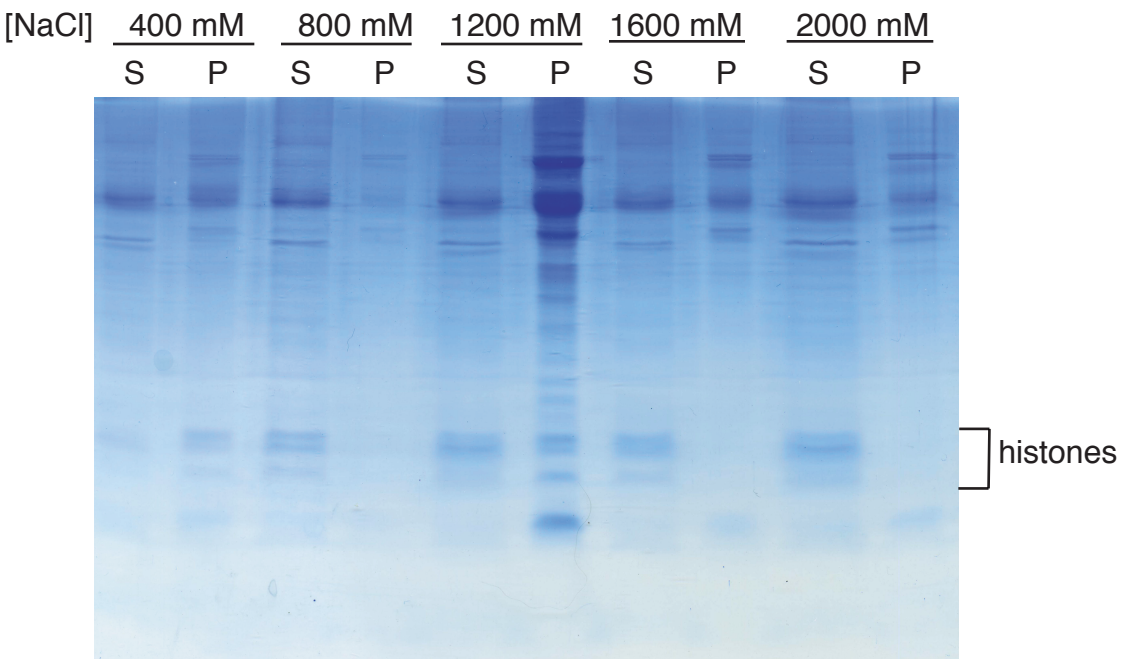


Figure 3.2. Optimization of salt concentration. SDS PAGE following histone isolation at salt concentrations ranging from 0.4-2.0 M NaCl. Histones are maximally released from the nuclear pellet into the supernatant at 2.0 M NaCl. S - supernatant, P - pellet

The yield at this stage was approximately 30% of our starting material in the crude nuclear pellet (Fig. 3.3).

In order to separate histones among themselves and from other remaining basic proteins, the sample was fractionated by RP-HPLC in an acetonitrile gradient (performed in collaboration with S. Hake, Fig. 3.4), lyophilized, and redissolved in H_2O. Analysis of HPLC fractions by Coomassie staining of polyacrylamide gels and by MALDI-TOF MS demonstrated the high purity of the final histone preparations (Fig. 3.5). Preliminary experimental analysis of all HPLC fractions showed that histones were only present in fractions 45 through 62. Histones H2A and H3 co-eluted on the HPLC (MW 14.2 and 14.7 kDa, respectively), which I unsuccessfully attempted to resolve in several ways. First, HPLC fractions were collected every half minute instead of every minute, but H2A and H3 still eluted in the same fractions. Secondly, H2A- and H3-containing fractions were separated on an AU gel, individual bands were excised, and histones were eluted from the gel in 5% acetic acid. Disappointingly, this produced a yield of less than 10%. Finally, I considered running these fractions on a second HPLC column, but had insufficient material to do so. Although it would be convenient to have H2A and H3 in separate fractions, I did not attempt to separate them further because the presence of both histones in a single sample did not interfere with MS or Edman analysis. Only the H2A N-terminus was susceptible to Edman sequencing, because the H3 N-terminal Serine (S1) was blocked, as previously shown in *T. brucei* PF [165].

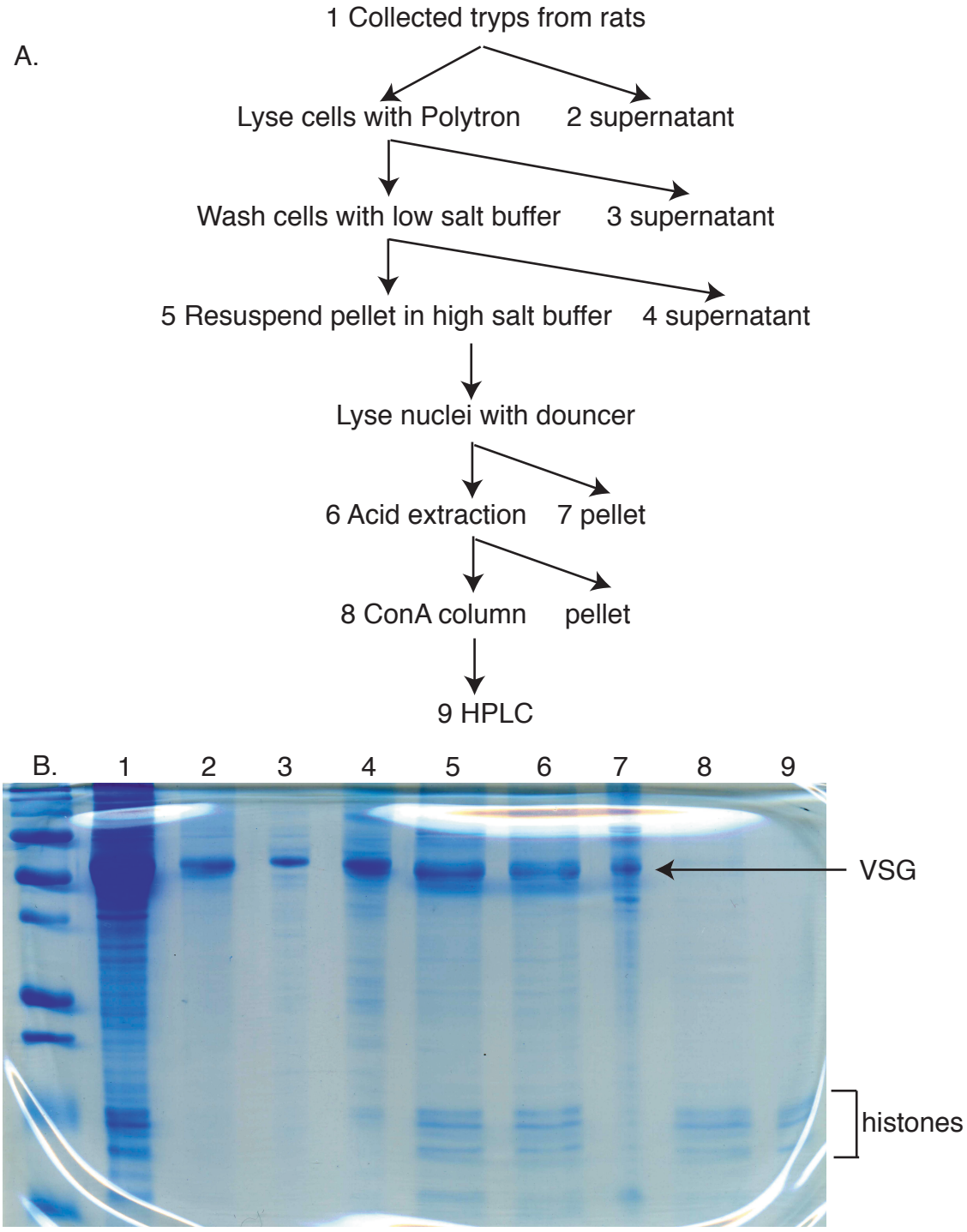


Figure 3.3. Final yield of crude histone purification. (A) Flow diagram of purification protocol. (B) SDS-PAGE of sample from each step of the purification process. Lanes are as follows: (1) Starting material, (2) supernatant following initial centrifugation, (3) soluble fraction following cell lysis, (4) low salt buffer wash, (5) crude nuclei resuspended in high salt buffer, (6) soluble fraction following mechanical disruption of nuclei, (7) insoluble fraction following mechanical disruption of nuclei, (8) proteins solubilized in sulfuric acid, (9) ConA flow through.

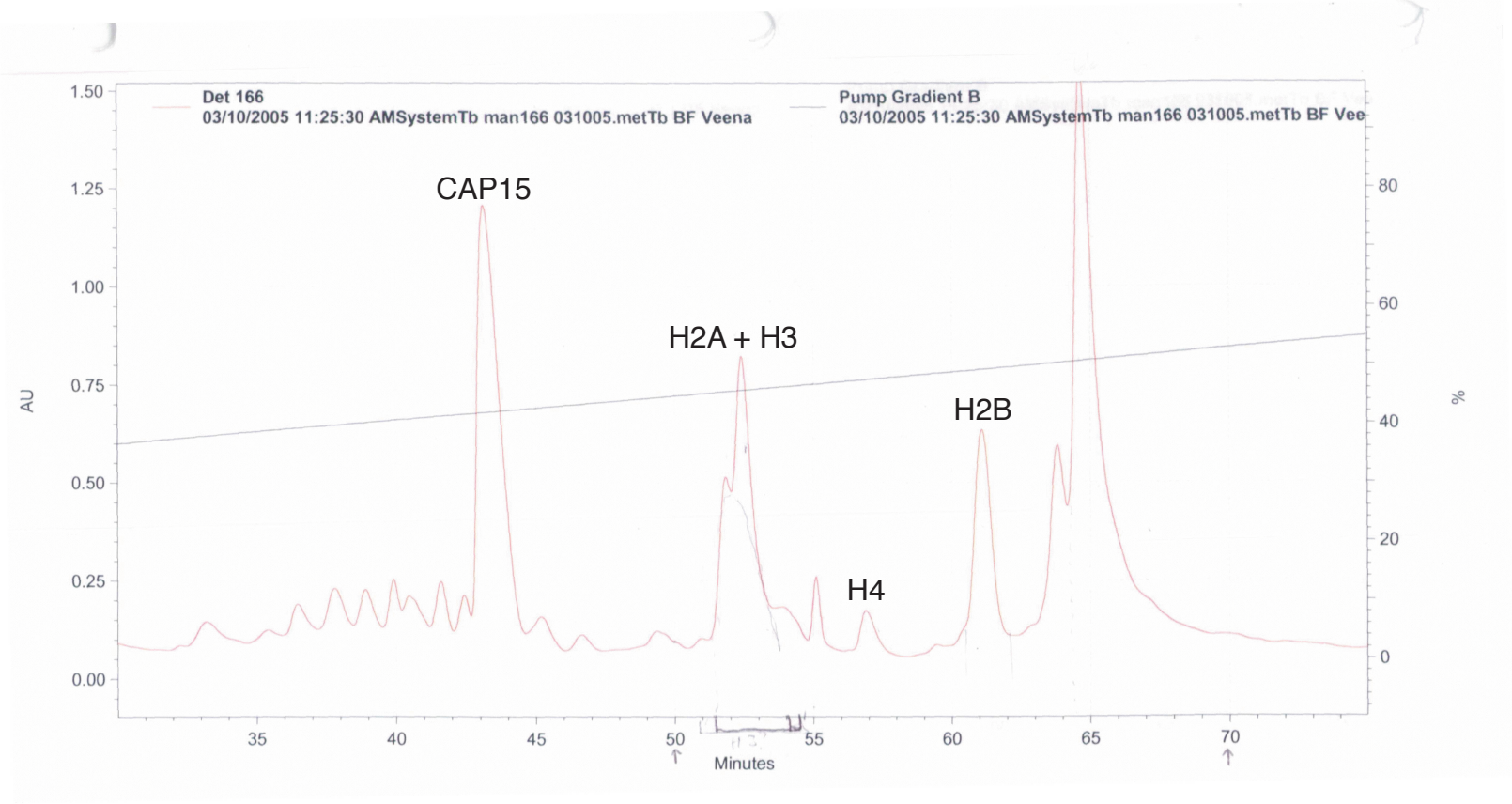


Figure 3.4. HPLC Elution Profile of BF 221 Histones. Crude histone preparation was fractionated by RP-HPLC. Absorbance 214 nm. H2B and H4 were eluted individually, but H2A and H3 eluted in overlapping fractions. Additional peaks did not contain histones. CAP15 = corset-associated protein 15.

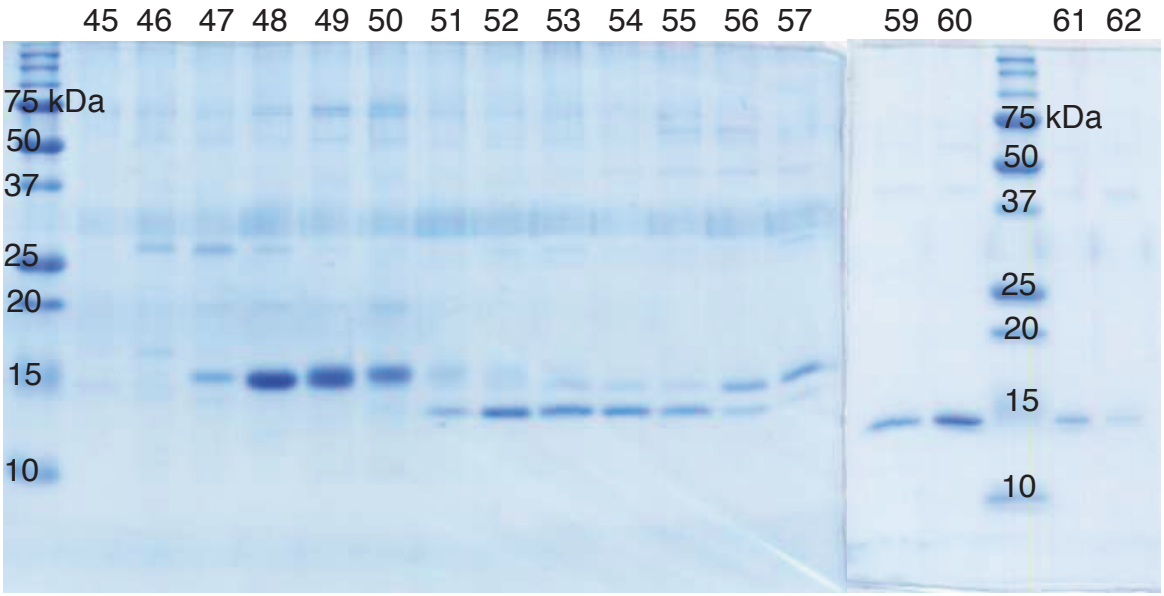


Figure 3.5. SDS-PAGE of HPLC-purified Histones. One fraction was collected each minute. One-tenth of each HPLC fraction was loaded on an SDS-PAGE and stained with Coomassie. Fractions 47-50 contain H2A and H3, fractions 51-56 contain H4, and fractions 56-62 contain H2B. Histones were identified by MALDI-TOF MS. Only fractions containing histones were loaded. Additional HPLC fractions did not contain histones.

When NP-40 was used in the protocol, our final yield of histones was approximately 30%. These samples were analyzed by Edman degradation, but NP-40 contamination of the purified histones prevented us from acquiring MS data. Substituting octyl glucoside for NP-40 prevented detergent contamination of the final samples, but resulted in a lower yield (~10%).

Edman degradation and MS analysis

A combination of MS techniques, including MALDI-TOF and ESI-MS/MS, and Edman degradation, was used to analyze *T. brucei* histones for PTMs. Edman degradation provided quantitative data for PTMs on the N-terminal 20–30 amino acids. For MS, histones were first digested with trypsin. Histones contain many arginine and lysine residues, resulting in excessive fragmentation by trypsin. In some experiments, therefore, histones were modified with propionic anhydride to prevent trypsin cleavage at lysine residues by propionylation (+56 Da) of unmodified (K) and monomethylated (Kme1) lysines, and the N-terminus. Dimethylated (Kme2), trimethylated (Kme3) and acetylated (Kac) lysines were not propionylated, but are also not cleaved by trypsin. Histone H3 was also digested with endoproteinases Asp-N and Glu-C, and H4 was digested with Glu-C.

MALDI-TOF analysis provided the molecular mass of the digested peptides. Mass shifts of 14, 28, and 42 Da may represent peptides with a

monomethyl, dimethyl, and trimethyl or acetyl group, respectively. The peptides were then sequenced by tandem MS to assign PTMs to their parent residue. Acetyl (42.0105 Da) and trimethyl (42.0468 Da) groups are difficult to distinguish by tandem MS. I relied on data generated by Edman degradation to help us assign mass shifts of 42 Da. Acetylated lysines also produce a characteristic immonium ion at m/z 126, and trimethylated lysines have a characteristic ion at MH^+-59, both of which were used to make assignments [180].

H2A Modifications

Analysis of the first 22 amino acids of H2A by Edman degradation demonstrated that 60% of A1 is monomethylated, and ~1% of K4 is acetylated (Fig. 3.6A, Table 3). The monomethylation of alanine may be a novel histone modification and was also found in *T. brucei* PF histones and *T. cruzi* [165,181].

Tandem MS data were acquired for trypsin-digested H2A (Fig. 3.7A). Peptides that were analyzed are shown in red. Trypsin digestion produced several spectra representing modified C-terminal peptides, demonstrating that K115, K119, K120, K122, K125, and K128 can be acetylated (Fig. 3.8-12, Table 1). In PF, all but one of these acetylated lysines (K122) were identified. I assigned these modifications as acetylation rather than trimethylation, because studies of PF histones showed that H2A exists in multiple acetylation states [165]. The presence of the immonium ion for acetylated lysine at m/z 126

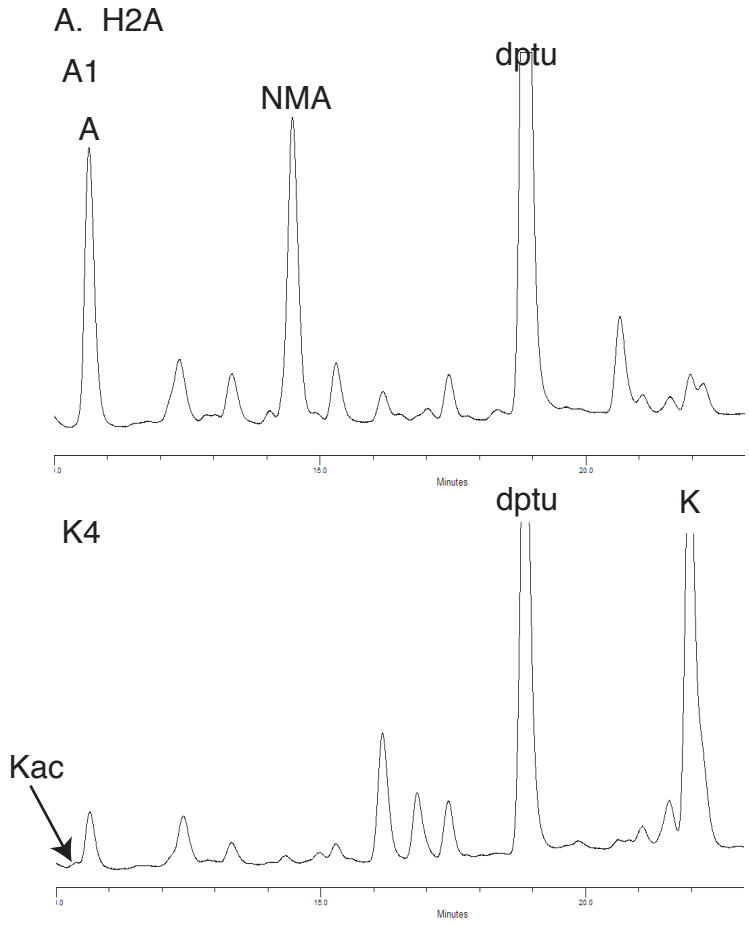
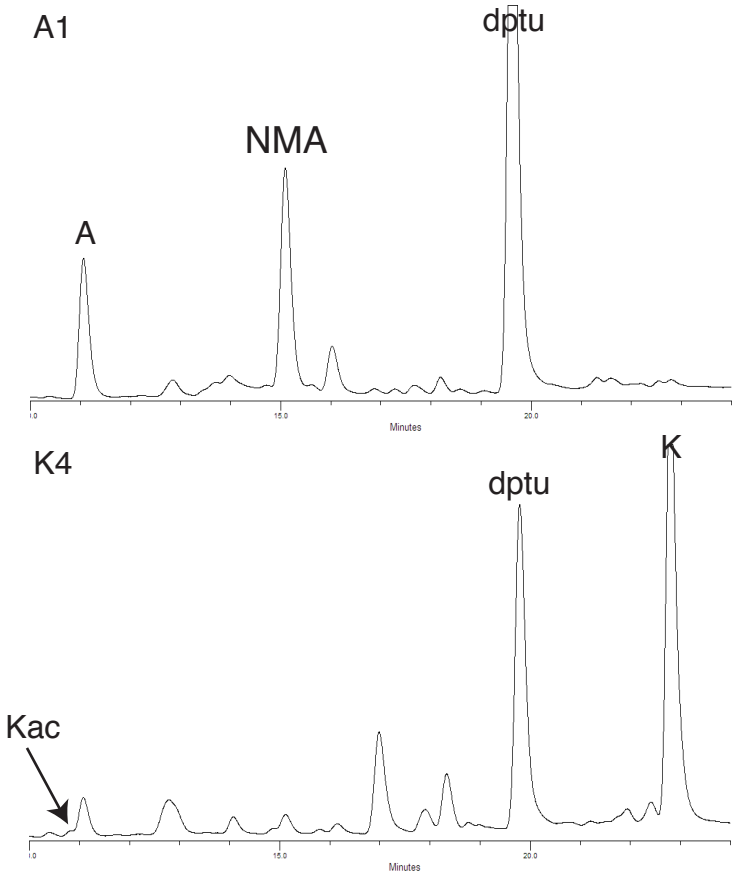


Figure 3.6. Edman degradation identifies PTMs at the N-termini of H2A and H2B. The PTMs were verified by comparison to known standards. (A) The PTH HPLC trace for cleavage cycle 1 of H2A shows that ~60% of A1 is N-methylated (NMA). The trace for cycle 4 shows that ~1% of K4 is acetylated. (B) The trace for cycle 1 of H2B shows that ~60% of A1 is N-methylated. The trace for cycle 4 shows that ~1% of K4 is acetylated. dptu is diphenylthiourea.

B. H2B



A. H2A trypsin
ATPKQAVKKASKGGSSRSVKAGLIFPVGRVGTLLRRGQYARRIGASGAVYMAA
VLEYLTAELLELSVKAAAQQTKKTKRLTPRTVTLAVRHDDDLGALLRNVTMSRG
GVMPSLNKALAKKQKSGKHAKATPSV

B. H2A PA-trypsin
ATPKQAVKKASKGGSSRSVKAGLIFPVGRVGTLLRRGQYARRIGASGAVYMAA
VLEYLTAELLELSVKAAAQQTKKTKRLTPRTVTLAVRHDDDLGALLRNVTMSRG
GVMPSLNKALAKKQKSGKHAKATPSV

C. H2B PA-trypsin
ATPKSTPAKTRKEAKKTRRQRKRTWNVYVSRSLRSINSQMSMTSRTMKIVNSF
VNDLFERIAAEAATIVRVNRKRTLGARELQTAVRLVLPADLAKHAMAEGTKAVS
HASS

D. H3 PA-trypsin
SRTKETARTKKTITSKKSKKASKGSDAASGVKTAQRRWRPGTVALREIRQFQRS
TDLLLQKAPFQRLVREVSGAQKEGLRFQSSAILAAQEATESYIVSLLADTNRACI
HSGRVTIQPKDIHLALCLRGERA

E. H4 PA-trypsin
AKGKKSGEAKGSQKRQKKVLRENVRGITRGSIRRLARRGGVKRISGVIYDEVRG
VLKSFVEGVVRDATAYTEYSRKKTVTAVDVVNALRKRGKILYGYA

Figure 3.7. *T. brucei* histone sequences with regions for which tandem MS data were acquired following endoproteinase digestion highlighted in red. (A) H2A trypsin digest. (B) H2A PA-trypsin digest. (C) H2B PA-trypsin digest. (D) H3 PA-trypsin digest. (E) H4 PA-trypsin digest.

TABLE 1. Posttranslational modifications on histones H2A, H2B, and H3		
Histone	Digest[a]	Modified Peptide[b]
H2A	Trypsin	GGVMPSLN^{115}**K**$_{ac}$ALAK (107-119)
		GGVMPSLN^{115}**K**$_{ac}$ALA^{119}**K**$_{ac}$ ^{120}**K**$_{ac}$QK (107-122)
		ALA^{119}**K**$_{ac}$ ^{120}**K**$_{ac}$Q^{122}**K**$_{ac}$SGK (116-125)
		SG^{125}**K**$_{ac}$HA^{128}**K**$_{ac}$ATPSV (123-133)
		HA^{128}**K**$_{ac}$ATPSV (126-133)
	PA-Trypsin	^{1}**A**$_{ac}$TPKQAVKKASKGGSSR (1-17)
		^{1}**A**$_{me1}$TPKQAVKKASKGGSSR (1-17)
H2B	PA-Trypsin	^{1}**A**$_{me1}$TPKSTPAKTR (1-11)
		^{1}**A**$_{me2}$TPKSTPAKTR (1-11)
		^{12}**K**$_{ac}$EAKKTR (12-18)
		KEAK^{16}**K**$_{ac}$TR (12-18)
H3	Trypsin	AS^{23}**K**$_{ac}$GSDAASGVK (21-32)
		GSDAASGV^{32}**K**$_{me3}$TAQR (24-36)
	PA-Trypsin	^{1}**S**$_{ac}$RTKETAR (1-8)
		EVSGAQ^{76}**K**$_{me1}$EGLR (70-80)
		EVSGAQ^{76}**K**$_{me2}$EGLR (70-80)
		EVSGAQ^{76}**K**$_{me3}$EGLR (70-80)

[a] Histones, which in some cases were modified with propionic anhydride, were digested with trypsin.

[b] Tryptic peptides were sequenced by tandem MS, and peptides that have covalent modifications are listed here.

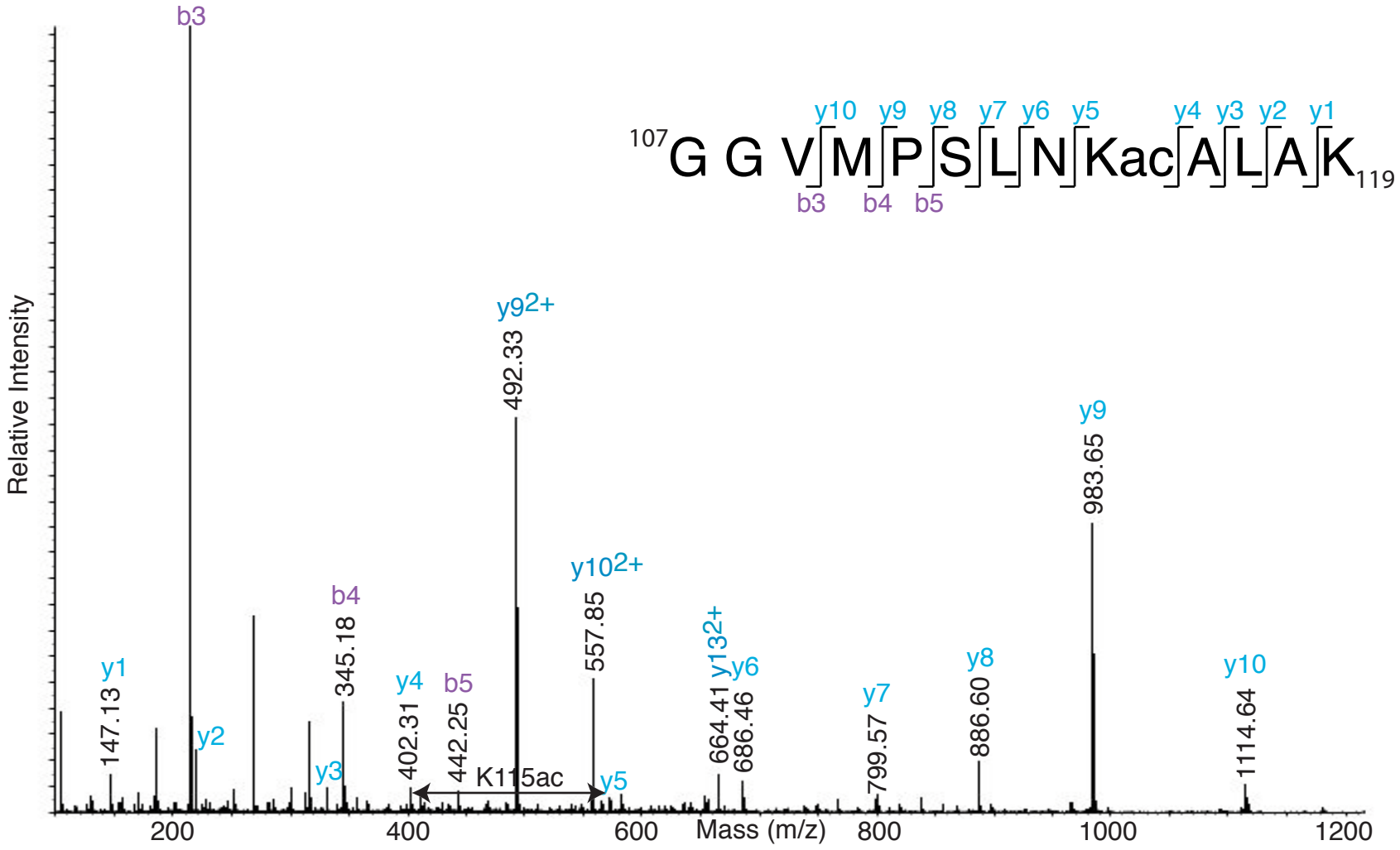


Figure 3.8. Histone H2A lysine 115 is acetylated. H2A tryptic peptide 107-119 (GGVMPSLNKALAK) was sequenced by tandem MS. Ion pair y4-y5 demonstrates that K115 is acetylated. m/z 664.4^{2+}; theoretical MW = 1327.7 Da; measured MW = 1327.8 Da.

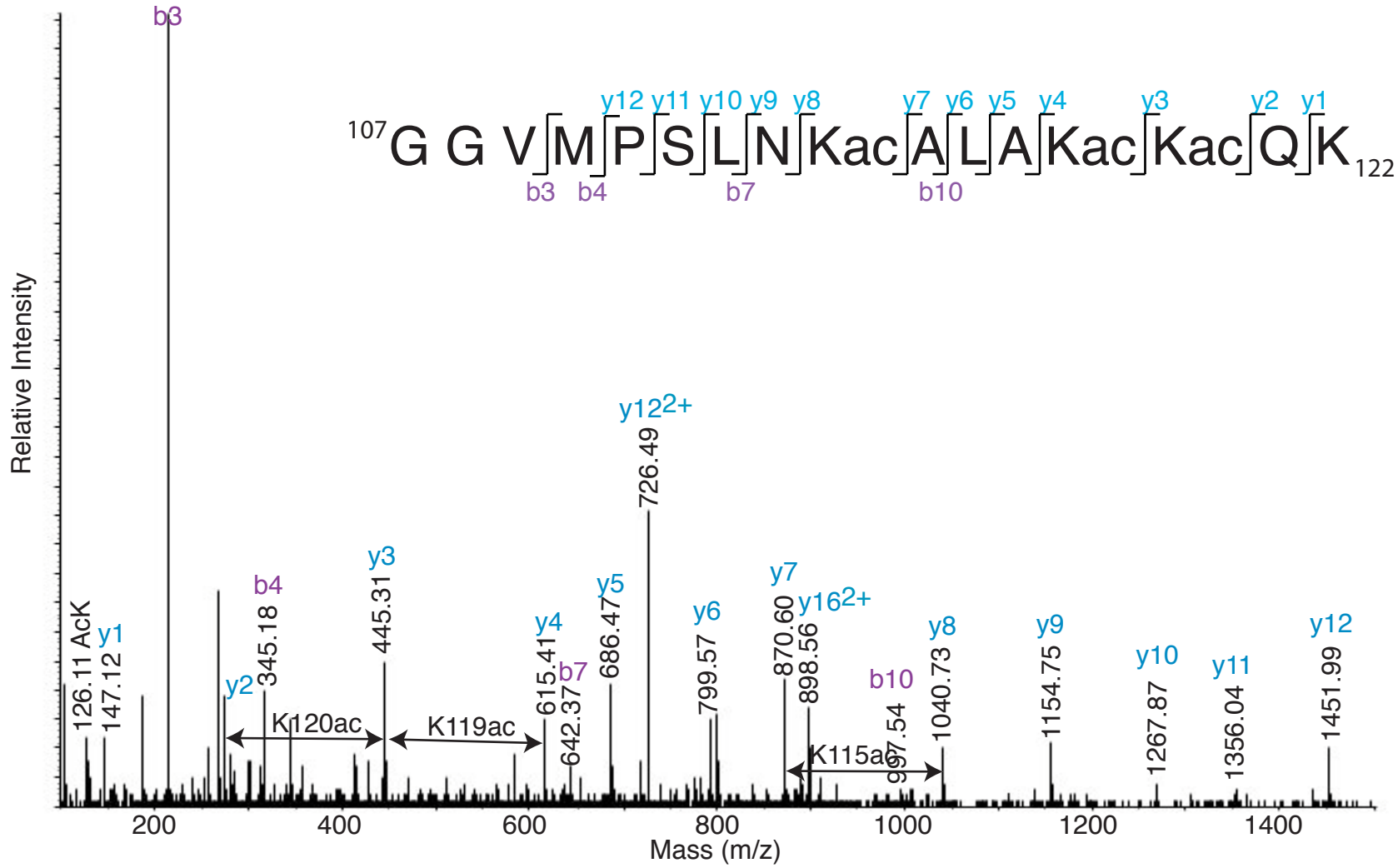


Figure 3.9. Histone H2A lysines 115, 119, and 120 are acetylated. H2A tryptic peptide 107-122 (GGVMPSLNKALAKKQK) was sequenced by tandem MS. Ion pairs y2-y3, y3-y4, and y7-y8 show that K120, K119, and K115 are acetylated, respectively. m/z 898.91^{2+}; theoretical MW = 1796.0 Da; measured MW = 1796.8 Da.

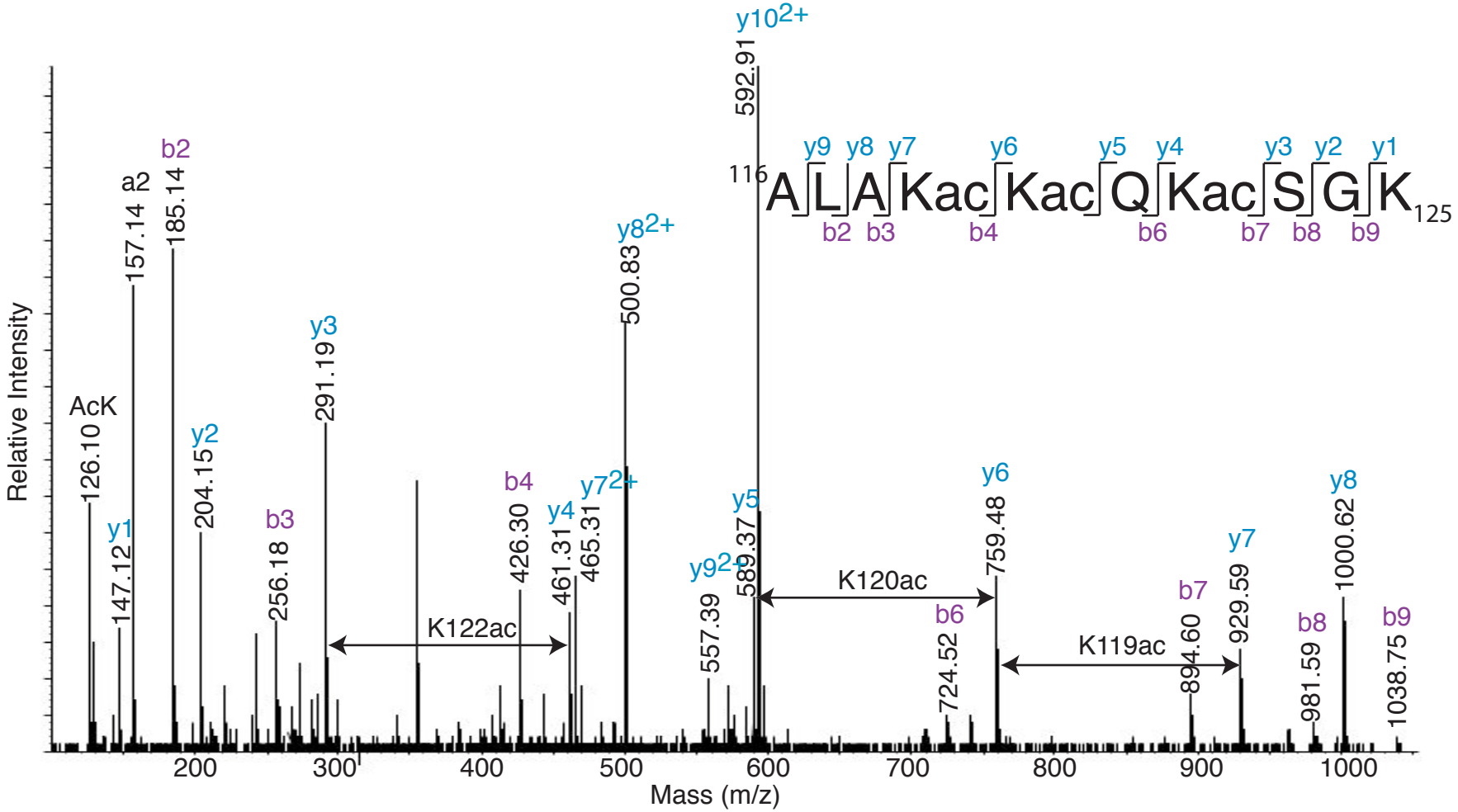


Figure 3.10. Histone H2A lysines 119, 120, and 122 are acetylated. H2A tryptic peptide 116-125 (ALAKKQKSGK) was sequenced by tandem MS. Ion pairs y6-y7 and b3-b4 demonstrate that K119 is acetylated; y5-y6 demonstrates that K120 is acetylated; y3-y4 and b6-b7 demonstrate that K122 is acetylated. The presence of the acetyl lysine immonium ion (m/z 126) supports the claim that the +42 Da shift for each lysine is due to acetylation, not trimethylation. m/z 592.9^{2+}; theoretical MW = 1184.7 Da; measured MW = 1184.8 Da.

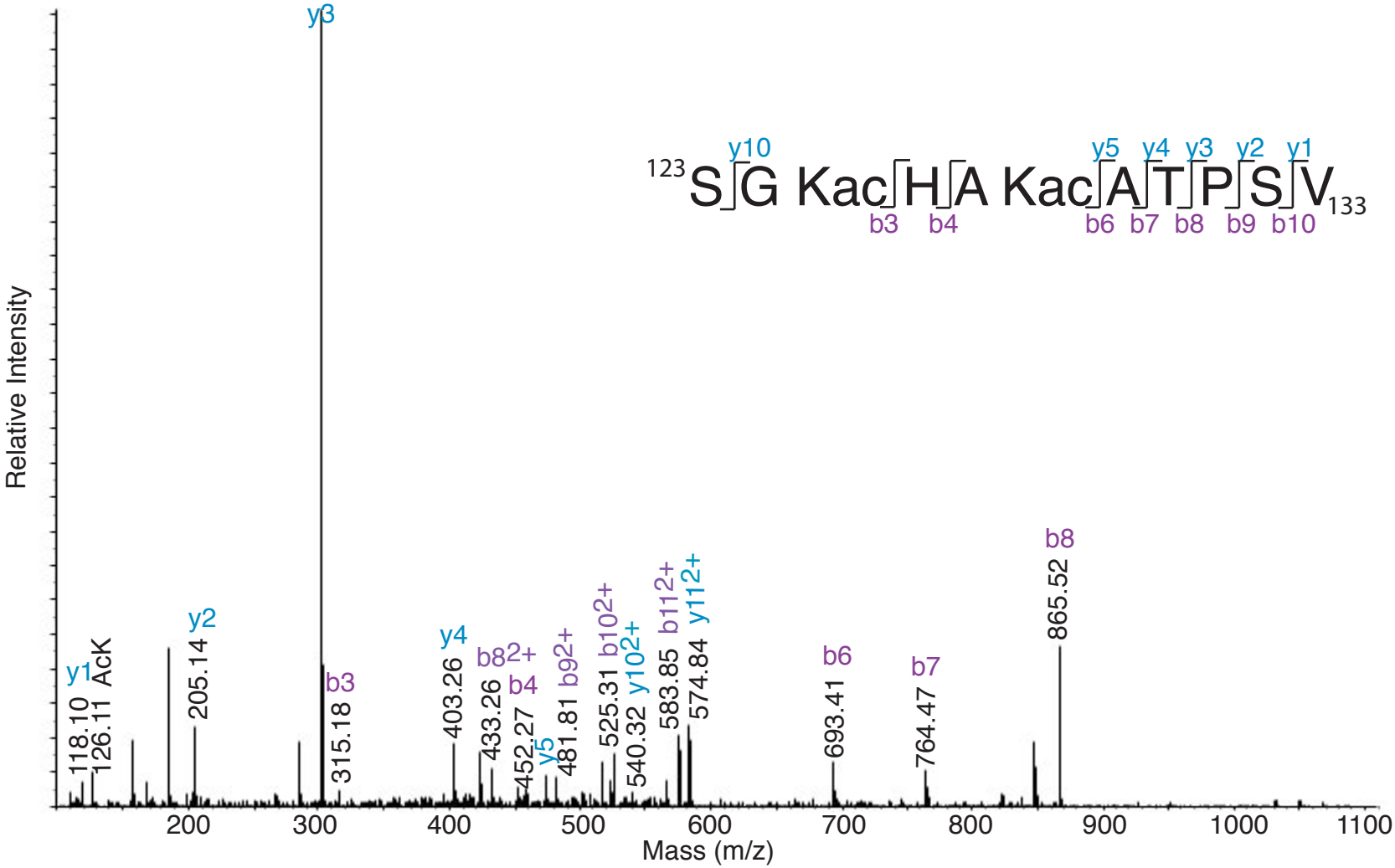


Figure 3.11. Histone H2A lysines 125 and 128 are acetylated. H2A tryptic peptide 123-133 (SGKHAKATPSV) was sequenced by tandem MS. Sequencing of this peptide is incomplete, but the b series (b3, b4, b6) support the assignment of K125 and K128 acetylation. m/z 583.9^{2+}; theoretical MW = 1166.6 Da; measured MW = 1166.8 Da.

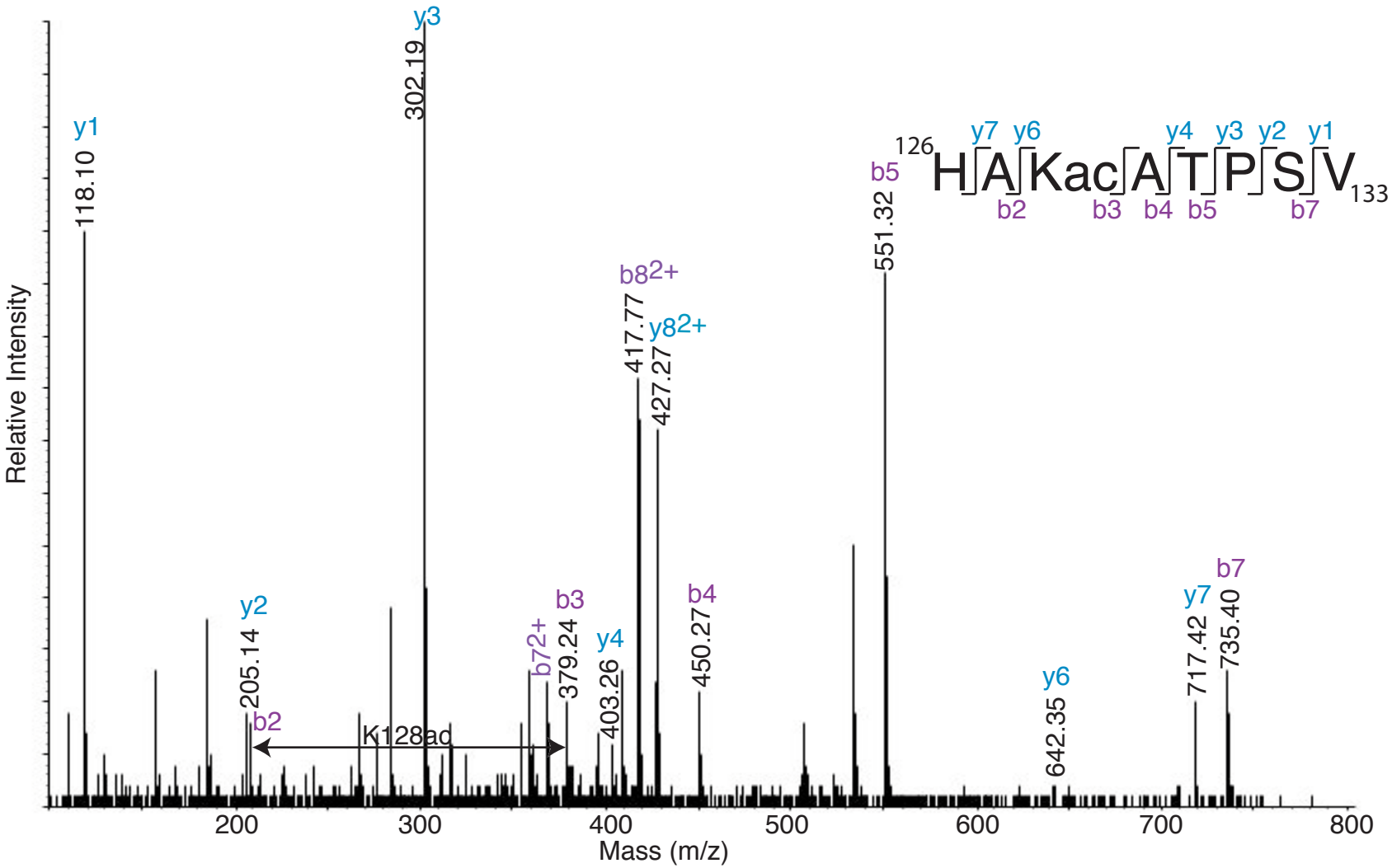


Figure 3.12. Histone H2A lysine 128 is acetylated. H2A tryptic peptide 126-133 (HAKATPSV) was sequenced by tandem MS. Ion pairs b2-b3 and and y4-y6 demonstrate that K128 is acetylated. m/z 426.8^{2+}; theoretical MW = 852.4 Da; measured MW = 852.6 Da.

supported this assignment. Multiple modifications of the H2A C-terminus have not been reported in other organisms. No peptides corresponding to the N-terminus were found after trypsin digest, most likely because there are five lysines and one arginine present that would result in excessive fragmentation of this sequence.

Next, H2A was modified with propionic anhydride prior to trypsin digest (hereafter referred to as PA-trypsin digest) to acquire information about the N-terminus (Fig. 3.7B). MALDI-TOF analysis of the PA-trypsin-digested H2A demonstrated that the N-terminus is present in three modified states (Fig. 3.13A). Tandem MS sequencing of these species demonstrated that A1 can be monomethylated, acetylated, or unmodified (Fig. 3.13B-D, Table 1). The y series of these spectra are identical. The mass shift from modified A1 can be observed by subtracting y16 from the MH^+ value of the spectra. Also, the b series are shifted by the mass of the modification (see the b6 ion in all three spectra). Premature cleavage of peptide 1-17 resulted in peptide 3-17, which was then sequenced, adding several additional peaks representing the y series of peptide 3-17 to these spectra. Peptides representing the C-terminus (H2A 107-133) were also found, but were too large to interpret.

