Published in final edited form as:

J Proteome Res. 2010 January ; 9(1): 585–594. doi:10.1021/pr900864s.

MS/MS/MS reveals false positive identification of histone serine methylation

Junmei Zhang¹, Yue Chen², Zhihong Zhang², Gang Xing¹, Joanna Wysocka³, and Yingming Zhao^{2,‡}

¹Protein Chemistry Technology Center, University of Texas Southwestern Medical Center at Dallas, Dallas, TX 75390-8816

²Ben May Department for Cancer Research, The University of Chicago, Chicago, IL 60637

³Department of Chemical and Systems Biology, Stanford University School of Medicine, Stanford, CA 94305

Abstract

Methylation of lysine and arginine residues is known to play a key role in regulating histone structure and function. However, methylation of other amino acid residues in histones has not been previously described. Using exhaustive nano-HPLC/MS/MS and blind protein sequence database searches, we tentatively assigned methylation to serine 28 of histone H3 from calf thymus. The assignment was in agreement with our stringent manual verification rules, co-elution in HPLC/MS/MS with its corresponding synthetic peptide, the dynamic nature of such methylation in distinct cell lines, and isotopic labeling. However, careful inspection of the MS/MS and MS/MS/MS spectra of a series of synthetic peptides confirmed that methylation actually occurs on K27 rather than on S28. The misassignment was caused by the fact that the ($y_9 + 14$) of the putative S28-methylated peptide and ($b_9 + 18$) ions of the K27 methylated peptide share the same m/z value (m/z 801). This MS/MS peak was used as the major evidence to assign methylation to S28 (consecutive y_8 and ($y_9 + 14$) ions). MS/MS/MS analysis revealed the false positive nature of serine methylation: the ambiguous ion at m/z 801 is indeed ($b_9 + 18$), an ion resulting from an *in vitro* reaction in the gas phase during collisionally activated dissociation (CAD). When lysine (K27) was acetylated, the degree of such *in vitro* reactions was greatly reduced, and such reactions were completely eliminated when the C-terminus was blocked by carboxylic group derivatization. Moreover, such side-chain assisted C-terminal rearrangement was found to be charge dependent. In aggregate, these results suggest that extra caution should be taken in interpretation of post-translational modification (PTM) data and that MS/MS as well as MS/MS/MS of synthetic peptides are needed for verifying the identity of peptides bearing a novel PTM.

Keywords

MS/MS/MS analysis; false positive identification; protein methylation; charge dependence; side-chain assisted C-terminal rearrangement; C-terminal elimination; loss of C-terminal; histone modifications

‡Correspondence: Dr. Yingming Zhao, Ben May Department for Cancer Research, The University of Chicago, 924 E 57th St, Knapp R120, Chicago, IL 60637 Phone: (773) 834-1561 Yingming.Zhao@uchicago.edu.

Supporting Information Available: This material is available free of charge via the Internet at <http://pubs.acs.org>.

Introduction

Protein methylation comprises a major group of protein post-translational modifications (PTMs). In histones, methylation of lysine and arginine residues plays important roles in regulating transcription, maintaining genomic integrity, contributing to epigenetic memory, and regulating diseases.^{1,2,3,4} In addition to methylation at lysine and arginine, methylation at other amino acid residues such as aspartate, glutamate, histidine, asparagine, glutamine, and cysteine has been reported. Nevertheless, substrates and functions of these modifications have not yet been carefully examined.⁵ Recently, we identified methylation at the side chains of aspartate and glutamate, suggesting the presence of these two modifications in eukaryotic cells.⁶

Since a variety of amino acids can be methylated,^{5,7} special care should be taken in experimental design and interpretation of MS/MS data. Ong *et al* described a metabolic labeling strategy using SILAC (stable isotope labeling by amino acids in cell culture) to label methylated proteins and to facilitate their identification and quantification.⁸ Heavy methyl SILAC is a general approach that can be used for identification of methylated residues in circumstances where the methylation is catalyzed by a S-adenosyl-L-methionine (SAM)-dependent methyltransferase. The existence of potential false positives also demands careful manual verification of MS/MS data, exclusive localization of methylation sites,^{9,10} and confirmation of peptide identification with synthetic peptides and HPLC co-elution.

In addition to MS/MS, multi-stage mass spectrometry analysis such as MS/MS/MS (MS3) in an ion trap mass spectrometer has been applied to improve accuracy of peptide identification in proteomic analysis.^{11,12,13,14} MS3 helps resolve ambiguity in sequence alignments resulting from overlapping fragment ions in low resolution MS/MS data.¹¹ In phosphor-proteomics studies, neutral-loss MS3 analysis allows further fragmentation of the peptide backbone and facilitates accurate localization of modification sites.^{15,16,17} MS3 analysis has also been applied in top-down protein identification and provides an additional level of evidence for confident peptide identification.¹⁸

Here, we present a case study that used MS3 analysis to reveal a false-positive identification of serine methylation in histone H3. The false-positive identification cannot be distinguished by common verification approaches including MS/MS of a synthetic peptide, HPLC co-elution, stable isotope labeling with heavy-isotope labeled SAM, dynamic status in different cell lines, or high resolution MS/MS analysis. Careful MS/MS/MS analysis revealed that the false positive PTM identification was caused by a side-chain assisted C-terminal rearrangement which occurred during collisionally activated dissociation (CAD). More importantly, such rearrangement is highly charge dependent, which has not been reported previously according to the best of our knowledge. The study highlights the importance of careful evaluation of PTM identification results and suggests that MS3 analysis is a valuable tool in the verification of novel protein modifications.

