Mapping Post-translational Modifications of the Histone Variant MacroH2A1 Using Tandem Mass Spectrometry*

Feixia Chu‡§¶‖**, Dmitri A. Nusinow¶‖‡‡, Robert J. Chalkley‡§‖, Kathrin Plath§§, Barbara Panning‡¶ ¶¶, and Alma L. Burlingame‡§‖‖

Post-translational histone modifications modulate chromatin-templated processes and therefore affect cellular proliferation, growth, and development. Although post-translational modifications on the core histones have been under intense investigation for several years, the modifications on variant histones are poorly understood. We used tandem mass spectrometry to identify covalent modifications on a histone H2A variant, macroH2A1.2. MacroH2A1.2 can be monoubiquitinated; however, the site of monoubiquitination has not been documented. In this study we used green fluorescent protein-tagged macroH2A1.2 to determine that Lys^{115} is a site of ubiquitination. In addition, we found that this variant H2A is methylated on the ϵ amino group of lysine residues Lys^{17}, Lys^{122}, and Lys^{238} and phosphorylated on Thr^{128}. Three of these modifications were also found to be present in the endogenous protein by mass spectrometric analysis. These results provide the first direct evidence that multiple post-translational modifications are imposed on macroH2A1.2, suggesting that, like canonical H2A, this variant H2A is subject to regulation by combinatorial use of covalent modifications. ***Molecular & Cellular Proteomics 5:194–203, 2006.***

All DNA-based processes in mammalian cells must contend with chromatin. DNA is wrapped around an octamer of four core histones, H2A, H2B, H3, and H4, to form nucleosomes, which in turn are subject to higher order packaging to form chromatin. Histones are subject to covalent modifications such as acetylation, methylation, phosphorylation, and ubiquitination. Different combinations of histone modifications characterize transcriptionally active euchromatin and transcriptionally silent heterochromatin (1, 2). These epigenetic marks provide binding sites for proteins that regulate gene expression and chromosome structure. Replacement of core histones with variant histones provides further structural diversity that may be used to modulate chromatin structure. The combination of regulated alterations in post-translational modifications of core histones and incorporation of histone variants affects chromatin architecture to modulate transcription, facilitate DNA repair and chromosome segregation, and define epigenetic states. Recently tandem mass spectrometry has proven advantageous in dealing effectively with the detection and assignment of this "combinatorial" post-translational complexity for individual sites as well as multisite occupancies at the protein level (3–5).

H2A is the core histone that has the most variants, and these variants are implicated in a number of processes. H2A.X is involved in DNA damage repair (6). H2A.Z is required for proper chromosome segregation and HP1α localization (7). H2A.BBD is associated with active chromatin (8), and macroH2A1.1, macroH2A1.2, and macroH2A2 are enriched on the facultative heterochromatin of the inactive X chromosome (9–11). MacroH2A1.1 and macroH2A1.2, which are produced by alternative splicing of the macroH2A1 gene (*H2AFY*, according to Human Genome Organization nomenclature), were first identified as histone variants that replace canonical H2A in an estimated 3% of nucleosomes (12). These are vertebrate-specific variants that contain a histone domain that is highly similar to H2A and a large, globular non-histone domain. The two isotypes of macroH2A1 are identical except for a 30-amino acid region in the non-histone domain. This non-histone domain is related to a family of proteins that includes polynucleotide hydrolases and nicotinamide dinucleotide metabolite-binding proteins (13, 14). MacroH2A1.1 and macroH2A1.2 differ in their expression patterns during development (15). MacroH2A1.1, but not macroH2A1.2, binds *O*-acetyl-ADP-ribose, a product of the nicotinamide dinucleotide-dependent SirT1 deacetylation reaction (14). In addition to enrichment on the inactive X chromosome, macroH2A1.2 is also enriched in regions of heterochromatin in senescent cells, suggesting a role in repression of transcription (9, 16). Depletion of macroH2A1.1 and macroH2A1.2 in female cells causes reactivation of a transgene on the inactive X chromosome, demonstrating a role for these variants in the maintenance of silenced chromatin (17). When macroH2A1.2 is incorporated into nucleosomes, transcription factor binding is

From the ‡Mass Spectrometry Facility and Departments of §Pharmaceutical Chemistry and ¶Biochemistry and Biophysics, University of California, San Francisco, California 94143 and the §§Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, Massachusetts 02139

Received, September 2, 2005, and in revised form, September 29, 2005

Published, MCP Papers in Press, October 5, 2005, DOI 10.1074/mcp.M500285-MCP200

inhibited, and the activity of chromatin remodeling factors is blocked, suggesting mechanisms by which this variant histone may facilitate transcriptional repression (18).

Canonical H2A is subject to several post-translational modifications (19–23). This suggests that the macroH2A family, which shows significant conservation with H2A, would also exhibit a similar range of covalent modifications. MacroH2A1.2 is ubiquitinated (17). In this work, we used capillary HPLC electrospray tandem mass spectrometry to identify a site of ubiquitination at Lys115 on immunopurified epitope-tagged macroH2A1.2. In addition, we detected methylation of Lys^{17}, Lys^{122}, and Lys^{238} and phosphorylation of Thr^{128}. Finally we showed the presence of three of these modifications on endogenous macroH2A1.2 by mass spectrometry. Thus, like the core histone H2A it replaces, macroH2A1.2 is subject to a variety of post-translational modifications. We speculate that the covalent modifications of macroH2A1.2 are a further elaboration of the histone code, providing additional mechanisms to regulate chromatin structure.

EXPERIMENTAL PROCEDURES

Cell Lines—Full-length human macroH2A1.2 cDNA was generated using the polymerase chain reaction and cloned into a backbone originating from pBOS-H2B-GFP (Clontech), generating a construct that expresses macroH2A1.2 tagged with green fluorescent protein (GFP)[1] on the C terminus. This plasmid was transfected into human embryonic kidney cells (293 cells) using FuGENE 6 (Roche Applied Science), and stable lines were selected using blasticidin S (Invitrogen). All 293 and 293-derived cell lines were maintained in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum and antibiotics.

Immunopurification—293 cells or 293:macroH2A1.2-GFP cells were washed in PBS and resuspended in lysis buffer (50 mM Tris, pH 7.4, 250 mM NaCl, 1 mM EDTA, 0.1% Triton X-100, 10 mM sodium butyrate, 1 mM phenylmethylsulfonyl fluoride, 1× Phosphatase Inhibitor Mixtures I and II (Sigma), 0.5 μg/ml leupeptin, 1 μg/ml aprotinin, and 0.7 μg/ml pepstatin). The resulting lysate was sonicated, and debris were removed by centrifugation. GFP-tagged macroH2A1.2 was affinity-purified from the supernatant from 293:macroH2A1.2-GFP cells using a rabbit polyclonal anti-GFP antibody (Abcam). Endogenous macroH2A1 was purified from the supernatant of 293 cells using rabbit polyclonal antibodies that recognize antibodies macroH2A1.1 and -1.2.[2] Immunoprecipitations were performed at 4 °C for 2 h and captured on protein A-coated magnetic beads (Dynal), and the beads were washed three times in room temperature lysis buffer.

