Published in final edited form as:

J Am Chem Soc. 2017 July 12; 139(27): 9090–9096. doi:10.1021/jacs.7b03430.

Emerging Chemistry Strategies for Engineering Native Chromatin

Yael David^{*,†} and Tom W. Muir^{*,‡}

[†]Chemical Biology Program, Memorial Sloan Kettering Cancer Center, New York, New York 10065, United States

[‡]Department of Chemistry, Princeton University, Princeton, New Jersey 08544, United States

Abstract

Chromosomes present one of most challenging of all substrates for biochemical study. This is because genomic DNA is physically associated with an astonishing collection of nuclear factors, which serve to not only store the nucleic acid in a stable form, but also grant access to the information it encodes when needed. Understanding this complex molecular choreography is central to the field of epigenetics. One of the great challenges in this area is to move beyond correlative type information, which is now in abundant supply, to the point where we can truly connect the dots at the molecular level. Establishing such causal relationships requires precise manipulation of the covalent structure of chromatin. Tools for this purpose are currently in short supply, creating an opportunity that, as we will argue in this Perspective, is well suited to the sensibilities of the chemist.

Introduction

Genomic DNA is packaged in the nuclei of eukaryotic cells in the form of chromatin. The fundamental repeating unit of this nucleo-protein complex, the nucleosome, is composed of ~150 base pairs of DNA coiled around an octameric spool of histone proteins comprising two copies each of H2A, H2B, H3, and H4 (Figure 1). The compact, disc-like shape of the nucleosome allows for the establishment of higher order chromatin structure in which genomic information becomes increasingly protected from DNA transactions such as transcription as a function of the level of compaction.¹ Indeed, the degree to which chromatin is folded differentially across a genome is thought to be a fundamental component of gene regulation.² As a consequence, understanding the mechanisms by which genomic access is controlled at the level of chromatin is key to unlocking how cellular phenotypes are established and maintained, and is a central focus of the field of epigenetics.

Epigenetic regulation of eukaryotic genomes is disarmingly complex and is known to involve the coordinated action of many players, including general and specific transcription factors,³ chromatin remodeling enzymes,⁴ non-coding RNAs,⁵ and DNA methylation.⁶ In

^{*}Corresponding Authors: davidshy@mskcc.org, muir@princeton.edu.

Notes: The authors declare no competing financial interest.

addition, post-translational modification (PTM) of the histone proteins has emerged as a key epigenetic modulator that can function either by directly altering DNA accessibility or through the recruitment of transacting factors.^{7,8} Together these regulatory inputs help choreograph normal cellular fate, and, not surprisingly, their dysregulation leads to aberrant gene expression, with far-reaching implications for human biology and disease.

Recent advances in high-throughput genomic analysis, including ChIP-seq (for Chromatin Immuno-Precipitation with next generation DNA sequencing⁹), ATAC-seq (assay for transposase-accessibility of chromatin with next generation DNA sequencing¹⁰), mRNA-seq,¹¹ and Hi-C (chromosome conformation capture),¹² have revolutionized our understanding of the complex landscape of the epigenome.¹³ Moreover, sophisticated proteomics methods, fueled by advances in mass spectrometry instrumentation, have allowed for the detection and, in many cases, quantification of over 100 different PTMs (often termed “marks” in the field) on histones. Collectively, these methods have revealed how epigenetic inputs such as histone PTMs, transcription factor occupancy, and chromatin architecture correlate with transcriptional outputs,¹⁴ and have led to the idea of functionally distinct “epigenetic states”.¹⁵ Major efforts are now underway toward understanding the hierarchy of epigenetic events, including identifying what role histone marks play in establishing and maintaining epigenetic states.¹⁶ This important undertaking requires moving beyond correlative relationships, and is made incredibly challenging by the highly interconnected and dynamic nature of the underlying biochemical processes.

Histone PTM Writers, Readers, and Erasers

From proteomics efforts, we now know histones are decorated with a staggering number of PTMs, over 30 different chemotypes and counting.¹⁷ These span the gamut from small chemical modifications (e.g., acetylation, methylation, and phosphorylation) to the addition of large proteins (e.g., ubiquitination and sumoylation). Histone marks can directly alter the chromatin structure by weakening the electrostatic interaction between the cationic histone proteins and anionic DNA, as in the case of lysine acetylation. Alternatively, they can induce changes in the accessibility of the DNA by recruiting downstream effector proteins, for example ATP-dependent chromatin remodelers,¹⁸ that contain so-called “reader” domains that bind to marks in specific sequence contexts.^{19,20} Notably, these binding modules, which include the well-known bromo and chromo domains that recognize acetyl- and methyllysine moieties, respectively, are frequently found within the enzymes that install (often termed “writers”) or remove (“erasers”) histone PTMs. This situation creates the potential for these enzymes to sense the modification status of the chromatin template they act on, leading to the generation of positive feedback loops and PTM cross-talks in which the presence of one mark on the chromatin substrate modulates the addition (or removal) of a different mark.²¹ Further adding to the intricacy of this regulation, many chromatin effectors contain multiple reader domains, thereby offering the possibility for the integration of multiple signals, for example through multivalent binding mechanisms in tuning biochemical outputs.²²

