Characterization of *Tetrahymena* Histone H2B Variants and Posttranslational Populations by Electron Capture Dissociation (ECD) Fourier Transform Ion Cyclotron Mass Spectrometry (FT-ICR MS)*

K. F. Medzihradszky‡, X. Zhang‡, R. J. Chalkley‡, S. Guan‡, M. A. McFarland§, M. J. Chalmers§, A. G. Marshall§, R. L. Diaz¶, C. D. Allis¶, and A. L. Burlingame‡‖

This work describes the nature and sequence information content of the electron capture dissociation mass spectra for the intact *Tetrahymena* histone H2B. Two major variants of this protein were present bearing nominal modifications of both +42 and +84 Da. This work describes identification of the nature of these two modifications. For example, using gas-phase selection and isolation of the +42-Da modified species, from a background of two H2B variants each present in six or more posttranslationally modified isoforms, we were able to determine that this +42-Da modification isoform bears trimethylation rather than acetylation. LC-CIDMS analysis was also employed on digested preparations to obtain complementary detail of the nature of site-specific posttranslational modifications. This study establishes that integration of the information from these two datasets provides a comprehensive map of posttranslational occupancy for each particular covalent assemblage selected for structural investigation. *Molecular & Cellular Proteomics 3: 872–886, 2004.*

Regulation of many cellular processes requires the controlled assembly and disassembly of protein complexes to carry out specific functions in certain cellular locations at defined times in concert with other members of a pathway or network (1). Hence, co-ordinated recruitment of new subunits to a protein machine to effect biological activity is required while maintaining other protein interactions that permit communication both upstream and downstream of any given protein machine in the pathway (2). Much of the regulation of biological activity is mediated by epigenetic processes that control the occupancy of posttranslational modification sites and hence assembly or communication via adapter or effector protein modules. While single modifications such as phosphorylation of tyrosine residues provide docking sites for SH2 domains of other proteins in the pathway (2–4), there is growing recognition that certain protein-protein interactions are governed by multidentate clusters of site-specific posttranslational modifications. Examples include histones (5), p53 (6), p130cas (7), and so on.

While there is still a need to explore and develop better methodology even for the widely studied posttranslational protein modification, phosphorylation, significant progress has in fact been made in the application of MS/MS for the detection and localization of phosphorylation sites (8, 9).

However, proteins may be modified with a plethora of different modifications, and these modifications are not necessarily mutually exclusive and independent of each other. For example, multisite phosphorylation can govern a wide range of effects (10). There is also clear interplay between phosphorylation and *O*-GlcNAc modification as many sites of these two modifications are close by or even on the same residue (11, 12). Histones display a wide repertoire of posttranslational modifications including methylation, dimethylation, trimethylation, acetylation, phosphorylation, and ubiquitination. Although functions for some of these modifications have been postulated, these modifications are believed to interact with each other such that combinations of modifications are proposed to act synergistically to effect regulation of gene expression (5). It is also interesting to note that *O*-GlcNAc transferase interacts with a histone deacetylase complex by binding to the corepressor mSin3A (13).

Current techniques focus on the identification of posttranslational modifications through analysis of the modified peptides in protein digests by MS/MS. However, in addition to possessing these effective methodologies to pinpoint sites of posttranslational modification unambiguously, there is a serious need for development of methods that can provide a

From the ‡Department of Pharmaceutical Chemistry and Mass Spectrometry Facility, University of California, San Francisco, CA 94143-0446; §Ion Cyclotron Resonance Program, National High Magnetic Field Laboratory, Florida State University, Tallahassee, FL 32310-3706; and ¶Laboratory of Chromatin Biology, Rockefeller University, New York, NY 10021

Received March 24, 2004, and in revised form June 9, 2004

Published, MCP Papers in Press, June 15, 2004, DOI 10.1074/mcp.M400041-MCP200

"bird's eye" mass balance of integrated posttranslational occupancy at the intact protein level. This includes having the ability to establish a scaffold of backbone covalent occupancy in a protein variant- or isoform-specific manner. Several groups have combined intact protein molecular mass analysis with peptide fragmentation analysis for comprehensive posttranslational modification analysis (14, 15), or even limited fragmentation at the protein level to determine the nature of posttranslational modifications prior to peptide analysis (16).

However, the recent discovery of electron capture dissociation (ECD)[1] (17) has opened a new era that will facilitate the direct structural characterization of intact proteins in the gas phase. Unlike other gas-phase ion dissociation methods, such as CID and infrared multiphoton dissociation, electron capture, and thus energy deposition and bond cleavage, occurs primarily at protonated peptide backbone bonds. Because the multiple protonation observed during ESI is statistically distributed along the entire length of the protein backbone, these peptide bond cleavages can provide comprehensive protein sequence information compared with collisional activation. In addition, nonergodic internal energy deposition does not vibronically randomize within adjacent residues prior to cleavage, so chemically labile substituents such as γ-carboxyglutamic acid remain intact (18). Thus, observation of the fragmentation patterns from these rather extensive ECD peptide bond cleavages permits determination of modification sites from the intact protein (19). Finally, an additional advantage of FT-ICR MS is its ability to isolate individual protein components at high resolution by stored waveform inverse FT axial excitation/ejection (SWIFT) (20) in an ICR trap prior to fragmentation analysis induced by thermal electron capture.

We wish to report studies of a histone H2B preparation isolated from *Tetrahymena* using these FT-ICR MS-based technologies together with corresponding capillary HPLC ESI tandem mass spectrometric analyses of proteolytic digests with a QqTOF geometry instrument. Histone H2B is expressed as two closely related proteins H2B.1 and H2B.2, differing by only three amino acids (21). The only posttranslational modification reported about the *Tetrahymena* proteins is the trimethylation of their N-terminal alanine (22). However, the bovine histone H2B is reportedly acetylated at Lys-5, -12, -15, and -20 as well as ubiquitinated at Lys-120. It was also recently published that the bovine protein is methylated and extensively acetylated in its C-terminal region (23). However, these assignments were based on peptide mass measurements alone and were not established unambiguously by fragmentation analysis.

Information derived from intact protein MS and MS/MS analysis combined with the more classical peptide LC-MS analysis approach are complementary, and we demonstrate that combining these two data types provides a more complete map of the occupancy of each particular covalent assemblage. Results are presented from three different mass spectrometers: protein digests were analyzed with a QSTAR (QqTOF) mass spectrometer; measurement of intact protein molecular masses was carried out using an LTQ-FT mass spectrometer. Also CID spectra produced by protein fragmentation in the linear ion trap were recorded using the LTQ-FT mass spectrometer. However, this instrument currently does not have the ability to perform ECD fragmentation in the FT-ICR cell. Hence, all ECD data were acquired at a later date on a home-built 9.4T FT-ICR mass spectrometer (24).

TABLE I
Proteins detected in the histone H2B sample

	Deconvolved MH^{+}[a]	$MH^{+}_{calculated}$	Tentative assignment[b]
1.	13467.5	13464.5	H2B.1 + 14 Da[c]
2.	13480.5	13478.5	H2B.1 + 28 Da
3.	13494.5[d]	13492.5	H2B.1 + 42 Da and/or
		13496.5	H2B.2
4.	13511.5	13510.5	H2B.2 + 14 Da and/or
		13506.5	H2B.1 + 56 Da
5.	13525.5	13524.5	H2B.2 + 28 Da and/or
		13520.5	H2B.1 + 70 Da
6.	13538.5	13538.5	H2B.2 + 42 Da and/or
		13534.5	H2B.1 + 84 Da
7.	13554.5	13552.5	H2B.2 + 56 Da
8.	13568.5	13566.5	H2B.2 + 70 Da
9.	13581.5	13580.5	H2B.2 + 84 Da
10.	13594.5	13594.5	H2B.2 + 98 Da
11.	13607.5	13608.5	H2B.2 + 112 Da
12.	13621.5	13622.5	H2B.2 + 126 Da
13.	13638.5	13636.5	H2B.2 + 140 Da

[a] Automated deconvolution program used—the identification of the monoisotopic peak was problematic due to the overlapping isotope clusters (see Fig. 1).

[b] Assignments are of the major protein species. Other modified species may also be present at lower levels.

[c] Histone H2B has two isoforms (see sequences in Scheme 1).

[d] Starting from here the isotope clusters definitely represent multiple components (see Fig. 2). These may be isobaric, *i.e.* acetyl *vs*. Me_3 or overlapping clusters of the two isoforms. The major (most abundant) components are underlined.

EXPERIMENTAL PROCEDURES

Histone H2B Purification—Tetrahymena thermophila strains CU427 and CU428, kindly provided by Peter Bruns (Cornell University, Ithaca, NY), were grown to log phase (cell density, 2.0–2.5 × 10^5 cells/ml) under standard conditions as described (25) and then starved overnight in 10 mM Tris-HCl (pH 7.4) at 30 °C gently shaking (50 rpm). Highly purified macronuclei were prepared as described (25) except for the addition of 1 mM PMSF and 10 mM butyric acid as well as the omission of spermidine from the isolation buffer. Macronuclei were washed in pH 7.5 nucleus wash buffer (250 mM sucrose, 10 mM Tris, 3 mM calcium chloride, 1 mM magnesium chloride, 10 mM butyric acid, 1 mM PMSF) and either used immediately or stored at −80 °C. Total histones from starved macronuclei were TCA-extracted, dissolved in deionized water, and chromatographed on a C8 column (220 × 4.6 mm Aquapore RP-300; Perkin Elmer, Wellesley, MA). Histone H2B was eluted at 1.0 ml/min with a linear ascending gradient

[1] The abbreviations used are: ECD, electron capture dissociation; SWIFT, stored waveform inverse FT axial excitation/ejection.

