Expression Patterns and Post-translational Modifications Associated with Mammalian Histone H3 Variants

Covalent histone modifications and the incorporation of histone variants bring about changes in chromatin structure that in turn alter gene expression. Interest in non-allelic histone variants has been renewed, in part because of recent work on H3 (and other) histone variants. However, only in mammals do three non-centromeric H3 variants (H3.1, H3.2, and H3.3) exist. Here, we show that mammalian cell lines can be separated into two different groups based on their expression of H3.1, H3.2, and H3.3 at both mRNA and protein levels. Additionally, the ratio of these variants changes slightly during neuronal differentiation of murine ES cells. This difference in H3 variant expression between cell lines could not be explained by changes in growth rate, cell cycle stages, or chromosomal ploidy, but rather suggests other possibilities, such as changes in H3 variant incorporation during differentiation and tissue- or species-specific H3 variant expression. Moreover, quantitative mass spectrometry analysis of human H3.1, H3.2, and H3.3 showed modification differences between these three H3 variants, suggesting that they may have different biological functions. Specifically, H3.3 contains marks associated with transcriptionally active chromatin, whereas H3.2, in contrast, contains mostly silencing modifications that have been associated with facultative heterochromatin. Interestingly, H3.1 is enriched in both active and repressive marks, although the latter marks are different from those observed in H3.2. Although the biological significance as to why mammalian cells differentially employ three highly similar H3 variants remains unclear, our results underscore potential functional differences between these and reinforce the general view that H3.1 and H3.2 in mammalian cells should not be treated as equivalent proteins.

Eukaryotic organisms depend on complex and highly regulated mechanisms to activate or silence genes in response to a variety of stimuli, including environmental changes, cell cycle regulators, and developmental cues. An increasing body of evidence suggests that epigenetic mechanisms involving chromatin remodeling alter the accessibility of proteins, such as transcription factors, to the DNA template.

The fundamental repeating unit of chromatin is the nucleosome core particle, which consists of DNA in close association with an octameric unit of core histones (H2A, H2B, H3, and H4). However, in some instances, specialized histone variants are found in place of the canonical histones, enabling the encoding of epigenetic information through defined or “specialized” nucleosome arrays (reviewed in Ref. 1).

Histones are subject to a diverse array of covalent modifications that occur mostly at the N- and C-terminal tails. The histone “code” hypothesis (2, 3) has been put forward to explain the seemingly complex nature of the reported patterns of histone modifications. Formally, this hypothesis states that a specific histone modification, or combination of modifications, can affect distinct downstream cellular events by altering the structure of chromatin and/or generating a binding platform for effector proteins, which specifically recognize the modification(s) and initiate events that lead to gene transcription or silencing. Expanding the scope of this code, a large number of variant histones has been identified, including some that are unique to vertebrates and some that are highly conserved among all eukaryotes (reviewed in Ref. 4). It has been shown that replacement of the replication-dependent (RD) histone H3 (formally H3.2, see supplemental Fig. 1, A) with its replication-independent (RI) variant H3.3 in Drosophila cells occurs at transcriptionally active loci (5, 6). Furthermore, characterization of Drosophila and Arabidopsis histones by mass spectrometry (MS) revealed enrichment of modifications associated with transcriptional activity, such as methylation of lysine 4 (Lys4) and Lys79 and acetylation of Lys14, Lys18, and Lys23, in H3.3 compared with H3.2 (7, 8). These results suggest that, at least in plant and Drosophila cells, H3.2 and its variant H3.3 play different roles in remodeling chromatin, in part by altering covalent histone modification patterns associated with transcriptional silencing and activation.

Unlike Drosophila, which contains only two different histone H3 molecules, mammalian cells contain three non-centromeric H3 variants: H3.1, H3.2, and H3.3, which differ only in a few amino acids (see supplemental Fig. 1, A). The function of these three mammalian H3s, especially H3.1 and H3.2, is poorly understood. In this report, we investigate the expression patterns and post-translational modifications associated with these three mammalian H3 variants. Analyses of multiple mammalian cell lines revealed that they can be divided into two groups based upon the relative amounts of the individual H3 variants in chromatin. Although the functional significance of this grouping remains unclear, cellular differentiation appears to alter these ratios in at least one ES cell line in a modest, but reproducible, fashion. We also show that these cell line-specific differences in H3 variant expression do...
not originate from changes in growth rate, cell cycle stages, or chromosomal ploidy. Possible mechanisms are discussed. Additionally, the different human H3 variants were subjected to quantitative tandem MS analyses. As expected from studies in Drosophila (7), transcriptionally active marks are associated with H3.3; those often associated with gene silencing, e.g. Lys4 di- and trimethylation, are found on H3.2. Surprisingly, H3.1 is enriched both with modifications that are largely associated with gene silencing, e.g. Lys9 dimethylation, as well as those linked to gene activation, e.g. Lys14 acetylation. These data reinforce the general view that alterations in the covalent modification patterns associated with histone variants provide additional regulatory options for epigenomic “indexing” of biological processes, many of which remain unclear. As well, our data lend support to a poorly appreciated notion that H3.1 and H3.2 variants, while highly similar at the level of post-translational modifications. Thus, our studies underscore the potential need for caution when interpreting H3-related studies in mammalian models.

MATERIALS AND METHODS

Cell Lines and Culture—All mammalian cell lines, with the exception of mouse L2 cells were grown in Iscove’s Dulbecco’s modified Eagle’s medium supplemented with 10% fetal calf serum and penicillin/streptomycin at 37 °C and 5% CO2. Cell lines used in this study are described under supplemental material and methods.

Cell Synchronization—HeLa cells were grown to 70% confluency and treated with 3 mM thymidine for 15 h. Thymidine containing medium was removed, cells washed once with fresh thymidine-free medium and grown for 9 h in regular medium. This process was repeated and then, after release from the double thymidine block, cells were harvested from individual plates every 2 h by trypsinization, washed with PBS, and split into three samples to prepare for cell cycle analysis by FACS and isolation of RNA and histones (see below).

