II. MASS SPECTROMETRY REVEALS AN INDEX OF CELL TYPE-SPECIFIC MODIFICATIONS ON H3 AND H4*

Received for publication, September 19, 2008, and in revised form, October 24, 2008. Published, JBC Papers in Press, October 28, 2008, DOI 10.1074/jbc.M807274200

Joshua J. Nicklay†, David Shechter‡, Raghu K. Chitta‡, Benjamin A. Garcia§, Jeffrey Shabanowitz‡, C. David Allis‡, and Donald F. Hunt†,‡,§,2

From the †Laboratory of Chromatin Biology, The Rockefeller University, New York, New York 10065 and the Departments of ‡Chemistry and §Pathology, University of Virginia, Charlottesville, Virginia 22901

Epigenetic information is hypothesized to be encoded in histone variants and post-translational modifications. Varied cellular and locus-specific combinations of these epigenetic marks are likely contributors to regulation of chromatin-templated transactions, including transcription, replication, recombination, and repair. Therefore, the relative abundance of histone modifications in a given cell type is a potential index of cell fate and specificity. Here, we utilize mass spectrometry techniques to characterize the relative abundance index of cell type-specific modifications. Notably, we have shown that the modifications on histone H4 in a variety of cell types, resolving 26 distinctly modified H3 peptides. We employed the electron transfer dissociation fragmentation technique in a “middle-down” approach on the H4 N-terminal tail to explore the overlap of post-translational modifications. We observed 66 discrete isoforms of the H4 1–23 fragment in four different cell types. Isolation of the stored, predeposition histone H4 from the frog egg also revealed a more varied pattern of modifications than the previously known diacetylation on Lys6 and Lys12. The developmental transitions of modifications on H3 and H4 were strikingly varied, implying a strong correlation of the histone code with cell type and fate. Our results are consistent with a histone code index for each cell type and uncover potential cross-talk between modifications on a single tail.

Histone proteins compose approximately half of the mass of chromatin, the physiological form of the genome. Two molecules each of H2A, H2B, H3, and H4 are wrapped with DNA into a nucleosome (1); nucleosomes are further compacted in larger fibers and ultimately into chromosomes. Post-translational modifications (PTMs) of histone proteins, including, but not limited to, acetylation, methylation, phosphorylation, ubiquitylation, sumoylation, and ribosylation, primarily on the disordered N-terminal “tails” (2), have been hypothesized to constitute part or all of the histone code (3, 4). These histone modifications are thought to participate in the regulation of the usage of the underlying genetic information. Post-translational modifications of histones have been shown to play many roles in chromatin biology, including cis-acting roles for histone PTMs such as acetylation, in which the neutralization of the positively charged lysine ε-amine alters chromatin compaction, and trans-acting roles for modifications, in which binding partners are recruited to facilitate enzymatic action (5). These varied roles for histone modifications result in downstream chromatin positive and negative regulatory events, including, but not limited to, transcription, replication, and repair.

Historically, these PTMs have been individually identified and characterized by a variety of techniques, including radioactive labeling and amino acid analysis, modification-specific antibodies, “bottom-up” mass spectrometry, and most recently by “middle-down” and “top-down” mass spectrometry. Mass spectrometry approaches have allowed the relative quantitation of individual modifications on histones. Recently, techniques have been developed to identify modifications on a single, large peptide fragment, effectively a “middle-down” approach to MS analysis. A large piece of the intact protein can be analyzed separately for accurate mass and for sequence information/PTM localization. The development of electron capture dissociation (6) and electron transfer dissociation (ETD) (7, 8) fragmentation has allowed middle-down analysis of larger proteolytic fragments of histone tails as a result of the charge state dependence of electron capture (6, 9, 10). Furthermore, relatively labile post-translational modifications, such as phosphorylation, remain intact when a peptide is subjected to electron capture dissociation or ETD fragmentation, making these techniques very useful for novel modification identification. However, we do note that the MS/MS spectra generated by these techniques are very complex, especially when generated by coeluting, isobaric parent ions, and therefore require extensive and time-intensive analysis.

* This work was supported, in whole or in part, by National Institutes of Health Grant National Research Service Award/Kirschstein Fellowship GM075486 (to D. S.), Grant GM37537 (to D. F. H.), and Grant GM53512 (to C. D. A.). This work was also supported by Rockefeller University. 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.

† Both of these authors contributed equally to this work.

‡ To whom correspondence should be addressed: McCormick Rd., Charlottesville, VA, 22901. Tel.: 434-924-3610; E-mail: dhf@virginia.edu.

§ The abbreviations used are: PTM, post-translational modification; MS, mass spectrometry; HPLC, high pressure liquid chromatography; ETD, electron transfer dissociation; PTR, proton transfer charge reduction; FT, Fourier transform; CAD, collisionally activated dissociation.
Mass Spectrometry Analysis of Histones in X. laevis

In particular, the single isof orm of histone H4 has been par-
ticularly amenable to middle-down analysis of its N-terminal
1–23 fragment (11–13). However, other histones have also
been studied using these approaches, including H2A (14), H2B
(9), and H3 (12, 15–17). ETD and electron capture dissociation
have been used to explicitly link H3K4 methylation to H3 acety-
lation on a single N-terminal tail (15, 16), an observation
that strongly supports the combinatorial, or multivalent, extension
to the histone code hypothesis (18).