SDS-PAGE Separation and In-gel Tryptic Digestion—Affinity-purified macroH2A1-containing samples were eluted from beads with sample buffer (0.24 M Tris, 8% SDS, 2.88 M β-mercaptoethanol, 40% glycerol, and 0.4% bromphenol blue) and loaded onto 4–20% SDS-PAGE Criterion Ready Gels (Bio-Rad) for separation. Proteins were visualized with Coomassie Brilliant Blue G-250. In-gel digestions on histone bands were performed utilizing a procedure described at ms-facility.ucsf.edu/ingel.html. Typically 50 ng of trypsin (porcine, side chain-protected; Promega, Madison, WI) was used for each gel band, and digestions were carried out at 37 °C for 2 h. Peptides were extracted from gel pieces with 50 μl of 50% acetonitrile, 2% formic acid thrice, and the extraction solution was dried down to ~10 μl.

On-line Capillary LC-MS and LC-MS-MS Analysis—A 1-μl aliquot of the digestion mixture was injected into an Ultimate capillary LC system via a FAMOS autosampler (LC Packings, Sunnyvale, CA) and separated by a 75-μm × 15-cm reverse-phase capillary column at a flow rate of ~330 nl/min. The HPLC eluent was connected directly to the micro-ion electrospray source of a QSTAR XL mass spectrometer (Applied Biosystems/MDS Sciex, Foster City, CA). Typical performance characteristics were >8000 resolution with 30-ppm mass measurement accuracy in both MS and CID spectra. LC-MS data were acquired in an information-dependent acquisition mode, cycling between 1-s MS acquisition followed by 3-s low energy CID data acquisition. The centroided peak lists of the CID spectra were searched against the National Center for Biotechnology Information (NCBI) protein database using Batch-Tag, a program in the in-house version of the University of California San Francisco ProteinProspector package (24), considering protein N terminus and lysine acetylation; lysine mono-, di-, and trimethylation; arginine methylation; phosphorylation; and lysine ubiquitination as variable modifications. The CID spectra with putative post-translational modifications were further inspected manually.

Immunofluorescence and Fluorescence in Situ Hybridization—Cells were cultured on glass coverslips, fixed in 4% paraformaldehyde for 10 min, extracted in PBS containing 0.5% Triton X-100 for 10 min, washed in PBS containing 0.2% Tween 20 (PBST), and then placed in a block solution (PBS, 0.2% cold water fish skin gelatin, 5% normal goat serum, and 0.2% Tween 20) for 20 min. The cells were then incubated with anti-macroH2A1 antibody diluted into block solution (1:500) for 1 h, washed three times in PBST, and then incubated with Texas Red-conjugated anti-rabbit antibody (Vector Laboratories) for 30 min. Coverslips were then washed in PBST, stained with 4,6-diamidino-2-phenylindole dihydrochloride in PBST, washed in PBST, and mounted in Vectashield (Vector Laboratories). Fluorescence *in situ* hybridization for *XIST* RNA using directly labeled DNA probes was carried out as described previously (25).

RESULTS

Expression and Purification of GFP-tagged MacroH2A1.2—To generate a source of macroH2A1.2 for affinity purification and mass spectrometry, we generated 293:macroH2A1.2-GFP, a transformed human embryonic kidney cell line stably expressing macroH2A1.2 tagged with GFP on the C terminus. To determine whether the addition of GFP to macroH2A1.2 adversely affected expression or localization, GFP fluorescence was used to analyze the intracellular localization of the fusion protein. For comparison, the distribution of endogenous macroH2A1 in the parental 293 cell line was determined by immunofluorescence using a rabbit polyclonal antibody that recognizes macroH2A1.1 and macroH2A1.2. Like endogenous macroH2A1, macroH2A1.2-GFP was distributed uniformly throughout most of the nucleus and was enriched in one to two large contiguous domains, which coincided with the two *XIST* RNA-coated inactive X chromosomes in this aneuploid cell line (Fig. 1*A*). Therefore addition of GFP to macroH2A1.2 did not noticeably alter the localization properties of this variant histone.

To determine whether macroH2A1.2-GFP was correctly incorporated into chromatin, macroH2A1.2-GFP was immuno-

[1] The abbreviations used are: used: GFP, green fluorescent protein; XIC, extracted ion chromatogram; me1, monomethylation; me2, dimethylation; P, phosphorylation; E3, ubiquitin-protein isopeptide ligase.

[2] K. Plath and B. Panning, in preparation.

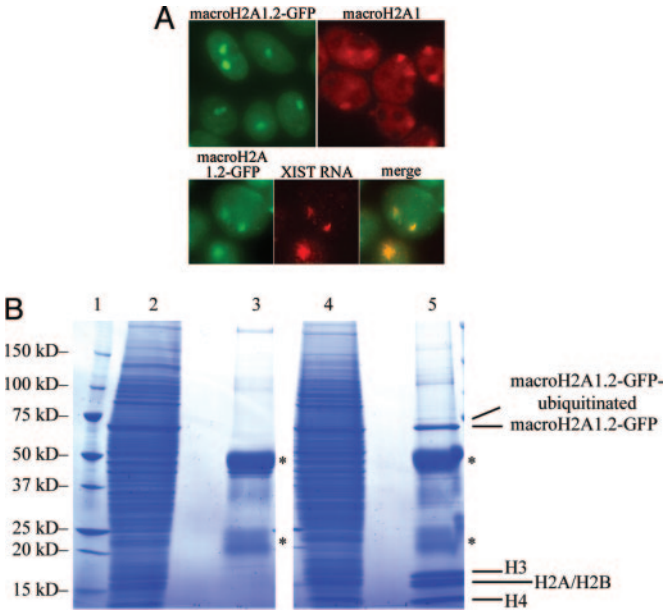

FIG. 1. **Localization and purification of GFP-tagged macroH2A1.2.** *A*, macroH2A1.2-GFP is enriched on the inactive X chromosome. *Upper panels*, GFP fluorescence indicates the enrichment of the macroH2A1.2-GFP (*left*, *green*) in two large domains. Immunostaining shows similar distribution of endogenous protein (*right*, *red*) in 293 cells. *Lower panels*, macroH2A1.2-GFP fluorescence (*green*) and *XIST* RNA fluorescence *in situ* hybridization (*red*). Merge (*right panel*) shows that the enriched regions of macroH2A1.2-GFP coincide with the inactive X chromosome. *B*, GFP immunopurification from 293 cells expressing macroH2A1.2-GFP. Immunoprecipitation of macroH2A1.2-GFP also co-precipitates other core histones, demonstrating that macroH2A1.2-GFP is incorporated into chromatin (*lane 5*). Control immunoprecipitation from the untagged parental 293 cells does not pull down either macroH2A1 or histones (*lane 3*). * is IgG heavy chain and light chain.

purified using an anti-GFP antibody (Fig. 1*B*). SDS-PAGE separation of the proteins that bound to anti-GFP antibody yielded two bands of ~70 kDa, the expected size for the macroH2A1.2-GFP fusion protein, that were both identified as macroH2A1.2 by Western blotting and LC-MS-MS analysis (data not shown). In addition to macroH2A1.2-GFP, there was the quartet of low molecular weight bands that is characteristic of the core histones and that were identified as H3, H2A, H2B, and H4 via tandem mass spectrometric analysis (data not shown). Co-purification of the core histone proteins with macroH2A1.2-GFP indicates that tagged macroH2A1.2-GFP was appropriately incorporated into chromatin. Because addition of a C-terminal GFP tag to macroH2A1.2 did not alter the characteristics of this variant histone as assayed by its incorporation into chromatin and localization, we used affinity-purified macroH2A1.2-GFP to identify post-translational modifications by tandem mass spectrometry.