Preparation of Histones—Nuclei and histones were isolated as described earlier (9). Cell nuclei were isolated by hypotonic lysis in buffer containing 10 mM Tris-HCl, pH 8.0, 1 mM KCl, 1.5 mM MgCl2, 1 mM dithiothreitol, 0.4 mM phenylmethylsulfonyl fluoride, protease and phosphatase inhibitors. Pelleted nuclei were acid-extracted using 0.4 N sulfuric acid, precipitated with trichloroacetic acid and resuspended in water.

Reverse Phase HPLC (RP-HPLC)—Separation of mammalian core histones by RP-HPLC was done as described (9). Briefly, acid-extracted histones were separated by RP-HPLC on a C8 column (220 by 4.6 mm Aquagaur P300, PerkinElmer Life Sciences) using a linear ascending gradient of 35–60% solvent B (solvent A: 5% acetonitrile, 0.1% trifluoroacetic acid, solvent B: 90% acetonitrile) over 75 min at 1.0 ml/min on a Beckman Coulter System Gold 126 Pump Module and 166/168 Detector. Under these conditions histones H3 split into two peaks. The H3-containing fractions were dried under vacuum and stored at −80 °C. RP-HPLC fractions were resuspended in water, analyzed by SDS-PAGE and control-stained with Coomassie Brilliant Blue. The identified fractions were then subjected to MS analysis.

Two-dimensional Triton-Acid-Urea (TAU) Gels—Total histones were dried under vacuum and resuspended in loading buffer (6 M urea, 0.02% (w/v) pyronin Y, 5% (v/v) acetic acid, 12.5 mg/ml protamine sulfate). Samples were separated on TAU mini-gels (15% PAGE, 6 M urea, 5% acetic acid, 0.37% Triton X-100; 300 V in 5% acetic acid for 1.5 h). Lanes containing the samples were cut out, adjusted in 0.125 M Tris, pH 6.6, and the TAU gel slice was assembled on top of a 15% SDS-PAGE mini-gel. After the run, the gel was stained with Coomassie Blue and destained with 5% methanol, 7.5% acetic acid overnight. The gels were scanned and quantified using Image Gauge software (Science Lab), with the subtraction of background staining.

Growth Rate Analysis—1 × 106 cells (HeLa and HEK293) were grown in 6-well plates, and every 24 h samples were collected and counted, with exclusion of dead cells (staining with Trypan Blue). Half of the cells were discarded to avoid contact inhibition of the cells as they become too confluent, fresh medium was added, and the cells were allowed to grow for another 24 h before the next sample was collected. With this method, we could measure the doubling time of HeLa and HEK293 cells while still maintaining them at normal confluency (i.e. the same confluency at which we grow them for all other experiments). Cell numbers were calculated to present the cell growth over days of both cell lines. HeLa and HEK293 cells were seeded as duplicates, and this experiment was performed twice.

Cell Cycle Analysis by FACS—1 × 106 cells were collected, washed with PBS, and fixed overnight at −20 °C in 70% ethanol, diluted in PBS. The next day, cells were washed with PBS and incubated for 30 min at 37 °C in PBS containing RNase A (10 µg/ml), followed by the addition of propidium iodide (10 µg/ml) and another incubation for 30 min at 37 °C. Stained cells were analyzed on a FACS sort instrument (BD Immunocytometry Systems) with the exclusion of doublets. Analysis of the results was performed with CellQuest software (BD Biosience).

Immunoblots—Acid-extracted histones were separated on 15% SDS-PAGE mini-gels and either stained with Coomassie Brilliant Blue or transferred onto poly(vinylidene difluoride) membranes (Millipore), and stained with Ponceau S (Sigma) to ensure proper protein transfer. After incubation with primary antibody (anti-H3 S28P: 1:1000 (Upstate Biotechnology) or anti-H3: 1:10000; Abcam)) and addition of a horseradish peroxidase-conjugated secondary antibody (Amersham Biosciences), membranes were incubated with ECL-Plus substrate (Amer sham Biosciences), and proteins were detected by exposure to x-ray film (Amersham Biosciences).

Sample Preparation of Histone H3 Variants for MS—Purified histone H3 protein from pooled RP-HPLC fractions were derivatized by treatment with propionyl anhydride reagent (8). The reagent was created using 75 µl of MeOH and 25 µl of propionic anhydride (Aldrich, Milwaukee, WI). Equal volumes of reagent and each H3 variant were mixed and allowed to react at 51 °C for 15 min. Propionylated histone H3s were then digested with trypsin (Promega) at a substrate/enzyme ratio of 20:1 for 5 h at 37 °C after dilution of the sample with 100 mM ammonium bicarbonate buffer solution (pH 8). The reaction was quenched by the addition of concentrated acetic acid and freezing. A second round of propionylation was then performed to propionylate the newly created peptide N termini.

Mass Spectrometry—Propionylated histone digest mixtures were loaded onto capillary precolumns (360 µm outer diameter × 75 µm inner diameter, Polymicro Technologies, Phoenix, AZ) packed with irregular C18 resin (5–20 µm, YMC Inc., Wilmington, NC) and washed with 0.1% acetic acid for 10 min. Precolumns were connected with Teflon tubing to analytical columns (360 µm outer diameter × 50 µm inner diameter, Polymicro Technologies) packed with regular C18 resin (5 µm, YMC Inc.) structured with an integrated electrospray tip as previously described (10). Samples were analyzed by nanoflow HPLC-μ-electrospray ionization on a linear quadrupole ion trap-Fourier Transform Ion Cyclotron Resonance (LTQ-FT-ICR) mass spectrometer (Thermo Electron, San Jose, CA). The gradient used on an Agilent 1100 series HPLC solvent delivery system (Palo Alto, CA) consisted of 0–40% B in 60 min, 45–100% B in 15 min (A, 0.1% acetic acid, B, 70% acetonitrile in
0.1% acetic acid) or other similar gradients. The LTQ-FT mass spectrometer was operated in the data-dependent mode with the 10 most abundant ions being isolated and fragmented in the linear ion trap. All MS/MS spectra were manually interpreted.