While it seems likely that the regulatory inputs alluded to above contribute to the establishment of epigenetic states, deducing just how it all works at the molecular level is far from straightforward. Ideally, this requires experimentally manipulating the levels of PTMs,

an undertaking where both genetic and pharmacological approaches have significant limitations. For instance, many writer and eraser enzymes have several substrates, and as a consequence, removing their activity, either genetically or using small-molecule inhibitors, will lead to multiple perturbations, thus making phenotypic effects hard to interpret.²³ By the same token, overexpressing a given writer or eraser enzyme is unlikely to afford a clean answer to whether a particular histone mark drives a given output. Mutagenesis of the histone proteins themselves is also not as useful as one might initially think. First, higher eukaryotes have sequence variants of canonical histones, each with its own specialized function and each coded by several genes; for example, there are three histone H3 variants in humans (H3.1, 3.2, and 3.3) collectively encoded by 15 genes located on disparate genomic loci. In addition, many histone residues are modified in multiple ways. For example, lysine 27 on histone H3, H3K27, can be either trimethylated or acetylated, with the two marks correlated with opposing effects on transcription.¹⁶ Thus, mutating a histone residue such as H3K27 such that it cannot be modified will likely perturb multiple pathways. In the face of these challenges, new methods are desperately needed. Ideally, an epigenetic engineer would have the ability to control, with surgical precision, the pattern of histone marks at any given genomic locus. Doing so would allow causal effects to be explored in an unambiguous manner. Since this problem requires the manipulation of a biopolymer, i.e., the chromatin fiber, it seems natural that chemistry-driven tools should be part of the solution. In the following sections, we highlight nascent efforts to develop such methods.

Chemical Approaches for Biochemical Analysis of Histone PTMS

Before discussing ongoing endeavors to engineer chromatin in cells, it is worth briefly reviewing the remarkable advances that have been made in preparing modified chromatin in a test tube. Progress in this area has helped lay the foundation for at least some of the emerging cell-based initiatives. Broadly speaking, chemical approaches to the manufacture of modified chromatin fall into three main categories: cysteine modification, protein ligation, and un-natural amino acid mutagenesis (for a comprehensive review on the topic, see ref 24). Each of these methods allows for the site-specific introduction of PTMs (or analogues thereof) into histones, which can then be biochemically assembled into chromatin using established methods. Importantly, these strategies have complementary strengths and weaknesses, but together create a powerful toolbox for the introduction of PTMs into histones. The cysteine modification strategy, for example, offers a straightforward route to the introduction of several PTM types into proteins—in practice, a histone is mutated to have a unique cysteine at the position of interest and then converted into a PTM analogue through, for example, alkylation of the sulfhydryl group (Figure 2A). This approach has proven especially powerful for introducing analogues of lysine methylation into histones.²⁵ On the other hand, many histone marks are not accessible using this simple strategy, and it is poorly suited to the introduction of multiple different PTMs into a histone polypeptide. Protein ligation methods, by contrast, offer a great deal more flexibility in terms of the type and number of PTMs that can be installed into a histone.²⁶ Both total synthesis and semisynthesis schemes can be used, in each case exploiting traceless versions of native chemical ligation to assemble the target histone from recombinant and/or synthetic peptide segments (Figure 2B).²⁶ These strategies are especially well suited to installing PTMs into

the flexible N- or C-terminal tails of histones, where in fact the majority of marks reside, but are somewhat more cumbersome when it comes to the modifying residues in the middle of histones (although this is still possible). Unnatural amino acid (UAA) mutagenesis (Figure 2C), which in its most commonly used incarnation relies on the amber suppression strategy,²⁷ does not suffer from any sequence restrictions and, like cysteine-based bioconjugation, is operationally straightforward to perform, at least once the necessary orthogonal tRNA/tRNA synthetase pair for the UAA of interest has been developed. Moreover, and as discussed later, this methodology is well positioned for cell-based applications. However, like the other two strategies, the UAA mutagenesis approach is not without its limitations: it cannot currently be applied to every PTM type, and it is nontrivial to incorporate multiple marks into the same histone.²⁸