Identification of Post-translational Modifications on GFP-tagged MacroH2A1.2—LC-MS-MS analysis of the tryptic digestion mixture eluted from the upper macroH2A1.2-GFP band identified macroH2A1.2 peptides and ubiquitin, suggesting ubiquitination of macroH2A1.2. Modified peptides can be identified by their characteristic mass changes due to modification, and sites of modification can be determined through fragmentation analysis. ProteinProspector reported one spectrum was potentially of a ubiquitinated peptide. The CID spectrum of a peptide with *m/z* 743.42^{3+} (Table I and Fig. 2) displayed partial C- and N-terminal sequence ion ladders (y_2 to y_{17} and b_2 to b_4) that matched a tryptic peptide of macroH2A1.2 spanning residues Gly^{96} to Lys^{116}. However, the molecular weight of this peptide was 114 amu higher than the unmodified peptide Gly^{96}–Lys^{116}, which equaled the mass of two glycine residues that remain after tryptic digestion of the ubiquitin appended to the ϵ amino group of a lysine residue (26). The presence of the y_1 ion at *m/z* 147.11 showed the C-terminal lysine (Lys^{116}) was unmodified. Meanwhile the y_2 ion and all subsequent y ions had a mass 114 amu higher than predicted for the unmodified peptide. The 242-amu mass difference between y_2 and y_1 matches the residue mass of a ϵ-(Gly-Gly)-amino lysine, thus identifying Lys^{115} as a site of ubiquitination. The difference in size and amount between the more abundant, faster migrating form of macroH2A1.2-GFP and the slower migrating, ubiquitin-containing form of macroH2A1.2-GFP is consistent with monoubiquitination of a subset of the total macroH2A1.2-GFP.

TABLE I
Modified peptides observed in the digestion mixture of macroH2A1.2-GFP

Experimental *m/z*	Observed mass error	Peptide sequences	Modified residues and modifications
	ppm		
743.4297^{3+}	6	Gly^{96}–Lys^{116}	$Lys^{115}ub^{a}$
405.9119^{3+}	12	Ser^{15}–Arg^{26}	$Lys^{17}me1$
497.6343^{3+}	−19	Gly^{121}–Lys^{134}	$Lys^{122}me2$
514.9697^{3+}	28	Gly^{121}–Lys^{134}	$Thr^{128}P$
530.2969^{3+}	−10	Lys^{235}–Arg^{248}	$Lys^{238}me1$
534.9685^{3+}	−10	Lys^{235}–Lys^{248}	$Lys^{238}me2$

[a] ub stands for ubiquitination.

LC-MS-MS analysis of the tryptic digest of the major macroH2A1.2-GFP band of slightly lower molecular weight showed 89 CID spectra matching the macroH2A1.2 sequence with no ubiquitin peptides. ProteinProspector reported several potentially modified peptides, and manual inspection of the CID spectra of these peptides revealed monomethylation of Lys^{17}, dimethylation of Lys^{122}, phosphorylation of Thr^{128}, and mono- and dimethylation of Lys^{238} on macroH2A1.2-GFP (Table I).

In the CID spectrum of a peptide with a *m/z* value of 405.91^{3+} (Fig. 3*A*), the mass values of the C-terminal sequence ions up to y_9 (y_1 to y_9) matched the anticipated values of the unmodified peptide Ser^{15}–Arg^{26}, suggesting all residues C-terminal from Ala^{18} were unmodified. Meanwhile the masses of N-terminal sequence ions starting from b_4 (b_4 to b_8)

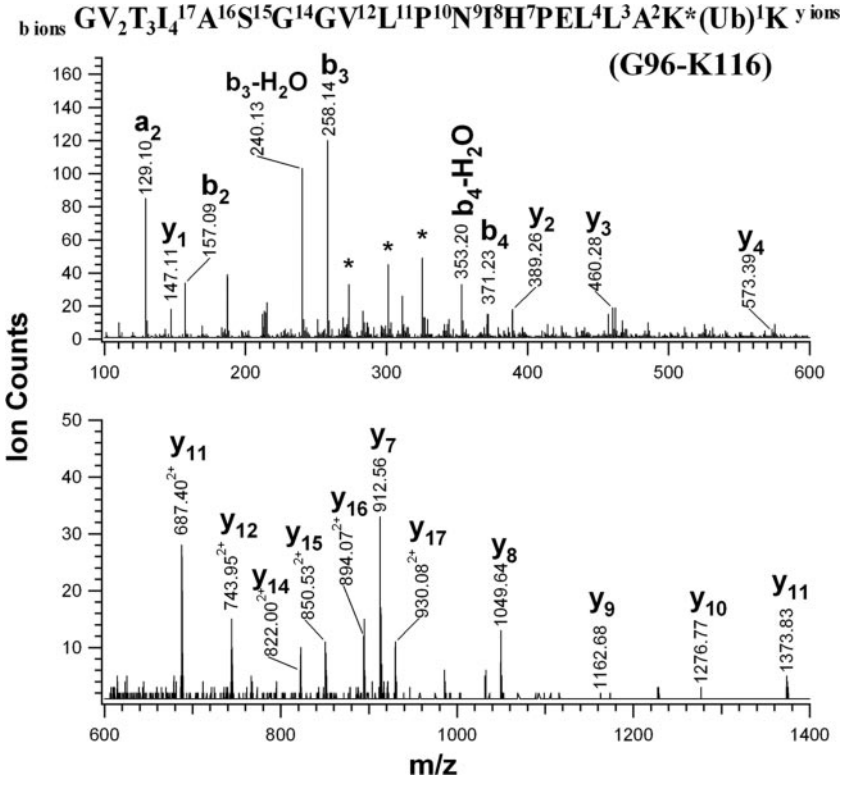

FIG. 2. **Structure and low energy CID spectrum of the modified peptide with an *m/z* values at 743. 43^{3+} from the digestion mixture of tagged macroH2A1.2-GFP.** The ordinates for *panels A* and *C* represent the number of iron counts.

displayed a mass change of +14 amu, suggesting a monomethylation within the N-terminal ^{15}SAKG^{18} moiety. Combining information derived from y and b ions, Lys^{17} can be assigned as the monomethylation site.

