

Point of View

Unraveling the histone's potential

A proteomics perspective

Justin Brumbaugh,^{1,2,†} Doug Phanstiel^{1,†} and Joshua J. Coon^{1,3,*}

¹Department of Chemistry; ²Integrated Program in Biochemistry; and ³Department of Biomolecular Chemistry; University of Wisconsin–Madison; Madison, Wisconsin USA

[†]These authors contributed equally to this work.

Key words: histone, post-translational modification, proteomics, mass spectrometry, epigenetics

Post translational modification (PTM) of histones has long been associated with epigenetic regulation. Although genomic approaches have established correlation between a handful of histone PTMs and transcriptional states, only recently have advancements in proteomics provided the tools necessary to study histone proteins and their relevant modifications in this context. Using mass spectrometry, researchers have demonstrated the ability to determine the full repertoire of histone PTMs, their residue specific location, the combinations in which they exist, and the proteins that interact with these combinations. Moving forward it will be imperative to develop novel approaches that combine proteomic and genomic technologies to determine the functional significance of these combinations of modifications. Assays with increased specificity will resolve more focused biological questions and determine to what extent, and by what mechanisms, histones influence transcription.

Over the last decade a variety of research has established a firm link between chromatin and transcription.¹⁻⁵ One such example is the extensive work on various histone post translational modifications (PTMs), which correlate strongly with dynamic transcriptional activity.⁶⁻⁸ Since different histone PTMs associate with specific biological states in development and disease, untangling the role of a given modification or pattern of modifications is of critical importance to both fundamental science and medicine.^{7,9-12} At this juncture, however, the mechanisms underlying this potential form of epigenetic regulation remain poorly understood—that is, how does a particular mark elicit a change in transcription on a molecular level or is histone modification merely a byproduct of other regulatory mechanisms? To answer these and related questions about the potential regulatory consequences of histone modifications it is imperative to develop new technologies and to combine genomic and proteomic approaches. Here, we offer our views on both state-of-the-art and emerging technologies with an emphasis

on proteomics-based approaches for unraveling the role of histones in transcription.

Although it is difficult to establish direct causal links between histone modification and gene transcription, a number of key PTMs generally correlate with transcriptional activation or repression. For example, acetylation is largely associated with actively transcribed genes while methylation of H3K27 appear to localize largely to the promoters of repressed genes.^{6,10,13,14} For a more exhaustive review, see Berger et al.¹⁵ These generalizations must be treated with care, however, since some modifications can exert different transcriptional outcomes depending on the surrounding microenvironment.¹⁵⁻¹⁷ Considering the nucleosome alone, histone N-terminal tails are subject to myriad modifications—e.g., over 3 million combinations are possible on the histone H4 tail—and it is unclear how the cross-talk between different combinations of these modifications impact transcription. For example, what happens if both an activating and repressive mark are found on the same histone tail? The added complexity of histone variants, the extent to which DNA is packaged (i.e., is DNA fully saturated with histone?), and the dynamic nature of histone modifications convolute the situation. Historically, antibodies were used to identify histone modifications and, when used in conjunction with expression profiling, can determine a modifications' possible effects on transcription. Genomic approaches like chromatin immunoprecipitation (ChIP) have provided valuable insight into functionality associated with a given mark. That said, antibody-based approaches require a priori knowledge of a modification, can suffer from cross reactivity and epitope occlusion, and rarely detect more than a single PTM on a histone tail. Imagine histone modification patterns as complex sentences. Antibodies can perhaps recognize a few letters or a word, but this isolated piece of information lacks the context needed to interpret the sentence. To fully understand the complex regulatory consequences of a particular histone modification, we must also take the surrounding modifications into account.