In other organisms, the C-termini of H2A and H2B are ubiquitinated [182,183]. Alignment of the H2A C-terminus (Fig. 3.28B) suggests that *T. brucei* H2A K122 might be the homologue of the ubiquitinated H2A K119 in other organisms. We therefore performed a directed search for ubiquitinated H2A

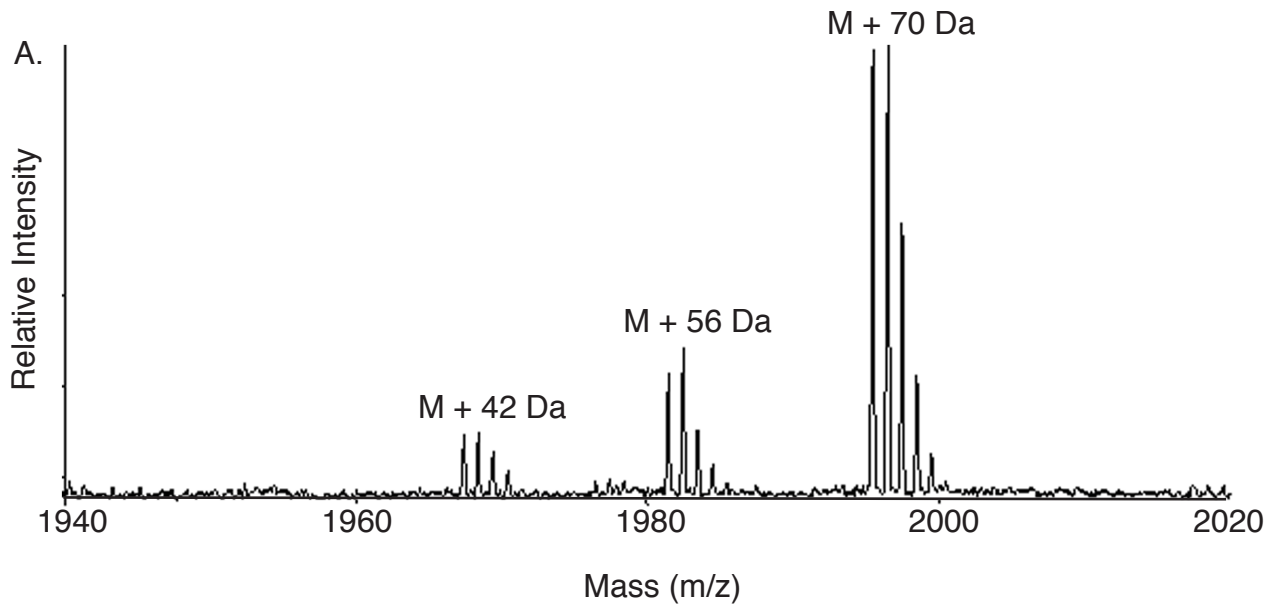
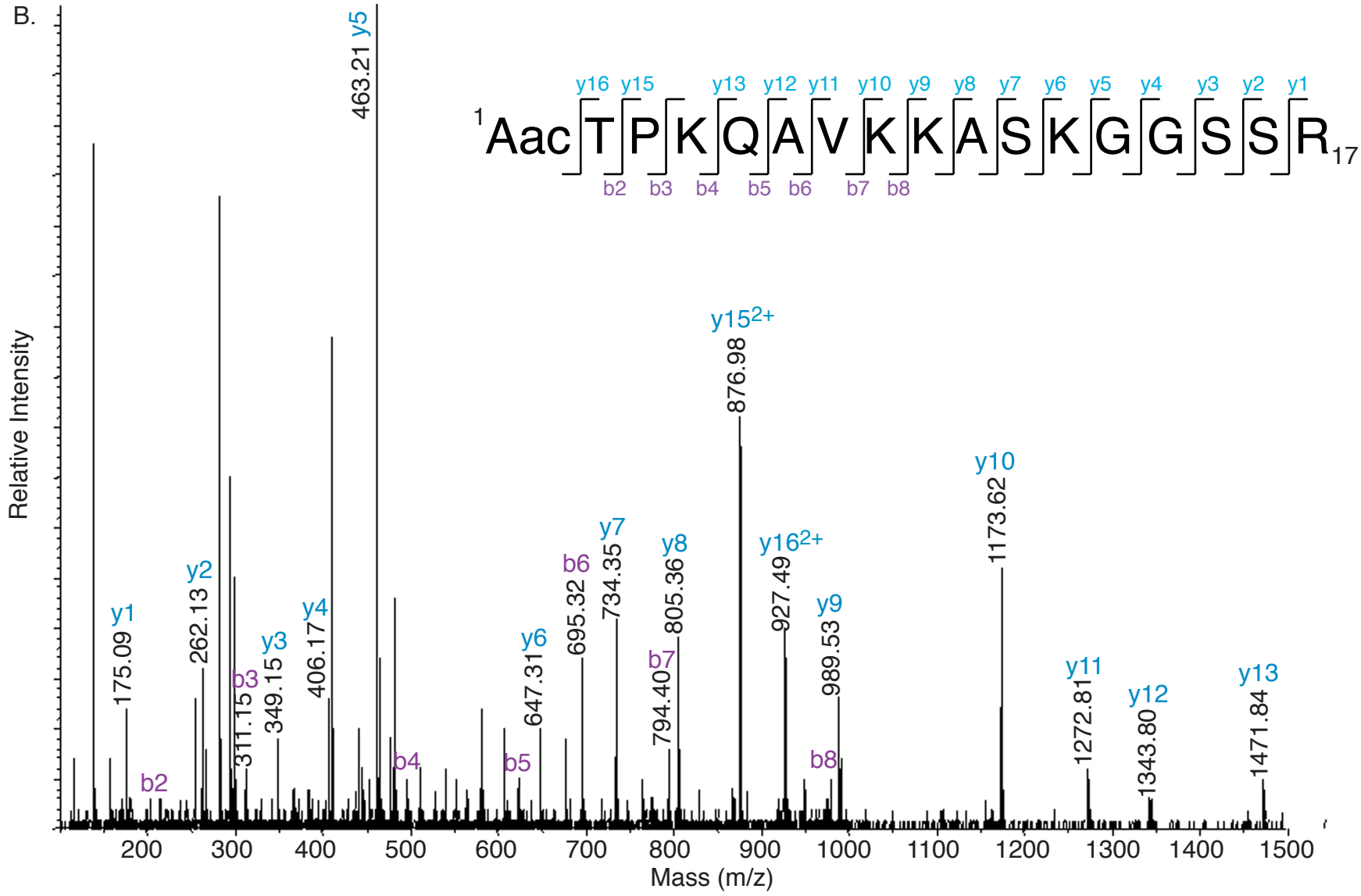
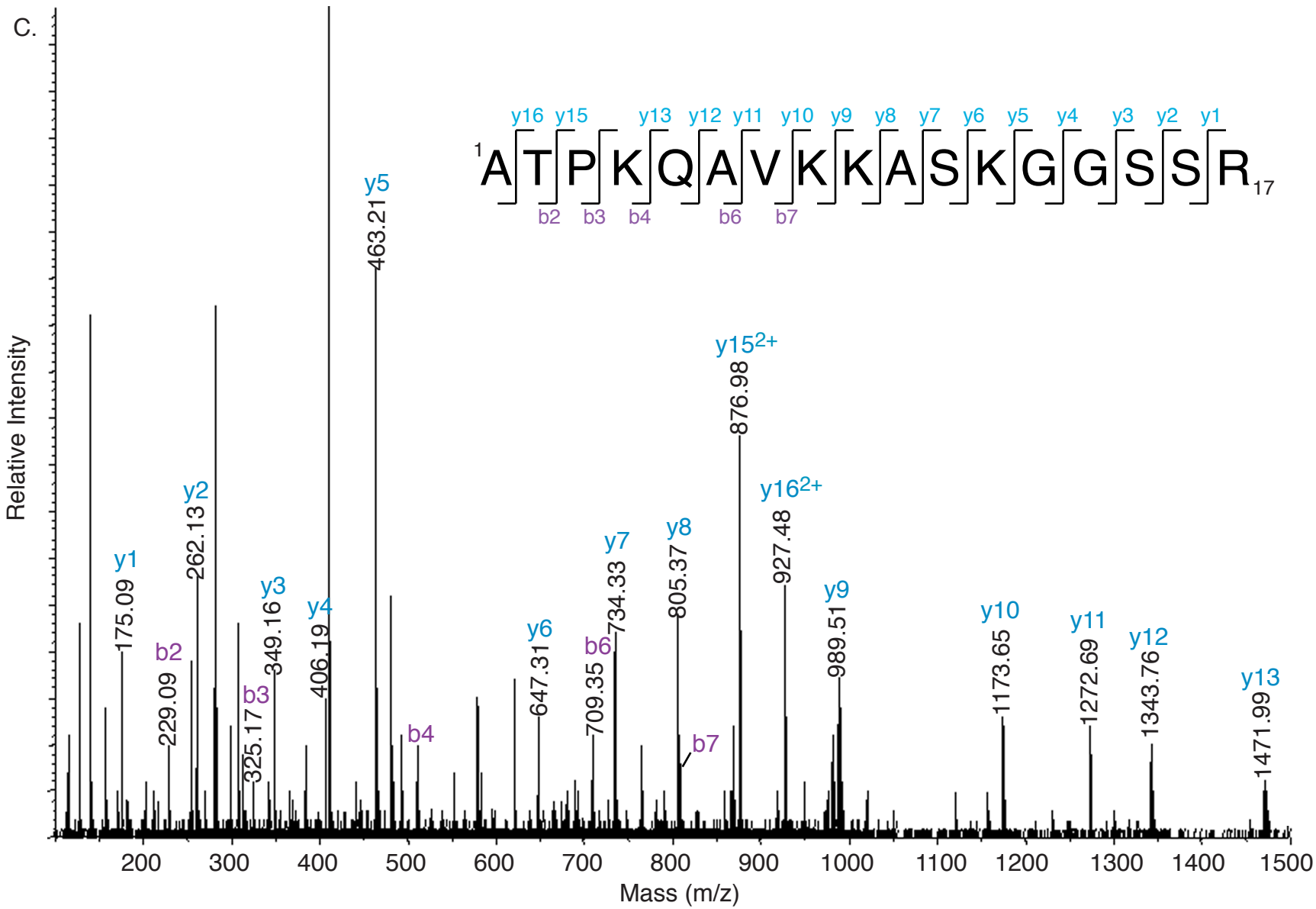
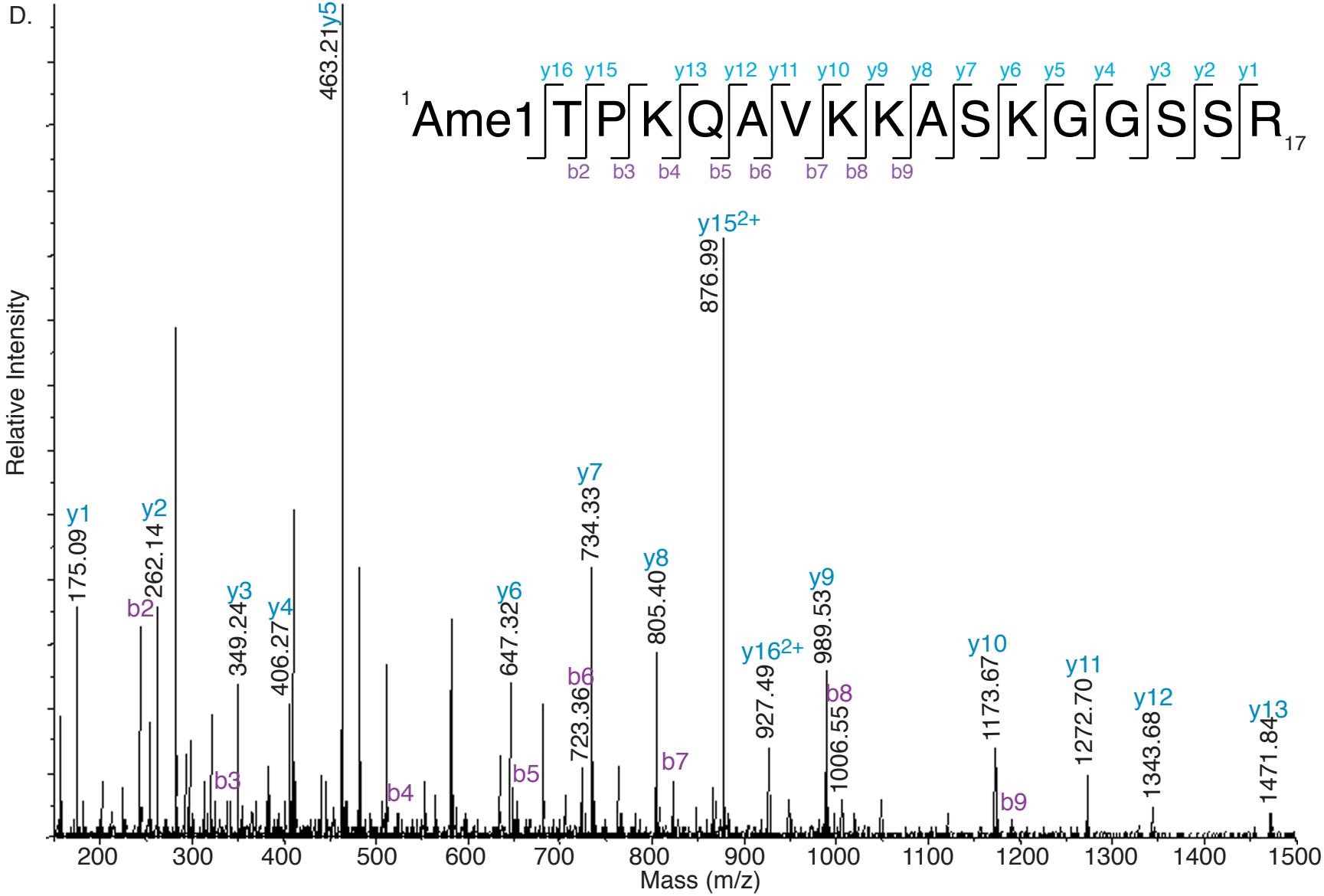


Figure 3.13. Histone H2A alanine 1 may be unmodified, acetylated, or monomethylated. Tryptic peptide 1-17 (ATPKQAVKKASKGGSSR) of propionylated H2A was subjected to MALDI-TOF analysis and sequencing by tandem MS. (A) MALDI-TOF shows that peptide 1-17 is present in three modification states. M is the molecular mass ion of peptide 1-17 with all lysines propionylated (1924 Da). Tandem MS sequencing of peaks labeled M + 42 Da, M + 56 Da, and M + 70 Da species indicate that H2A A1 may be acetylated, propionylated (meaning it is unmodified), or both propionylated and monomethylated, respectively. (B) MS/MS shows that H2A A1 is acetylated (m/z 984.1^{2+}). Theoretical MW = 1967.0 Da; measured MW = 1967.2 Da. (C) MS/MS shows that H2A A1 is unmodified (m/z 991.1^{2+}). Theoretical MW = 1981.0 Da; measured MW = 1981.2 Da. (D) MS/MS shows that H2A A1 is monomethylated (m/z 998.1^{2+}). Theoretical MW = 1995.0 Da; measured MW = 1995.2 Da. These assignments were made by looking at the b ions, which are all shifted relative to the unmodified species. Also, the difference between the MH+ of the spectrum and y16 provides the mass of the modified A1.


B.
Relative Intensity
Mass (m/z)
200
300
400
500
600
700
800
900
1000
1100
1200
1300
1400
1500
463.21 y5
175.09 y1
262.13 y2
311.15 b3
349.15 y3
406.17 y4
647.31 y6
695.32 b6
734.35 y7
794.40 b7
805.36 y8
876.98 y15 2+
927.49 y16 2+
989.53 y9
1173.62 y10
1272.81 y11
1343.80 y12
1471.84 y13
b2
b4
b5
b8
1 Aac T P K Q A V K K A S K G G S S R 17
y16 y15 y13 y12 y11 y10 y9 y8 y7 y6 y5 y4 y3 y2 y1
b2 b3 b4 b5 b6 b7 b8


C.
y16 y15 y13 y12 y11 y10 y9 y8 y7 y6 y5 y4 y3 y2 y1
1ATPKQAVKKASKGGSSR17
b2 b3 b4 b6 b7
Relative Intensity
y5
463.21
y15 2+
876.98
y16 2+
927.48
y2
262.13
y1
175.09
y3
349.16
y4
406.19
y6
647.31
y7
734.33
y8
805.37
y9
989.51
y10
1173.65
y11
1272.69
y12
1343.76
y13
1471.99
b2
229.09
b3
325.17
b4
b6
709.35
b7
200 300 400 500 600 700 800 900 1000 1100 1200 1300 1400 1500
Mass (m/z)


D.
1Ame1TPKQAVKKASKGGSSR17
y16 y15 y13 y12 y11 y10 y9 y8 y7 y6 y5 y4 y3 y2 y1
b2 b3 b4 b5 b6 b7 b8 b9
Relative Intensity
Mass (m/z)
463.21y5
y15 2+ 876.99
y1 175.09
y2 262.14
b2
y3 349.24
y4 406.27
b3
b4
y6 647.32
b5
y7 734.33
b6 723.36
y8 805.40
b7
y16 2+ 927.49
y9 989.53
b8 1006.55
y10 1173.67
b9
y11 1272.70
y12 1343.68
y13 1471.84
200 300 400 500 600 700 800 900 1000 1100 1200 1300 1400 1500

K122. MS can identify sites of ubiquitination by looking for a GG dinucleotide linkage (114 Da) to lysine, which results from tryptic cleavage of covalently linked ubiquitin (-RLRGG) [184]. The tandem MS spectrum of a peptide with m/z 1029.26^{3+} displayed partial y-series (y1-y6, y8, y15) and b-series (b3, b12, b22-24) ions, demonstrating that the spectrum represents propionylated H2A peptide 107–133 (Fig. 3.14). The molecular weight of this species and the y- and b-ions indicate that K115 and K128 are propionylated, and, of the remaining 4 lysines (K119, K120, K122, and K125), one may be ubiquitinated and three are acetylated.

H2B Modifications

Edman degradation of the first 25 amino acids of H2B demonstrated that 60% of A1 is monomethylated, and ~1% of K4 is acetylated (Fig. 3.6B, Table 3). Both H2A and H2B begin with the sequence ATPK, and both are modified at A1 and K4 to the same degree. PA-trypsin digestion of H2B provided MS/MS data for the lysine-rich N-terminus (Fig. 3.7C). Quadropole-quadropole-TOF (QqTOF) analysis of the H2B digest showed that N-terminal peptide 1-11 can be present in three modification states (Fig. 3.15A). Tandem MS sequencing of these peaks showed that A1 can be unmodified, mono-, or di-methylated (Fig. 3.15B-D, Table 1). Again, the y series of the three spectra are identical. The modification of A1

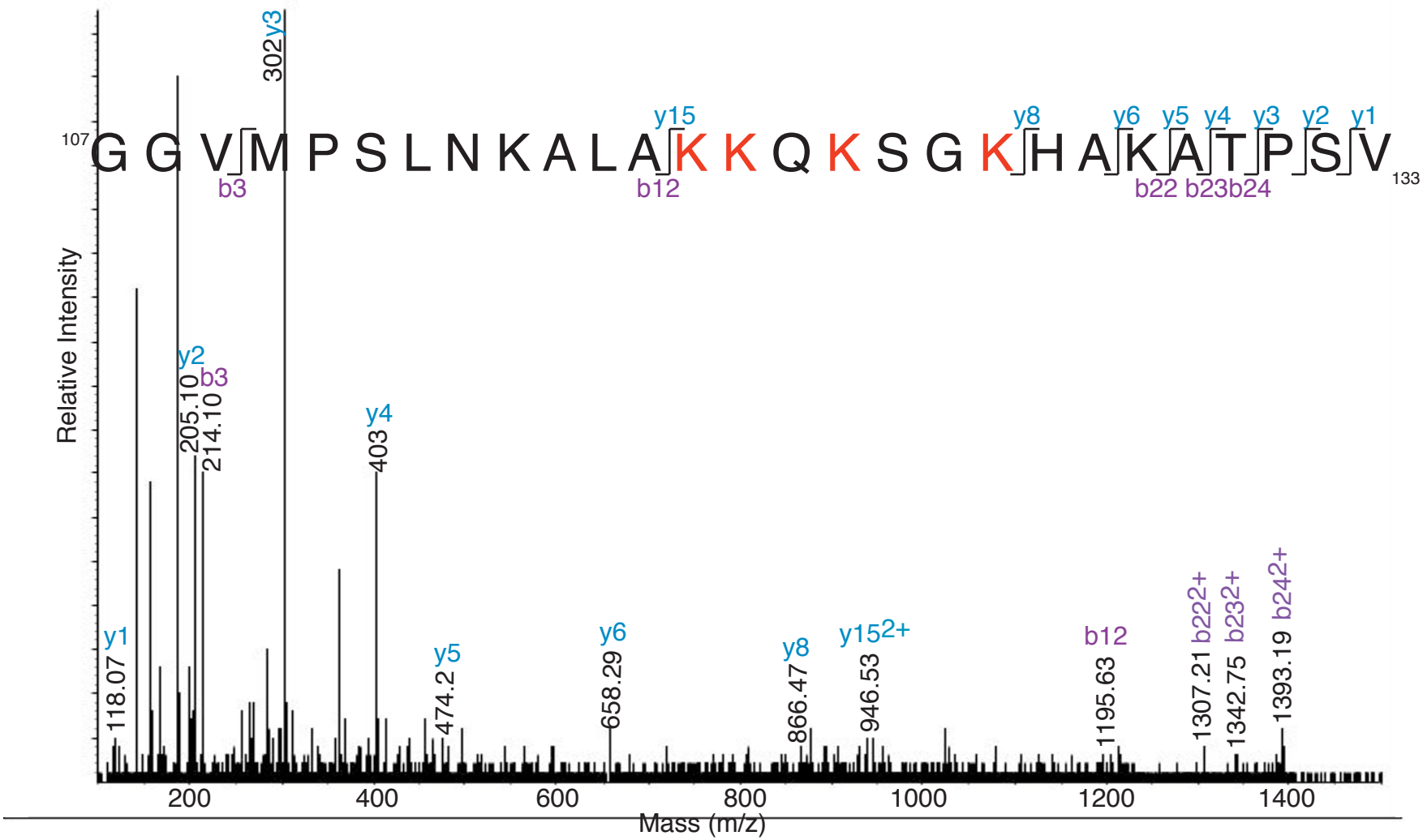


Figure 3.14. The C-terminus of histone H2A is ubiquitinated. Propionylated H2A was digested with trypsin, producing peptide 107-133 (GGVMPSLNKALAKKQKSGKHAKATPSV). Sequencing of this peptide by tandem MS shows that three lysines are acetylated and one is ubiquitinated (modified lysines shown in red). m/z 1029.26^{2+} Theoretical MW = 2057.5 Da. measured MW = 2057.5 Da. See text for details.

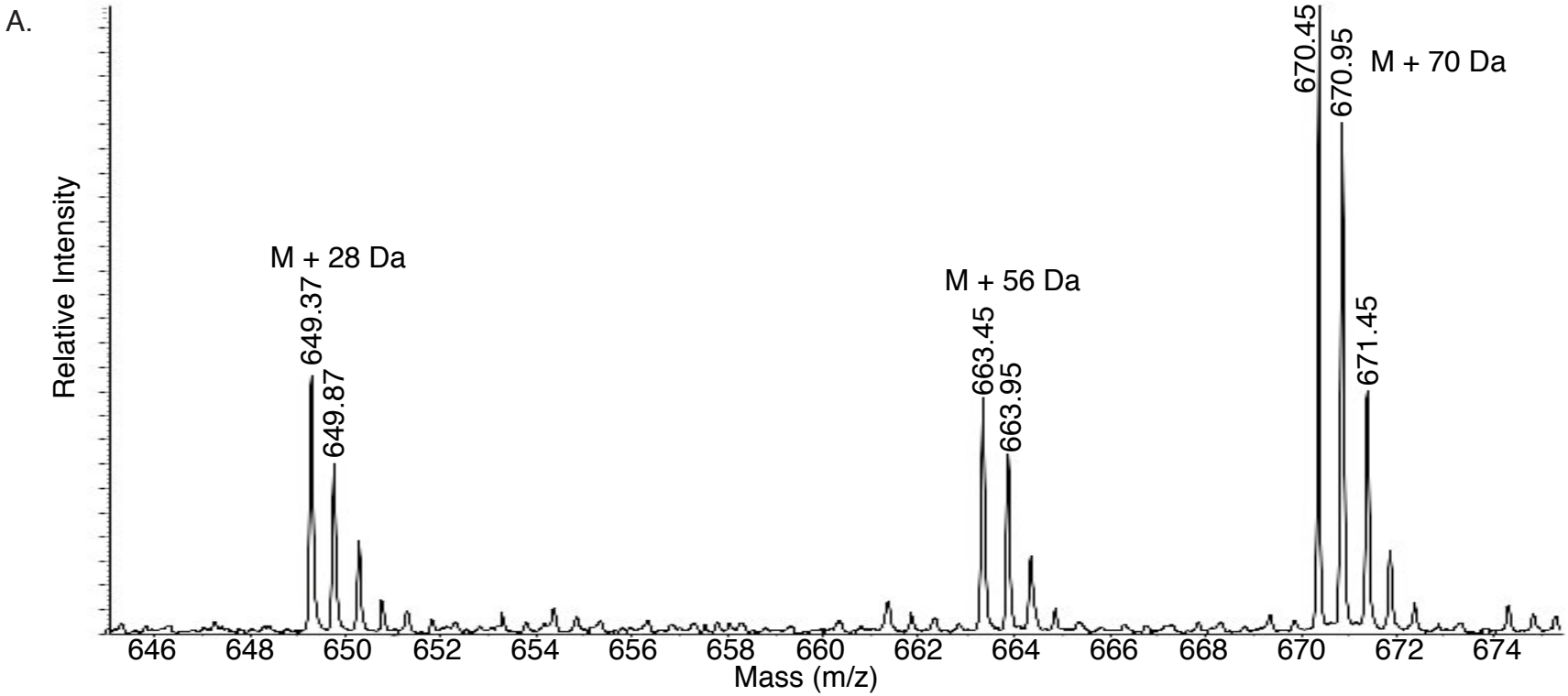
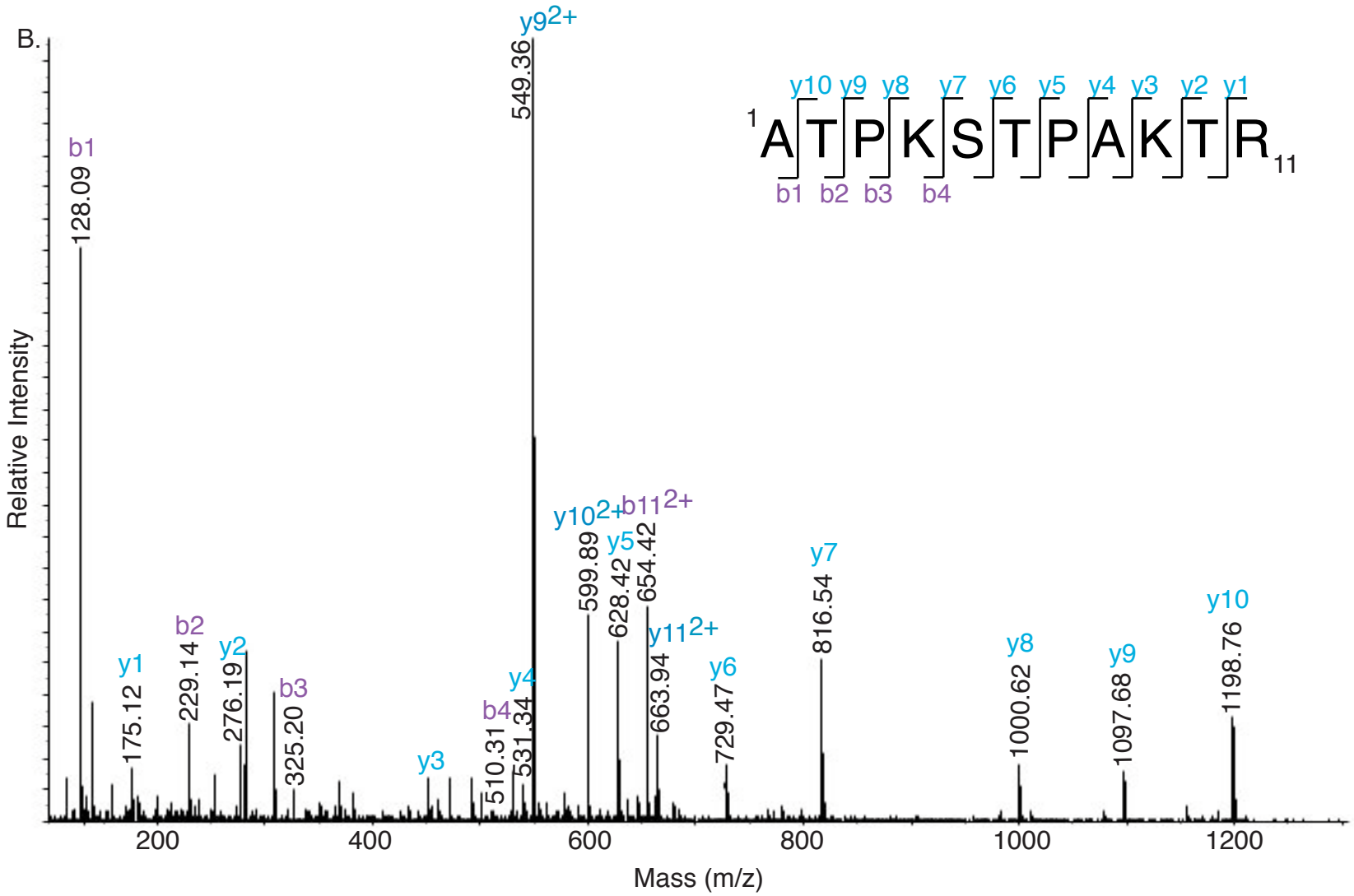
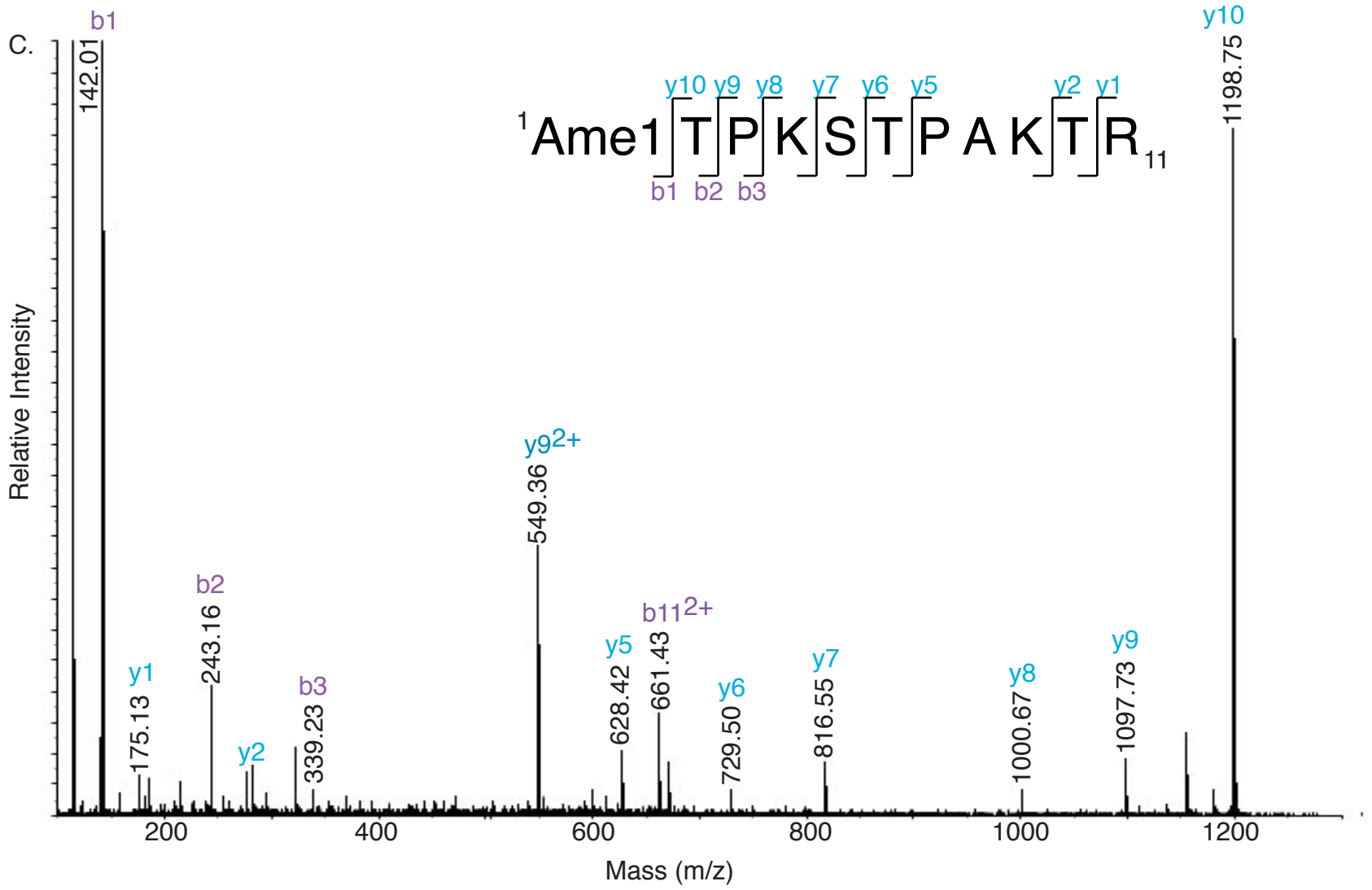
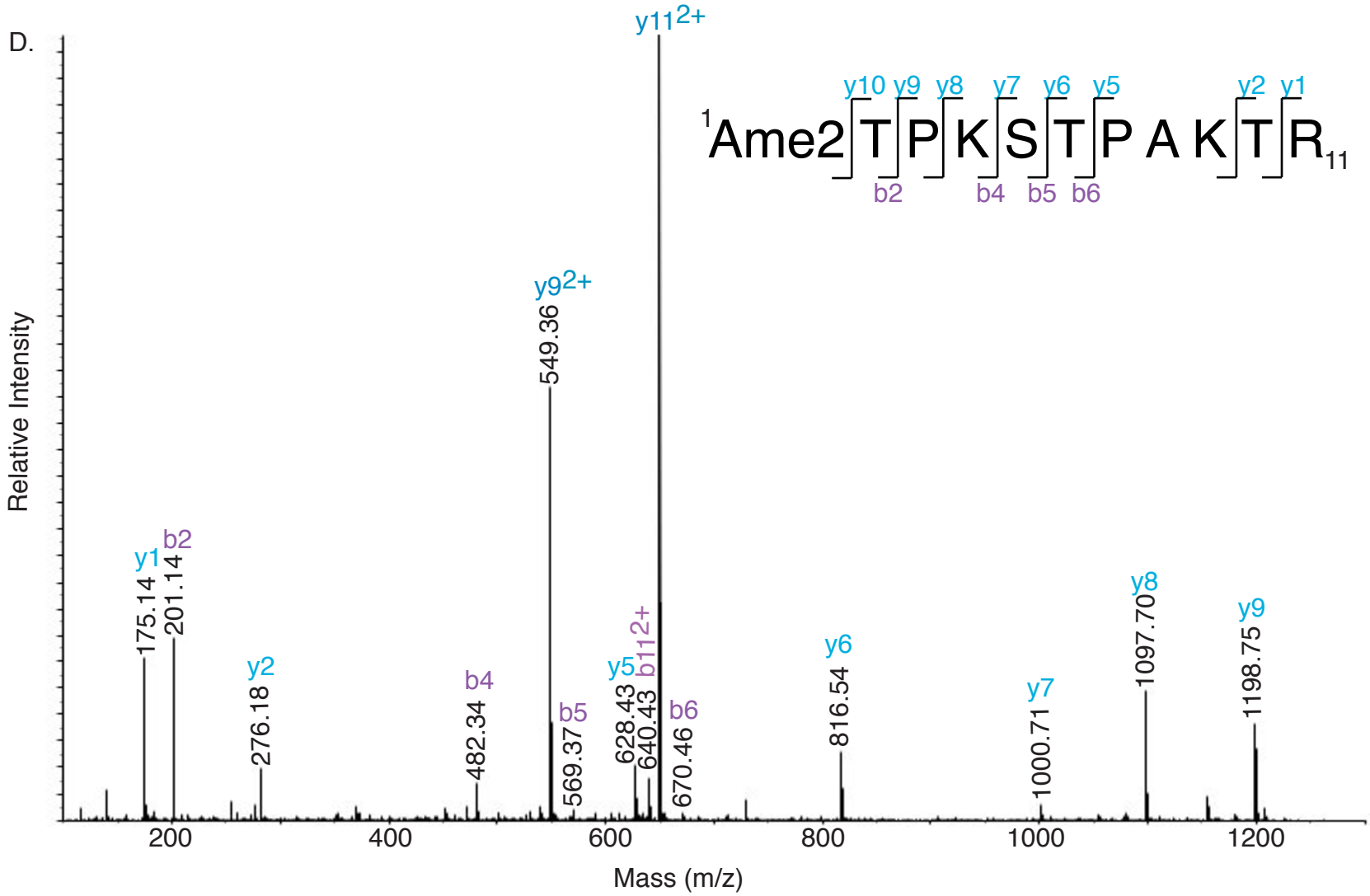


Figure 3.15. Histone H2B alanine A1 can be unmodified, mono-, or dimethylated. Propionylated H2B was digested with trypsin, producing peptide 1-11 (ATPKSTPAKTR). (A) QqTOF data shows that peptide 1-11 is present in three modification states. M is the molecular mass ion of the peptide when all of the lysines are propionylated. Sequencing by tandem MS confirmed that these peaks correspond to H2B peptide 1-11. The isotope pattern demonstrates that these peaks represent the +2 species. (B) MS/MS shows that the peak labeled M + 56 Da represents unmodified A1, meaning that it is propionylated (m/z 663.4^{2+}). Theoretical MW = 1325.7 Da; measured MW = 1325.8 Da. (C) MS/MS shows that the peak labeled M + 70 Da represents monomethylated and propionylated A1 (m/z 670.4^{2+}). Theoretical MW = 1339.7 Da; measured MW = 1339.8 Da. (D) MS/MS shows that the peak labeled M + 28 Da represents dimethylated A1 (m/z 649.4^{2+}). Theoretical MW = 1325.7 Da; measured MW = 1325.8 Da. These assignments can be observed from the b ions, which are all shifted relative to the unmodified species. The mass difference between the MH+ of the spectrum and y10 also provides the mass of the modified A1.


B.
Relative Intensity
Mass (m/z)
200
400
600
800
1000
1200
y10 y9 y8 y7 y6 y5 y4 y3 y2 y1
1ATPKSTPAKTR11
b1 b2 b3 b4
y9 2+
549.36
b1
128.09
y1
175.12
b2
229.14
y2
276.19
b3
325.20
y3
y4
b4
510.31
531.34
y10 2+
b11 2+
599.89
y5
628.42
654.42
y11 2+
663.94
y6
729.47
y7
816.54
y8
1000.62
y9
1097.68
y10
1198.76


C.
1Ame1TPKSTPAKTR11
b1 b2 b3
y10 y9 y8 y7 y6 y5 y2 y1
Relative Intensity
Mass (m/z)
200
400
600
800
1000
1200
142.01
b1
175.13
y1
243.16
b2
339.23
b3
y2
549.36
y9 2+
628.42
y5
661.43
b11 2+
729.50
y6
816.55
y7
1000.67
y8
1097.73
y9
1198.75
y10


D.
Relative Intensity
Mass (m/z)
200
400
600
800
1000
1200
y11 2+
y9 2+
549.36
y1
175.14
b2
201.14
y2
276.18
b4
482.34
b5
569.37
y5
628.43
b11 2+
640.43
b6
670.46
y6
816.54
y7
1000.71
y8
1097.70
y9
1198.75
y10 y9 y8 y7 y6 y5 y2 y1
b2 b4 b5 b6
1Ame2TPKSTPAKTR11

can be observed by subtracting y10 from the MH+ value or from the shift in the b series (see the b2 ion in all three spectra).

One spectrum was obtained representing two species of peptide 12–18, which showed that K12 or K16 is acetylated (Fig. 3.16, Table 1). Ions b2-b3 and y4-y6 are present for both species, which allows us to estimate the abundance of the modified species based on the intensity of these peaks. I estimate that acetylated K12 represents approximately 30% of the modified species, and acetylated K16 is ~70%. The PTMs were designated as acetylated lysines because the Kac immonium ion was present at m/z 126. These marks are presumably minor, because they were not observed by Edman analysis.

H3 Modifications

Multiple attempts were made to acquire information on the modification state of H3. Many of the peptides produced by the PA-trypsin digest of H3 were sequenced by tandem MS (Fig. 3.7D). QqTOF analysis of the H3 PA-trypsin digest showed multiple modification states for the lysine-rich peptide 9–36 (Fig. 3.17A). Sequencing by tandem MS showed that all 8 lysines are propionylated in the species represented as M + 448 Da. Tandem MS of the three remaining peaks showed that they represent modified species of peptide 9–36, but modifications could not be assigned to specific lysines. The presence of these peaks suggests that several lysines at the H3 N-terminus can be modified.

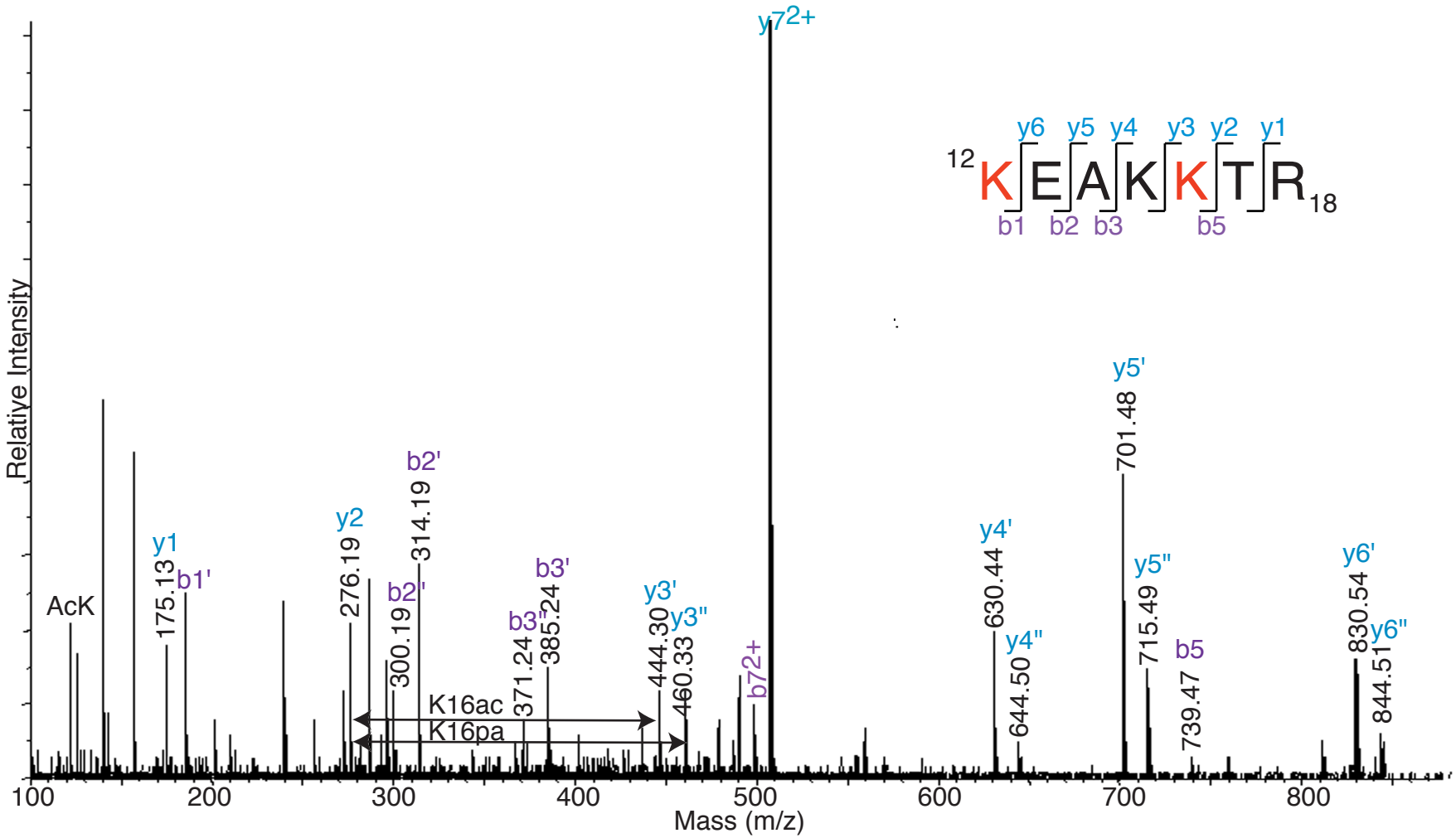


Figure 3.16. Fig. 4. Histone H2B lysines 12 and 16 are acetylated. H2B was propionylated and digested by trypsin. Peptide 12-18 (KEAKKTR) was sequenced by tandem MS. A single MS spectrum contains two modified species, demonstrating that either K12 (b2"-b3" and y3"-y6") or K16 (b1'-b3' and y3'-y6') are acetylated. Theoretical MW = 860.5 Da; measured MW = 860.5 Da.

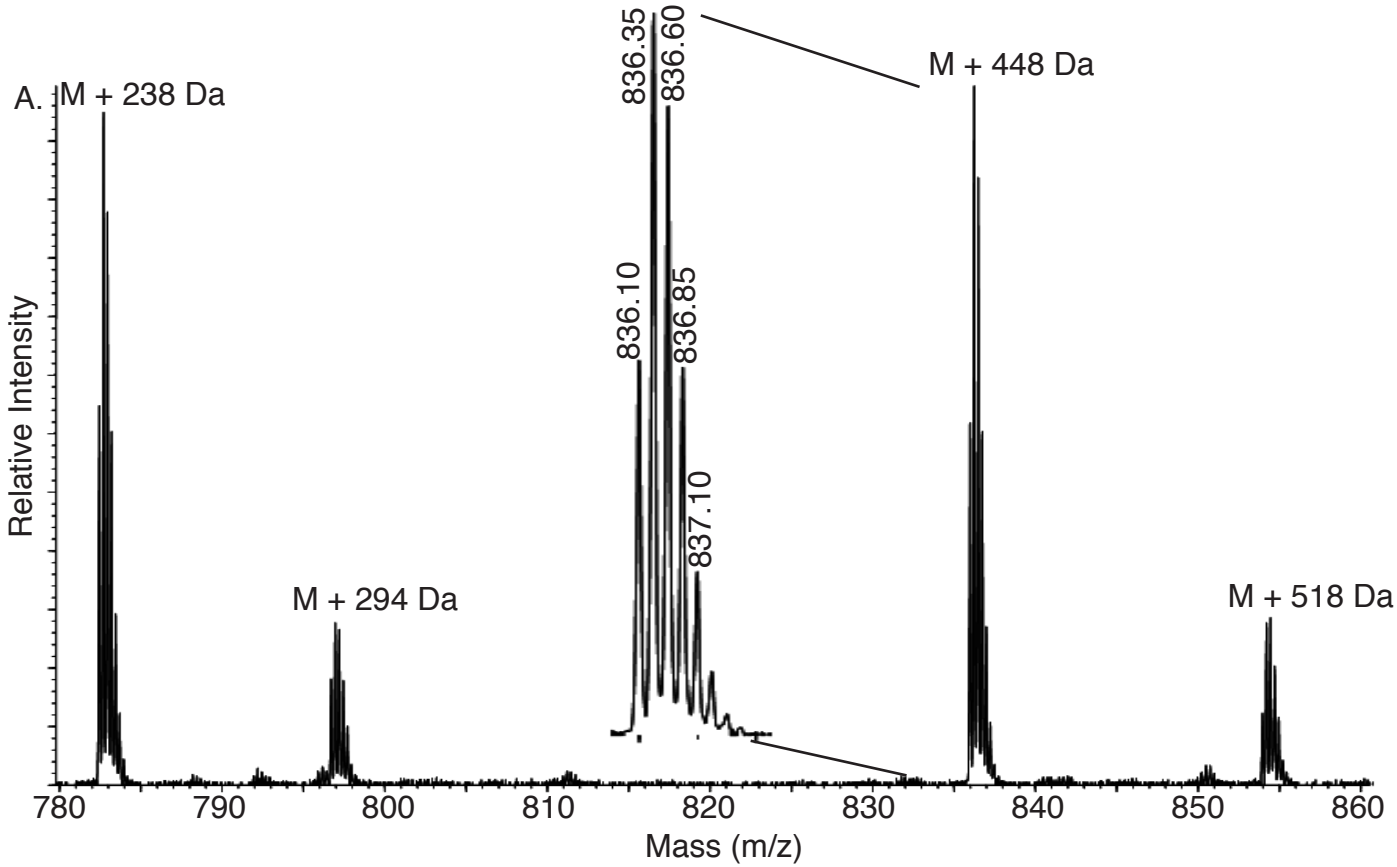
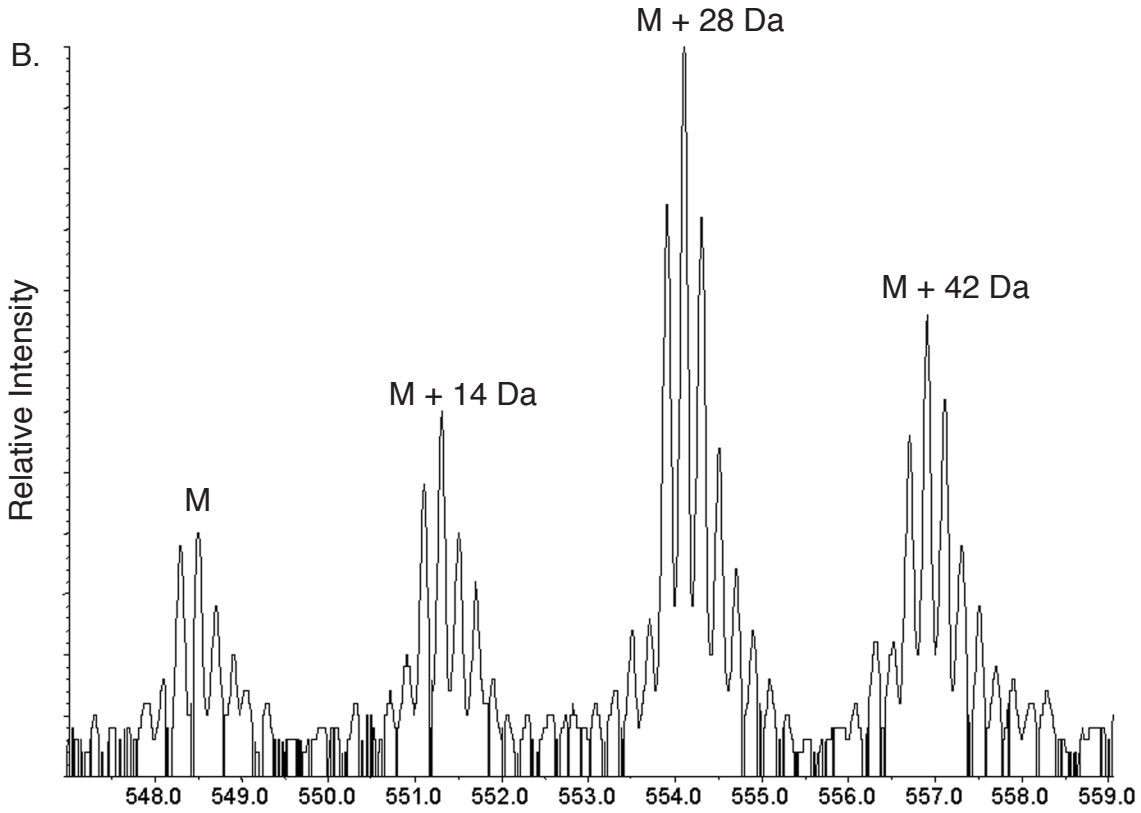


Figure 3.17. The N-terminus of histone H3 is present in multiple modification states. Quadrupole-quadrupole-TOF data were acquired for H3 digested with endoproteinases trypsin and Asp-N. M is the molecular ion mass of each peptide. (A) Propionylated H3 digested with trypsin produces peptide 9-36 (TKKTITSKKSKKASKGSDAASGVKTAQR), which was observed as [MH4]4+(m/z 723.9). Sequencing by tandem MS of the peptide labeled M + 448 Da shows that this is the species for which all 8 lysines are propionylated. Magnification of peak M + 448 Da (inset) shows the isotope pattern of this peptide, which enabled us to determine that this is a +4 species. Partial sequencing by tandem MS of the remaining three peaks confirms that they correspond to peptide 9-36, but modifications could not be assigned to specific parent residues.



3.17B. Asp-N-digested H3 produces peptide 1-25 (SRTKETARTKKTITSKKSKKASKGS), which was observed at m/z 548.3 for [MH5]5+. Sequencing by tandem MS confirms that the peak labeled M is unmodified peptide 1-25. Partial sequencing by tandem MS of the three modified species M + 14 Da, M + 28 Da, and M + 42 Da confirms that they correspond to peptide 1-25, but modifications could not be assigned to specific parent residues.

A spectrum corresponding to the N-terminus, peptide 1–8, showed that S1 can be acetylated (Fig. 3.18, Table 1). Peptide 1-8 was the result of an incomplete trypsin digest at R2, possibly due to the proximity of the cleavage site to the N-terminus. Peptide 3-8 was also sequenced by MS/MS, showing that K4 was propionylated (data not shown). However, peptide 3-8 with a methylated K4 was not observed. QqTOF of the H3 PA-trypsin digest also showed that peptide 70-80 can be present in four modification states (Fig. 3.19A). Tandem MS sequencing of these peaks showed that K76 can be unmodified, mono-, di-, or tri-methylated (Fig. 3.19B-E, Table 1). The spectra representing di- and tri-methyl K76 are incomplete, but a shift in the y series at y5 and the b series at b7 represents the mass shift from modified K76 (see y6-y9 and b9 for dimethyl K76, and y7-y9 and b7-b10 for trimethyl K76). This modification has also been characterized in PF [165].

Although the N-terminus of H3 was blocked to analysis by Edman degradation, I attempted to acquire Edman degradation data by separating the peptides from the PA-trypsin digest by HPLC and sequencing peptide 9–36. A weak profile was obtained, but no PTMs were detectable. Interpretation of the Edman degradation data was further complicated by the simultaneous sequencing of an H2A peptide.