Stable Isotope Labeling for Relative Quantitative Analysis—For a differential expression comparison of histone post-translational modifications from the three H3 variants, stable isotope labeling based on conversion of peptide carboxylic groups to their corresponding methyl esters was used (11). First, all samples were dried to dryness by lyophilization. Aliquots of solutions from propionylated histone peptides from H3.1, H3.2, or H3.3 were converted to d<sub>c</sub>-methyl esters by reconstituting the lyophilized sample in 100 μl of 2 M d<sub>c</sub>-methanol/HCl, or converted to d<sub>e</sub>-ethyl esters by reconstituting the lyophilized sample in 100 μl of 2 M d<sub>e</sub>-methanol/HCl. Reaction mixtures were allowed to stand for 1 h at room temperature. Methyl ester solvent was removed from each sample by lyophilization, and the procedures were repeated using a second 100-μl aliquot of methyl ester reagents. Solvent was then removed again by lyophilization, and samples were dissolved in 20 μl of 0.1% acetic acid. Aliquots of each solution were then equally mixed for comparative analysis by MS.

Quantitative PCR—Total RNA isolation was performed using TRIzol Reagent (Invitrogen). Single-stranded cDNA was generated with the Superscript First-Strand Synthesis kit (Invitrogen). Quantitative PCR was performed with SYBR green dye according to the manufacturer’s instructions (Stratagene). HeLa cDNA was used to generate a standard curve from which the amount of cDNA amplified in each sample was determined as indicated. mRNA levels were normalized to H3.2 mRNA expression. All oligos were synthesized by Sigma, and the sequences of the primer pairs for quantitative PCR used in this study are listed under supplemental material and methods.

RESULTS

Only mammalian cells contain three different non-centromeric H3 variants: RD H3.1 and H3.2 variants versus a single RI H3.3. Other organisms contain only one type of H3, H3.3 (Saccharomyces cerevisiae), or two, H3.2 and H3.3 (e.g. Arabidopsis, Xenopus, and Drosophila) (supplementary Fig. 1A, top). The three mammalian H3 variants are almost identical in their amino acids sequence (see supplemental Fig. 1, A, bottom). H3.1 and H3.2 have only a single amino acid difference (amino acid 96, cysteine/serine, respectively; gray box), whereas H3.1 and H3.3 differ in five amino acids (amino acid 31, alanine/serine; amino acid 87, serine/alanine; amino acid 89, valine/isoleucine; amino acid 90, methionine/glycine; amino acid 96, cysteine/serine, respectively).

Because there are no available antibodies that distinguish the three mammalian H3 variants, the investigation of endogenous expression of these variants is restricted largely to chromatographic and electrophoretic separations (7, 9, 12). Thus, to test whether the different mammalian H3 variants are present in similar abundance in different cell types, we first turned to a chromatographic method. Acid-extracted total histones from different mammalian cell lines were resolved by RP-HPLC; two distinctive H3 peaks were typically observed (supplemental Fig. 1, A). Based on the H3 RP-HPLC profiles, we were able to separate mammalian cell lines into two groups (A and B) based on their peak height (absorbance) and peak area differences. Interestingly, members of group A show an absorbance that is higher for peak 2 than peak 1; a reverse relationship is evident for group B (i.e., an absorbance that is higher for peak 1 than peak 2 (supplemental Fig. 1, C).

To gain further insight into the protein compositions of the two RP-HPLC peaks, MS was employed (supplemental Fig. 2). Our MS analyses demonstrate that the first fractions of peak 1 contain H3.2, and the later "shoulder" fractions contain H3.3; peak 2, in contrast, contains almost entirely H3.1 variant (supplemental Fig. 2 and supplemental Tables 1–3).
Mammalian Histone H3 Variants

To confirm the observed cell type-specific H3 differences and to additionally identify the expression levels of each H3 variant in the different cell lines examined, we separated acid-extracted histones on two-dimensional TAU gels (Fig. 1A) and visualized the histones by staining with Coomassie Blue. Three of the histone spots also stained with H3 specific antibodies (data not shown). To determine the identity of the variant in each of the three H3 spots, each protein sample was digested in gel and the resulting peptides were then characterized by tandem MS as described above (data not shown and Fig. 1A).

Two-dimensional TAU gels were then employed to examine the distribution of H3.1, H3.2, and H3.3 in six different mammalian cell lines, three from each of the group A and B categories (Fig. 1B). We find that cell lines from group A are enriched in H3.1 and those in group B are enriched in H3.3. Next, we quantified the proportion of protein in each of the different H3 spots using Image Gauge software (Fig. 1C). Interestingly, while the proportion of H3.2 (dark gray bars) did not change dramatically between the two different groups, cell lines from group A were enriched in H3.1 (light gray bars). In contrast, cell lines from group B contained less or equal proportions of H3.1 compared with H3.3 protein (black bars). Furthermore, as seen before by RP-HPLC analysis (supplemental Fig. 1), several modest changes in H3 variant composition were observed in murine ES cells (LF2) that were treated with retinoic acid (RA) to induce neuronal differentiation (Fig. 1C, see LF2 columns). First, an immediate increase in H3.3 and a corresponding drop in H3.1 occurred during the first 6 days of treatment, whereas H3.2 levels remained largely the same. However, by day 10 post-RA treatment, when the majority of the ES cells have taken on a neuronal phenotype, the levels of H3.2 increased marginally and the levels of H3.3 dropped slightly, whereas the levels of H3.1 remained about the same. These results confirm and extend the observations made by RP-HPLC (supplemental Fig. 1, C).

Because all cell lines from group A are derived from cancer cells, we wondered whether the high level of H3.1 expression arises from differences in chromosomal ploidy. Therefore, we used RP-HPLC to separate histones from mouse embryonic fibroblast cells where the chromosomal status was either diploid (P-CUT MEF) or tetraploid (10T1/2). Results of this experiment are shown in Fig. 1D. Both of these cell lines had equal peak area ratios in RP-HPLC analysis and were assigned to group B because the observed ratio of peak areas (peak 1/peak 2) was 3. We then separated the H3 variants by two-dimensional TAU gels and found that the levels of H3.1, H3.2, and H3.3 were very similar between diploid (P-CUT MEF) and tetraploid (10T1/2) cell lines (Fig. 1E). H3.3 was the most highly expressed variant in both cell lines, followed by H3.1 and then H3.2. These results suggest that ploidy is not responsible for the different expression levels observed for H3 variant proteins. However, because differences in the peak 1/peak 2 ratio were observed in human (group A) versus mouse (group B) species (Fig. 1 and supplemental Fig. 1), we cannot rule out the formal possibility that variable copy numbers of the H3.1 and H3.2 genes might contribute, at least in part, to some of the differences in H3 variant expression profiles. Another explanation for the above observations is that cells from embryonic origin contain high levels of H3.3, and cells derived from adult tissue have more H3.1.