Materials and Methods

Preparation of core histones from calf thymus

Core histones were prepared according to a procedure described previously.¹⁹ About 20 g of fat-free calf thymus were sliced into 1–2 cm³ cubes, soaked in 16 mL of 0.5 M sucrose solution for 3 min, and subsequently mixed with 144 mL of homogenization buffer (0.25 M sucrose, 3.3 mM CaCl₂). The mixture was homogenized for 30 sec twice in an Oster 12-speed blender at the lowest speed setting. The homogenate was then filtered through two layers of cheesecloth, and the filtrate was centrifuged at 1,000 × g for 10 min to obtain cell pellets. The pellets were resuspended in 4 volumes of hypotonic buffer (50 mM Tris-Cl, pH 7.9, 2.5 mM MgCl₂, 10

mM KCl, 0.5 mM DTT, and 0.5 mM PMSF) and slowly stirred for 30 min. The suspension was centrifuged at $1,600 \times g$ for 10 min to collect pelleted nuclei. The core histones were then extracted twice using 3–4 volumes of 0.4 N H_2SO_4 overnight followed by centrifugation at $22,000 \times g$. The extract was dialyzed sequentially against H_2O and 50 mM Tris buffer (pH 7.3) for 8 h each. The core histone preparation was then subjected to HPLC separation using a C4 column. Finally, each core histone protein peak was collected, dried in a SpeedVac, and dissolved in water.

Isolation of core histones from cultured cells

Cells were harvested by centrifugation, washed twice with ice-cold phosphate buffered saline (PBS, Mediatech, Herndon, VA) containing 5 mM sodium butyrate, and lysed in Triton extraction buffer (TEB: PBS containing 0.5% Triton X-100, 0.1 M PMSF, 50 mM sodium butyrate and 30 mM nicotinamide). After centrifugation, the supernatant was discarded; the nuclear pellet was washed again with TEB and the core histones were extracted with 0.4 N H_2SO_4 on ice overnight. After centrifugation, histones were precipitated by trichloroacetic acid precipitation method. Histone pellets were collected by centrifugation, and washed sequentially with acidified acetone (0.1% HCl in acetone) followed by two more washes with acetone. After drying at room temperature for 5–15 min, the pellets were dissolved in water. The histones were then separated by SDS-PAGE and stained with Coomassie blue.

Isotopic labeling of HeLa cells

HeLa cells were grown in DMEM culture medium (Mediatech, Herndon, VA) for 3 days. The isotopic labeling experiment using a heavy form of S-adenosylmethionine (^{13}C , 2H -labeled SAM, Sigma-Aldrich, St. Louis, MO) was carried out as we previously described.²⁰ Core histones from the cells were extracted using the method described above.

In-gel Digestion

Protein bands of interest were destained in a destaining solution (ethanol/water (50%:50%, v/v)) and then with water for twenty minutes. The protein bands were cut into 1 mm^3 cubes, dehydrated in acetonitrile and dried in a SpeedVac. The dried gel pieces were rehydrated and covered with 50 mM ammonium bicarbonate solution containing 10 ng/ μL trypsin and subjected to overnight digestion at 37°C . The resulting peptides were cleaned with C_{18} ZipTips (Millipore, Bedford, MA) according to the manufacturer's instructions prior to nano-HPLC/mass spectrometric analysis.

HPLC-MS Analysis

Each sample was dissolved in 4 μL of HPLC buffer A (0.1% formic acid/2% acetonitrile/97.9% H_2O (v/v/v)) and 1 μL was injected into the Agilent 1100 nano flow HPLC system. Mass analysis was performed on a LTQ – 2D ion trap spectrometer (ThermoFisher Scientific, San Jose, CA) equipped with a nano-electrospray ionization source. The capillary column (10 cm length $\times$ 75 μm ID) was home packed with Luna C_{18} resin (5 μm particle size, 100 $\text{\AA}$ pore diameter) (Phenomenex, Torrance, CA). Peptides were eluted from the column using a gradient from 8% to 90% buffer B (0.1% formic acid/90% acetonitrile/9.9% H_2O (v/v/v)) in a two-hour cycle. The eluted peptides were directly electro-sprayed into the LTQ spectrometer with MS/MS spectra acquired in a data dependent mode that cycled between MS and MS/MS of the 10 strongest parent ions.

Protein sequence alignment and manual validation of peptide identifications

The LC/MS/MS dataset was searched against the corresponding protein sequence with PTMap,²¹ an in-house developed software, to identify all possible protein modifications. PTMap was specified to identify protein modifications with mass shifts ranging from -100 Da to $+200 \text{ Da}$,

in 1-Da increments. When searching, trypsin was specified as the proteolytic enzyme and 3 missing cleavages were allowed. Mass errors of precursor and product ions were set at ± 4 Da and ± 0.6 Da, respectively. Each modification site was exclusively localized in the peptide sequence by PTMap.²¹ All peptide identifications were manually validated with high stringency according to previously published criteria.⁹

Peptide derivatization

The tryptic peptides of interest were reacted with 1-(2-pyrimidyl) piperazine (PP), 1-(3-Dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (EDC), and 1-hydroxy-7-azabenzotriazole (HOAt) (Sigma, St. Louis, MO) according to a procedure reported previously.²² The carboxylic acid group of the C-terminal (or acidic residues, aspartic acid and glutamic acid) can react readily with the secondary amine group of the piperazine to form an amide.
22 The PP-derivatized peptides were desalted with C18 ZipTips prior to LC/MS/MS analysis.