Tandem MS sequencing of trypsin-digested H3 uncovered two modifications at the N-terminus — K23ac and K32me3 (Table 1). K23 was shown to be acetylated by the presence of the immonium ion at m/z 126 (Fig. 3.20). The

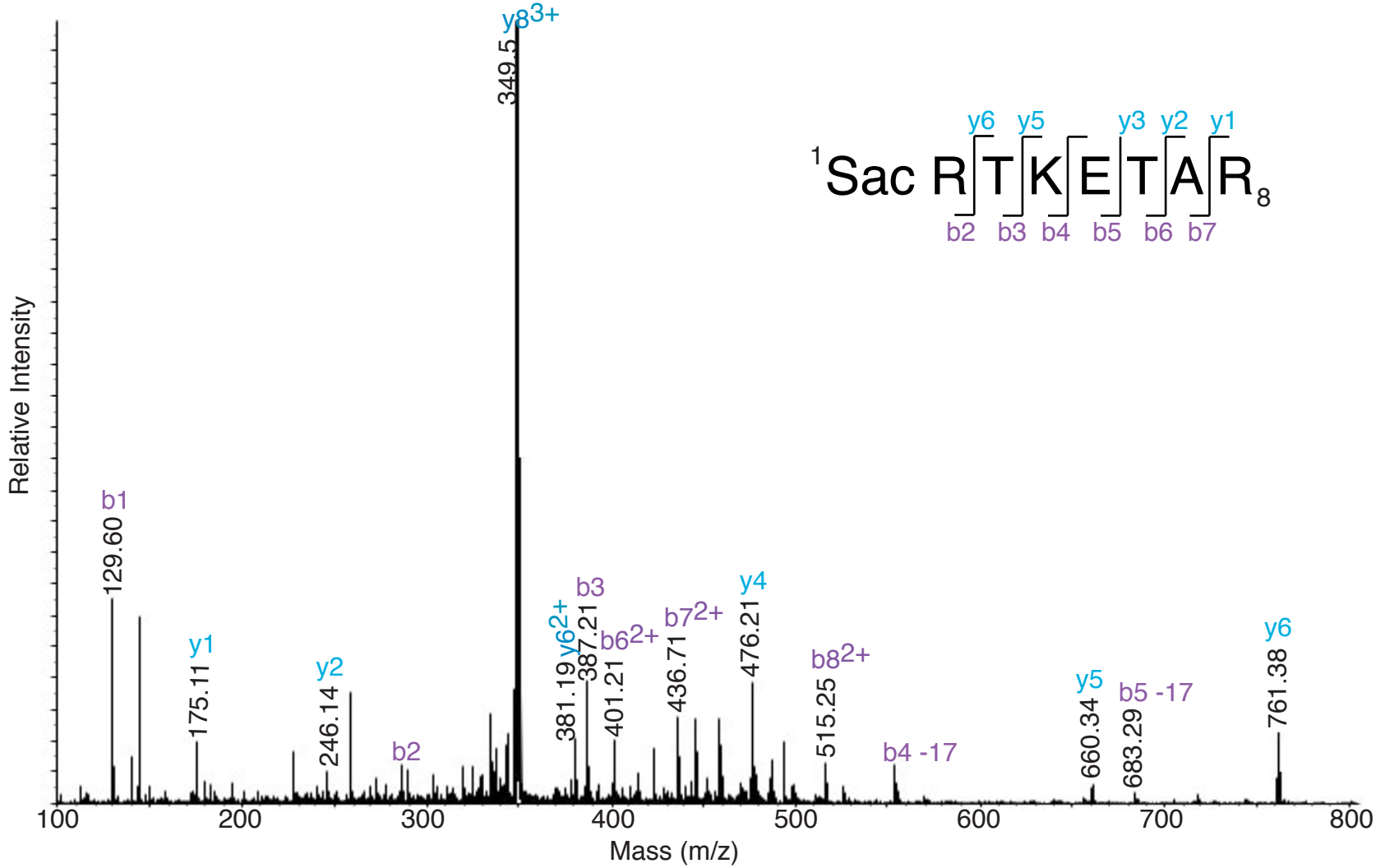


Figure 3.18. Histone H3 serine 1 is acetylated. Propionylated H3 was digested with trypsin, producing peptide 1-8 (SRTKETAR). Tandem MS sequencing demonstrates the modification - all b ions are shifted relative to the unmodified peptide. m/z 349.5^{3+}; theoretical MW = 1046.5 Da; measured MW = 1046.5 Da.

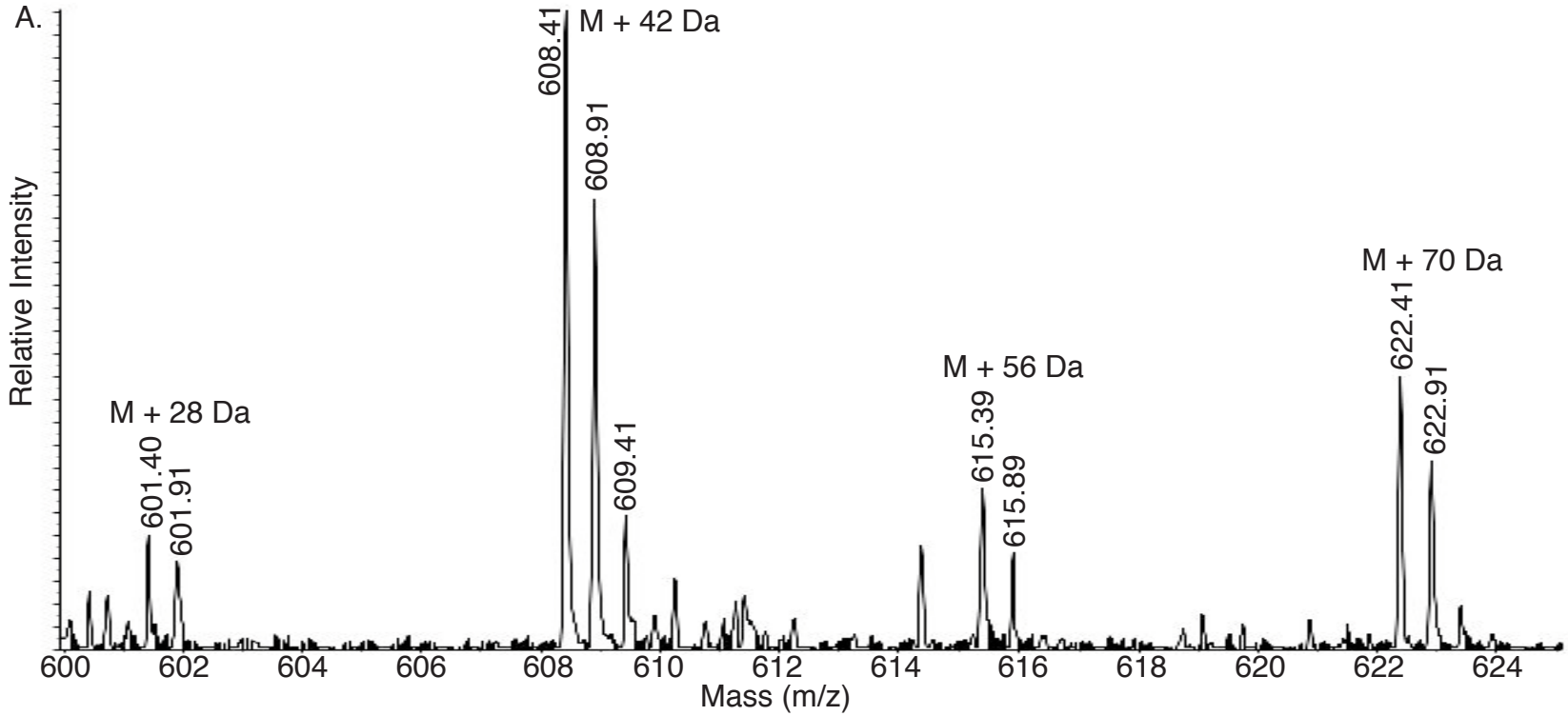
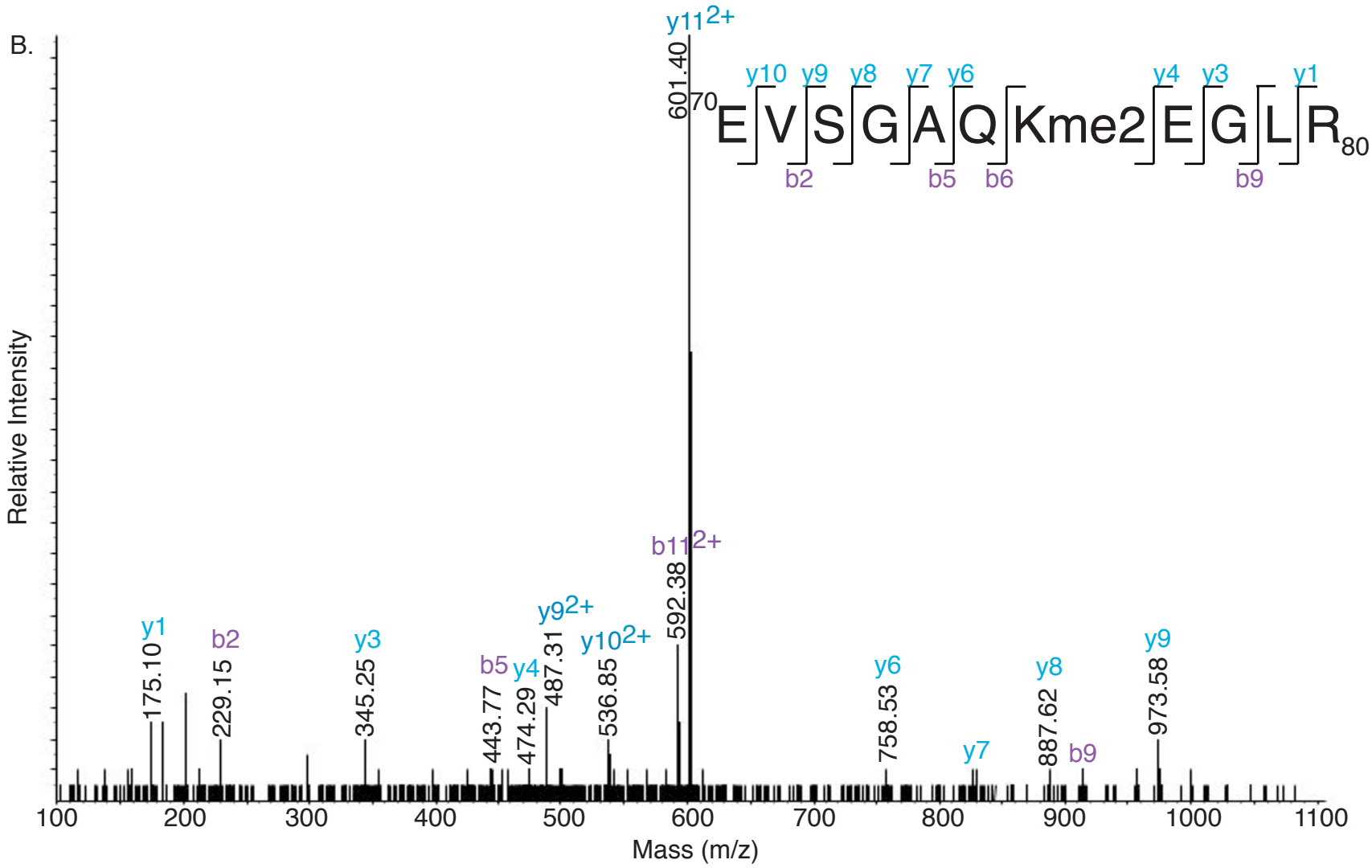
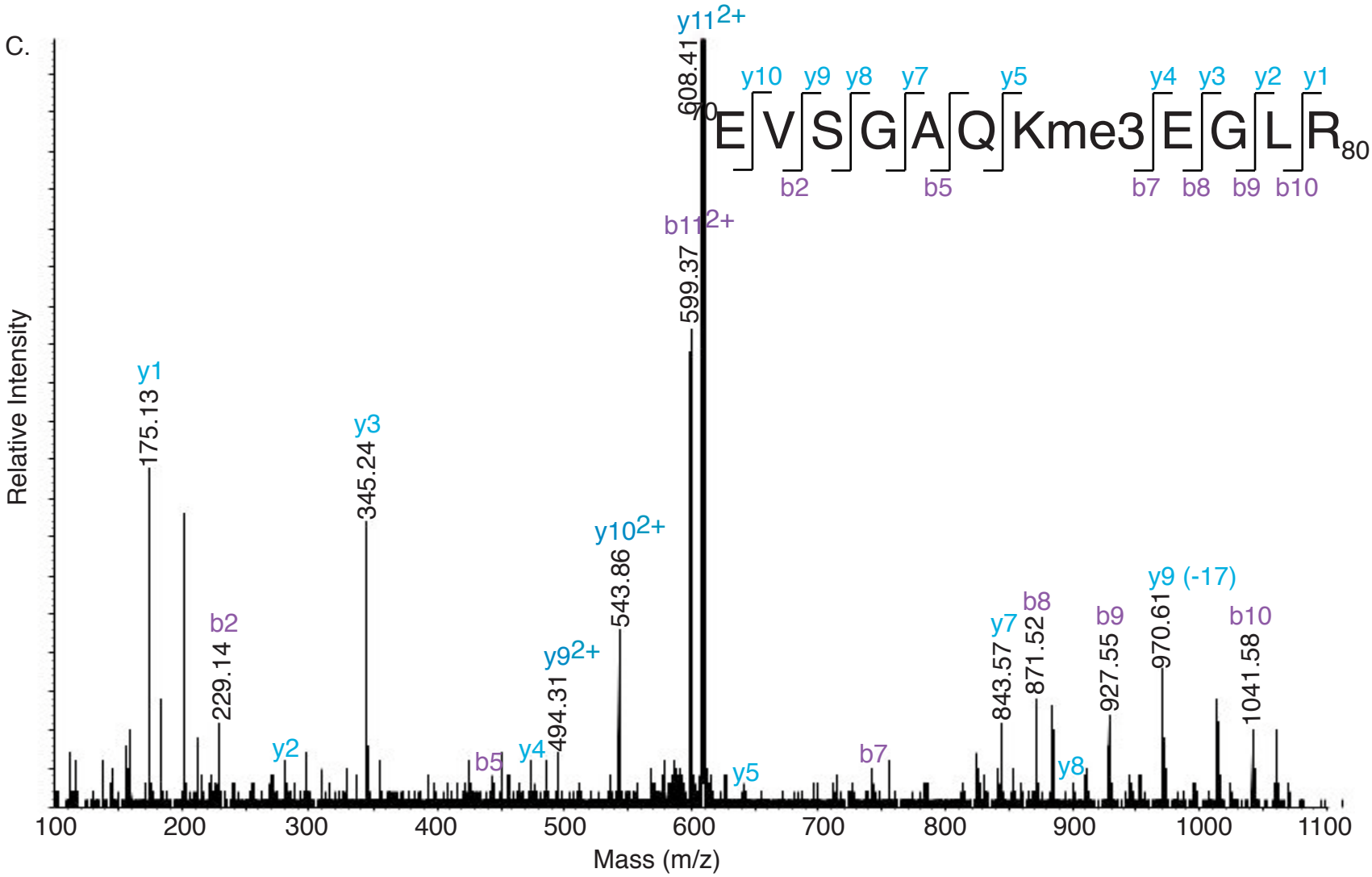
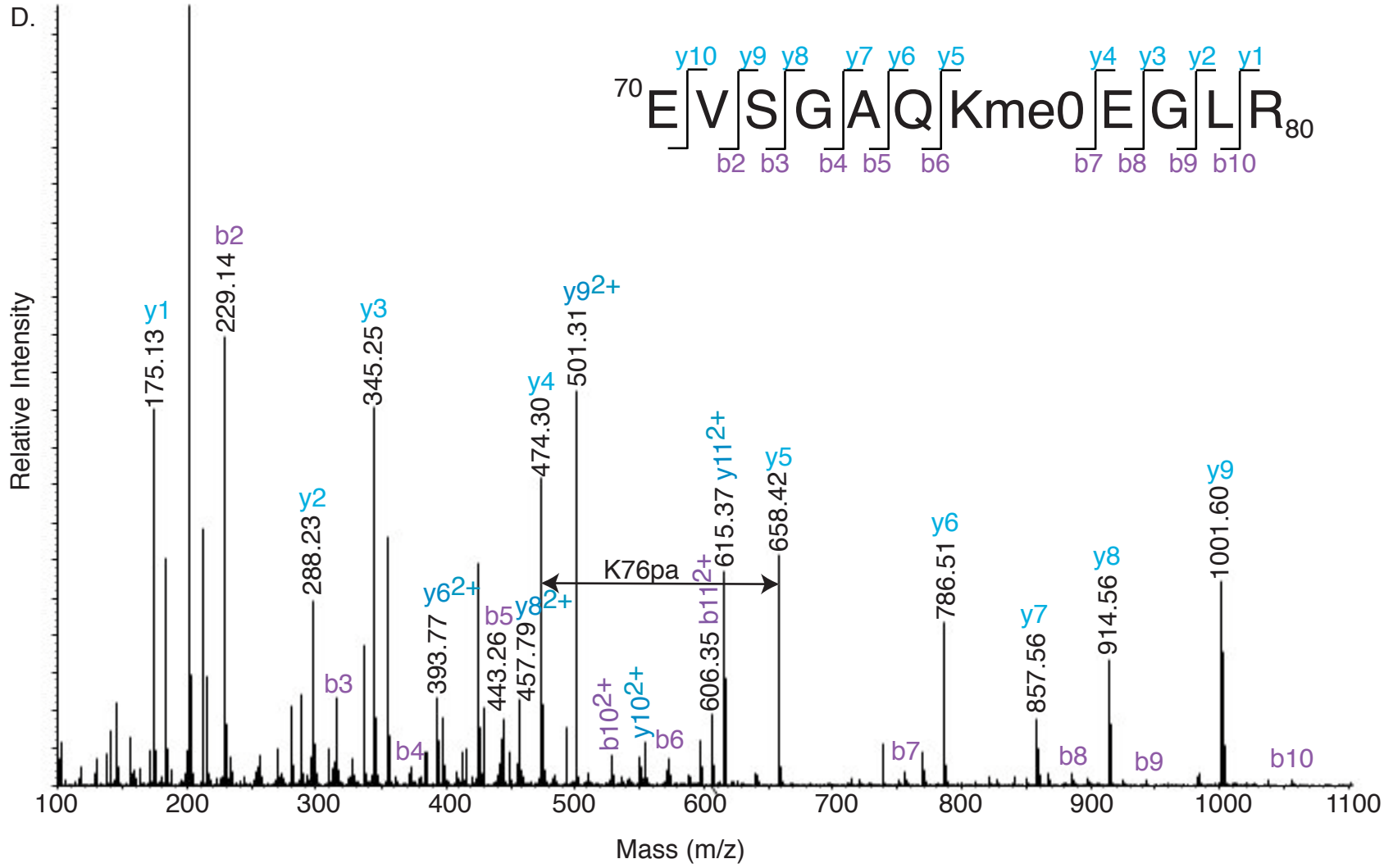
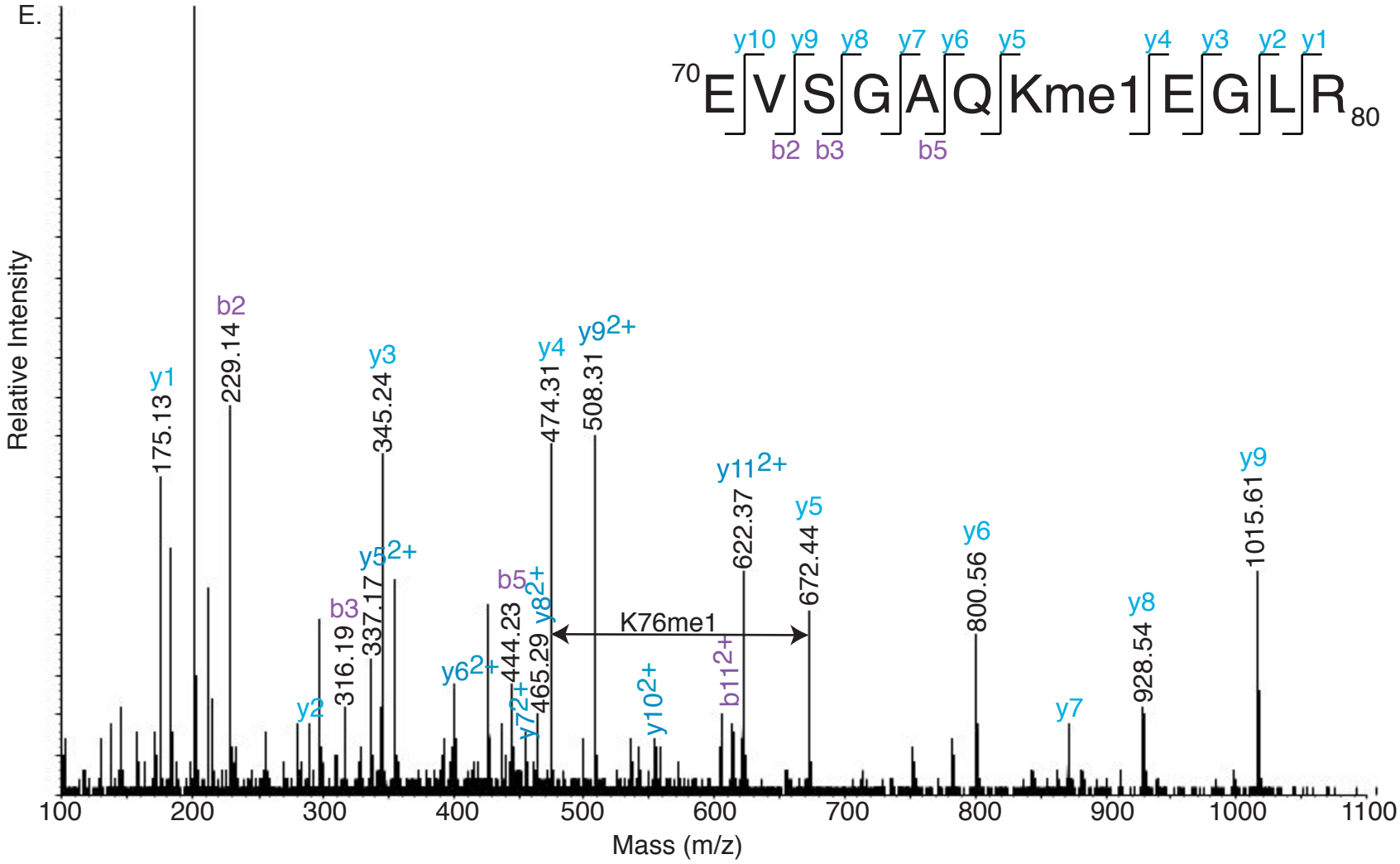


Figure 3.19. Histone H3 lysine K76 may be mono-, di-, or trimethylated. Propionylated H3 was digested with trypsin, producing peptide 70-80 (EVSGAQKEGLR). (A) Qq-TOF data shows that peptide 70-80 is present in 4 possible modification states. M is the molecular mass of the unmodified peptide. The isotope pattern demonstrates that these peaks represent the +2 species. Sequencing by tandem MS confirms that these peaks correspond to H3 peptide 70-80. (B) MS/MS data of the peak labeled M + 28 Da shows that H3 K76 is dimethylated (m/z 601.4^{2+}). Theoretical MW = 1201.6 Da; measured MW = 1201.8 Da. (C) MS/MS data of the peak labeled M + 42 Da shows that H3 K76 is trimethylated (m/z 608.4^{2+}). Theoretical MW = 1215.6 Da; measured MW = 1215.8 Da. (D) MS/MS data of the peak labeled M + 56 Da shows that H3 K76 can be unmodified (propionylated) (m/z 615.4^{2+}). Theoretical MW = 1229.6 Da; measured MW = 1229.8 Da.(E) MS/MS data of the peak labeled M + 70 Da shows that H3 K76 can be monomethylated and propionylated (m/z 622.4^{2+}). Theoretical MW = 1243.6 Da; measured MW = 1243.8 Da. These assignments can be made by looking at ion pairs b6-b7 and y4-y5, which provide the mass of the modified lysine.


B.
y11 2+
601.40
70 E V S G A Q Kme2 E G L R 80
y10 y9 y8 y7 y6 y4 y3 y1
b2 b5 b6 b9
Relative Intensity
y1
175.10
b2
229.15
y3
345.25
b5
443.77
y4
474.29
y9 2+
487.31
y10 2+
536.85
b11 2+
592.38
y6
758.53
y7
y8
887.62
b9
y9
973.58
100 200 300 400 500 600 700 800 900 1000 1100
Mass (m/z)


C.
y11 2+
608.41
y10 y9 y8 y7 y5 y4 y3 y2 y1
70 E V S G A Q Kme3 E G L R 80
b2 b5 b7 b8 b9 b10
b11 2+
599.37
Relative Intensity
y1
175.13
y3
345.24
y10 2+
543.86
b2
229.14
y9 2+
494.31
y2
b5
y4
y5
b7
y7
843.57
b8
871.52
y8
b9
927.55
y9 (-17)
970.61
b10
1041.58
100 200 300 400 500 600 700 800 900 1000 1100
Mass (m/z)


D.
70EVSGAQKme0EGLR80
y10 y9 y8 y7 y6 y5 y4 y3 y2 y1
b2 b3 b4 b5 b6 b7 b8 b9 b10
Relative Intensity
175.13 y1
229.14 b2
288.23 y2
345.25 y3
393.77 y6 2+
443.26 b5
457.79 y8 2+
474.30 y4
501.31 y9 2+
606.35 b11 2+
615.37 y11 2+
658.42 y5
786.51 y6
857.56 y7
914.56 y8
1001.60 y9
b3
b4
b10 2+
y10 2+
b6
b7
b8
b9
b10
K76pa
100 200 300 400 500 600 700 800 900 1000 1100
Mass (m/z)


E.
70EVSGAQKme1EGLR80
y10 y9 y8 y7 y6 y5 y4 y3 y2 y1
b2 b3 b5
Relative Intensity
Mass (m/z)
100 200 300 400 500 600 700 800 900 1000 1100
y1 175.13
b2 229.14
y3 345.24
y4 474.31
y9 2+ 508.31
y11 2+ 622.37
y5 672.44
y6 800.56
y9 1015.61
y8 928.54
y7
y2
b3 316.19
y5 2+ 337.17
b5
y8 2+
y6 2+ 444.23
y7 2+ 465.29
y10 2+
b11 2+
K76me1

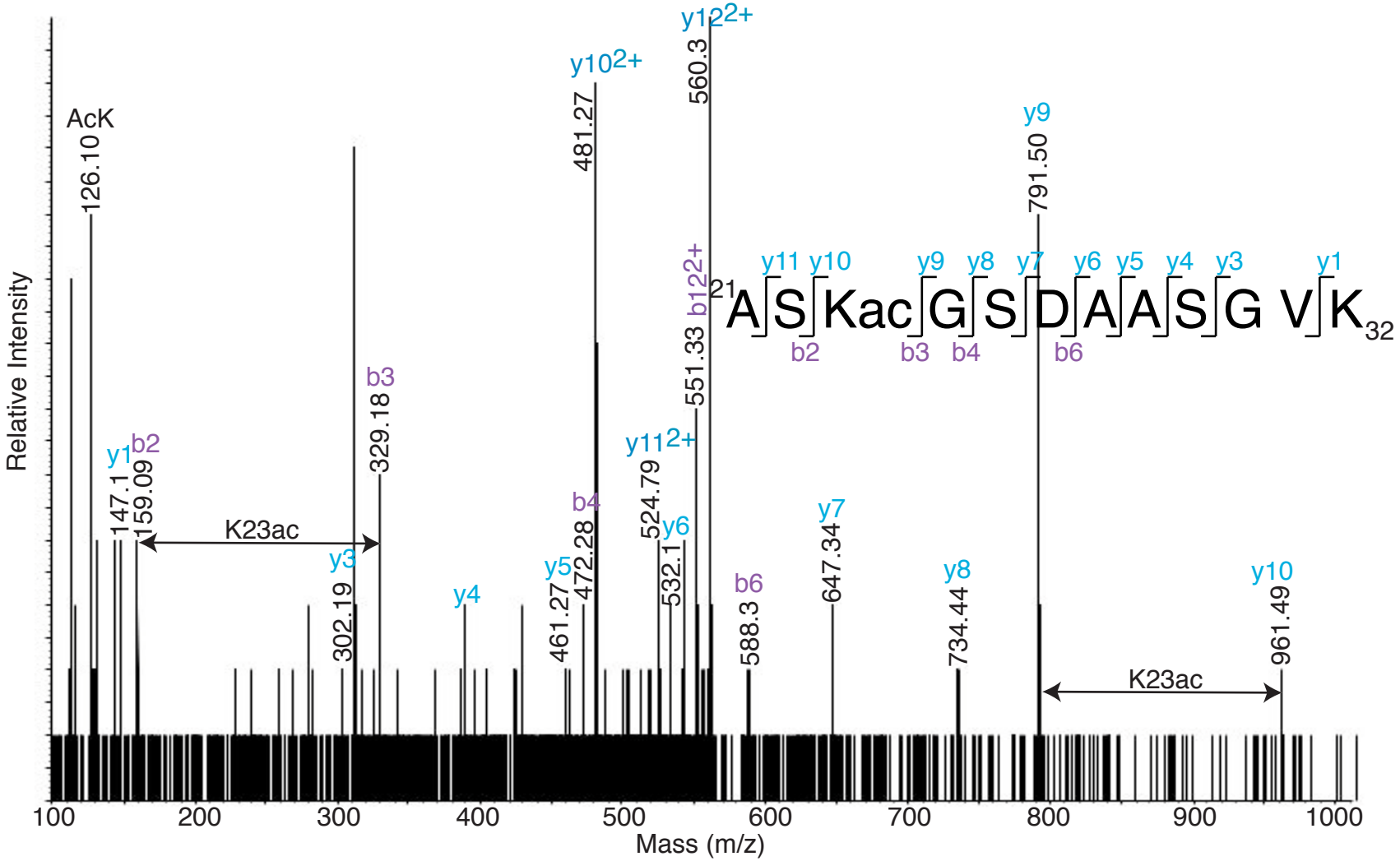


Figure 3.20. Histone H3 lysine 23 is acetylated. H3 was digested with trypsin, producing peptide 21-32 (ASKGSDAASGVK). Acetylation is demonstrated by ion pairs b2-b3 and y9-y10 and the presence of the acetyl lysine immonium ion (m/z 126). m/z 560.3^{2+}; theoretical MW = 1119.6 Da; measured MW = 1119.6 Da.

spectrum representing modified K32 is incomplete. The presence of the y4 and y7-y12 ions demonstrates that there is a mass shift of +42 Da on the GVK sequence, which I assign to K32. K32 was shown to be methylated by the presence of MH^+-59 ion (Fig. 3.21). The density of basic amino acids at the H3 N-terminus prevented other modified lysines from being discovered following trypsinization.

Finally, I attempted MS analysis of Asp-N- and Glu-C-digested H3. QqTOF analysis of Asp-N-digested H3 confirmed that multiple modification-states exist for N-terminal peptide 1–25 (Fig. 3.17B). Some MS/MS data were acquired for modified peptide 1-25, but were uninterpretable, primarily due to the number of lysines present and the observation of these peptides as +5 species. Glu-C-digested H3 was expected to produce peptides 1-5 and 6-47, but neither were visible by MS. I concluded our study of H3 with the knowledge that PTMs were present at the N-terminus that could not be assigned to specific amino acids.

H4 Modifications

Edman degradation of the first 22 amino acids showed that, unlike H2A and H2B, the H4 N-terminus has numerous PTMs (Fig. 3.22, Table 3). K2, K5, K10, and K14, are acetylated in less than 10% of H4. K4ac is an abundant mark that is found in 73% of H4. Both K17 and K18 can be mono-, di-, or tri-methylated and 60% of A1 is monomethylated.

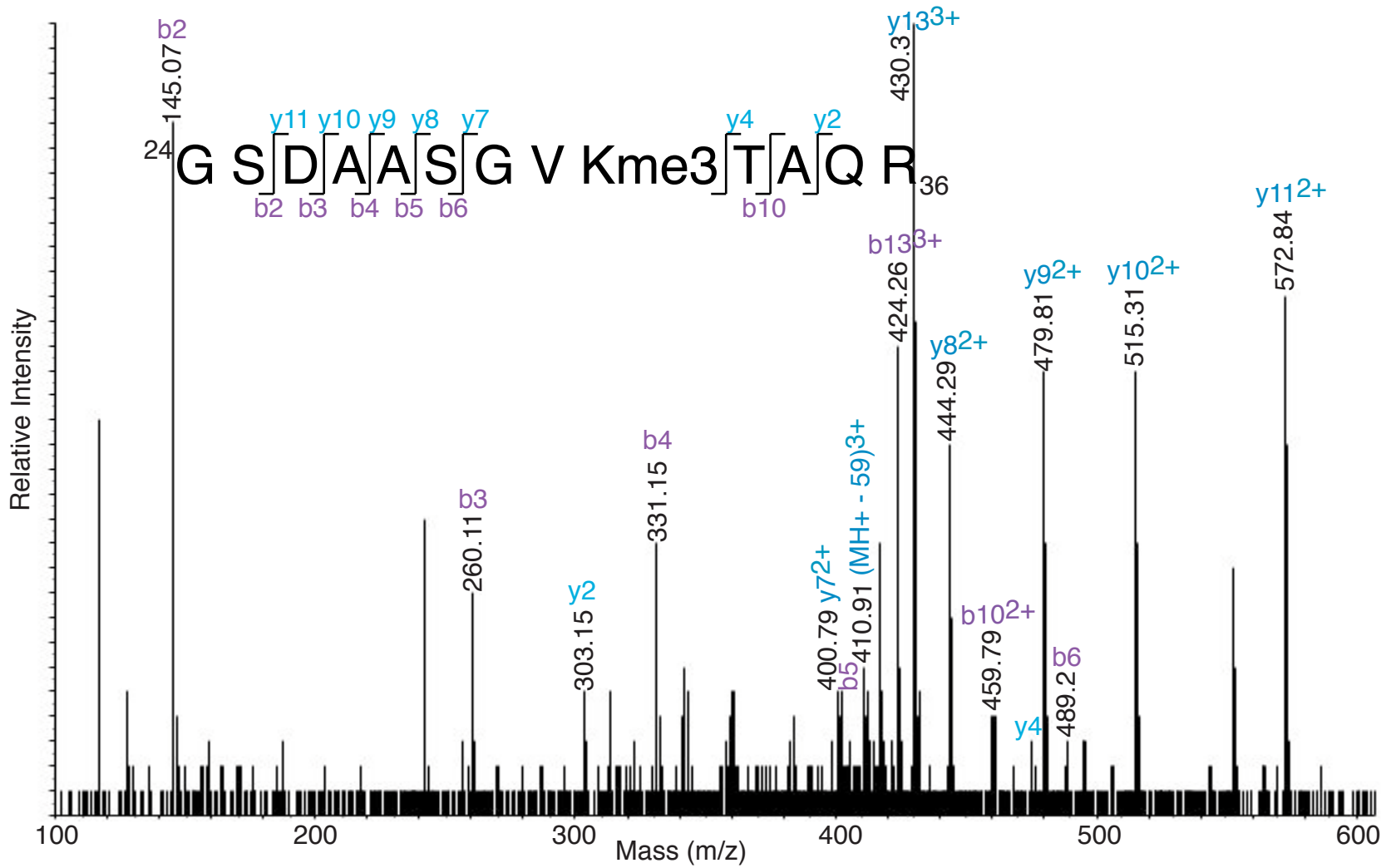


Figure 3.21. Histone H3 lysine 32 is trimethylated. H3 was digested with trypsin, producing peptide 24-36 (GSDAASGVKTAQR). Modification of K32 is demonstrated by ion pair y4-y7. The assignment of trimethylation was supported by the presence of the MH+ - 59 ion. m/z 430.3^{3+}; theoretical MW = 1289.6 Da; measured MW = 1288.9 Da.

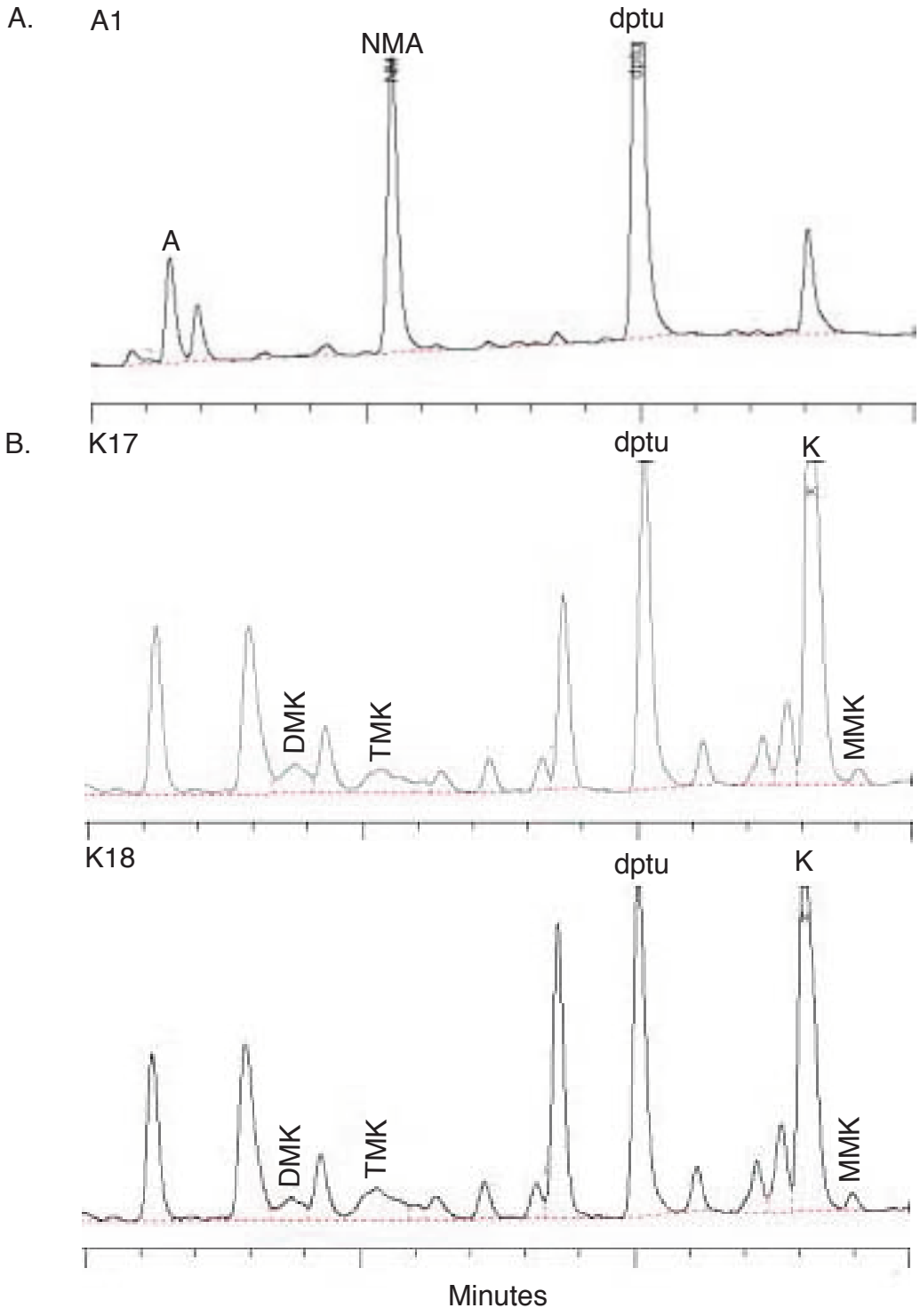
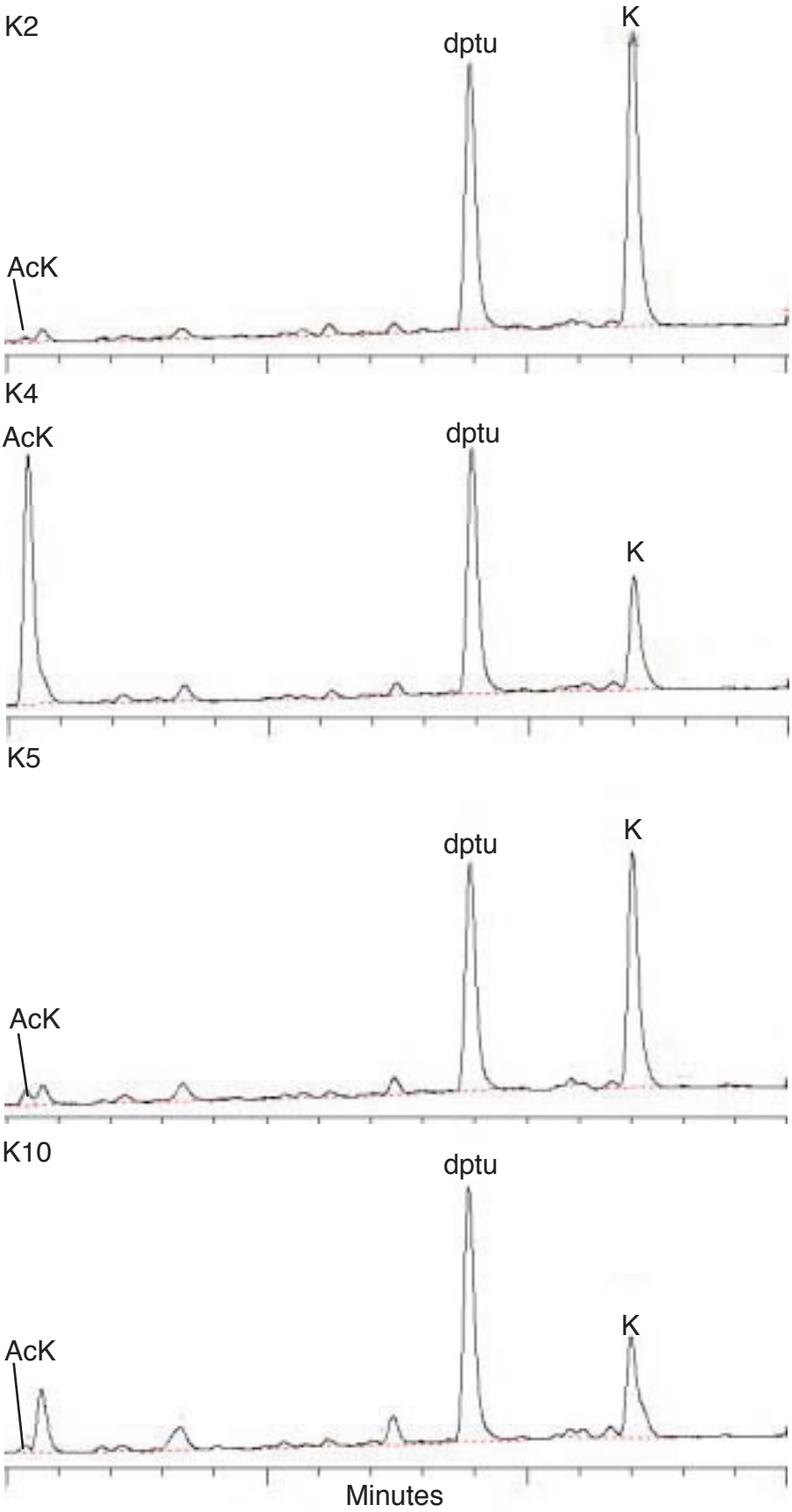


Figure 3.22. Edman degradation identifies PTMs at the H4 N-terminus. The PTMs were verified by comparison to known standards. (A) The PTH HPLC trace for cleavage cycle 1 of H4 shows that ~60% of A1 is N-methylated (NMA). (B) The trace for cycles 17 and 18 shows that K17 and K18 are mono-, di-, and tri-methylated. (C) The trace for cycles 2, 4, 5, and 10 show that 2% of K2, 73% of K4, 7% of K5, and 7% of K10 are acetylated. dptu is diphenylthiourea.

C.



I examined H4 in more detail by MS, to confirm the PTMs observed by Edman degradation as well as to determine whether less abundant PTMs were present. H4 Glu-C, trypsin-, and PA-trypsin-digested peptides were fractionated by HPLC, and individual fractions were analysed by MALDI-TOF. Fractionating the digest uncovered a number of peaks that were not visible prior to fractionation, presumably because multiple modifications of a single peptide divided the signal or stronger peptides in the mix were suppressing it. Peptides were then sequenced by tandem MS.

Sequencing by tandem MS of trypsin-digested H4 provided minimal information due to excessive fragmentation of the N-terminus. I discovered one modified species at the H4 N-terminus, demonstrating that K5 and K10 are acetylated (Fig. 3.23, Table 2).

Tandem MS of PA-trypsin-digested H4 provided extensive sequence coverage (Fig. 3.7E). The data did not suggest the presence of additional PTMs in the core domain or C-terminus of H4. MALDI-TOF demonstrated multiple modification states for peptides 1–15 and 16–21 (Fig. 3.24A and 25A, Table 2). Peptide 1-15 had the following sequence: AKGKKSGEAKGSQKR, which contains 5 lysines and an N-terminal alanine, all of which can be modified. The number of modifiable amino acids made it difficult to interpret these spectra, so I relied on Edman data to make assignments. For example, modifications were first assigned to A1 and K4 because these amino acids are found predominantly

TABLE 2. Posttranslational modifications are present at the N-terminus of Histone H4	
Digest	Modified Peptide[a]
Glu-C	$^{1}\mathbf{A}_{me1}$KG$^{4}\mathbf{K}_{ac}$KSGE (1-8)
	A$^{10}\mathbf{K}_{ac}$GSQKRQKKVLRE (9-22)
	AKGSQKRQ$^{17}\mathbf{K}_{me3}$KVLRE (9-22)
	A$^{10}\mathbf{K}_{ac}$GSQKRQK$^{18}\mathbf{K}_{me1}$VLRE (9-22)
	AKGSQKRQ$^{17}\mathbf{K}_{me3}$$^{18}\mathbf{K}_{me2}$VLRE (9-22)
	AKGSQKRQ$^{17}\mathbf{K}_{me3}$$^{18}\mathbf{K}_{me3}$VLRE (9-22)
	A$^{10}\mathbf{K}_{ac}$GSQKRQ$^{17}\mathbf{K}_{me3}$$^{18}\mathbf{K}_{me3}$VLRE (9-22)
Trypsin	$^{5}\mathbf{K}_{ac}$SGEA$^{10}\mathbf{K}_{ac}$GSQK (5-14)
PA-Trypsin	AKG$^{4}\mathbf{K}_{ac}$KSGEAKGSQKR (1-15)
	$^{1}\mathbf{A}_{me1}$KGKKSGEAKGSQKR (1-15)
	$^{1}\mathbf{A}_{me1}$KG$^{4}\mathbf{K}_{ac}$KSGEAKGSQKR (1-15)
	$^{1}\mathbf{A}_{me1}$$^{2}\mathbf{K}_{ac}G^{4}\mathbf{K}_{ac}$KSGEAKGSQKR (1-15)
	$^{1}\mathbf{A}_{me1}$$^{2}\mathbf{K}_{ac}G^{4}\mathbf{K}_{ac}$KSGEA$^{10}\mathbf{K}_{ac}$GSQKR (1-15)
	$^{1}\mathbf{A}_{me1}$$^{2}\mathbf{K}_{me2}G^{4}\mathbf{K}_{ac}$KSGEAKGSQKR (1-15)
	Q$^{17}\mathbf{K}_{me1}$KVLR (16-21)
	QK$^{18}\mathbf{K}_{me1}$VLR (16-21)
	QK$^{18}\mathbf{K}_{me2}$VLR (16-21)
	Q$^{17}\mathbf{K}_{me3}$KVLR (16-21)
	QK$^{18}\mathbf{K}_{me3}$VLR (16-21)
	Q$^{17}\mathbf{K}_{me1}$$^{18}\mathbf{K}_{me1}$VLR (16-21)

[a] N-terminal peptides resulting from each digest were sequenced by tandem MS. Peptides that have covalent modifications are listed here.

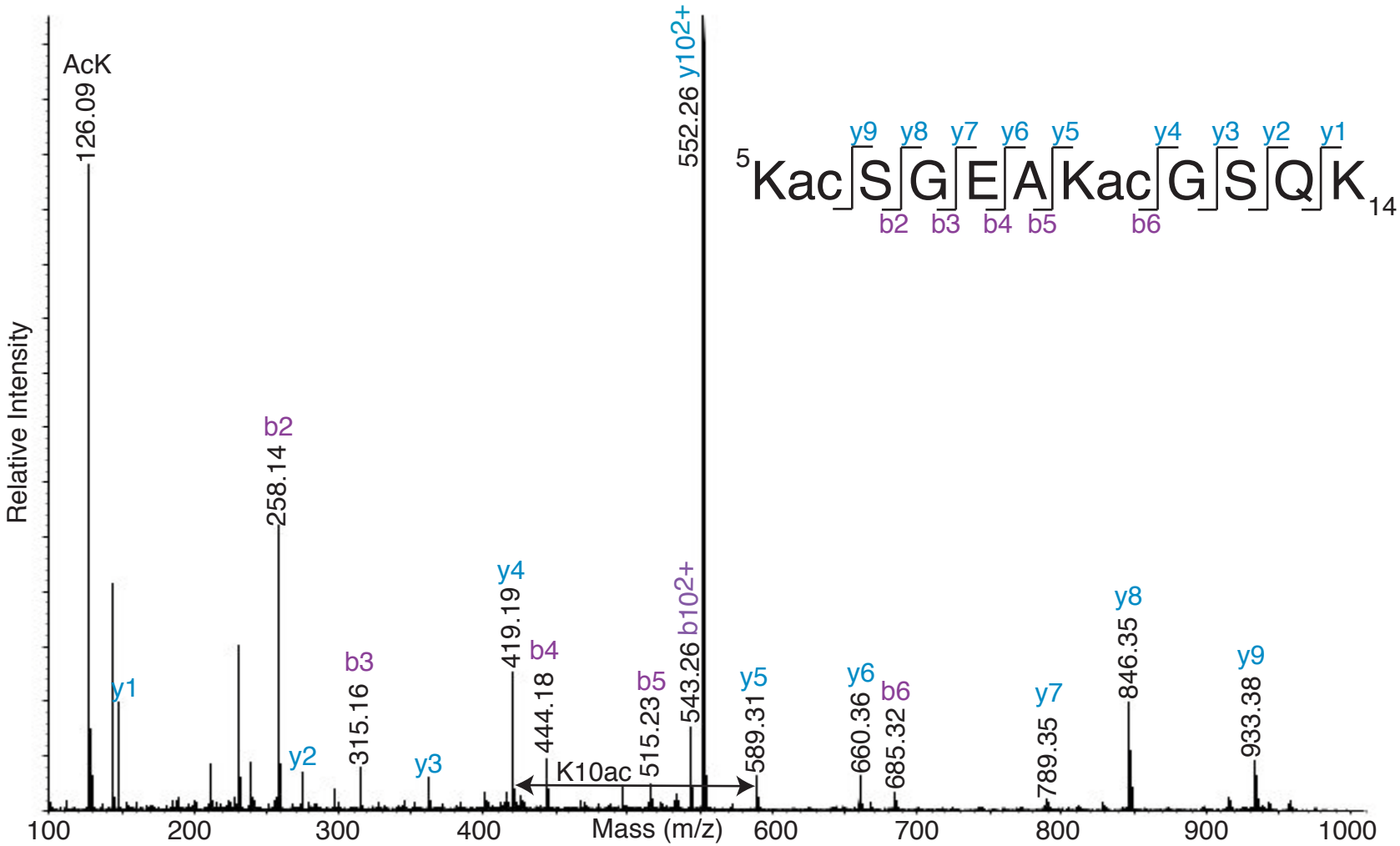


Figure 3.23. Histone H4 lysines 5 and 10 are acetylated. H4 tryptic peptide 5-14 (KSGEAKGSQK) was sequenced by tandem MS. Ion pairs y4-y5 and b5-b6 show that K10 is acetylated. Ion pair y9-y10 shows that K5 is acetylated. The presence of the acetyl immonium ion (m/z 126.09) supports the assigment of acetylation. m/z 552.3^{2+}; theoretical MW = 1103.5 Da; measured MW = 1103.6 Da.