Mass spectrometry (MS) is an effective tool for examining histone modifications because it (1) does not require a priori knowledge of a modification, (2) can provide quantitative information and (3) can unambiguously identify combinatorial modification patterns (i.e., multiple PTMs on a single molecule).¹⁸⁻³² The position of a given PTM is determined via tandem MS (MS/MS), since it provides direct identification of a protein's primary sequence and the presence of a PTM alters the residue mass of an associated amino acid in a

*Correspondence to: Joshua J. Coon; University of Wisconsin–Madison; Departments of Chemistry and Biomolecular Chemistry; 1101 University Avenue; Madison, Wisconsin 53706 USA; Email: jcoon@chem.wisc.edu

Submitted: 06/30/08; Accepted: 09/17/08

Previously published online as an *Epigenetics* E-publication:
<http://www.landesbioscience.com/journals/epigenetics/article/7005>

predictable way (i.e., phosphorylation adds 80 Daltons to serine, threonine or tyrosine).^{20,33-38} A number of groups have employed proteolytic digestion of histone proteins to produce peptides that are sufficiently short for sequencing via collision activated dissociation (CAD) MS/MS.^{30,39-42} The availability of high mass accuracy mass spectrometers aided this approach and allowed for the distinction between PTMs of similar mass, such as acetylation and trimethylation of lysines which differ by a mere 0.03638 Daltons.^{21,42-45} These enzyme-based approaches allowed researchers to catalog sites of modifications on histones from a number of species but provided little information regarding the combinations of modifications present on a given histone molecule; that information is largely lost upon proteolytic digestion. Some groups omitted enzymatic digestion and tandem mass spectrometry altogether and relied on high mass accuracy detection of intact histone proteins to infer the modifications present.⁴⁶ Although this approach provides information on the presence of all PTMs on a given histone molecule, it can rarely be used to determine which residues are modified. The introduction of electron-based fragmentation methods—electron capture dissociation (ECD) and electron transfer dissociation (ETD)—now allow researchers to sequence large, basic and highly modified proteins like histones.^{23,26,28,33,47-50} By employing these fragmentation methods to large, n-terminal, tail peptides or even entire proteins it is possible to study the complete combinatorial space of histone modifications.

A handful of groups have used ECD and ETD mass spectrometry to investigate specific patterns of histone modifications. In 2004, Kellher et al. sequenced histone H4 proteins using a top down approach (i.e., without enzymatic digestion prior to analysis).⁴⁹ A high mass accuracy Fourier transform-ion cyclotron resonance mass spectrometer (FT-ICR MS) capable of ECD enabled the group to unambiguously assign ten discrete modification patterns that span the entire H4 tail. Not only did this work place multiple H4 modifications in context of one another, but it also pioneered a method for relative quantification of each H4 isoform. Thus, although this work identified only a modest number of isoforms, it served as an important starting point for further studies. Building upon this work, our own group and others have focused on the entire n-terminal tails of histones, where the majority of the modifications are present. Since a majority of the PTMs of interest are located within the first 20 amino acids of H4 or first 37 amino acids of H3, proteases were used to specifically cleave the N-terminal portion of these proteins. Hunt et al. pioneered this approach in 2005 on histone H3 from HeLa cells and later used it to identify more than fifty H3 isoforms in *Tetrahymena thermophila*.^{48,51} A year later Garcia et al. identified over 150 n-terminal tail isoforms in HeLa cells using a similar approach.⁵² And recently, our own study yielded 74 distinct H4 isoforms in human embryonic stem cells.¹⁹

Closer inspection of these data sets revealed interesting patterns of histone modification. For instance, H3K4me3 and H3 hyperacetylation, both of which are commonly associated with transcriptional activation, are often found together in HeLa cells. Alternatively, H4R3 methylation was only observed in the presence of multiply methylated H4K20, an interesting observation since these marks are thought to have competing effects.¹⁹ One explanation for this is that opposing marks temper the transcriptional output of either mark alone. On the other hand, di- or tri- methylation of H4K20 may be prerequisite for the recruitment of factors that modify

H4R3. The exact nature of these relationships and the subsequent transcriptional consequences remain to be determined; however, such findings provide evidence that histones are modified in a concerted and purposeful manner and beg the question: do specific patterns of modifications exhibit distinct function? Moving forward, it will be critical to seek out cross-disciplinary work that will apply the current technology to answer relevant biological questions like that posed above.