Results

Initial identification of Ser methylation through unrestrictive sequence alignment

To examine whether methylation at residues other than lysine and arginine exist in core histones, we performed exhaustive HPLC/MS/MS analysis of tryptic digests of histones along with blind protein sequence database searching. About 5 μ g core histones from calf thymus were resolved in SDS-PAGE gel. The protein bands that corresponded to the molecular weights of histones H1, H2A, H2B, H3, and H4 were excised from the gel and digested with trypsin. The tryptic peptides from each histone were analyzed in a nano-HPLC/LTQ mass spectrometer using a two-hour gradient for exhaustive peptide identification. The resulting MS/MS data were analyzed by PTMap, an algorithm enabling identification of all possible PTMs with high sensitivity and high accuracy.²¹

Protein methylation can be induced *in vitro*. For example, the presence of methanol in a protein or peptide sample may induce methylation of aspartate and glutamate residues.⁶ To prevent such *in vitro* methylation, methanol was avoided in each step of sample handling, including extraction of core histones from cells, in-gel digestion, and HPLC/MS/MS analysis.

The methylated peptide candidates identified by PTMap were manually verified using a stringent verification procedure described previously.⁹ Because protein methylation can potentially occur at amino acid residues with nucleophilic, polar side-chains, which account for about 50% of ribosomally coded amino acid residues, a methylation site must be exclusively located to avoid false identification. To meet this criterion, trustworthy identification of a methylated peptide requires the PTM site to be identified by consecutive ions in the b or y ion series, or by the simultaneous appearance of modified b and y ions in which the modified residue is the terminal residue of each fragment. By this approach, methylated peptides with ambiguous methylation sites were removed from further consideration. Use of a comprehensive validation procedure such as this is critical to ensure the accuracy of peptide identification.

A putative serine-methylated peptide was identified from histone H3 using this screening strategy (Fig. 1A). Careful inspection of the MS/MS spectrum verified the peptide identification based on three lines of evidence: (i) all major peaks in the MS/MS spectrum could be explained by the peptide sequence (Fig. 1A); (ii) a methylated serine residue (H3S28 in the histone H3 sequence) could be exclusively located by the consecutive y ions (y_8 , ($y_9 + 14$)); and (iii) the mass shift of the peptide containing H3S28 was not caused by a polymorphism, because the unmodified peptide containing H3S28 was also identified in the same sample (Fig. 1B).

Efforts to verify histone Ser methylation

MS/MS of a synthetic peptide and co-elution in HPLC/MS/MS—The chemical nature of an identified peptide can be confirmed by MS/MS of its corresponding synthetic peptide, a gold standard for verification of peptide identification and chemical identity. To confirm the initial identification of the H3S28 methylated peptide, we synthesized an identical peptide, with the same sequence and same methylated residue (all the synthetic peptides were obtained from Genemed Synthesis Inc, San Antonio, TX), and performed MS/MS of the synthetic peptide under the same HPLC conditions. The MS/MS spectrum of the synthetic peptide (Supporting Information F1A) matched exactly with that from the *in vivo*-derived peptide (Fig. 1 A) and matched the MS/MS spectrum from a mixture of the synthetic peptide and the *in vivo*-derived peptides (Supporting Information F1B). The SRM (selected reaction monitoring) chromatogram (Supporting Information F2A, B, and C) also showed that the elution times were almost identical among the three samples, providing additional evidence for the peptide identification.

Isotopic labeling of the methyl group—To further distinguish whether methylation of H3S28 occurs *in vivo* or *in vitro*, we performed stable isotope labeling using heavy isotope-labeled SAM in cell cultures.^{23,24,25} HeLa cells were cultured with media supplemented with ¹³CD₃-labeled SAM. Core histones were isolated by acid extraction and resolved by SDS-PAGE. Histone H3 was analyzed using the same method as described above. The H3S28-methylated tryptic peptide was again observed with a mass shift relative to the unmodified peptide by 18 Da when isolated from cells cultured with media containing ¹³CD₃-SAM (Supporting Information F3). These 14- and 18-Da increases correspond to addition of ¹²CH₃ and ¹³CD₃ to serine 28 of histone H3 (Fig. 1A and Supporting Information F3), respectively. These results provide additional evidence that serine methylation was not an *in vitro* modification that occurred during sample preparation under our experimental conditions.

Dynamic analysis of H3S28 methylation status in different cell lines—To determine if H3S28 methylation is present in histone H3 from other cell lines, we extracted core histones from A431, MCF7, and HeLa cells. Core histones were analyzed using the same procedure as used for analysis of calf thymus histones. The S28 methylation site was identified by tandem mass spectrometry in histone H3 from all three cell lines. We further performed semi-quantification analysis using the spectral counting method^{26,27} to evaluate the relative portion of peptides bearing Ser methylation. We postulated that different cell lines would show different levels of Ser methylation due to their distinct genetic background and epigenetic program. If the methylation reaction was an artifact arising *in vitro* during sample handling, then the methylation status should be comparable in each cell type, because the same procedures were used to identify H3S28 methylated peptides in all four cell types. Indeed, our analysis revealed that the methylation status differed among calf thymus, A431, MCF7, and HeLa cells (Supporting Information F4). These results suggested that the observed methylation was likely to be dynamic in these cells and the identified methylation was not an artifact arising from inappropriate sample preparation.