The availability of multiple protein engineering strategies for introducing PTMs means that synthetic routes can now be found for the majority of modified histone targets one might wish to study²⁴—these routes can, if necessary, involve the integrated use of these methods.^{29,30} Thus, as we write this piece, many dozens of modified histones have been generated using chemistry-driven methods. These reagents have been used to study various aspects of chromatin biochemistry,^{31,32} in many cases allowing mechanistic hypotheses generated from cell-based genomic data to be tested. For example, we have used semisynthesis to show that ubiquitylation of lysine 120 in histone H2B (H2B-Ub) directly activates the H3K79 methyltransferase, hDOT1L,^{33,34} a crosstalk that was previously suggested by genetic studies in yeast.³⁵ More recently, these biochemical strategies have evolved to the point where they can be used as hypothesis-generating tools. This notable development is largely due to the ability to perform unbiased screens using large libraries of either modified histone peptides or nucleosomes (Figure 3). The former have proven extremely useful for defining the binding preferences of isolated reader domains,³⁶ and for testing the specificity of anti-PTM antibodies,³⁷ which are routinely used in the epigenetics field. The manufacture of modified nucleosomes is significantly more involved than for short peptides, even with the robust protein chemistries described above. Using these reagents in a library format might, therefore, seem impractical were it not for two innovations, namely the development of a streamlined nucleosome assembly approach and the incorporation of identity barcodes into the nucleosomal DNA.³⁸ The former significantly shortens the time it takes to reconstitute nucleosomes from their component parts, whereas the latter allows next-generation sequencing (which has fueled the genomics field in recent years) to be used as a highly sensitive yet quantitative readout of biochemical assays. In other words, thousands of binding and/or enzyme activity measurements can be made in parallel using extremely small (picomole level) amounts of precious library material.³⁸

Chemical Approaches for Modifying Chromatin in Cells

Irrespective of whether the methods discussed above are used to test existing hypotheses or to generate new ones, a reductionist approach to studying chromatin biology is always employed. This is both a strength and a weakness—use of a molecularly defined system allows causality to be confidently assigned but does create the possibility of artifacts; i.e., key players might be missing from the reconstituted system. This situation presents somewhat of a conundrum since ideally one would want to validate the *in vitro* finding in

cells, yet the very reason that the reduction approach was taken in the first place is because it is exceedingly difficult to test molecular mechanisms as they relate to epigenetics in the cellular environment. It is this problem that has driven the work discussed below.

Unnatural Amino Acid Mutagenesis

Of the protein engineering strategies described above, arguably the UAA mutagenesis method offers the most obvious route to tailoring the structure of chromatin in cells. This is because the approach is based on the ribosomal incorporation of an UAA, typically in response to an amber (stop) codon introduced site-specifically into the gene of interest using standard molecular biology methods.³⁹ Key to the method is the development of a dedicated tRNA/tRNA-synthetase pair for the UAA of interest. Importantly, this pair must not cross-react with any of the endogenous amino acids or translation machinery in the cell; i.e., it must be orthogonal. Generally, this is achieved by importing a heterologous tRNA/tRNA-synthetase pair into the cell-type of interest—directed protein evolution schemes are typically employed to fine-tune both orthogonality and UAA selectivity. The cognate amino acid substrate for a tRNA/tRNA-synthetase pair largely dictates the structure of the UAA that it can be engineered to accept. For example, the tRNA^{Tyr}/tRNA pair from *Escherichia coli* can be engineered to use benzoylphenylalanine (Bpa),⁴⁰ whereas the tRNA^{Pyl}/tRNA pair from *Methanosaccina berkeri*, which normally incorporates the non-canonical amino acid pyrrolysine, can be engineered to accept various lysine derivatives, including *N*-ε-acetyl-lysine.⁴¹

UAA mutagenesis systems are now available for a range of important cell types,^{42–45} including mammalian cells,⁴⁶ allowing an ever-growing collection of biochemical and biophysical probes to be site-specifically introduced into proteins. A few recent examples serve to illustrate the unique power of this expanded genetic code for studying epigenetic processes (Figure 4A). The first of these involves use of a photo-cross-linking strategy to determine how chromatin structure changes during the cell cycle.⁴⁷ Specifically, Neumann and co-workers were interested in the role of a unique binding surface on the nucleosome disc, termed the acidic patch. Based on previous biochemical and structural studies,⁴⁸ it was known that this negatively charged surface could bind the positively charged tail of histone H4 from another nucleosome, and that this engagement favored chromatin compaction through a stapling-type mechanism. Incorporation of the photoreactive UAA, Bpa, into the acidic patch using amber suppression allowed the importance of this stapling interaction to be confirmed in the context of yeast cells. Moreover, by exploiting the temporal control conferred by use of the photo-cross-linker, the investigators found that the stapling interaction was most pronounced during mitosis. This set the stage for the discovery of a novel regulatory mechanism in which removal of a histone PTM that blocks the stapling interaction, acetylation of H4 on lysine 16 (H4K16ac), is stimulated by the presence of a second PTM, phosphorylation of H3 on serine 10 (H3S10ph), known to be present during the M-phase. While Bpa is certainly a powerful tool for identifying protein–protein interactions, the steric bulk of this UAA may make it unsuitable for some applications. Fortunately, other photoreactive UAAs are available. These include a diazirine-containing analogue of lysine, termed photolysine,⁴⁹ which generates a reactive carbene species upon UV irradiation. This photoreactive amino acid can be nonspecifically incorporated into