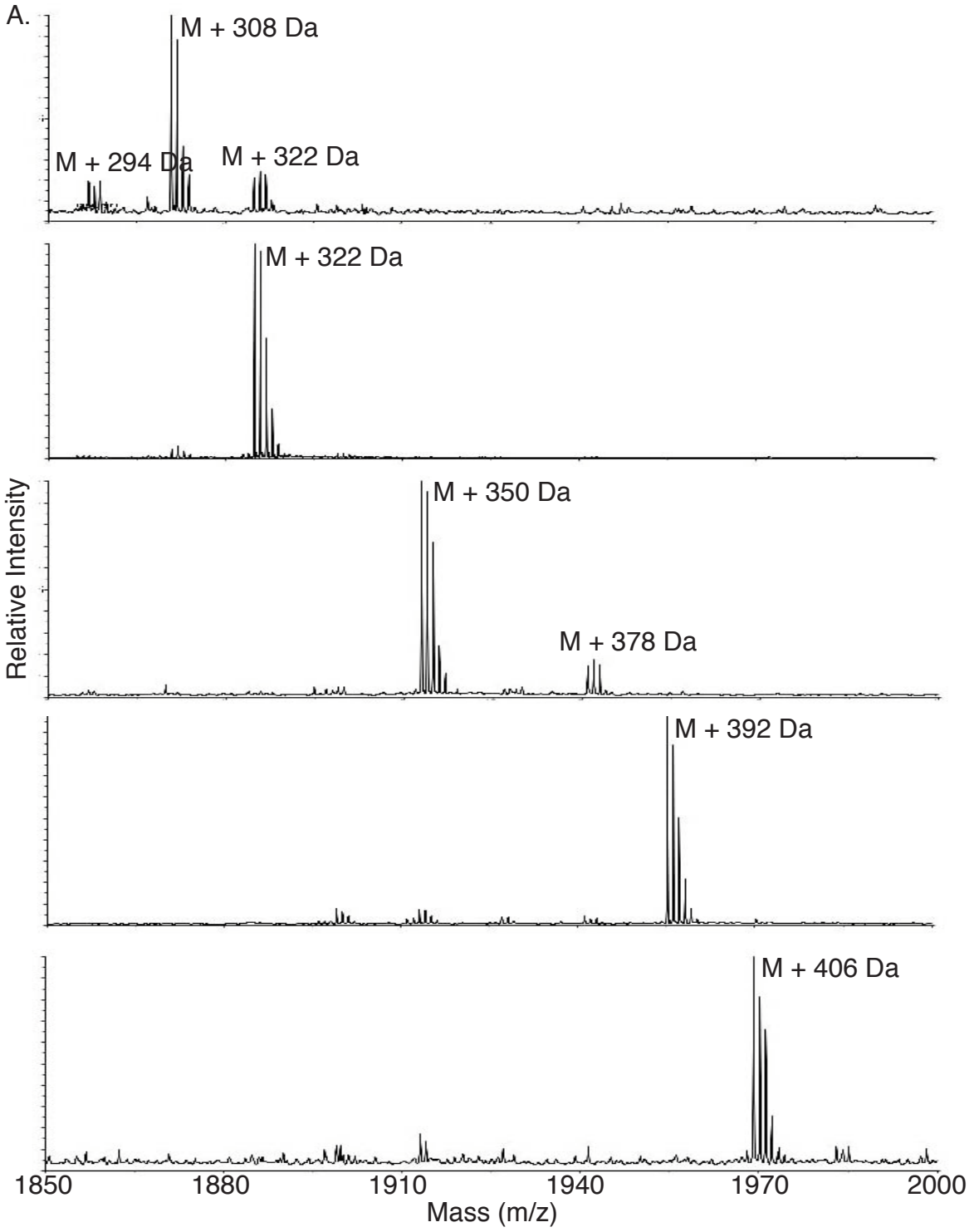


Figure 3.24. Histone H4 peptide 1-15 is present in multiple modification states. Propionylated H4 was digested with trypsin and fractionated by HPLC. (A) MALDI-TOF data were acquired for 5 HPLC fractions. M is the molecular ion mass of peptide 1-15 (AKGKKSGEAKGSQKR, MW 1559 Da). Sequencing by tandem MS of 5 of these peaks (B-F) confirm that they represent peptide 1-15. The remaining peaks were incompletely sequenced.

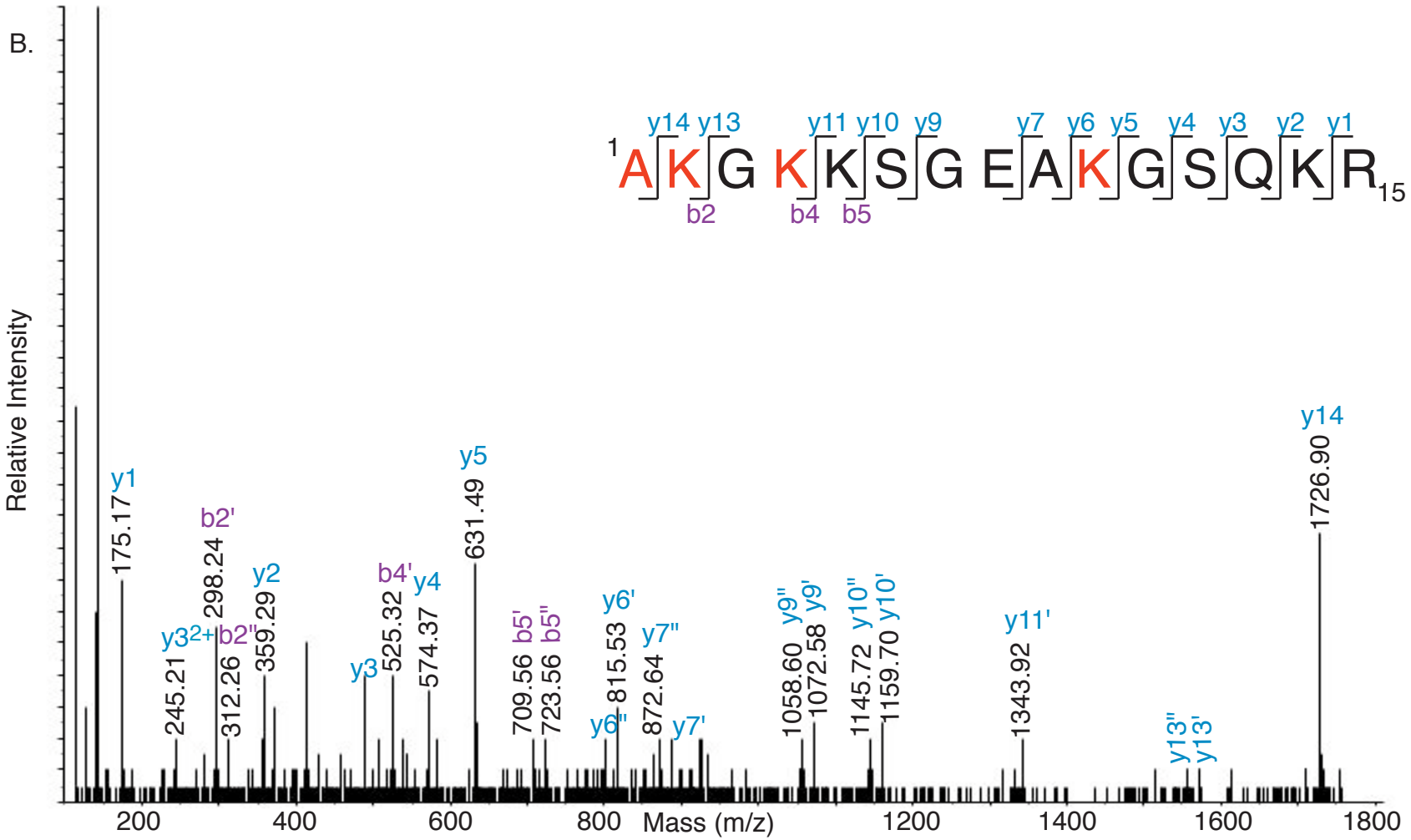


Figure 3.24B. Histone H4 alanine 1 and lysines 2, 4, and 10 are modified. This spectrum of the peak labeled M + 308 Da represents two modified species of peptide 1-15: A1me1K2acGK4acKSGEAK10acGSQKR (y6", y7", y9", y10", y13", b2", and b5") and A1me1K2me2GK4acKSGEAKGSQKR (y6', y7', y9', y10', y11', y13', b2', b4', b5'). The b2 ion and y13-14 ion pair demonstrate that lysine 2 can be either acetylated or dimethylated. Modified residues are shown in red. m/z 934.7^{2+}; theoretical MW = 1867.9 Da; measured MW = 1868.4 Da.

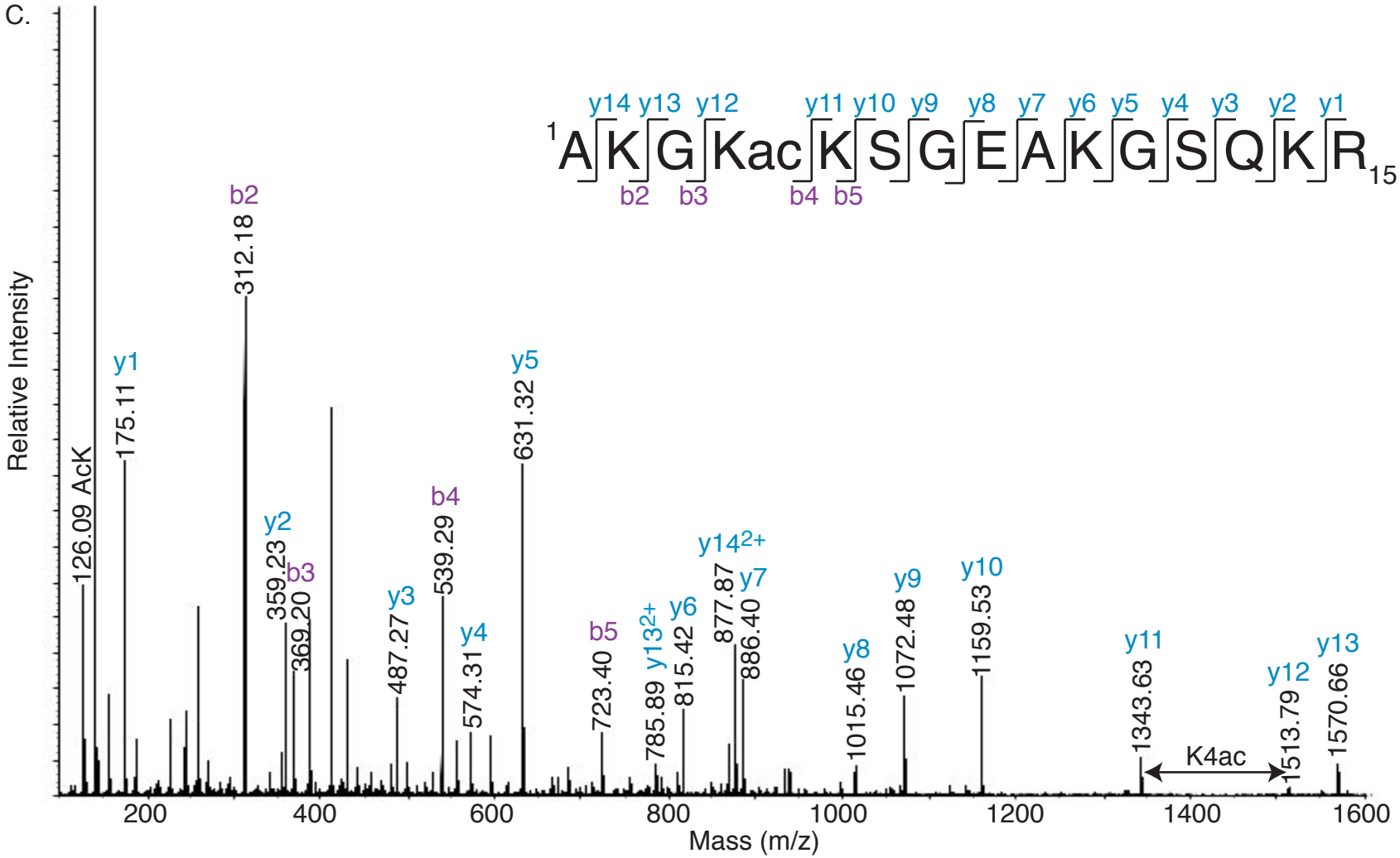


Figure 3.24C. Histone H4 lysine 4 is acetylated. Tandem MS sequencing of the peak labeled M + 322 Da shows that K4 may be acetylated. The ion pair y11-y12 demonstrates this modification. m/z 941.5^{2+}; theoretical MW = 1881.9 Da; measured MW = 1882.0 Da.

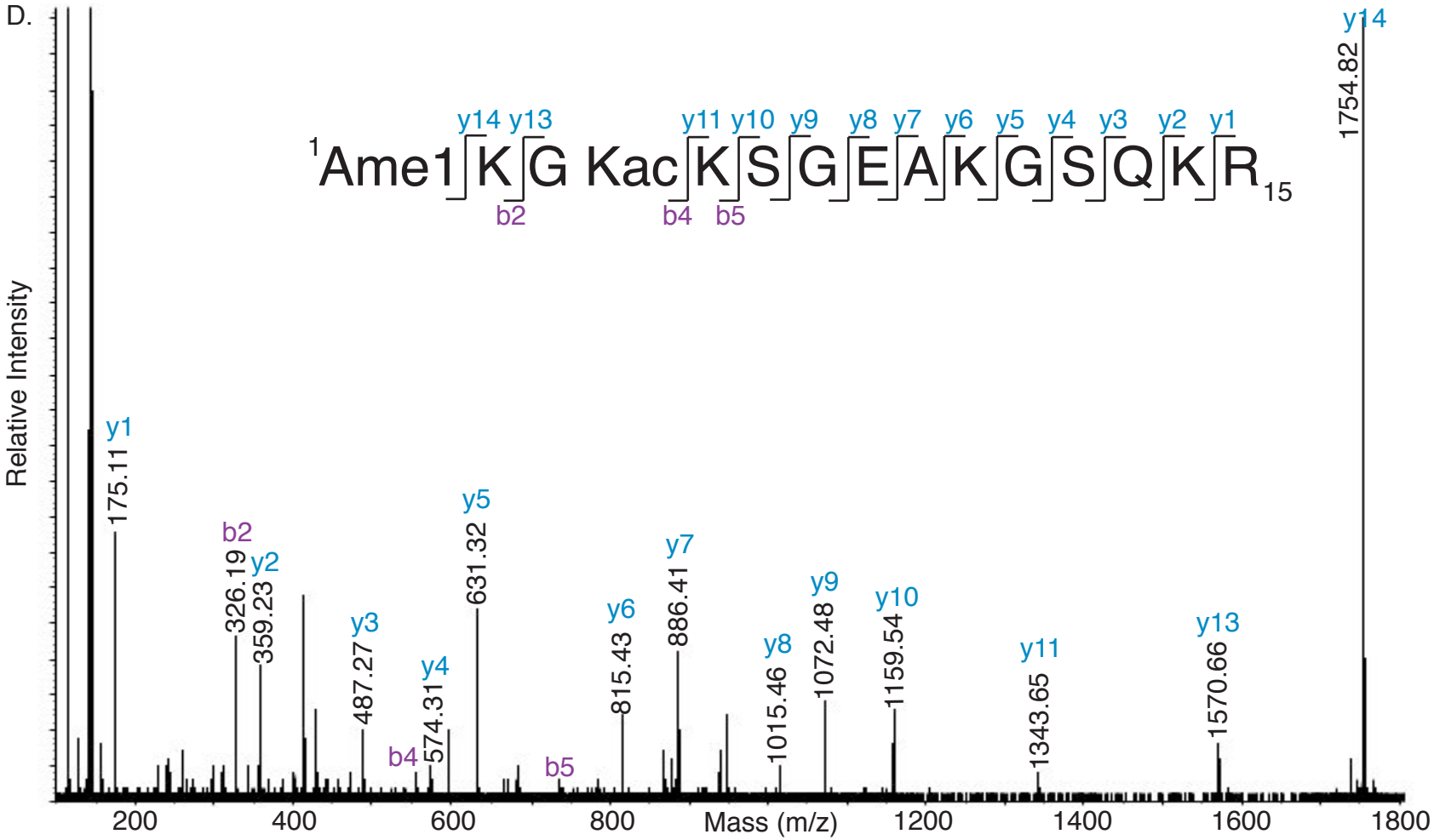


Figure 3.24D. Histone H4 alanine 1 is monomethylated and lysine 4 is acetylated. The peak labeled M + 336 Da represents a modified species of peptide 1-15 in which both A1 and K4 are modified. Compared to (C), the b ions are shifted +14 in this spectrum due to A1 monomethylation (see b2). m/z 948.5^{2+}; theoretical MW = 1895.9 Da; measured MW = 1896.0 Da.

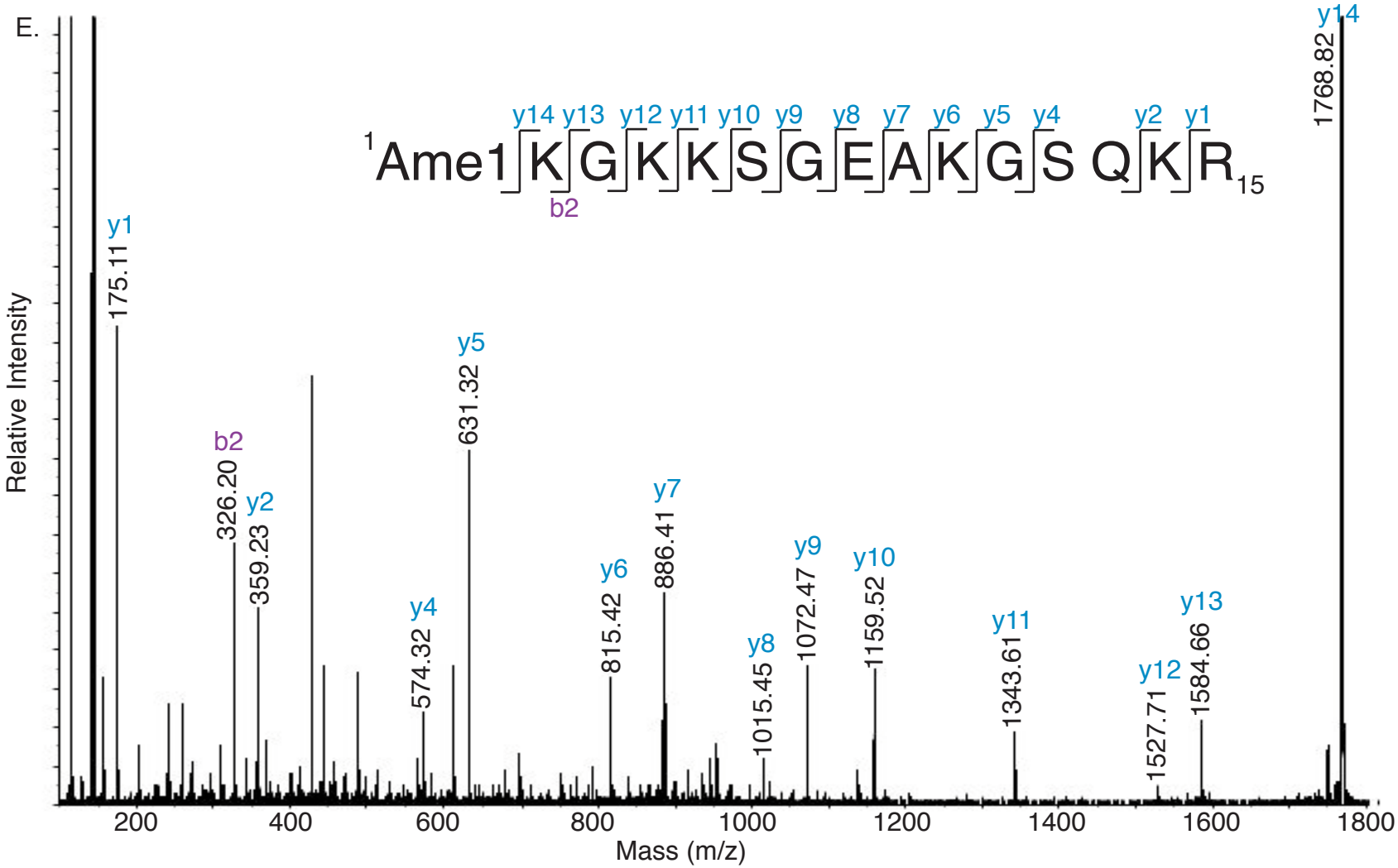


Figure 3.24E. Histone H4 alanine 1 is monomethylated. The peak labeled M + 350 Da represents a modified species of peptide 1-15 in which alanine 1 is monomethylated. Compared to (C) and (D), y12-y14 are shifted +14 because K4 is propionylated (+56 Da) instead of acetylated (+42 Da). m/z 955.5^{2+}; theoretical MW = 1909.9 Da; measured MW = 1910.0 Da.

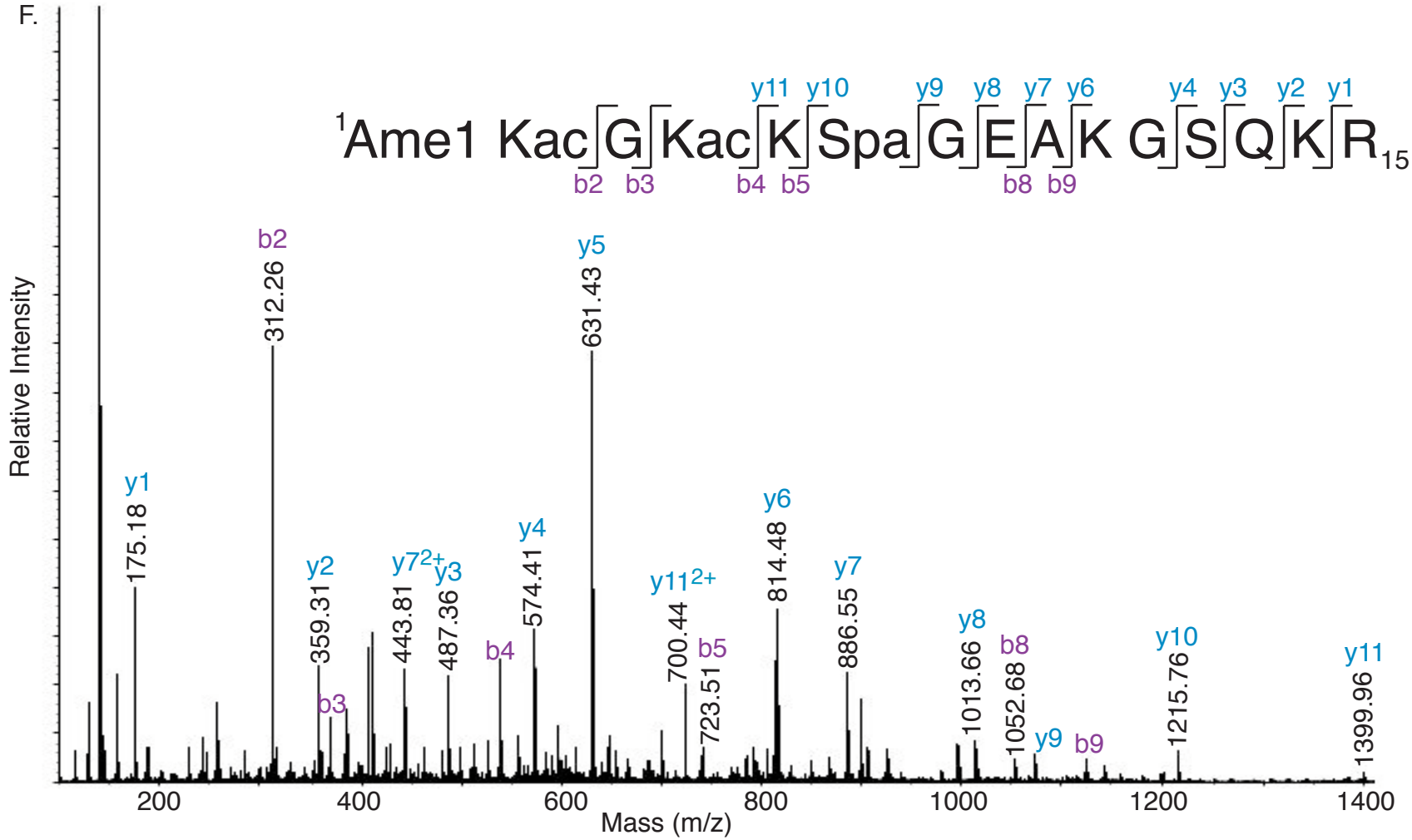


Figure 3.24F. Histone H4 alanine 1, lysines 2 and 4, and serine 6 are modified. Tandem MS sequencing of the peak labeled M + 378 shows that propionic anhydride may also modify serines and threonines. Propionylation of S6 is demonstrated by ion pair y9-y10. m/z 969.5^{2+}; theoretical MW = 1937.9 Da; measured MW = 1938.0 Da.

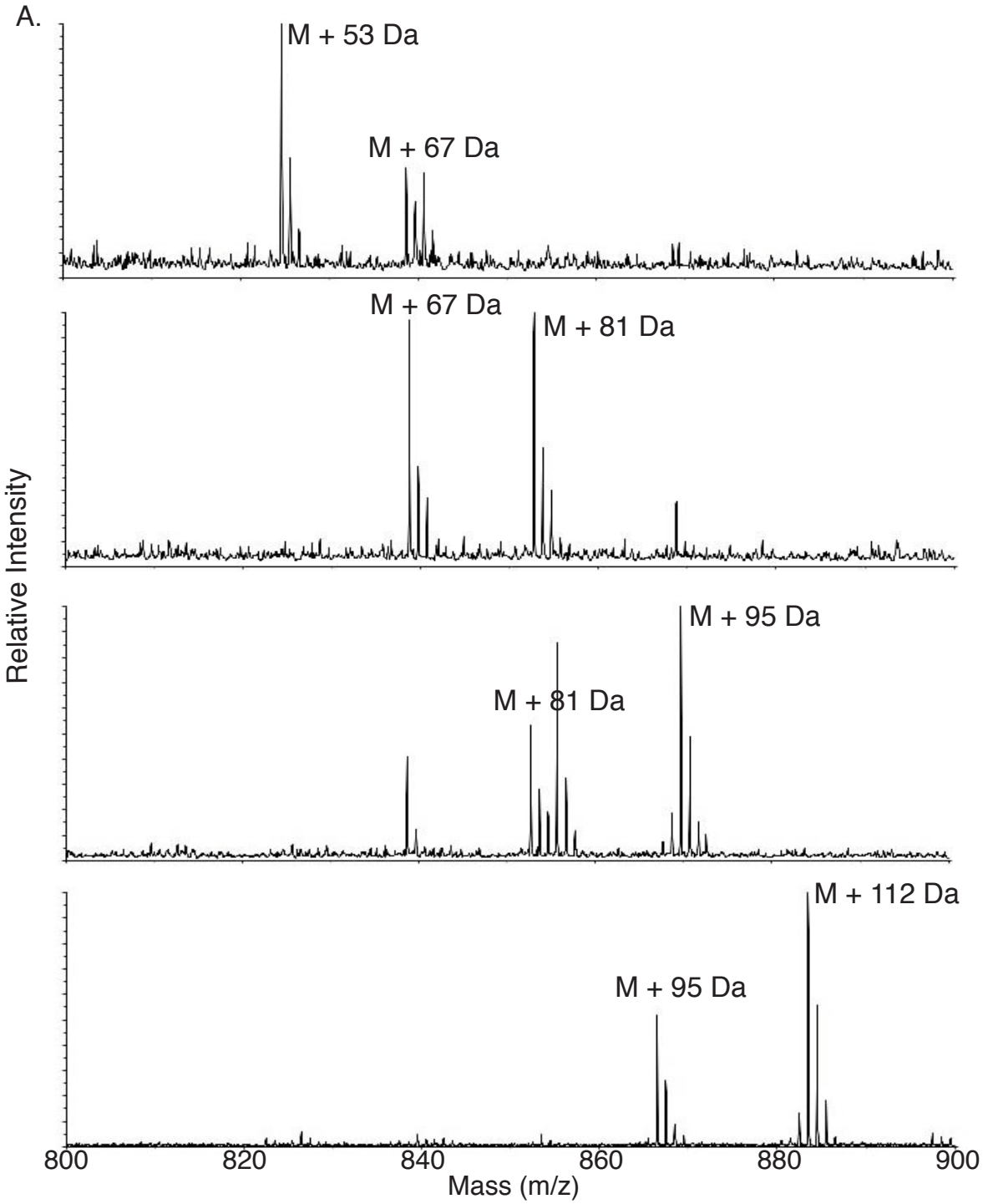


Figure 3.25. Histone H4 peptide 16-21 is present in multiple modification states. Propionylated H4 digested with trypsin was fractionated by HPLC.
(A) MALDI-TOF data were acquired for 4 HPLC fractions, demonstrating that peptide 16-21 is present in several modification states. Sequencing by tandem MS (B-F) confirmed that these peaks correspond to peptide 16-21 (QKKVLR), although, in some cases, glutamine rearranged to form pyroglutamic acid, resulting in the loss of ammonia (-17 Da). M is the molecular mass of unmodified peptide 16-21 (771.51 Da).

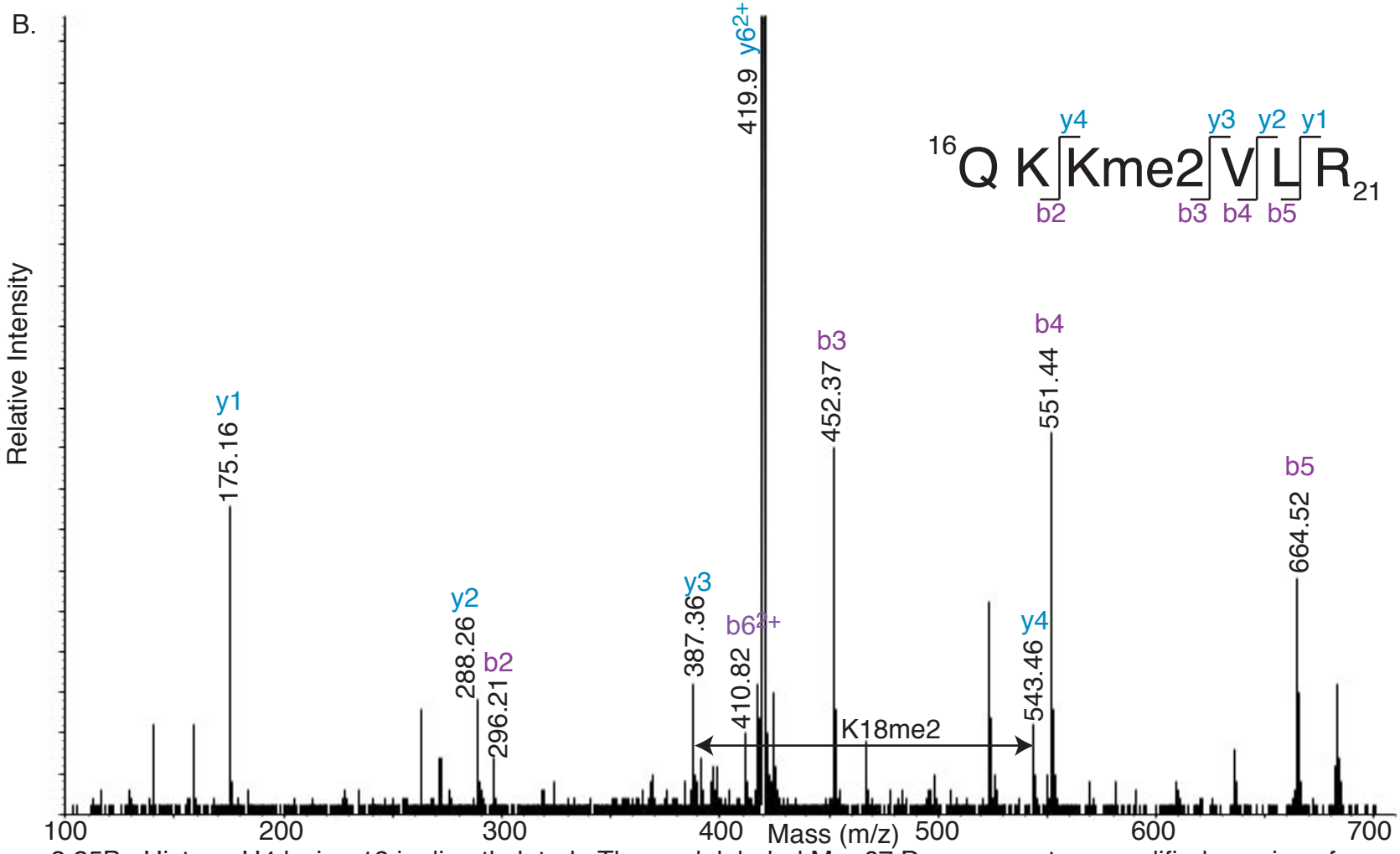


Figure 3.25B. Histone H4 lysine 18 is dimethylated. The peak labeled M + 67 Da represents a modified species of peptide 16-21 in which K17 is propionylated (+56), K18 is dimethylated (+28), and glutamine has rearranged to form pyroglutamic acid (-17). The ion pairs y3-y4 and b2-b3 demonstrate the modification on K18. m/z 419.9^{2+}; theoretical MW = 838.5 Da; measured MW = 838.8 Da.

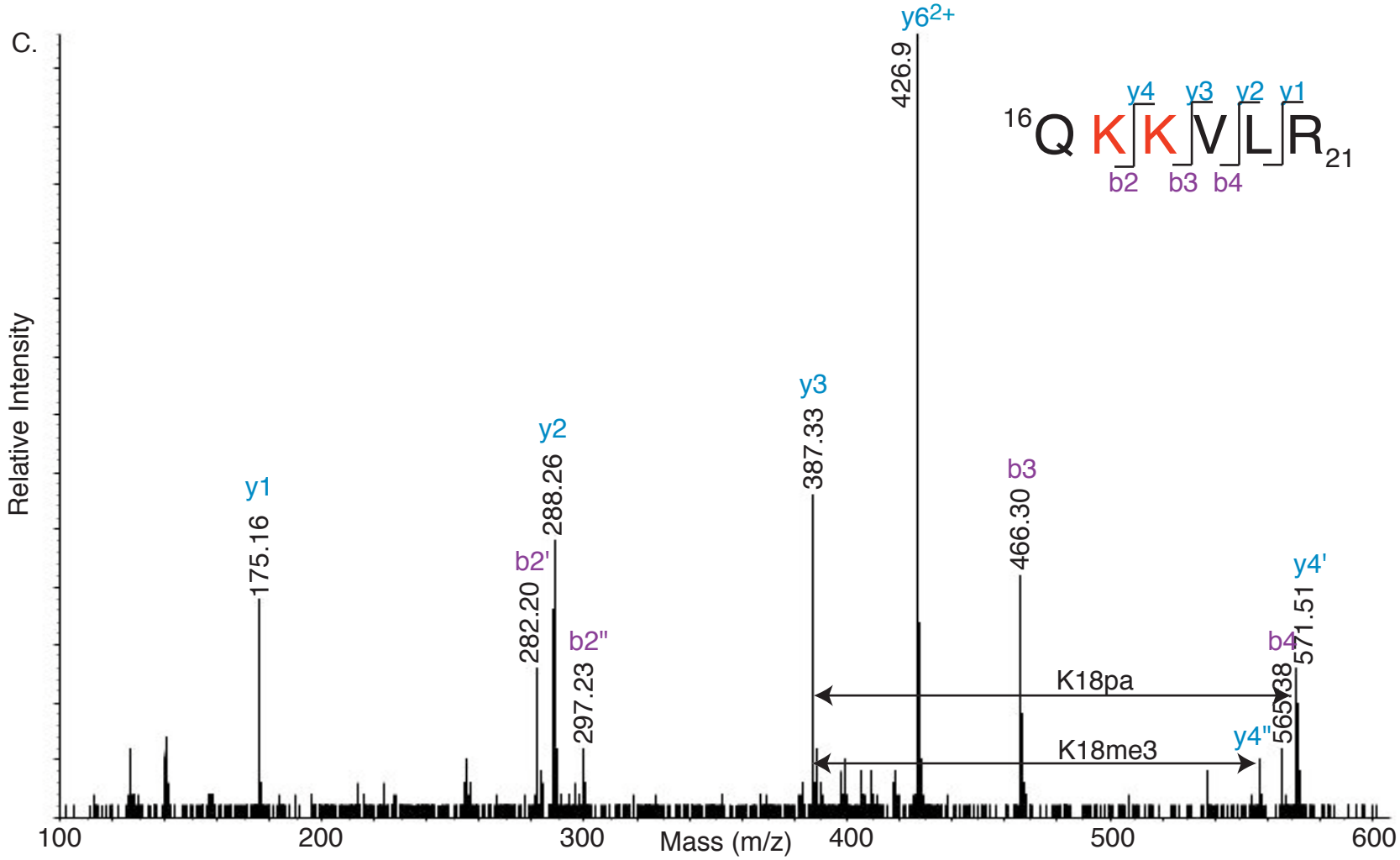


Figure 3.25C. Histone H4 lysines 17 and 18 are trimethylated. The peak labeled M + 81 Da represents two modified species of peptide 16-21: QK17me3KVLR (b2' and y4') and QKK18me3VLR (b2" and y4"). The lysine that is not trimethylated is propionylated (+56), and glutamine has rearranged to form pyroglutamic acid (-17). Modified lysines are shown in red. m/z 426.9^{2+}; theoretical MW = 851.5 Da; measured MW = 851.8 Da.

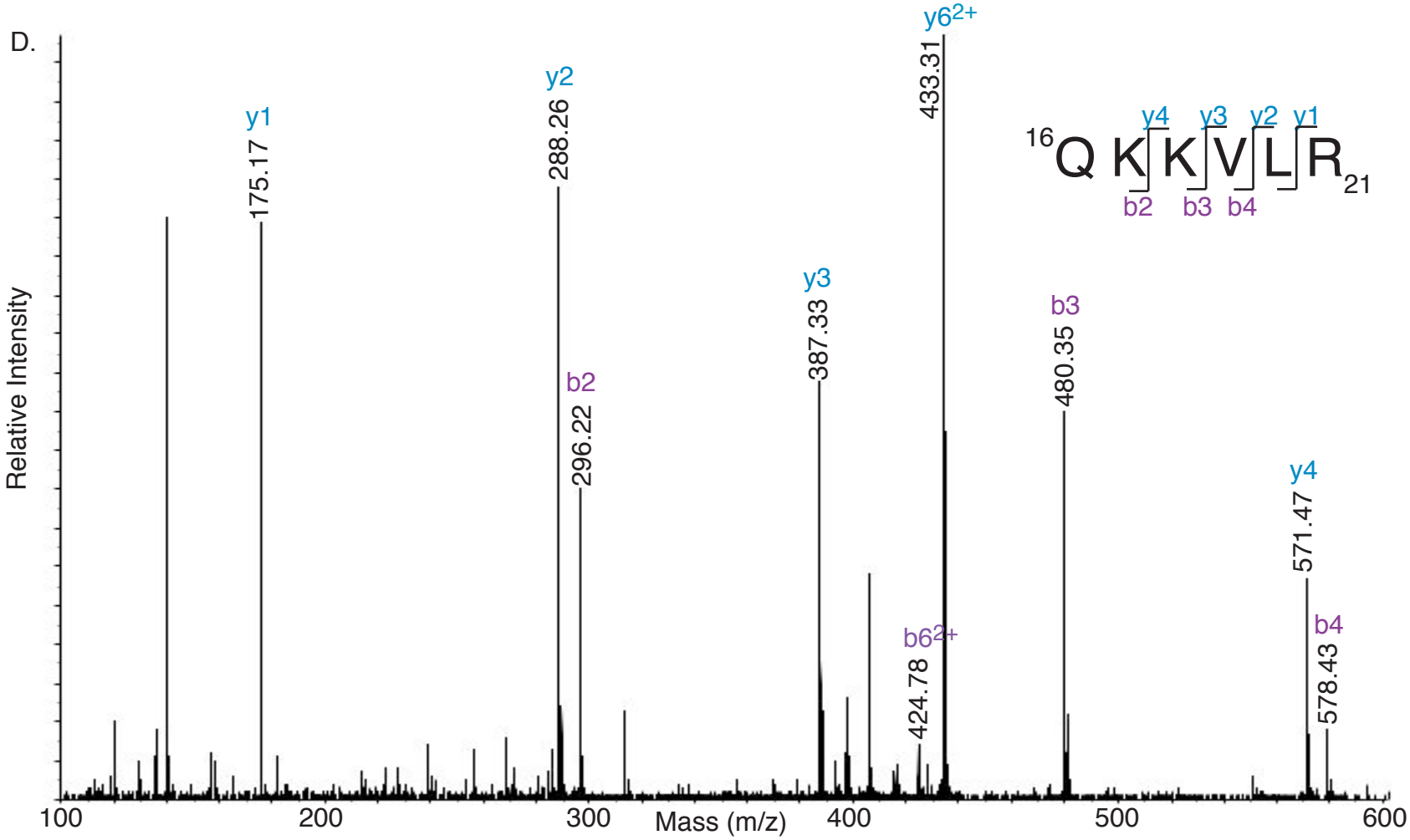


Figure 3.25D. Histone H4 glutamine 16 rearranges to form pyroglutamic acid. The peak labeled M + 95 represents a species of peptide 16-21 in which both lysines are propionylated (unmodified). Compared to the unmodified peptide, all of the b ions are shifted -17 Da, which is due to the loss of ammonia from glutamine. m/z 433.4^{2+}; theoretical MW = 865.5 Da; measured MW = 865.8 Da.

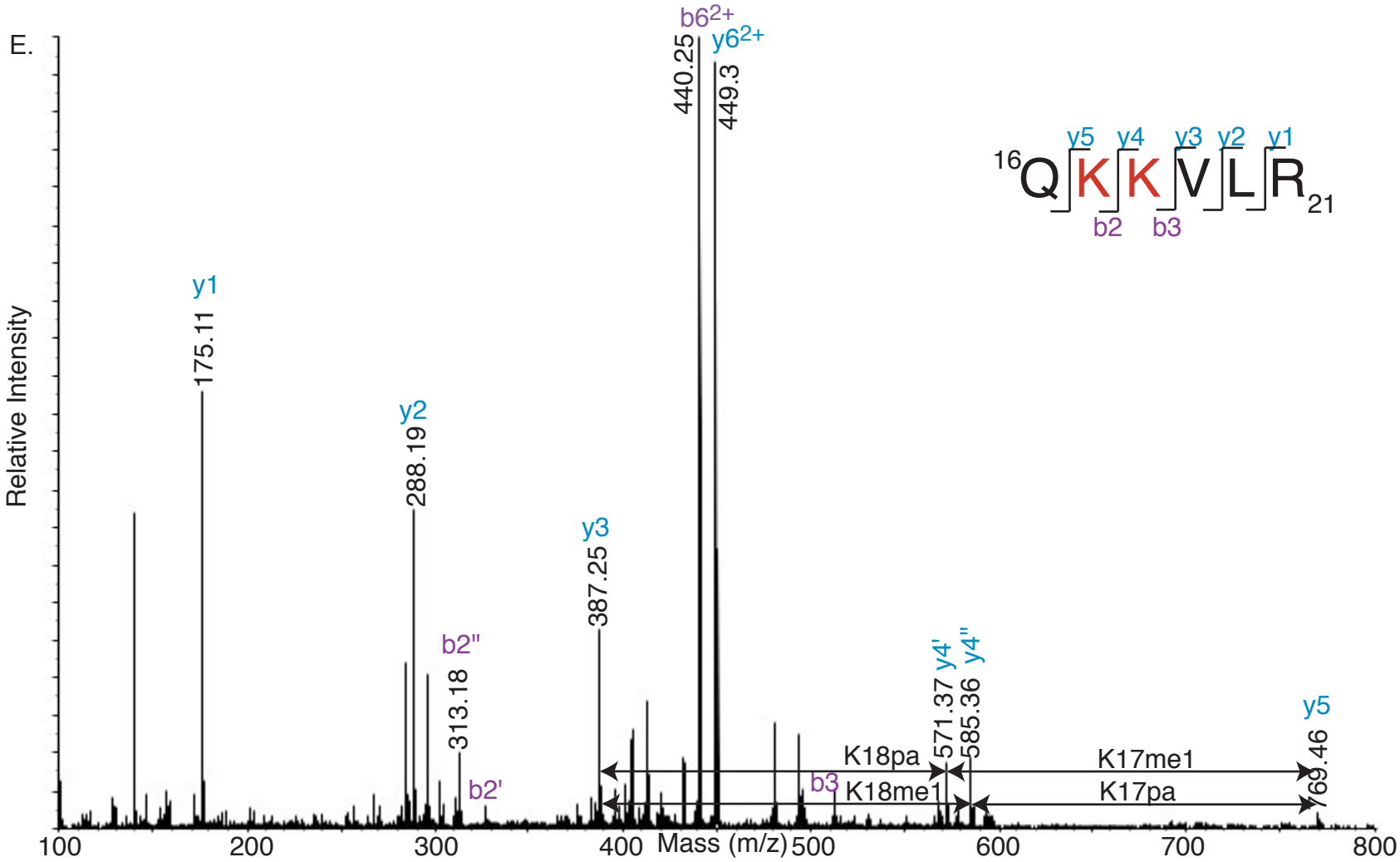


Figure 3.25E. Histone H4 lysines 17 and 18 are monomethylated. This tandem MS spectrum represents two modified species of peptide 16-21: QK17me1KVLR (b2' and y4') and QKK18me1VLR (b2" and y4"). The lysine that is not monomethylated is unmodified (propionylated). Modified lysines shown in red. m/z 449.3^{2+}; theoretical MW = 897.5 Da; measured MW = 897.6 Da.

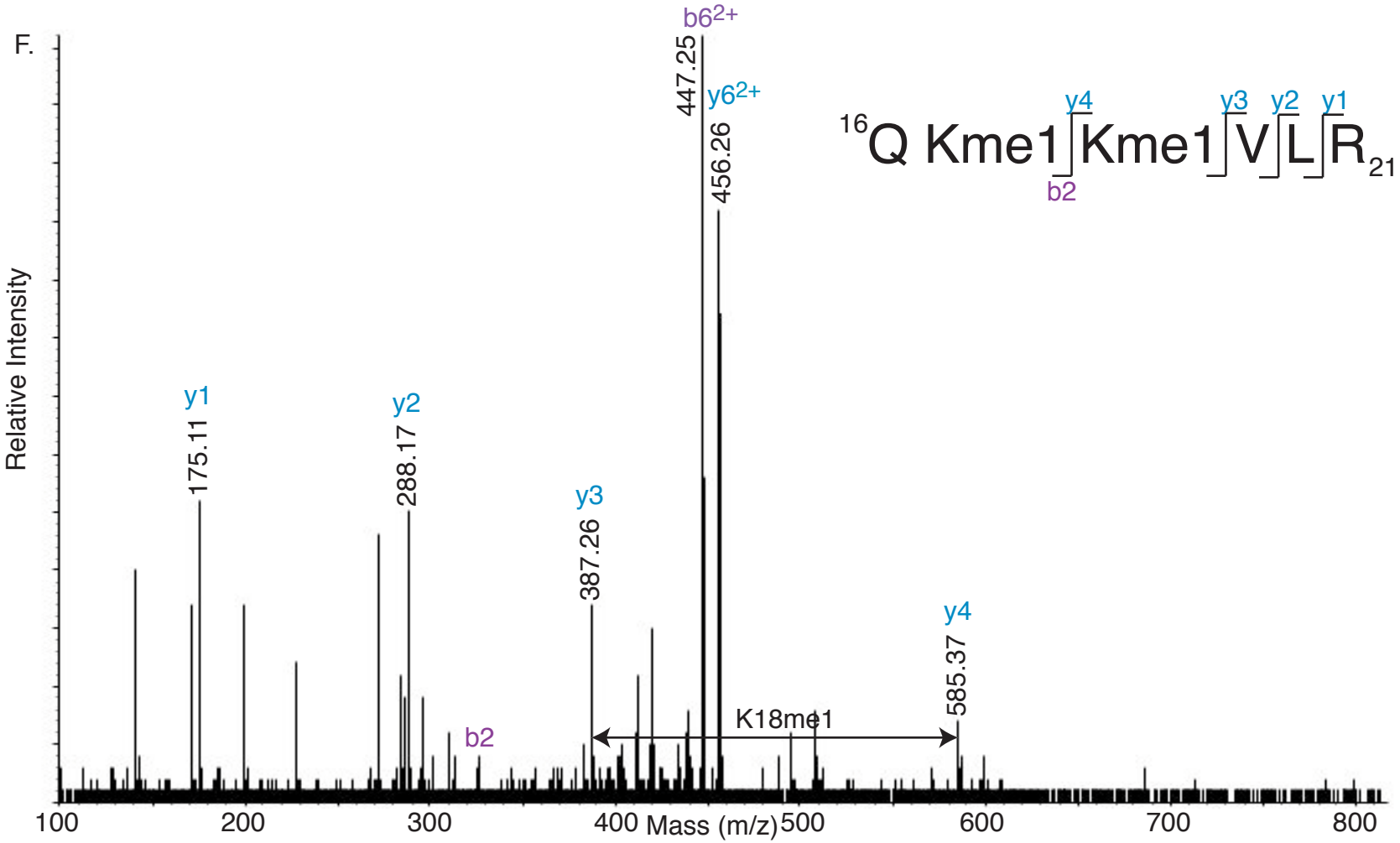


Figure 3.25F. Histone H4 lysines 17 and 18 are monomethylated. This spectrum shows that there is modified species of peptide 16-21 in which both K17 and K18 are monomethylated. m/z 456.3^{2+}; theoretical MW = 911.5 Da; measured MW = 911.6 Da.

in their modified states. Also, acetyl and trimethyl marks, both of which result in a +42 Da mass shift, were assigned based on Edman data.

Sequencing of peptide 1–15 by tandem MS revealed a modification, K2me2, which was not observed by Edman degradation (Fig. 3.24B). Two species of peptide 1-15 are represented on this spectrum at m/z 934.7^{2+}, A1me1K2acGK4acKSGEAK10acGSQKR and A1me1K2me2GK4acKSGEAKGSQKR. Only one set of y1-y5 ions is present, indicating that the two species are equivalent, and unmodified, at the sequence GSQKR (H4 11-15). There is also a single y14 ion, which, when subtracted from MH^+, shows that both species are monomethylated at A1. The two b2 ions demonstrate that K2 is dimethylated in one species and acetylated in the other. Two b5 ions show that both species have an acetylated lysine, which I assigned to K4. Finally, two y6 ions are present 14 Da apart, and the mass difference between y5 and the two y6 ions indicates that K10 is acetylated in one species and propionylated in the other.