We have presented in the accompanying article (40) an an-
alysis of histone modifications and histone variants present in
seven distinct cell types from the anuran Xenopus laevis,
including the stored, predeposition histones in the oocytes and
in the eggs, sperm histones, early embryo equivalent pronuclei
histones, A6 (kidney cell lineage) cultured cells, S3 (neurula cell
lineage) cultured cells, and nucleated adult erythrocytes. In
the accompanying article (40), we present evidence that histone
modifications and variants are distinctly enriched in each cell
type, providing a “histone code” index that may serve to char-
acterize each distinct tissue.

Here, we present the mass spectrometric analysis of histone
H3, using conventional MS/MS collisionally activated disso-
ciation (CAD) fragmentation, from Xenopus sperm, early embryo
equivalent pronuclei, erythrocytes, and S3 cells. We also pres-
ent the mass spectrometric analysis of the histone H4 N-termi-
nal tail, in a middle-down approach using ETD fragmentation,
from sperm, stored (predeposition) egg, early embryo equiva-
 lent pronuclei, and S3 cells. These results and the similar results
presented in the accompanying article (40) strongly support the
histone code hypothesis, since they provide an index of modi-
fications for each cell type and are further correlated with the
particular biological phenomena of the given cell.

EXPERIMENTAL PROCEDURES

Xenopus Egg Extract Preparation, Pronuclei Assembly, and
Histone Purification

Xenopus interphase egg extract was prepared as described
(19) and in the accompanying article (40). Pronuclei were
assembled and isolated as described in the accompanying
article (40). Egg histones were purified on heparin-Sepha-
rose, as described, whereas chromatin-bound histones from
pronuclei, erythrocytes, and tissue culture cells were acid-
extracted as described, all in the presence of phosphatase
inhibitor mixture (Roche PhosSTOP), protease inhibitor
mixture (Roche Complete), and 10 mM sodium butyrate (20).
Core histones were chromatographically separated on a
4.6-μm C8 reversed phase column as described (20). Frac-
tions containing each core histone were identified based on
known retention times (20) and verified by Coomassie stain-
ing and Western blotting.

Enzymatic Digestion

H4—C8 reverse phase-HPLC-purified fractions correspond-
ing to histone H4 from Xenopus sperm, egg, early embryo, or
neurula (S3) cultured cells were combined (~20 pmol of total
protein) and incubated with endoproteinase AspN (Roche
Applied Science) (1:20 enzyme/protein) for 7 h at 37 °C and in
NH₂HCO₃ pH 8. The digestion was quenched with glacial ace-
etic acid and stored at −35 °C prior to analysis.

H3—Purified H3 (~25 pmol of total protein) from Xenopus
sperm, early embryo, neurula (S3) cultured cells, and erythro-
cyte was reduced with dithiothreitol and alkylated with iodoac-
etamide (21). Following alkylation, H3 fractions were treated
with propionylation reagent, digested with trypsin (Promega,
Madison, WI) (1:20 enzyme/protein) for 7 h at 37 °C, and
immediately retreated with trypsin, as previously described (22).

Mass Spectrometry

Roughly 10–15 pmol of enzymatically digested mixtures of
either H3 or H4 peptides were loaded onto self-packed micro-
capillary precolumns. Briefly, a fused silica precolumn (360 ×
75 μm, outer diameter × inner diameter; Polymicro Technol-
ologies, Phoenix, AZ) was packed with 8–10 cm of irregular 5–20
μm C₁₈ beads (YMC, Kyoto, Japan) and connected via a Teflon
sleeve (0.060 × 0.012 inch, outer diameter × inner diameter;
Zeus Industrial Products, Orangeburg, SC) to a fused silica
microcapillary analytical column (360 × 50 μm, outer dia-
meter × inner diameter) packed with 6–8 cm of regular 5 μm C₁₈
beads (YMC) and equipped with a laser-pulled (P-2000, Sutter
Instruments, Novato, CA) electrospores emitter tip (2–5 μm in
diameter).

Peptides were gradient-eluted using an Agilent 1100 series
HPLC (Palo Alto, CA) into either a Finnigan LTQ Quadrupole
Ion Trap (Thermo Fisher Scientific, San Jose, CA) modified for
ETD/proton transfer charge reduction (PTR) (6), a Thermo
Electron LTQ-Orbitrap (Thermo Fisher Scientific, Bremen,
Germany), or a Finnigan LTQ-FT (Thermo Fisher Scientific,
Bremen, Germany) utilizing a gradient of 0–60% B in 60 min,
60–100% B in 3 min, and held at 100% B for 4 min (A = 0.1%
acetic acid; B = 70% acetonitrile in 0.1% acetic acid). The LTQ-
Orbitrap and LTQ-FT were operated in a data-dependent mode in
which the MS1 scan was taken from m/z 300–2,000, in the
Orbitrap (r = 30,000 at m/z 400) or ion cyclotron resonance
(r = 100,000 at m/z 400) mass analyzer, followed by 10 CAD
MS/MS scans (normalized collision energy of 35 with a mass
window of 3 atomic mass units) in the linear ion trap mass
analyzer. A dynamic exclusion of 30 s was used with a repeat
count of 2. The ETD/PTR-modified LTQ was operated in a
data-dependent mode in which the MS1 scan was taken from
m/z 300–2,000, followed by six MS/MS scans. For ETD, the
radical anion of fluoranthene (m/z 202) accumulated in the lin-
ear ion trap for 2–5 ms and subsequently reacted with the par-
et ion from the MS1 scan for 90–120 ms. For PTR, the benzo-
ate anion (m/z 121) accumulated in the linear ion trap for 2–3
ms and reacted with the ETD-generated fragment ions for
100–150 ms.