Investigation of histone isoforms in various biological samples has allowed for correlation with cell states. For example, Kelleher et al. used MS to examine the effects of cell cycle on histone isoforms in HeLa cells.^{50,53} The group found that newly synthesized histone H4 is progressively methylated at lysine 20 as cells cycle and that this process is largely unaffected by acetylation. Meanwhile, our own group has quantitatively tracked histone H4 tail modifications as human embryonic stem (ES) cells differentiate.¹⁹ We found that acetylation on H4 tails decreases as cells progress through differentiation, while di-methyl isoforms increase steadily and unmethylated H4 drops precipitously. Note that these changes happen on the same time scale as ES cell commitment. Each of these experiments provide interesting ties between specific patterns of histone PTMs and biological states; however, none of these studies claim a causal relationship between the identified marks and the biology.

Correlative studies are a good foundation for answering biological questions; however, once a histone modification pattern is identified, more in depth studies will be necessary to better understand how these modifications elicit changes in transcription. There are different ways that modification of histones can alter gene expression. It is believed, for example, that addition or removal of acetylation alters chromatin structure, making it more or less accessible to transcriptional machinery.⁵⁴⁻⁵⁶ Also, histones and their respective modifications may serve as binding partners for transcriptional regulators that, upon recognizing a certain mark, then exert their function at that specific location.^{2,4,57-59} Several recent publications have examined the latter possibility in more depth. For example, Heo et al. overexpressed a tagged portion of histone H3 (residues 1–40), pulled down the tagged histone and its associated protein complexes, and analyzed the interacting proteins via mass spectrometry. The group observed a number of relevant proteins, including PRMT5, an arginine methyl-transferase, and several constituents of the histone deacetylase complex.⁶⁰ Unfortunately, however, since the group had no knowledge of the recombinant histones' modifications, it was not possible to determine which histone PTMs correspond to the interacting proteins. Using a more directed approach, Wysocka et al. synthesized biotin-linked histone H3 tails with known modifications. These tails were then incubated with cell lysate and, following washing and elution, the resulting samples were analyzed via mass spectrometry to characterize proteins that bind the respective H3 isoforms. From this study, the group identified WDR5, a protein that binds with high specificity to hypermethylated forms of H3K4.⁶⁰ Interestingly, WDR5 appears to be involved in activation of certain *HOX* genes and injection of a WDR5 morpholino into *Xenopus* embryos resulted in developmental defects.⁵⁷ By characterizing a set of known histone modification-specific interactions and examining the binding partners' cellular functions, more direct connections can be made between a given histone mark and the resulting biological function. Once this is established, it is possible to gain even more information

by applying quantitative mass spectrometry to compare lysates from two different biological treatments or conditions (i.e., are particular binding partners increased in disease vs. normal tissue?).

A major problem with the techniques described above is the global nature inherent in the experiments. More simply put, the entire population of proteins, microenvironments and genomic loci within the cell are sampled together, making it difficult to assign functional meaning to the results. ChIP is a notable exception and the genome-specific information obtained from this kind of analysis is certainly one reason that the technique is so powerful. However, even ChIP is limited since typically, only a single histone modification can be probed at a time. Once again, this limits analysis since the modification's context in relation to the surrounding PTMs is unknown. The next step in this field may be finding a way to perform proteomics on the level of a single genomic locus. This would open the door to numerous meaningful experiments, whether considering histone modification patterns at insulator regions or examining histone binding proteins at intron-exon junctions. As analytical instrumentation improves, these experiments are becoming more feasible and we inch closer to unraveling the histone regulatory role. More specifically, improved technology and cross-disciplinary work should help to determine whether particular histone modifications exert specific transcriptional regulation or simply correlate with gene expression. It seems likely that both occur and we must be cognizant that a given modification's function may change based on cell line, the cellular environment, cell cycle, etc. Improving technology, together with combined genomic and proteomic approaches will help to fill in these gaps.