cellular proteins by simply including the UAA in the cell culture medium—the high lysine content of histones makes them well suited to this approach. Indeed, metabolic labeling with photolysine in combination with quantitative proteomic workflows was shown to be a powerful strategy for identifying histone-binding factors in mammalian cells. Notably, an amber suppression system for a caged version of photolysine (as well as photolysine bearing a crotonyl mark) was recently reported,⁵⁰ further expanding the utility of this system in a cellular context.

The prospect of directly introducing PTMs into cellular proteins using amber suppression is obviously very exciting when it comes to applications in the epigenetic area. As noted earlier, engineered tRNA/tRNA-synthetase pairs for several modified amino acids are now available. Recently, Chin and coworkers described the generation of engineered mammalian cells that allow for the improved incorporation of UAAs, including *N*- ϵ -acetyl-lysine, into recombinant proteins.⁵¹ By stably integrating engineered tRNA/tRNA-synthetase pairs into the host cell genome, as opposed to using transiently transfected cells, the authors were able to dramatically increase the efficiency of UAA incorporation. Impressively, this allowed the incorporation of *N*- ϵ -acetyl-lysine site-specifically at several positions in the H3 variant, H3.3 (Figure 4A). The mRNA-seq method was then used to assess the effect of increasing the levels H3.3 acetylation on transcription. Notably, the authors found that certain genes are selectively sensitive to the presence of specific acetylation marks—for example, the gene for the non-coding XIST RNA, which is required for X-chromosome inactivation, was activated by acetylation on lysine 56 on H3.3.⁵¹ This insight highlights the power of combining genomic tools based on next-generation sequencing with the ability to precisely manipulate chromatin structure, in this case through amber suppression.

Protein Fusion Strategies

Chemical biologists have developed highly versatile methods for manipulating cellular proteins based on the use of fused effector domains, i.e., analogous to GFP fusions. These domains can, for instance, be specifically modified with cell-permeable small molecules bearing a suitably reactive warhead linked to a biochemical or biophysical probe.⁵² Examples include the powerful SNAP and halo-tag systems that are widely used to covalently attach synthetic cargoes such as fluorophores to cytosolic proteins.^{53–55} Alternatively, these domains can mediate the recruitment of designated cellular factors in response to pharmacological or even optical inputs.^{56,57} Well-known approaches such as the rapamycin/FKBP/FRB three-hybrid system and the LOV domain-based photoactivation system permit protein activity to be controlled through a variety of strategies, including regulating cellular location and triggering degradation.^{58–60} These genetic fusion strategies are increasingly being applied to study chromatin, allowing the attachment of synthetic fluorophores compatible with super-resolution methods such as STORM,⁶¹ or temporal control over epigenetic processes through changes in cellular localization⁶² or recruitment of enzymatic activities.⁶³

Among the various options available for chemically modifying cellular proteins using fusion domains, the use of split inteins presents unique opportunities for directly incorporating epigenetic marks into chromatin. Inteins are autocatalytic proteins that spontaneously excise

themselves from a host protein in a traceless manner—a post-translational process termed protein splicing.⁶⁴ To the first approximation, inteins are promiscuous with respect to the host protein, a feature that has made them powerful tools in chemical biology.^{65,66} While most inteins are contiguous polypeptides embedded within a single protein, a small number of these domains are naturally fractured into two pieces.⁶⁵ These split inteins must first associate before protein trans-splicing (PTS) can occur—in essence they behave like single-turnover protein ligases.