Data Analysis

High resolution H4 data were inspected manually using Qual
Browser (Thermo Electron Corp.) software for all masses corre-
adia c h i n g to H4 1–23 with any combination of 0–5 acetyla-
tions (Δ42.0106 Da), 0–5 methylations (Δ14.0157 Da), and 0–1
phosphorylation (Δ79.9663 Da); H3 analyses were manually
inspected by accurate mass using Qual Browser for all combinations of up to six methylations and two acetylations on each the five tryptic peptides listed in Table 1. In addition, spectra were manually inspected for any additional highly abundant modifications and adducts. Masses that agreed to within 5 ppm of the theoretical mass were analyzed to determine their relative abundance with respect to all other modifications and combinations of modifications. Relative abundance information for each species was determined by taking the area under the curve of the selected ion chromatogram for the most abundant isotope for every charge state present utilizing a ±0.01 Da mass window and by comparing it with the sum of the areas for every unique species of the given peptide. For all species of H4 1–23 found in the high resolution analysis, ETD/PTR data were manually interrogated for sequence validation and for PTM localization; potential modifications observed on the peptides of H3 were localized using CAD data acquired in the linear ion trap of the LTQ-Orbitrap or LTQ-FT.

RESULTS

Histones were isolated from X. laevis eggs (predeposition histones complexed with chaperones), sperm, early embryo equivalent pronuclei, erythrocytes, and S3 cultured cells. The mixed histone population was further purified on a C8 reversed phase HPLC column to resolve each core histone in an individual fraction (Fig. 1). The peak fractions of histone H3 and H4 were then subjected to enzymatic proteolysis and subsequent mass spectrometry as described below.

We propionylated histone H3 both before and after proteolysis and prior to mass spectrometric analysis. Conversion of unmodified and dimethylated lysine residues as well as protein and peptide N termini to their propionylated derivatives (an addition of 56.0262 Da per lysine and/or N termini), prevents trypsin from cleaving at lysines and also results in an increase in the overall hydrophobicity of derivatized peptides, allowing for enhanced retention of smaller H3 peptides. Propionylation also tends to reduce the overall charge of a given peptide, allowing for efficient MS/MS analysis of the typically basic histone proteins via CAD.

Histone H3—Table 1 shows an overview of all of the post-translational modifications observed on H3 from the four cell types analyzed.

H3 species containing Lys\(^9\) and Lys\(^{14}\) from sperm and early embryo show Lys\(^9\) monomethylation at a level roughly equivalent to that of the nonmethylated form. This is in striking contrast to H3 from erythrocyte, in which a majority of Lys\(^9\) exists in the trimethylated state, as well as to S3 H3, which displays a preference for both mono- and dimethylated Lys\(^9\). Interestingly, although all cell types contain low to moderate levels of acetylation, only erythrocyte and S3 cells contain any combination of Lys\(^9\) methylation and Lys\(^{14}\) acetylation. For the mono-acetylated species observed in all cell types, analysis of various CAD MS/MS spectra shows the presence of both acetylated Lys\(^9\) and acetylated Lys\(^{14}\) (see Fig. 2). Although these two species are unique from one another (i.e. the modifications are not simultaneously present on the same peptide), they were not chromatographically resolved under the experimental conditions, and separate relative abundances could not be efficiently obtained. Therefore, the relative abundance noted in Table 1 for monoacetylated H3 9–17 includes the summation of both Lys\(^9\) acetylation and Lys\(^{14}\) acetylation.

Acetylation distributed over Lys\(^{18}\) and Lys\(^{23}\) appears to be a relatively major modification seen on H3 of all cell types, existing on roughly the same order of magnitude as the unmodified form of the 18–26 H3 peptide. Diacetylation over both Lys\(^{18}\) and Lys\(^{23}\) was also detected in three of the four cell types and exists at roughly an order of magnitude below the acetylated species in sperm, early embryo, and S3 cells.

MS/MS Analysis of H3 27–40 reveals the apparent mutually exclusive nature of methylation upon Lys\(^{27}\) and Lys\(^{36}\) in all cell types except for sperm; interestingly, only sperm H3 27–40 displays any combinatorial methylation marks over Lys\(^{27}\) and Lys\(^{36}\). Also of note is the observation that trimethylation was shown to exist on Lys\(^{27}\) in all cell types, yet no Lys\(^{36}\) trimethylation was seen.

MS signals for the H3 peptide containing Lys\(^4\) display a clear preference in each cell type for the unmodified form of the peptide, with the monomethylated form existing at roughly an order of magnitude below the unmodified form.

Histone H4—Following enzymatic digestion with endopeptidase AspN, a 23-residue peptide corresponding to the extreme N terminus of histone H4 was generated, providing complete coverage for the most heavily modified region of H4 and thus allowing for the comparative analysis of all combinations of N-terminal modifications of H4.

H4 peptides were subjected to C18 reversed phase HPLC, which provides separation based upon increasing numbers of acetylations (differing by 42.0106 Da mass units). High resolution MS1 scans on the LTQ-Orbitrap of the various species of H4 1–23 containing, for example, a total of three acetylations, revealed distinct differences of 14.0157/z mass units, indicative of variously methylated forms of triacetylated H4 1–23. The same mass differences were seen for mono-, di-, and tetra-acetylated H4 1–23.