Because H3.1 and H3.2 are expressed in S-phase whereas H3.3 is expressed and incorporated into chromatin independent of the cell cycle (5), we next wondered whether the observed differences in H3 variant expression arise from differences in growth rates and/or time spent in S-phase. Therefore, we tested a representative cell line from group A and B (HeLa and HEK293, respectively) in growth assays and cell cycle analyses. While somewhat variable, these cells showed a similar growth curve (Fig. 2A), and similar numbers of cells in S-phase by FACS analysis (Fig. 2B). It is also important to note that the growth rates of cell lines within a single group was very different, e.g. within group A, HT-29 grew extremely slowly, whereas Raji cells grew extremely fast (data not shown). Therefore, we conclude that the observed differences in the proportions of the three H3 variant proteins between groups A and B are not likely explained by the RD expression of H3.1 and H3.2 alone.

We wondered whether differences in the proportion of H3.1, H3.2, and H3.3 between group A and B cells originated from differences in steady state levels of mRNA expression. To address this possibility, we performed quantitative analyses of mRNA expression levels of five human cell lines used in this study. We could not include other cell lines from group B in this study, because these are of mouse origin and differ...
Mammalian Histone H3 Variants

in their nucleotide sequence from human H3 variant genes. Fig. 2, C and D show the mRNA expression levels of one H3.2, nine different H3.1, and both H3.3A and H3.3B genes normalized to 18 S rRNA expression. Because we do not know if the 18 S rRNA expression level is the same in all human cell lines examined, we also normalized our data to H3.2 mRNA expression because the proportion of H3.2 protein did not change as drastically between groups A and B compared with H3.1 and H3.3 protein (supplemental Fig. 3). Because it is still possible that different cell lines express different amounts of 18 S rRNA, these results should be viewed with caution. Nonetheless, we observed a similar pattern in H3 variant gene expression when normalized to 18S rRNA expression (Fig. 2, C and D) as when normalized to H3.2 gene expression (supplemental Fig. 3). Interestingly, H3.1 genes of the five different human cell lines were expressed at relative low level and did not exhibit dramatic differences in their expression, with the exception of CEM cells, which seem to express H3.1C. On the other hand, HEK293 cells, which belong to group B, showed a reproducible increase in the expression level of H3.3A (almost 2-fold compared with other human cell lines from group A). The lack of significant differences in growth rates from HEK293 and HeLa cells together with the mRNA expression data suggest that the differences in H3.1, H3.2, and H3.3 protein expression that we observed by both RP-HPLC and two-dimensional TAU gel analyses might originate at the transcriptional level and are independent of growth rates.

Because we observed slight differences in the proportions of cells in G1 or G2/M between HeLa and HEK293 cells, we wondered if cell cycle phases could account for the observed differences in H3.1, H3.2, and H3.3 proportions between HeLa and HEK293 cell lines. We therefore performed a detailed analysis of H3 variant expression on both mRNA and protein levels during G1, S, and G2/M phases in HeLa cells. The results from one of two independently conducted, highly reproducible experiments are shown in Fig. 3. HeLa cells were synchronized in G1 by a double thymidine block and released from this block to continue through different cell cycle phases. Every 2 h, cells were harvested and samples prepared for cell cycle analysis by FACS, mRNA isolation, and cDNA generation or acid-extraction of histones. Fig. 3A shows the cell cycle profile of these cells analyzed by FACS, and quantitative analysis of the amount of cells in each cell phase is depicted in Fig. 3B. The majority of asynchronously growing cells was found to be in G1 (~75%), but also cells in S and G2/M phase were observed. Treatment of HeLa cells with Nocodazole led to an arrest in G2/M (~50%). ~75% of cells were found to be in G1, 2 h after the release from the double thymidine block. 6–8 h after the release, the amount of cells in G1 dropped and more cells in S-phase were found (~55%). 10–12 h after the thymidine release cells were found to move into G2/M (~20 to 40%). These results show that we were able to enrich for cells in specific cell cycle phases.

We also isolated histones by acid-extraction from cells at different time points and tested for mitosis-specific histone modifications by immunobLOTS as an indication of a successful enrichment of cells in G2/M and synchronization by thymidine block (Fig. 3C). Unfortunately, we could not test for the enrichment of cells in G1 or S-phase, because histone modifications specific for these cell cycle stages have not been identified. We used an antibody against the well characterized mitosis-specific H3 Ser28 phosphorylation mark and found that histones from cells arrested with nocodazole stained strongly for this mark and also that histones from the 10 and 12 h time points were positive for H3 Ser28 phosphorylation (Fig. 3C, top). The blot was stripped and stained with an antibody against the C-terminal tail of H3 as a loading control (Fig. 3C, bottom). These results confirm the data we obtained by FACS analysis and show that we successfully enriched for cells in different cell cycle phases.

Next we asked if the expression of H3 variant genes changes during cell cycle phases. To test for these possibilities, we isolated RNA from these cells, generated cDNA and tested for H3.3A, H3.3B (Fig. 3D, left), and H3.2, H3.1H, and H3.1L (Fig. 3D, right) gene expression by quantitative PCR. Surprisingly, we found that both the RD H3 variant genes encoding H3.1 and H3.2 as well as the RI H3.3 genes (A and B) increased in their expression 8 h after the release of thymidine. These data suggest that during S-phase the expression of all H3 variant genes increases.