References

- Pray-Grant MG, Daniel JA, Schieltz D, Yates JR, Grant PA. Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation. *Nature* 2005; 433:434-8.
- Jenuwein T, Allis CD. Translating the histone code. *Science* 2001; 293:1074-80.
- Turner BM. Cellular memory and the histone code. *Cell* 2002; 111:285-91.
- Strahl BD, Allis CD. The language of covalent histone modifications. *Nature* 2000; 403:41-5.
- Pokholok DK, Harbison CT, Levine S, Cole M, Hannett NM, Lee TI, et al. Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell* 2005; 122:517-27.
- Bernstein BE, Meissner A, Lander ES. The mammalian epigenome. *Cell* 2007; 128:669-81.
- Lee TI, Jenner RG, Boyer LA, Guenther MG, Levine SS, Kumar RM, et al. Control of developmental regulator's by polycomb in human embryonic stem cells. *Cell* 2006; 125:301-13.
- Rice JC, Allis CD. Histone methylation versus histone acetylation: new insights into epigenetic regulation. *Curr Opin Cell Biol* 2001; 13:263-73.
- Fraga MF, Ballestar E, Villar-Garea A, Boix-Chornet M, Espada J, Schotta G, et al. Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nat Genet* 2005; 37:391-400.
- Pan GJ, Tian SL, Nie J, Yang CH, Ruotti V, Wei HR, et al. Whole-genome analysis of histone H3 lysine 4 and lysine 27 methylation in human embryonic stem cells. *Cell Stem Cell* 2007; 1:299-312.
- Torres-Padilla ME, Parfitt DE, Kouzarides T, Zernicka-Goetz M. Histone arginine methylation regulates pluripotency in the early mouse embryo. *Nature* 2007; 445:214-8.
- Selgson DB, Horvath S, Shi T, Yu H, Tze S, Grunstein M, et al. Global histone modification patterns predict risk of prostate cancer recurrence. *Nature* 2005; 435:1262-6.
- Meshorer E, Misteli T. Opinion—Chromatin in pluripotent embryonic stem cells and differentiation. *Nat Rev Mol Cell Biol* 2006; 7:540-6.
- Roth SY, Denu JM, Allis CD. Histone acetyltransferases. *Annu Rev Biochem* 2001; 70:81-120.
- Berger SL. The complex language of chromatin regulation during transcription. *Nature* 2007; 447:407-12.
- Guenther MG, Levine SS, Boyer LA, Jaenisch R, Young RA. A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* 2007; 130:77-88.
- Bernstein BE, Mikkelsen TS, Xie XH, Kamal M, Huebert DJ, Cuff J, et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 2006; 125:315-26.
- Gingras AC, Gstaiger M, Raught B, Aebersold R. Analysis of protein complexes using mass spectrometry. *Nat Rev Mol Cell Biol* 2007; 8:645-54.
- Phanstiel D, Brumbaugh J, Berggren WT, Conard K, Feng X, Levenstein ME, et al. Mass spectrometry identifies and quantifies 74 unique histone H4 isoforms in differentiating human embryonic stem cells. *Proc Natl Acad Sci USA* 2008; 105:4093-8.
- Witze ES, Old WM, Resing KA, Ahn NG. Mapping protein post-translational modifications with mass spectrometry. *Nat Methods* 2007; 4:798-806.
- Garcia BA, Hake SB, Diaz RL, Kauer M, Morris SA, Recht J, et al. Organismal differences in post-translational modifications in histones H3 and H4. *J Biol Chem* 2007; 282:7641-55.
- Pesavento JJ, Kim YB, Taylor GK, Kelleher NL. Shotgun annotation of histone modifications: A new approach for streamlined characterization of proteins by top down mass spectrometry. *J Am Chem Soc* 2004; 126:3386-7.
- Boyne MT, Pesavento JJ, Mizzen CA, Kelleher NL. Precise characterization of human histories in the H2A gene family by top down mass spectrometry. *J Proteome Res* 2006; 5:248-53.
- Garcia BA, Siuti N, Thomas CE, Mizzen CA, Kelleher NL. Characterization of neurohistone variants and post-translational modifications by electron capture dissociation mass spectrometry. *Int J Mass Spectrom* 2007; 259:184-96.
- Garcia BA, Joshi S, Thomas CE, Chitta RK, Diaz RL, Busby SA, et al. Comprehensive phosphoprotein analysis of linker histone H1 from *Tetrahymena thermophila*. *Mol Cellular Proteomics* 2006; 5:1593-609.