Comparison of the N terminus of H4 from the four cell types under analysis revealed distinct patterns of acetylation, methylation, and phosphorylation for each biological state. Fig. 3 shows MS1 scans from the LTQ-Orbitrap, demonstrating the unique patterns of methylation for triacetylated H4 1–23 for each cell type analyzed. For example, triacetylated sperm H4 1–23 displays abundant dimethylation but very minor mono- and trimethylation. This pattern of methylation appears drastically different from the methylation states of triacetylated egg, early embryo, and S3 H4 1–23. Egg H4 1–23 shows relatively equivalent amounts of the nonmethylated, monomethylated, and dimethylated forms of triacetylated H4 1–23, whereas early embryo displays fairly abundant amounts of nonmethylated, triacetylated H4 1–23 relative to the mono- and dimethylated forms. S3 also exhibits abundant amounts of dimethylation in combination with triacetylation of H4 1–23 but also contains relatively abundant amounts of nonmethylation and monomethylation.

An ETD/PTR tandem mass spectrum of the [M + 6H]\(^{6+}\) species of H4 1–23 containing two internal acetylations and two methylations (at m/z 419.9237) from Xenopus sperm is shown in Fig. 4. Manual sequencing of the fragment ions...
allowed for the full characterization of every potential PTM observed in the LTQ-Orbitrap analysis. Localization of both methylations to Lys\textsuperscript{20} was facilitated by the observation of the z\textsubscript{3} (m/z 371) and z\textsubscript{4} (m/z 527) ions, the latter of which displays a mass shift of 28 Da, indicative of lysine dimethylation. Full acetylation of Lys\textsuperscript{16} is indicated by the shift in mass of 42 Da above that of an unmodified lysine residue, observed on both the c\textsubscript{16} (m/z 1539) ion and the z\textsubscript{8} (m/z 1146) ion. Full N-terminal acetylation is observed from the 42 Da shift of the first c' ion observed (c\textsubscript{3}, m/z 360) and is consistent with the known amino acid sequence of the N terminus of H4. Localization of the third acetyl group was more difficult due to the presence of two isobaric species within the isolation window: H4 1–23, containing N-terminal, Lys\textsuperscript{5}, and Lys\textsuperscript{16} acetylation in combination with Lys\textsuperscript{20} dimethylation, and H4 1–23, containing N-terminal, Lys\textsuperscript{8}, and Lys\textsuperscript{16} acetylation in combination with Lys\textsuperscript{20} dimethylation. We differentiated these two peptide species by the presence of two fragment ions, separated by 42 Da, for each of the c\textsubscript{5}, c\textsubscript{6}, and c\textsubscript{7} ions (as well as for the complementary z\textsubscript{16}, z\textsubscript{17},

![HPLC reversed phase chromatograms of isolated histones used for mass spectrometry analysis](image-url)

**FIGURE 1.** HPLC reversed phase chromatograms of isolated histones used for mass spectrometry analysis. Acid-extracted histones (chromatin-bound in the sperm, egg, pronuclei, S3, and erythrocyte cells) or heparin-Sepharose isolated (chaperone-bound in the egg) were applied on a reversed phase C8 HPLC column and eluted. The protein chromatogram (absorbance at 214 nm) for each column is shown from 35 to 85 min. a, sperm histones. b, stored, predeposition egg histones. c, early embryo equivalent pronuclei histones. d, S3 cultured cell histones. e, erythrocyte histones.
and z18 ions). We note that since the Lys5 acetyl species and the Lys8 acetyl species do not separate chromatographically on a C18 column, it is not possible to obtain separate relative abundance information.

A summary of all combinations of PTMs observed on H4 1–23 from each cell type is shown in Table 2. Full N-terminal acetylation is observed in every combination in all cell types and therefore will be disregarded when the specific sites of modification are discussed.

As seen in Table 2, analysis of H4 1–23 from transcriptionally silent sperm chromatin displays a large amount of Lys20 dimethylation, a known silencing mark, in combination with anywhere from two to four acetylations, with the majority of Lys20 dimethylation occurring in combination with acetylation over

**TABLE 1**

Summary of high resolution analyses of PTMs seen on H3 peptides from Xenopus sperm, egg, erythrocyte, and S3 cultured cells

Five H3 propionylated tryptic peptides (TKQTAR; KSTGGKAPR; QKLATKAAR; KSAPATGGVKKPHR; EIAQDFKTLDR) are shown in the left column. Each apparent modification on the lysines in those peptides is highlighted in the second column. The relative abundances of each modification in each histone source are displayed in the final four columns. Note that chromatographic separation of species containing Lys27 methylation from species with Lys36 methylation was not possible in erythrocyte or S3 cells. Therefore, relative abundance information for mono- and dimethylation are reported as a single value.