The discovery of naturally split inteins that associate with low nanomolar affinity and then undergo extremely rapid PTS ($t_{1/2} < 1$ min) has opened the way to the efficient semisynthesis of proteins in living cells.⁶⁷ Indeed, our laboratory has recently exploited this capability to incorporate a variety of probes into cellular chromatin.⁶⁸ By genetically fusing one fragment of the intein pair (Int^N, 120 amino acids) to a histone and delivering the second fragment (Int^C, 35 amino acids) with a defined cargo, we were able to generate site-specifically labeled histones in the chromatin of live mammalian cells (Figure 4B). Note, the traceless nature of PTS means that the final labeled protein no longer contains the fused intein, a feature that differentiates the approach from other fusion strategies where the bulky domain remains attached post-labeling. Moreover, this attribute allowed implementation of a quencher-fluorophore strategy that ensured only the labeled histone product was fluorescent, i.e., no background.⁶⁸ The scope of this PTS strategy for chemically customizing chromatin was further illustrated by the successful installation of ubiquitin (among the biggest and most complex of the PTMs found on histones) into histone H2B at lysine 120 (H2B-Ub). In this case, the PTS reaction was performed in isolated nuclei rather than live cells. Nonetheless, the generation of semisynthetic H2B-Ub in native chromatin using PTS led to an increase in the levels H3K79 methylation, a result expected from prior biochemical studies. In a follow-up study, we then adapted this strategy to allow specific installation of a mutant version of ubiquitin into H2B, and in the process verify the importance of a surface epitope on the protein needed for stimulating the activity of the Dot1L methyl-transferase.⁶⁹ This example illustrates the true power of using chemical methods to engineer native chromatin: It is highly unlikely that the mutant version of ubiquitin would not have been incorporated into chromatin by the endogenous ligase enzymes, and thus the importance of ubiquitin epitope in Dot1L stimulation could not have been tested by simply expressing the mutant in mammalian cells. Thus, the chemical approach provided a unique solution to the problem.

Toward Locus-Specific Chromatin Tailoring

Approaches such as UAA mutagenesis and PTS open the door to customizing the covalent structure of histones with the flexibility of organic chemistry. However, in the examples published so far, little or no control was possible over where the modifications occurred in the genome—the modified histones were broadly distributed in the cellular chromatin.^{51,68} Ideally, one would want to direct the modifications to specific genomic loci such that their impact on processes such as gene expression could be evaluated in a more controlled fashion. Only a few years ago this would have seemed like a very tall order; however, we are fortunate to live in an era when targeted genome engineering has become a reality. Programmable DNA-binding proteins such as zinc-finger nucleases, transcription activator-

like effector nucleases (TALENs), and, most recently, CRISPR (clustered regularly interspaced short palindromic repeats)/Cas9 nuclease have revolutionized the business of genome editing.⁷⁰ All of these systems can be adapted to epigenome engineering by directly or indirectly localizing writer or eraser activities to specific loci (Figure 5A). For example, in a pioneering study, Pabo and co-workers demonstrated that fusing an H3K9 methyltransferase domain to a zinc-finger targeting vector allowed repression of specific genes *in vivo*.⁷¹ There have been many variations of this experiment since, allowing the levels of both DNA and histone marks to be manipulated through recruitment of appropriate writer/eraser and even reader activities to designated genomic loci.⁷²

The nuclease-deficient version of Cas9, dCas9, is rapidly becoming the platform of choice for epigenome engineering, due to the ease with which it can be programmed with appropriate guide-RNAs.⁷³ Indeed, this targeting system has proven to be remarkably amenable to integration with the various fusion domain technologies developed by chemical biologists in recent years, some of which were introduced earlier. This has, for example, allowed the development of conditional versions of dCas9 tools that can be turned on and off in response to small-molecule or optical inputs.⁷⁴ In addition, a variety of dCas9-fusion strategies have been introduced for genomic imaging applications based on the attachment of protein or small-molecule fluorophores.^{57,75} As a further illustration of the scope of this technology, we recently succeeded in delivering a variety of synthetic cargoes to genomic loci by combining dCas9 with the aforementioned PTS ligation approach.³⁰ In this case, a “one-pot” procedure was developed that allowed the attachment of a synthetic molecule to the protein in the cell culture media, followed by delivery of the resulting conjugate to live cells (Figure 5B). The approach was compatible with a variety of epigenetic probes, including molecules that bind to specific reader domains within epigenetic factors. Notably, this allowed the recruitment of endogenous epigenetic proteins and macromolecular effector complexes, each harboring the reader domain of interest, to specified genetic loci.

These are certainly heady times for chemical biologists, synthetic biologists, and bioengineers interested in manipulating eukaryotic genomes. It should be remembered that the CRISPR/Cas9 revolution is still in its infancy—it is hard to imagine that the current frenzy of activity in this area will slow down any time soon. The dCas9 toolbox, for example, is still rapidly expanding, with new applications of the system being published at a remarkable rate. With respect to the main theme of this Perspective, it will be interesting to see whether genomic targeting and protein engineering strategies can be further integrated such that precise chemical tailoring of chromatin proteins such as histones is possible in a locus-specific manner. The development of such second-generation “chemo-epigenetic” tools will certainly challenge the creative talents of chemical biologists. Nonetheless, we would argue that the potential payoff of such work is likely to be considerable, and so it seems likely that this area will continue to attract significant attention from the chemical biology community.