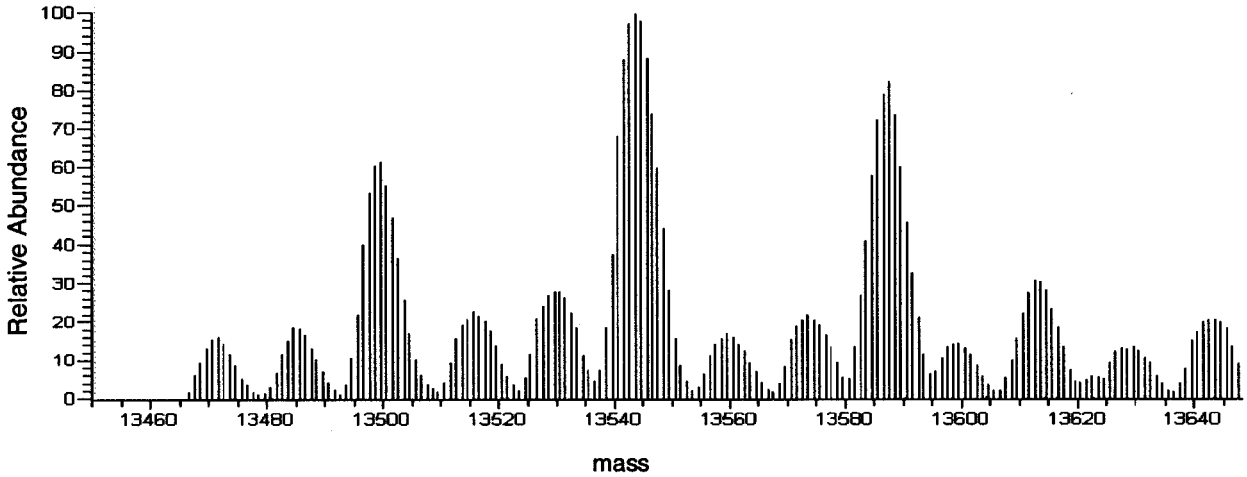


FIG. 1. **Deconvolved mass spectrum of H2B analyzed by FT-ICR MS.** Components of the mixture are listed in Table I.

of 35–60% ACN and 0.1% TFA over 55 min on a Waters 625LC (Millipore, Bedford, MA).

Proteolytic Digestions—An aliquot of the histone H2B sample (in 25 mM ammonium bicarbonate buffer, pH ~7.8) was denatured by boiling prior to the digestion, then incubated for 2 h at 37 °C with 40 ng trypsin (porcine, side-chain protected; Promega, Madison, WI) or 16 ng endoproteinase AspN (*Pseudomonas fragi*, sequencing grade; Roche, Indianapolis, IN), respectively. The digestions were terminated by acidifying the mixtures.

Capillary HPLC-ESI MS/MS—The digests were analyzed by LC/MS using a C18 PepMap 75-μm × 150-mm column on an Eksigent nanoHPLC pump (Eksigent, Livermore, CA) linked with a FAMOS autosampler (LC Packings, San Francisco, CA). Solvent A was 0.1% formic acid in water, and B was 0.1% formic acid in ACN, at a flow rate of ~300 nl/min. One microliter of the digests was injected at 5% B, then the organic content of the mobile phase was increased linearly to 50% over 30 min. The column effluent was directed to a QSTAR Pulsar tandem mass spectrometer of QqTOF geometry (Applied Biosystems/MDS Sciex, Toronto, CA). Throughout the chromatographic separation, 1-s MS acquisitions were followed by two 3-s CID experiments for computer-selected precursor ions in information-dependent acquisition mode. The collision energy was set according to the *m/z* value and charge state of any given precursor ion. The CID spectra were submitted for database searching using in-house Protein Prospector (prospector.ucsf.edu) permitting acetylation and mono-, di-, and trimethylation as variable modifications. In addition, the CID spectra of all modified peptides were manually inspected.

FT-ICR MS—The intact histone proteins as well as an aliquot of the tryptic digest were analyzed by capillary HPLC-ESI-FT-ICR MS using a Surveyor HPLC pump interfaced to an LTQ-FT mass spectrometer (both Thermo, San Jose, CA). Separation was performed using a 150-μm × 10-cm C18 column (Microtech, Sunnyvale, CA). Solvent A was 0.1% formic acid and solvent B was 0.1% formic acid in ACN, and the gradient was 2% B for the first 5 min, then a gradient to 40% B over the next 40 min, followed by a gradient up to 90% B over the next 10 min at a flow rate of 800 nl/min. Spraying was from an uncoated 15-μm inner diameter spraying needle (New Objective, Woburn, MA). All MS data were acquired in the ICR cell; the amount of ions injected into the ICR trap was optimized by monitoring ion counts in the linear trap prior to injection into the ICR trap. In addition, the intact proteins were also analyzed by nanospray sample introduction at a concentration of ~5 pmol/μl. The deconvolution program employed is a tool in the Finnigan Bioworks toolbox (Thermo).

ECD experiments on the intact histone mixture were performed on a homebuilt 9.4 Tesla passively shielded ESI-Q-FT-ICR mass spectrometer (24) controlled by a modular ICR data acquisition system (MIDAS) (26). The protein concentration employed was the same as above. A Nanomate chip system (Advion BioSciences, Inc., Ithaca, NY) consisting of a 10 × 10 grid of reproducibly formed 10-μm inner diameter spray nozzles with low nanoliter per minute flow rates was used for sample introduction. For external ion accumulation (27), ions were held in a focusing octopole followed by precursor ion mass selection in a quadrupole mass filter and subsequent accumulation in a linear octopole ion trap. For selection of the single species at peak 1125.114 (12+), further separation was achieved by SWIFT isolation (20) in the open-ended cylindrical ICR trap. The instrument configuration, operating conditions, and accuracy of mass measurement (<5 ppm) for ECD have been previously described (28). Briefly, electrons for ECD are generated from a 10-mm diameter dispenser cathode (no. 1109; HeatWave, Watsonville, CA) mounted on the central axis of the system. During the 10-ms ECD event, the trap plates were set at +10 V, the extraction grid at +5 V, and the cathode was biased to –2 V. Following ECD, a 100-ms event was carried out in which the trap plates are dropped to +2 V, the grid pulsed to +5 V, and the cathode biased to +10 V to purge any remaining electrons from the ICR cell. Following dissociation, ions undergo frequency-sweep excitation and broadband detection (512 Kword data points).

RESULTS

For Intact Protein—The protein preparation was introduced intact by reversed-phase LC/MS using conditions similar to those described for analysis of the digests, and as anticipated practically no chromatographic separation was observed among the isoforms present. Thus, the sample was nanospray-introduced into the LTQ-FT mass spectrometer, which permitted extended acquisition of MS data and CID analysis of multiple components. Automated deconvolution of the data indicated the presence of 13 proteins in the mixture (Table I, Fig. 1). Due to the complexity of the sample even with the resolution and mass accuracy provided by an FT-ICR MS instrument, not all components detected could be resolved (Table I, Figs. 1 and 2).

After molecular mass analysis, CID experiments were per-

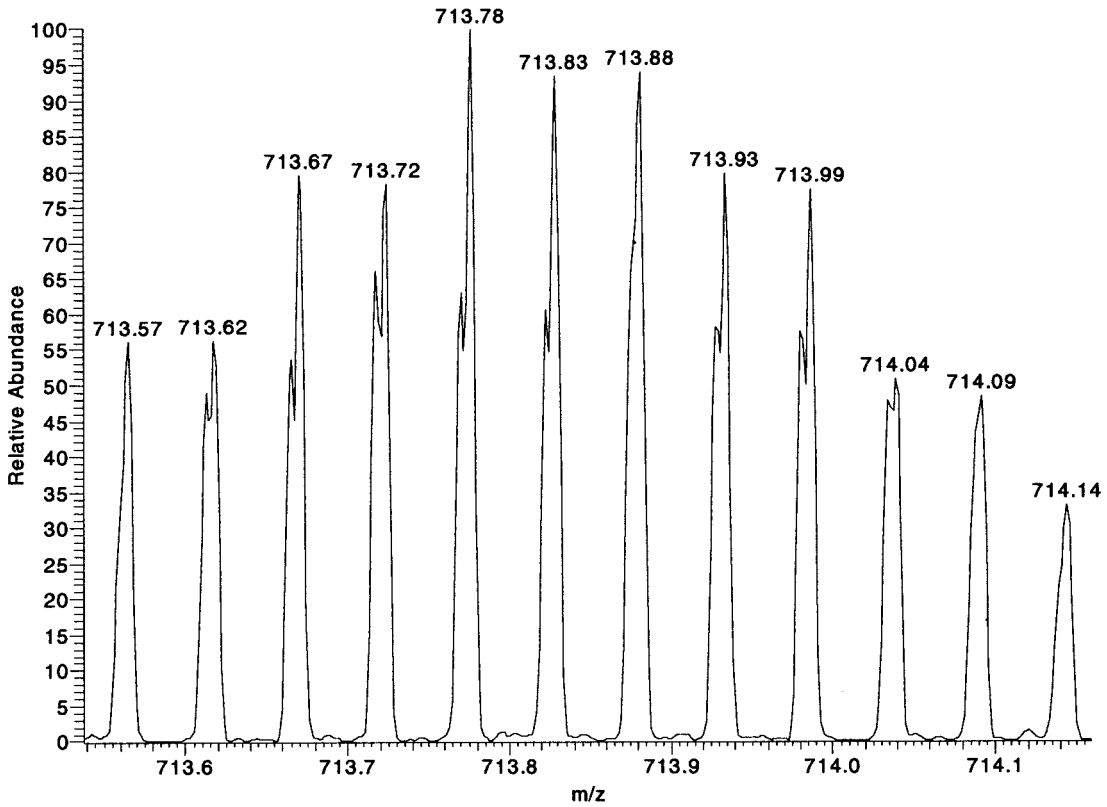


FIG. 2. **Isotope cluster of the most abundant ions in the 19+ ion cluster during FT-ICR MS mass measurement.** The "shoulders" on the peaks suggest the presence of multiple unresolved components.

formed on a series of components. Ions were isolated in the linear trap and an ~3-Da wide window was employed for the precursor ion selection. Thus, each experiment yielded information on multiple components. As one might anticipate, the most abundant species (no. 6 in Table I) yielded the best CID spectrum (Fig. 3). However, analysis of these low-energy CID spectra provided only limited information, because the proteins fragmented along a few peptide bonds situated approximately in the middle of the protein sequence (see sequence coverage in Scheme 1). However, the **y** fragments observed confirmed the presence of both H2B isoforms, and **b** ions of histone H2B.2 established that the +42-Da modification was located somewhere within the N-terminal 57 residue moiety. Accurate mass values of the **b** ions were consistent with a trimethyl modification rather than acetylation. For example, the predicted protonated mass values for modified $\mathbf{b_{57}}$ are 6418.6719 and 6418.7082, acetylated or trimethylated, respectively. The measured ion at *m/z* 642.77672 (10+) (Fig. 4) translates to a singly charged mass of 6418.6968 Da that corresponds to a 1.6-ppm error for a trimethylated peptide, but a 3.8-ppm difference from the theoretical mass of an acetylated species. The high abundance and signal-to-noise of this peak ensures reliable accurate mass measurement.