Next, we wondered if the proportions of H3 variant proteins change in different cell cycle phases, particularly in S or G2/M. Therefore, we isolated histones from asynchronously growing cells, nocodazole-arrested cells, and cells harvested at two (G1) and eight (S) hours after release from the thymidine block and then separated and visualized them by two-dimensional TAU gels with Coomassie Blue (Fig. 3E). As discussed above, H3 variant protein levels were quantified, and the results of two independent experiments are shown in Fig. 3F. We did not observe a significant change in the proportions of H3 variants that would explain the observed differences between group A and B cells. These data suggest that the proportions of H3 variants remain constant regardless of cell cycle phases, and that the observed H3 variant proportion differences between group A and B cell lines are cell intrinsic phenomena.

The above data suggest that different cell lines contain different steady-state levels of H3.1, H3.2, and H3.3 proteins, although the functional significance for these differences remains unclear. To determine whether these variants might have different biological functions revealed by distinct post-translational modification “signatures,” tandem MS was employed to identify covalent modifications present on each of the three variants. The use of MS for both the qualitative and quantitative analysis of post-translational modifications also circumvents problems associated with the use of site-specific antibodies such as specificity, cross-reactivity and epitope occlusion through interference by closely neighboring modifications (13–15).

Treatment of histone H3 with propionic anhydride reagent converts amino groups on the N terminus and internal lysine residues (endogenously unmodified and mono-methylated residues only) to propionyl amides. The consequence of this procedure is removal of charge from lysine residues and increased hydrophobicity of histone peptides, thus facilitating their analysis by MS. Additionally, propionylation of histone proteins blocks trypsin from cleaving residues on the C-terminal side of lysine (di- and trimethylated lysine residues are not cleaved by trypsin as well). Therefore, upon digesting propionylated histones with trypsin, cleavage only occurs C-terminal to arginine, generating a fairly uniform set of peptides from highly modified H3 protein and allowing for a more straightforward monitoring of post-translational modifications (16).

Histones from HEK293 cells were isolated from nuclei by acid extraction, suggesting that a majority of the histones purified in this study originated from nuclear (presumably chromatin-incorporated) histones. Individual histone H3 variants from HEK293 cells were derivatized with propionic anhydride, digested with trypsin (cleavage C-terminal to Arg residues), and the N termini of the newly formed peptides were also derivatized with propionic anhydride. The resulting mixture was then analyzed by a combination of LC-MS/MS on a linear ion trap/Fourier transform mass spectrometer (11). Stable isotopic labeling was employed to estimate the relative abundances of the post-translational modifications on each variant. For example, to compare modifications on H3.3 and H3.1, peptides from the former were converted to d40-methyl esters and those from the latter were converted to d34-methyl...
FIGURE 3. Analysis of histone H3 variant expression in different cell cycle phases. A, HeLa cells were synchronized in G1 phase by a double thymidine block, released and analyzed every 2 h by FACS. B, quantification of FACS analysis results from the experiment shown in A. C, immunoblots with histones isolated from HeLa cells, described in A, using antibodies against H3 S28 phosphorylation (H3 S28P, top) and C-terminal tail of H3 (H3, bottom) as loading control. D, quantification of H3.3A and H3.3B (left) and H3.2, H3.1H, and H3.1L (right)
FIGURE 4. Comparative analysis of post-translational modifications on histones H3.1 and H3.3.

A. mass spectrum recorded on a mixture of peptide d$_3$ and d$_4$-methyl esters derived from histone H3.3 and H3.1, respectively. Zoom inset shows a pair of [M+2H]$^{2+}$ ions at m/z 738.4093 and 744.4470 that correspond to a post-translationally modified peptide containing four carboxylic acid groups. The signal for the d$_3$-labeled peptide (H3.3) is enriched by a factor of 3 over the d$_4$-labeled peptide. B. MS/MS spectrum recorded on the precursor ion at m/z 738.4093. Fragment ions of type b and y define the peptide sequence as EIAQDFKMe$_2$TDLR (residues 73–83 of both H3.3 and H3.1).

Mammalian Histone H3 Variants

esters. Equal amounts of the two samples were then mixed and analyzed on the above mass spectrometer (11). As a result of the above derivatization, ions corresponding to peptides from the two variants that contain the same post-translational modification appear in the mass spectrum as doublets. These doublet peaks are separated by multiples of 3 mass units per carboxylic acid group (C terminus plus Asp and Glu residues) for singly charged ions and 1.5 mass units for doubly charged ions. As an example, a comparative analysis of peptides derived from histones H3.3 and H3.1 is shown in Fig. 4. The figure inset shows a magnification of the mass range from m/z 737–747. Signals at m/z 738.4039 and 744.4039 (6-mass unit separation) correspond to [M+2H]$^{2+}$ ions for the same isotopically labeled peptide (containing 4 carboxylic acid groups) from H3.3 and H3.1, respectively. The MS/MS spectrum recorded on the ion at m/z 738.4093 is shown in Fig. 4B and defines the sequence of the peptide to be EIAQDFKMe$_2$TDLR (residues 73–83 of both H3.3 and H3.1). This peptide is chemically modified by the addition of a propionyl amide group on the N terminus and four methyl ester groups on the carboxylic acid groups. A comparison of the areas under the signals for the pairs of [M+2H]$^{2+}$ ions indicate that the modified peptide is about 4-fold more abundant in H3.3 than in H3.1. Dimethylated Lys$^9$ has been observed on hyperacetylated histone H3 (17) and is known to be associated with transcriptional activation (18).