| Peptide | Modification | Sperm | Early Embryo | Erythrocyte | S3 |
|---------|--------------|-------|--------------|-------------|----|
| TKQTAR  | Unmodified   | ⬤      | ⬤            | ⬤           | ⬤  |
|         | K4 Monomethyl| ⬤      | ⬤            | ⬤           | ⬤  |
|         | K4 Dimethyl  | ⬤      | ⬤            | ⬤           | ⬤  |
|         | K4 Trimethyl | ⬤      | ⬤            | ⬤           | ⬤  |
| KSTGGKAPR| Unmodified | ⬤      | ⬤            | ⬤           | ⬤  |
|         | K9 Monomethyl| ⬤      | ⬤            | ⬤           | ⬤  |
|         | K9 Dimethyl  | ⬤      | ⬤            | ⬤           | ⬤  |
|         | K9 Trimethyl | ⬤      | ⬤            | ⬤           | ⬤  |
|         | K9/K14 Monoacetyl| ⬤ | ⬤ | ⬤ | ⬤ |
|         | K9 Monomethyl + K14 Monoacetyl| ⬤ | ⬤ | ⬤ | ⬤ |
|         | K9 Dimethyl + K14 Monoacetyl| ⬤ | ⬤ | ⬤ | ⬤ |
|         | K9 Trimethyl + K14 Monoacetyl| ⬤ | ⬤ | ⬤ | ⬤ |
| K15QLATK23AAR| Unmodified | ⬤      | ⬤            | ⬤           | ⬤  |
|         | K18/K23 Monomethyl| ⬤ | ⬤ | ⬤ | ⬤ |
|         | K18/K23 Monoacetyl| ⬤ | ⬤ | ⬤ | ⬤ |
|         | K18 + K23 Diacyl| ⬤ | ⬤ | ⬤ | ⬤ |
| K25SAPATGGV36KPHR*| Unmodified | ⬤      | ⬤            | ⬤           | ⬤  |
|         | K27/K36 Monomethyl| ⬤ | ⬤ | ⬤ | ⬤ |
|         | K27/K36 Dimethyl| ⬤ | ⬤ | ⬤ | ⬤ |
|         | K27 Trimethyl| ⬤ | ⬤ | ⬤ | ⬤ |
|         | K27 Monoacetyl| ⬤ | ⬤ | ⬤ | ⬤ |
|         | K27 Monomethyl + K36 Monomethyl| ⬤ | ⬤ | ⬤ | ⬤ |
|         | K27 Dimethyl + K36 Dimethyl| ⬤ | ⬤ | ⬤ | ⬤ |
|         | K27/K36 Monomethyl + Dimethyl**| ⬤ | ⬤ | ⬤ | ⬤ |
|         | K27/K36 Monomethyl + Trimethyl**| ⬤ | ⬤ | ⬤ | ⬤ |
|         | K27/K36 Dimethyl + Trimethyl**| ⬤ | ⬤ | ⬤ | ⬤ |
| EIAQDFK79TDLR| Unmodified | ⬤      | ⬤            | ⬤           | ⬤  |
|         | K79 Methyl| ⬤      | ⬤            | ⬤           | ⬤  |
|         | K79 Dimethyl| ⬤      | ⬤            | ⬤           | ⬤  |

* modifications on K37 were not observed
† relative abundance combines K27 and K36 monomethylation
** both possibilities were observed by MS/MS

- Not Detected (or <1% relative abundance)
  - 1-10% relative abundance
  - 10-50% relative abundance
  - 50-75% relative abundance
  - >75% relative abundance
either Lys\(^{12}\) or Lys\(^{16}\). Additionally, sperm H4 uniquely contains slight amounts of Lys\(^{20}\) trimethylation as well as low level Arg\(^{3}\) monomethylation. A distinct absence of Ser\(^{1}\) phosphorylation is also apparent.

H4 1–23 from stored egg histones shows Lys\(^{5}\) and Lys\(^{12}\) diacetylation present at 50% relative abundance. In addition, little methylation and phosphorylation is observed, and the total number of combinations is drastically reduced compared with the other cell types analyzed; specifically, only 10 unique combinations of modifications were observed on stored egg H4 1–23.

A dramatic increase in the overall number of combinations of modifications is observed upon progression from the gametogenic cell stages (sperm and egg) to the early embryo cell type. Specifically, 34 unique combinations of modifications are observed, with a majority containing dimethylation over either Arg\(^{3}\) or Lys\(^{20}\). Interestingly, these two marks do not exist on the same N-terminal tail and thus appear to be mutually exclusive to one another. Also notable is the massive increase in Ser\(^ {1}\) phosphorylation in the early embryo, a known mark of mitosis and DNA replication (23), which is observed in multiple different combinations. Additionally, diacetylation over Lys\(^{5}\), Lys\(^{6}\), Lys\(^{12}\), and Lys\(^{16}\) is present to a significant extent, some of which may be remnants of the massive amounts of Lys\(^{5}\) and Lys\(^{12}\) diacetylation seen in stored egg histones.

An apparent loss of overall acetylation on S3 histone H4 1–23 is seen in Table 2, with a substantial amount of the protein existing with a single, N-terminal acetylation. A loss of both Ser\(^{1}\) phosphorylation and Arg\(^{3}\) dimethylation from the early embryo stage is clear, and an increase in Lys\(^{20}\) dimethylation is evident.

**DISCUSSION**

Here we presented the mass spectrometry analysis of histones H3 and H4 from the eggs, sperm, early embryo, erythrocytes, and S3 cultured cells of the frog *X. laevis*. Along with our accompanying article (40), this is the first reported “life span” analysis of global histone modifications from gametic, embryonic, and somatic cell types. Our mass spectrometry analysis is mostly consistent with and complementary to the immunoblotting reported in the accompanying article (40). It also provides a thorough and more quantitative analysis of the H3 modifications than does immunoblotting alone.
thermore, our use of ETD fragmentation of histone H4 from the sperm, egg, early embryo, and S3 cells allowed us to track H4 species with multiple modifications on a single N-terminal tail, providing a unique picture of the modification transitions in developmental time and space.