MS/MS/MS analysis reveals the false positive identification of the Ser-methylated peptide

Although the identity of H3S28 methylation was supported by stringent manual verification including MS/MS and co-elution of a synthetic peptide, isotopic labeling, and dynamic change of the PTM status, two observations raised the possibility of false identification of the serine methylated peptide: (i) The putative serine methylation site could not be identified at the +2 charge state. The peptide has two basic lysine residues. Accordingly, the ions at the +2 charge state would be dominant compared to the ions at the +1 charge state (Fig. 2A). (ii) This serine methylation was identified only in the peptide “K^{Me}SAPATGGVK” which has one missing

trypsin cleavage, but was not identified in the fully cleaved tryptic peptide “^{Me}SAPATGGVK” while the unmodified fully tryptic peptide was identified (Supporting Information F5).

Histone H3K27 methylation is a well characterized histone PTM. We suspected that the newly identified H3S28 methylation was actually an artifact resulting from H3K27 methylation. To investigate any possible misassignment of methylation site from MS/MS spectra, we carefully compared the fragmentation patterns of the two synthetic peptides, “^{Me}KSAPATGGVK” and “^K^{Me}SAPATGGVK”. At the +2 charge state, all the major fragment ions of the lysine methylated peptide could be assigned as b or y ions (Fig. 3A, 3B), except that a minor ion at m/z 801 (5–10% relative intensity) could not be assigned. This m/z value is the same as the m/z value of a methylated y9 ion “y9 + 14” (y9 has m/z 787) or of a “b9 + 18”. At the +1 charge state this ambiguous ion became much more dominant (70–80% relative intensity) (Fig. 3C) resulting in a fragmentation pattern that is very similar to that of a serine methylated peptide at the +1 charge state (Fig. 3C and Supporting Information F1A). This unusual phenomenon was not detected by the computer software and may have led to the false identification of H3S28 methylation when the MS/MS spectrum from the H3K27 methylated peptide at the +1 charge state was analyzed.

The appearance of the ambiguous ion at 801 m/z in the fragmentation of H3K27 methylated peptide initially led us to suspect a gas-phase methyl group transfer from the N-terminal Lys to its neighboring Ser, which has been reported previously.²⁸ To test the possibility that the ion at m/z 801 actually represented a “y9+14” ion, the singly charged synthetic peptide ^{Me}KSAPATGGVK was subjected to MS/MS/MS. The resulting MS/MS/MS spectrum of m/z 801 was significantly different from the spectrum expected for the fragment ion (^{Me}SAPATGGVK ((y9 + 14))). The MS/MS/MS spectrum could actually be assigned as a truncated peptide, ^{Me}KSAPATGGV, derived from the parent peptide by loss of the C-terminal residue lysine (Fig. 3D). These results indicate that m/z 801 should be assigned as (b9 + 18) for the parent peptide ^{Me}KSAPATGGVK, rather than y9 for ^K^{Me}SAPATGGVK. Because the two ions have the same elemental composition, even high-resolution MS/MS could not resolve this structure. Nonetheless, using MS/MS/MS analysis, we were able to conclusively establish that the identification of H3S28 methylation was a false-positive caused by unexpected fragmentation behavior of the H3K27 methylated peptide.

It should be pointed out that a portion of the unmodified peptide KSAPATGGVK also undergoes such *in vitro* reactions. When the fragment ion at m/z 787 of the unmodified peptide were subjected to MS/MS/MS, the resulting spectrum can be assigned as a mixture of KSAPATGGV (b9 + H₂O), and SAPATGGVK (y9) (Supporting Information F6).

The false positive is caused by a charge-dependent, side-chain assisted C-terminal rearrangement

Loss of internal amino acid residue(s)^{29,30,31} from protonated peptides (singly or doubly charged) as well as loss of the C-terminal residue from both protonated^{32,33,34,35–41} and metal cationized peptides^{42–45} have been reported for singly charged peptides that do not carry any post-translational modifications. Several mechanisms were proposed for such phenomena, and all of them involved formation of a cyclic b ion intermediate that reopens at preferential sites, which leads either to the loss of the C-terminal amino acid (resulting in a truncated peptide without the original C-terminal residue) or to the loss of internal amino acid(s) (with the observation of sequence scrambling in MS/MS). However, the mechanisms differ from each other by two key points: (1) how the cyclic b ion intermediate is formed, and (2) whether such an intermediate formation requires assistance of a non-C-terminal basic residue.

To determine if a similar mechanism was responsible for the observed fragmentation pattern of H3K27 methylated peptide, and to probe structural features that may promote such gas-

phase reactions, we carried out extensive studies of MS/MS and MS/MS/MS spectra of a series of synthetic peptides with modified N or C-termini.

Analysis of synthetic peptide ^{Ac}KSAPATGGVK—To test if the lysine residue and its basic, positive side chain were involved in the C-terminal rearrangement, we synthesized the peptide with an acetylated ε-amine group on the N-terminal lysine residue. There are two major differences between lysine acetylation and lysine mono-methylation. First, acetylation on the side chain of the lysine residue (not N-terminal amine) in this peptide eliminates the nucleophilicity of the lysine side chain, which is not the case for lysine mono-methylation. Second, an acetyl group is bulkier than a methyl group. These two factors would be expected to influence the rearrangement reaction if the N-terminal lysine residue is involved in formation of a cyclic b ion intermediate. As expected, the fragment ion ($b_9 + 18$) was not detectable in the MS/MS spectrum of the doubly charged precursor ion (Fig. 4B), and this ion is 3–4 times less abundant at the +1 charge state (Fig. 4C) than that of the peptide without lysine acetylation (Fig. 3C) (also see Table 1). MS/MS and MS/MS/MS of the peptide carrying both singly charged and doubly charged ions confirmed both the peptide sequence and the lysine acetylation (Fig 4B, 4C and 4D). In contrast, mutations at the residue adjacent to the N-terminal lysine had little effect on the rearrangement reaction, as demonstrated by the mass spectrometric data of the synthetic peptides ^{Me}KTAPATGGVK, ^{Me}KAAPATGGVK, and ^{Me}K^{Ac}SAPATGGVK (Table 1, Supporting Information F7–F9). Thus, our results demonstrate that both the nucleophilicity status and the steric group of the N-terminal lysine residue are key factors determining the extent of C-terminal elimination.