The identity and structure of the peptide with an *m/z* value of 497.63^{3+} was determined in a similar fashion (Fig. 3*B*). Essentially the complimentary and continuous C- and N-terminal fragment ion ladders (y_1 to y_8 and b_2 to b_8, respectively) provided sufficient information for the conclusive assignment of the peptide sequence. However, the molecular weight of the precursor ion was 28 amu higher than that of the unmodified peptide Gly^{121}–Lys^{134}. The masses of the C-terminal sequence ions (y_1 to y_8) were the anticipated values, showing that all residues C-terminal of Ile^{126} were unmodified. On the other hand, the addition of 28 amu on b ions (b_2 to b_8), relative to the unmodified peptide, revealed that the first two N-terminal residues of this peptide, Gly^{121} and Lys^{122}, were either dimethylated or doubly monomethylated. The absence of a side chain in glycine residues precludes the possibility for modifications on Gly^{121}, indicating that Lys^{122} is dimethylated.

The precursor ion at *m/z* 514.97^{3+} had a mass increment of +80 amu compared with that of the unmodified peptide Gly^{121}–Lys^{134}, indicating phosphorylation or sulfation of this peptide (Fig. 3*C*). The presence of a 98-amu neutral loss from the precursor ion at *m/z* 482.32^{3+} suggested phosphorylation rather than sulfation (27). The mass values of the first six C-terminal sequence ions (y_1 to y_6) and the first six N-terminal sequence ions (b_2 to b_7) were identical to those of the unmodified peptide; thus residues from Pro^{129} to Lys^{134} and from Gly^{121} to Ile^{127} were not modified. Therefore, Thr^{128} remained the only site for phosphorylation. As expected, C- and N-terminal sequence ions containing Thr^{128} all had the characteristic H_3PO_4 neutral loss (y_7-98 to $y_{12}-98$ and b_8-98). The absence of any Thr^{128}-containing fragment ions that matched the masses of the unmodified peptide further confirmed this assignment. This peptide also existed in the tryptic digest of monoubiquitinated macroH2A1.2-GFP, suggesting these two post-translational modifications are not mutually exclusive.

The CID spectra of the peptides with *m/z* of 530.30^{3+} and 534.97^{3+} (Fig. 4) correspond in mass to singly and doubly methylated versions of the peptide Lys^{235}–Arg^{248}. Both spectra exhibited the first seven C-terminal sequence ions (y_1 to y_7) with mass values identical to those of the unmodified peptide, suggesting that the Glu^{242}–Arg^{248} peptide moiety is not modified. Similarly, although weak, the mass of b_3 ion at *m/z* 243.2 demonstrated that the first three N-terminal residues, Lys^{235}–Gly^{237}, remained unmodified. The mass differences between b_4 and b_3 ions, 142 amu for the peptide with an *m/z* of 530.30^{3+} (Fig. 4*A*) and 156 amu for the peptide with an *m/z* of 534.97^{3+} (Fig. 4*B*), correspond to the residue mass of mono- and dimethylated lysines, respectively. Furthermore, when compared with those of the unmodified peptide, b ions containing Lys^{238} (b_4 to b_9) all displayed a mass increment of 14 amu for the peptide with an *m/z* of 530.30^{3+} or 28 amu for the peptide with an *m/z* of 534.97^{3+} (Fig. 4, *A* and *B*). Therefore,

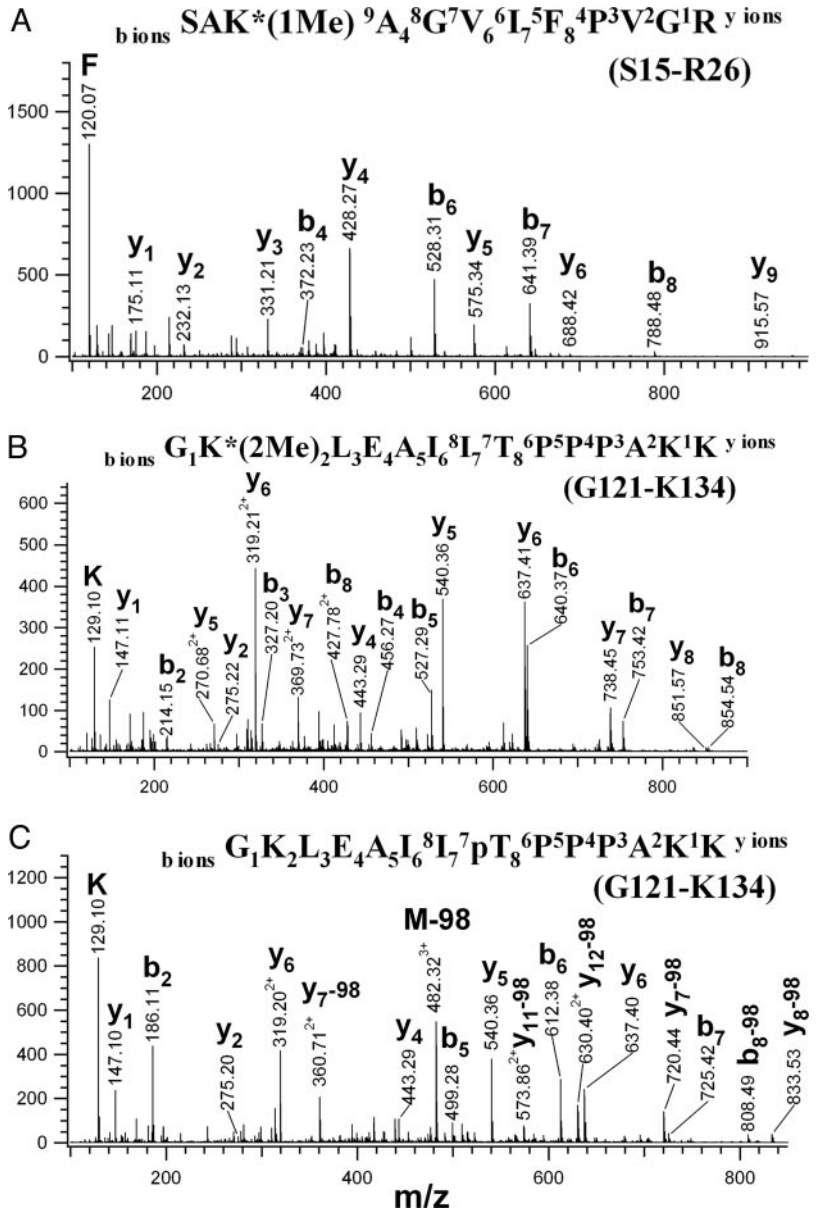

FIG. 3. **Structure and low energy CID spectra of the modified peptides with *m/z* values at 405. 91^{3+} (*A*), 497.63^{3+} (*B*), and 514.97^{3+} (*C*) from the digestion mixture of tagged macroH2A1.2-GFP.** The ordinates for *panels A–C* represent the number of ion counts.

we conclude that Lys^{238} occurs in mono- and dimethylated forms.