A second purification of histone H2B from *T. thermophila* was subjected to ECD analysis on the Tallahassee custom-built FT-ICR MS instrument. This time sample introduction was accomplished using the Nanomate chip described above. This system is more efficient than nanospray needles in sample consumption due to the ease of saving any solution that was not sprayed and returning to a 96-well microtiter plate. This introduction system may be used in a completely automated fashion as well (29). Two different techniques were employed to select precursor ions for the ECD experiments. SWIFT isolation permitted almost single species selection. The most abundant component of the mixture, 1125.3343 (12+), was SWIFT-selected for ECD fragmentation (Fig. 5). Interpretation of the ECD spectra established this species as H2B.1 modified by 42 Da nominally. The smallest fragment ion detected was $\mathbf{c_3}$ at *m/z* 356.2662. This mass already contains the 42-Da mass shift and could be assigned based on its accurate measured mass as trimethylation. The calculated mass for this structure (*m/z* 356.2661) is within 1 ppm of the mass observed. All further N-terminal **c** fragments displayed the nominal 42-Da mass shift, while none of the C-terminal **z** ions present indicated any modification. Hence, the trimethylation must be present either at the N-terminal amino function or on the ε-amino group of Lys-3, or as a mutually exclusive mixture of both possible structures. The ECD experiment was repeated with a quadrupole-selected, wider precursor ion window. The whole charge state envelope was isolated in order to

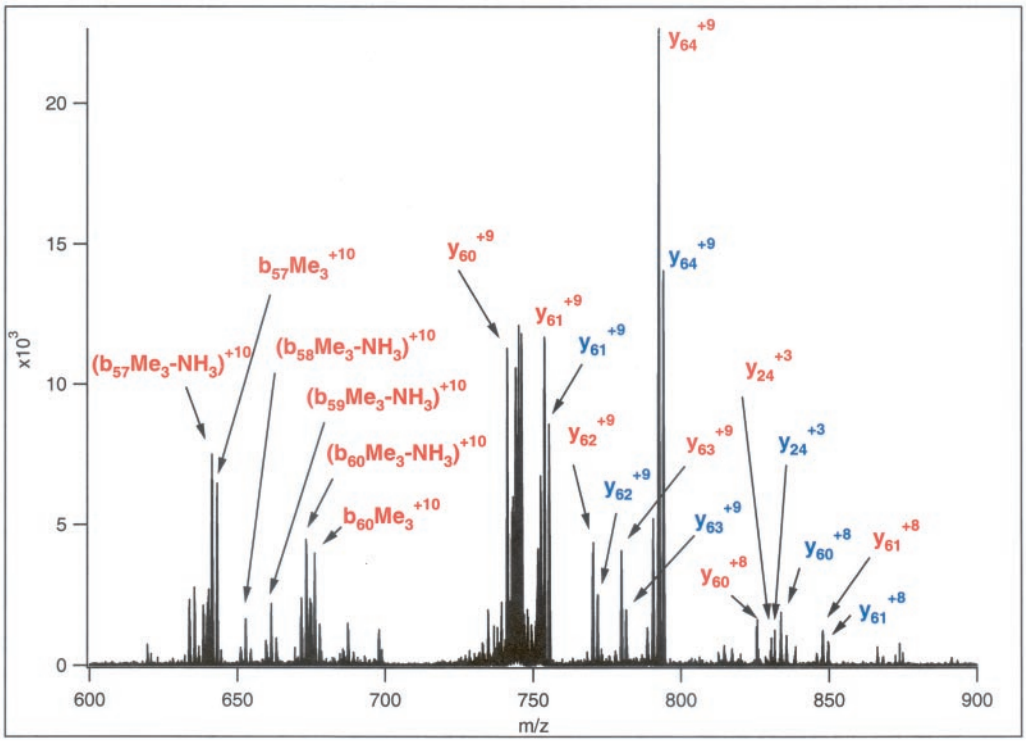


FIG. 3. **FT-CID spectrum of *m/z* 713 (19+).** Duplicate pairs of **y** fragments revealed the presence of both H2B.1 (P08993, labeled in *blue*) and H2B.2 (P08994, labeled in *red*), while **b** ions were detected only for H2B.2.

SCHEME 1. **CID sequence coverage of H2B.1 and 2.** Sequence differences in *bold*.

H2B.1

1
APKKAPAA**AA** EKKVKKAPTT EKKNKKKRSE TFAIYIFKVL KQVHPDVGIS

51
KKAMNIMNSF INDSFERIAL ESSKLVRFNK RRTLSSREVQ TAVKLLLPGE

101
LARHAISEGT KAVTKFSSS**T** N

H2B.2

1
APKKAPAA**TT** EKKVKKAPTT EKKNKKKRSE TFAIYIFKVL KQVHPDVGIS

51
KKAMNIMNSF INDSFERIAL ESSKLVRFNK RRTLSSREVQ TAVKLLLPGE

101
LARHAISEGT KAVTKFSSS**S** N

maximize the quantity of precursor ions, leading to better signal-to-noise for fragment ions and better mass accuracy (see Fig. 6, Table II, and supplemental data).

The peaks observed in this ECD spectrum (detailed data presented in supplemental data) featured mostly **c** and **z** fragments, starting again with $\mathbf{c_3}$, modified by 42 Da. Accurate mass measurement confirmed that this 42 Da represents modification by trimethylation rather than the isobaric acetylation. Starting from Lys-4, doubly modified (+84 Da) **c** fragments were also detected. Accurate mass measurement established the second modification as acetylation. For example, the observed mass for $\mathbf{c_4}$, *m/z* 526.3735 is within 3 ppm of the calculated mass for the singly acetylated, singly trimethylated structure: *m/z* 526.3718. The deviation would be 72 ppm and −66 ppm relative to the calculated values of the alternative doubly acetylated or doubly trimethylated fragments, respectively. Because the H2B.1 and H2B.2 protein sequences are identical up to the ninth residue (see Scheme

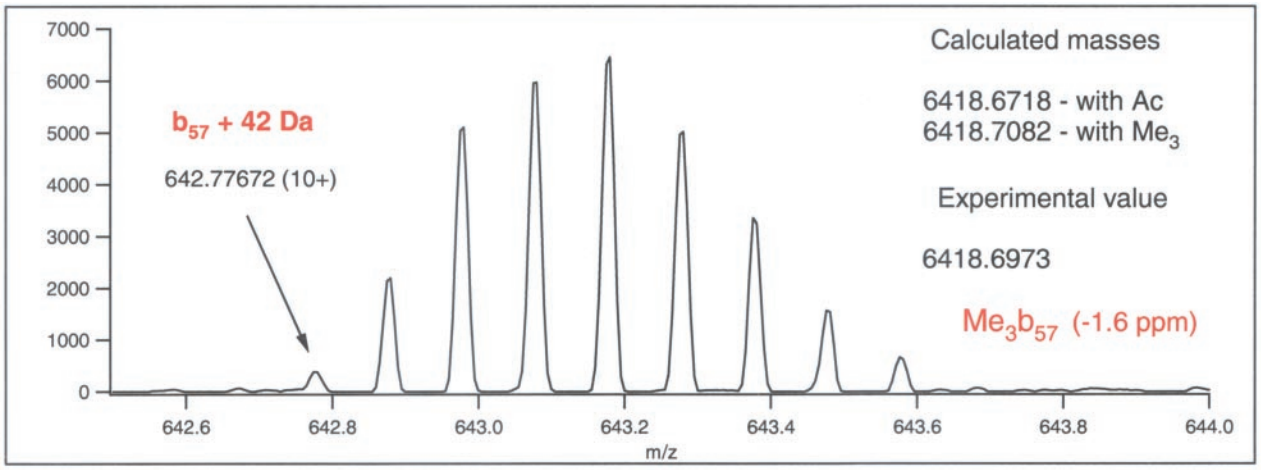
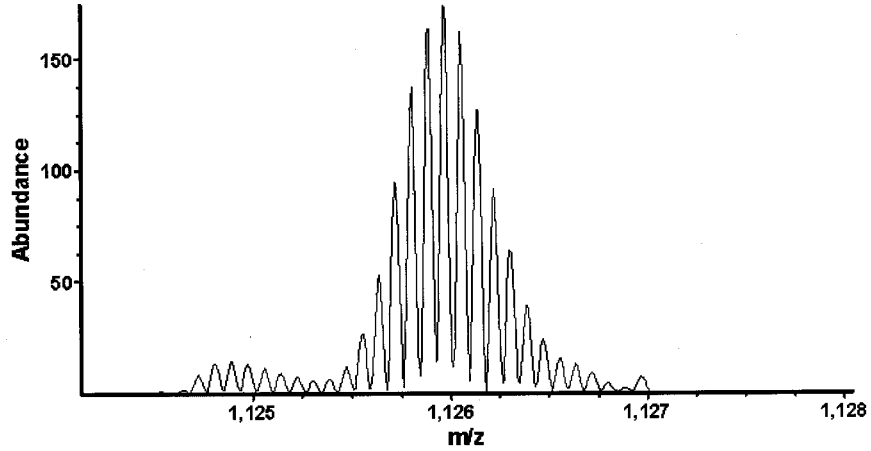


FIG. 4. **Fragment ion in the CID spectrum acquired by LTQ-FT mass spectrometer.** The accurate mass of this **b** fragment was consistent with a trimethyl modification rather than acetylation.

FIG. 5. **SWIFT-selected precursor ion for FT-ICR MS-ECD experiment.**



1), the modified low-mass **c** fragments may obviously represent contributions from both proteins. SWIFT-ECD analysis of the most abundant species indicated that H2B.1 is trimethylated either at its N terminus or on Lys-3. The ECD data for the full charge state envelope shows that H2B.2 displays the same modification, as witnessed by the presence of a trimethylated $\mathbf{c_{12}}$ of this isoform at *m/z* 597.3750 (2+). H2B.2 is also acetylated at Lys-4: its $\mathbf{c_{11}}$ ion was detected doubly modified at *m/z* 584.3435 (2+). This mass also reflects the presence of both an acetylation and a trimethylation. The C-terminal **z** fragments residue range of both isoforms did not indicate the presence of acetylation or methylation. However, the presence of **z** fragments shifted by 16 Da in the ECD spectrum suggested oxidation of H2B.1. The $\mathbf{z_{70}}$ was observed with a 16-Da shift, at *m/z* 1304.6990 (6+), within 3 ppm of the theoretical mass. Combined with the observation of an unmodified $\mathbf{c_{54}}$ this suggests oxidation of Met-57. Summaries of the sequence coverage obtained by ECD are presented in Schemes 2 and 3.