Mass spectrometric analysis of the synthetic peptide ^{Ac}KSAPATGGVK derivatized with 1-(2-pyrimidyl) piperazine (PP)—To determine if the C-terminal carboxyl group affects C-terminal rearrangement, we derivatized the C-terminal carboxylic acid of the peptide with 1-(2-pyrimidyl) piperazine (PP) (Fig. 5A).²² When the derivatized peptide was subjected to MS/MS, no fragment ion corresponding to ($b_9 + 18$) (m/z 829) was observed at either the +2 or +1 charge state (Fig. 5B and 5C, Table 1), clearly demonstrating that a free C-terminal carboxyl group was required for C-terminal elimination reactions.

Possible mechanism for the C-terminal rearrangement reaction—The fact that C-terminal elimination of the synthetic peptide ^{Ac}KSAPATGGVK was reduced 3–4 times at the +1 charge state compared to its methylated counterpart (Figs. 3–4, Table 1) indicates that the amine group of the N-terminal lysine side chain actively participates in the rearrangement reaction as shown in Scheme 1. In this scheme, a nucleophilic nitrogen of the side chain of a basic residue (the N-terminal lysine in this study) that is flexible enough to be in close proximity to the C-terminus attacks the C-terminal carboxylic carbon to form a cyclic intermediate stabilized by a salt bridge. The cyclic intermediate may undergo further rearrangements, leading to loss of the C-terminal residue (resulting in the observed ($b_9 + 18$) ion, and two small molecule products, CO and $R_n\text{-CH=NH}$).

Based on this mechanism, we can now explain the experimental observations that the extent of the C-terminal rearrangement is affected by both side-chain modification of the basic residue and peptide charge state. When the N-terminal lysine side chain becomes methylated, its nucleophilicity is increased due to the electron donor property of the added methyl group, which leads to enhanced C-terminal rearrangement (Fig. 3D vs. Supporting Information F6). In contrast, acetylation of the N-terminal lysine side chain greatly reduces its nucleophilicity, which then in turn is responsible for the significant decrease of C-terminal rearrangement for the acetylated peptide (Fig. 3C, Fig. 4C, Table 1). And the reason that such C-terminal elimination is more prominent in singly charged ions than in the corresponding doubly charged counterparts (Figs. 3–4, Supporting Information F7–F9, Table 1), is because one proton in a doubly charged ion has a good chance of staying on the N-terminal lysine side chain,

eliminating its nucleophilicity. In singly charged ions, however, the side chain of the N-terminal monomethylated lysine seems to remain largely free and nucleophilic.

We conclude that the MS/MS spectrum shown in Figure 1A was generated from the well known H3K27 methylated peptide and not from a H3S28 serine methylated peptide. A side-chain assisted C-terminal rearrangement reaction occurred during tandem mass spectrometry of the H3K27 methylated peptide, producing a truncated peptide ion ($b_9 + 18$) which happens to have the same m/z value of ($y_9 + 14$) (indicating serine methylation) (Fig. 3). This rearrangement reaction prefers both a free C-terminus and a free nucleophilic nitrogen flexible enough to be in close proximity to the C-terminus (Figs. 4 and 5, Table 1, Scheme 1). In addition, the phenomenon is charge dependent and occurs preferentially when the ions are in the +1 charge state (Figs. 3 and 4, Supporting Information F7–F9, Table 1). The proposed mechanism (Scheme 1) readily explains why we have never observed serine methylation at the +2 charge state, even though this peptide occurs primarily as a doubly charged ion rather than as a singly charged ion (Fig. 2). Because of their similar structure and hydrophobicity, the *in vivo* derived lysine methylated peptide co-elutes with the synthetic serine methylated peptide (Supporting Information F2). Our isotopic labeling experiment (Supporting Information F3) labeled the methyl group at H3K27 and not H3S28. Finally, the observed dynamic nature of methylated peptides (Supporting Information F4) was actually a characteristic of the well known H3K27 lysine methylation and was not a characteristic of H3S28 methylation.

Discussion

Mass spectrometry has become an indispensable tool for the identification of posttranslational modifications due to its unparalleled sensitivity and high speed. However, incomplete or ambiguous peptide fragmentation often results in false positive identification or misassignment of PTM sites. When non-restrictive protein sequence alignment is involved, the chance of making a false positive identification rises further due to increasing database size and highly homologous candidate sequences. The results reported here demonstrate an unusual case of false positive identification of a putative Ser methylated peptide that could be verified using (i) stringent manual verification, (ii) MS/MS of a synthetic peptide, (iii) co-elution in HPLC/MS/MS, (iv) metabolic labeling with stable-isotope labeled SAM, and (v) dynamic analysis in different cell lines. Nevertheless, it was only through MS/MS/MS analysis on ambiguous fragment ions, coupled with careful analysis of a series of synthetic peptides, that the false positive identification was revealed and shown to be caused by a C-terminal assisted gas-phase elimination reaction that occurred during collisionally activated dissociation (CAD).