Table 1 provides a compilation of pre-translational modifications detected and enriched on the H3 variants isolated from two samples independently purified from HEK293 cells. With the exception of acetylation of Lys$^9$ on H3.1, all modifications were detected on each of these three variants. Marks found to be enriched by a factor of at least 2.7 (+++ in Table 1) on H3.3 in both of the above samples include: acetylation of Lys$^9$, Lys$^{14}$, Lys$^{27}$, and Lys$^{18}$ together with Lys$^{23}$, mono- and dimethylation of Lys$^{27}$, and dimethylation of Lys$^{23}$. These modifications have been described in Drosophila and partly in Arabidopsis (7, 8), and are consistent with the general view that the H3.3 variant is involved in the establishment of “active” chromatin. In contrast, marks greatly enriched (+++ in Table 1) in both samples of H3.2 were Lys$^{27}$ di- and trimethylation. Both states of this methylation mark are also found on H3.2 in Arabidopsis (8) and are often implicated in “silent” chromatin (19, 20), specifically in the formation and maintenance of facultative heterochromatin. Thus, H3.3 and H3.2 appear to carry covalent modification “signatures” that largely denote “active” versus “inactive” chromatin, respectively. In contrast, dimethylation of Lys$^9$, often considered an “off” mark (21), and acetylation of Lys$^{14}$, often considered an “on” mark (22), are both significantly enriched (+++ in Table 1) on H3.1 in both independent experiments. These marks were determined to be enriched on H3.1 on separate peptides because quantitative analysis of peptides containing both marks simultaneously revealed that Lys$^9$ dimethylation together with Lys$^{14}$ acetylation is found on only 9% of all H3.1, 6% of all H3.2, and 9% of all H3.3 peptides. Thus, a clear cut difference between “on” versus “off” covalent modification signatures is less clear with the H3.1 variant. Several pre-translational modifications were only found enriched in one of the two samples (+++ in Table 1). Although of variable nature, these, too, follow the same observed trend; enrichment of active marks on H3.3, silent marks on H3.2 and a combination of both on H3.1. Also enriched on H3.1 is a previously unidentified modification, monomethylation of Lys$^{23}$.

Relative enrichments of the observed modifications are displayed in pair-wise fashion in Fig. 5A as follows: (I-H3.1/H3.2) Di- and trimethy-
luation of Lys\textsuperscript{27} and monomethylation of Lys\textsuperscript{36} are 2–5-fold enriched on H3.2 compared with H3.1, whereas H3.1 contains 3–4-fold more acetylation and monomethylation of Lys\textsuperscript{9} and Lys\textsuperscript{14}, respectively, than H3.2; (II-H3.1/H3.3) A 3–5-fold increase in acetylation of Lys\textsuperscript{9}, Lys\textsuperscript{14}, and Lys\textsuperscript{18} is observed on H3.3 compared with H3.1, whereas H3.1 is found to contain more dimethylation of Lys\textsuperscript{9} (3-fold) and monomethylation (3-fold) of Lys\textsuperscript{64} than H3.3. H3.3 also contains more mono- and dimethylation of Lys\textsuperscript{36} and dimethylation of Lys\textsuperscript{79} (4–5-fold) than H3.1; (III-H3.2/H3.3) A 2–7-fold enrichment of acetylation on H3.3 is observed at Lys\textsuperscript{9}, Lys\textsuperscript{14}, Lys\textsuperscript{9} + Lys\textsuperscript{14}, Lys\textsuperscript{18} + Lys\textsuperscript{21}, and Lys\textsuperscript{27}. Dimethylation of Lys\textsuperscript{36} and Lys\textsuperscript{79} is also enriched on H3.3 compared with H3.2. In keeping with its “euchromatic” silencing function, histone H3.2 contains more Lys\textsuperscript{27} trimethylation (3-fold) than H3.3. Fig. 5B schematically summarizes the marks that are enriched at least 2-fold in each individual experiment and depicts the enrichment of active marks on H3.3, silent marks on H3.2, and both active and silent on H3.1.

These collective results show that the three mammalian H3 variants are enriched in different post-translational modifications and in different patterns of these modifications, suggesting that they have different biological functions. From the relative abundance data, we conclude that H3.3 is involved in gene activation and that H3.2 is used primarily in euchromatic gene silencing. The function of H3.1 is yet to be defined since our analyses show that it is enriched in marks that are largely associated with both gene silencing and gene activation. The combination and non-overlapping nature of these modification patterns clearly distinguishes H3.1 from H3.2 and H3.3 (see Fig. 5B). These data also underscore the need to not combine H3.1 and H3.2 together as H3 in mammalian models.

**DISCUSSION**

Previous studies showed that epigenetic indexing mechanisms help determine whether a gene is maintained in a silent or active state. Histone modifications clearly play a role in this process, as does the incorporation of specialized histone variants into nucleosomes, the latter being particularly important for chromatin remodeling. One histone variant in particular, H3.3, has been associated with transcriptional activation. In Drosophila and mammalian cells, H3.3 is closely associated with transcriptionally active foci (5, 6, 23), and found to be enriched in active marks (7). Additionally, recent reports suggest that transcriptional activation triggers the deposition and removal of H3.3 from chromatin in Drosophila cells (24).

Because several of the above studies used Drosophila cells as their experimental model, less is known about the potentially different functions of mammalian H3 variants. Moreover, mammalian cells are unique in that they contain, in addition to RI H3.3, also H3.1 and H3.2, both of which assemble by RD mechanisms. This special feature of mammalian cells has been largely ignored, in part because H3.1 and H3.2 differ only in one amino acid in the histone core region (see supplemental Fig. 1, A). Nevertheless, the post-translational modification signatures differ significantly between these highly similar proteins, suggesting that H3.1 and H3.2 are likely to differ in function. Our data hint at the intriguing possibility that the unique mammalian H3 variant, H3.1, may play a specialized role in chromatin biology that may corre-