For Protein Digests—In order to gain more information on the location(s) and distribution of posttranslational modifications in histone H2B, the protein mixture was digested with two endoproteinases, trypsin and AspN, and the corresponding digests were subjected to LC/MS analyses on a QqTOF mass spectrometer in an information-dependent acquisition manner. The CID data were evaluated for the presence of covalent modifications; *i.e.* for acetylation or methylation of lysine residues because these were the modifications expected based on the mass spectrometric measurement observed for the intact proteins. The modified peptides that have been identified from the tryptic digest are listed in Table III. From these experiments, the acetylation of Lys-4 (Fig. 7) and Lys-41 (data not shown) was established unambiguously. In addition, the complex modification occupancy of the N-terminal "tail" was also confirmed. In the case of those peptides containing double modification, *i.e.* +84 Da, the CID data spectra obtained from multiply charged ions of peptides [1–13] and [1–15] for both isoforms usually did not permit site assignments. This situation is due to the fact that the first N-terminal fragment in these CID spectra are $\mathbf{a_5}$ and $\mathbf{b_5}$ ions,

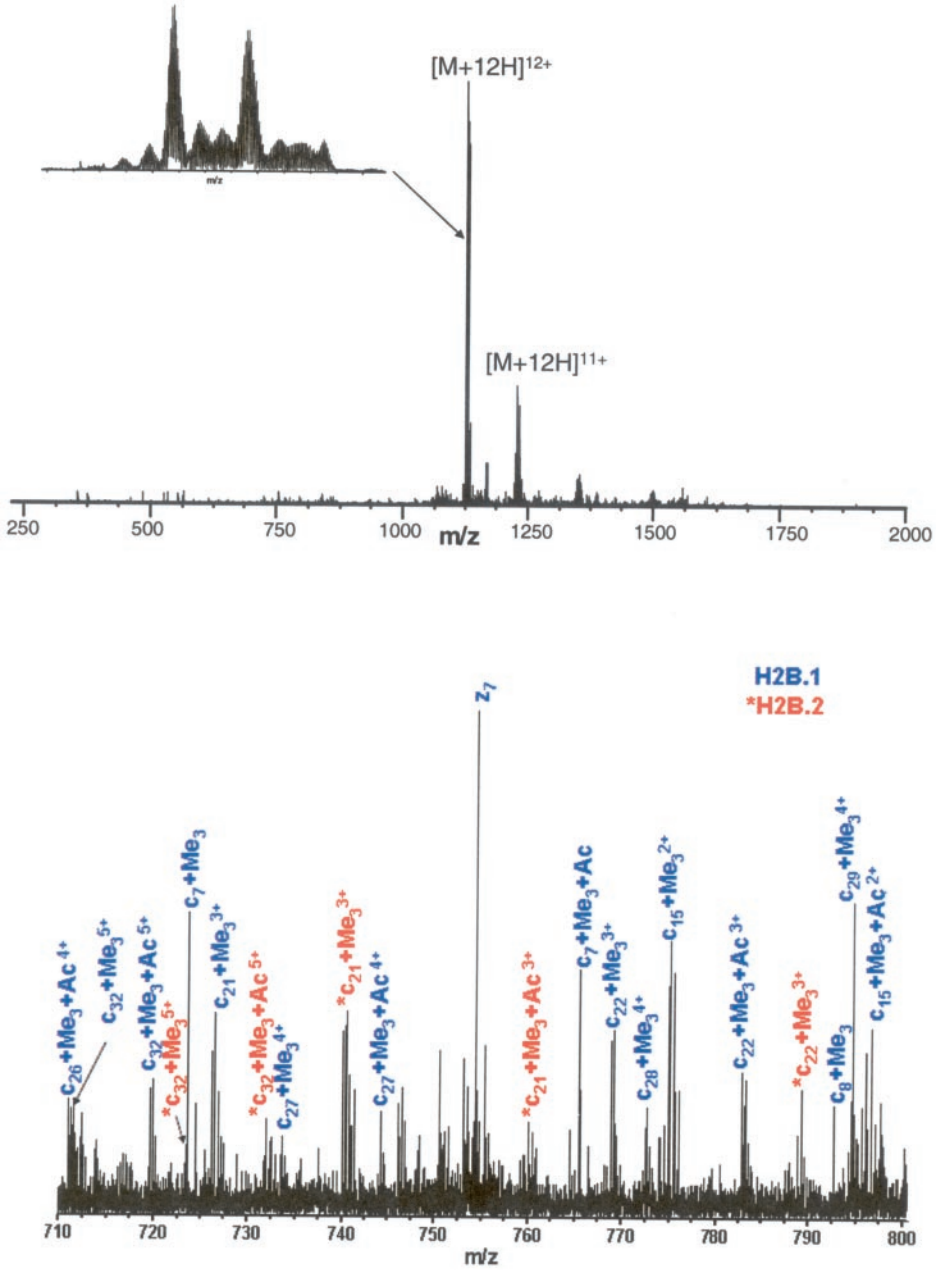


FIG. 6. **ECD spectrum of the quadrupole-selected precursor ions.** The lower panel shows a 90-Da mass range in detail to illustrate the complexity of the data. (Supplemental Table I shows the full fragment list.)

at *m/z* 552.35 and 580.39, respectively, and these accurate mass values reflect the presence of both an acetyl and a trimethyl modification. These two different modifications could occupy two of three possible sites; namely the N terminus itself and the side chains of Lys-3 and -4. In two CID spectra acquired from doubly modified ^{1}APKKAPAAAAEKKVK^{15} and ^{1}APKKAPAATTEKKVK^{15}, $\mathbf{y_{14}}$ fragments ions were observed modified by 42 Da. The presence of these $\mathbf{y_{14}}$ fragments, *m/z* 493.65 (3+) and 513.60 (3+), respectively, point to the modification of the N terminus itself. In the CID spectrum of the H2B.2 [1–15] peptide [*m/z* 413.773 (4+)], its $\mathbf{b_3}$ fragment was detected at *m/z* 339.24. This mass suggested trimethylation, either at the N terminus or Lys-3, while the presence of acetylation at Lys-4 could be established from additional fragment ions. In the CID spectrum of the equivalent H2B.1 peptide [*m/z* 398.770 (4+)], the first N-terminal ion observed was $\mathbf{b_4}$ bearing both the acetyl and trimethyl modification. Thus, we could assign the modification sites only for the other isoform. Interestingly, the singly modified [1–15] analog that would represent a component from the major protein population according to the intact protein analysis, was not detected in the tryptic digest, probably because when either or both Lys-3 and Lys-4 are not modified trypsin will cleave at those sites forming modified tri- or tetrapeptides that may not be retained on the reversed-phase column.

Only one modification was detected in the C-terminal part

of the protein. Unfortunately, the CID spectrum obtained was comprised of two coeluting, isobaric molecules at *m/z* 631.668 (3+). The spectrum, though weak, displayed ions verifying the presence of both a dimethylated peptide spanning residues [104–121] from H2B.1 and a trimethylated (or acetylated) peptide of residues [104–121] from H2B.2 (see Fig. 8). Thus the presence of the C-terminal H2B.1 peptide was established by fragments occurring at *m/z* 234.11, 871.4, 970.5, and 1041.4, corresponding to $\mathbf{y_2}$, $\mathbf{y_8}$, $\mathbf{y_9}$, and $\mathbf{y_{10}}$, respectively; while the presence of the trimethylated or acetylated other isoform was established from $\mathbf{y_2}$ at *m/z* 220.1 as well as a series of 42-Da-shifted **b** ions including $\mathbf{b_9}$ and $\mathbf{b_{10}}$ at *m/z* 937.5 and 1036.5, respectively.

Results from the analysis of the endoproteinase AspN digest revealed modification of both Lys-3 and Lys-4 (see Table IV). For example, the CID spectrum of the peptide at *m/z* 507.94 (5+) is shown in Fig. 9. It represents residues [1–23] bearing modification by +84 Da. Based on the occurrence of particular internal fragments, it is clear from its assignment to the partial sequence PKKAP that this +84-Da shift occupies the ϵ amino functions of both the Lys-3 and Lys-4 [see *m/z* 289.71 (2+)]. In addition, doubly modified internal fragments were observed for the H2B.1 peptide [1–10] and are listed in supplemental data. While these abundant internal fragment ions helped to verify that both Lys residues are modified in fact, their presence also interfered with the site assignment. For example, *m/z* 339.24 could represent a trimethyl-$\mathbf{b_3}$ assignment as well as one of these moieties: KKV/KVK/VKK-NH_3 (see Fig. 9).

Interestingly, a series of larger polypeptides, corresponding to residues [1–45] for both isoforms, was detected in this digest that established the presence of a variable degree of methylation (Fig. 10). CID analysis of the peptides confirmed their identity. In addition, the different mass shifts of $\mathbf{b_5}$ ions confirmed that the major modifications occupy either the N terminus or the side-chains of Lys-3 and -4. From the relative abundances, these differently methylated species show good agreement with the relative abundances observed for the intact protein population, confirming the majority of the post-translational modifications on these proteins are concentrated among the N-terminal residues.

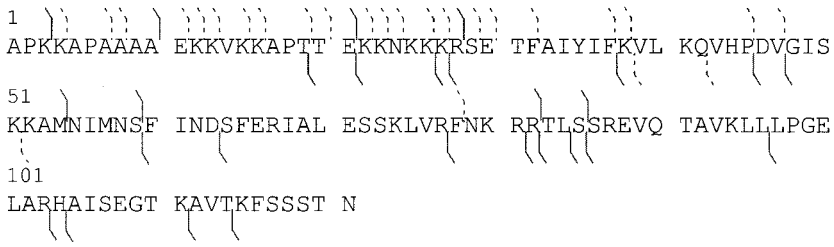
TABLE II
Quadrupole-selected precursor ions for an ECD-experiment (shown in Fig. 6)

	Ion detected	$MH^+_{determined}$	$MH^+_{calculated}$	Assignment
A.	1121.5547(12+)	13447.6	13450.5	H2B.1
B.	1122.9312(12+)	13464.1	13464.5	H2B.1 + Me
C.	1124.0862(12+)	13477.9	13478.5	H2B.1 + Me_2
D.	1125.8904(12+)[a]	13491.6	13492.5	H2B.1 + 42
E.	1126.5431(12+)	13507.4	13508.5	H2B.1 + 42+O
F.	1127.6409(12+)	13520.6	13520.5	H2B.1 + 42 + 28
G.	1128.8056(12+)	13534.6	13534.5	H2B.1 + 2 × 42
H.	1130.2974(12+)	13552.5	13552.5	H2B.2 + 42 + 14
I.	1131.4025(12+)	13565.7	13566.5	H2B.2 + 42 + 28
J.	1133.3576(12+)	13577.2	13576.5	H2B.1 + 3 × 42

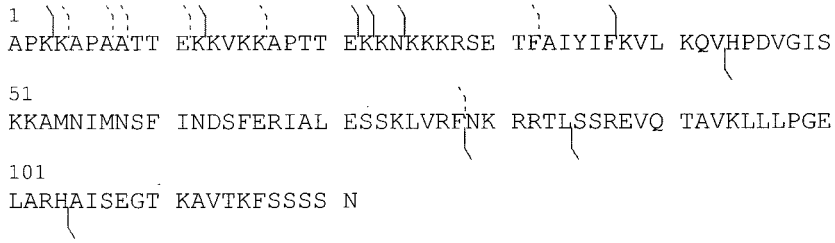
[a] The most abundant components are underlined.