Based on the proposed mechanism of such C-terminal rearrangement, a simple rule can be learned – when the peptide sequence matches “B...B” (where B refers to any basic residue such as Lys, Arg or His) and the N-terminal residue is modified (with the proviso that the modification does not eliminate its nucleophilicity), C-terminal rearrangement could result in the misassignment of PTM on the second residue from the N-terminus, if ($b_{n-1} + H_2O$) ions are not considered during unrestrictive sequence alignment. This phenomenon may be more prominent when the number of charges carried by the peptide is less than the number of available basic groups in the sequence. After incorporating this rule into updated PTMap software, we were able to completely eliminate this type of false positive PTM identification.

It should be noted that some commercial search engine, such as Mascot (www.matrixscience.com), also do not consider the possibility for formation of ($b_{n-1} + H_2O$) fragment ions. For example, when we used Mascot (version 2.2) to search the same data with both serine methylation and lysine methylation as variable modifications, it assigned the MS/MS spectrum of the singly charged precursor ion m/z 929 to serine methylated peptide KSAPATGGVK, and that of the doubly charged precursor m/z 465 to lysine methylated

version of the same peptide with very similar scores (Supporting Information F10). This suggests that the consideration of ($b_{n-1} + H_2O$) type of fragment ions and the charge state dependent nature of such ions should be incorporated into protein sequence alignment software in the future to improve the accuracy of peptide and PTM identification.

Our work highlights the possibility of PTM misassignment with almost perfect sequence alignment. As a result, one would have to examine mass spectrometric data very carefully and seek additional means of validation. Synthetic peptides may still serve as a gold standard if both MS/MS and MS/MS/MS spectra of the synthetic peptides agree with those of the *in vivo* derived peptides of interest at all charge states. Our case study demonstrates that multi-stage mass spectrometry analysis can be a powerful approach for revealing the identity of true peptide sequences/modification sites. In view of the increased application of blind sequence alignment of mass spectrometric data, the significance of the results from this study should not be overlooked.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by a grant from the NIH (CA126832 to Y.Z.).

Abbreviations

H3K27	lysine 27 of histone H3
H3S28	serine 28 of histone H3
HPLC	high performance liquid chromatography
MS	mass spectrometry
MS/MS	tandem mass spectrometry or mass spectrometry/mass spectrometry
CAD	collisionally activated dissociation
PBS	phosphate-buffered saline
PTM	protein post-translational modification
SAM	S-adenosyl-L-methionine
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SRM	selected reaction monitoring
TEB	Triton extraction buffer
TFA	trifluoroacetic acid

References

1. Strahl BD, Allis CD. The language of covalent histone modifications. *Nature* 2000;403(6765):41–45. [PubMed: 10638745]
2. Ng SS, Yue WW, Oppermann U, Klose RJ. Dynamic protein methylation in chromatin biology. *Cellular and Molecular Life Sciences* 2009;66(3):407–422. [PubMed: 18923809]
3. Martin C, Zhang Y. The diverse functions of histone lysine methylation. *Nature Reviews Molecular Cell Biology* 2005;6(11):838–849.

4. Bedford MT. Arginine methylation at a glance. *Journal of Cell Science* 2007;120(24):4243–4246. [PubMed: 18057026]
5. Walsh CT, Garneau-Tsodikova S, Gatto GJ. Protein posttranslational modifications: The chemistry of proteome diversifications. *Angew. Chem.-Int. Edit* 2005;44(45):7342–7372.
6. Sprung R, Chen Y, Zhang K, Cheng D, Zhang T, Peng J, Zhao Y. Identification and validation of eukaryotic aspartate and glutamate methylation in proteins. *Journal of Proteome Research* 2008;7(3): 1001–1006. [PubMed: 18220335]
7. Clarke S. Protein methylation. *Current Opinion in Cell Biology* 1993;5(6):977–983. [PubMed: 8129951]
8. Ong S-E, Mann M. Identifying and quantifying sites of protein methylation by heavy methyl SILAC. *Curr Protoc Protein Sci* 2006;Chapter 14:Unit 14.9. [PubMed: 18429299]
9. Chen Y, Kwon SW, Kim SC, Zhao YM. Integrated approach for manual evaluation of peptides identified by searching protein sequence databases with tandem mass spectra. *Journal of Proteome Research* 2005;4(3):998–1005. [PubMed: 15952748]
10. Chen Y, Zhang J, Xing G, Zhao Y. Mascot-derived false positive peptide identifications revealed by manual analysis of tandem mass spectra. *Journal of Proteome Research* 2009;8:3141–3147. [PubMed: 19368407]
11. Olsen JV, Mann M. Improved peptide identification in proteomics by two consecutive stages of mass spectrometric fragmentation. *Proceedings of the National Academy of Sciences of the United States of America* 2004;101(37):13417–13422. [PubMed: 15347803]
12. Bandeira N, Olsen JV, Mann JV, Mann M, Pevzner PA. Multi-spectra peptide sequencing and its applications to multistage mass spectrometry. *Bioinformatics* 2008;24(13):i416–23. [PubMed: 18785330]
13. Mann K, Poustka AJ, Mann M. In-depth, high-accuracy proteomics of sea urchin tooth organic matrix. *Proteome Science* 2008;6
14. Schenk S, Schoenhals GJ, de Souza G, Mann M. A high confidence, manually validated human blood plasma protein reference set. *BMC Med Genomics* 2008;1:41. [PubMed: 18793429]
15. Olsen JV, Blagoev B, Gnad F, Macek B, Kumar C, Mortensen P, Mann M. Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* 2006;127(3):635–648. [PubMed: 17081983]
16. Lee J, Xu YD, Chen Y, Sprung R, Kim SC, Xie SH, Zhao YM. Mitochondrial phosphoproteome revealed by an improved IMAC method and MS/MS/MS. *Molecular & Cellular Proteomics* 2007;6(4):669–676. [PubMed: 17208939]
17. Ulintz PJ, Bodenmiller B, Andrews PC, Aebersold R, Nesvizhskii AI. Investigating MS2/MS3 matching statistics. *Molecular & Cellular Proteomics* 2008;7(1):71–87. [PubMed: 17872894]
18. Macek B, Waanders LF, Olsen JV, Mann M. Top-down protein sequencing and MS3 on a hybrid linear quadrupole ion trap-orbitrap mass spectrometer. *Molecular & Cellular Proteomics* 2006;5(5): 949–958. [PubMed: 16478717]
19. Shechter D, Dormann HL, Allis CD, Hake SB. Extraction, purification and analysis of histones. *Nature Protocols* 2007;2(6):1445–1457.
20. Jung SY, Li YH, Wang Y, Chen Y, Zhao YM, Qin J. Complications in the assignment of 14 and 28 Da mass shift detected by mass spectrometry as in vivo methylation from endogenous proteins. *Analytical Chemistry* 2008;80(5):1721–1729. [PubMed: 18247584]
21. Chen Y, Chen W, Cobb MH, Zhao YM. PTMap-A sequence alignment software for unrestricted, accurate, and full-spectrum identification of post-translational modification sites. *Proceedings of the National Academy of Sciences of the United States of America* 2009;106(3):761–766. [PubMed: 19136633]
22. Xu YW, Zhang LJ, Lu HJ, Yang PY. Mass spectrometry analysis of phosphopeptides after peptide carboxy group derivatization. *Analytical Chemistry* 2008;80(21):8324–8328. [PubMed: 18821779]
23. Oda Y, Huang K, Cross FR, Cowburn D, Chait BT. Accurate quantitation of protein expression and site-specific phosphorylation. *Proceedings of the National Academy of Sciences of the United States of America* 1999;96(12):6591–6596. [PubMed: 10359756]
24. Ong SE, Mittler G, Mann M. Identifying and quantifying in vivo methylation sites by heavy methyl SILAC. *Nature Methods* 2004;1(2):119–126. [PubMed: 15782174]