The analysis of histones from these varied sources in the frog has an additional advantage over the more typically studied cultured cells: the sperm, early embryo, and erythrocytes are all transcriptionally competent, whereas the S3 cultured cells are exponentially growing, transcriptionally active cells. Other histone modification analyses have been performed on transcriptionally active samples (24, 25). Ongoing transcription is tightly correlated with changes in histone modification (26), temporally both before and after transcription. Therefore, our analysis of permanently silenced (erythrocytes), transcriptionally “poised” (sperm and egg), and transcriptionally active (S3 cells) provides a snapshot of distinct chromatin states. For instance, purified egg histones are unique in that the combinations of modifications observed do not connote a specific transcriptional state, but rather are so-called “predeposition marks.” These marks are indicative of nonchromatin, chaperone-bound localization of histones in preparation for nuclear remodeling of sperm chromatin upon fertilization and the subsequent rapid cell cycles during early development.

Our analysis of histone H3 modifications revealed some intriguing global patterns of modification. We note that K4me1 was apparent in all cell types, regardless of transcriptional competence, whereas only the S3 cells had appreciable levels of K4me2 and K4me3, consistent with the known role of higher states of Lys4 methylation and transcriptional activation (27). Also noteworthy is the distribution of K9me1, highly abundant across all four cell types and probably indicative of a variety of cellular processes, including ongoing transcription and heterochromatin formation (28). Intriguingly, K9me3 was present in all of the cell types, but most abundant in erythrocytes, consistent with the permanent silencing of the erythrocyte nuclei; despite this enriched Lys9 methylation and heterochromatin formation, *Xenopus* erythrocytes appear not to have HP1 protein (29).

We also observed a disparate range of histone H3 Lys27 and Lys36 methylation patterns. K27me has been correlated with silencing mediated by Polycomb group effector proteins (30), yet it has also been shown to be part of a so-called “bivalent” mark when coexisting on a chromatin locus with K4me (31). In this bivalent context, the chromatin is thought to be poised for a developmental switch between activation and repression.
Intriguingly, the sperm and the early embryo have lower total levels of Lys\(^{27}\) methylation than do the somatic erythrocyte and S3 cells, perhaps consistent with a developmentally specified lineage commitment and concomitant increase in Lys\(^{27}\) methylation. In the absence of a top- or middle-down analysis of the H3 tail in these samples, we cannot determine to what extent the Lys4 and Lys27 methyl marks constitute a bivalent domain. These domains are specified in a locus-specific fashion, so a global analysis of this sort is a crude tool for parsing these differences.

Intriguingly, we only observed an overlap of Lys\(^{27}\) and Lys\(^{36}\) methylation on the H3 27–40 tryptic peptide on the sperm H3. Lys36 methylation is highly correlated with ongoing transcription (28, 32), whereas Lys27 methylation is correlated with transcriptional silencing (30). We have previously observed an overlap of Lys27 and Lys36 methylation (25) in somatic HeLa and 293 cultured cells. The curious overlap between two modifications that have putatively opposite biological effects may suggest that these two modifications may encode a “bivalent” domain, similar to that between K4me and K27me (31). Clearly, future work will be necessary to determine the functional role of the simultaneous presence of these modifications. It would also be interesting to ask whether K4me, K27me, and K36me all simultaneously exist on the same H3 tail. We do also note that the relative abundances of the individual H3 Lys\(^{27}\) and Lys\(^{36}\) methylation species represented in Table 1 are pooled, since they are isobaric and cannot be chromatographically separated. Therefore, we cannot make direct comparisons with the slightly different pattern of enrichment of H32K27me species in the immunoblotting in the accompanying article (40), since the abundances presented here may be more due to the presence of K36me.

Analysis of the H3 peptide containing Lys\(^{79}\) reveals an interesting trend in that both sperm and early embryo contain relatively significant amounts of both mono- and dimethylation, whereas erythrocyte H3 is highly enriched in monomethylation but not dimethylation and S3 H3 contains only dimethylation and no monomethylation. The significance of these observations is unclear.

Finally, we note that although we observed H3S10 phosphorylation in the accompanying article (40) by immunoblotting of the oocyte, egg, and early embryo histones, thorough searching of the mass spectra did not reveal an appreciable quantity of S10ph. Although our immunoblotting results are consistent...
with the known appearance of H3S10ph during oocyte meiotic maturation (33), the discrepancy with the mass spectrometry analysis is probably explained by the greater sensitivity of chemiluminescent detection.

Our analysis of histone H4 using ETD uncovered a number of patterns of modification coexistence and negative cross-talk. The most striking observation is the relative dearth of modifications on the somatic and transcriptionally active S3 cell histones. The bulk of S3 histone H4 proteins only contain one or two total modifications on the 1–23 tail. In contrast, the early embryo histone H4 contains three, four, and even five total modifications on the 1–23 tail. We did not observe any preference for particular sites of H4 tail lysine acetylation, except for a marked enrichment of Lys⁵ and Lys¹² acetylation in the egg, in contrast to the explicit preference observed in Drosophila (for Lys¹²) and in Tetrahymena (for Lys¹₁, the

**TABLE 2**

**Representation of every PTM combination observed on the N terminus of H4 from Xenopus sperm, egg, early embryo, and S3 cultured cells**

Each row in the table represents a single discrete H4 1–23 peptide species. The left columns indicate the potentially modified residues on each peptide; empty boxes indicate that the residue is unmodified on that species, whereas colored boxes, as annotated in the key at the bottom, show the presence of a particular modification on that species. All observed peptides were acetylated on the N terminus. The relative abundances of each modified peptide species in each histone source are displayed in the final four columns. Isobaric species that do not separate chromatographically are grouped by brackets. Relative abundances between the isobaric species could not be adequately obtained; thus, each unique species within a bracketed group that was observed via ETD/PTR MS/MS is reported as existing at equivalent relative abundance.