Post-translational Modifications on Endogenous MacroH2A1—The phosphorylation and methylation of macroH2A1.2-GFP may reflect biologically interesting modifications that occur on the endogenous protein or may be modifications that are consequences of overexpression or tagging of macroH2A1.2. To determine whether endogenous material is subject to these post-translational modifications, we used tandem mass spectrometry to analyze the tryptic peptides produced from macroH2A1 immunopurified from 293 cells with a rabbit polyclonal anti-macroH2A1 antibody, which recognizes macroH2A1.1 and macroH2A1.2 (data not shown). Although the macroH2A1 antiserum was not as efficient for immunoprecipitation as the GFP antiserum and less total protein was purified, LC-MS-MS analysis revealed macroH2A1 as the principal component in the sample.

Evidence for three of the modifications detected in the tagged protein was found in the analysis of the endogenous protein. The peptide containing monomethylated Lys^{17} was detected and identified by MS-MS fragmentation. In addition, peptides with the masses of modified peptides containing Lys^{122} dimethylation and Thr^{128} phosphorylation were also detected, although they were not intense enough for fragmentation analysis (Table I). For example, Fig. 5*A* shows an extracted ion chromatogram (XIC) for ions of *m/z* 405.9. An XIC plot shows when during the LC run ions of a particular *m/z* were observed. The *m/z* 405.9 corresponds to the *m/z* of the triply charged peptide containing monomethylated Lys^{17} that was observed in the macroH2A1.2-GFP digest, and a peak is

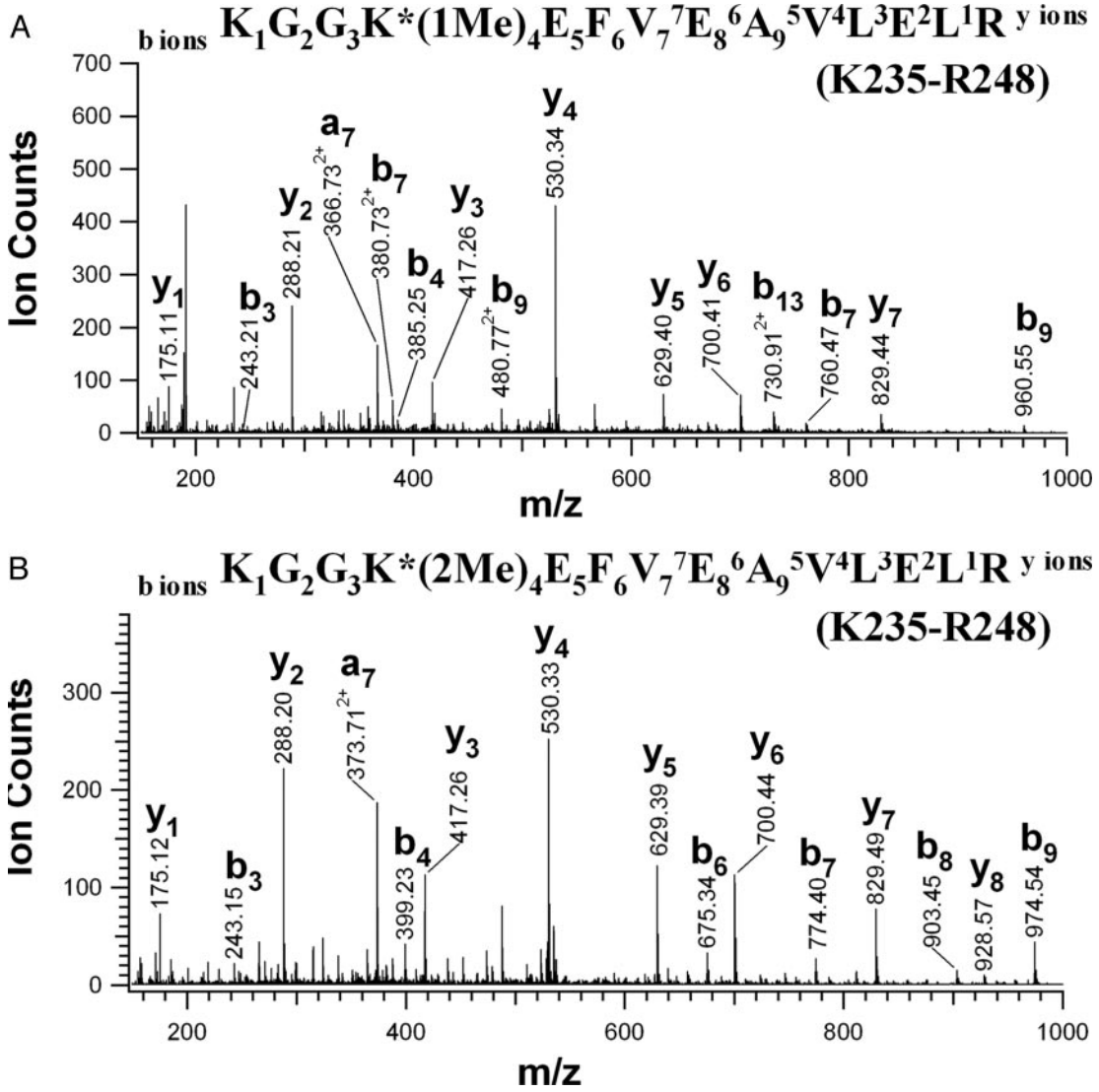

FIG. 4. **Structure and low energy CID spectrum of the modified peptide with an *m/z* value at 530. 30^{3+} (*A*) and 534.97^{3+} (*B*) from the digestion mixture of tagged macroH2A1.2-GFP.**

clearly identified at ≈23.8 min (Fig. 5*A*). The mass spectrum of the peptides eluting at this time point shows that *m/z* 405.9^{3+} co-eluted with peptides giving signals at *m/z* 445.6^{3+}, 519.0^{4+}, and 601.3^{2+} (Fig. 5*A*, *inset*). Similarly the XIC for *m/z* 405.9 from the endogenous macroH2A1 digest also exhibited a peak at ≈23.8 min (Fig. 5*B*) that co-eluted with peptides giving signals at *m/z* 445.6^{3+}, 519.0^{4+}, and 601.3^{2+} (Fig. 5*B*, *inset*). This indistinguishable retention time observed in the chromatographic elution profiles of the digestion mixtures of macroH2A1.2-GFP and macroH2A1 suggests the identical chemical composition for the species with *m/z* 405.9^{3+}. In addition, the tandem mass spectrum of the peptide at *m/z* 405.9^{3+} in the macroH2A1 digest (Fig. 5*D*) contained the same major fragment ions as the MS-MS spectrum of this peptide from the macroH2A1.2-GFP digest (Fig. 5*C*). Therefore, monomethylation at Lys^{17} was confidently identified in endogenous macroH2A1. Peptides with the same masses as the peptides containing dimethylated Lys^{122} and phosphorylated Thr^{128} observed in the tagged protein were also detected in macroH2A1 with the same retention times. Although not definitive proof, this is strong evidence that these two modifications are also present in endogenous protein. The peptides corresponding to mono- and dimethylation of Lys^{238} were not observed in the macroH2A1 digest.