Tandem MS sequencing of other modified species of peptide 1-15 confirmed that K2, K4, and K10 can be acetylated (Fig. 3.24B-F). No modifications were observed on K5 or K14 with this digest. Additional ions found on these spectra were due to cleavage at K2 and independent fragmentation of the peptide (KGKKSGEAK) resulting from this cleavage.

Tandem MS analysis of peptide 16–21 (QKKVLR) demonstrated that K17 is mono- and tri-methylated, and K18 is mono-, di-, and tri-methylated (Fig.

3.25B-F). The presence of only two modifiable lysines made the data easier to interpret than the spectra representing peptide 1-15; however, because the two lysines are adjacent to each other, only two ions, b2 and y4, can be used to distinguish between modification states.

Tandem MS spectra representing peptide 16–21 were initially difficult to interpret because the mass of some of the modified peptides was 17 Da less than the theoretical mass (Fig. 3.25B-D). Closer examination of these spectra demonstrate that glutamine 16 (Q16) forms a cyclic compound between its free amine and amide carbon, resulting in the loss of ammonia and consequent molecular weight decrease of 17 Da. The cyclization of glutamine to form pyroglutamic acid has been characterized previously by MS [185].

MS analysis of Glu-C-digested H4 was focused on N-terminal peptides 1–8 and 9–22. I was unable to acquire MALDI-TOF data for peptide 1–8 and found only one modified species by tandem MS, demonstrating that A1 is monomethylated and K4 is acetylated (Fig. 3.26, Table 2). Edman degradation suggests that this is the most abundant species of peptide 1–8. MALDI-TOF analysis of the fractionated Glu-C digest showed that peptide 9–22 is present in multiple modification-states (Fig. 3.27A). Tandem MS sequencing of peptide 9–22 illustrated the combinations in which K10ac, K17me, and K18me may be present (Fig. 3.27B-F, Table 2). Modified peptide 9-22 was observed as either a +4 or +5 species, which complicated the interpretation of these spectra. I was unable to sequence all of the modified species of peptide 9–22, so PTMs for the

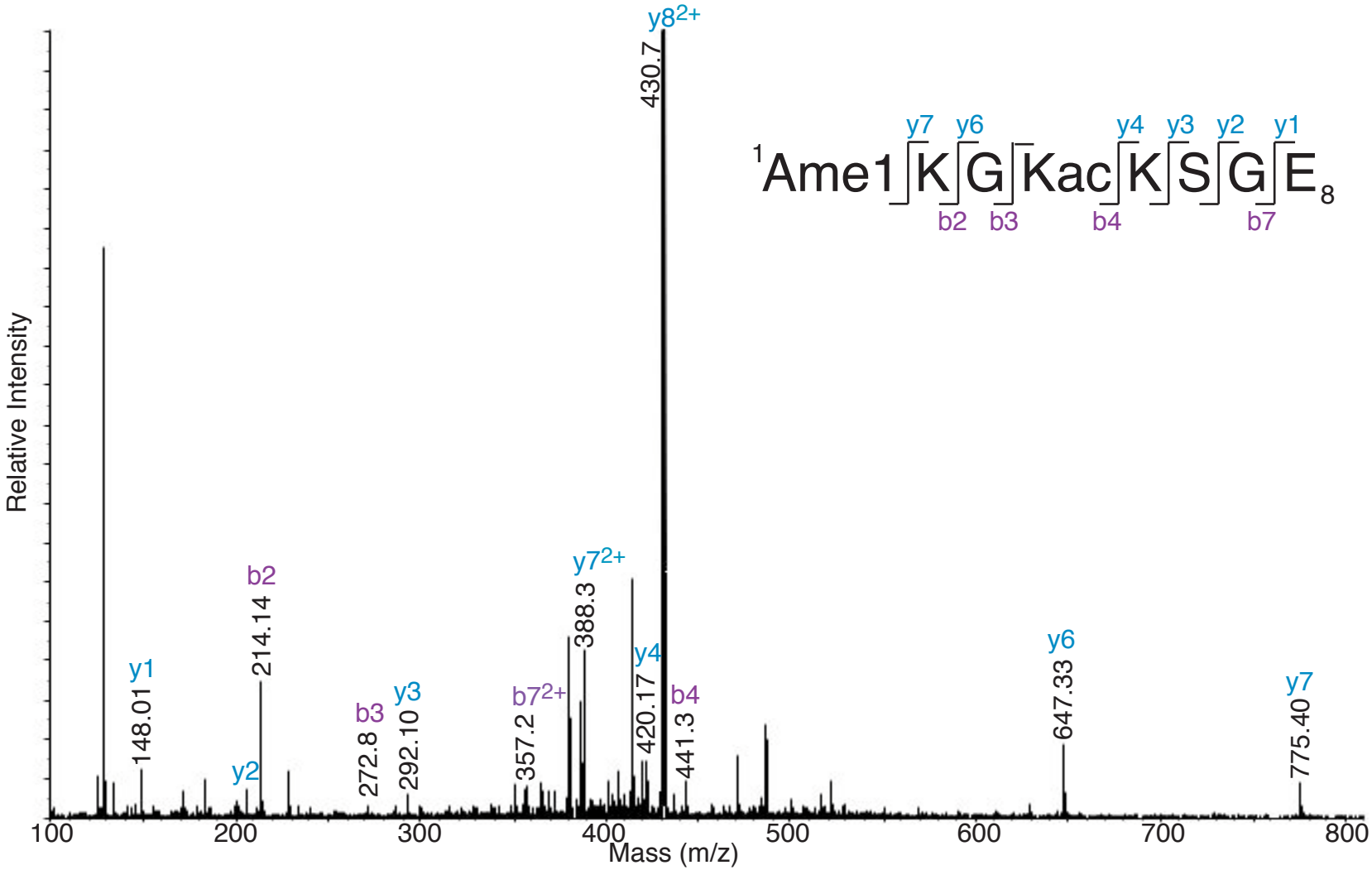


Figure 3.26. Histone H4 alanine 1 is monomethylated and lysine 4 is acetylated. H4 was digested with endoproteinase Glu-C, producing peptide 1-8 (AKGKKSGE). A1 monomethylation is demonstrated by ion pair y7-y8, and K4 acetylation is demonstrated by ion pairs b3-b4 and y4-y6. m/z 430.7^{2+}; theoretical MW = 860.5 Da; measured MW = 860.4 Da.

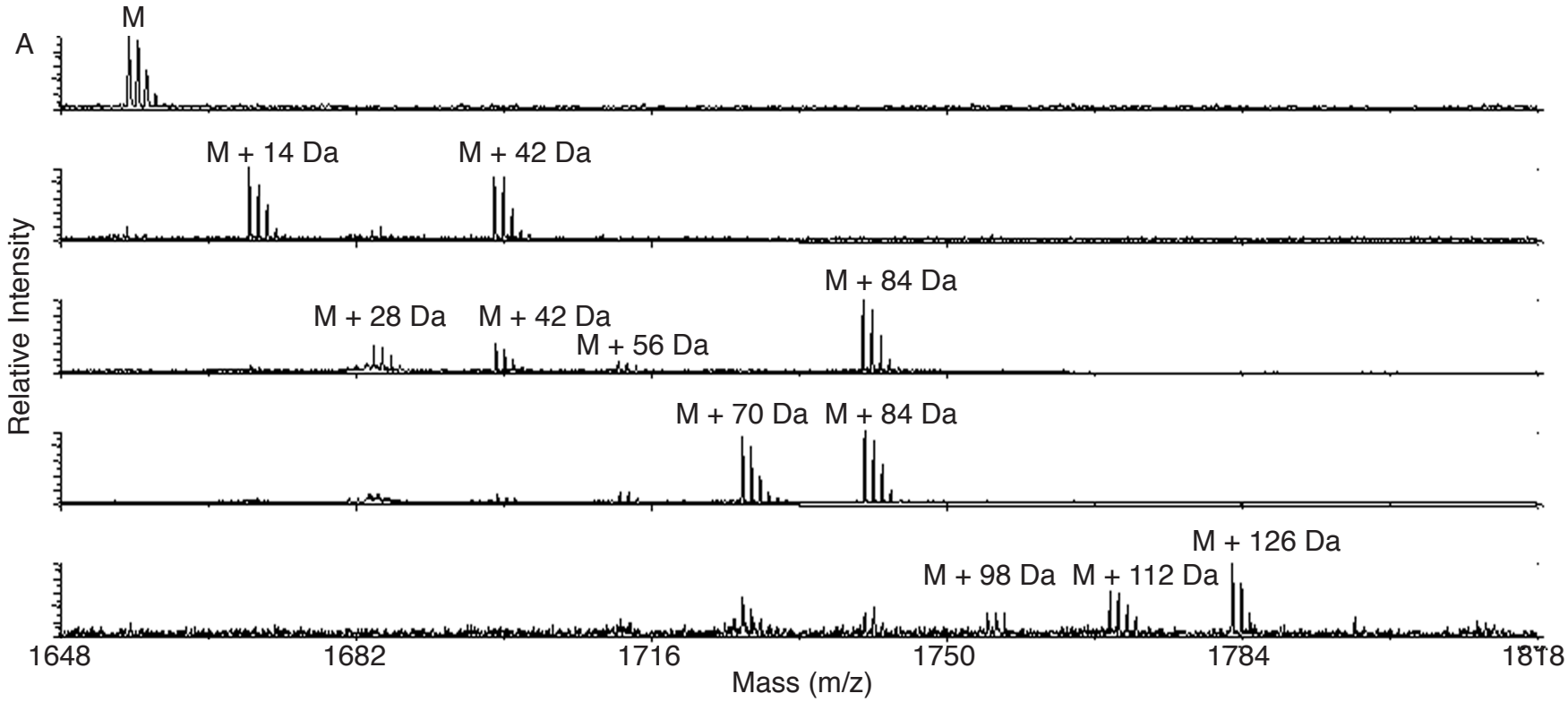


Figure 3.27. Histone H4 peptide 9-22 is present in multiple modification states. H4 digested by endoproteinase Glu-C was fractionated by HPLC. (A) MALDI-TOF data were acquired for 5 consecutive HPLC fractions. Peaks are separated by 14 Da, the molecular mass of a methyl group. M is the molecular ion mass of peptide 9-22 (1655 Da). Sequencing by tandem MS (B-F) confirmed that peaks M + 42 Da, M + 56 Da, M + 70 Da, M + 84 Da, and M + 126 Da correspond to peptide 9-22 (AKGSQKRQKKVLRE). The remaining peaks were incompletely sequenced.

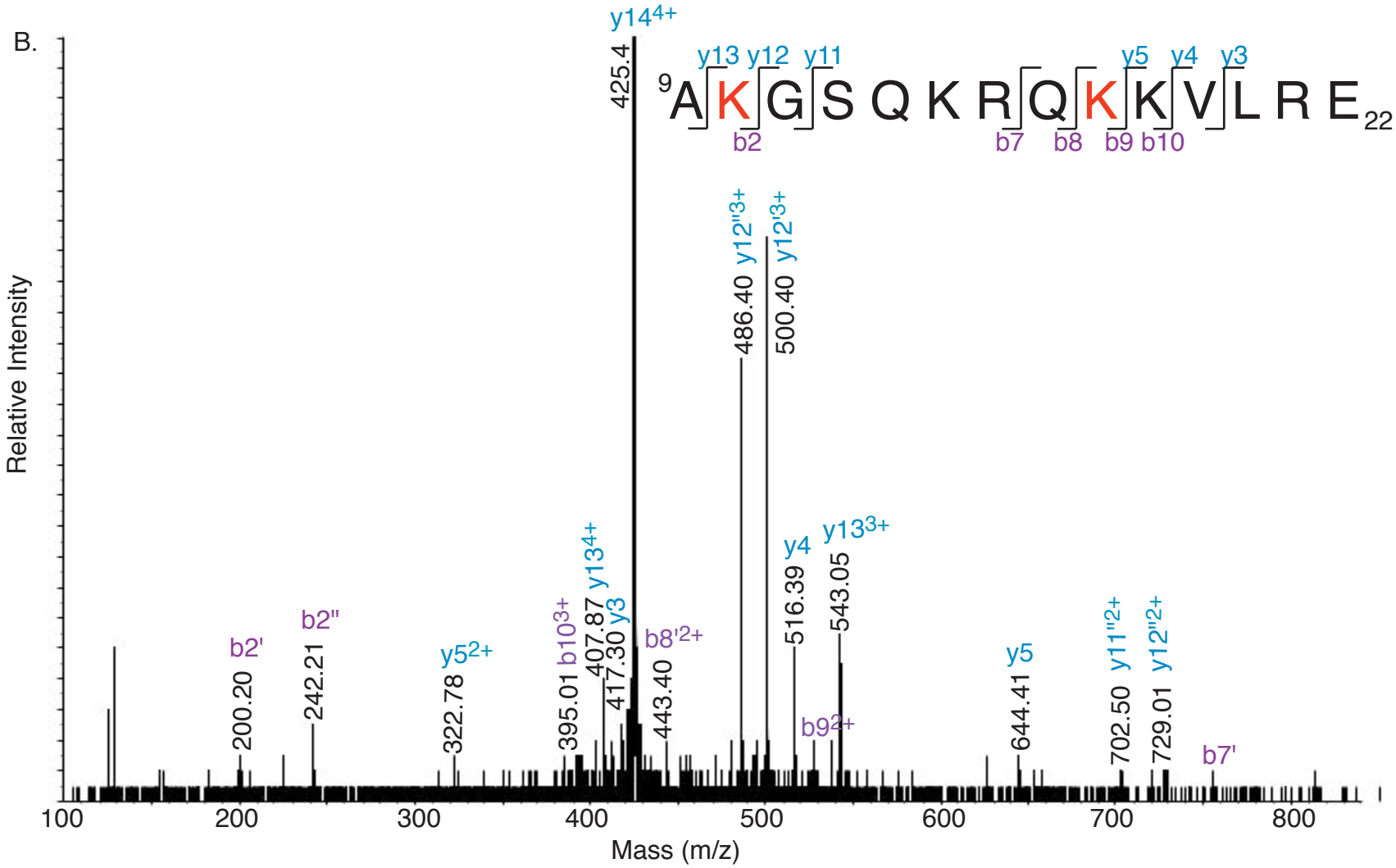


Figure 3.27B. Histone H4 lysine 10 is acetylated and lysine 17 is trimethylated. The peak labeled M + 42 Da represents two modified species of peptide 9-22: AK10acGSQKRQKKVLRE (b2", y11", y12") and AKGSQKRQK17me3KVLRE (b2', b7', b8', y12'). m/z 425.4^{4+}; theoretical MW = 1698.0 Da; measured MW = 1698.6 Da.

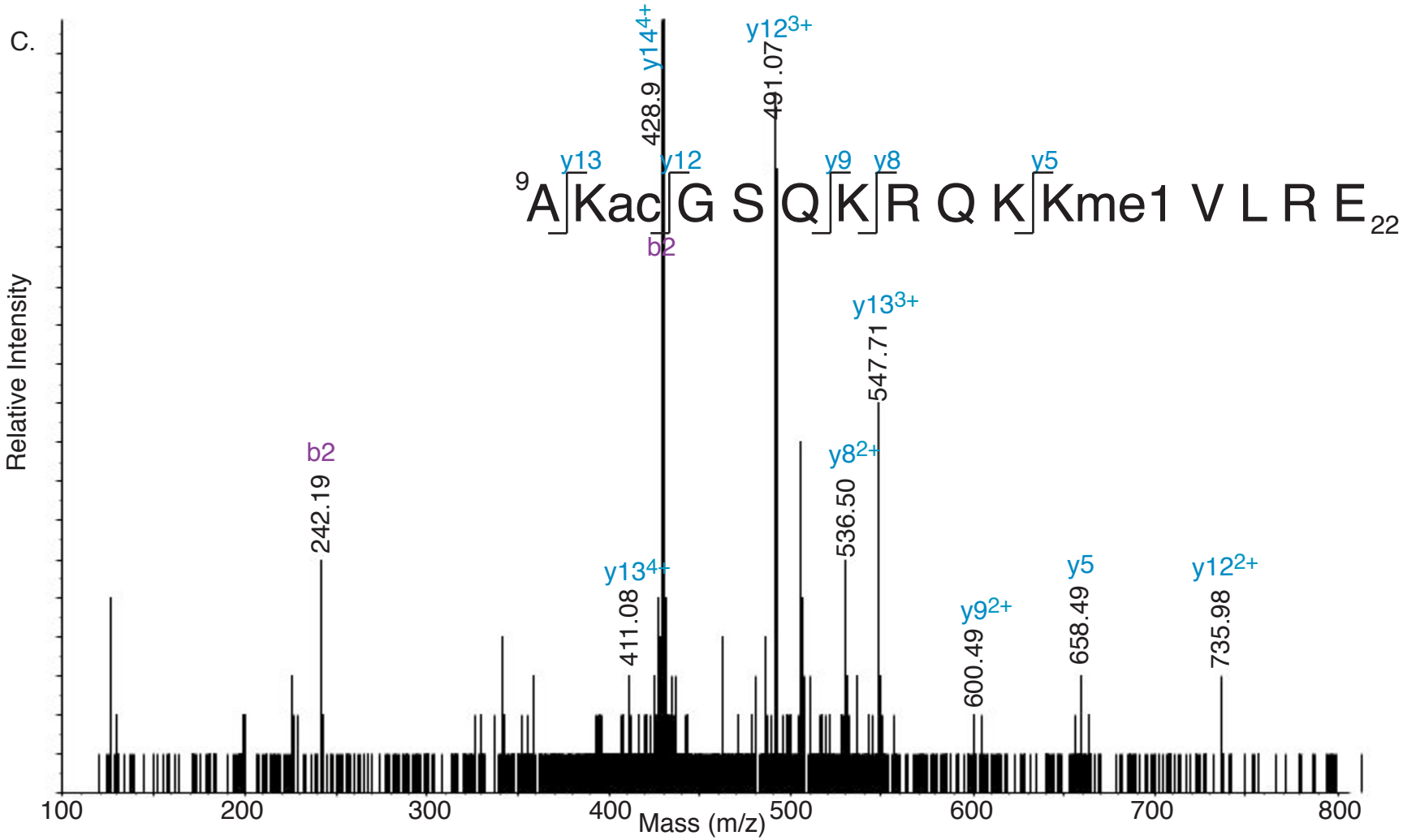


Figure 3.27C. Histone H4 lysine 10 is acetylated and lysine 18 is monomethylated. The peak labeled M + 56 Da represents a modified species of peptide 9-22 in which both K10 and K18 are modified. Although this spectrum is incomplete, the presence of the b2 and y12-y13 ions indicate that K10 is acetylated, and the y5 ion suggests that K18 is monomethylated. m/z 428.9^{4+}; theoretical MW = 1712.0 Da; measured MW = 1712.6 Da.

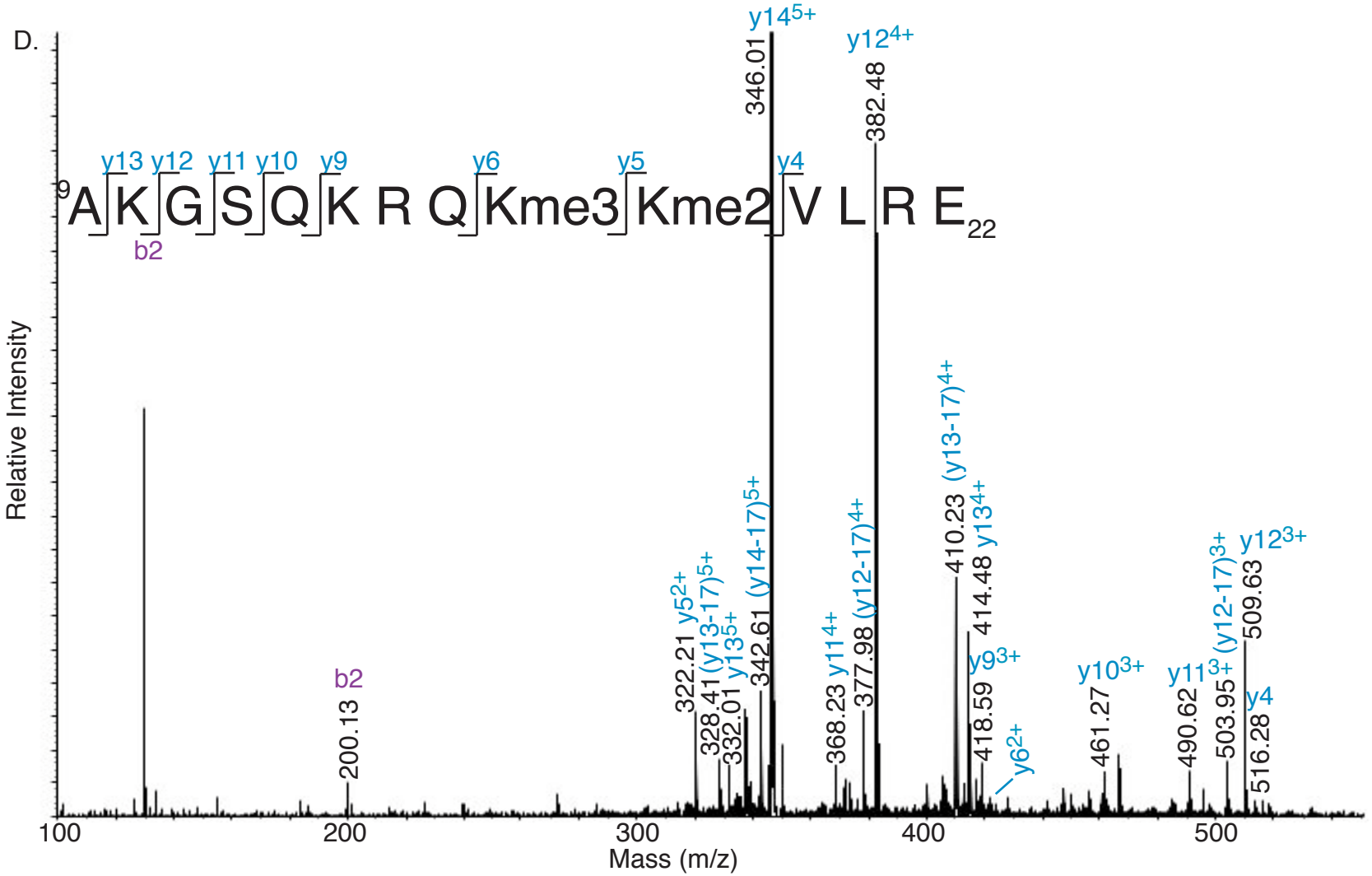


Figure 3.27D. Histone H4 lysine 17 is trimethylated and lysine 18 is dimethylated. The peak labeled M + 70 Da represents a modified species of peptide 9-22 in which both lysine 17 and 18 are modified. These modifications are confirmed by the y4-y5 and y5-y6 ion pairs. m/z 346.0^{5+}; theoretical MW = 1726.0 Da; measured MW = 1726.0 Da.

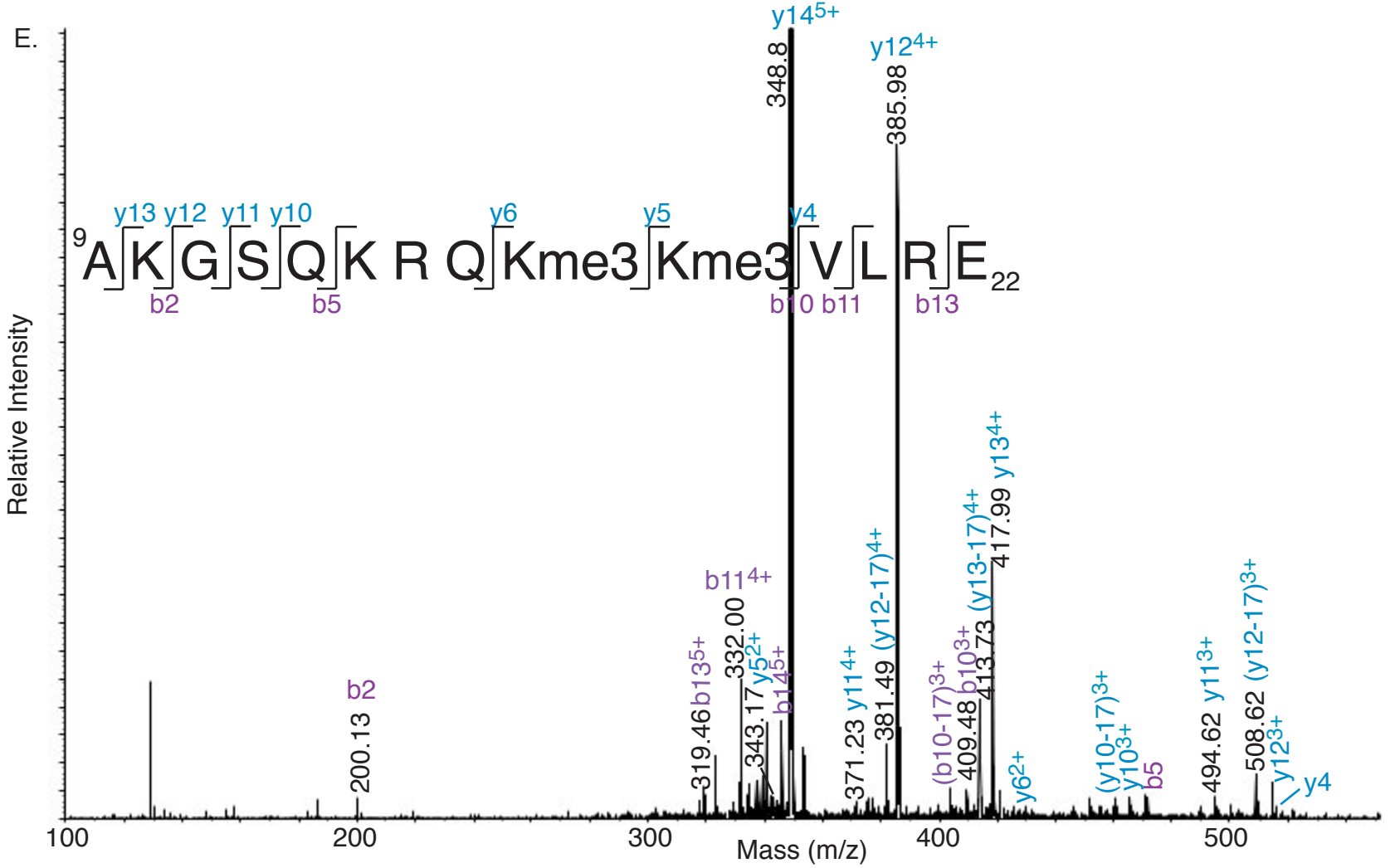


Figure 3.27E. Histone H4 lysines 17 and 18 are trimethylated. The peak labeled M + 84 Da was sequenced, and ion pairs y4-y5 and y5-y6 demonstrate that K17 and K18 are modified. m/z 348.8^{5+}; theoretical MW = 1740.0 Da; measured MW = 1740.0 Da.

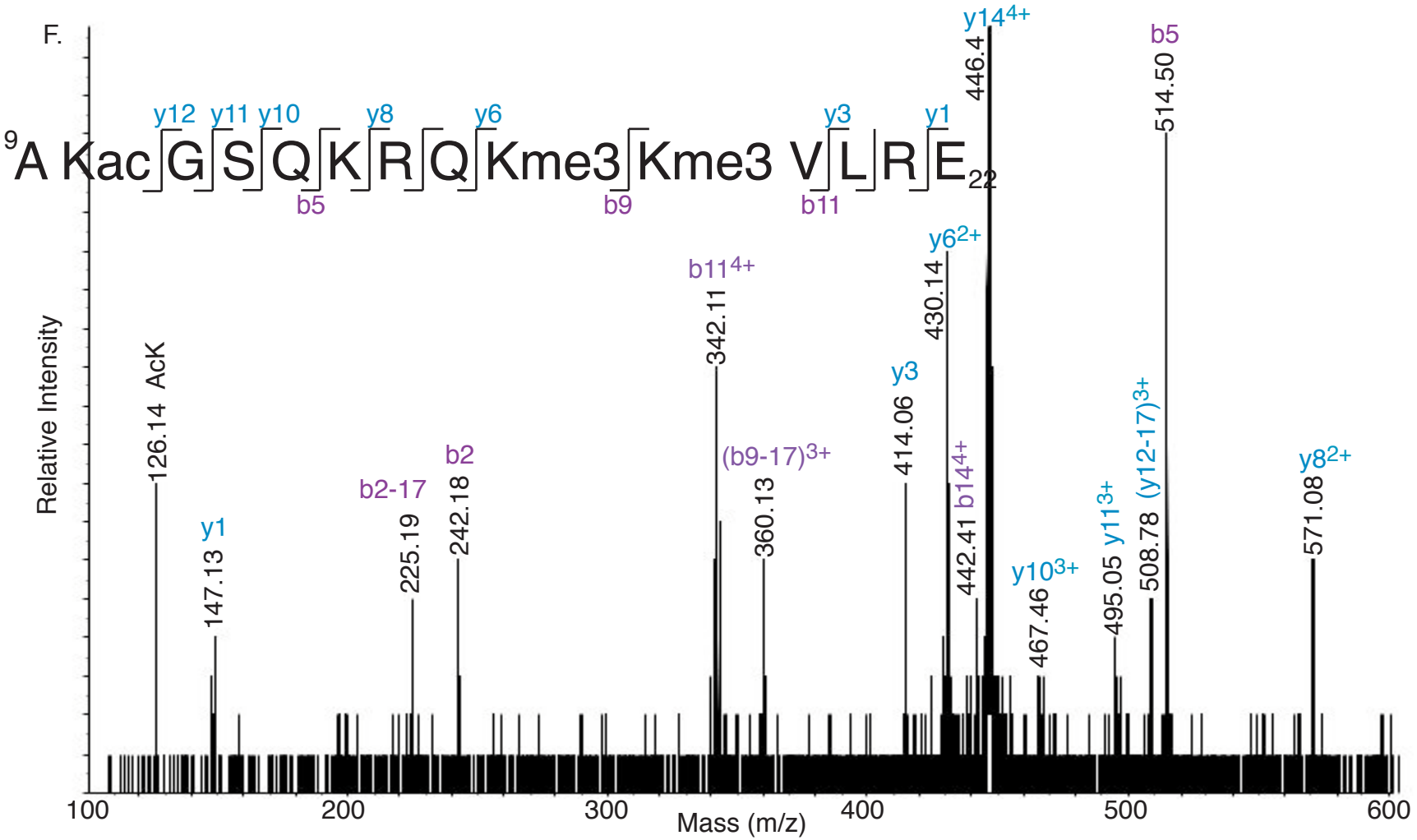


Figure 3.27F. Histone H4 lysine 10 is acetylated and lysines 17 and 18 are trimethylated. The peak labeled M + 126 Da was sequenced by tandem MS, and the b2 ion shows that K10 is acetylated. The y3 and y6 ions show that the additional +84 Da fall on the sequence KKV, which we assign as trimethyl K17 and K18. m/z 446.4^{4+}; theoretical MW = 1782.0 Da; measured MW = 1782.6 Da.

following species were not assigned: M + 14 Da, M + 28 Da, M + 98 Da, and M + 112 Da.

Two modified species of peptide 9-22 were observed at 425.4^{4+}, represented by the M + 42 Da peak: AK10acGSQKRQKKLVRE and AKGSQKRQK17me3KLVRE (Fig. 3.27B). The presence of two b2 and y12 ions shows that K10 is acetylated in one species and unmodified in the other. The presence of the b8 ion for the species unmodified at K10 shows that the additional +42 Da must fall on the sequence KKVLRE (H4 17-22). Only one set of y3-y5 ions is present, indicating that the two are equivalent, and unmodified, at the sequence KVLRE (H4 K18-22). Therefore, K17 must be trimethylated in the second modified species.

The significance of these PTMs and the combinations in which they are present is currently unknown, but functional studies of individual marks and their corresponding histone-modifying enzymes will enable us to interpret these data in a meaningful way in the future.

Comparison of Posttranslational Modifications between Bloodstream-Form and Procyclic-Form Histones

All of the modifications observed in PF were also identified in BF, but the analysis of BF histones was more extensive and uncovered additional PTMs (Table 3) [165]. I cannot dismiss the possibility that these PTMs may also be

TABLE 3. Summary of Histone Modifications in Trypanosomes						
	Bloodstream Forms			Procyclic Forms		
Histone	N-terminus	Acetyl Lysine	Methyl Lysine	N-terminus	Acetyl Lysine	Methyl Lysine
H2A	$A1_{me1}$ (60%)[a, b] $A1_{ac}$[b]	K4 (1%)[a] K115[b], K119[b], K120[b], K122[b], K125[b], K128[b]	none	$A1_{me1}$ (>80%)[a]	K115[b], K119[b] K120[b] K125[b] K128[b]	none
H2B	$A1_{me1}$ (60%)[a, b] $A1_{me2}$[b]	K4 (1%)[a] K12[b] K16[b]	none	$A1_{me1}$ (>80%)[a]	none	none
H3	$S1_{ac}$[b]	K23[b]	$K32_{me3}$[b] $K76_{me1, me2, me3}$[b]	$S1_{ac}$[b]	none	$K76_{me1, me2, me3}$[b]
H4	$A1_{me1}$ (60%)[a, b]	K2 (2%)[a, b] K4 (73%)[a, b] K5 (7%)[a, b] K10 (7%)[a, b]	$K2_{me2}$[b] $K17_{me1}$ (4%), $K17_{me2}$ (8%), $K17_{me3}$ (9%)[a, b] $K18_{me1}$ (6%), $K18_{me2}$ (4%), $K18_{me3}$ (18%)[a, b]	$A1_{me1}$ (>80%)[a, b]	K4 (80%)[a, b] K10 (10%)[a]	$K17_{me1}$ (<10%), $K17_{me2}$ (<5%)[a] $K18_{me1}$ (<10%), $K18_{me3}$ (<5%)[a]

[a] PTMs observed by Edman degradation. The levels of acetylation and methylation, shown in parenthesis, were quantified by comparison to known standards. Three trials were performed for BF H4, providing an error of +/-2%.
[b] PTMs observed by tandem MS

present in PF given the opportunity to perform additional studies, so a direct comparison may not be possible with the current MS data. However, Edman degradation data allowed us to draw some distinctions. Analysis of histone PTMs showed that H2A K4, H2B K4, and H4 K2 and K5 are acetylated at low levels in BF, but not PF. H4 K2 and K5 are adjacent to modified amino acids A1 and K4, respectively, so it is possible that K2 and K5 acetylation modulates the effects of the major marks in BF, but not in PF, possibly by interfering with the binding of effector proteins. This phenomenon is well-documented in other organisms [186]. Similarly H2A and H2B K4 acetylation may affect interactions between monomethylated A1 with its binding partner in BF. Also, monomethylation of A1, which was identified in both PF and BF, was found to a greater extent in PF: more than 80% of PF H2A, H2B, and H4 were monomethylated at A1, while only 60% of BF histones were monomethylated at A1. These results suggest that the downstream biological effects of H2A, H2B, and H4 N-terminal PTMs, especially N-methylation of A1, may be subtly different in PF and BF.

Methylation of K17 and K18 also demonstrated lifecycle differences: trimethylation of K17 and dimethyation of K18 were not observed in PF histones. Additionally, less than 5% of PF H4 K18 was trimethylated, while 18% of BF H4 K18 was trimethylated. It is possible that different permutations of K17 and K18 methylation result in the binding of different effector proteins, or that they modulate the binding efficiency of a single effector protein. Regardless,

differences in K17 and K18 methylation levels in PF and BF H4 suggest different biological outcomes for the two lifecycles.

It is not unexpected that the differences between BF and PF are subtle. The observation that different PTMs are present in the two lifecycle stages warrants further investigation. However, it should be noted that Edman degradation is not a sensitive technique, so PTMs observed in low abundance exclusively in BF may have simply been overlooked in PF.

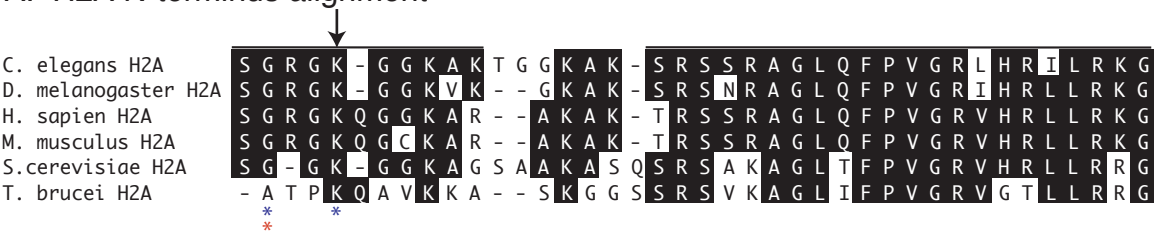
Discussion

In this study, I attempted to identify the major covalent modifications present on the four core histones of BF *T. brucei*. All of the identified PTMs and the methods used to discover them are summarized in Table 3. The major obstacle to this endeavor was H3, because sequencing of this histone was incomplete. The N-terminus was blocked to Edman degradation, and minimal tandem MS data were acquired for modified N-terminal peptides. However, Qq-TOF data of the H3 PA-trypsin and Asp-N digests suggested that several lysines at the H3 N-terminus can be modified. MALDI-TOF data of the PF H3 N-terminus also suggested that a number of PTMs were present in this lifecycle stage, although no tandem MS data were acquired to assign PTMs to their specific amino acids (C. Janzen, unpublished data).

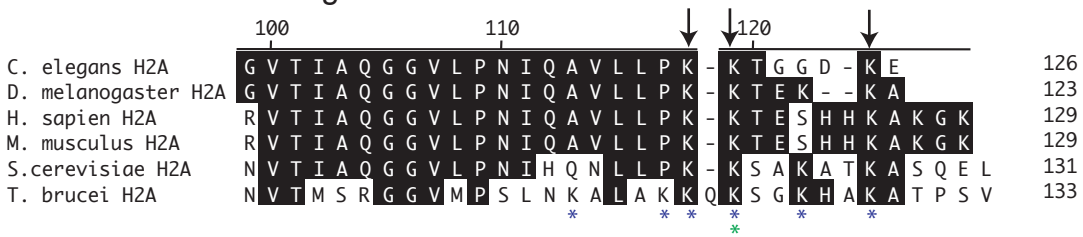
Although there is no *T. brucei* homologue to H3 K9, which is associated with heterochromatin when methylated, there are predicted homologues of modified lysines 4, 18, 23, 27, and 36 (*T. brucei* K4, K16, K19, K23, and K32) (Fig. 3.28D). Sequencing of trypsin-digested H3 by tandem MS showed that *T. brucei* K23 is acetylated and K32 is trimethylated, the homologues of K27 and K36 in other organisms. K27 can be both methylated and acetylated, and K36 is methylated. In humans and other organisms, methylation of H3 K27 recruits the Polycomb group proteins, which mediate transcriptional silencing [187]. In *S. cerevisiae,* methylation of H3 K36 by Set2 is correlated with transcriptional activation [188]. Set2 has been shown to affect the process of transcription elongation through its association with the RNA polymerase II elongation complex [189].

I also found that H3 K76 in the core domain could be mono-, di-, or tri-methylated. K76 is the homologue of H3 K79 in *S. cerevisiae* and has been associated with euchromatin when methylated. H3 K76 di- and trimethylation and the responsible histone methyltransferases, DOT1A and DOT1B, have been studied in detail in *T. brucei* [166]. Deletion of trypanosome DOT1 proteins caused defects in cell cycle regulation and differentiation between lifecycle stages. This study demonstrated that, while *T. brucei* may have homologues to histone modifications and corresponding histone-modifying enzymes in *S. cerevisiae*, the biological functions of these marks may be organism-specific.

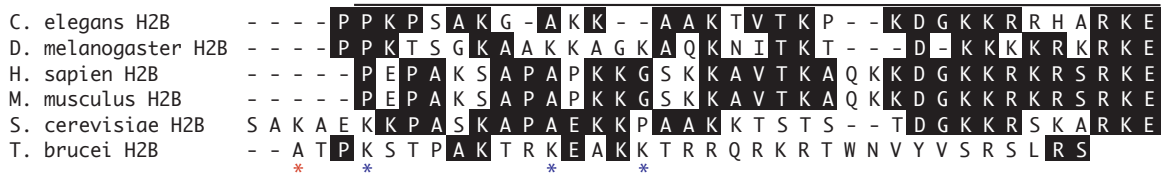
A. H2A N-terminus alignment



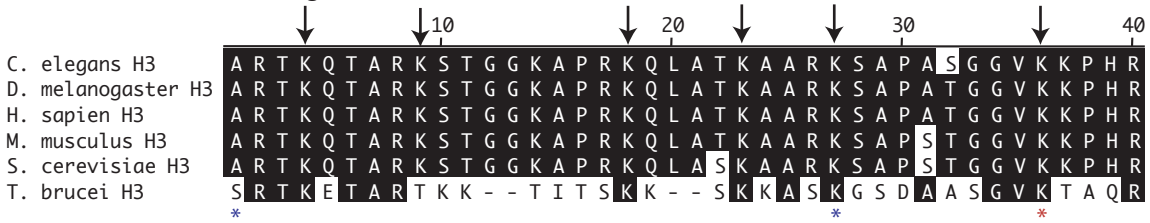
B. H2A C-terminus alignment



C. H2B N-terminus alignment



D. H3 N-terminus alignment



E. H4 N-terminus alignment

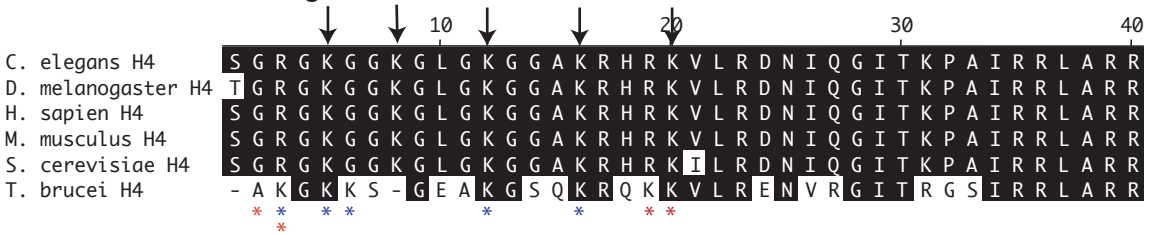


Figure 3.28. Fig. 1. Sequence alignments of the N-termini of each core histone and the C-terminus of H2A. Identical residues are shaded. Important sites of homology, which are referred to in the text, are indicated by arrows. The sequence ruler, which is included when the sequences are concordant, is numbered according to the *Homo sapiens* sequence. Residues found to be acetylated in T. brucei histones are indicated with an (*) and methylated residues with a (*). A putative site of ubiquitination at the H2A C-terminus is indicated with a (*). (A) H2A N-terminus. (B) H2A C-terminus. (C) H2B N-terminus. (D) H3 N-terminus. (E) H4 N-terminus.

In a later chapter, I began to characterize trimethyl H3 K4 using a specific antibody. Therefore, our efforts in this chapter were targeted toward identifying H3K4me3 by MS. I was unsuccessful for two reasons: 1) the nature of the peptides produced by the various digests were not amenable to sequencing by tandem MS, and 2) the mark is most likely present in low abundance, so the signal may have been suppressed by more abundant peptides. H3 peptide 1-25, which was produced by the Asp-N digest, has 2 arginines and 8 lysines and was observed as a +5 species. For these reasons, MS/MS data for modified peptide 1-25 were uninterpretable. The Glu-C (H3 1-5, SRTKE) and trypsin (H3 3-8, TKETAR) digests both produced K4-containing peptides that were small and hydrophilic. It has been suggested that these peptides might elute immediately from the HPLC column prior to being analyzed by the mass spectrometer. The unmodified PA-trypsin peptide 3-8 was observed, indicating that perhaps neutralizing the charge of K4 by propionylation reduced the hydrophilicity of the peptide, causing it to be retained on the HPLC column. The low abundance of the K4 methyl mark also contributed to its elusiveness and was further complicated by the presence of H2A peptides, which were present at a higher concentration than H3 peptides.

Next I investigated the PTMs on *T. brucei* histone H4. It has been shown that the acetylation of the *S. cerevisiae* H4 N-terminus at K5, K8, K12, and K16 is classically associated with transcriptional activation for two reasons [9]. First, acetylation functions to decrease the positive charge of the histone tail, thereby

disrupting DNA-histone interactions and allowing for increased access to DNA by transcription factors [11]. Second, acetylated and unacetylated lysines serve as binding sites for effector proteins involved in transcriptional silencing or activation. According to our alignment, *S. cerevisiae* K5, K12, and K16 correspond to *T. brucei* K4, K10, and K14, all of which are acetylated (Fig. 3.28E). Although the *T. brucei* H4 sequence diverges considerably from the extremely conserved H4 N-terminus of other organisms, 3 of the 4 lysines appear to be conserved. Acetylation of *T. brucei* K4 and K10 suggests that the link between H4 N-terminal acetylation and transcriptional activity is conserved in an evolutionarily divergent organism.

Importantly, in addition to the conserved acetylated lysines found at the H4 N-terminus, there is an apparent homologue of the methylated H4 K20, *T. brucei* H4 K18 (Fig. 3.28E). Trimethyl H4 K20 is a marker of constitutive and facultative heterochromatin and is in competition with the hyperacetylated H4 N-terminus [190]. Monomethyl H4 K20 is present at active gene promoters in association with H4 hyperacetylation, which demonstrates that the degree of methylation may be associated with distinct transcriptional states [191]. I found that *T. brucei* H4 K18 and the adjacent K17 may both be mono-, di-, or tri-methylated. K17 and K18 exist together in a number of different modification states (Table 2). It is unlikely that these different methylation levels are created by different methyltransferases because there are only 3 identifiable putative SET domain-containing methyltransferases in the *T. brucei* genome [148]. Instead, it is more

likely that a single methyltransferase acts on both K17 and K18, and the degree of methylation is determined by cofactors acting on the methyltransferase. I cannot predict whether the degree of methylation or the combination of modification states in which K17 and K18 co-exist will result in distinct biological readouts. If so, the amount of variability would allow for intricately regulated downstream effects.

Two modified lysines in the *T. brucei* H4 N-terminus, K2 and K5, did not have homologues in other organisms. I was especially interested in K2, which can be either acetylated or dimethylated. The modification of H4 K2 is very minor: Edman data showed that ~2% of H4 K2 is acetylated. MS/MS data showed that H4 K2 was dimethylated, but this modification was not observed during Edman analysis, so I expected that H4K2me2 was present in less than 1% of H4. Histones from other organisms possess a number of lysines that may be either acetylated or methylated. Chicken H3 K9, for example, is subject to both types of covalent modifications. Chromatin immunoprecipitation experiments showed that H3K9me3 is associated with inactive chromatin, whereas H3K9ac is found at the boundaries of heterochromatin [192]. The authors suggest that acetylation of H3 K9 acts as an insulator against the propagation of heterochromatin, which is initiated by H3K9me3 and H3K9me3 binding proteins. This study demonstrates that acetylation and methylation of a single lysine may have antagonistic roles in the establishment of different chromatin states. Whether *T. brucei* H4 K2 is involved in determining chromatin structure is unknown, but I propose that the

dual modification status of this lysine makes it an interesting target for further characterization.

The final point of interest at the H4 N-terminus is the N-methylation of A1, which was also observed in H2A and H2B. This is an unusual histone modification discovered recently in trypanosomatids [165,181]. Here, I also showed that A1 can be acetylated in H2A or dimethylated in H2B. I speculate that N-methylation of A1 in H2A, H2B, and H4 acts as a binding platform for an effector protein, and that binding may be modulated by neighboring marks (H4 K2 and H2A and H2B K4 acetylation).

Histones H2A and H2B are notable for the relatively few modifications found at their N-termini compared to other organisms. Both H2A and H2B begin with the sequence ATPK in *T. brucei*, and both are N-methylated at ~60% of A1 and acetylated at ~1% of K4. Methylation of A1, but not acetylation of K4, was confirmed by tandem MS for both H2A and H2B. *T. brucei* H2A K4 is homologous to human H2A K5, a residue that is acetylated by Hat1 and may be associated with the deposition of newly synthesized histones into chromatin (Figure 3.28A) [193]. Additionally, H2B is acetylated at K12 and K16, neither of which is a homologue of lysines in other organisms (Fig. 3.28C).

The H2A C-terminus contains 6 acetylated lysines (K115, K119, K120, K122, K125, and K128). Three of these, K120, K122, and K128, are probable homologues of lysines in other organisms, but the homologues of *T. brucei* K120 and K128 do not appear to be modified (Fig. 3.28B). *T. brucei* H2A K122 aligns

with human H2A K119, the site of ubiquitination on this histone [183]. In other organisms, H2A ubiquitination is associated with transcriptional silencing. During mouse spermatogenesis, H2A ubiquitination is enriched in a region of the murine nucleus known as the sex body, a structure containing the heterochromatic X and Y chromosomes [194]. It was shown in HeLa cells that the H2A E3 ubiquitin ligase complex is composed of four members of the Polycomb group proteins, thereby suggesting a link between Polycomb-mediated silencing and H2A ubiquitination [195]. There is some evidence to suggest that the *T. brucei* H2A C-terminus is ubiquitinated — the molecular weight of an H2A peptide 107–133 corresponds to a modified species that has one ubiquitinated and three acetylated lysines. Interestingly, *T. brucei* K122 is the only lysine at the H2A C-terminus that is adjacent to a possible phosphorylation site, S123. In other organisms, ubiquitinated lysines at the H2A and H2B C-termini are adjacent to phospho-acceptor sites. It has been suggested that phosphorylation influences ubiquitination and de-ubiquitination of its neighboring lysine and also influences whether effector proteins bind ubiquitinated lysine [186]. Based on our data and the alignment of the H2A C-terminus with the C-terminus of H2A from other organisms, I propose that *T. brucei* H2A K122 can be ubiquitinated.

A similar study identifying covalent histone modifications in PF was published recently (Table 3) [165]. Analysis by Edman degradation allowed us to draw some direct comparisons (see Results), but I cannot make many comparisons based on the MS data. A number of acetylated lysines were found

in BF histones by MS that were not present in PF histones (H2A K122, H2B K12 and K16, H4 K2 and K5), but I did not attribute these to lifecycle differences. Sodium butyrate, a histone deacetylase inhibitor, was used during histone purification from BF trypanosomes, but not from PF, which may account for the observed variation. I also refrained from drawing comparisons between BF and PF of the H4 N-terminus because it was examined in more detail in this study: the less abundant methyl mark found in BF (K2) was only uncovered following HPLC fractionation of H4 Glu-C and PA-trypsin digests. A more direct comparison of histone modifications in BF and PF would require isotopic labeling and mixing experiments, so that histones from both stages may be co-purified and analyzed simultaneously.

While this study was in progress, a paper characterizing the PTMs on H4 from *Trypanosoma cruzi* was published [181]. The N-terminus of H4 is nearly identical in the two closely related organisms. Using MS/MS, the authors found that alanine 1 was monomethylated, lysines 4, 10, 14, and 57 were acetylated, lysine 18 is monomethylated, and arginine 53 is dimethylated. I examined our data for evidence of R53 methylation and K57 acetylation, but neither mark was found. Perhaps this reflects species-specific variation of histone PTMs. A number of marks discovered in our study were not found in *T. cruzi*, including K2 and K5 acetylation, K2 dimethylation, K17 methylation, and K18 di- and tri-methylation. As with PF histones, it is hard to say whether these differences are

real or whether the marks were overlooked by the MS methods employed by the authors.