| N-term | S1   | R3   | K5   | K8   | K12  | K16  | K20  | Sperm | Egg  | Early Embryo | S3  |
|--------|------|------|------|------|------|------|------|-------|------|--------------|----|
| 0 Ac   |      |      |      |      |      |      |      |       |      |              |    |
| 1 Ac   |      |      |      |      |      |      |      |       |      |              |    |
| 2 Ac   |      |      |      |      |      |      |      |       |      |              |    |
| 3 Ac   |      |      |      |      |      |      |      |       |      |              |    |

- Not Detected
- 1-10% Relative abundance
- 10-50% Relative abundance
- 50% Relative abundance

Mass Spectrometry Analysis of Histones in X. laevis

JANUARY 9, 2009 • VOLUME 284 • NUMBER 2

JOURNAL OF BIOLOGICAL CHEMISTRY

1083
Mass Spectrometry Analysis of Histones in X. laevis
equivalent residue to Lys\textsuperscript{12} in metazoans) (34). This difference is probably due to the higher sensitivity of mass spectrometry employed in our study.

H4 Lys\textsuperscript{8} and Lys\textsuperscript{12} acetylation are known to be predeposition modifications, highly enriched prior to chromatin assembly during S-phase (35, 36), yet no known molecular function has been ascribed to these modifications. Tellingly, mutation of all four of the lysine residues in the yeast H4 N-terminal tail resulted in lethality, whereas no single mutation had a pronounced effect (37), consistent with a combinatorial role for lysine modifications. Indeed, a later study identified the crucial residues for chromatin assembly and viability as Lys\textsuperscript{5}, Lys\textsuperscript{8}, and Lys\textsuperscript{12} (38). Furthermore, our observation of many more predeposition H4 isoforms in the egg, including lysine-acetylated (on each of Lys\textsuperscript{5}, Lys\textsuperscript{8}, and Lys\textsuperscript{12}), lysine-triacetylated (including on Lys\textsuperscript{5}, Lys\textsuperscript{8}, and Lys\textsuperscript{12}), serine (Ser\textsuperscript{1})-phosphorylated, and arginine (Arg\textsuperscript{3}) methylated (See Table 2), suggest an even more complex predeposition modification pattern than suspected. These observations were only possible due to our use of ETD fragmentation to analyze the combinatorial pattern of modifications on a single histone tail.

Finally, we note the entirely negative correlation between H4 arginine 3 methylation and serine 1 phosphorylation. We did not see a single species with both Ser\textsuperscript{1} phosphorylation and Arg\textsuperscript{3} methylation modifications. The significance of this observation is unclear. However, we do know that a potent arginine methyltransferase activity in Xenopus eggs methylates Arg\textsuperscript{3} and is inhibited by S1ph,\textsuperscript{4} suggesting that S1ph may serve a negative regulatory function for subsequent histone modifications.

Our observations strongly support the histone code hypothesis (3), especially in light of our recent extension of the hypothesis to explicitly include the tremendous benefits of a multivalent readout of combinatorial modifications by effector proteins (18). However, we are aware that these observations do not imply causality between relative abundances of histone modifications and histone variants and molecular and cellular phenotypes. We therefore interpret our observation of such striking differences in global enrichment of histone modifications as, at minimum, support for a syntactic definition of the histone code, in which histone modifications provide the syntax (or context) in which chromatin-mediated DNA transactions occur. We anticipate that such histone modification indexing, both on a global cell-specific scale as we have presented here and using ChIP-Seq (39) and other techniques, will have wide ranging significance for the study of epigenetic phenomena.

Acknowledgments—We are grateful to Hiro Funabiki and members of the Funabiki laboratory for generously sharing the frog colony and Jim Maller for Xenopus S3 cells. We thank Sandra Hake and Laura Banaszynski for constructive comments on the manuscript. We also thank Scott Ficarro for instruction on making Lithisil frits.