---

**TABLE 1**

| Residue/peptide | Modification | Function | H3.1 | H3.2 | H3.3 |
|----------------|-------------|----------|------|------|------|
| TK\textsuperscript{QTAR} | Monomethyl | A\textsuperscript{a} | + | + | + |
| K\textsuperscript{*STGK*GKAPR} | Monomethyl | S\textsuperscript{b} | + | + | + |
| Dimethyl | S | + | + | + |
| Trimethyl | S | + | + | + |
| Acetyl | A | + | + | + |
| K\textsuperscript{*STGK*GK*APR} | Acetyl | A | + | + | + |
| Dimethyl + acetyl | NK | + | + | + |
| Trimethyl + acetyl | NK | + | + | + |
| Acetyl + acetyl | NK | + | + | + |
| K\textsuperscript{*QLATKAAR} | Monomethyl | NK | + | + | + |
| Acetyl | A | + | + | + |
| K\textsuperscript{*QLATK*K3}AAR | Acetyl | A | + | + | + |
| K\textsuperscript{*QLATK*K3}AAR | Acetyl + acetyl | NK | + | + | + |
| K\textsuperscript{*SAPATG*GGKKPHR} | Monomethyl | S | + | + | + |
| Dimethyl | S | + | + | + |
| Trimethyl | S | + | + | + |
| Acetyl | A | + | + | + |
| K\textsuperscript{*SAPATGGVK*KPHR} | Monomethyl + monomethyl | NK | + | + | + |
| Dimethyl | A | + | + | + |
| K\textsuperscript{*SAPATGGVK*KPHR} | Monomethyl + dimethyl | NK | + | + | + |
| Monomethyl + trimethyl | NK | + | + | + |
| Dimethyl + monomethyl | NK | + | + | + |
| Dimethyl + dimethyl | NK | + | + | + |
| Trimethyl + monomethyl | NK | + | + | + |
| Trimpethyl + dimethyl | NK | + | + | + |
| K\textsuperscript{*LPFQR} | Monomethyl | NK | + | + | + |
| Monomethyl | NK | + | + | + |
| Dimethyl | A | + | + | + |
| VTIMPK\textsuperscript{DIQLAR} | Monomethyl | NK | + | + | + |

\textsuperscript{a} A, transcriptional activation.
\textsuperscript{b} S, transcriptional silencing.
\textsuperscript{c} NK, unknown function.
\textsuperscript{d} ND, not detected.
late with differentiation or cell origin determination. However, this possibility remains to be shown, in part because of limited reagents that distinguish H3.1 from H3.2.

Post-translational modifications of histones have been shown to be important in establishing and maintaining chromatin remodeling events leading to gene activation or silencing. Different modifications have different biological read-outs, and the marks on histones can therefore point toward a potential function. Using a combination of isotopic labeling and quantitative MS, we show that human H3.1, H3.2, and H3.3 variants are enriched in different post-translational modifications, suggesting separate biological functions for each of the variants. As has been shown previously in Drosophila and Arabidopsis, H3.3 is enriched in modifications associated with transcriptional activation (7, 8). These observations are both interesting and important, because they suggest that the function of H3.3 has been evolutionarily conserved.

These studies also serve as a key internal control for our MS/MS analysis of human H3.1 and H3.2, where no data are available to date. H3.2 is found in all eukaryotes except budding yeast and has been implicated in gene silencing. Our data support these observations, as we find that H3.2 is enriched in Lys27 di- and trimethylation. These generally repressive marks have been associated with gene silencing and the formation of facultative heterochromatin (reviewed in Ref. 25). Unexpectedly, we find that H3.1 has evolved to contain a distinct covalent modification spectrum as compared with H3.2 and H3.3. H3.1 is enriched in

FIGURE 5. Relative enrichment of post-translational modifications on human histone H3 variants determined by stable isotope labeling and quantitative MS. A, histogram bars show enrichment between H3.1 and H3.2 (panel I), H3.1 and H3.3 (panel II), and H3.2 and H3.3 (panel III). MS analyses were done on two independently prepared H3 samples from HEK293 cells. B, schematic representation and summary of results found by MS (see A). Modifications that are found reproducibly enriched on the H3 variants are depicted as follows: methylation as one (monomethyl), two (dimethyl), three (trimethyl) red circles, and acetylation as blue circles. Modifications that are enriched on the same peptide are connected by a line.
Mammalian Histone H3 Variants

Lys3 dimethylation, a modification associated with gene silencing (reviewed in Ref. 26) as well as Lys14 acetylation, a modification we find on H3.3, and a novel mark, Lys64 monomethylation. These data show that the three human H3 variants differ in their post-translational modifications and therefore suggest that each variant is likely to perform a different biological function.

We show that mammalian cell lines (human and mouse origin) can be divided into two groups (A and B) that differ in their expression levels of H3 variants. Our data suggest that neither the ploidy status of the cell nor its growth rate is an explanation for the variant usage detected in our studies. As expected, we found that H3.3A and B gene expression is high also outside of S-phase, whereas H3.1 and H3.2 gene expression is low, in accordance with the general notion that H3.3 is a RI and H3.1 and H3.2 are RD-expressed genes. One can envision that H3.3 is expressed at all cell cycle stages to allow its incorporation into chromatin, and the subsequent activation of previously silenced genes upon appropriate outside stimuli. Interestingly, however, we also found that the RI H3.3 variant genes, which are described by many groups to be deposited into chromatin in a replication-independent manner (5) are also up-regulated during S-phase, similar to H3.1 and H3.2 genes. These results suggest that during S-phase, when the DNA content is doubled, the expression of all H3 variants is up-regulated to provide the cell with the materials to heritably maintain its nucleosomal composition in both daughter cells.

Having ruled out other possibilities, our results suggest that H3 variant composition correlates either with the tissue-, species-, or most interestingly, developmental origin of the cell in each group. Cells of embryonic origin contain more H3.3 compared with H3.1 and H3.2, whereas cells derived from adult tissue have more H3.1 protein than H3.2 and H3.3. One intriguing possibility would be that during the differentiation of certain cell types the ratio of these variants changes. In the case of neuronal cells the proportion of H3.3 increases during differentiation, as has been described previously (27), and is similar to what we observe in RA-treated embryonic stem cells (see Fig. 1C, bracket). However, other cell types might behave differently from neuronal cells. A previous study reported that, during terminal differentiation of murine erythrocytoma cells, incorporation of H2A variants, but not H3.3, into chromatin rapidly increases although these cells stopped dividing (28). In support, Urban and Zweidler (29) found changes in the proportion of H3.2 and H3.3 during chicken development. Dramatic increases of H3.3 were found in liver and kidney, but not other tissues, where the amounts of H3.2 protein remained relatively high (29). However, we cannot rule out the possibility of other more trivial explanations. Variant gene copy numbers between human and mouse, for example, could account for the observed differences in H3 variant proportions. Despite these uncertainties, our data underscore the importance to distinguish the three H3 variants from each other in future studies.