- Thomas CE, Kelleher NL, Mizzen CA. Mass spectrometric characterization of human histone H3: A bird's eye view. *J Proteome Res* 2006; 5:240-7.
- Siuti N, Roth MJ, Mizzen CA, Kelleher NL, Pesavento JJ. Gene-specific characterization of human histone H2B by electron capture dissociation. *J Proteome Res* 2006; 5:233-9.
- Medzihradsky KF, Zhang X, Chalkley RJ, Guan S, McFarland MA, Chalmers MJ, et al. Characterization of *Tetrahymena* histone H2B variants and posttranslational modifications by electron capture dissociation (ECD) Fourier transform ion cyclotron mass spectrometry (FT-ICR MS). *Mol Cellular Proteomics* 2004; 3:872-86.
- Su XD, Zhang LW, Lucas DM, Davis ME, Knapp AR, Green-Church KB, et al. Histone H4 acetylation dynamics determined by stable isotope labeling with amino acids in cell culture and mass spectrometry. *Anal Biochem* 2007; 363:22-34.
- Zhang LW, Freitas MA. Comparison of peptide mass mapping and electron capture dissociation as assays for histone posttranslational modifications. *Int J Mass Spectrom* 2004; 234:213-25.
- Ong SE, Mann M. Mass spectrometry-based proteomics turns quantitative. *Nat Chem Biol* 2005; 1:252-62.
- Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol* 1999; 17:994-9.
- Syka JEP, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc Natl Acad Sci USA* 2004; 101:9528-33.
- Hunt DF, Yates JR, Shabanowitz J, Winston S, Hauer CR. Protein sequencing by tandem mass-spectrometry. *Proc Natl Acad Sci USA* 1986; 83:6233-7.
- Pandey A, Mann M. Proteomics to study genes and genomes. *Nature* 2000; 405:837-46.
- Ficarro SB, McClelland ML, Stukenberg PT, Burke DJ, Ross MM, Shabanowitz J, et al. Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae*. *Nat Biotechnol* 2002; 20:301-5.
- Oda Y, Nagasu T, Chait BT. Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. *Nat Biotechnol* 2001; 19:379-82.
- Zhou HL, Watts JD, Aebersold R. A systematic approach to the analysis of protein phosphorylation. *Nat Biotechnol* 2001; 19:375-8.
- Tackett AJ, Dilworth DJ, Davey MJ, O'Donnell M, Aitchison JD, Rout MP, et al. Proteomic and genomic characterization of chromatin complexes at a boundary. *J Cell Biol* 2005; 169:35-47.
- Janzen CJ, Fernandez JP, Deng HT, Diaz R, Hake SB, Cross GAM. Unusual histone modifications in *Trypanosoma brucei*. *FEBS Lett* 2006; 580:2306-10.
- Ren C, Zhang LW, Freitas MA, Ghoshal K, Parthun MR, Jacob ST. Peptide mass mapping of acetylated isoforms of histone H4 from mouse lymphosarcoma cells treated with histone deacetylase (HDACs) inhibitors. *J Am Soc Mass Spectrom* 2005; 16:1641-53.
- Garcia BA, Busby SA, Shabanowitz J, Hunt DF, Mishra N. Resetting the epigenetic histone code in the MRL-lpr/lpr mouse model of lupus by histone deacetylase inhibition. *J Proteome Res* 2005; 4:2032-42.
- Syka JEP, Marto JA, Bai DL, Horning S, Senko MW, Schwartz JC, et al. Novel linear quadrupole ion trap/FT mass spectrometer: Performance characterization and use in the comparative analysis of histone H3 post-translational modifications. *J Proteome Res* 2004; 3:621-6.
- Garcia BA, Barber CM, Hake SB, Ptak C, Turner FB, Busby SA, et al. Modifications of human histone H3 variants during mitosis. *Biochemistry (Mosc)* 2005; 44:13202-13.
- Garcia BA, Busby SA, Barber CM, Shabanowitz J, Allis CD, Hunt DF. Characterization of phosphorylation sites on histone H1 isoforms by tandem mass spectrometry. *J Proteome Res* 2004; 3:1219-27.
- Zhang LW, Freitas MA, Wickham J, Parthun MR, Klisovic MI, Marcucci G, et al. Differential expression of histone post-translational modifications in acute myeloid and chronic lymphocytic leukemia determined by high-pressure liquid chromatography and mass spectrometry. *J Am Soc Mass Spectrom* 2004; 15:77-86.
- Zubarev RA, Kelleher NL, McLafferty FW. Electron capture dissociation of multiply charged protein cations. A nonergodic process. *J Am Chem Soc* 1998; 120:3265-6.