Conclusions

We have argued in this Perspective that there is a pressing need for new methods that allow us to move beyond the correlative-type data that currently dominate the field of epigenetics.

Establishing causal biochemical relationships in a cellular context is a necessary next step if we are to truly appreciate how epigenetic processes relate to gene expression programs and ultimately cellular phenotypes. Exactly how unique patterns of epigenetic marks are installed, spatially confined, and maintained across the genome remains very poorly understood, as does the role that these “states” play in tuning DNA transactions through the formation of fundamental genomic structures. We believe that addressing these questions requires the development of tools that allow native cellular chromatin to be manipulated with the same level of precision and synthetic flexibility that is already possible in the test tube. In particular, the ability to install native histone PTMs in cells, with single-nucleosome resolution, would open the way to a deeper understanding of how epigenetic features such as enhancers, promoters, and heterochromatin are generated and maintained. Coupling such methods with emerging single-cell sequencing and imaging capabilities would allow the precise cellular responses to epigenetic stimuli to be studied in unprecedented detail. While this is a daunting undertaking, the first tentative steps in this direction are now being made, and we hope that these early efforts will inspire and encourage others to work on this fundamental biological problem.

Acknowledgments

We would like to thank members of the Muir lab for many helpful discussions. Some of the work discussed in this Perspective was conducted in the Muir lab and was supported by U.S. National Institutes of Health grants R37-GM086868, R01-GM107047, and P01-CA196539. The David lab is supported by the CCSG core grant P30 CA008748 and the Josie Robertson Young Investigator Award.

References

1. Sewitz SA, Fahmi Z, Lipkow K. *Curr Opin Struct Biol.* 2017; 42:162–168. [PubMed: 28284913]
2. Li G, Reinberg D. *Curr Opin Genet Dev.* 2011; 21:175–186. [PubMed: 21342762]
3. Kadonaga JT. *Cell.* 2004; 116:247–257. [PubMed: 14744435]
4. Bartholomew B. *Annu Rev Biochem.* 2014; 83:671–696. [PubMed: 24606138]
5. Shafik A, Schumann U, Evers M, Sibbritt T, Preiss T. *Biochim Biophys Acta, Gene Regul Mech.* 2016; 1859:59–70.
6. Ambrosi C, Manzo M, Baubec TJ. *Mol Biol.* 2017; 429:1459.
7. Smith E, Shilatfard A. *Mol Cell.* 2010; 40:689–701. [PubMed: 21145479]
8. Bowman GD, Poirier MG. *Chem Rev.* 2015; 115:2274–2295. [PubMed: 25424540]
9. Furey TS. *Nat Rev Genet.* 2012; 13:840–852. [PubMed: 23090257]
10. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. *Nat Methods.* 2013; 10:1213–1218. [PubMed: 24097267]
11. Han Y, Guo S, Muegge K, Zhang W, Zhou B. *Bioinf Biol Insights.* 2015; 9:29–46.
12. Dekker J, Mirny L. *Cell.* 2016; 164:1110–1121. [PubMed: 26967279]
13. Mahony S, Pugh BF. *Crit Rev Biochem Mol Biol.* 2015; 50:269–283. [PubMed: 26038153]
14. The ENCODE (ENCyclopedia Of DNA Elements) Project. The ENCODE Project Consortium. *Science.* 2004; 306:636–640. [PubMed: 15499007]
15. Soshnev AA, Josefowicz SZ, Allis CD. *Mol Cell.* 2016; 62:681–694. [PubMed: 27259201]
16. Bonasio R, Tu S, Reinberg D. *Science.* 2010; 330:612–616. [PubMed: 21030644]
17. Huang H, Lin S, Garcia BA, Zhao Y. *Chem Rev.* 2015; 115:2376–2418. [PubMed: 25688442]
18. Venkatesh S, Workman JL. *Nat Rev Mol Cell Biol.* 2015; 16:178–189. [PubMed: 25650798]
19. Musselman CA, Lalonde ME, Cote J, Kutateladze TG. *Nat Struct Mol Biol.* 2012; 19:1218–1227. [PubMed: 23211769]