DISCUSSION

Several studies of intact histones have been carried out by MS (30–33). These studies were performed using MALDI-TOF, LC-ESI-MS on a quadrupole instrument or LC-ESI-MS with a QqTOF mass spectrometer. These instruments provide information on the global picture of the protein's posttranslational state but due to their limited mass resolution and mass measurement accuracy they struggle to resolve species that are of similar molecular masses, *e.g.* the difference between a species bearing methylation (+14 Da) *versus* an oxidation (+16 Da), etc. The higher resolution of an FT-ICR MS instrument has the potential to achieve this, although when there are overlapping isotope clusters (as shown in Table I and Figs. 1 and 2) it will not always provide unambiguous component resolution. However, for these intact protein studies cited above, no fragmentation data was reported at the protein or



SCHEME 2. **ECD sequence coverage for H2B.1.** N-terminal fragment ions were **c** type ions; *bold* fragments were observed modified by 42-Da; *dashed* fragments were observed modified by 42 and 84 Da. Some **a+1** fragments were observed as well. C-terminal fragment ions were **z** type ions; *bold* fragments were observed unmodified; *dashed* fragments were also observed modified by 16 Da.



SCHEME 3. **ECD sequence coverage for H2B.2.** N-terminal fragment ions were **c** type ions; *bold* fragments were observed modified by 42 Da; *dashed* fragments were observed modified by 42 and 84 Da. C-terminal fragment ions were **z** type and were all observed unmodified.

TABLE III
Modified peptides identified from the tryptic digest (listed in elution order)

Ion detected	$MH^+_{calculated}$	Structure	Δ[ppm]
418.608 (3+)	1253.758	^{4}K(Ac)APAAAAEKKVK^{15}	+40
438.611 (3+)	1313.779	^{4}K(Ac)APAATTEKKVK^{15}	+29
341.978 (4+)	1364.889	^{1}APKK(Ac,Me$_3$)APAAAAEKK^{13}[a]	+44
413.773 (4+)	1652.011	^{1}APKK(Ac,Me$_3$)APAATTEKKVK^{15}[b]	+35
		Me$_3$^{1}APKK(Ac)APAATTEKKVK^{15}[c]	
410.269 (4+)	1637.995	Me$_2$^{1}APKK(Ac)APAATTEKKVK^{15}[d]	+35
398.770 (4+)	1591.990	^{1}APKK(Ac,Me$_3$)APAAAAEKKVK^{15}[a]	
		Me$_3$^{1}APKK(Ac)APAAAAEKKVK^{15}[e]	+42
395.266 (4+)	1577.974	Me$_2$^{1}APKK(Ac)APAAAAEKKVK^{15}[e]	+42
487.961 (3+)	1461.843	^{39}VLK(Ac)QVHPDVGISK^{51}	+17
631.668 (3+)	1892.972	^{104}HAISEGTK(Me$_2$)AVTKFSSSTN^{121}[f]	+9
		^{104}HAISEGTK(Me$_3$)AVTKFSSSSN^{121}	
	1892.935	^{104}HAISEGTK(Ac)AVTKFSSSSN^{121}	+28
636.323(3+)[g]	1906.950	^{104}HAISEGTK(Me$_3$)AVTKFSSSTN^{121}	+2

[a] AcK is present, Me$_3$ is either on K or on the N terminus.
[b] In the CID spectrum of the first eluting peptide internal ions indicate that both Lys residues are modified.
[c] In the CID of the second eluting peptide **y$_{14}$** was detected modified by only +42 Da, b$_3$ shows the presence of Me$_3$, and the Lys immonium ion indicates acetylation. The two differently modified species were not separated chromatographically.
[d] Same CID as above, very weak **b$_3$** present.
[e] From the presence of the 42-modified **y$_{14}$** it is obvious that the N terminus is modified. However the first **b** fragment detected is **b$_4$** doubly modified.
[f] These structures represent identical masses, the CID spectrum displays **y** ions from H2B.1, and also features **b** fragments modified by 42 Da (Fig. 8).
[g] No CID data.

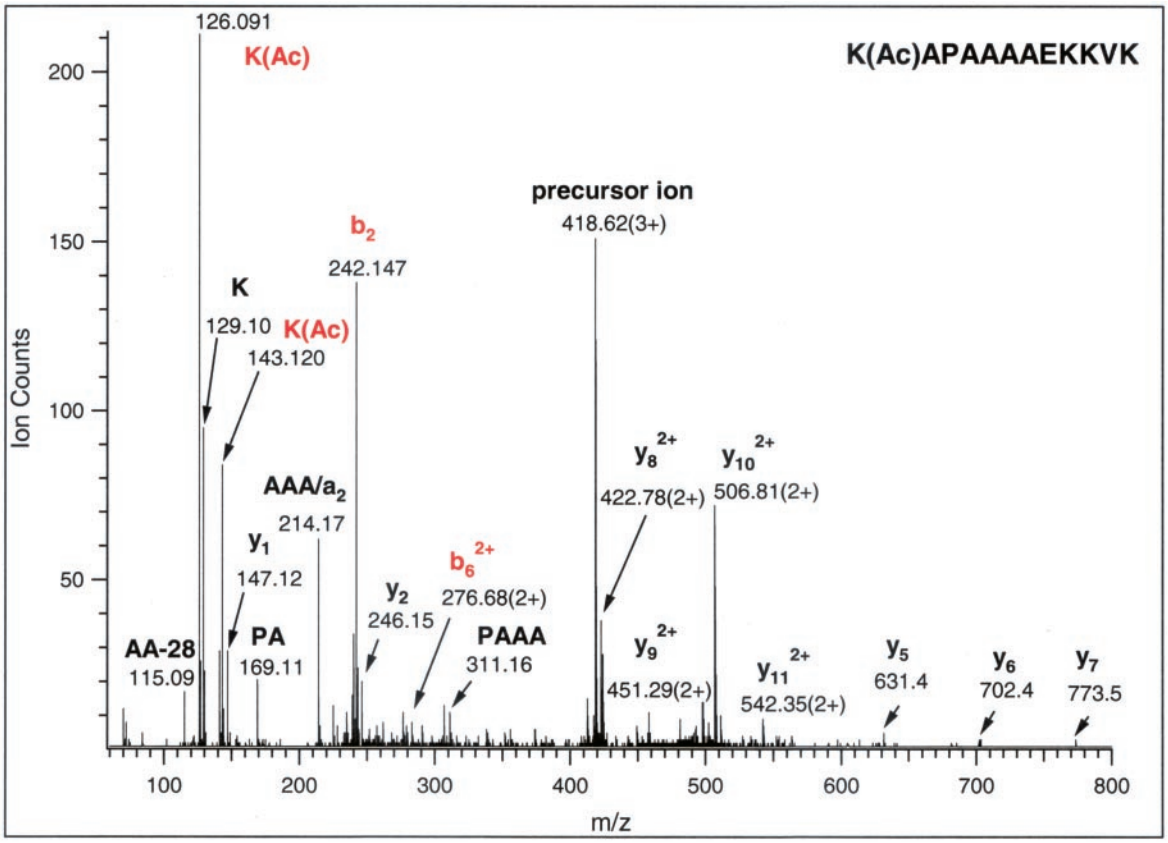


FIG. 7. **Low energy CID data of ^{4}Lys(Ac)-Ala-Pro-Ala-Ala-Ala-Ala-Glu-Lys-Lys-Val-Lys^{15} of histone H2B.1.** The immonium ions for the modified residue and the b$_2$ fragment ion unambigously identify the modification as an acetyl group instead of a trimethyl derivative. The calculated masses for these ions are: 126.092, 143.119, and 242.151 Da. Thus, the mass deviations between the calculated and measured values are −8, +7, and −17 ppm.

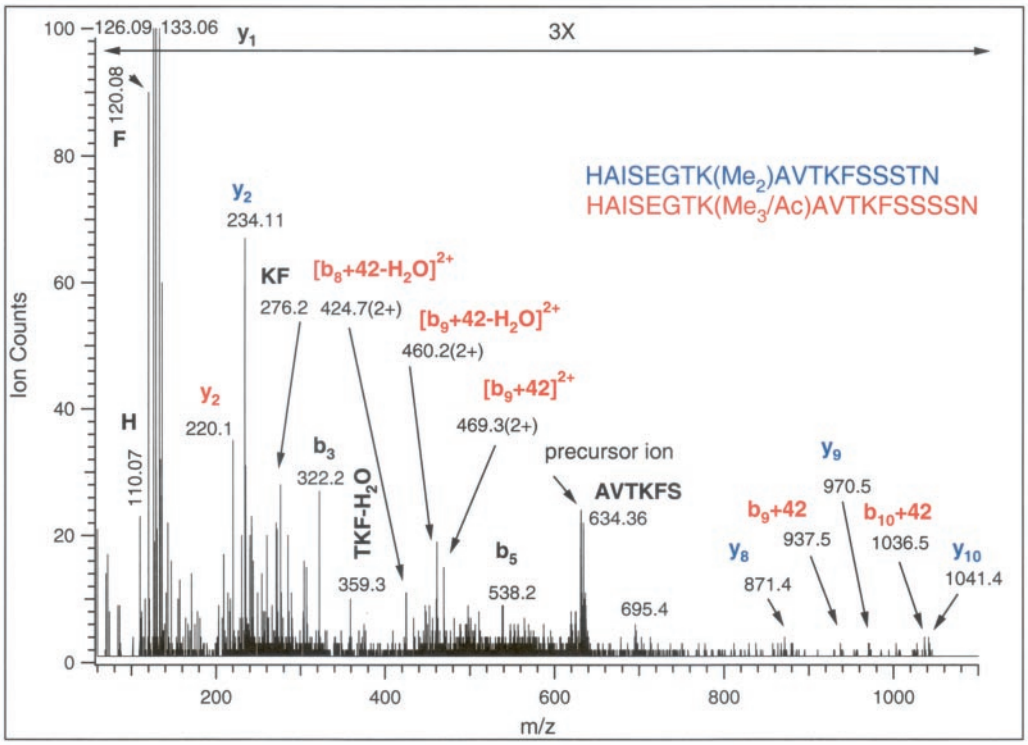


FIG. 8. **Low-energy CID of two isobaric C-terminal peptides.** The precursor ion was *m/z* 631.668 (3+), and contains both the structures listed here (see Table III). Fragments that belong to the H2B.1 isoform are labeled in *blue*, fragments that indicate the presence of the H2B.2 peptide are labeled in *red*. Some fragments are common to both peptides.

peptide levels and thus the sites of modification assigned are somewhat speculative.