DISCUSSION

Post-translational modifications of histones affect most processes that involve chromatin, including transcription, replication, DNA repair, and DNA segregation. This is also true of histone variants, which recruit DNA repair machinery to lesions, differentiate chromatin at centromeres, and define transcriptionally active or silent domains. However, it remains

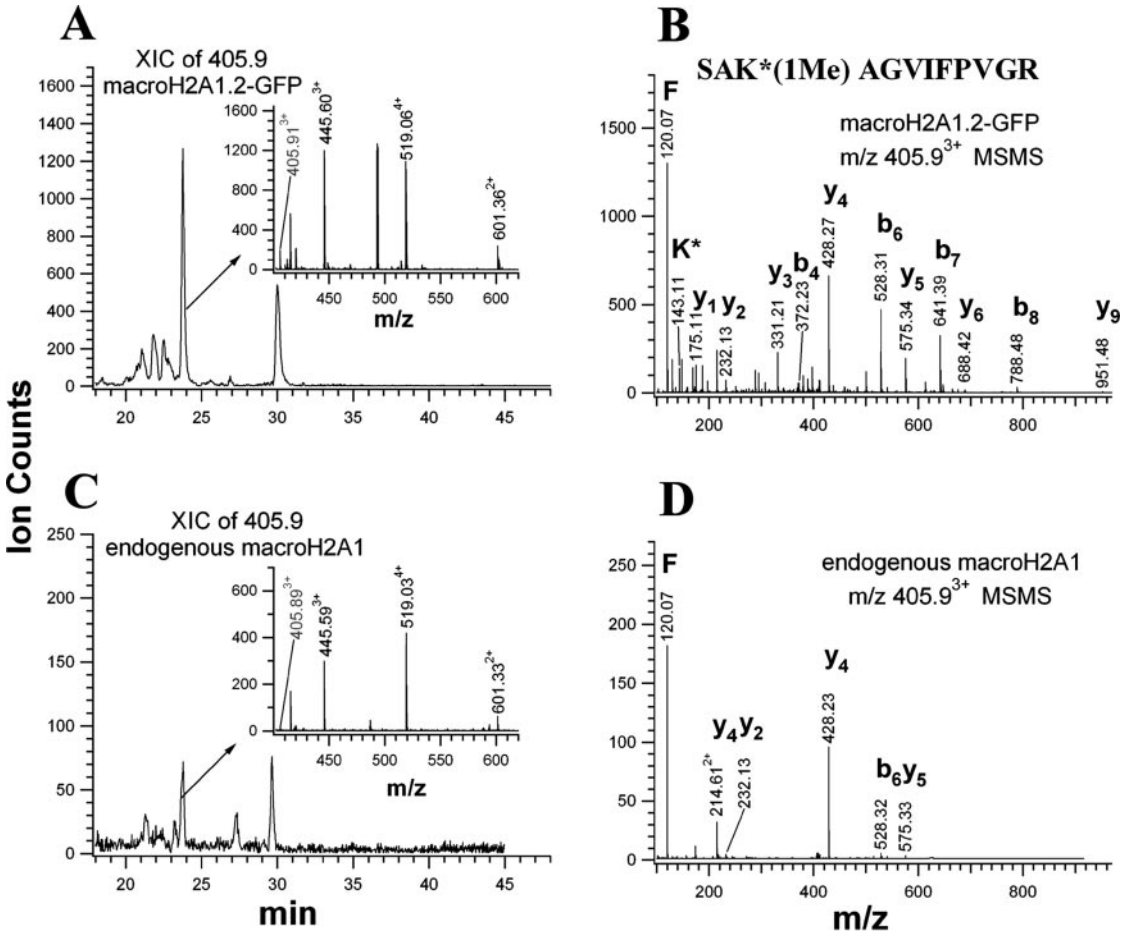

FIG. 5. **The peptide containing monomethylated Lys^{17} with an *m/z* value of 405. 9^{3+} existed both in the tagged macroH2A1.2-GFP digest and the endogenous macroH2A1 digest.** *A* and *C*, XIC of *m/z* value 405.9 from the tryptic digestion mixture of the tagged macroH2A1.2-GFP (*A*) and the endogenous macroH2A1 (*B*). *Insets* are the mass spectra of the peptides eluted at ≈23.8 min. *B* and *D*, CID spectra of the peptide from the tagged macroH2A1.2-GFP digest (*B*) and the endogenous macroH2A1 digest (*D*).

unclear whether variant histones, which show considerable sequence conservation with the core histones they replace (28), use the same set of modifications as their canonical counterparts. In this study, we investigated covalent modifications on a histone H2A variant, macroH2A1.2, using liquid chromatographic and mass spectrometric analysis of affinity-tagged material. MacroH2A1.2-GFP was ubiquitinated on Lys^{115}. The difference in electrophoretic mobility of the ubiquitinated and unubiquitinated forms was consistent with monoubiquitination of this protein. Monoubiquitination has been reported for macroH2A1 (17), suggesting that Lys^{115} is the site of this modification on the endogenous protein. In addition, Lys^{17} monomethylation (Lys^{17}me1), Lys^{122} dimethylation (Lys^{122}me2), and Thr^{128} phosphorylation (Thr^{128}P), and Lys^{238} mono- and dimethylation (Lys^{238}me1 and Lys^{238}me2) were identified in macroH2A1.2-GFP.

We confirmed that Lys^{17}me1, Lys^{122}me2, and Thr^{128}P occurred on endogenous material using tandem mass spectrometry to analyze protein immunopurified with antiserum that recognizes both isoforms of macroH2A1. Separation of macroH2A1.1 and macroH2A1.2 prior to mass spectrometric analysis will be necessary to determine whether these modifications occur on one or both of these isoforms. The peptides containing Lys^{238}me1 and Lys^{238}me2 were not detected on endogenous macroH2A1. The peptides containing these modifications were less abundant in the macroH2A1.2-GFP preparation, and limiting quantities of immunopurified macroH2A1 may have precluded their detection. MacroH2A1.2-GFP was correctly incorporated into chromatin and displayed its hallmark enrichment on the inactive X chromosome, suggesting that Lys^{238}me1 and Lys^{238}me2 are also likely to be of biological relevance.