The study of *T. cruzi* histones also looked at bulk acetylation and methylation by *in vivo* labeling of live trypanosomes with acetyl and methyl donors [^{3}H]-acetic acid and L-[methyl-^{3}H]-methionine, respectively [181]. It was shown that H2A, H4, and, to a lesser extent, H3, were labeled with [^{3}H]-acetic acid; H3, H2B, and, to a lesser extent, H4, were labeled with L-[methyl-^{3}H]-methionine. Our results are in agreement with the findings for H2A, H3, and H4; however, methylation of H2B was not observed in *T. brucei*, so perhaps this reflects species-specific variation of histone PTMs as well.

In higher eukaryotes, histones are highly conserved proteins. The extreme sequence divergence of *T. brucei* histones makes them interesting targets for phylogenetic comparisons. I demonstrated that, despite the sequence divergence, the locations of several modified residues are similar. The information compiled here provides a number of starting points for functional studies, some of which we are beginning to undertake. Future work will focus on identifying the modifications that play key roles in transcriptional regulation in trypanosomes.

Chapter 4. Investigating the Target of the Putative Deacetylase TbSIR2RP-1

Introduction

Sequencing of the *T. brucei* genome identified 7 putative histone deacetylases, including 3 SIR2-related proteins, based on sequence homology [148]. To be able to analyze the function of individual histone modifications, I first needed to identify the histone-modifying enzymes that create these marks. I chose to focus on *T. brucei* SIR2 related protein-1 (TbSIR2RP-1) for two reasons. First, *S. cerevisiae* SIR2, an H4 K16ac deacetylase, is involved in establishing and propagating heterochromatin at telomeres and subtelomeric regions [29]. Because many *VSG*s are subtelomeric and are subject to strictly controlled transcriptional silencing, it seemed reasonable to think that TbSIR2RP-1 might play a role in regulating VSG expression, although a recent study showed that disruption of *TbSIR2RP-1* does not cause de-repression of silent VSGs [170]. Slight de-repression at non-ES subtelomeric RNA pol I-driven reporter cassettes was observed. Secondly, it was shown that cells over-expressing TbSIR2RP-1 demonstrated increased sensitivity to micrococcal nuclease following DNA damage, compared to wild-type cells, indicating that this enzyme influences chromatin structure [172]. The same study showed that TbSIR2RP-1 has *in vitro* deacetylase activity against *T. brucei* histones. I analyzed H4 from *T. brucei*

mutants lacking the SIR2-related protein 1, and conclude that TbSIR2RP-1 is most likely not an H4 deacetylase.

Results

Isolating Histones from a ΔTbsir2rp-1 Cell Line

A *ΔTbsir2rp-1* cell line was created by replacing endogenous *TbSIR2RP-1* alleles with drug resistance markers, which were targeted to the *TbSIR2RP-1* locus using endogenous UTRs (Fig. 4.1; constructs provided by M. Hoek). The *ΔTbsir2rp-1* cells were amplified in rats, collected from rat blood by centrifugation following exsanguination, and were used as starting material for the purification of histones using the protocol developed for BF 221wt cells. The elution profile of *ΔTbsir2rp-1* histones obtained during fractionation by HPLC did not demonstrate significant differences from the BF 221wt profile (Fig. 4.2). Any differences in the elution profiles were most likely due to the use of different columns for each sample. To confirm the purity of the histones, HPLC fractions were analyzed by Coomassie stain (Fig. 4.3). The identity of the histones in each HPLC fraction was confirmed by MALDI-TOF.

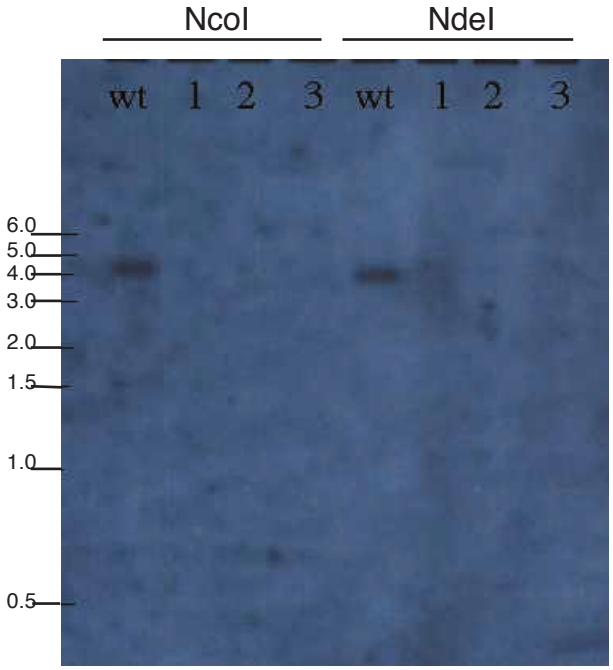


Figure 4.1. Replacement of two TbSIR2RP-1 Alleles Demonstrated by Southern blot. To confirm knockout, gDNA from BF 221 wt cells and 3 clones, MH166.6.167.1-3, were digested with NcoI and NdeI and probed with TbSIR2RP-1. Expected bands at 5.0 and 4.6 kb were observed for NcoI and NdeI digest, respectively, in wt cells, but no bands corresponding to TbSIR2RP-1 were observed for clones.

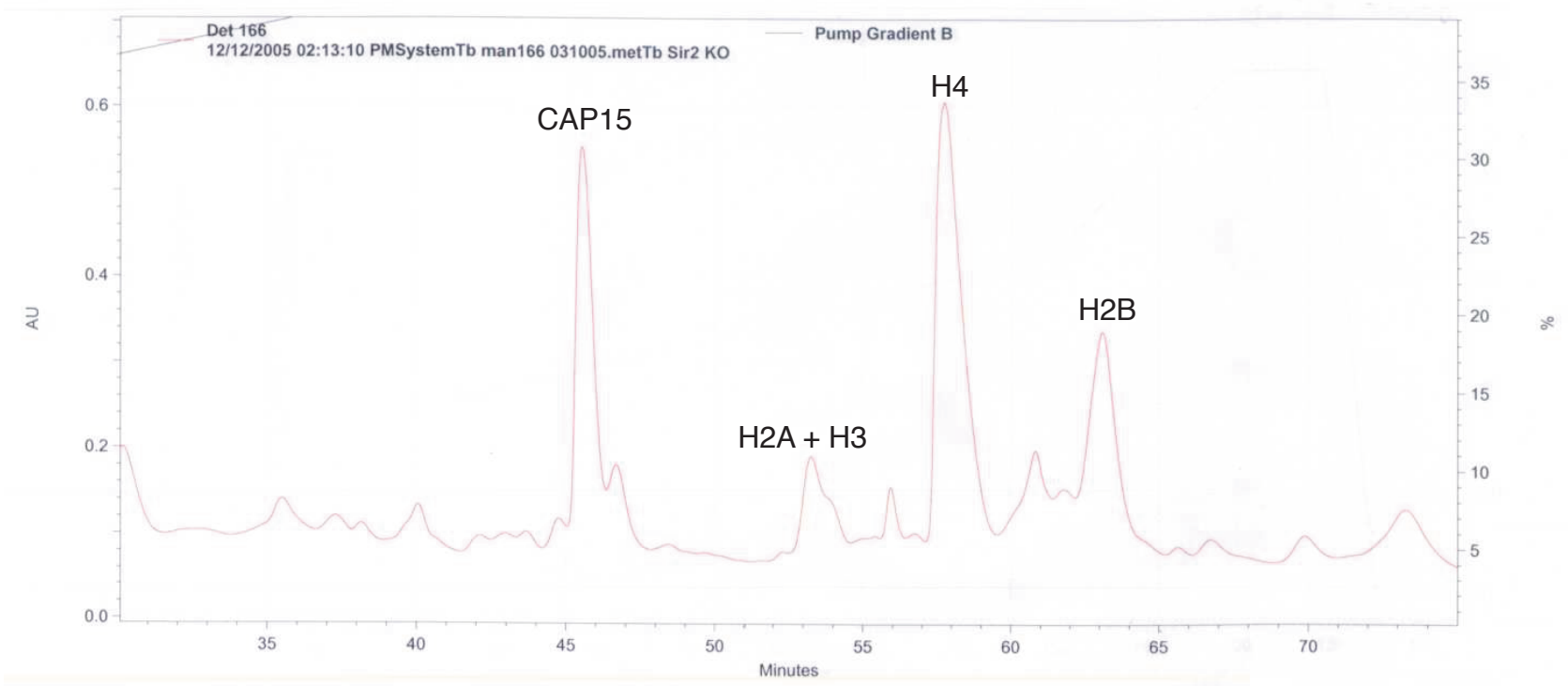


Figure 4.2. HPLC Elution Profile of Δ*Tbsir2rp-1* Knockout Histones. Crude histone preparation was fractionated by RP-HPLC. Absorbance 214 nm. Additional peaks did not contain histones. CAP15 = corset-associated protein 15.

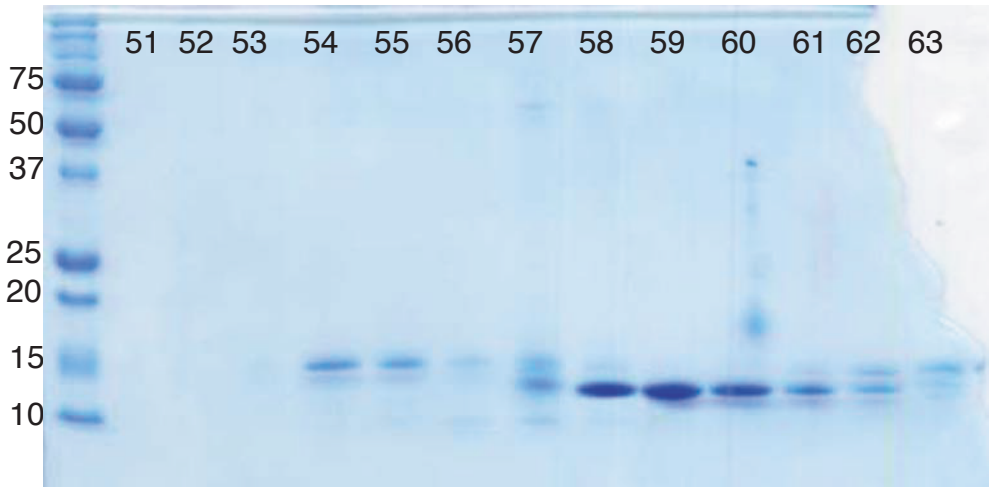


Figure 4.3. SDS-PAGE of HPLC-purified Δ*Tbsir2rp-1* histones. Fractions 54 and 55 contain H2A and H3; fractions 58-63 contain H4; fractions 62 and 63 contain H2B. Confirmed by MALDI-TOF.

Investigating the Target of TbSIR2RP-1 Deacetylase

SIR2 in yeast and other organisms is an H4 deacetylase, so *ΔTbsir2rp-1* H4 was examined first for increased levels of acetylation. *ΔTbsir2rp-1* H4 was analyzed by Edman degradation, but, surprisingly, no differences in the acetylation pattern were observed (Table 4, Fig. 4.4). No additional PTMs were identified on H4 besides those at the N-terminus, so I concluded that TbSIR2RP-1 is either not an H4 deacetylase or its loss is very efficiently compensated for by another deacetylase.

To determine whether any of the histones were the *in vivo* target of the putative TbSIR2RP-1 deacetylase, purified histones from 221wt and *ΔTbsir2rp-1* were analyzed on a TAU gel (Fig. 4.5). A TAU gel allows for the separation of proteins based both on size and charge. Therefore, phosphorylation and acetylation, both of which modify the charge of a protein, would be visible as distinct bands from the unmodified species. The more positively-charged unacetylated species migrates closer to the bottom of the gel than the acetylated species. I expected that *ΔTbsir2rp-1* histones would be more highly acetylated, which would be demonstrated by an upward shift on the gel.

H2A, H2B, and H3 were analyzed on a TAU gel with 15% polyacrylamide. H2B is represented on the gel by only one band, and there is no apparent difference in the migration pattern of H2B from wt versus *ΔTbsir2rp-1* cells. The band most likely represents unmodified H2B because the three modified lysines

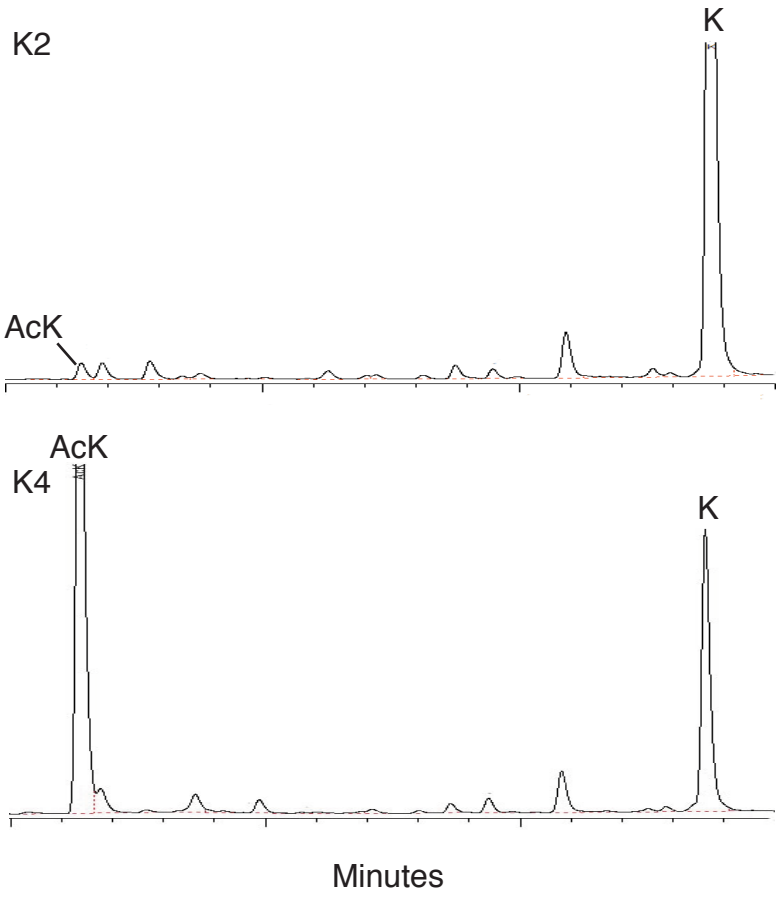


Figure 4.4. Edman degradation of histone H4 from Δ*Tbsir2rp-1* cells shows that TbSIR2RP-1 is not an H4 deacetylase. The PTH HPLC trace of cycles 2, 4, 5, and 10 show that 2% of K2, 73% of K4, 5% of K5, and 7% of K10 are acetylated. These values do not represent a significant change from the acetylation levels of H4 from wt cells, suggesting that TbSIR2RP-1 does not deacetylate the H4 N-terminal lysines. Lysine and acetyl lysine peaks were verified by comparison to known standards.

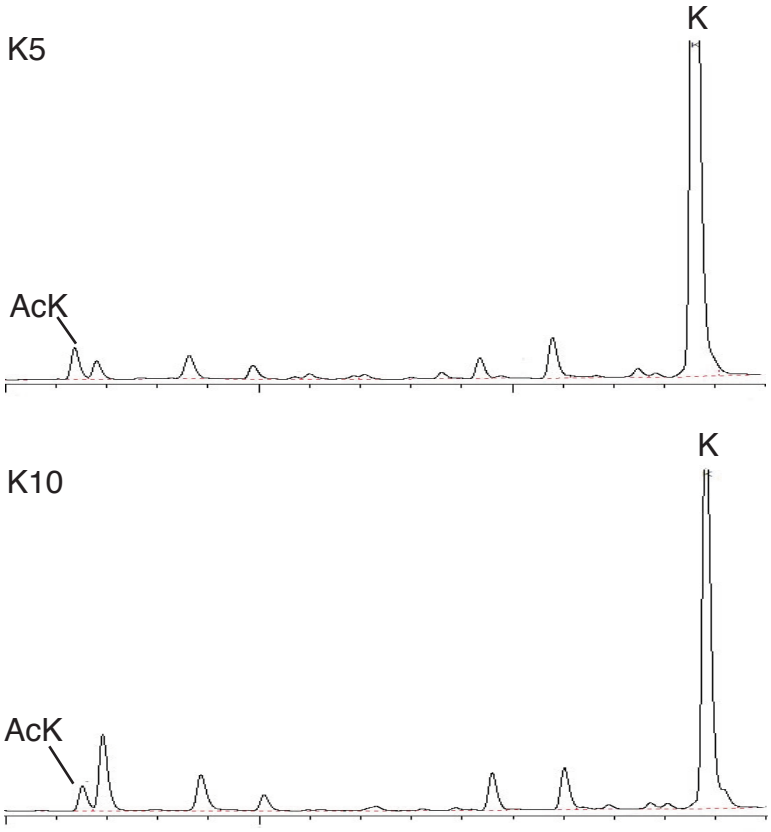

K5
K
AcK
K10
K
AcK

TABLE 4. Quantification of posttranslational modifications at the H4 N-terminus from wild-type and *ΔTbsir2rp-1* cells[a]		
AA Position	WT[b] (% modified)	*ΔTbsir2rp-1*[c] (% modified)
A1me1	60	60
K2ac	2	2
K4ac	73	73
K5ac	7	5
K10ac	7	7
K17me1, 2, 3	4, 8, 9	--
K18me1, 2, 3	6, 4, 18	--

[a] H4 purified from wild-type and Δ*Tbsir2rp-1* cells was sequenced by Edman degradation. The levels of acetylation and methylation were quantified by comparison to known standards.

[b] Data for wild-type H4 are the average of three trials.

[c] Δ*Tbsir2rp-1* H4 was sequenced once using conditions that were not optimal for the quantification of methyl lysines, so these values are not listed.

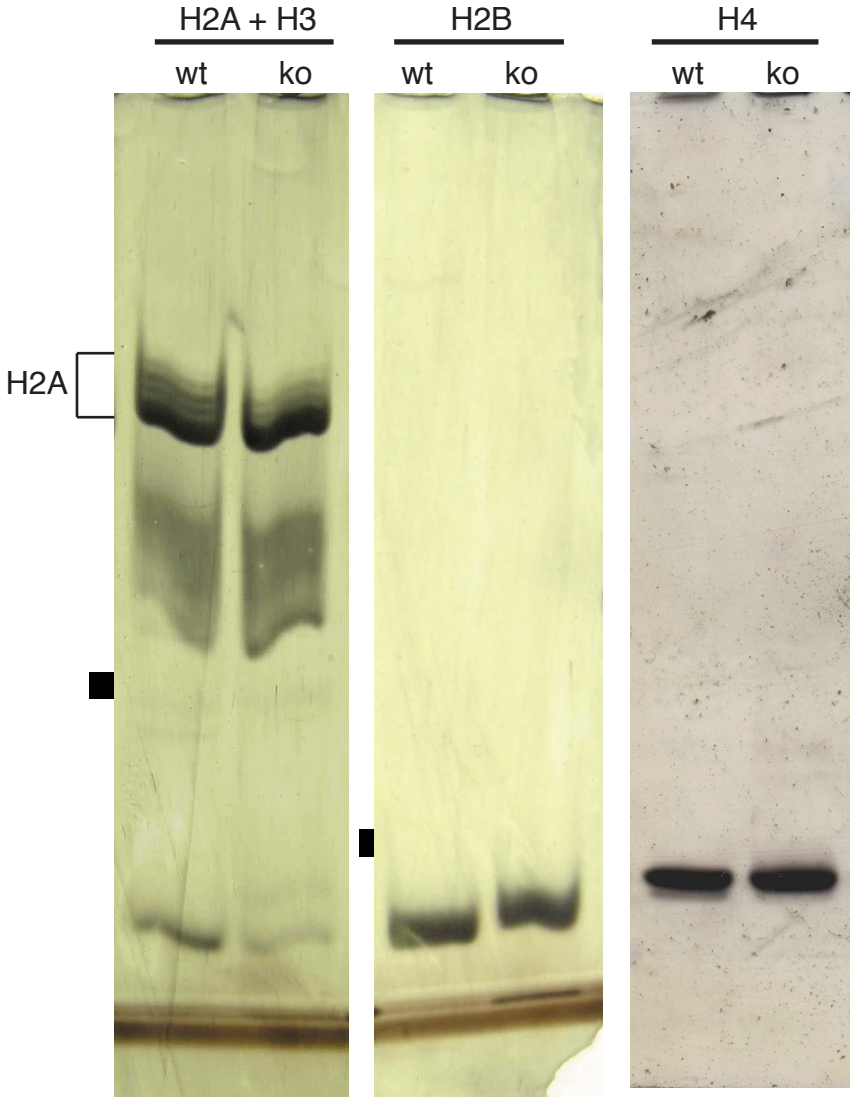


Figure 4.5. Histones from BF 221wt cells (wt) and BF Δ*Tbsir2rp-1* cells (ko) on TAU gel (H2A/H3 and H2B on 15% PAGE; H4 on 20% PAGE). No obvious differences were observed between the migration pattern of 221wt and Δ*Tbsir2rp-1* histones.

(K4, K12, and K16) that I previously identified were acetylated at very low levels (Table 3).

H2A and H3 are present in the same lane because they could not be separated from each other by HPLC. Multiple bands at the top of this lane most likely represent H2A, which has 6 acetylated lysines at the C-terminus that can be found in different combinations. Upward mobility shifts of two subtle bands, which cannot be seen on this reproduction, between wt and *ΔTbsir2rp-1* H2A and H3 suggest that TbSIR2RP-1 may act as a deacetylase for one or both of these histones. The likely candidate is H2A because it has multiple acetylated lysines, while H3 has only one.

H4 was analyzed on a TAU gel with 20% polyacrylamide because separation could not be achieved on a 15% polyacrylamide gel. Two bands are present for both wt and *ΔTbsir2rp-1* H4. The darker band most likely represents the singly acetylated species of H4. H4 analysis by Edman degradation showed that 73% of K4 is acetylated, so the singly acetylated H4 is the major species. The lighter band, which migrates directly beneath the darker one, most likely represents the minor, unacetylated species. No differences in the two bands were observed between wt and *ΔTbsir2rp-1* H4, which confirms that TbSIR2RP-1 is not an H4 deacetylase.

Discussion

To begin characterizing histone modifying enzymes in trypanosomes, I chose to study the SIR2 homologue, TbSIR2RP-1. H4 from a *ΔTbsir2rp-1* cell line was analyzed by Edman degradation and had the same acetylation levels as wildtype H4, indicating that TbSIR2RP-1 is not an H4 deacetylase. There was a possibility that two additional SIR2-related proteins in *T. brucei*, TbSIR2RP-2 and 3, might perform functions redundant to TbSIR2RP-1, masking the effects of knocking out the putative deacetylase. However, it was shown that TbSIR2RP-2 and 3 localize to the mitochondrion, while only TbSIR2RP-1 is found in the nucleus [170]. Therefore, I think it is unlikely that TbSIR2RP-2 and 3 perform functions that are redundant to TbSIR2RP-1. I concluded that TbSIR2RP-1 is not an H4 deacetylase.

An earlier study of TbSIR2RP-1 showed *in vitro* that the purified enzyme acts as a deactylase of trypanosome histones [172]. Purified histones labeled with a radioactive acetyl donor, [^{3}H] acetyl-CoA, by yeast H4 acetyltransferase, HAT1, were incubated with TbSIR2RP-1. The release of [^{3}H]acetyl groups was observed for each of the four major histones. I did not think our results contradict this work because it is well-known that enzymes demonstrate different, and perhaps broader, specificity *in vitro* than they do *in vivo*. This study shows that TbSIR2RP-1 has deacetylase activity, but its target has yet to be identified.

Comparing purified histones from BF 221wt and *ΔTbsir2rp-1* cells on a TAU gel demonstrated minor differences in H2A and H3, but not H4 or H2B. The two upward band shifts that were observed in *ΔTbsir2rp-1* H2A and H3 were consistent with the presence of more highly acetylated species. However, more work is needed to verify that either H2A or H3 is the target of TbSIR2RP-1. For example, purified H2A from wildtype and *ΔTbsir2rp-1* cells may be labeled with propionic acid or deuterated propionic acid and mixed. MALDI-TOF analysis of the trypsinized sample will show C-terminal H2A peptide 107-133 in its various acetylated states. The peaks corresponding to either wildtype or *ΔTbsir2rp-1* H2A will be evident from their masses, and the intensity of the peaks may be compared to determine whether acetylated species are less abundant in *ΔTbsir2rp-1* H2A. This would illustrate whether TbSIR2RP-1 is an H2A deacetylase.

Finally, although this study was initiated with the idea that TbSIR2RP-1 might be involved in the regulation of VSG expression, two separate studies have shown that TbSIR2RP-1 does not affect antigenic variation [170,171]. In the first study, both alleles of *TbSIR2RP-1* were disrupted in a cell line containing a luciferase marker downstream of an inactive ES promoter [171]. There was no increase in luciferase expression in the *ΔTbsir2rp-1* cell line, indicating that TbSIR2RP-1 is not involved in silencing at the inactive ES in BF. It is thought that different mechanisms control silencing of the ES in BF and PF [149]. The authors differentiated wildtype and *ΔTbsir2rp-1* cells to PF, but again found no

difference in luciferase expression at the ES between the two cell lines. The second study similarly used a reporter construct in both wildtype and *ΔTbsir2rp-1* cells to show that TbSIR2RP-1 is not involved in ES silencing [170]. However, using a reporter construct downstream of a repressed promoter 2 kb from a telomere, the authors show that reporter expression was increased following disruption of both *TbSIR2RP-1* alleles, indicating that TbSIR2RP-1 does play a role in telomeric silencing. It seems that, although TbSIR2RP-1 is not an H4 deacetylase, it remains a player in the establishment of silent chromatin at the telomere.

Chapter 5. Histone variant H2Bv replaces H2B in nucleosomes enriched for H3 lysine 4 and 76 trimethylation

Introduction

In other organisms, histone PTMs have been studied for their roles in a variety of biological processes. Causal relationships between the various histone modifications led to the 'trans histone' cross-talk hypothesis, which in part states that PTMs on one histone may prohibit or be required for PTMs on another histone [61]. In one famous instance of cross-talk between histones, ubiquitination of histone H2B lysine 123 in *S. cerevisiae* was shown to be required for H3 lysine 4 and 79 methylation [62,63]. Furthermore, these modifications are linked to transcription elongation, as it has been shown that the elongating RNA polymerase II and its binding partners form a platform for the recruitment of H3 K4 and K79 methyltransferases Set1 and Dot1, respectively, following H2B ubiquitination (reviewed in [65]).

The addition of PTMs to a histone is one mechanism for providing diversity to a nucleosome and allowing for a variety of biological readouts. A second mechanism for providing diversity is the substitution of histone variants for canonical histones in the nucleosome [71]. Histone variants have previously been identified for each of the four major histones in trypanosomes [173,174]. Preliminary work on *T. brucei* histone variant H2Bv showed that *H2Bv* is a single

copy gene that is essential for viability [174]. Also, histone variant H2AZ co-immunoprecipitates with H2Bv, demonstrating that the two variants form a dimer within a nucleosome. I continued the characterization of H2Bv by identifying PTMs present on the histone variant.

After identifying many of the histone PTMs in *T. brucei*, I began to examine the functional relationships between individual marks using sequence alignments as a guide. Here, I examined the sequence homology between *T. brucei* H2B, H2Bv, and H2B from other organisms and found that H2Bv has a homologue (lysine 129) to *S. cerevisiae* H2B K123, but *T. brucei* H2B does not (Fig. 5.3). Based on the alignment, I hypothesized that H2Bv K129 is ubiquitinated and required for H3 K4 and K76 methylation. In order to assay for H3 K4 methylation, I raised a specific antibody against an H3K4me3-containing peptide and began to characterize this mark.

Although I found that H2Bv K129 is not ubiquitinated, I was able to show that H2Bv specifically associates with H3 that is trimethylated at K4 and K76. These data suggest that H3 K4 and K76 methylation in trypanosomes is regulated by a novel mechanism, possibly involving the replacement of major H2B with H2BV in the nucleosome.

Results

Anti-H3K4me0 and H3K4me3 Antibodies

Although I was unable to identify many PTMs at the *T. brucei* H3 N-terminus in Chapter 3, a sequence alignment of the H3 N-termini suggested that *T. brucei* H3 K4, a possible homologue of the well-studied methyl H3 K4 in other organisms, was methylated (Fig. 3.28D). Therefore, I raised antibodies against H3 peptide 2-9, RTKETART, that was unmethylated or trimethylated at K4. Purification of the anti-H3K4me3 antibody was complicated by its overlapping specificity for H3K4me3- and H3K76me3-containing peptides. To purify anti-H3K4me3 antibody with diminished H3K76me3 binding affinity, H3K4me3 antiserum was twice depleted on a column coupled to an H3K76me3 peptide. H3K4me0 and H3K4me3 antisera were then affinity purified on columns coupled to their corresponding peptides.

The specificity of the two antibodies was confirmed by Western blot analysis (Fig. 5.1). First, a blot containing purified histones, the identity of which were verified by MS, was incubated with the anti-H3K4me3 antibody to show that the antibody only recognized H3, but not H2A, H2B, or H4 (Fig. 5.1A). Second, the two antibodies were subject to peptide competition by pre-incubating them with no peptide, H3K4me0-, H3K4me3-, H3K76me0-, and H3K76me3-containing peptides (Fig. 5.1B). A blot containing whole cell lysate from PF 427wt cells was

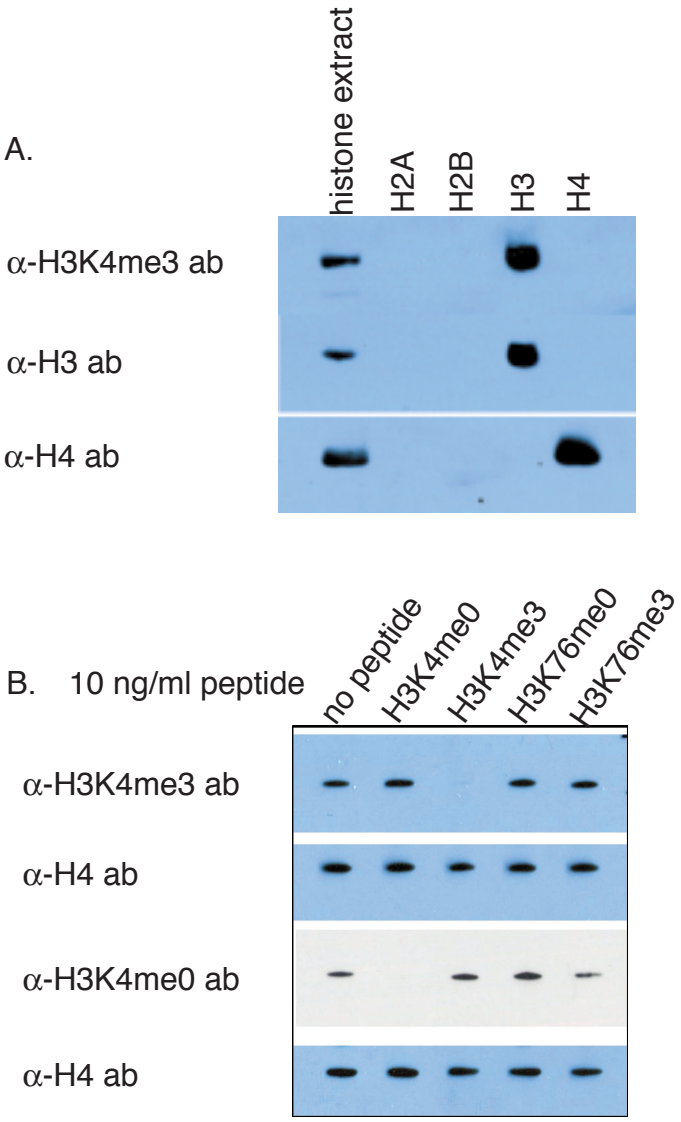


Figure 5.1. Specificity of anti-Histone H3 lysine 4 unmethylated and trimethyl antibodies. (A) Crude histone extract and purified histones (verified by MS) were loaded on a Western blot and probed with anti-H3K4me3, anti-H3, and anti-H4 antibodies. The blot shows that the anti-H3K4me3 antibody only recognizes H3, and not other histones. (B) Anti-H3K4me3 and me0 antibodies were pre-incubated with no peptide, H3K4me0, H3K4me3, H3K76me0, and H3K76me3 peptides prior to incubation on a Western blot. Both antibodies bind only their own peptide, which demonstrates their specificity. An anti-H4 antibody is used as a loading control.

incubated with the antibodies, which showed that both antibodies bind only their corresponding peptide.

Localization of H3 with Unmethylated and Trimethylated Lysine 4

Immunofluorescence using the anti-H3K4me0 and H3K4me3 antibodies in wildtype BF and PF cells illustrated the localization of these marks. In BF cells, both H3K4me0 and H3K4me3 were distributed throughout the nucleus, with the exception of the nucleolus (Fig. 5.2A). Staining with the anti-H3K4me3 antibody produced a punctate staining pattern in BF cells, indicating that there are some regions within the nucleus where trimethyl H3 K4 is concentrated. In PF cells, again, both H3K4me0 and H3K4me3 are distributed throughout the nucleus (Fig. 5.2B). However, both demonstrate regions of higher intensity staining, indicating that there may be foci where unmethylated and trimethyl H3 K4 are concentrated.

Sequence Homology of T. brucei Histones H2B and H2Bv

T. brucei histone variant H2Bv shares 38% sequence homology with H2B [174]. The region of greatest similarity is at the C-terminus, where both histones also share significant sequence homology with H2B from other organisms (Fig. 5.3). Interestingly, the C-terminal tails of H2B (5´ –SHASS) and H2Bv (5´ –AKYNASREEAYSKVL) differ such that H2Bv shares greater sequence homology

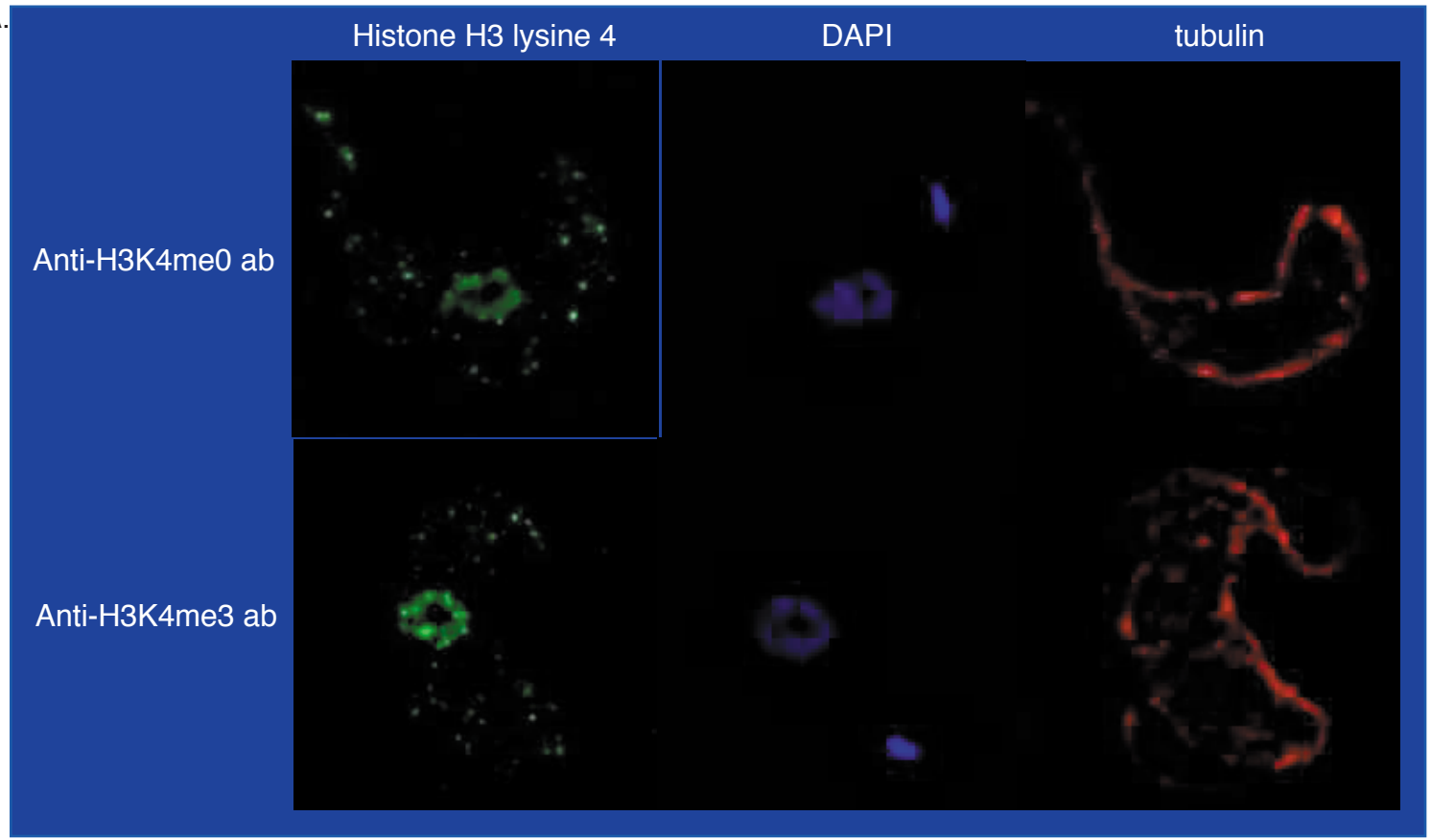


Figure 5.2. Localization of unmethylated and trimethyl histone H3 lysine 4. The cell in the top panel is stained with anti-H3K4me0 antibody, and the cell in the bottom panel is stained with anti-H3K4me3 antibody. Cells are also stained with DAPI and tubulin so that DNA and the cytoskeleton, respectively, may be visualized. (A) Wildtype bloodstream form cells are stained with the two antibodies, demonstrating that unmethylated H3 K4 is evenly distributed throughout the nucleus. Trimethyl H3 K4 may also be found throughout the nucleus, but has a more speckled appearance.

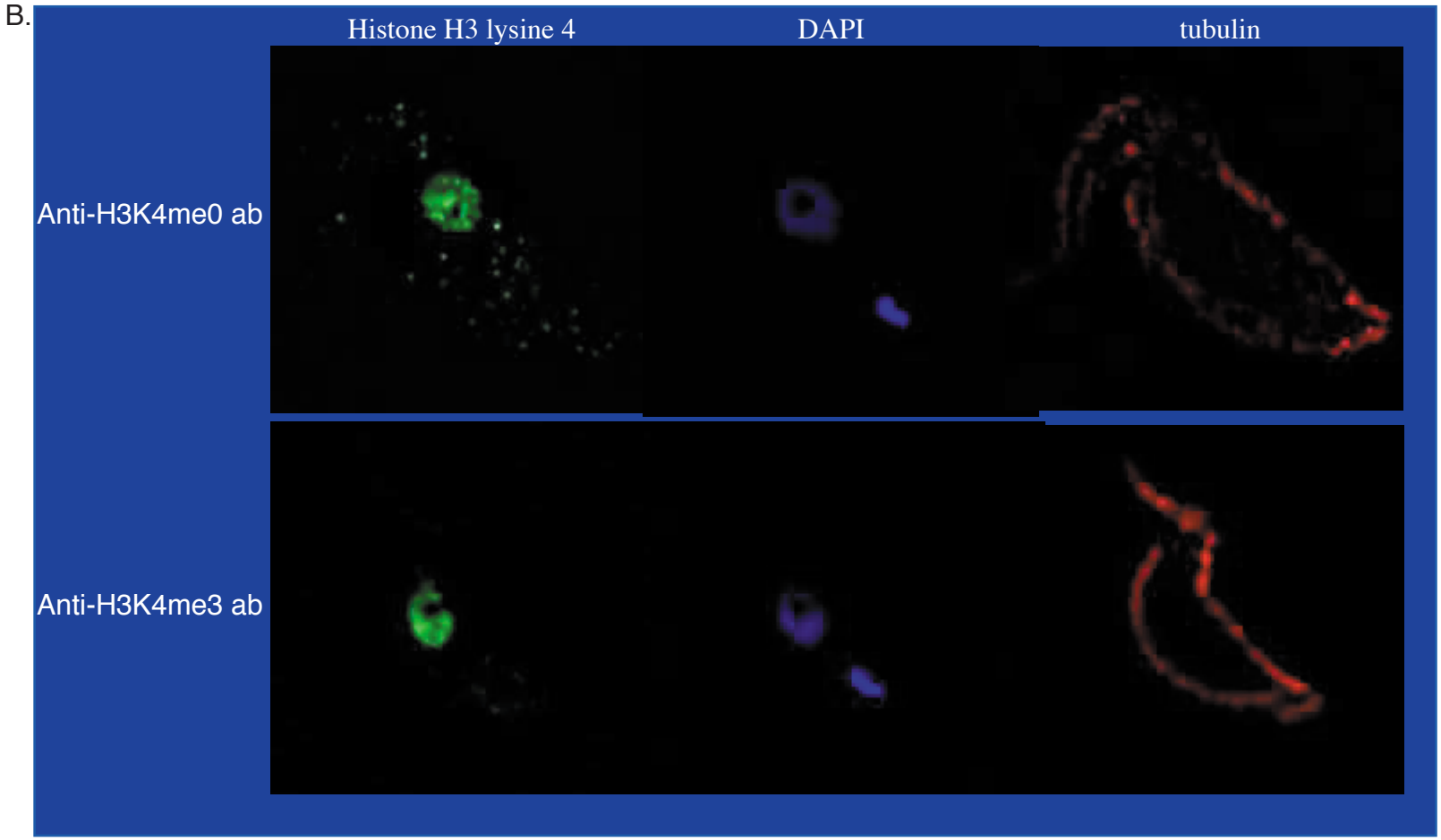


Figure 5.2B. Wildtype procyclic form cells are stained with anti-H3K4me0 and me3 antibodies. Both unmethylated and trimethyl H3 K4 are found distributed unevenly throughout the nucleus.

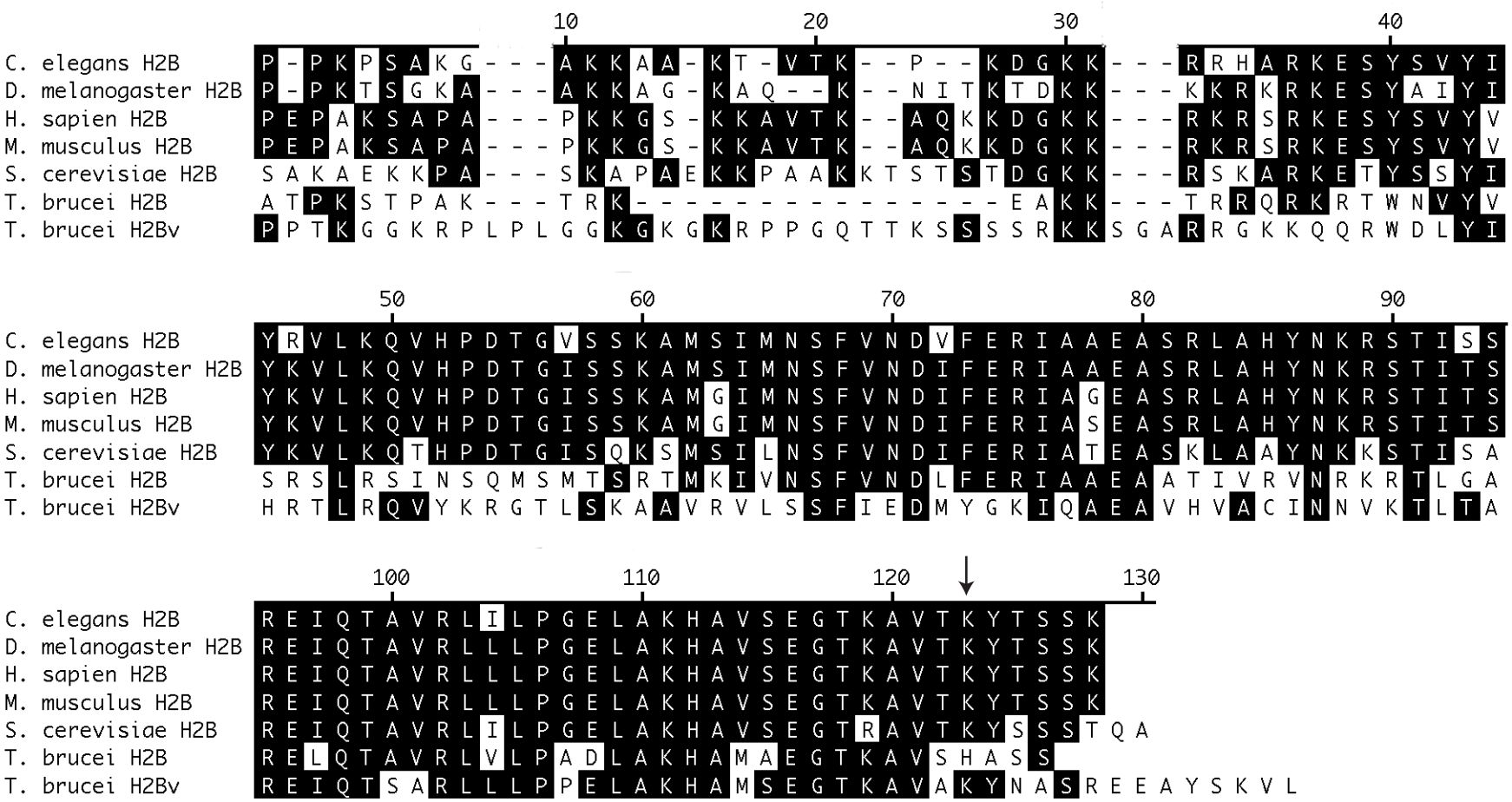


Figure 5.3. Sequence alignment of *T. brucei* histones H2B and H2Bv with H2B from other organisms. Identical residues are shaded. The sequence ruler is numbered according to the *S. cerevisiae* sequence. An arrow points to *T. brucei* H2Bv K129, which is homologous to the ubiquitinated K123 from *S. cerevisiae*.

with the H2B consensus sequence from other organisms than *T. brucei* H2B. Apart from the C-terminus, H2Bv shares limited sequence homology with the H2B consensus.

S. cerevisiae H2B lysine 123 is famously known as a site of ubiquitination, which is required for H3 K4 and K79 methylation [62,63]. The H2B sequence alignment shows that *T . brucei* H2Bv lysine 129 is homologous to *S. cerevisiae* H2B K123, but no such homologue exists in *T. brucei* H2B. Based on this alignment, I hypothesized that H2Bv lysine 129 could be ubiquitinated and required for H3 K4 and K76 methylation.

Purification of H2Bv and Identification of Posttranslational Modifications

To investigate whether H2Bv K129 is ubiquitinated, H2Bv-FLAG was first affinity purified from $1.0x10^{10}$ BFJEL18 cells using the purification protocol described in Chapter 3. Following acid extraction, the sample was incubated on anti-FLAG affinity resin and eluted by boiling the resin in 2x SDS-PAGE loading buffer. The entire sample was then analyzed by Coomassie stain (Fig. 5.4). Several bands were present, all of which were identified by MS, including H2Bv-FLAG (16.7 kDa), the heavy (50 kDa) and light (25 kDa) chains of the immobilized M2 antibodies that dissociated from the gel during the elution process, and histones (11-15 kDa) that co-immunoprecipitated with H2Bv-FLAG. In order to look for ubiquitinated (25.1 kDa) or otherwise modified H2Bv-FLAG,

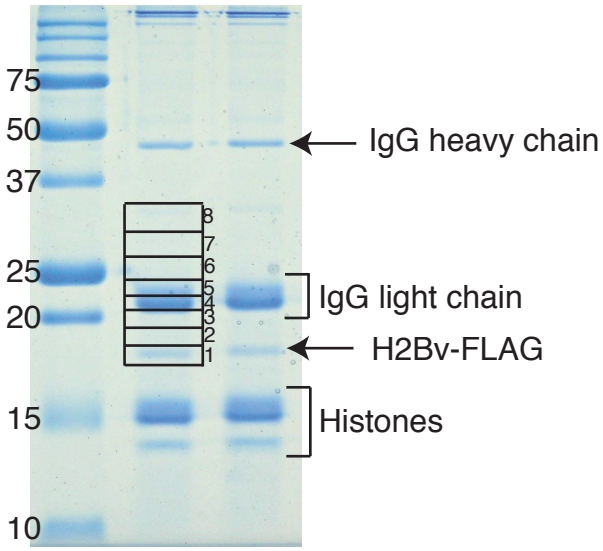


Figure 5.4. Purification of Histone H2B variant. FLAG-tagged H2Bv was affinity purified on resin conjugated to anti-FLAG M2 antibody. H2Bv was eluted by boiling the resin in SDS loading buffer. The eluate was loaded on an SDS-PAGE gel and stained with Coomassie. The identity of the bands on the gel were confirmed by MS. In addition to H2Bv-FLAG (16.7 kDa), histones and anti-FLAG ab were also eluted from the resin. To look for ubiquitinated H2Bv-FLAG (~25.1 kDa), eight gel slices were cut between the band representing H2Bv-FLAG and ~35 kDa. These gel slices were digested with trypsin and analyzed by MS/MS. Ubiquitinated H2Bv was not found, but several PTMs at the H2Bv N-terminus were identified.

which would migrate more slowly than the unmodified H2Bv-FLAG, the region of the gel between ~17 and 35 kDa was cut into 8 gel slices, trypsinized, and analyzed by tandem MS. Gel slices #1-7 contained H2Bv-FLAG tryptic peptides and provided adequate sequence coverage by MS (Fig. 5.5). Unmodified K129 was observed in H2Bv peptide 118-129, HAMSEGTKAVAK, which was produced as the result of a partial digest.

Posttranslational modifications were identified at the H2Bv N-terminus, including acetylation of lysines 7 and 19 and methylation of lysines 15 and 17 (Fig. 5.6A-G, Table 5). H2Bv peptide 8-17 was observed at m/z 525.52^{2+} and 532.63^{2+}, which represents a +28 and +42 Da mass shift from the unmodified peptide (Fig. 5.6C-D). The +28 Da mass shift was due to the presence of a dimethyl group on K15. The +42 Da mass shift was also assigned to K15, and was identified as a trimethyl group by the presence of the MH^{+}-59 ion. H2Bv peptide 8-19, observed at m/z 646.56^{2+}, showed that K17 was trimethylated because the y2-y4 ions provide a mass difference (m/z 227.17) expected for a trimethyl lysine (m/z 227.17), not an acetyl lysine (m/z 227.13)(Fig. 5.6E). No PTMs were identified on H2Bv apart from at the N-terminus. This may be due to the limitations of the trypsin digest, which excessively fragments this basic protein. However, the H2Bv results are consistent with our findings that the four major histones also have few PTMs, most of which are at the N-terminus.

Ubiquitin covalently attaches to lysine at its C-terminus. Ubiquitination can be identified by MS by looking for a GG dipeptide linkage (114 Da) to lysine,

PPTKGGKRPLPLGGKGKGKRPPGQTTKSSSSRKKSGARRGKKQQRWDLYIHR
TLRQVYKRGTLSKAAVRVLSSFIEDMYGKIQAEAVHVACINNVKTLTAREIQTSA
RLLLPPELAKHAMSEGTKAVAKYNASREEAYSKVL

Figure 5.5. Histone H2Bv sequence with regions for which tandem MS data were acquired following trypsin digestion shown in red.