20. Patel DJ, Wang Z. *Annu Rev Biochem.* 2013; 82:81–118. [PubMed: 23642229]
21. Zhang T, Cooper S, Brockdorff N. *EMBO Rep.* 2015; 16:1467–1481. [PubMed: 26474904]
22. Su Z, Denu JM. *ACS Chem Biol.* 2016; 11:564–574. [PubMed: 26675328]
23. James LI, Frye SV. *Curr Opin Chem Biol.* 2016; 33:135–141. [PubMed: 27348158]
24. Muller MM, Muir TW. *Chem Rev.* 2015; 115:2296–2349. [PubMed: 25330018]
25. Simon MD, Shokat KM. *Methods Enzymol.* 2012; 512:57–69. [PubMed: 22910202]
26. Holt M, Muir T. *Annu Rev Biochem.* 2015; 84:265–290. [PubMed: 25784050]
27. Lang K, Chin JW. *Chem Rev.* 2014; 114:4764–4806. [PubMed: 24655057]
28. Lammers C, Hahn LE, Neumann H. *ChemBioChem.* 2014; 15:1800–1804. [PubMed: 24890611]
29. Virdee S, et al. *J Am Chem Soc.* 2011; 133:10708–10711. [PubMed: 21710965]
30. Liszczak GP, et al. *Proc Natl Acad Sci U S A.* 2017; 114:681–686. [PubMed: 28069948]
31. Neumann H, et al. *Mol Cell.* 2009; 36:153–163. [PubMed: 19818718]
32. Dhall A, et al. *J Biol Chem.* 2014; 289:33827–33837. [PubMed: 25294883]
33. McGinty RK, Kim J, Chatterjee C, Roeder RG, Muir TW. *Nature.* 2008; 453:812–816. [PubMed: 18449190]
34. Zhou L, et al. *Nat Commun.* 2016; 7:10589. [PubMed: 26830124]
35. Ng HH, Xu RM, Zhang Y, Struhl K. *J Biol Chem.* 2002; 277:34655–34657. [PubMed: 12167634]
36. Rothbart SB, Krajewski K, Strahl BD, Fuchs SM. *Methods Enzymol.* 2012; 512:107–135. [PubMed: 22910205]
37. Nishikori S, et al. *J Mol Biol.* 2012; 424:391–399. [PubMed: 23041298]
38. Nguyen UT, et al. *Nat Methods.* 2014; 11:834–840. [PubMed: 24997861]
39. Liu CC, Schultz PG. *Annu Rev Biochem.* 2010; 79:413–444. [PubMed: 20307192]
40. Chin JW, Martin AB, King DS, Wang L, Schultz PG. *Proc Natl Acad Sci U S A.* 2002; 99:11020–11024. [PubMed: 12154230]
41. Neumann H, Peak-Chew SY, Chin JW. *Nat Chem Biol.* 2008; 4:232–234. [PubMed: 18278036]
42. Yanagisawa T, et al. *ChemBioChem.* 2014; 15:1830–1838. [PubMed: 25067793]
43. Wals K, Ovaa H. *Front Chem.* 2014; 2:15. [PubMed: 24790983]
44. Hancock SM, Uprety R, Deiters A, Chin JW. *J Am Chem Soc.* 2010; 132:14819–14824. [PubMed: 20925334]
45. Bianco A, Townsley FM, Greiss S, Lang K, Chin JW. *Nat Chem Biol.* 2012; 8:748–750. [PubMed: 22864544]
46. Gautier A, et al. *J Am Chem Soc.* 2010; 132:4086–4088. [PubMed: 20218600]
47. Wilkins BJ, et al. *Science.* 2014; 343:77–80. [PubMed: 24385627]
48. Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ. *Nature.* 1997; 389:251–260. [PubMed: 9305837]
49. Yang T, Li XM, Bao X, Fung YM, Li XD. *Nat Chem Biol.* 2016; 12:70–72. [PubMed: 26689789]
50. Xie X, et al. *J Am Chem Soc.* 2017; 139:6522–6525. [PubMed: 28459554]
51. Elsasser SJ, Ernst RJ, Walker OS, Chin JW. *Nat Methods.* 2016; 13:158–164. [PubMed: 26727110]
52. Hinner MJ, Johnsson K. *Curr Opin Biotechnol.* 2010; 21:766–776. [PubMed: 21030243]
53. Ishitsuka Y, et al. *J Phys Chem B.* 2015; 119:6611–6619. [PubMed: 25978145]
54. Dunleavy EM, Almouzni G, Karpen GH. *Nucleus.* 2011; 2:146–157. [PubMed: 21738837]
55. Murrey HE, et al. *J Am Chem Soc.* 2015; 137:11461–11475. [PubMed: 26270632]
56. Umeda N, Ueno T, Pohlmeyer C, Nagano T, Inoue T. *J Am Chem Soc.* 2011; 133:12–14. [PubMed: 21142151]
57. Schmidt JC, Zaug AJ, Cech TR. *Cell.* 2016; 166:1188–1197. [PubMed: 27523609]
58. Winter GE, et al. *Science.* 2015; 348:1376–1381. [PubMed: 25999370]
59. Cho U, et al. *PLoS One.* 2013; 8:e72393. [PubMed: 23991108]
60. Pratt MR, Schwartz EC, Muir TW. *Proc Natl Acad Sci U S A.* 2007; 104:11209–11214. [PubMed: 17563385]
61. Wombacher R, et al. *Nat Methods.* 2010; 7:717–719. [PubMed: 20693998]