Comparison of the post-translational modifications that occur on the histone domain of macroH2A1 and canonical histone H2A reveals a divergent modification map without any identical modifications between these histones (Fig. 6). This lack of overlap in covalent modifications is in part due to the discrepancy in their primary amino acid sequences. Six of the 10 documented modification sites in canonical H2A, Lys^{5}, Lys^{13}, Lys^{15}, Lys^{95}, Arg^{99}, and Thr^{120}, are not conserved in

FIG. 6. **Sequence alignments and post-translational modifications of H2A and macroH2A1.** *A*, sequence alignment of human histone H2A.1 and the histone domain of variant macroH2A1.1 and macroH2A1.2. Reported post-translational modification sites on H2A (*above*) and new modifications sites on the histone domain of macroH2A1 (*below*) are shown. In *red* are residues that are different between H2A and macroH2A1. *B*, sequence alignment of the non-histone domains of macroH2A1.2 and macroH2A1.1. Post-translational modifications on macroH2A1.2 are annotated *above* the alignment. Shaded in *blue* is the region of non-identity between macroH2A1.1 and macroH2A1.2. *Ub*, ubiquitination; *Ac*, acetylation; *Me*, methylation; *P*, phosphorylation.

macroH2A1, and one of the three modification sites in the histone domain of macroH2A1, Lys^{17}, is not conserved in H2A. Lys^{36} and Lys^{119} ($H2ALys^{36}$ and $H2ALys^{119}$) are conserved, and only the peptides with unmodified residues were observed in the LC-MS analysis of macroH2A1. Lys^{36} resides in the region where the major structural difference between macroH2A1- and H2A-containing nucleosomes occurs (29), suggesting that this region is important for regulation of chromatin structure. The sequences surrounding $H2ALys^{36}$ are not conserved in macroH2A1, suggesting that the histone acetyltransferase that recognizes $H2ALys^{36}$ may not recognize the corresponding residue in macroH2A1. In addition, the non-histone domain of macroH2A1.2 interacts with histone deacetylases 1 and 2 (29), which may inhibit accumulation of acetylation on macroH2A1. Consistent with this idea, acetylation was not detected on macroH2A1.2-GFP or macroH2A1 and was detected on H2A peptides (data not shown). The lack of acetylation on macroH2A1 may also be due to its enrichment in silent chromatin, which is generally hypoacetylated (29).

$H2ALys^{119}$ can be ubiquitinated and acetylated, and ubiquitination was detected on Lys^{115} of macroH2A1.2-GFP, which is adjacent to the residue of macroH2A1 that corresponds to $H2ALys^{119}$. Ubiquitination of $H2ALys^{119}$ is mediated by Ring1 and Ring1b, two mammalian Polycomb proteins that are enriched on the inactive X chromosome (21, 30–32). H2A ubiquitinated at Lys^{119} is enriched on the inactive X chromosome, and this enrichment first occurs when transcriptional silencing is initiated (21, 31, 32). This suggests the possibility that macroH2A1 Lys^{115} ubiquitination may also be mediated by Polycomb proteins and may contribute to silencing of the inactive X chromosome. In addition to monoubiquitination, macroH2A1.2 is also subject to poly- or multiubiquitination mediated by a CULLIN3/SPOP ubiquitin E3 ligase complex (17). With the identification of the monoubiquitination site, it will be possible to determine whether the same residue in macroH2A1.2 is used as the site for both types of ubiquitination. Lys^{122}me2 and Thr^{128}P reside in the region of macroH2A1 that binds SPOP, which plays a role in

the correct localization of this variant histone (17, 33). These modifications are correctly positioned to regulate the macroH2A1.2-SPOP interaction and modulate localization of macroH2A1.2.

Histone variants are under intense investigation and are recognized for their roles in transcription, DNA segregation, DNA repair, and chromatin structure. As modifications of canonical histones are also involved in many aspects of chromatin metabolism, the post-translational modifications on histone variants will likely prove to be equally important. Already H2A.X phosphorylation is recognized as a critical step during DNA damage repair (34), and acetyl-lysine in H3.3 is associated with domains of active transcription in *Drosophila* (35). Vertebrate-specific macroH2A1 is unique among histone variants due to its large C-terminal non-histone domain. It is associated with facultative heterochromatin and plays a role in transcriptional silencing. Determining how these newly discovered modifications affect localization and function of macroH2A1 will provide new insights into this very interesting family of H2A variants.

Acknowledgments—We thank Tom G. Fazzio and Judith A. Sharp for helpful discussions.

* This work was supported by National Institutes of Health National Center for Research Resources (NCRR) Grants 01614, 12961, and 15804 (to A. L. B.) and by National Institutes of Health Grant RO1 GM63671 and the Pew Biomedical Scholars Program (to B. P.). 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 article (available at http://www.mcponline.org) contains supplemental material.

‖ These authors contributed equally to this work.

** A Rett Syndrome Research Foundation postdoctoral fellow.

‡‡ Recipient of the University of California Dissertation Year Fellowship.

¶¶ To whom correspondence may be addressed: Dept. of Biochemistry and Biophysics, Genentech Hall S372B, 600 16th St., University of California, San Francisco, CA 94143-2200. E-mail: bpanning@biochem.ucsf.edu.

‖‖ To whom correspondence may be addressed: Dept. of Pharmaceutical Chemistry, 521 Parnassus Ave., Rm. C18, University of California, San Francisco, CA 94143-0446. E-mail: alb@cgl.ucsf.edu.