TABLE 5. Posttranslational modifications are present at the N-terminus of Histone H2Bv	
Digest	Modified Peptide[a]
Trypsin	GG^{7}**K**$_{ac}$RPLPLGGK
	GG^{7}**K**$_{ac}$RPLPLGG^{15}**K**$_{me3}$GK
	RPLPLGG^{15}**K**$_{me2}$GK
	RPLPLGG^{15}**K**$_{me3}$GK
	RPLPLGG^{15}**K**$_{me3}$G^{17}**K**$_{me3}$GK
	G^{17}**K**$_{me3}$G^{19}**K**$_{ac}$RPPGQTTK
	G^{19}**K**$_{ac}$RPPGQTTK

[a] Tryptic peptides were sequenced by tandem MS, and peptides that have covalent modifications are listed here.

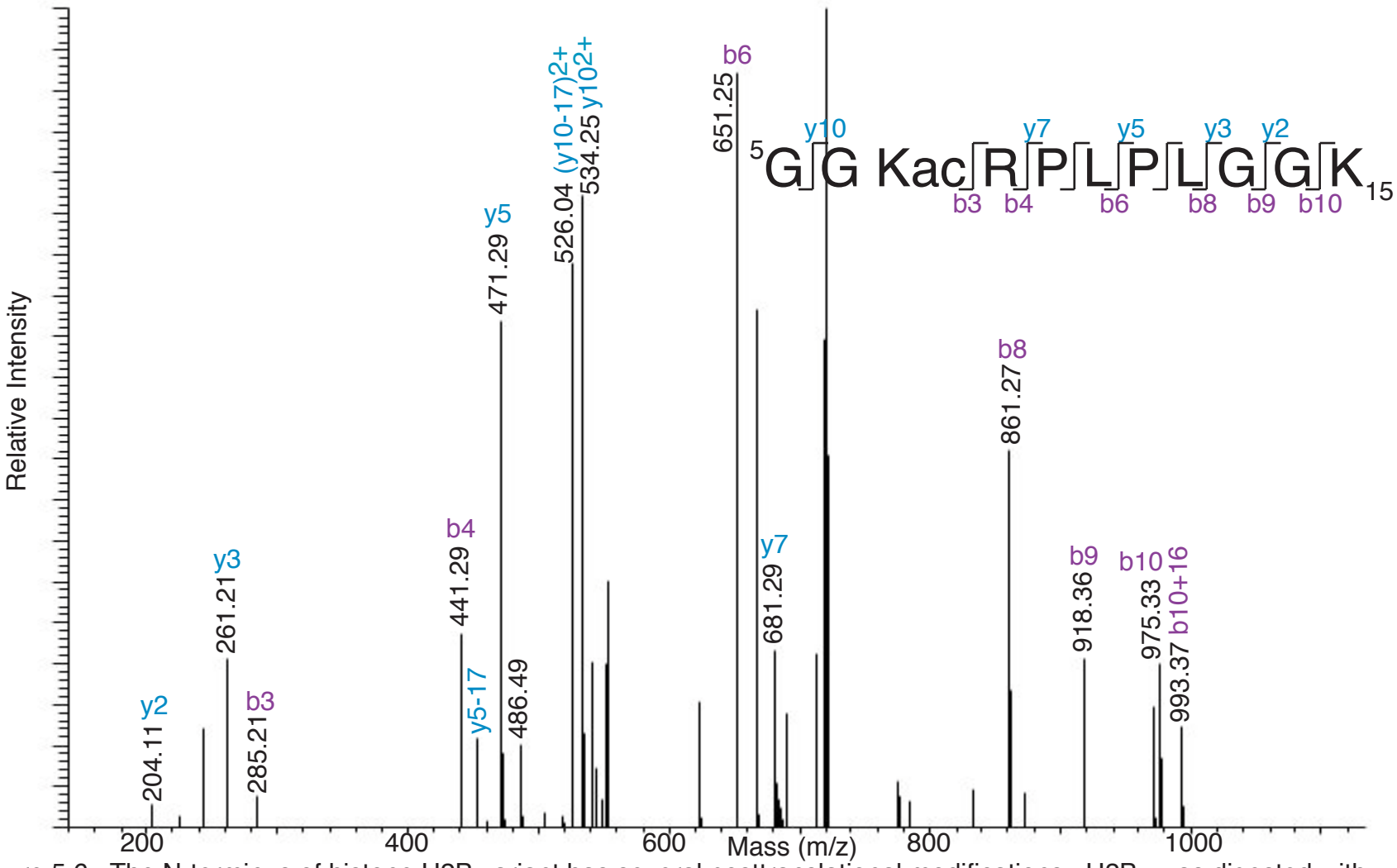


Figure 5.6. The N-terminus of histone H2B variant has several posttranslational modifications. H2Bv was digested with trypsin, and the resulting peptides were sequenced by tandem MS. (A) Sequencing of peptide 5-15 (GGKRPLPLGGK) shows that K7 is acetylated. This is demonstrated by b3 and y7-y10 ions. m/z 561.5^{2+}; theoretical MW = 1121.7 Da; measured MW = 1122.0 Da.

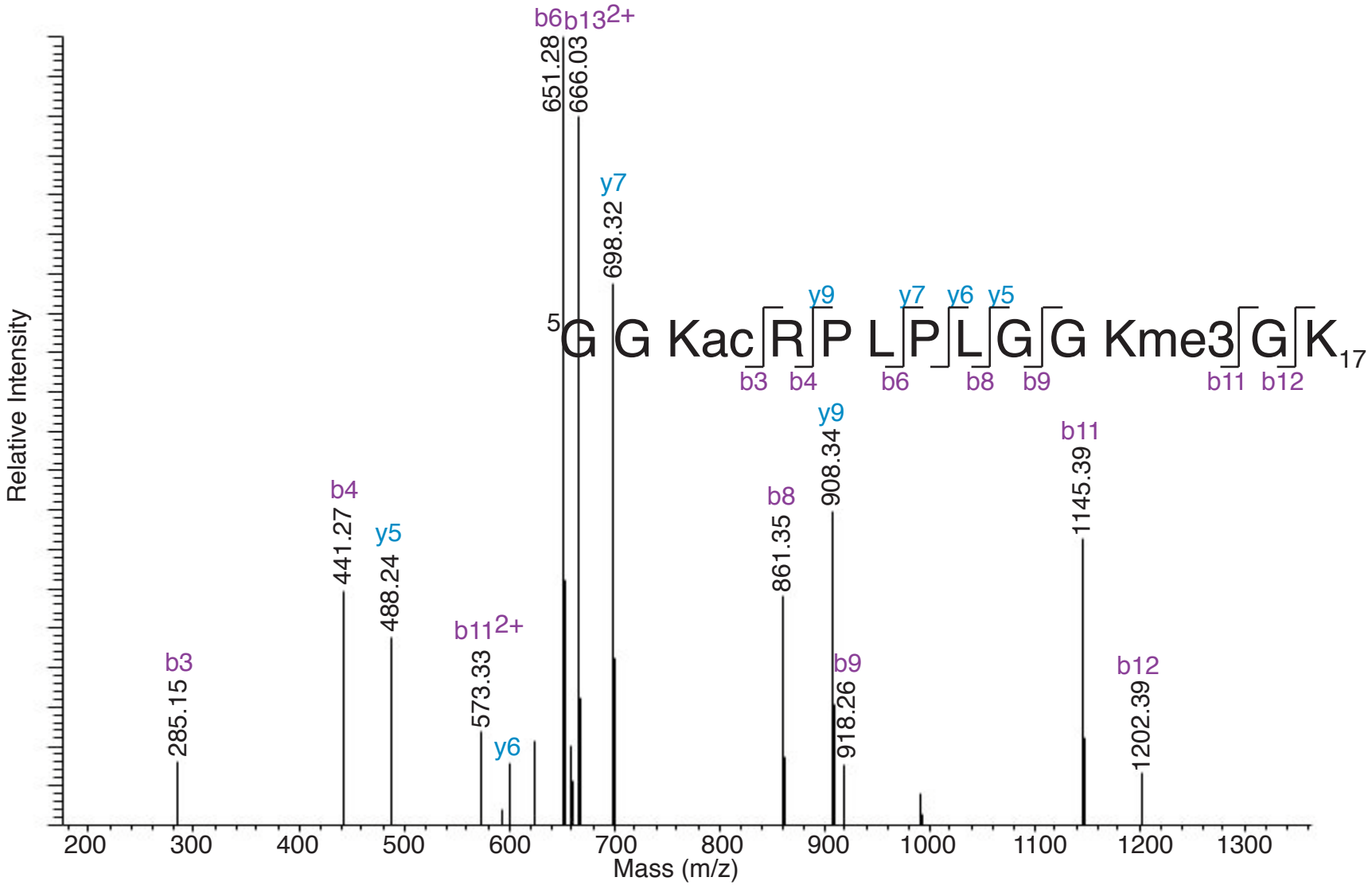


Figure 5.6B. Sequencing of H2Bv tryptic peptide 5-17 (GGKRPLPLGGKGK) shows that K7 is acetylated and K15 is trimethylated. m/z 675.4^{2+}; theoretical MW = 1348.8 Da; measured MW = 1349.8 Da.

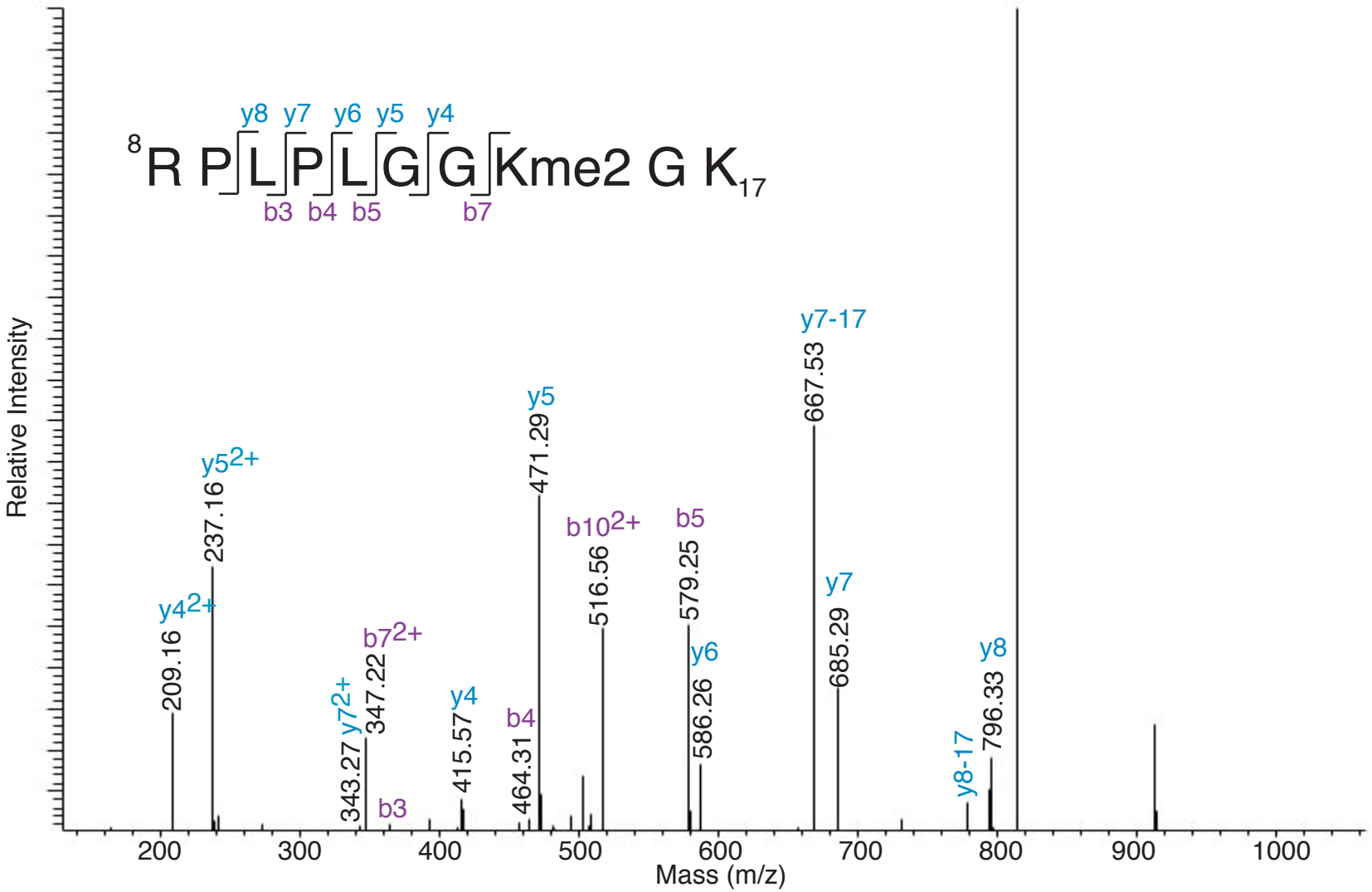


Figure 5.6C. Sequencing of H2Bv tryptic peptide 8-17 (RPLPLGGKGK) shows that K15 is dimethylated. m/z 525.52^{2+}; theoretical MW = 1698.0 Da; measured MW = 1698.6 Da.

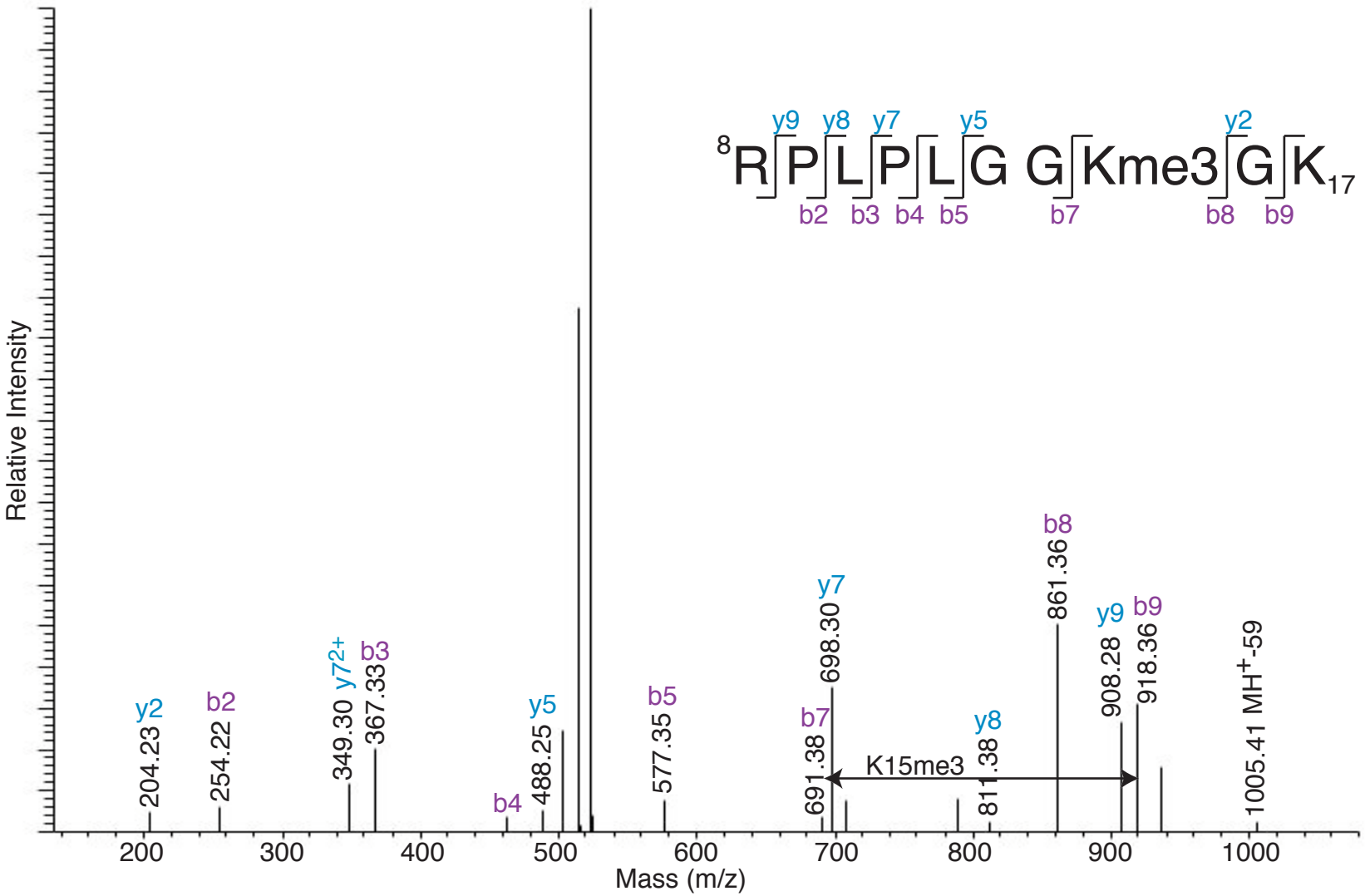


Figure 5.6D. Sequencing of H2Bv tryptic peptide 8-17 (RPLPLGGKGK) shows that K15 is trimethylated, which is demonstrated by ion pair b7-b8. The presence of the MH^+-59 ion supports the assignment of trimethylation. m/z 532.6^{2+}; theoretical MW = 1064.6 Da; measured MW = 1064.2 Da.

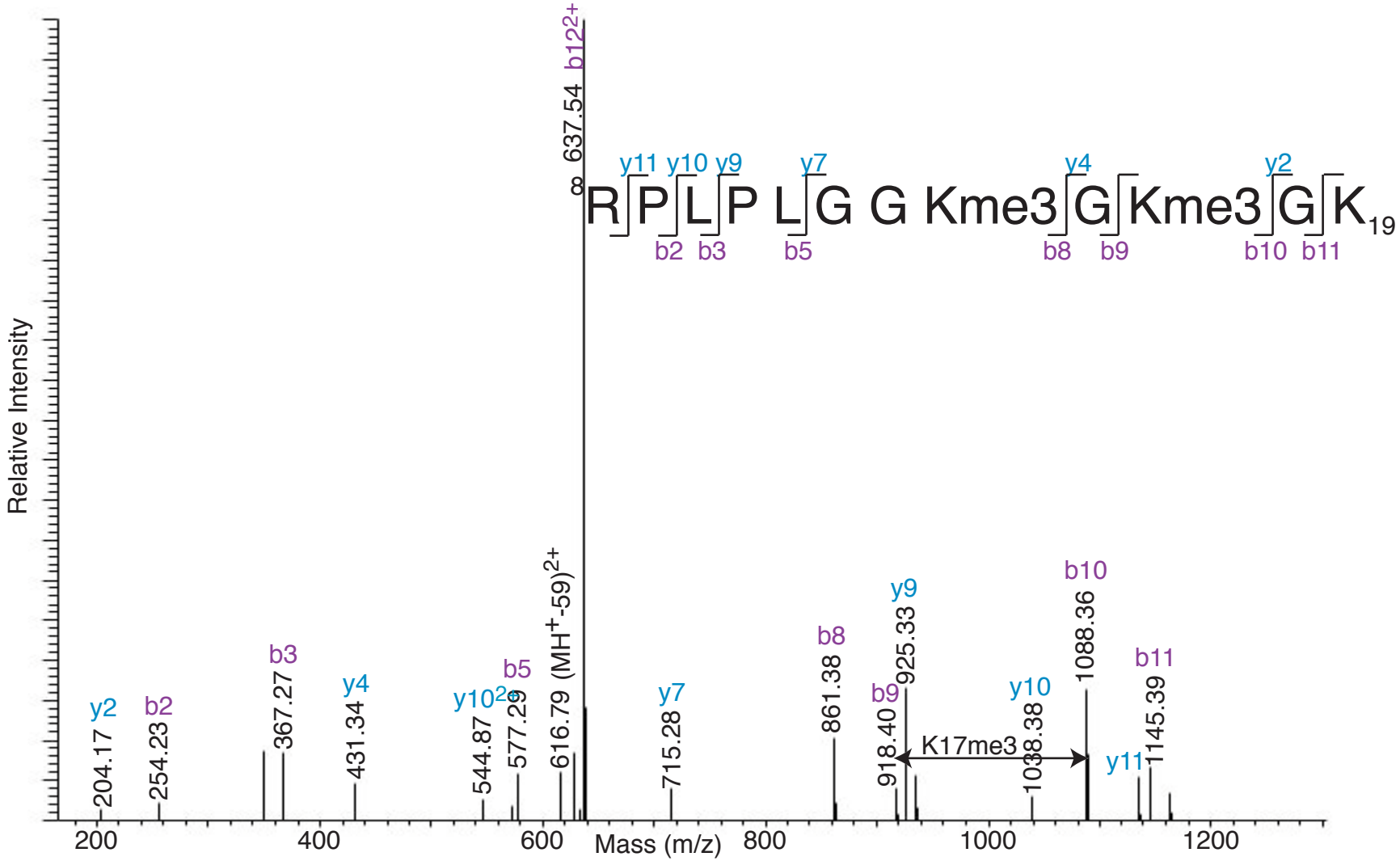


Figure 5.6E. Sequencing of H2Bv tryptic peptide 8-19 (RPLPLGGKGKGK) showed that both K15 and K17 are trimethylated. K15 trimethylation is demonstrated by ion pairs y4-y7 and b5-b8, and K17 trimethylation is demonstrated by ion pairs y2-y4 and b9-b10. The presence of the MH+-59 supports the assignment of trimethylation. m/z 646.6^{2+}; theoretical MW = 1291.8 Da; measured MW = 1292.2 Da.

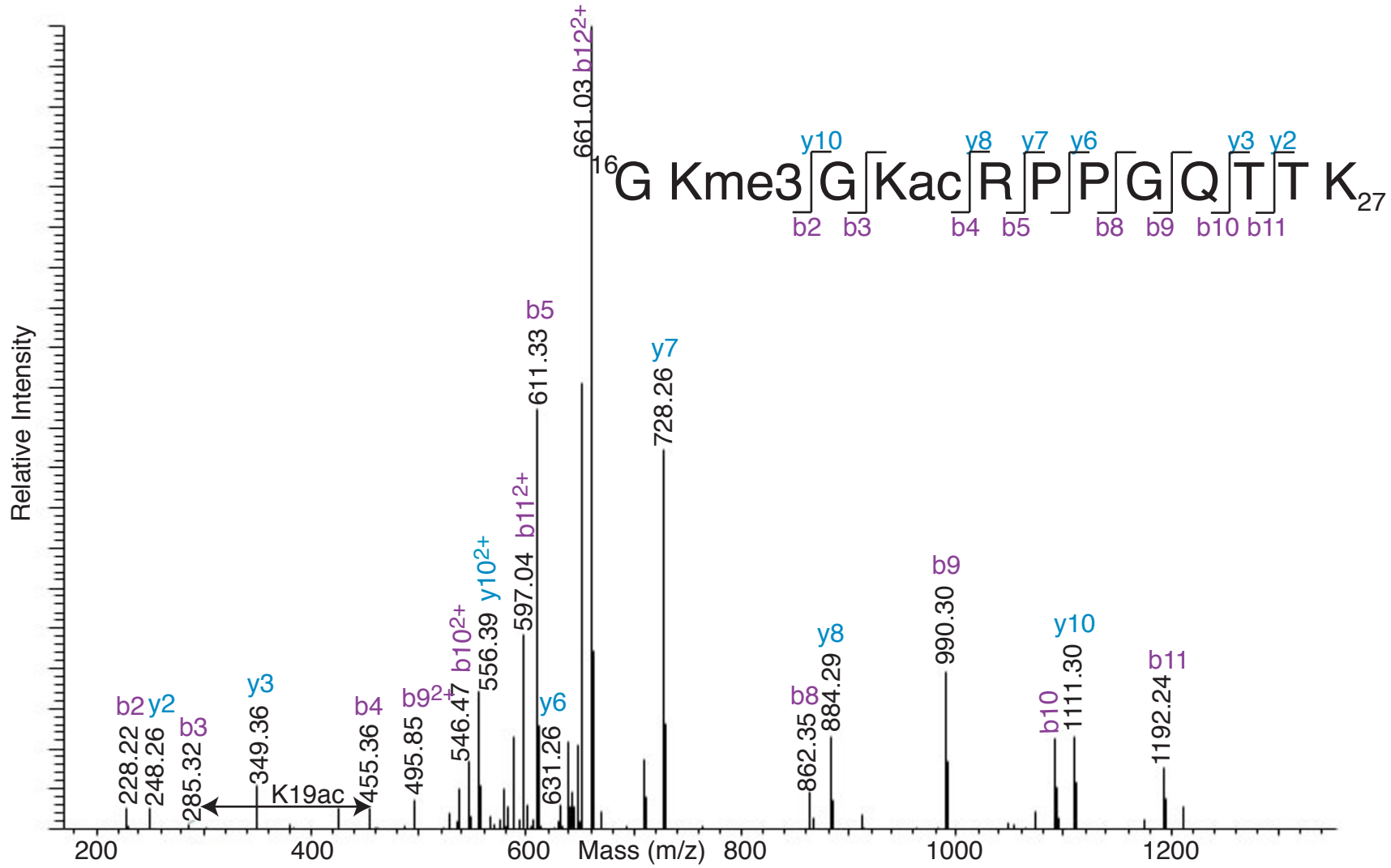


Figure 5.6F. Sequencing of H2Bv trytic peptide 16-27 (GKGKRPPGQTTK) shows that K17 is trimethylated and K19 is acetylated. The b2 ion demonstrates that K17 is trimethylated, and the b3-b4 and y8-y10 ion pairs show that K19 is acetylated. m/z 669.9^{2+}; theoretical MW = 1338.7 Da; measured MW = 1338.8 Da.

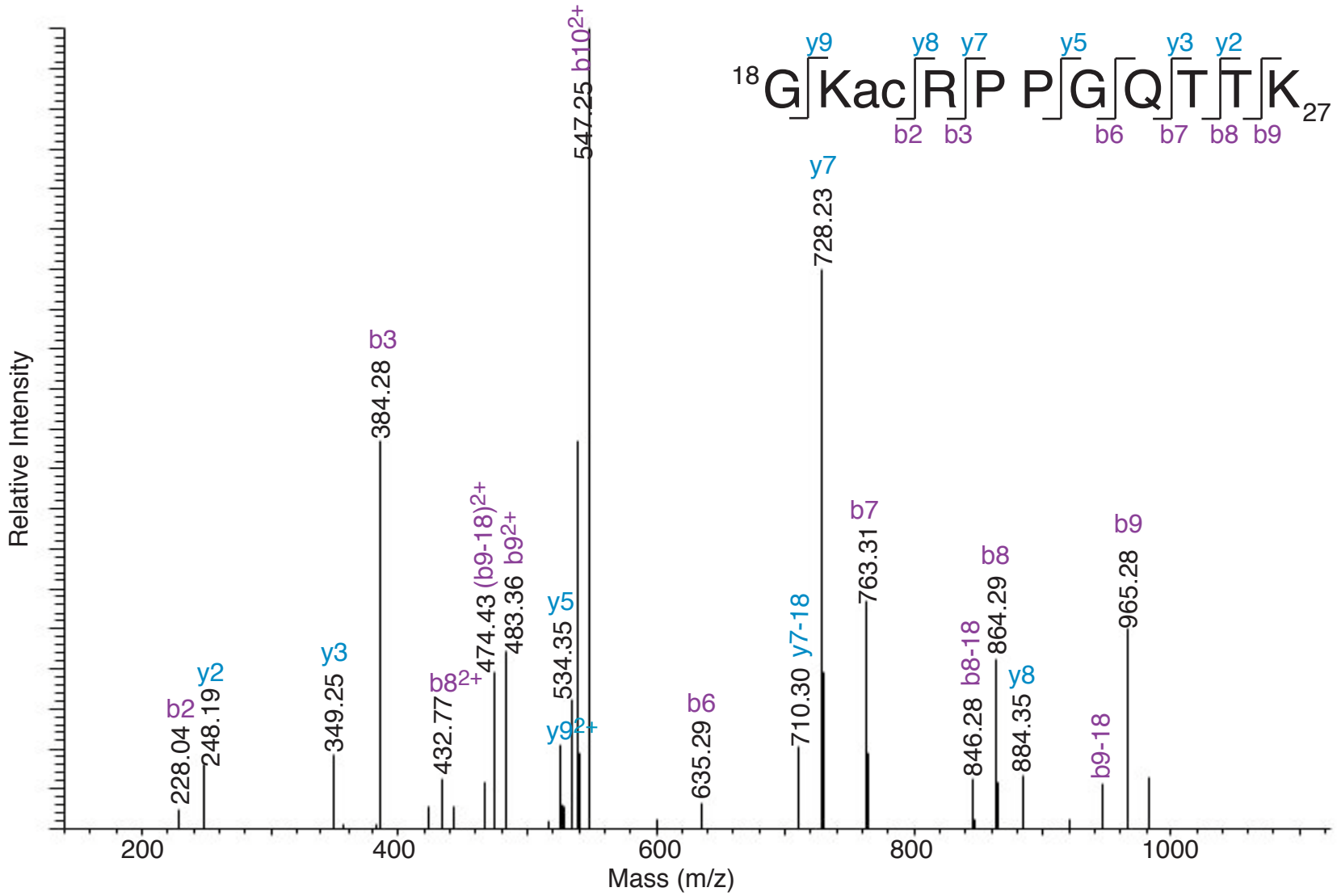


Figure 5.6G. H2Bv tryptic peptide 18-27 (GKRPPGQTTK) was sequenced, and b2 and y8-y9 ions demonstrate that K19 is acetylated. m/z 556.5^{2+}; theoretical MW = 1111.6 Da; measured MW = 1112.0 Da.

which results from tryptic cleavage of ubiquitin. Although a targeted search for the ubiquitinated H2Bv peptide 126-134 (AVAKYNASR) containing lysine 129 was performed on all 8 gel slices, no evidence of H2Bv ubiquitination was found.

There were two factors preventing us from concluding that H2Bv was not ubiquitinated. First, a single ubiquitin tryptic peptide was identified in gel slices #6 and 7, which ranged from ~25 to 32 kDa. Second, H2Bv tryptic peptides were found in gel slices #6 and 7 as well. This suggests that ubiquitinated H2Bv could be present, but was not identified by these methods. Also, it was possible that the IgG light chain (25 kDa) was co-migrating with ubiquitinated H2Bv (25.1 kDa). IgG was a darker band than the unmodified H2Bv and would certainly suppress the signal of a minor modified H2Bv species. Therefore, an independent method would be required to verify that H2Bv was not ubiquitinated.

Western Blot Analysis of Purified H2Bv-FLAG

To further test whether H2Bv-FLAG was or was not ubiquitinated, purified material was analyzed on a Western blot using an anti-ubiquitin antibody (Fig. 5.7). To avoid contaminating the sample with IgG from the anti-FLAG M2 affinity resin, H2Bv-FLAG was eluted with FLAG peptide instead of boiling the resin in SDS-PAGE loading buffer. This resulted in a much lower yield because only ~10% of the bound material eluted with FLAG peptide. Immunoprecipitated H2Bv-FLAG from $2.0x10^8$ cells was analyzed on a blot alongside purified

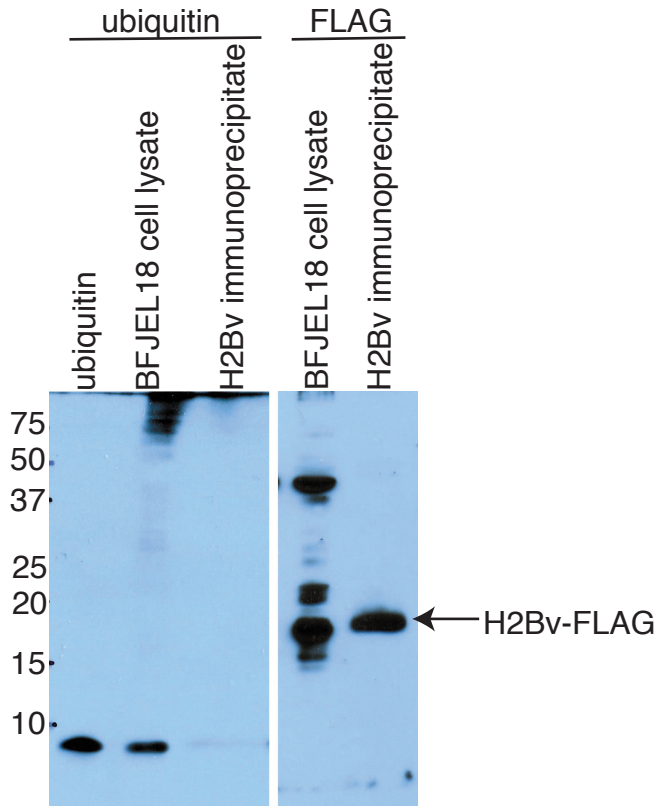


Figure 5.7. Histone H2B variant is not ubiquitinated. Purified ubiquitin (Sigma), cell lysate from a cell line containing FLAG-tagged H2Bv, and immunoprecipitated H2Bv-FLAG are analyzed on a Western blot. Incubating the blot with anti-ubiquitin antibody (P4D1 from Santa Cruz Biotechnology) demonstrates that this antibody recognizes trypanosome ubiquitin (8.4 kDa) in the cell lysate lane. The lane containing immunoprecipitated H2Bv-FLAG from $2.0x10^8$ cells does not have ubiquitinated proteins. The cell lysate and immunoprecipitated H2Bv-FLAG from $2.0x10^7$ cells are also probed with anti-FLAG antibody to show that FLAG-tagged H2Bv (16.7 kDa) is present in the immunoprecipitate lane.

ubiquitin and cell lysate from cells containing FLAG-tagged H2Bv. The blot shows no evidence of ubiquitinated proteins in the H2Bv-FLAG purification. The signal below 10 kDa in the cell lysate lane is presumably ubiquitin itself, which demonstrates that the antibody recognizes trypanosome ubiquitin. A smear of bands in the higher molecular weight range suggests that there are many ubiquitinated proteins in *T. brucei*. This becomes more evident when more cell lysate is loaded or the exposure time is increased.

Based on the MS and Western analysis of purified H2Bv-FLAG, I concluded that H2Bv is not ubiquitinated.

H3 K4 and K76 Methylation is enriched in H2Bv-containing Nucleosomes

Although there was no evidence of H2Bv ubiquitination, I continued to pursue the hypothesis that H2Bv affected H3 K4 and K76 methylation. Both H3 methyl marks could be assayed using specific antibodies. H3 K76 methylation was evaluated using an H3K76me3-specific antibody that was characterized previously [166].

To investigate whether H2Bv specifically co-localized with H3 that is methylated at K4 and K76, I prepared mononucleosomes from cell lines containing FLAG-tagged H2B or H2Bv (Fig. 5.8A). Mononucleosomes were immunoprecipitated with anti-FLAG affinity resin, and H2B- and H2Bv-containing mononucleosomes were evaluated for levels of H3 K4 and K76 methylation using

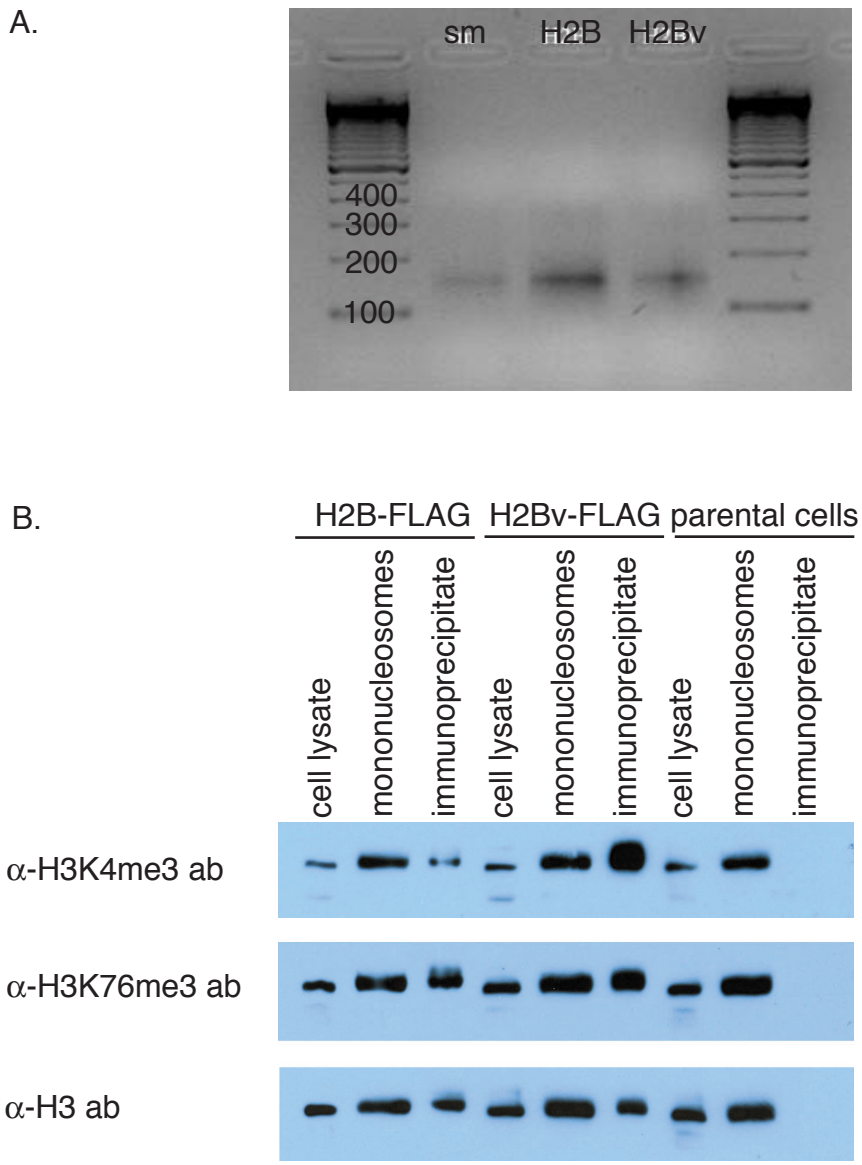


Figure 5.8. Histone H2Bv co-immunoprecipitates with Histone H3 that is trimethylated at lysines 4 and 76. (A) DNA extracted from the nucleosomes (prepared from cell lines containing FLAG-tagged H2B or H2Bv) is approximately 150 bp, demonstrating that mononucleosomes were prepared. (B) Mononucleosomes were immunoprecipitated with anti-FLAG affinity resin and evaluated by Western blot with anti-H3K4me3 and K76me3 antibodies. H2Bv-containing mononucleosomes were highly enriched for H3 that is trimethylated at K4 compared to H2B-containing mononucleosomes. H2Bv-containing mononucleosomes were slightly, but reproducibly, enriched for H3 that is trimethylated at K76. The general anti-H3 antibody demonstrates that the immunoprecipitate was loaded equally in the H2B and H2Bv lanes.

specific antibodies (Fig. 5.8B). H2Bv-containing mononucleosomes were highly enriched for H3 K4 methylation compared to H2B-containing mononucleosomes. H2Bv-containing mononucleosomes were also slightly, but very reproducibly, enriched for H3 K76 methylation. A general H3 antibody showed that the amount of H3 loaded from the H2B and H2Bv immunoprecipitate was approximately equal.

These results showed that there was a specific association between H2Bv and trimethyl H3 K4 and K76. This was consistent with the relationship observed in yeast between the H2B C-terminus and H3 K4 and K79 methylation, although, because I have shown that H2Bv is not ubiquitinated, it is clear that *T. brucei* H2Bv affects H3 methylation by an independent pathway.

H2Bv K129 does not affect H3 K4 or K76 Methylation

To investigate whether H2Bv K129 played a role in the enrichment of H3 K4 and K76 methylation in H2Bv-containing mononucleosomes, K129 was mutated to alanine (K129A) or arginine (K129R). Cell lysate from cell lines that replaced the two endogenous *h2bv* alleles with FLAG-tagged H2Bv K129A or K129R was collected and analyzed by Western blot using the H3K4me3- and H3K76me3-specific antibodies (Fig. 5.9). Mutation of H2Bv K129 appeared to have no effect on the level of H3 K4 and K76 methylation compared to wildtype H2Bv. This was consistent with our previous results that H2Bv K129 is not

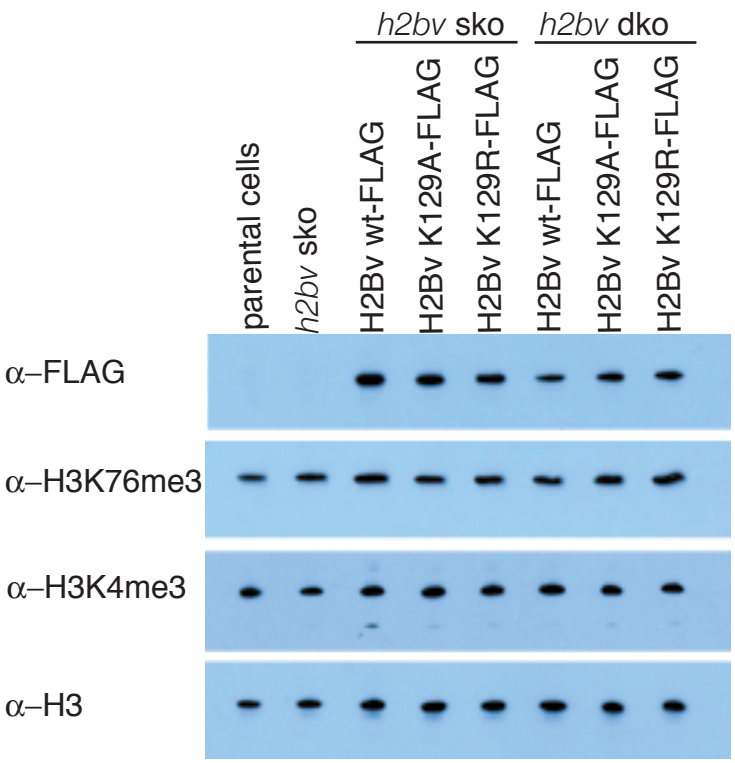


Figure 5.9. Mutation of histone H2Bv lysine 129 does not interfere with H3 lysine 4 or 76 methylation. Cell lines with ectopic copies of FLAG-tagged H2Bv wt, H2Bv K129A, or H2Bv K129R were introduced in *h2bv* single knockout (sko) or double knockout (dko) backgrounds. Cell lysate from these cell lines were loaded on a Western blot and probed with anti-FLAG, anti-H3K4me3, anti-H3K76me3, and anti-H3 antibodies. H2Bv K129 mutants did not demonstrate any change in H3 K4 or K76 methylation compared to wildtype H2Bv.

ubiquitinated, so I concluded that H2Bv K129 does not play a role homologous to yeast H2B K123.

Next, I collected mononucleosomes containing FLAG-tagged H2Bv K129A and K129R and immunoprecipitated them with anti-FLAG affinity gel. Mutation of H2Bv K129 did not prevent enrichment of H3 K4 and K76 methylation in H2Bv-containing nucleosomes (Fig. 5.10). These results show that H2Bv lysine 129 is not involved in co-localization of H2Bv with methylated H3.

The C-terminus of H2Bv is Essential

T. brucei H2B and H2Bv share significant sequence homology at the C-terminus, except for the last 5-15 amino acids. I hypothesized that the H2Bv C-terminal tail might be responsible for the enrichment of H3 methylation in H2Bv-containing mononucleosomes, even if K129 was not. The N-terminus and core domain of the two histones share minimal sequence homology, so it is also possible that these domains might be responsible for the co-immunoprecipitation results, but I did not pursue this further.

To study the relationship between the H2Bv C-terminal tail and H3 methylation, a FLAG-tagged H2Bv mutant was made in which the C-terminal tail (5´ –AKYNASREEAYSKVL) was removed. Unusually, when the FLAG-tagged H2BvΔCTD mutant replaced one allele of endogenous *h2bv*, the FLAG fusion protein was not observed in any of the four clones (Fig. 5.11). Proper integration

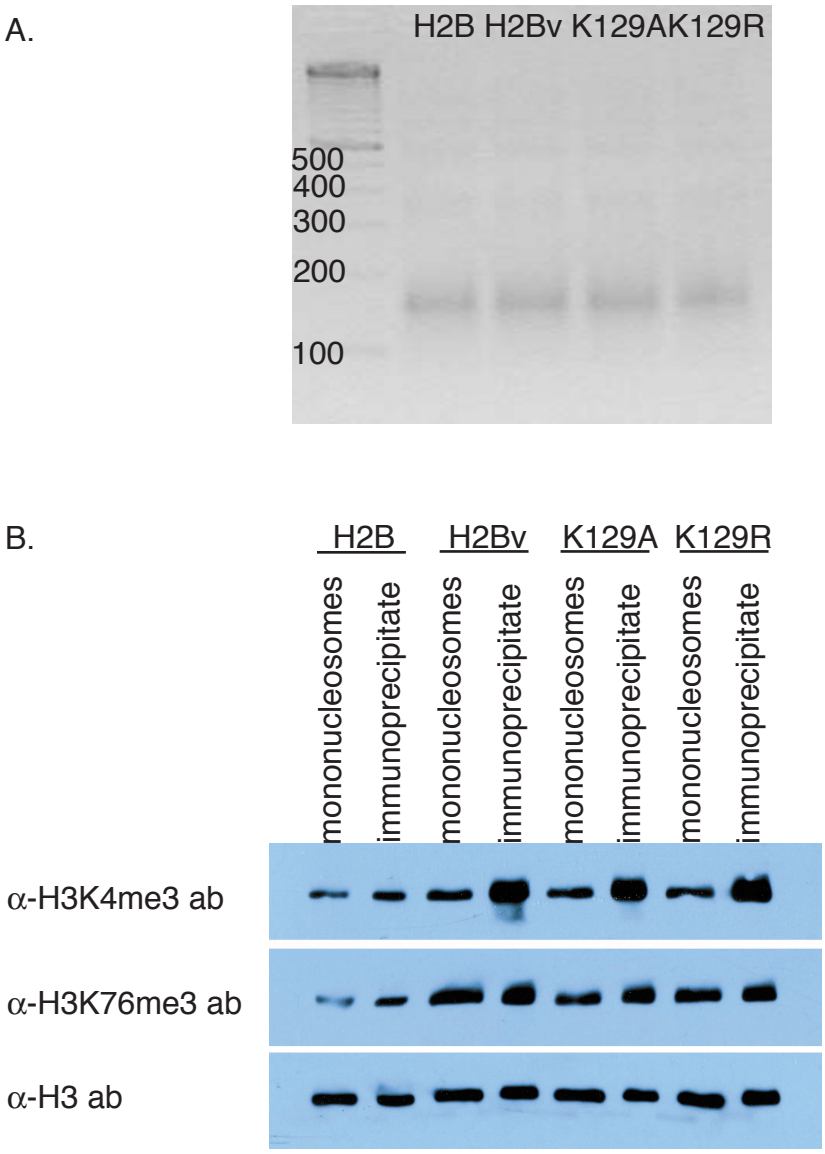


Figure 5.10. Mutation of histone H2Bv lysine 129 does not prevent H3 that is methylated at K4 and K76 from co-immunoprecipitating with H2Bv. (A) Mononucleosomes were prepared from cell lines with FLAG-tagged H2B, H2Bv, H2Bv K129A, and H2Bv K129R. DNA was extracted from the mononucleosomes and analyzed on an agarose gel. (B) Mononucleosomes were immunoprecipitated on anti-FLAG resin and analyzed by Western blot with anti-H3K4me3, H3K76me3, and H3 antibodies. Immunoprecipitate from H2Bv K129 mutants showed no change in the level of H3 K4 and K76 methylation associated with H2Bv-containing mononucleosomes compared to wildtype H2Bv.

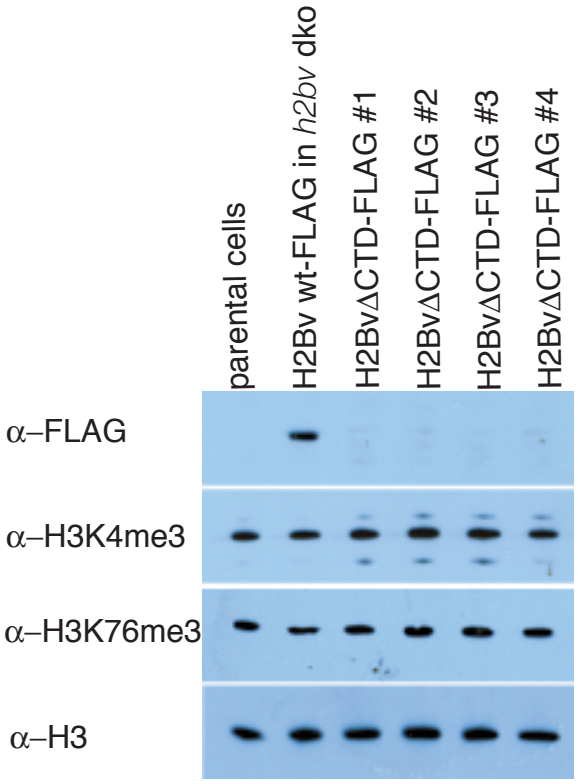


Figure 5.11. Deletion of the C-terminal tail of histone H2Bv lysine 129 does not interfere with H3 lysine 4 or 76 methylation. In *h2bv* single knockout (sko) cells, an ectopic copy of FLAG-tagged H2Bv lacking its C-terminal tail (H2BvΔCTD) was introduced, and cell lysate was collected from four representative clones. These cell lysates were analyzed by Western blot using anti-FLAG, anti-H3K4me3, anti-H3K76me3, and anti-H3 antibodies. H2BvΔCTD mutants did not express their FLAG fusion protein. These mutants did not demonstrate any change in H3 K4 or K76 methylation compared to wildtype H2Bv.

of the mutant was confirmed by sequencing. The ectopic copy of H2Bv was downstream of an inducible T7 promoter, so the H2Bv mutant must be transcribed by the T7 polymerase, and either the mRNA or protein is rapidly degraded. This is an interesting result because, even though an endogenous copy of *h2bv* is present, the cell will not tolerate the presence of the H2Bv mutant, suggesting that the C-terminal tail plays an essential role in the nucleosome. Also, the second allele of *h2bv* could not be replaced, confirming that the C-terminal tail of H2Bv is essential.

Discussion

The remarkable sequence conservation between *T. brucei* H2Bv and the H2B consensus sequence led to our initial hypothesis that lysine 129 was ubiquitinated and required for H3 K4 and K76 methylation. Ubiquitinated histones are frequently present in low abundance, so they may be overlooked when performing a survey of histone PTMs by MS. However, I performed a targeted MS search for ubiquitinated H2Bv K129 by specifically enriching for the mass corresponding to the theoretical mass of ubiquitinated H2Bv peptide 126-134. The presence of a single ubiquitin peptide in gel slices #6 and 7 was most likely due to ubiquitinated H2A, a PTM that was identified in a previous chapter. H2A co-immunoprecipitated with H2Bv and was also found migrating in a higher

molecular weight range (25-32 kDa) than would be expected based on its molecular weight (14.1 kDa).

I also analyzed purified H2Bv-FLAG by Western blot with an anti-ubiquitin antibody and found there were no bands present corresponding to ubiquitinated H2Bv. Both methods enabled us to conclude that H2Bv K129 is not ubiquitinated. This is supported by the observation that mutation of K129 is completely acceptable to the cells and does not disrupt H3 K4 or K76 methylation. These experiments show that H2Bv K129 does not play a role homologous to yeast H2B K123, despite the sequence similarities.