Numerous studies have used MS to study histone post-translational modifications at the peptide digest level (34–39). These studies have identified several modification sites but provided no information about the global state of the protein in terms of posttranslational modifications and relative stoichiometries of such modifications. The only previously published modification of *Tetrahymena* histone H2B is trimethylation of the N-terminal alanine residue (22). While our results confirmed this assignment, they have also established in addition that the protein's N-terminal alanine occurs free and with heterogeneous occurrence of mono- and dimethyl populations as well as the trimethyl reported much earlier (22).

The present study presented here is the first attempt to combine the two approaches to provide a complete characterization of the protein posttranslational cluster occupancy. It clearly illustrates the advantages and drawbacks of both approaches. As shown in the summary results (Fig. 11), analysis of the digests revealed the presence of multiple modifications on *Tetrahymena* H2B. However, from information obtained from analysis of the digest datasets we do not know the relative amounts or anything about the co-occurrence of the different modifications on the protein. In contrast, molecular mass characterization of our protein preparation provided critical information on the relative distribution of protein populations, clearly identifying the major protein variants and posttranslational occupancies; *i.e.* a trimethylated species together with trimethylated and acetylated doubly modified species for both H2B.1 and H2B.2. Taking advantage of ECD and accurate mass measurement, the acetylation site (Lys-4) could be assigned, while two potential sites (N-terminal Ala or Lys-3) for the single trimethylation were determined. So in order to differentiate between these two potential trimethylation sites, interpretation of CID analyses from both proteolytic digests was necessary. From these analyses, conclusive evidence was obtained for trimethylation of the N-terminal amino function as well as the presence of an additional moiety, *i.e.* acetylation, located on either Lys-3 or Lys-4. Additionally, analysis of other modified N-terminal peptides established both Lys-3 and Lys-4 to be modified (acetylation and trimethylation), but for these particular peptides their N terminus was found to be free. Thus, our results reveal that the modification state at the N terminus and at Lys-3 must be mutually exclusive. It should be noted that tryptic peptides that would represent the major component of the protein population observed above (Me_3 modification only) were not detected in the digest. However, results from the tryptic digest did reveal the presence of modifications at Lys-41 and Lys-111, which appear to be present at much lower stoichiometry. Thus, while the analysis of the tryptic digest helped to pinpoint some posttranslational modifications, in fact it failed

TABLE IV
Modified peptides identified from the endoprotease AspN digest

Ion detected	MH^+ calculated	Structure	Δ [ppm]
490.323 (2+)[a,b]	979.594	^{1}APKK(Me$_3$, Ac)APAAAA^{10}	+44
483.325 (2+)	965.578	^{1}APKK(Me$_2$, Ac)APAAAA^{10}	+66
476.315 (2+)	951.563	^{1}APKK(Me, Ac)APAAAA^{10}	+62
520.332 (2+)[a,b]	1039.615	^{1}APKK(Me$_3$, Ac)APAATT^{10}	+47
507.941 (5+)[a,b]	2535.524	^{1}APKK(Me$_3$, Ac)APAATTEKKVKKAPTTEKK^{23}	+59
505.139 (5+)	2521.508	^{1}APKK(Me$_2$, Ac)APAATTEKKVKKAPTTEKK^{23}	+62
502.335 (5+)	2507.493	^{1}APKK(Me, Ac)APAATTEKKVKKAPTTEKK^{23}	+60
495.932 (5+)[a,b]	2475.503	^{1}APKK(Me$_3$, Ac)APAAAAEKKVKKAPTTEKK^{23}	+51
493.120 (5+)	2461.487	^{1}APKK(Me$_2$, Ac)APAAAAEKKVKKAPTTEKK^{23}	+33
490.324 (5+)[a,b]	2447.472	^{1}APKK(Me, Ac)APAAAAEKKVKKAPTTEKK^{23}	+47
563.26 (9+)	5061.34	[1–45] H2B.1+Me	–12
564.82 (9+)	5075.38	[1–45] H2B.1+Me$_2$	–12
566.36 (9+)[a,c]	5089.24	[1–45] H2B.1+Me$_3$	–12
567.92 (9+)	5103.28	[1–45] H2B.1+Ac,Me	–12
569.45 (9+)	5117.05	[1–45] H2B.1+Ac,Me$_2$	–12
569.90 (9+)	5121.10	[1–45] H2B.2+Me	–12
571.03 (9+)	5131.27	[1–45] H2B.1+Ac,Me$_3$	–12
573.03 (9+)[a,c,d]	5149.27	[1–45] H2B.2+Me$_3$	–12
574.56 (9+)	5163.04	[1–45] H2B.2+Ac,Me	–12
576.14 (9+)	5177.26	[1–45] H2B.2+Ac,Me$_2$	–12
577.70 (9+)[a,c]	5191.30	[1–45] H2B.2+Ac,Me$_3$	–12

[a] Identity confirmed by CID.
[b] Doubly modified internal fragments establish Lys-3, and -4 as the sites of modification.
[c] In all spectra **b$_5$** is the first N-terminal fragment detected. No "diagnostic" internal fragments were observed.
[d] Dimethyl derivative also detected overlapping with the Ac,Me$_3$-modified H2B.1 peptide ions.

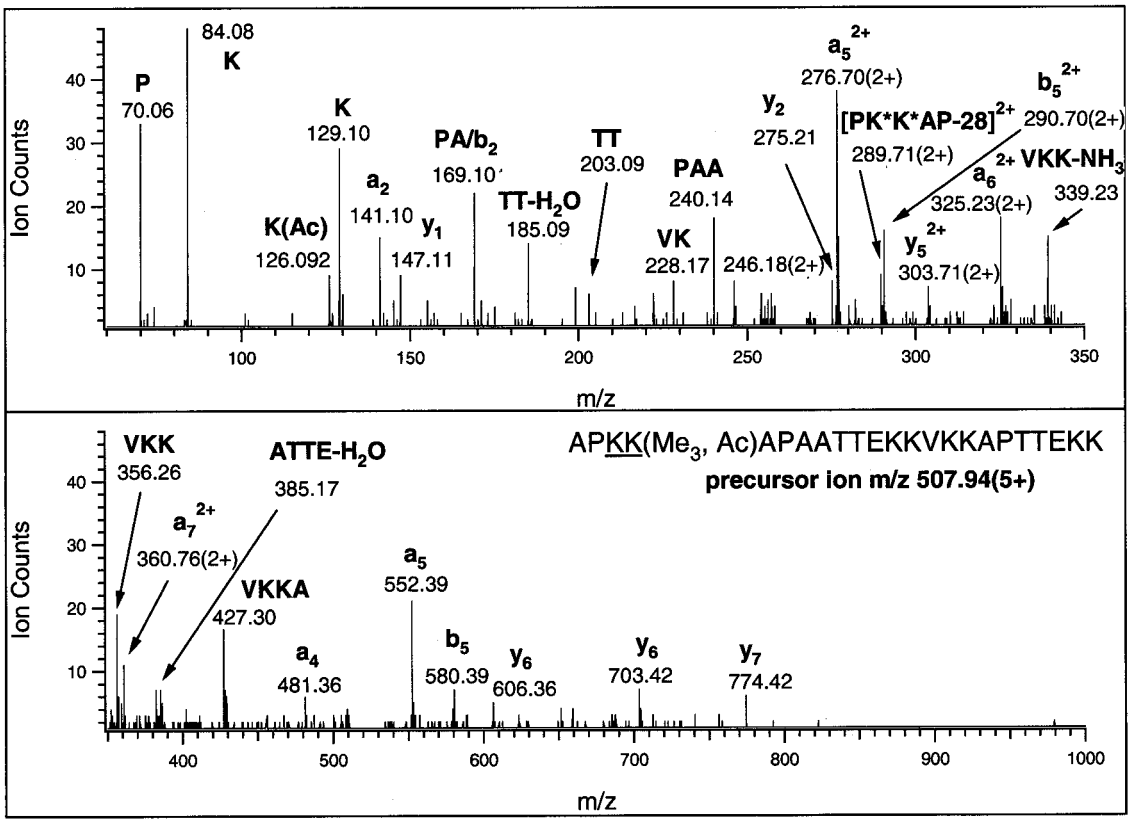


FIG. 9. **Low-energy CID spectrum of an H2B.2 N-terminal peptide spanning residues [1–23] formed during the endoproteinase AspN digestion.** The presence of the doubly modified internal ion PKKAP at *m/z* 289.71 (2+) demonstrates modification of both Lys-3 and Lys-4 residues.