REFERENCES

1. Jenuwein, T., and Allis, C. D. (2001) Translating the histone code. *Science* **293,** 1074–1080
2. Fischle, W., Wang, Y., and Allis, C. D. (2003) Histone and chromatin cross-talk. *Curr. Opin. Cell Biol.* **15,** 172–183
3. Freitas, M. A., Sklenar, A. R., and Parthun, M. R. (2004) Application of mass spectrometry to the identification and quantification of histone post-translational modifications. *J. Cell. Biochem.* **92,** 691–700
4. Medzihradszky, K. F., Zhang, X., Chalkley, R. J., Guan, S., McFarland, M. A., Chalmers, M. J., Marshall, A. G., Diaz, R. L., Allis, C. D., and Burlingame, A. L. (2004) Characterization of *Tetrahymena* histone H2B variants and posttranslational populations by electron capture dissociation (ECD) Fourier transform ion cyclotron mass spectrometry (FT-ICR MS). *Mol. Cell. Proteomics* **3,** 872–886
5. Burlingame, A. L., Zhang, X., and Chalkley, R. J. (2005) Mass spectrometric analysis of histone posttranslational modifications. *Methods* **36,** 383–394
6. Foster, E. R., and Downs, J. A. (2005) Histone H2A phosphorylation in DNA double-strand break repair. *FEBS J.* **272,** 3231–3240
7. Rangasamy, D., Greaves, I., and Tremethick, D. J. (2004) RNA interference demonstrates a novel role for H2A.Z in chromosome segregation. *Nat. Struct. Mol. Biol*. **11,** 650–655
8. Chadwick, B. P., and Willard, H. F. (2001) A novel chromatin protein, distantly related to histone H2A, is largely excluded from the inactive X chromosome. *J. Cell Biol.* **152,** 375–384
9. Costanzi, C., and Pehrson, J. R. (1998) Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. *Nature* **393,** 599–601
10. Costanzi, C., and Pehrson, J. R. (2001) MACROH2A2, a new member of the MARCOH2A core histone family. *J. Biol. Chem.* **276,** 21776–21784
11. Chadwick, B. P., and Willard, H. F. (2001) Histone H2A variants and the inactive X chromosome: identification of a second macroH2A variant. *Hum. Mol. Genet*. **10,** 1101–1113
12. Pehrson, J. R., and Fried, V. A. (1992) MacroH2A, a core histone containing a large nonhistone region. *Science* **257,** 1398–1400
13. Allen, M. D., Buckle, A. M., Cordell, S. C., Lowe, J., and Bycroft, M. (2003) The crystal structure of AF1521 a protein from Archaeoglobus fulgidus with homology to the non-histone domain of macroH2A. *J. Mol. Biol.* **330,** 503–511
14. Kustatscher, G., Hothorn, M., Pugieux, C., Scheffzek, K., and Ladurner, A. G. (2005) Splicing regulates NAD metabolite binding to histone macroH2A. *Nat. Struct. Mol. Biol.* **12,** 624–625
15. Pehrson, J. R., Costanzi, C., and Dharia, C. (1997) Developmental and tissue expression patterns of histone macroH2A1 subtypes. *J. Cell. Biochem.* **65,** 107–113
16. Zhang, R., Poustovoitov, M. V., Ye, X., Santos, H. A., Chen, W., Daganzo, S. M., Erzberger, J. P., Serebriiskii, I. G., Canutescu, A. A., Dunbrack, R. L., Pehrson, J. R., Berger, J. M., Kaufman, P. D., and Adams, P. D. (2005) Formation of MacroH2A-containing senescence-associated heterochromatin foci and senescence driven by ASF1a and HIRA. *Dev. Cell* **8,** 19–30
17. Hernandez-Munoz, I., Lund, A. H., van der Stoop, P., Boutsma, E., Muijrers, I., Verhoeven, E., Nusinow, D. A., Panning, B., Marahrens, Y., and van Lohuizen, M. (2005) Stable X chromosome inactivation involves the PRC1 Polycomb complex and requires histone MACROH2A1 and the CULLIN3/SPOP ubiquitin E3 ligase. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 7635–7640
18. Angelov, D., Molla, A., Perche, P. Y., Hans, F., Cote, J., Khochbin, S., Bouvet, P., and Dimitrov, S. (2003) The histone variant macroH2A interferes with transcription factor binding and SWI/SNF nucleosome remodeling. *Mol. Cell* **11,** 1033–1041
19. 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
20. Zhang, Y., Griffin, K., Mondal, N., and Parvin, J. D. (2004) Phosphorylation of histone H2A inhibits transcription on chromatin templates. *J. Biol. Chem.* **279,** 21866–21872
21. de Napoles, M., Mermoud, J. E., Wakao, R., Tang, Y. A., Endoh, M., Appanah, R., Nesterova, T. B., Silva, J., Otte, A. P., Vidal, M., Koseki, H., and Brockdorff, N. (2004) Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. *Dev. Cell* **7,** 663–676
22. Turner, B. M. (2002) Cellular memory and the histone code. *Cell* **111,** 285–291
23. Barber, C. M., Turner, F. B., Wang, Y., Hagstrom, K., Taverna, S. D., Mollah, S., Ueberheide, B., Meyer, B. J., Hunt, D. F., Cheung, P., and Allis, C. D. (2004) The enhancement of histone H4 and H2A serine 1 phosphorylation during mitosis and S-phase is evolutionarily conserved. *Chromosoma* **112,** 360–371
24. Chalkley, R. J., Baker, P. R., Huang, L., Hansen, K. C., Allen, N. P., Rexach, M., and Burlingame, A. L. (2005) Comprehensive analysis of a multidimensional liquid chromatography mass spectrometry dataset acquired on a quadrupole selecting, quadrupole collision cell, time-of-flight mass spectrometer: II. new developments in Protein Prospector allow for reliable and comprehensive automatic analysis of large datasets. *Mol. Cell. Proteomics* **4,** 1194–1204
25. Plath, K., Fang, J., Mlynarczyk-Evans, S. K., Cao, R., Worringer, K. A., Wang, H., de la Cruz, C. C., Otte, A. P., Panning, B., and Zhang, Y. (2003) Role of histone H3 lysine 27 methylation in X inactivation. *Science* **300,**

131–135
26. Marotti, L. A., Jr., Newitt, R., Wang, Y., Aebersold, R., and Dohlman, H. G. (2002) Direct identification of a G protein ubiquitination site by mass spectrometry. *Biochemistry* **41,** 5067–5074
27. Medzihradszky, K. F., Darula, Z., Perlson, E., Fainzilber, M., Chalkley, R. J., Ball, H., Greenbaum, D., Bogyo, M., Tyson, D. R., Bradshaw, R. A., and Burlingame, A. L. (2004) *O*-Sulfonation of serine and threonine: mass spectrometric detection and characterization of a new posttranslational modification in diverse proteins throughout the eukaryotes. *Mol. Cell. Proteomics* **3,** 429–440
28. Malik, H. S., and Henikoff, S. (2003) Phylogenomics of the nucleosome. *Nat. Struct. Biol.* **10,** 882–891
29. Chakravarthy, S., Gundimella, S. K., Caron, C., Perche, P. Y., Pehrson, J. R., Khochbin, S., and Luger, K. (2005) Structural characterization of the histone variant macroH2A. *Mol. Cell. Biol.* **25,** 7616–7624
30. Plath, K., Talbot, D., Hamer, K. M., Otte, A. P., Yang, T. P., Jaenisch, R., and Panning, B. (2004) Developmentally regulated alterations in Polycomb repressive complex 1 proteins on the inactive X chromosome. *J. Cell Biol.* **167,** 1025–1035
31. Wang, H., Wang, L., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R. S., and Zhang, Y. (2004) Role of histone H2A ubiquitination in Polycomb silencing. *Nature* **431,** 873–878
32. Fang, J., Chen, T., Chadwick, B., Li, E., and Zhang, Y. (2004) Ring1b-mediated H2A ubiquitination associates with inactive X chromosomes and is involved in initiation of X inactivation. *J. Biol. Chem.* **279,** 52812–52815
33. Takahashi, I., Kameoka, Y., and Hashimoto, K. (2002) MacroH2A1.2 binds the nuclear protein Spop. *Biochim. Biophys. Acta* **1591,** 63–68
34. Fernandez-Capetillo, O., Lee, A., Nussenzweig, M., and Nussenzweig, A. (2004) H2AX: the histone guardian of the genome. *DNA Repair* (*Amst.*) **3,** 959–967
35. McKittrick, E., Gafken, P. R., Ahmad, K., and Henikoff, S. (2004) Histone H3.3 is enriched in covalent modifications associated with active chromatin. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 1525–1530