48. Coon JJ, Ueberheide B, Syka JEP, Dryhurst DD, Ausio J, Shabanowitz J, et al. Protein identification using sequential ion/ion reactions and tandem mass spectrometry. *Proc Natl Acad Sci USA* 2005; 102:9463-8.
49. Pesavento JJ, Mizzen CA, Kelleher NL. Quantitative analysis of modified proteins and their positional isomers by tandem mass spectrometry: Human histone H4. *Anal Chem* 2006; 78:4271-80.
50. Pesavento JJ, Yang H, Kelleher NL, Mizzen CA. Certain and progressive methylation of histone H4 at lysine 20 during the cell cycle. *Mol Cell Biol* 2008; 28:468-86.
51. Taverna SD, Ueberheide BM, Liu YF, Tackett AJ, Diaz RL, Shabanowitz J, et al. Long-distance combinatorial linkage between methylation and acetylation on histone H3N termini. *Proc Natl Acad Sci USA* 2007; 104:2086-91.
52. Garcia BA, Pesavento JJ, Mizzen CA, Kelleher NL. Pervasive combinatorial modification of histone H3 in human cells. *Nat Methods* 2007; 4:487-9.
53. Pesavento JJ, Bullock CR, Leduc RD, Mizzen CA, Kelleher NL. Combinatorial modification of human histone H4 quantitated by two-dimensional liquid chromatography coupled with top down mass spectrometry. *J Biol Chem* 2008; 283:14927-37.
54. Durrin LK, Mann RK, Kayne PS, Grunstein M. Yeast Histone H4 N-Terminal Sequence Is Required for Promoter Activation In vivo. *Cell* 1991; 65:1023-31.
55. Struhl K. Histone acetylation and transcriptional regulatory mechanisms. *Genes Dev* 1998; 12:599-606.
56. Yoshida M, Horinouchi S, Beppu T. Trichostatin-a and Trapoxin—Novel Chemical Probes for the Role of Histone Acetylation in Chromatin Structure and Function. *Bioessays* 1995; 17:423-30.
57. Wysocka J, Swigut T, Milne TA, Dou YL, Zhang X, Burlingame AL, et al. WDR5 associates with histone H3 methylated at K4 and is essential for H3K4 methylation and vertebrate development. *Cell* 2005; 121:859-72.
58. Hecht A, Laroche T, Strahlbolsinger S, Gasser SM, Grunstein M. Histone H3 and H4 N-Terminal Interact with Sir3 and Sir4 Proteins—a Molecular-Model for the Formation of Heterochromatin in Yeast. *Cell* 1995; 80:583-92.
59. Edmondson DG, Smith MM, Roth SY. Repression domain of the yeast global repressor Tup1 interacts directly with histones H3 and H4. *Genes Dev* 1996; 10:1247-59.
60. Heo K, Kim B, Kim K, Choi J, Kim H, Zhan YX, et al. Isolation and characterization of proteins associated with histone H3 tails in vivo. *J Biol Chem* 2007; 282:15476-83.