62. Yumerefendi H, et al. *Nat Chem Biol.* 2016; 12:399–401. [PubMed: 27089030]
63. Hathaway NA, et al. *Cell.* 2012; 149:1447–1460. [PubMed: 22704655]
64. Vila-Perello M, Muir TW. *Cell.* 2010; 143:191–200. [PubMed: 20946979]
65. Shah NH, Muir TW. *Chem Sci.* 2014; 5:446–461. [PubMed: 24634716]
66. Giriat I, Muir TW. *J Am Chem Soc.* 2003; 125:7180–7181. [PubMed: 12797783]
67. Borra R, Dong D, Elnagar AY, Woldemariam GA, Camarero JA. *J Am Chem Soc.* 2012; 134:6344–6353. [PubMed: 22404648]
68. David Y, Vila-Perello M, Verma S, Muir TW. *Nat Chem.* 2015; 7:394–402. [PubMed: 25901817]
69. Holt MT, et al. *Proc Natl Acad Sci U S A.* 2015; 112:10365–10370. [PubMed: 26240340]
70. Kungulovski G, Jeltsch A. *Trends Genet.* 2016; 32:101–113. [PubMed: 26732754]
71. Snowden AW, Gregory PD, Case CC, Pabo CO. *Curr Biol.* 2002; 12:2159–2166. [PubMed: 12498693]
72. Laufer BI, Singh SM. *Epigenet Chromatin.* 2015; 8:34.
73. Mohr SE, et al. *FEBS J.* 2016; 283:3232–3238. [PubMed: 27276584]
74. Wang H, La Russa M, Qi LS. *Annu Rev Biochem.* 2016; 85:227–264. [PubMed: 27145843]
75. Leonetti MD, Sekine S, Kamiyama D, Weissman JS, Huang B. *Proc Natl Acad Sci U S A.* 2016; 113:E3501–3508. [PubMed: 27274053]

Figure 1. Packaging of genomic DNA into chromatin. Schematic illustrating some of the epigenetic modifications associated with euchromatic (“active”) and heterochromatic (“silenced”) states.

Figure 2.

Chemical methods for the synthesis of site-specifically modified histones *in vitro*. (A) Synthesis of Kme3 (left), Kac (middle), and Rme/Rme2a analogues from cysteine-containing histones. (B) Expressed protein ligation: synthesis of site-specifically modified histone through traceless native chemical ligation of recombinant and synthetic fragments. (C) Amber suppression: ribosomal incorporation of an unnatural amino acid (UAA) into a protein of choice using an orthogonal tRNA/tRNA synthetase pair. In panels B and C, green star corresponds to PTM.

Figure 3.

High-throughput methods for studying chromatin biochemistry. (A) Microarrays of histone-derived peptides carrying specific PTMs (stars) are incubated with a protein of interest (POI), followed by immunodetection. For clarity, only two rows of peptides are illustrated. (B) Modified histones are assembled into a DNA barcoded nucleosome library. This library is then used to profile the activity of a chromatin effector, employing next-generation DNA sequencing (Next Gen Seq) as a highly sensitive yet quantitative readout.

Figure 4.

Methods for the synthesis of site-specifically modified histones in cells. (A) Application of amber suppression to chromatin (Figure 1C). The requisite tRNA/tRNA synthetase pair is heterologously expressed in the cell type of interest. Addition of the unnatural amino acid of interest (star) leads to the generation of the site-specifically modified protein of interest, e.g., a histone. (B) Using an ultrafast split intein pair, one attached to a histone and genetically fused and the other exogenously added, to perform histone semisynthesis on native chromatin. Star: chemical modification.

Figure 5. Locus-specific chromatin modification. (A) Expression of a targeting domain (TALEN, ZNF, dCas9) fused to chromatin effector domain to induce a local histone modification. (B) One-pot protein synthesis and delivery. Step 1: dCas9 is chemically modified by protein trans-splicing. Step 2: Protein delivery into cells. Star: chemical modification.