Although our original hypothesis proved to be incorrect, I was able to show, using specific antibodies, that H2Bv was found in mononucleosomes enriched for H3 K4 and K76 methylation. H3 K4 and K76 methylation are typically associated with transcriptionally active chromatin, although this relationship has not been demonstrated in trypanosomes [196]. Based on our results, I hypothesized that H2Bv replaces H2B in nucleosomes, which then permits H3 K4 and K76 to be methylated. Replacing modified canonical histones with histone variants may be a mechanism for relieving transcriptional repression and would be in keeping with one of the roles of histone variants in other organisms.

Previous work on *T. brucei* H2Bv and histone variant H2AZ showed that the two variants co-immunoprecipitate in the same nucleosome [174]. H2AZ does not co-immunoprecipitate with canonical H2A or H2B, suggesting that

nucleosomes containing H2Bv consist of two H2AZ/H2Bv dimers. As an extension of our previous hypothesis, perhaps H2AZ/H2Bv dimers replace canonical H2A/H2B dimers in the nucleosome as a mechanism for relieving transcriptional repression. Canonical H2A and H2B may repress transcription or may simply prevent H3 from being methylated at sites associated with active transcription. In this way, histone variants may play a role in the regulation of gene expression in trypanosomes.

The same study showed that H2AZ and H2Bv did not co-localize with sites of active transcription, which were visualized by BrUTP incorporation [174]. However, the authors suggest that perhaps the H2AZ/H2Bv-containing nucleosomes are localized to the promoter regions of active genes and are displaced during transcription by RNA polymerase II, so the variants may not be readily visible at actively transcribed chromatin. It seems unlikely, although still a possibility, that H2Bv would be found in nucleosomes enriched for H3 K4 and K76 methylation, marks traditionally associated with transcriptional activity, and would not be present at sites of active transcription.

Since H2AZ and H2BV appear to be present in the same nucleosome, another alternative pathway leading to H3 K4 and K76 methylation could involve ubiquitination at the H2AZ C-terminus. Like canonical H2A, trypanosome H2AZ has 6 lysines at its C-terminus. However, none of these aligns with the ubiquitinated H2A K119 observed in other organisms (discussed in Chapter 3).

This hypothesis can be readily addressed by epitope-tagging and purifying H2AZ and analyzing the sample by Western blot with a ubiquitin antibody and MS.

H2Bv is an essential gene. A deletion mutant of the first 23 amino acids of *H2Bv* was viable, indicating that the N-terminus is not essential [174]. A FLAG-tagged H2Bv mutant in which C-terminal residues 128-142 were removed was not expressed in an *h2bv* single knockout background. The second allele of *h2bv* could not be removed in this cell line, which illustrates that residues 128-142 are essential. It is difficult to draw any conclusions concerning the role of a mutation when the phenotype is lethality, but I hypothesized that residues 128-142 are involved in the enrichment of H3 methylation in H2Bv-containing nucleosomes. I based this hypothesis on the conservation of some residues in the H2Bv C-terminal tail with H2B in other organisms, but also on the absence of sequence homology between *T. brucei* H2B and H2Bv C-terminal tails when the remainder of the C-terminus is highly conserved. Although the mechanism has not been delineated, it is clear that trypanosome H3 K4 and K76 methylation are governed by a different pathway than what is observed in yeast.

Chapter 6. Conclusions

Future Steps in Trypanosome Chromatin Biology

A comprehensive list of PTMs on the four core histones provides a necessary starting point for the study of chromatin in *T. brucei*. Based on this information, additional approaches may be established to aid in understanding the relationship between chromatin structure and gene expression in trypanosomes.

Site-Directed Mutagenesis of H3 and H4. One approach to determining how various PTMs are involved in gene expression or other biological processes would be to introduce point mutations in the amino acids of interest and assay for a phenotype. Our lab has chosen to focus on histones H3 and H4. Both *H3* and *H4* are present in tandem arrays, which must be deleted and replaced with a single ectopic copy of each gene. The ectopic copies should be introduced with endogenous UTRs because the expression of at least one histone has been shown to be cell cycle regulated [197]. Also, *H3* and *H4* would probably need to be introduced downstream of an inducible T7 promoter so that they may be expressed at levels comparable to their expression level in wild-type cells. In yeast and other organisms, the N-termini of these histones have been shown to be functionally redundant in some cases, so it would be preferable to make one cell line in which both *H3* and *H4* were present in single copies. For example, if a

single point mutation in *H4* did not demonstrate a phenotype, the N-terminus of *H3* could be deleted to determine if it was functionally redundant to *H4*. With this cell line, point mutations in the modified amino acids could be assayed for specific phenotypes, including de-repression of *VSG* silencing at the inactive ES.

Histone modifying enzymes. A number of putative histone modifying enzymes have been identified in the trypanosome genome project, some of which are beginning to be characterized (see Chapter 1)[148]. Determining which histone modifying enzyme creates each PTM could be a straightforward endeavor. For non-essential enzymes, histones would be purified from knockout cell lines according to the protocol described in Chapter 3. Alternatively, if an enzyme is essential, an RNAi knockdown construct could be created. Purified histones could then be modified with deuterated propionic acid, mixed with propionylated histones from wt BF or PF, and trypsinized. Changes in the abundance of PTMs in mutant cells could be determined by comparing the intensity of their corresponding MALDI-TOF peaks to wt peaks. Also, if specific antibodies were available, mutant cell lines could be more easily tested by Western blot.

Once the targets of histone modifying enzymes have been identified, deletion mutants of the enzymes or histone target mutations could be studied for defects in lifecycle stage differentiation, cell cycle progression, ES silencing, and non-ES transcriptional silencing and activation. The resulting data, and the

phenotypes of mutagenized cells described above, could be used to determine which histone PTMs are involved in these biological processes.

Specific antibodies. Antibodies specific for histone PTMs would be valuable tools for studying marks that have interesting preliminary functional data. They may be used for analysis by Western blot, immunofluorescence, and chromatin immunoprecipitation. I previously established that commercially available antibodies are not specific for divergent trypanosome histones, so antibodies must be raised against modified trypanosome peptides and rigorously purified and validated.

PTM-binding proteins. The downstream biological effects of histone PTMs are mediated through the binding of effector proteins [2]. In order to identify effector proteins that bind PTMs on trypanosome histones, modified trypanosome peptides may be used for affinity purification of binding partners from trypanosome cell extracts. These proteins may be resolved by SDS-PAGE and identified by MS. Effector proteins must be distinguished from proteins that bind the unmodified peptide, although proteins that bind the unmodified peptide, but not the modified peptide, may also be of interest. The epitope-tagged effector protein may also be used to pulldown additional proteins that potentially act together in a complex. Finally, PTM-binding proteins may be deleted to determine which biological processes are affected.

Antigenic Variation

Heterochromatin formation at the 3´ end of inactive ES may be involved in prohibiting transcription by RNA pol I, which initiates transcription at many ES, but only transcribes to the 3´ end of a single ES [111]. The two pathways of heterochromatin assembly described in Chapter 1, mediated by HP1 and SIR2, do not appear to be involved in ES silencing. HP1 and H3 K9 do not have homologues in trypanosomes. A SIR2-related protein, TbSIR2RP-1, plays a role in establishing telomeric silencing at non-ES loci, but not at ES loci [170]. Perhaps a novel pathway is responsible for heterochromatin assembly at the ES. The tools described above may help determine whether histone PTMs, histone modifying enzymes, and effector proteins are involved in establishing transcriptionally restrictive or permissive chromatin at the ES.

It should be noted that heterochromatin formation is not the only mechanism thought to be regulating ES silencing. As mentioned in Chapter 1, the active ES is localized to the ESB, so expulsion of the inactive ES from the ESB may also inhibit transcription of these loci. This is important because some studies have already been published in which deletion of factors, including histone modifying enzymes, have been assayed for de-repression of the inactive ES, but no major effects on ES silencing have been observed. These factors may be involved in heterochromatin formation, which could be inhibited in deletion mutants, but the inactive ES in the mutants would still not have access to

factors required to promote transcription to the 3´ end of the ES, so no ES phenotype would be readily observed. Therefore, better assays must be devised to test a candidate protein for involvement in ES silencing.

In order to better understand what makes the ESB the exclusive site of VSG transcription, we should first identify the factors that comprise the ESB. It is predicted that the ESB will contain both elongation factors and RNA processing machinery, since it was shown that only transcripts from the active ES were properly processed [111]. Elongation factors in other organisms include PTM-binding effector proteins and histone modifying enzymes, so these should be epitope-tagged in trypanosomes and visualized by immunofluorescence for localization to the ESB. This would provide a number of candidates that could be involved in ES transcription regulation, although this approach might be challenging if the critical factors are present in very small quantities.

Conservation in an Evolutionarily Divergent Organism

In the previous chapter, I showed that H2Bv shared significant sequence homology with yeast H2B and is present in nucleosomes enriched for H3 K4 and K76 methylation. However, ubiquitination of the H2Bv C-terminus was not required for this enrichment. I propose that, while mechanisms of histone cross-talk are not conserved, the same endpoint is achieved in both organisms. In yeast, H2B K123 ubiquitination acts as a trigger for downstream events that

include H3 K4 and K79 methylation, all of which are associated with transcription elongation. In trypanosomes, perhaps the same trigger is achieved by replacing H2A/H2B dimers with H2AZ/H2Bv dimers in nucleosomes. The H2AZ/H2Bv dimers may act to relieve repression imposed by H2A/H2B or may somehow actively facilitate H3 K4 and K76 methylation. H2AZ and H2BV share 43% and 38% sequence identity with *T. brucei* H2A and H2B, respectively [174]. Sequences that diverge between the variant and its corresponding major histone may be sufficient to induce H3 methylation, even in the absence of ubiquitination. Based on our data and sequence alignments, a likely sequence candidate would be the C-terminal tail of H2BV (aa 128-142), which is essential.

There are three components to our model, all of which can be addressed. First, I propose that the H2Bv/H2AZ dimer is deposited in the nucleosome in a replication-independent manner. There is ample support for replication-independent deposition of histone variants in the literature. FLAG-tagged H2Bv could be used to pulldown interacting proteins, which could be identified by MS. Putative chromatin remodeling factors or histone chaperones would make excellent candidates; these, in turn, could be deleted to look for defects of H2Bv/H2AZ incorporation into the nucleosome.

Second, I propose that H2Bv/H2AZ deposition allows for H3 K4 and K76 methylation. This process may be mediated through the C-terminal tail of H2Bv. In addition to possessing several conserved residues from the yeast H2B C-terminus, the C-terminal tail is a good candidate because the crystal structure of

the nucleosome predicts that it is in close proximity to H3 K76, which would facilitate interactions between the two domains [198]. Introducing an ectopic copy of H2B with the H2Bv C-terminal tail may result in enrichment of H3 K4 and K76 methylation in H2B-containing nucleosomes. This would show that the H2Bv C-terminal tail was sufficient to direct H3 K4 and K76 methylation.

Third, I propose that H3 K4 and K76 methylation is associated with actively transcribed chromatin. Chromatin immunoprecipitation using specific antibodies against trimethyl H3 K4 and K76 could show whether these marks are associated with transcriptionally active loci. Primers against active and inactive *VSG*s could be used to distinguish transcriptionally active and silent loci, but loci that are transcribed by RNA pol II should also be examined.

Our efforts to study chromatin in trypanosomes illustrate that, while some histone sequences and posttranslational modifications are conserved, the functional relationships between marks may not be. Work in divergent organisms will help to define some of the general principles underlying chromatin biology.

References

[1] Kouzarides, T. (2007). Chromatin modifications and their function. Cell 128, 693-705.

[2] Jenuwein, T. and Allis, C.D. (2001). Translating the histone code. Science 293, 1074-80.

[3] Lachner, M., O'Carroll, D., Rea, S., Mechtler, K. and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. Nature 410, 116-20.

[4] Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C. and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature 410, 120-4.

[5] Jacobs, S.A., Taverna, S.D., Zhang, Y., Briggs, S.D., Li, J., Eissenberg, J.C., Allis, C.D. and Khorasanizadeh, S. (2001). Specificity of the HP1 chromo domain for the methylated N-terminus of histone H3. EMBO J 20, 5232-41.

[6] Owen, D.J. et al. (2000). The structural basis for the recognition of acetylated histone H4 by the bromodomain of histone acetyltransferase gcn5p. EMBO J 19, 6141-9.

[7] Dhalluin, C., Carlson, J.E., Zeng, L., He, C., Aggarwal, A.K. and Zhou, M.M. (1999). Structure and ligand of a histone acetyltransferase bromodomain. Nature 399, 491-6.

[8] Winston, F. and Allis, C.D. (1999). The bromodomain: a chromatin-targeting module? Nat Struct Biol 6, 601-4.

[9] Grunstein, M. (1997). Histone acetylation in chromatin structure and transcription. Nature 389, 349-52.

[10] Pokholok, D.K. et al. (2005). Genome-wide map of nucleosome acetylation and methylation in yeast. Cell 122, 517-27.

[11] Brower-Toland, B., Wacker, D.A., Fulbright, R.M., Lis, J.T., Kraus, W.L. and Wang, M.D. (2005). Specific contributions of histone tails and their acetylation to the mechanical stability of nucleosomes. J Mol Biol 346, 135-46.

[12] Jacobson, R.H., Ladurner, A.G., King, D.S. and Tjian, R. (2000). Structure and function of a human TAFII250 double bromodomain module. Science 288, 1422-5.

[13] Tschiersch, B., Hofmann, A., Krauss, V., Dorn, R., Korge, G. and Reuter, G. (1994). The protein encoded by the Drosophila position-effect variegation suppressor gene Su(var)3-9 combines domains of antagonistic regulators of homeotic gene complexes. EMBO J 13, 3822-31.

[14] Eissenberg, J.C., Morris, G.D., Reuter, G. and Hartnett, T. (1992). The heterochromatin-associated protein HP-1 is an essential protein in Drosophila with dosage-dependent effects on position-effect variegation. Genetics 131, 345-52.

[15] Rea, S. et al. (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature 406, 593-9.

[16] Nielsen, A.L., Sanchez, C., Ichinose, H., Cervino, M., Lerouge, T., Chambon, P. and Losson, R. (2002). Selective interaction between the chromatin-remodeling factor BRG1 and the heterochromatin-associated protein HP1alpha. EMBO J 21, 5797-806.

[17] Aagaard, L. et al. (1999). Functional mammalian homologues of the Drosophila PEV-modifier Su(var)3-9 encode centromere-associated proteins which complex with the heterochromatin component M31. EMBO J 18, 1923-38.

[18] Lachner, M. and Jenuwein, T. (2002). The many faces of histone lysine methylation. Curr Opin Cell Biol 14, 286-98.

[19] Grewal, S.I. and Moazed, D. (2003). Heterochromatin and epigenetic control of gene expression. Science 301, 798-802.

[20] Nakayama, J., Rice, J.C., Strahl, B.D., Allis, C.D. and Grewal, S.I. (2001). Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. Science 292, 110-3.

[21] Verdel, A., Jia, S., Gerber, S., Sugiyama, T., Gygi, S., Grewal, S.I. and Moazed, D. (2004). RNAi-mediated targeting of heterochromatin by the RITS complex. Science 303, 672-6.

[22] Hall, I.M., Shankaranarayana, G.D., Noma, K., Ayoub, N., Cohen, A. and Grewal, S.I. (2002). Establishment and maintenance of a heterochromatin domain. Science 297, 2232-7.

[23] Volpe, T.A., Kidner, C., Hall, I.M., Teng, G., Grewal, S.I. and Martienssen, R.A. (2002). Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. Science 297, 1833-7.

[24] Cam, H.P., Sugiyama, T., Chen, E.S., Chen, X., FitzGerald, P.C. and Grewal, S.I. (2005). Comprehensive analysis of heterochromatin- and RNAi-mediated epigenetic control of the fission yeast genome. Nat Genet 37, 809-19.

[25] Rusche, L.N., Kirchmaier, A.L. and Rine, J. (2002). Ordered nucleation and spreading of silenced chromatin in Saccharomyces cerevisiae. Mol Biol Cell 13, 2207-22.

[26] Moretti, P., Freeman, K., Coodly, L. and Shore, D. (1994). Evidence that a complex of SIR proteins interacts with the silencer and telomere-binding protein RAP1. Genes Dev 8, 2257-69.

[27] Luo, K., Vega-Palas, M.A. and Grunstein, M. (2002). Rap1-Sir4 binding independent of other Sir, yKu, or histone interactions initiates the assembly of telomeric heterochromatin in yeast. Genes Dev 16, 1528-39.

[28] Hoppe, G.J., Tanny, J.C., Rudner, A.D., Gerber, S.A., Danaie, S., Gygi, S.P. and Moazed, D. (2002). Steps in assembly of silent chromatin in

yeast: Sir3-independent binding of a Sir2/Sir4 complex to silencers and role for Sir2-dependent deacetylation. Mol Cell Biol 22, 4167-80.

[29] Imai, S., Armstrong, C.M., Kaeberlein, M. and Guarente, L. (2000). Transcriptional silencing and longevity protein Sir2 is an NAD-dependent histone deacetylase. Nature 403, 795-800.

[30] Gottschling, D.E., Aparicio, O.M., Billington, B.L. and Zakian, V.A. (1990). Position effect at S. cerevisiae telomeres: reversible repression of Pol II transcription. Cell 63, 751-62.

[31] Bryk, M., Banerjee, M., Murphy, M., Knudsen, K.E., Garfinkel, D.J. and Curcio, M.J. (1997). Transcriptional silencing of Ty1 elements in the RDN1 locus of yeast. Genes Dev 11, 255-69.

[32] Smith, J.S. and Boeke, J.D. (1997). An unusual form of transcriptional silencing in yeast ribosomal DNA. Genes Dev 11, 241-54.

[33] Rine, J. and Herskowitz, I. (1987). Four genes responsible for a position effect on expression from HML and HMR in Saccharomyces cerevisiae. Genetics 116, 9-22.

[34] Suka, N., Luo, K. and Grunstein, M. (2002). Sir2p and Sas2p opposingly regulate acetylation of yeast histone H4 lysine16 and spreading of heterochromatin. Nat Genet 32, 378-83.

[35] Kimura, A., Umehara, T. and Horikoshi, M. (2002). Chromosomal gradient of histone acetylation established by Sas2p and Sir2p functions as a shield against gene silencing. Nat Genet 32, 370-7.

[36] Kurdistani, S.K. and Grunstein, M. (2003). Histone acetylation and deacetylation in yeast. Nat Rev Mol Cell Biol 4, 276-84.

[37] Kayne, P.S., Kim, U.J., Han, M., Mullen, J.R., Yoshizaki, F. and Grunstein, M. (1988). Extremely conserved histone H4 N terminus is dispensable for growth but essential for repressing the silent mating loci in yeast. Cell 55, 27-39.

[38] Thompson, J.S., Ling, X. and Grunstein, M. (1994). Histone H3 amino terminus is required for telomeric and silent mating locus repression in yeast. Nature 369, 245-7.

[39] Li, B., Carey, M. and Workman, J.L. (2007). The role of chromatin during transcription. Cell 128, 707-19.

[40] Marshall, N.F., Peng, J., Xie, Z. and Price, D.H. (1996). Control of RNA polymerase II elongation potential by a novel carboxyl-terminal domain kinase. J Biol Chem 271, 27176-83.

[41] Komarnitsky, P., Cho, E.J. and Buratowski, S. (2000). Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. Genes Dev 14, 2452-60.

[42] Buratowski, S. (2003). The CTD code. Nat Struct Biol 10, 679-80.

[43] Krogan, N.J. et al. (2002). RNA polymerase II elongation factors of Saccharomyces cerevisiae: a targeted proteomics approach. Mol Cell Biol 22, 6979-92.

[44] Pokholok, D.K., Hannett, N.M. and Young, R.A. (2002). Exchange of RNA polymerase II initiation and elongation factors during gene expression in vivo. Mol Cell 9, 799-809.

[45] Squazzo, S.L. et al. (2002). The Paf1 complex physically and functionally associates with transcription elongation factors in vivo. EMBO J 21, 1764-74.

[46] Krogan, N.J. et al. (2003). The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation. Mol Cell 11, 721-9.

[47] Wood, A., Schneider, J., Dover, J., Johnston, M. and Shilatifard, A. (2003). The Paf1 complex is essential for histone monoubiquitination by the Rad6-Bre1 complex, which signals for histone methylation by COMPASS and Dot1p. J Biol Chem 278, 34739-42.

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

[49] Belotserkovskaya, R., Oh, S., Bondarenko, V.A., Orphanides, G., Studitsky, V.M. and Reinberg, D. (2003). FACT facilitates transcription-dependent nucleosome alteration. Science 301, 1090-3.

[50] Ruthenburg, A.J., Allis, C.D. and Wysocka, J. (2007). Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark. Mol Cell 25, 15-30.

[51] Schubeler, D. et al. (2004). The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. Genes Dev 18, 1263-71.

[52] Schneider, R., Bannister, A.J., Myers, F.A., Thorne, A.W., Crane-Robinson, C. and Kouzarides, T. (2004). Histone H3 lysine 4 methylation patterns in higher eukaryotic genes. Nat Cell Biol 6, 73-7.

[53] Liang, G. et al. (2004). Distinct localization of histone H3 acetylation and H3-K4 methylation to the transcription start sites in the human genome. Proc Natl Acad Sci U S A 101, 7357-62.

[54] Briggs, S.D., Bryk, M., Strahl, B.D., Cheung, W.L., Davie, J.K., Dent, S.Y., Winston, F. and Allis, C.D. (2001). Histone H3 lysine 4 methylation is mediated by Set1 and required for cell growth and rDNA silencing in Saccharomyces cerevisiae. Genes Dev 15, 3286-95.

[55] Boa, S., Coert, C. and Patterton, H.G. (2003). Saccharomyces cerevisiae Set1p is a methyltransferase specific for lysine 4 of histone H3 and is required for efficient gene expression. Yeast 20, 827-35.

[56] Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J.R., Cole, P.A. and Casero, R.A. (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell 119, 941-53.

[57] Li, H., Ilin, S., Wang, W., Duncan, E.M., Wysocka, J., Allis, C.D. and Patel, D.J. (2006). Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. Nature 442, 91-5.

[58] Wysocka, J. et al. (2006). A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. Nature 442, 86-90.

[59] Pray-Grant, M.G., Daniel, J.A., Schieltz, D., Yates, J.R., 3rd and Grant, P.A. (2005). Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation. Nature 433, 434-8.

[60] Hampsey, M. and Reinberg, D. (2003). Tails of intrigue: phosphorylation of RNA polymerase II mediates histone methylation. Cell 113, 429-32.

[61] Fischle, W., Wang, Y. and Allis, C.D. (2003). Histone and chromatin cross-talk. Curr Opin Cell Biol 15, 172-83.

[62] Briggs, S.D., Xiao, T., Sun, Z.W., Caldwell, J.A., Shabanowitz, J., Hunt, D.F., Allis, C.D. and Strahl, B.D. (2002). Gene silencing: trans-histone regulatory pathway in chromatin. Nature 418, 498.

[63] Sun, Z.W. and Allis, C.D. (2002). Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. Nature 418, 104-8.

[64] Dover, J., Schneider, J., Tawiah-Boateng, M.A., Wood, A., Dean, K., Johnston, M. and Shilatifard, A. (2002). Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6. J Biol Chem 277, 28368-71.

[65] Shilatifard, A. (2006). Chromatin Modifications by Methylation and Ubiquitination: Implications in the Regulation of Gene Expression. Annu Rev Biochem 75, 243-69.

[66] Wood, A. et al. (2003). Bre1, an E3 ubiquitin ligase required for recruitment and substrate selection of Rad6 at a promoter. Mol Cell 11, 267-74.

[67] Ng, H.H., Dole, S. and Struhl, K. (2003). The Rtf1 component of the Paf1 transcriptional elongation complex is required for ubiquitination of histone H2B. J Biol Chem 278, 33625-8.

[68] Xiao, T., Kao, C.F., Krogan, N.J., Sun, Z.W., Greenblatt, J.F., Osley, M.A. and Strahl, B.D. (2005). Histone H2B ubiquitylation is associated with elongating RNA polymerase II. Mol Cell Biol 25, 637-51.

[69] Daniel, J.A., Torok, M.S., Sun, Z.W., Schieltz, D., Allis, C.D., Yates, J.R., 3rd and Grant, P.A. (2004). Deubiquitination of histone H2B by a yeast acetyltransferase complex regulates transcription. J Biol Chem 279, 1867-71.

[70] Henry, K.W. et al. (2003). Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8. Genes Dev 17, 2648-63.

[71] Bernstein, E. and Hake, S.B. (2006). The nucleosome: a little variation goes a long way. Biochem Cell Biol 84, 505-517.

[72] Varga-Weisz, P.D. and Becker, P.B. (1998). Chromatin-remodeling factors: machines that regulate? Curr Opin Cell Biol 10, 346-53.

[73] Meneghini, M.D., Wu, M. and Madhani, H.D. (2003). Conserved histone variant H2A.Z protects euchromatin from the ectopic spread of silent heterochromatin. Cell 112, 725-36.

[74] Guillemette, B., Bataille, A.R., Gevry, N., Adam, M., Blanchette, M., Robert, F. and Gaudreau, L. (2005). Variant histone H2A.Z is globally localized to the promoters of inactive yeast genes and regulates nucleosome positioning. PLoS Biol 3, e384.

[75] Li, B., Pattenden, S.G., Lee, D., Gutierrez, J., Chen, J., Seidel, C., Gerton, J. and Workman, J.L. (2005). Preferential occupancy of histone variant H2AZ at inactive promoters influences local histone modifications and chromatin remodeling. Proc Natl Acad Sci U S A 102, 18385-90.

[76] Zhang, H., Roberts, D.N. and Cairns, B.R. (2005). Genome-wide dynamics of Htz1, a histone H2A variant that poises repressed/basal promoters for activation through histone loss. Cell 123, 219-31.

[77] Zanton, S.J. and Pugh, B.F. (2006). Full and partial genome-wide assembly and disassembly of the yeast transcription machinery in response to heat shock. Genes Dev 20, 2250-65.

[78] Rangasamy, D., Berven, L., Ridgway, P. and Tremethick, D.J. (2003). Pericentric heterochromatin becomes enriched with H2A.Z during early mammalian development. EMBO J 22, 1599-607.

[79] Rangasamy, D., Greaves, I. and Tremethick, D.J. (2004). RNA interference demonstrates a novel role for H2A.Z in chromosome segregation. Nat Struct Mol Biol 11, 650-5.

[80] Mizuguchi, G., Shen, X., Landry, J., Wu, W.H., Sen, S. and Wu, C. (2004). ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. Science 303, 343-8.

[81] Krogan, N.J. et al. (2003). A Snf2 family ATPase complex required for recruitment of the histone H2A variant Htz1. Mol Cell 12, 1565-76.

[82] Kobor, M.S., Venkatasubrahmanyam, S., Meneghini, M.D., Gin, J.W., Jennings, J.L., Link, A.J., Madhani, H.D. and Rine, J. (2004). A protein complex containing the conserved Swi2/Snf2-related ATPase Swr1p deposits histone variant H2A.Z into euchromatin. PLoS Biol 2, E131.

[83] Ahmad, K. and Henikoff, S. (2002). The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. Mol Cell 9, 1191-200.

[84] Tagami, H., Ray-Gallet, D., Almouzni, G. and Nakatani, Y. (2004). Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. Cell 116, 51-61.

[85] Konev, A.Y. et al. (2007). CHD1 motor protein is required for deposition of histone variant H3.3 into chromatin in vivo. Science 317, 1087-90.

[86] Chow, C.M. et al. (2005). Variant histone H3.3 marks promoters of transcriptionally active genes during mammalian cell division. EMBO Rep 6, 354-60.

[87] McKittrick, E., Gafken, P.R., Ahmad, K. and Henikoff, S. (2004). Histone H3.3 is enriched in covalent modifications associated with active chromatin. Proc Natl Acad Sci U S A 101, 1525-30.

[88] Nakayama, T., Nishioka, K., Dong, Y.X., Shimojima, T. and Hirose, S. (2007). Drosophila GAGA factor directs histone H3.3 replacement that prevents the heterochromatin spreading. Genes Dev 21, 552-61.

[89] Ross, R. and Thomson, D. (1910). A case of sleeping sickness studied by precise enumerative methods: regular periodical increase of the parasites disclosed. Proc R Soc London Biol 82, 411-415.

[90] Vickerman, K. (1969). On the surface coat and flagellar adhesion in trypanosomes. J Cell Sci 5, 163-93.

[91] Cross, G.A.M. (1975). Identification, purification and properties of clone-specific glycoprotein antigens constituting the surface coat of Trypanosoma brucei. Parasitology 71, 393-417.

[92] Miller, E.N., Allan, L.M. and Turner, M.J. (1984). Topological analysis of antigenic determinants on a variant surface glycoprotein of Trypanosoma brucei. Mol Biochem Parasitol 13, 67-81.

[93] Cross, G.A.M. (1977). Isolation, structure and function of variant-specific surface antigens. Ann Soc Belg Med Trop 57, 389-402.

[94] Berriman, M. et al. (2005). The genome of the African trypanosome Trypanosoma brucei. Science 309, 416-22.

[95] Marcello, L. and Barry, J.D. (2007). Analysis of the VSG gene silent archive in Trypanosoma brucei reveals that mosaic gene expression is prominent in antigenic variation and is favored by archive substructure. Genome Res 17, 1344-52.

[96] Callejas, S., Leech, V., Reitter, C. and Melville, S. (2006). Hemizygous subtelomeres of an African trypanosome chromosome may account for over 75% of chromosome length. Genome Res 16, 1109-18.

[97] Melville, S.E., Leech, V., Navarro, M. and Cross, G.A.M. (2000). The molecular karyotype of the megabase chromosomes of Trypanosoma brucei stock 427. Mol Biochem Parasitol 111, 261-73.

[98] Van der Ploeg, L.H., Schwartz, D.C., Cantor, C.R. and Borst, P. (1984). Antigenic variation in Trypanosoma brucei analyzed by electrophoretic separation of chromosome-sized DNA molecules. Cell 37, 77-84.

[99] El-Sayed, N.M., Hegde, P., Quackenbush, J., Melville, S.E. and Donelson, J.E. (2000). The African trypanosome genome. Int J Parasitol 30, 329-45.

[100] Vickerman, K. (1965). Polymorphism and mitochondrial activity in sleeping sickness trypanosomes. Nature 208, 762-6.

[101] Roditi, I. et al. (1989). Procyclin gene expression and loss of the variant surface glycoprotein during differentiation of Trypanosoma brucei. J Cell Biol 108, 737-46.

[102] Crowe, J.S., Barry, J.D., Luckins, A.G., Ross, C.A. and Vickerman, K. (1983). All metacyclic variable antigen types of Trypanosoma congolense identified using monoclonal antibodies. Nature 306, 389-91.

[103] Matthews, K.R. (2005). The developmental cell biology of Trypanosoma brucei. J Cell Sci 118, 283-90.

[104] Ziegelbauer, K., Quinten, M., Schwarz, H., Pearson, T.W. and Overath, P. (1990). Synchronous differentiation of Trypanosoma brucei from bloodstream to procyclic forms in vitro. Eur J Biochem 192, 373-8.

[105] Overath, P., Czichos, J., Stock, U. and Nonnengaesser, C. (1983). Repression of glycoprotein synthesis and release of surface coat during transformation of Trypanosoma brucei. EMBO J 2, 1721-8.

[106] Pays, E., Hanocq-Quertier, J., Hanocq, F., Van Assel, S., Nolan, D. and Rolin, S. (1993). Abrupt RNA changes precede the first cell division during the differentiation of Trypanosoma brucei bloodstream forms into procyclic forms in vitro. Mol Biochem Parasitol 61, 107-14.

[107] Matthews, K.R. and Gull, K. (1994). Evidence for an interplay between cell cycle progression and the initiation of differentiation between life cycle forms of African trypanosomes. J Cell Biol 125, 1147-56.

[108] Priest, J.W. and Hajduk, S.L. (1994). Developmental regulation of Trypanosoma brucei cytochrome c reductase during bloodstream to procyclic differentiation. Mol Biochem Parasitol 65, 291-304.

[109] Matthews, K.R., Sherwin, T. and Gull, K. (1995). Mitochondrial genome repositioning during the differentiation of the African trypanosome between life cycle forms is microtubule mediated. J Cell Sci 108 (Pt 6), 2231-9.

[110] Gunzl, A., Bruderer, T., Laufer, G., Schimanski, B., Tu, L.C., Chung, H.M., Lee, P.T. and Lee, M.G. (2003). RNA polymerase I transcribes procyclin genes and variant surface glycoprotein gene expression sites in Trypanosoma brucei. Eukaryot Cell 2, 542-51.

[111] Vanhamme, L., Poelvoorde, P., Pays, A., Tebabi, P., Xong, H.V. and Pays, E. (2000). Differential RNA elongation controls the variant surface glycoprotein gene expression sites of *Trypanosoma brucei*. Mol Microbiol 36, 328-340.

[112] Navarro, M. and Cross, G.A.M. (1996). DNA rearrangements associated with multiple consecutive directed antigenic switches in *Trypanosoma brucei*. Mol Cell Biol 16, 3615-3625.

[113] Cross, G.A.M., Wirtz, L.E. and Navarro, M. (1998). Regulation of vsg expression site transcription and switching in Trypanosoma brucei. Mol Biochem Parasitol 91, 77-91.

[114] Zomerdijk, J.C., Ouellette, M., ten Asbroek, A.L., Kieft, R., Bommer, A.M., Clayton, C.E. and Borst, P. (1990). The promoter for a variant surface glycoprotein gene expression site in Trypanosoma brucei. Embo J 9, 2791-801.

[115] Kooter, J.M., van der Spek, H.J., Wagter, R., d'Oliveira, C.E., van der Hoeven, F., Johnson, P.J. and Borst, P. (1987). The anatomy and transcription of a telomeric expression site for variant-specific surface antigens in T. brucei. Cell 51, 261-72.

[116] Cully, D.F., Ip, H.S. and Cross, G.A.M. (1985). Coordinate transcription of variant surface glycoprotein genes and an expression site associated gene family in Trypanosoma brucei. Cell 42, 173-82.

[117] Cully, D.F., Gibbs, C.P. and Cross, G.A.M. (1986). Identification of proteins encoded by variant surface glycoprotein expression site-associated genes in Trypanosoma brucei. Mol Biochem Parasitol 21, 189-97.

[118] Schell, D., Evers, R., Preis, D., Ziegelbauer, K., Kiefer, H., Lottspeich, F., Cornelissen, A.W. and Overath, P. (1991). A transferrin-binding protein of Trypanosoma brucei is encoded by one of the genes in the variant surface glycoprotein gene expression site. EMBO J 10, 1061-6.

[119] Salmon, D., Geuskens, M., Hanocq, F., Hanocq-Quertier, J., Nolan, D., Ruben, L. and Pays, E. (1994). A novel heterodimeric transferrin receptor encoded by a pair of VSG expression site-associated genes in T. brucei. Cell 78, 75-86.

[120] van Luenen, H.G., Kieft, R., Mussmann, R., Engstler, M., ter Riet, B. and Borst, P. (2005). Trypanosomes change their transferrin receptor expression to allow effective uptake of host transferrin. Mol Microbiol 58, 151-65.

[121] Salmon, D., Paturiaux-Hanocq, F., Poelvoorde, P., Vanhamme, L. and Pays, E. (2005). Trypanosoma brucei: growth differences in different mammalian sera are not due to the species-specificity of transferrin. Exp Parasitol 109, 188-94.

[122] Barry, J.D. and McCulloch, R. (2001) Antigenic variation in trypanosomes: enhanced phenotypic variation in a eukaryotic parasite. In Advances in Parasitology, pp. 1-70. Academic Press Ltd, London.

[123] Robinson, N.P., Burman, N., Melville, S.E. and Barry, J.D. (1999). Predominance of duplicative VSG gene conversion in antigenic variation in African trypanosomes. Mol Cell Biol 19, 5839-5846.

[124] Bernards, A., Van der Ploeg, L.H., Frasch, A.C., Borst, P., Boothroyd, J.C., Coleman, S. and Cross, G.A.M. (1981). Activation of trypanosome surface glycoprotein genes involves a duplication-transposition leading to an altered 3' end. Cell 27, 497-505.

[125] Pays, E., Van Meirvenne, N., Le Ray, D. and Steinert, M. (1981). Gene duplication and transposition linked to antigenic variation in Trypanosoma brucei. Proc Natl Acad Sci U S A 78, 2673-7.

[126] Rudenko, G., McCulloch, R., Dirksmulder, A. and Borst, P. (1996). Telomere exchange can be an important mechanism of variant surface glycoprotein gene switching in *Trypanosoma brucei*. Molecular & Biochemical Parasitology 80, 65-75.

[127] Pays, E., Van Assel, S., Laurent, M., Darville, M., Vervoort, T., Van Meirvenne, N. and Steinert, M. (1983). Gene conversion as a mechanism for antigenic variation in trypanosomes. Cell 34, 371-81.

[128] Myler, P.J., Allison, J., Agabian, N. and Stuart, K. (1984). Antigenic variation in African trypanosomes by gene replacement or activation of alternate telomeres. Cell 39, 203-11.

[129] Turner, C.M.R. (1997). The rate of antigenic variation in fly-transmitted and syringe-passaged infections of *Trypanosoma brucei*. FEMS Microbiology Letters 153, 227-231.

[130] Iyer, L.M., Anantharaman, V., Wolf, M.Y. and Aravind, L. (2008). Comparative genomics of transcription factors and chromatin proteins in parasitic protists and other eukaryotes. Int J Parasitol 38, 1-31.

[131] Schlimme, W., Burri, M., Bender, K., Betschart, B. and Hecker, H. (1993). Trypanosoma brucei brucei: differences in the nuclear chromatin of bloodstream forms and procyclic culture forms. Parasitology 107, 237-47.

[132] Wirtz, E., Hoek, M. and Cross, G.A.M. (1998). Regulated processive transcription of chromatin by T7 RNA polymerase in Trypanosoma brucei. Nucleic Acids Res 26, 4626-34.

[133] Jenuwein, T., Forrester, W.C., Qiu, R.G. and Grosschedl, R. (1993). The immunoglobulin mu enhancer core establishes local factor access in nuclear chromatin independent of transcriptional stimulation. Genes Dev 7, 2016-32.

[134] Jenuwein, T., Forrester, W.C., Fernandez-Herrero, L.A., Laible, G., Dull, M. and Grosschedl, R. (1997). Extension of chromatin accessibility by nuclear matrix attachment regions. Nature 385, 269-72.

[135] Palenchar, J.B. and Bellofatto, V. (2006). Gene transcription in trypanosomes. Mol Biochem Parasitol 146, 135-41.

[136] Clayton, C.E. (2002). Life without transcriptional control? From fly to man and back again. EMBO J 21, 1881-1888.

[137] Campbell, D.A., Thomas, S. and Sturm, N.R. (2003). Transcription in kinetoplastid protozoa: why be normal? Microbes Infect 5, 1231-40.

[138] Myler, P.J. et al. (1999). Leishmania major Friedlin chromosome 1 has an unusual distribution of protein-coding genes. Proc Natl Acad Sci U S A 96, 2902-6.

[139] McDonagh, P.D., Myler, P.J. and Stuart, K. (2000). The unusual gene organization of Leishmania major chromosome 1 may reflect novel transcription processes. Nucleic Acids Res 28, 2800-3.

[140] Irmer, H. and Clayton, C. (2001). Degradation of the unstable EP1 mRNA in Trypanosoma brucei involves initial destruction of the 3'-untranslated region. Nucleic Acids Res 29, 4707-15.

[141] Gilinger, G. and Bellofatto, V. (2001). Trypanosome spliced leader RNA genes contain the first identified RNA polymerase II gene promoter in these organisms. Nucleic Acids Res 29, 1556-64.

[142] Bangs, J.D., Crain, P.F., Hashizume, T., McCloskey, J.A. and Boothroyd, J.C. (1992). Mass spectrometry of mRNA cap 4 from trypanosomatids reveals two novel nucleosides. J Biol Chem 267, 9805-15.

[143] Mair, G., Ullu, E. and Tschudi, C. (2000). Cotranscriptional cap 4 formation on the Trypanosoma brucei spliced leader RNA. J Biol Chem 275, 28994-9.

[144] Ullu, E., Matthews, K.R. and Tschudi, C. (1993). Temporal order of RNA-processing reactions in trypanosomes: rapid trans splicing precedes polyadenylation of newly synthesized tubulin transcripts. Mol Cell Biol 13, 720-5.

[145] LeBowitz, J.H., Smith, H.Q., Rusche, L. and Beverley, S.M. (1993). Coupling of poly(A) site selection and trans-splicing in Leishmania. Genes Dev 7, 996-1007.

[146] Smith, J.L., Levin, J.R., Ingles, C.J. and Agabian, N. (1989). In trypanosomes the homolog of the largest subunit of RNA polymerase II is encoded by two genes and has a highly unusual C-terminal domain structure. Cell 56, 815-27.

[147] Evers, R., Hammer, A., Kock, J., Jess, W., Borst, P., Memet, S. and Cornelissen, A.W. (1989). Trypanosoma brucei contains two RNA polymerase II largest subunit genes with an altered C-terminal domain. Cell 56, 585-97.

[148] Ivens, A.C. et al. (2005). The genome of the kinetoplastid parasite, *Leishmania major*. Science 309, 436-42.

[149] Navarro, M. and Cross, G.A.M. (1998). In situ analysis of a variant surface glycoprotein expression-site promoter region in Trypanosoma brucei. Mol Biochem Parasitol 94, 53-66.

[150] Rudenko, G., Blundell, P.A., Dirks-Mulder, A., Kieft, R. and Borst, P. (1995). A ribosomal DNA promoter replacing the promoter of a telomeric VSG gene expression site can be efficiently switched on and off in *T. brucei*. Cell 83, 547-553.

[151] Navarro, M. and Gull, K. (2001). A pol I transcriptional body associated with VSG mono-allelic expression in *Trypanosoma brucei*. Nature 414, 759-763.

[152] Pays, E., Lheureux, M. and Steinert, M. (1981). The expression-linked copy of a surface antigen gene in *Trypanosoma brucei* is probably the one transcribed. Nature 292, 265-7.

[153] Greaves, D.R. and Borst, P. (1987). *Trypanosoma brucei* variant-specific glycoprotein gene chromatin is sensitive to single-strand-specific endonuclease digestion. J Mol Biol 197, 471-83.

[154] Horn, D. and Cross, G.A.M. (1995). A developmentally regulated position effect at a telomeric locus in *Trypanosoma brucei*. Cell 83, 555-561.

[155] Horn, D. and Cross, G.A.M. (1997). Position-dependent and promoter-specific regulation of gene expression in *Trypanosoma brucei*. EMBO J 16, 7422-7431.

[156] Navarro, M., Cross, G.A.M. and Wirtz, E. (1999). *Trypanosoma brucei* variant surface glycoprotein regulation involves coupled activation/inactivation and chromatin remodeling of expression sites. EMBO J 18, 2265-2272.

[157] Hughes, K., Wand, M., Foulston, L., Young, R., Harley, K., Terry, S., Ersfeld, K. and Rudenko, G. (2007). A novel ISWI is involved in VSG expression site downregulation in African trypanosomes. EMBO J 26, 2400-10.

[158] Munoz-Jordan, J.L., Davies, K.P. and Cross, G.A.M. (1996). Stable expression of mosaic coats of variant surface glycoproteins in *Trypanosoma brucei*. Science 272, 1795-1797.

[159] Chaves, I., Rudenko, G., Dirks-Mulder, A., Cross, M. and Borst, P. (1999). Control of variant surface glycoprotein gene-expression sites in *Trypanosoma brucei*. EMBO J 18, 4846-4855.

[160] Glover, L. and Horn, D. (2006). Repression of polymerase I-mediated gene expression at Trypanosoma brucei telomeres. EMBO Rep 7, 93-9.

[161] Glover, L., Alsford, S., Beattie, C. and Horn, D. (2007). Deletion of a trypanosome telomere leads to loss of silencing and progressive loss of terminal DNA in the absence of cell cycle arrest. Nucleic Acids Res 35, 872-80.

[162] Dreesen, O. and Cross, G.A.M. (2006). Consequences of telomere shortening at an active VSG expression site in telomerase-deficient Trypanosoma brucei. Eukaryot Cell 5, 2114-9.

[163] Bernards, A., Michels, P.A., Lincke, C.R. and Borst, P. (1983). Growth of chromosome ends in multiplying trypanosomes. Nature 303, 592-7.

[164] Pays, E., Laurent, M., Delinte, K., Van Meirvenne, N. and Steinert, M. (1983). Differential size variations between transcriptionally active and inactive telomeres of Trypanosoma brucei. Nucleic Acids Res 11, 8137-47.

[165] Janzen, C.J., Fernandez, J.P., Deng, H., Diaz, R., Hake, S.B. and Cross, G.A.M. (2006). Unusual histone modifications in *Trypanosoma brucei*. FEBS Lett 580, 2306-10.

[166] Janzen, C.J., Hake, S.B., Lowell, J.E. and Cross, G.A.M. (2006). Selective di- or trimethylation of histone H3 lysine 76 by two DOT1 homologs is important for cell cycle regulation in *Trypanosoma brucei*. Mol Cell 23, 497-507.

[167] van Leeuwen, F., Gafken, P.R. and Gottschling, D.E. (2002). Dot1p modulates silencing in yeast by methylation of the nucleosome core. Cell 109, 745-56.

[168] Siegel, N., Kawahara, T., DeGrasse, J., Janzen, C.J., Horn, D. and Cross, G.A.M. (2007). Acetylation of histone H4K4 is cell-cycle regulated and mediated by HAT3 in *Trypanosoma brucei.* Mol Microbiol in press.

[169] Ingram, A.K. and Horn, D. (2002). Histone deacetylases in Trypanosoma brucei: two are essential and another is required for normal cell cycle progression. Mol Microbiol 45, 89-97.

[170] Alsford, S., Kawahara, T., Isamah, C. and Horn, D. (2007). A sirtuin in the African trypanosome is involved in both DNA repair and telomeric gene silencing but is not required for antigenic variation. Mol Microbiol 63, 724-36.

[171] Hoek, M. (2001) Functional Analysis of *Trypanosoma brucei* Expression-Site-Associated-Gene 8, pp. 204. The Rockefeller University, New York City.

[172] Garcia-Salcedo, J.A., Gijon, P., Nolan, D.P., Tebabi, P. and Pays, E. (2003). A chromosomal SIR2 homologue with both histone NAD-dependent ADP-ribosyltransferase and deacetylase activities is involved in DNA repair in *Trypanosoma brucei*. EMBO J 22, 5851-62.

[173] Lowell, J.E. and Cross, G.A.M. (2004). A variant histone H3 is enriched at telomeres in Trypanosoma brucei. J Cell Sci 117, 5937-47.

[174] Lowell, J.E., Kaiser, F., Janzen, C.J. and Cross, G.A.M. (2005). Histone H2AZ dimerizes with a novel variant H2B and is enriched at repetitive DNA in Trypanosoma brucei. J Cell Sci 118, 5721-30.

[175] Rout, M.P. and Field, M.C. (2001). Isolation and characterization of subnuclear compartments from *Trypanosoma brucei*. Identification of a major repetitive nuclear lamina component. J Biol Chem 276, 38261-71.

[176] Jin, M., Bateup, H., Padovan, J.C., Greengard, P., Nairn, A.C. and Chait, B.T. (2005). Quantitative analysis of protein phosphorylation in mouse brain by hypothesis-driven multistage mass spectrometry. Anal Chem 77, 7845-51.

[177] Hamana, K. and Iwai, K. (1976). Effects of triton X-100 on gel electrophoresis and gel chromatography of histones. Possible binding to helical regions. J Biochem 79, 125-9.

[178] Harlow, E. and Lane, D. (1999) Using Antibodies: a Laboratory Manual, Cold Spring Harbor Laboratory Press. New York.

[179] Wirtz, E., Leal, S., Ochatt, C., and Cross, G.A.M. (1999). A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in Trypanosoma brucei. Mol Biochem Parasitol 99, 89-101.

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

[181] da Cunha, J.P., Nakayasu, E.S., de Almeida, I.C. and Schenkman, S. (2006). Post-translational modifications of *Trypanosoma cruzi* histone H4. Mol Biochem Parasitol 150, 268-277.

[182] West, M.H. and Bonner, W.M. (1980). Histone 2B can be modified by the attachment of ubiquitin. Nucleic Acids Res 8, 4671-80.

[183] Goldknopf, I.L. and Busch, H. (1977). Isopeptide linkage between nonhistone and histone 2A polypeptides of chromosomal conjugate-protein A24. Proc Natl Acad Sci U S A 74, 864-8.

[184] Chu, F., Nusinow, D.A., Chalkley, R.J., Plath, K., Panning, B. and Burlingame, A.L. (2006). Mapping post-translational modifications of the histone variant MacroH2A1 using tandem mass spectrometry. Mol. Cell Proteomics 5, 194-203.

[185] Khandke, K.M., Fairwell, T., Chait, B.T. and Manjula, B.N. (1989). Influence of ions on cyclization of the amino terminal glutamine residues of tryptic peptides of streptococcal PepM49 protein. Resolution of cyclized peptides by HPLC and characterization by mass spectrometry. Int J Pept Protein Res 34, 118-23.

[186] Fischle, W., Wang, Y. and Allis, C.D. (2003). Binary switches and modification cassettes in histone biology and beyond. Nature 425, 475-9.

[187] Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R.S. and Zhang, Y. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science 298, 1039-43.

[188] Krogan, N.J. et al. (2003). Methylation of histone H3 by Set2 in Saccharomyces cerevisiae is linked to transcriptional elongation by RNA polymerase II. Mol Cell Biol 23, 4207-18.

[189] Kizer, K.O., Phatnani, H.P., Shibata, Y., Hall, H., Greenleaf, A.L. and Strahl, B.D. (2005). A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation. Mol Cell Biol 25, 3305-16.

[190] Sarg, B., Helliger, W., Talasz, H., Koutzamani, E. and Lindner, H.H. (2004). Histone H4 hyperacetylation precludes histone H4 lysine 20 trimethylation. J Biol Chem 279, 53458-64.

[191] Talasz, H., Lindner, H.H., Sarg, B. and Helliger, W. (2005). Histone H4-lysine 20 monomethylation is increased in promoter and coding regions of active genes and correlates with hyperacetylation. J Biol Chem 280, 38814-22.

[192] Litt, M.D., Simpson, M., Gaszner, M., Allis, C.D. and Felsenfeld, G. (2001). Correlation between histone lysine methylation and developmental changes at the chicken beta-globin locus. Science 293, 2453-5.

[193] Verreault, A., Kaufman, P.D., Kobayashi, R. and Stillman, B. (1998). Nucleosomal DNA regulates the core-histone-binding subunit of the human Hat1 acetyltransferase. Curr Biol 8, 96-108.

[194] Baarends, W.M., Hoogerbrugge, J.W., Roest, H.P., Ooms, M., Vreeburg, J., Hoeijmakers, J.H. and Grootegoed, J.A. (1999). Histone ubiquitination and chromatin remodeling in mouse spermatogenesis. Dev Biol 207, 322-33.

[195] Wang, H., Wang, L., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R.S. and Zhang, Y. (2004). Role of histone H2A ubiquitination in Polycomb silencing. Nature 431, 873-8.

[196] Sims, R.J., 3rd, Nishioka, K. and Reinberg, D. (2003). Histone lysine methylation: a signature for chromatin function. Trends Genet 19, 629-39.

[197] Garcia-Salcedo, J.A., Gijon, P. and Pays, E. (1999). Regulated transcription of the histone H2B genes of Trypanosoma brucei. Eur J Biochem 264, 717-23.

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