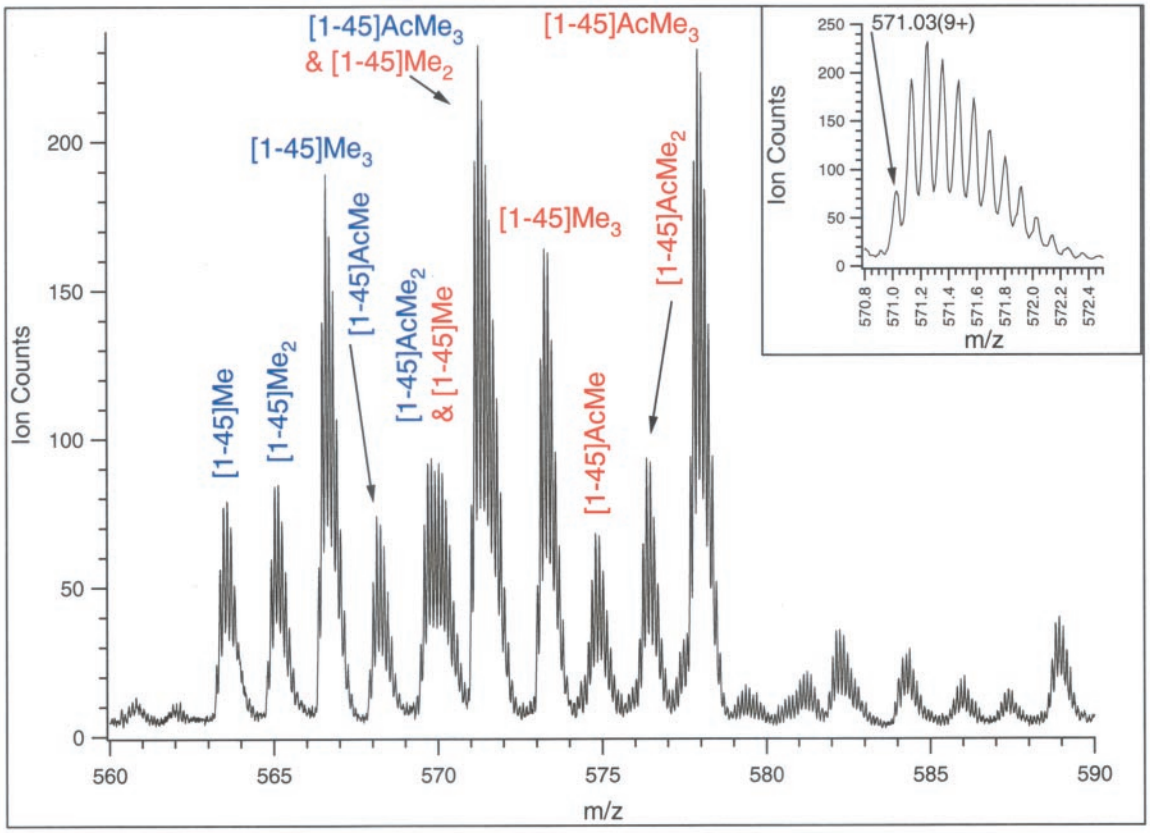


FIG. 10. **Mass spectrum of differently modified versions of the peptide spanning residues [1–45], detected in the endoproteinase AspN digest.** This spectrum represents a spectrum averaged over a period of 2 min during which these components eluted. The monoisotopic masses detected are listed in Table IV. The *inset* shows one of the isotope clusters. H2B.1 peptides are labeled in *blue*, H2B.2 peptides are labeled in *red*.

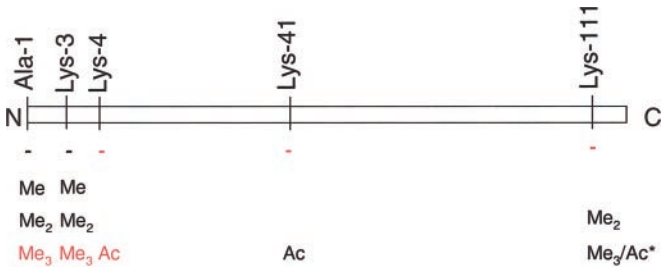


FIG. 11. **Summary of the posttranslational modifications identified on Histone H2B.** The major modification states (detected in ECD FT-ICR MS) are in *red*. Other modifications were only detected after enzymatic digestion. The *asterisk* indicates a mass accuracy not good enough to distinguish between acetylation or trimethylation of this residue. The same modification pattern was observed for both H2B.1 and H2B.2.

to reveal evidence for the most abundant form of the suite of modified proteins established much earlier (22). On this matter, we anticipated that digestion by the endoproteinase AspN would provide this information, but unfortunately none of the CID spectra acquired provided the conclusive evidence necessary for exact site assignments due to the presence of internal fragment ions isobaric with the expected discriminatory **b** ions sought in carrying out this digest.

Relative to core histones H3 and H4 (see for review Ref. 40 and references therein), considerably less information is known regarding the modification status of H2B in different organisms, and even less so regarding the modifications on *Tetrahymena* H2B. Information on the covalent modification profile of ciliate histones will be important as organisms such as *Tetrahymena* afford an opportunity to gain insights into the function of these marks through a combination of histone genetics and biochemistry. For example, Lys-9 methylation in H3 is required for programmed elimination of germ-line DNA sequences in a pathway guided by small RNAs (41). While *Saccharomyces cerevisiae* has often been the premier model for mutational analyses into histone function, it is becoming clear that budding yeast lacks some of the hallmark properties of heterochromatin exhibited by other eukaryotes, including Lys-9 methylation.

Many of the covalent modifications identified in this study, or patterns of modification, are novel, and thus it will be interesting to dissect the biological relevance of these modifications in an organism that permits detailed mutational analyses. For example, it has recently been shown that H2B is phosphorylated on Ser-14 by MstI kinase to signal the apoptotic pathway in vertebrate and mammalian cells (42). How-

ever, the functional significance of this phosphorylation mark remains unclear, and in particular no phospho-binding effector protein or complex has been identified that may dock on the phosphorylated H2B tail in much the same way that bromodomains and chromodomains have been shown to engage context-dependent acetyl-lysine and methyl-lysine histone tails, respectively, to bring about downstream events (43).

The H2B analyzed in the current study originated from macronuclei prepared from starved cells. However, during the sexual pathway known as conjugation, macronuclei, now known as "old macronuclei," are selectively eliminated by an apoptosis-like process (44). It will be interesting to determine if phosphorylation of H2B also correlates with the pronounced chromatin changes that accompany the elimination of this nucleus. To that end, we note that the second major isoform of H2B in *Tetrahymena* (H2B.2) differs from the other major isoform (H2B.1) in that alanines at positions 9 and 10 are replaced by threonines at 9 and 10. It will be interesting to determine if either or both of these threonines are phosphorylated in "old macronuclei" that become pycnotic and are being resorbed during conjugation. Despite this uncertainty, an example of functional conservation might exist with the findings of *Tetrahymena* trimethylated Lys-3 and acetylated Lys-4 described in this study. It has been reported that Lys-5 in bovine histone H2B is found in either an acetylated or monomethylated state (23). If these lysine residues are functionally equivalent, it may be that Ser-6 in vertebrate H2B, adjacent to Lys-5, may provide another example of a "methyl/phos" switch to displace a chromatin-bound effector(s) that remains to be identified (5). It will be interesting to determine if Ser-6 in vertebrate H2B or a functionally equivalent serine in H2B in invertebrates or unicellular H2Bs is involved in chromatin compaction and/or the pronounced chromatin remodeling that accompanies apoptosis in these models.

Further experiments are planned on this sample that include the SWIFT isolation of individual species within the protein population to characterize which of the lower stoichiometry modifications detected occur on particular individual protein species. Ideally, prior chromatographic separation of the differentially modified species would be beneficial. Unfortunately, in our hands the presence of different degrees of methylation has minimal effect on reversed-phase chromatographic behavior of the protein.

Currently routine analysis of proteolytic digests of proteins is roughly two orders of magnitude more sensitive than the intact protein ECD analysis described herein. Of course, lower sensitivity for ECD analysis of intact protein is partly expected due to the extensive fragmentation produced leading to distribution of ion signal among a large number of fragment ions with lower individual peak intensities. In addition, protein species bearing the same posttranslational modifications but occupying different sites will have the same exact mass so will contribute their individual fragments to the same spectrum. Hence, significantly more sample will be required to produce this kind of data on all intact protein species known to be present from this work. It is also of concern that interpretation of protein fragmentation data is significantly more complicated at this time. Nevertheless, the extensive sequence coverage afforded by this approach can provide confidence that all multiply modified occupancy isoforms of a protein have been fully characterized. One other factor that is hindering intact protein sequence analysis at the moment should be noted. Unlike the situation in the peptide analysis field, which has many software tools developed to assist in the analysis of fragmentation data, the tools available for the analysis and interpretation of ECD data produced in protein fragmentation experiments are inadequate in many respects. However, the development of new software is underway (45), so data analysis should gradually become easier.

The results from this study demonstrate that analysis of a protein digest has the inherent potential to overlook significant information about the totality of any particular posttranslational state and thus the actual structural nature of the intact physiologically active forms of a protein. By observing the distribution of intact molecular protein mass values, one is able to derive essential information about the modifications that are present on each protein variant and isoform, as well as their relative stoichiometries. However, to detect and determine the locations of lower stoichiometry modifications, the significantly higher sensitivity of present peptide analyses should provide an important advantage for some time as long as the required modified peptides are in fact observed during the LC/CID-MS experiment.

Finally, there is an obvious demand for development of new MS approaches that have the inherent capability and power to address the full plethora of protein covalent moieties present. Such methodology is essential to define, dissect, and understand the totality of structural motifs present that are regulated and modulated by epigenetic processes. There is mounting evidence that changes in posttranslational occupancies and/or cluster patterns are involved in motif recognition by different effectors that provide links in communication with alternative cell pathways and functions (1–7, 10). Hence, we are confident that the approach illustrated by this example, namely combining intact protein fragmentation with peptide fragmentation, outlines a critical new experimental route to defining and understanding the functional modulation of proteins by multidentate posttranslational occupancies.

Note—After completion of this work, a communication by Kelleher and coworkers has appeared on histone H4 modification using ECD on an FT-ICR MS instrument (46). This communication contains high-quality ECD data confirming several known histone H4 modification sites.

Acknowledgments—We are indebted to Vlad Zabrouskov (ThermoFinnigan) for making the FT-ICR MS-CID measurements. We thank Mark R. Emmett for helpful discussions.

* This work was supported by National Institutes of Health National Center for Research Resources Grant RR 01614 (to A. L. B.) and RR 15804 (to A. L. B.), National Institutes of Health Grant GM 63959 (to C. D. A.), and National Science Foundation Grant NSF CHE-99–09502 (to A. G. M.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[S] The on-line version of this manuscript (available at http://www.mcponline.org) contains supplemental material.

‖ To whom correspondence should be addressed: Department of Pharmaceutical Chemistry, University of California, 513 Parnassus Avenue, San Francisco, CA 94143-0446. Tel.: 415-476-5641; Fax: 415-476-0688; E-mail: alb@itsa.ucsf.edu.