25. Zhu HN, Pan SQ, Gu S, Bradbury EM, Chen X. Amino acid residue specific stable isotope labeling for quantitative proteomics. *Rapid Communications in Mass Spectrometry* 2002;16(22):2115–2123. [PubMed: 12415544]
26. Carvalho PC, Hewel J, Barbosa VC, Yates JR. Identifying differences in protein expression levels by spectral counting and feature selection. *Genetics and Molecular Research* 2008;7(2):342–356. [PubMed: 18551400]
27. Old WM, Meyer-Arendt K, Aveline-Wolf L, Pierce KG, Mendoza A, Sevinsky JR, Resing KA, Ahn NG. Comparison of label-free methods for quantifying human proteins by shotgun proteomics. *Molecular & Cellular Proteomics* 2005;4(10):1487–1502. [PubMed: 15979981]
28. Xiong L, Ping L, Yuan B, Wang Y. Methyl group migration during the fragmentation of singly charged ions of trimethyllysine-containing peptides: precaution of using MS/MS of singly charged ions for interrogating peptide methylation. *Journal of the American Society for Mass Spectrometry* 2009;20(6):1172–1181. [PubMed: 19303795]
29. Vachet RW, Bishop BM, Erickson BW, Glish GL. Novel peptide dissociation: Gas-phase intramolecular rearrangement of internal amino acid residues. *Journal of the American Chemical Society* 1997;119(24):5481–5488.
30. Harrison AG. Peptide sequence scrambling through cyclization of b(5) ions. *Journal of the American Society for Mass Spectrometry* 2008;19(12):1776–1780. [PubMed: 18639467]
31. Yague J, Paradela A, Ramos M, Ogueta S, Marina A, Barahona F, de Castro JAL, Vazquez J. Peptide rearrangement during quadrupole ion trap fragmentation: Added complexity to MS/MS spectra. *Analytical Chemistry* 2003;75(6):1524–1535. [PubMed: 12659218]
32. Thorne GC, Gaskell SJ. Elucidation of some fragmentations of small peptides using sequential mass spectrometry on a hybrid instrument. *Rapid Commun Mass Spectrom* 1989;3(7):217–21. [PubMed: 2520240]
33. Thorne GC, Ballard KD, Gaskell SJ. Metastable decomposition of peptide $[M + H]^+$ ions via rearrangement involving loss of the c-terminal amino-acid residue. *Journal of the American Society for Mass Spectrometry* 1990;1(3):249–257.
34. Ballard KD, Gaskell SJ. Intramolecular oxygen-18 isotopic exchange in the gas-phase observed during the tandem mass-spectrometric analysis of peptides. *Journal of the American Chemical Society* 1992;114(1):64–71.
35. Gonzalez J, Besada V, Garay H, Reyes O, Padron G, Tambara Y, Takao T, Shimonishi Y. Effect of the position of a basic amino acid on C-terminal rearrangement of protonated peptides upon collision-induced dissociation. *Journal of Mass Spectrometry* 1996;31(2):150–158. [PubMed: 8799268]
36. Tsaprailis G, Nair H, Somogyi A, Wysocki VH, Zhong WQ, Futrell JH, Summerfield SG, Gaskell SJ. Influence of secondary structure on the fragmentation of protonated peptides. *Journal of the American Chemical Society* 1999;121(22):5142–5154.
37. Vachet RW, Asam MR, Glish GL. Secondary interactions affecting the dissociation patterns of arginine-containing peptide ions. *Journal of the American Chemical Society* 1996;118(26):6252–6256.
38. Deery MJ, Summerfield SG, Buzy A, Jennings KR. A mechanism for the loss of 60 u from peptides containing an arginine residue at the C-terminus. *Journal of the American Society for Mass Spectrometry* 1997;8(3):253–261.
39. Dikler S, Kelly JW, Russell DH. Improving mass spectrometric sequencing of arginine-containing peptides by derivatization with acetylacetone. *Journal of Mass Spectrometry* 1997;32(12):1337–1349. [PubMed: 9423284]
40. Fang SP, Takao T, Satomi Y, Mo WJ, Shimonishi Y. Novel rearranged ions observed for protonated peptides via metastable decomposition in matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Journal of the American Society for Mass Spectrometry* 2000;11(4):345–351. [PubMed: 10757171]
41. She Y-M, Krokhn O, Spicer V, Loboda A, Garland G, Ens W, Standing KG, Westmore JB. Formation of $(b_{n-1} + H_2O)$ ions by collisional activation of maldi-formed peptide $[M + H]^+$ ions in a QqTOF mass spectrometer. *Journal of the American Society for Mass Spectrometry* 2007;18(6):1024–1037. [PubMed: 17418589]