Our interesting observation that during RA-treatment of murine ES cells the levels of H3 variants slightly change (modest increase of H3.3 during the first 6 days of treatment, and a slight drop of H3.3 and increase in H3.2 levels at day 10) parallels the report from a recent study by Chambery and Bickmore (30). This report describes the nuclear reorganization of the HoxB locus upon RA-treatment of murine ES cells (OS25). Interestingly, they suggest that higher-order chromatin structure regulates the expression of the HoxB gene cluster. Upon induction with RA, the HoxB1 locus loops out away from the chromosomal territories with kinetics that parallel those of its transcription, so that when HoxB1 expression is silenced after day 4, the frequency and extend of its looping also decreases. The later expressed gene locus of HoxB9 does not loop out until day 10. Chambery and Bickmore (30) also show that chromatin compaction and nuclear organization represent a level of chromatin structure that is not simply a reflection of underlying histone acetylation. The kinetics of HoxB locus reorganization parallels our observed modest changes of H3.2 and H3.3 levels over time in our RA-treated murine ES cells. One exciting explanation would be that H3 variants are involved in the nuclear organization of chromatin, with H3.1 associated with irreversibly silenced gene loci, H3.2 with reversibly silenced and H3.3 with active gene loci. Future studies will have to determine if H3 variants might play a role in the organization of the nuclear architecture.

This is, to our knowledge, the first comprehensive study of the three mammalian H3 variants, H3.1, H3.2, and H3.3, addressing both their level of expression and their post-translational modifications. Our data point to the existence of a regulatory mechanism in mammalian cells that is more complex than that in lower eukaryotes. We suggest that the three H3 variants might have different biological functions that are based on differences in covalent modification patterns. Our findings also suggest that a prevailing view, namely, that RI-coupled assembly leads to the incorporation of H3.3 into non-replicating chromatin thereby replacing H3.1 and H3.2 over time, may not account for all biological phenomena in which these H3 variants participate.

Acknowledgments—We thank members of the Allis and Hunt laboratories for insightful discussions. We especially thank Christian Janzen for critical review of the manuscript.

REFERENCES
1. Hake, S. B., Xiao, A., and Allis, C. D. (2004) Br. J. Cancer 90, 761–769
2. Strahl, B. D., and Allis, C. D. (2000) Nature 403, 45–45
3. Fischle, W., Wang, Y., and Allis, C. D. (2003) Curr. Opin. Cell. Biol. 15, 172–183
4. Malik, H. S., and Henikoff, S. (2003) Nat. Struct. Biol. 10, 882–891
5. Ahmad, K., and Henikoff, S. (2002) Mol. Cell. 9, 1191–1200
6. Smith, M. M. (2002) J. Biol. Chem. 277, 3847–3857
7. McKittrick, E., Gaflken, P. R., Ahmad, K., and Henikoff, S. (2004) Proc. Natl. Acad. Sci. U. S. A. 101, 1525–1530
8. Johnson, L., Mollah, S., Garcia, B. A., Muratore, T. L., Shanabawitz, J., Hunt, D. F., and Jacobsen, S. E. (2004) Nucleic Acids Res. 32, 6511–6518
9. Hake, S. B., Garcia, B. A., Kauer, M., Baker, S. P., Shanabawitz, J., Hunt, D. F., and Allis, C. D. (2005) Proc. Natl. Acad. Sci. U. S. A. 102, 6344–6349
10. Martin, S. E., Shanabawitz, J., Hunt, D. F., and Marto, J. A. (2000) Anal Chem. 72, 4266–4274
11. Syka, J. E., Marto, J. A., Bai, D. L., Horning, S., Senko, M. W., Schwartz, J. C., Ueberbierheide, B., Garcia, B., Busby, S., Muratore, T., Shanabawitz, J., and Hunt, D. F. (2004) J. Proteome Res. 3, 621–626
12. Waterborg, J. H. (1990) J. Biol. Chem. 265, 17157–17161
13. Zhang, L., Eagen, E. E., Parthun, M. R., and Freitas, M. A. (2003) Chromosoma 112, 77–86
14. Pasenza, J., Kim, Y. B., Taylor, G. K., and Kelleher, N. L. (2004) J. Am. Chem. Soc. 126, 3386–3387
15. Garcia, B. A., M. C., Hake, S. B., Prak, C., Turner, F. B., Busby, S. A., Shanabawitz, J., Moran, R. G., Allis, C. D., and Hunt, D. F. (2005) Biochemistry 44, 13202–13213
16. Clayton, A. L., Rose, S., Barratt, M. J., and Mahadevan, L. C. (2000) EMBO J. 19, 3714–3726
17. Zhang, K., Siino, J. S., Jones, P. R., Yau, P. M., and Bradbury, E. M. (2004) Proteomics. 4, 3765–3775
18. van Leeuwen, F., Gaflken, P. R., and Gotschling, D. E. (2002) Cell 109, 745–756
19. Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R. S., and Zhang, Y. (2002) Science 298, 1039–1043
20. Rouguille, C., Chaumeil, J., Sarma, K., Allis, C. D., Reinberg, D., Avner, P., and Heard, E. (2004) Mol. Cell. Biol. 24, 5475–5484
21. Stewart, M. D., Li, C., and Wong, J. (2005) Mol. Cell. Biol. 25, 2525–2538
22. Aglioti, T., Chen, G., and Thanos, D. (2002) Cell 111, 381–392
23. Chow, C. M., Georgiou, A., Sztoriz, H., Maia, E. S. A., Pombo, A., Barahona, I., Dargelo, E., Canzonetta, C., and Dillon, N. (2005) EMBO Rep. 6, 354–360
24. Schwartz, B. E., and Ahmad, K. (2005) Genes Dev. 19, 804–814
25. Chow, J. C., and Brown, C. J. (2003) Cell Mol. Life Sci. 60, 2586–2603
26. Lachner, M., and Junewein, T. (2002) Curr. Opin. Cell. Biol. 14, 286–298
27. Pina, B., and Suau, P. (1987) Dev. Biol. 123, 51–58
28. Grove, C. W., and Zweidler, A. (1994) Biochemistry 23, 4436–4443
29. Urban, M. K., and Zweidler, A. (1983) Dev. Biol. 95, 421–428
30. Chambery, S., and Bickmore, W. A. (2004) Genes Dev. 18, 1119–1130