REFERENCES

1. Pawson, T., and Nash, P. (2003) Assembly of cell regulatory systems through protein interaction domains. *Science* **300,** 445–452
2. Pawson, T., Raina, M., and Nash, P. (2002) Interaction domains: From simple binding events to complex cellular behavior. *FEBS Lett.* **513,** 2–10
3. Hunter, T. (2000) Signaling–2000 and beyond. *Cell* **100,** 113–127
4. Schlessinger, J. (2000) Cell signaling by receptor tyrosine kinases. *Cell* **103,** 211–225
5. Fischle, W., Wang, Y., and Allis, C. D. (2003) Binary switches and modification cassettes in histone biology and beyond. *Nature* **425,** 475–479
6. Appella, E., and Anderson, C. W. (2001) Post-translational modifications and activation of p53 by genotoxic stresses. *Eur. J. Biochem.* **268,** 2764–2772
7. Pellicena, P., and Miller, W. T. (2001) Processive phosphorylation of p130Cas by Src depends on SH3-polyproline interactions. *J. Biol. Chem.* **276,** 28190–28196
8. Shou, W., Verma, R., Annan, R. S., Huddleston, M. J., Chen, S. L., Carr, S. A., and Deshaies, R. J. (2002) Mapping phosphorylation sites in proteins by mass spectrometry. *Methods Enzymol.* **351,** 279–296
9. Kalume, D. E., Molina, H., and Pandey, A. (2003) Tackling the phosphoproteome: tools and strategies. *Curr. Opin. Chem. Biol.* **7,** 64–69
10. Cohen, P. (2000) The regulation of protein function by multisite phosphorylation—A 25 year update. *Trends Biochem. Sci.* **25,** 596–601
11. Wells, L., Vosseller, K., and Hart, G. W. (2001) Glycosylation of nucleocytoplasmic proteins: Signal transduction and *O*-GlcNAc. *Science* **291,** 2376–2378
12. Comer, F. I., and Hart, G. W. (2001) Reciprocity between *O*-GlcNAc and *O*-phosphate on the carboxyl terminal domain of RNA polymerase II. *Biochemistry* **40,** 7845–7852
13. Yang, X., Zhang, F., and Kudlow, J. E. (2002) Recruitment of *O*-GlcNAc transferase to promoters by corepressor mSin3A: Coupling protein *O*-GlcNAcylation to transcriptional repression. *Cell* **110,** 69–80
14. Loughrey Chen, S., Huddleston, M. J., Shou, W., Deshaies, R. J., Annan, R. S., and Carr, S. A. (2002) Mass spectrometry-based methods for phosphorylation site mapping of hyperphosphorylated proteins applied to Net1, a regulator of exit from mitosis in yeast. *Mol. Cell. Proteomics* **1,** 186–196
15. Kjeldsen, F., Haselmann, K. F., Budnik, B. A., Sorensen, E. S., and Zubarev, R. A. (2003) Complete characterization of posttranslational modification sites in the bovine milk protein PP3 by tandem mass spectrometry with electron capture dissociation as the last stage. *Anal Chem.* **75,** 2355–2361
16. Claverol, S., Burlet-Schiltz, O., Gairin, J. E., and Monsarrat, B. (2003) Characterization of protein variants and post-translational modifications: ESI-MSn analyses of intact proteins eluted from polyacrylamide gels. *Mol. Cell. Proteomics* **2,** 483–493
17. Zubarev, R. A., Kelleher, N. L., and McLafferty, F. W. (1998) Electron capture dissociation of multiply charged protein cations. A nonergodic process. *J. Am. Chem. Soc.* **120,** 3265–3266
18. Kelleher, N. L., Zubarev, R. A., Bush, K., Furie, B., Furie, B. C., McLafferty, F. W., and Walsh, C. T. (1999) Localization of labile posttranslational modifications by electron capture dissociation: The case of γ-carboxyglutamic acid. *Anal Chem.* **71,** 4250–4253
19. Shi, S. D., Hemling, M. E., Carr, S. A., Horn, D. M., Lindh, I., and McLafferty, F. W. (2001) Phosphopeptide/phosphoprotein mapping by electron capture dissociation mass spectrometry. *Anal Chem.* **73,** 19–22
20. Guan, S. H., and Marshall, A. G. (1996) Stored waveform inverse Fourier transform (SWIFT) ion excitation in trapped-ion mass spectrometry: Theory and applications. *Int. J. Mass Spectrom. Ion Proc.* **158,** 5–37
21. Nomoto, M., Imai, N., Saiga, H., Matsui, T., and Mita, T. (1987) Characterization of two types of histone H2B genes from macronuclei of *Tetrahymena thermophila. Nucleic Acids Res.* **15,** 5681–5697
22. Nomoto, M., Kyogoku, Y., and Iwai, K.,(1982) *N*-Trimethylalanine, a novel blocked N-terminal residue of *Tetrahymena* histone H2B. *J. Biochem.* **92,** 1675–1678
23. Zhang, L., Eugeni, E. E., Parthun, M. R., and Freitas, M. A. (2003) Identification of novel histone post-translational modifications by peptide mass fingerprinting. *Chromosoma* **112,** 77–86
24. Senko, M. W., Canterbury, J. D., Guan, S., and Marshall, A. G. (1996) A high-performance modular data system for FT-ICR mass spectrometry. *Rapid Commun. Mass Spectrom.* **10,** 1839–1844
25. Gorovsky, M. A., Yao, M. C., Keevert, J. B., and Pleger, G. L. (1975) Isolation of micro- and macronuclei of *Tetrahymena pyriformis*. *Methods Cell Biol.* **9,** 311–327
26. Senko, M. W., Hendrickson, C. L., PasaTolic, L., Marto, J. A., White, F. M., Guan, S. H., and Marshall, A. G. (1996) Electrospray ionization Fourier transform ion cyclotron resonance at 9.4 T. *Rapid Commun. Mass Spectrom.* **10,** 1824–1828
27. Hendrickson, C. L., Quinn, J. P., Emmett, M. R., and Marshall, A. G. (2001) Mass-selective external ion accumulation for Fourier transform ion cyclotron resonance mass spectrometry, in *49th ASMS Conference on Mass Spectrometry and Allied Topics.* Chicago, IL.
28. Håkansson, K., Chalmers, M. J., Quinn, J. P., McFarland, M. A., Hendrickson, C. L., and Marshall, A. G. (2003) Combined electron capture and infrared multiphoton dissociation for multistage MS/MS in a Fourier transform ion cyclotron resonance mass spectrometer. *Anal. Chem.* **75,** 3256–3262
29. Zhang, S., Van Pelt, C. K., and Henion, J. D. (2003) Automated chip-based nanoelectrospray-mass spectrometry for rapid identification of proteins separated by two-dimensional gel electrophoresis. *Electrophoresis* **24,** 3620–3632
30. Zhang, L., Freitas, M. A., Wickham, J., Parthun, M. R., Klisovic, M. I., Marcucci, G., and Byrd, J. C. (2004) Differential expression of histone post-translational modifications in acute myeloid and chronic lymphocytic leukemia determined by high-pressure liquid chromatography and mass spectrometry. *J. Am. Soc. Mass Spectrom.* **15,** 77–86
31. Galasinski, S. C., Louie, D. F., Gloor, K. K., Resing, K. A., and Ahn, N. G. (2002) Global regulation of post-translational modifications on core histones. *J. Biol. Chem.* **277,** 2579–2588
32. Galasinski, S. C., Resing, K. A., and Ahn, N. G. (2003) Protein mass analysis of histones. *Methods* **31,** 3–11
33. Bonaldi, T., Regula, J. T., and Imhof, A. (2004) The use of mass spectrometry for the analysis of histone modifications. *Methods Enzymol.* **377,** 111–130
34. Strahl, B. D., Briggs, S. D., Brame, C. J., Caldwell, J. A., Koh, S. S., Ma, H., Cook, R. G., Shabanowitz, J., Hunt, D. F., Stallcup, M. R., and Allis, C. D. (2001) Methylation of histone H4 at arginine 3 occurs *in vivo* and is mediated by the nuclear receptor coactivator PRMT1. *Curr. Biol.* **11,** 996–1000
35. Smith, C. M., Haimberger, Z. W., Johnson, C. O., Wolf, A. J., Gafken, P. R., Zhang, Z., Parthun, M. R., and Gottschling, D. E. (2002) Heritable chromatin structure: Mapping "memory" in histones H3 and H4. *Proc. Natl. Acad. Sci. U. S. A.* **99 Suppl 4,** 16454–16461
36. Smith, C. M., Gafken, P. R., Zhang, Z., Gottschling, D. E., Smith, J. B., and Smith, D. L. (2003) Mass spectrometric quantification of acetylation at specific lysines within the amino-terminal tail of histone H4. *Anal. Biochem.* **316,** 23–33
37. Zhang, K., Tang, H., Huang, L., Blankenship, J. W., Jones, P. R., Xiang, F., Yau, P. M., and Burlingame, A. L. (2002) Identification of acetylation and methylation sites of histone H3 from chicken erythrocytes by high-accuracy matrix-assisted laser desorption ionization time-of-flight, matrix-assisted laser desorption ionization-postsource decay, and nanoelectrospray ionization tandem mass spectrometry. *Anal. Biochem.* **306,** 259–269
38. Kouach, M., Belaiche, D., Jaquinod, M., Couppez, M., Kmiecik, D., Ricart,

G., Van Dorsselaer, A., Sautiere, P., and Briand, G. (1994) Application of electrospray and fast atom bombardment mass spectrometry to the identification of post-translational and other chemical modifications of proteins and peptides. *Biol Mass Spectrom.* **23,** 283–294
39. Cocklin, R. R., and Wang, M. (2003) Identification of methylation and acetylation sites on mouse histone H3 using matrix-assisted laser desorption/ionization time-of-flight and nanoelectrospray ionization tandem mass spectrometry. *J. Protein Chem.* **22,** 327–334
40. Fischle, W., Wang, Y., and Allis, C. D. (2003) Histone and chromatin cross-talk. *Curr. Opin. Cell Biol.* **15,** 172–183
41. Mochizuki, K., Fine, N. A., Fujisawa, T., and Gorovsky, M. A. (2002) Analysis of a piwi-related gene implicates small RNAs in genome rearrangement in tetrahymena. *Cell* **110,** 689–699
42. Cheung, W. L., Ajiro, K., Samejima, K., Kloc, M., Cheung, P., Mizzen, C. A., Beeser, A., Etkin, L. D., Chernoff, J., Earnshaw, W. C., and Allis, C. D. (2003) Apoptotic phosphorylation of histone H2B is mediated by mammalian sterile twenty kinase. *Cell* **113,** 507–517
43. Jenuwein, T., and Allis, C. D. (2001) Translating the histone code. *Science* **293,** 1074–1080
44. Davis, M. C., Ward, J. G., Herrick, G., and Allis, C. D. (1992) Programmed nuclear death: Apoptotic-like degradation of specific nuclei in conjugating *Tetrahymena*. *Dev. Biol.* **154,** 419–432
45. Taylor, G. K., Kim, Y. B., Forbes, A. J., Meng, F., McCarthy, R., and Kelleher, N. L. (2003) Web and database software for identification of intact proteins using "top down" mass spectrometry. *Anal Chem.* **75,** 4081–4086
46. Pesavento, J. J., Kim, Y. B., Taylor, G. K., and Kelleher, N. L. (2004) Shotgun annotation of histone modifications: A new approach for streamlined characterization of proteins by top down mass spectrometry. *J. Am. Chem. Soc.* **126,** 3386–3387