42. Tang XJ, Ens W, Standing KG, Westmore JB. Daughter ion mass-spectra from cationized molecules of small oligopeptides in a reflecting time-of-flight mass-spectrometer. *Analytical Chemistry* 1988;60(17):1791–1799. [PubMed: 3232814]
43. Renner D, Spiteller G. Linked scan investigation of peptide degradation initiated by liquid secondary ion mass-spectrometry. *Biomedical and Environmental Mass Spectrometry* 1988;15(2):75–77. [PubMed: 3349211]
44. Grese RP, Cerny RL, Gross ML. Metal ion-peptide interactions in the gas phase: a tandem mass spectrometry study of alkali metal cationized peptides. *Journal of the American Chemical Society* 1989;111(8):2835–42.
45. Grese RP, Gross ML. Gas-phase interactions of lithium ions and dipeptides. *Journal of the American Chemical Society* 1990;112(13):5098–104.

Figure 1.

Initial assignment of methylation to a serine residue in histone H3 based on tandem mass spectrometry data. (A) The MS/MS spectrum used for initial identification of $K^{Me}SAPATGGVK$ with m/z 929.53; (B) Unmodified counterpart ($KSAPATGGVK$) from calf thymus histone H3. The labels “b” and “y” designate the N- and C-terminal fragment ions, respectively, of the peptide produced by collision-induced fragmentation at the peptide bond in the mass spectrometer. The label “a” designates N-terminal fragments produced by breakage at the C-C bond adjacent to the peptide bond. The subscripted number (e.g., b_9 , y_9) represents the number of N- or C-terminal residues present in the peptide fragment. The label “ Δ ”

designates “b”, “y” or “a” ions with water and/or ammonia loss. The label “*” when present designates the precursor ion. The same nomenclature is used for all subsequent figures.

Figure 2.

The MS spectra of a synthetic peptide $K^{Me}SAPATGGVK$. (A) Full scan; (B) MS/MS spectrum of doubly charged parent ion with m/z 465.60.

Figure 4.

The MS spectra of a synthetic peptide AcKSAPATGGVK. (A) Full scan; (B) MS/MS spectrum of doubly charged parent ion with m/z 479.78; (C) MS/MS spectrum of singly charged parent ion with m/z 957.52; (D) MS/MS/MS spectrum of singly charged fragment ion with m/z 829.42 (957.52 → 829.42 →).

Figure 5.

The MS spectra of a synthetic peptide AcKSAPATGGVK derivatized with 1-(2-pyrimidyl) piperazine (PP). (A) Full scan; (B) MS/MS spectrum of doubly charged parent ion with m/z 552.86; (C) MS/MS spectrum of singly charged parent ion with m/z 1103.64.

Scheme 1.

A proposed mechanism for the side-chain assisted C-terminal rearrangement observed in this study.

Table 1

Relative abundance of the C-terminal rearrangement ions ($b_{n-1} + 18$) of protonated synthetic peptides. All the peptides studied have ten ($n = 10$) amino acids.

Peptide Sequence	Charge State	Precursor m/z	Fragment Ions			
			$b_9 + 18$	b_9		
		m/z	Relative Abundance (%)	m/z	Relative abundance (%)	Rearrangement Ratio ($b_9 + 18$) / b_9
MeKSAPATGGVK	+1	929.5	76	783.5	47	1.62
	+2	465.3	7		60	0.12
KTAPATGGVK	+1	943.5	59	797.6	25	2.36
	+2	472.7	1		16	0.06
KAAPATGGVK	+1	913.5	71	767.5	33	2.15
	+2	457.3	6		53	0.11
MeK ^{Ac} SAPATGGVK	+1	971.5	64	825.5	25	2.56
	+2	486.3	6		57	0.11
AcKSAPATGGVK	+1	957.5	3	811.4	7	0.43
	+2	479.5	n. d. ¹		3	0.00
AcKSAPATGGVK-pp2	+1	1103.6	n. d.	811.5	42	0.00
	+2	552.9	n. d.		55	0.00

¹“n. d.” means not detected.

²The G terminal carboxylic acid group of the peptide AcKSAPATGGVK was derivatized with 1-(2-pyrimidyl) piperazine (PP).