REFERENCES
1. Lugner, K., and Richmond, T. J. (1998) Curr. Opin. Struct. Biol. 8, 33–40
2. Kouzarides, T. (2007) Cell 128, 693–705
3. Strahl, B. D., and Allis, C. D. (2000) Nature 403, 41–45
4. Jenuwein, T., and Allis, C. D. (2001) Science 293, 1074–1080
5. Taverna, S., Li, H., Ruthenburg, A., Allis, C., and Patel, D. (2007) Nat. Struct. Mol. Biol. 14, 1025–1040
6. Zubarev, R. A., Horn, D. M., Fridriksson, E. K., Kelleher, N. L., Kruger, N. A., Lewis, M. A., Carpenter, B. K., and McLafferty, F. W. (2000) Anal. Chem. 72, 563–573
7. Coon, J., Ueberheide, B., Syka, J., Dryhurst, D., Ausio, J., Shabanowitz, J., and Hunt, D. (2005) Proc. Natl. Acad. Sci. U. S. A. 102, 9463–9468
8. Syka, J. E., Coon, J. J., Schroeder, M. J., Shabanowitz, J., and Hunt, D. F. (2004) Proc. Natl. Acad. Sci. U. S. A. 101, 9528–9533
9. Medzhuradszky, 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) Mol. Cell Proteomics 3, 872–886
10. Mackay, C. L., Ramsahoye, B., Burgess, K., Cook, K., Weidt, S., Creanor, J., Harrison, D., Langridge-Smith, P., Hupp, T., and Hayward, L. (2008) Anal. Chem. 80, 4147–4153
11. Phanstiel, D., Brumbaugh, J., Berggren, W. T., Conard, K., Feng, X., Lemberg, M. E., McAlister, G. C., Thomson, J. A., and Coon, J. J. (2008) Proc. Natl. Acad. Sci. U. S. A. 105, 4093–4098
12. Garcia, B. A., Pesavento, J. J., Mizzen, C. A., and Kelleher, N. L. (2007) Nat. Methods 4, 487–489
13. Pesavento, J. J., Bullock, C. R., LeDuc, R. D., Mizzen, C. A., and Kelleher, N. L. (2008) J. Biol. Chem. 283, 14927–14937
14. Boyne, M. T., II, Pesavento, J. J., Mizzen, C. A., and Kelleher, N. L. (2006) J. Proteome Res. 5, 248–253
15. Taverna, S. D., Ueberheide, B. M., Liu, Y., Tackett, A. J., Diaz, R. L., Shabanowitz, J., Chait, B. T., Hunt, D. F., and Allis, C. D. (2007) Proc. Natl. Acad. Sci. U. S. A. 104, 2086–2091
16. Jiang, L., Smith, J. N., Anderson, S. L., Ma, P., Mizzen, C. A., and Kelleher, N. L. (2007) J. Biol. Chem. 282, 27923–27934
17. Garcia, B. A., Thomas, C. E., Kelleher, N. L., and Mizzen, C. A. (2008) J. Proteome Res. 7, 4225–4236
18. Ruthenburg, A., Li, H., Patel, D., and Allis, C. (2007) Nat. Rev. Mol. Cell. Biol. 8, 983–994
19. Shechter, D., Costanzo, V., and Gautier, J. (2004) Nat. Cell Biol. 6, 648–655
20. Shechter, D., Dormann, H., Allis, C., and Hake, S. (2007) Nat. Protoc. 2, 1445–1457
21. Schroeder, M. J., Shabanowitz, J., Schwartz, J. C., Hunt, D. F., and Coon, J. J. (2004) Anal. Chem. 76, 3590–3598
22. Garcia, B. A., Busby, S. A., Shabanowitz, J., Hunt, D. F., and Mishra, N. (2005) J. Proteome Res. 4, 2032–2042
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) Chromosoma 112, 360–371
24. Garcia, B. A., Hake, S. B., Diaz, R. L., Kauer, M., Morris, S. A., Recht, J., Shabanowitz, J., Mishra, N., Strahl, B. D., Allis, C. D., and Hunt, D. F. (2007) J. Biol. Chem. 282, 7641–7655
25. Hake, S. B., Garcia, B. A., Duncan, E. M., Kauer, M., Delaire, G., Shabanowitz, J., Bazett-Jones, D. P., Allis, C. D., and Hunt, D. F. (2006) J. Biol. Chem. 281, 559–568
26. Berger, S. L. (2007) Nature 447, 407–412
27. Ruthenburg, A., Allis, C., and Wysocka, J. (2007) Mol. Cell 25, 15–30
28. Kim, A., Kiefer, C. M., and Dean, A. (2007) Mol. Cell. Biol. 27, 1271–1279
29. Gilbert, N., Boyle, S., Sutherland, H., de Las Heras, J., Allan, J., Jenuwein, T., and Bickmore, W. A. (2003) EMBO J. 22, 5540–5550
30. Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R. S., and Zhang, Y. (2002) Science 298, 1039–1043
31. Bernstein, B. E., Mikkelsen, T. S., Xie, X., Kamal, M., Huebert, D. J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., Jaenisch, R., Wagschal, A., Feil, R., Schreiber, S. L., and Lander, E. S. (2006) Cell 125, 315–326
32. Morris, S. A., Shibata, Y., Noma, K., Tsukamoto, Y., Warren, E., Temple, B., Grewal, S. I., and Strahl, B. D. (2005) Eukaryot. Cell 4, 1446–1454

\textsuperscript{4} D. Shechter, C. D. Allis, manuscript in preparation.
33. Schmitt, A., Gutierrez, G. J., Lenart, P., Ellenberg, J., and Nebreda, A. R. (2002) *FEBS Lett.* **518**, 23–28
34. Sobel, R. E., Cook, R. G., and Allis, C. D. (1994) *J. Biol. Chem.* **269**, 18576–18582
35. Chicoine, L. G., Schulman, I. G., Richman, R., Cook, R. G., and Allis, C. D. (1986) *J. Biol. Chem.* **261**, 1071–1076
36. Allis, C. D., Chicoine, L. G., Richman, R., and Schulman, I. G. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82**, 8048–8052
37. Megee, P. C., Morgan, B. A., Mittman, B. A., and Smith, M. M. (1990) *Science* **247**, 841–845
38. Ma, X. J., Wu, J., Altheim, B. A., Schultz, M. C., and Grunstein, M. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95**, 6693–6698
39. Barski, A., Cuddapah, S., Cui, K., Roh, T. Y., Schones, D. E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007) *Cell* **129**, 823–837
40. Shechter, D., Nicklay, J. J., Chitta, R. K., Shabanowitz, J., Hunt, D. F., and Allis, C. D. (2009) *J. Biol. Chem.* **284**, 1064–1074