

Manuscript Details

Manuscript number	COCHBI_2017_140_R1
Title	Synthetic post-translational modification of histones
Short title	Synthetic post-translational modification of histones
Article type	Review article

Abstract

Chromatin is the physiological template of genetic information in all eukaryotic cells, a highly organised complex of DNA and histone proteins central in regulating gene expression and genome organisation. A multitude of histone post-translational modifications (PTMs) have been discovered, providing a glance into the complex interplay of these epigenetic marks in cellular processes. In the last decade, synthetic and chemical biology techniques have emerged to study these modifications, including genetic code expansion, histone semisynthesis and post-translational chemical mutagenesis. These methods allow for the creation of histones carrying synthetic modifications which can in turn be assembled into designer nucleosomes. Their application in in vitro and in vivo studies is now beginning to have an important impact on chromatin biology. Efforts towards introducing multiple labile modifications in histones as well as expanding their use in cellular biology promise new powerful tools to study epigenetics.

Keywords Synthetic histones ; synthetic nucleosomes ; designer chromatin ; chemical modification ; histone post-translational modification ; genetic code expansion ; amber suppression ; protein semisynthesis ; native chemical ligation ; intein trans splicing ; split inteins ; chemical mutagenesis ; dehydroalanine ; tag-and-modify ; in vivo histone modification ; immobilised nucleosome pull-down ; mass spectrometry-based proteomics

Corresponding Author Benjamin G. Davis

Order of Authors Simon Nadal, Ritu Raj, Shabaz Mohammed, Benjamin G. Davis

Submission Files Included in this PDF

File Name [File Type]

COCHBI_2017_140_revised_highlighted.pdf [Revised Manuscript with Changes Marked]

Highlights.docx [Highlights]

COCHBI_2017_140_revised.docx [Manuscript File]

To view all the submission files, including those not included in the PDF, click on the manuscript title on your EVISE Homepage, then click 'Download zip file'.

Synthetic post-translational modification of histones

Simon Nadal¹, Ritu Raj¹, Shabaz Mohammed¹ and Benjamin G. Davis^{1*}

Current Opinion in Chemical Biology

Addresses

¹Department of Chemistry, University of Oxford, Chemistry Research Laboratory, Mansfield Road, Oxford OX1 3TA, UK

S.N. simon.nadal@chem.ox.ac.uk ; ORCID: 0000-0002-0748-0284

R.R. ritu.raj@chem.ox.ac.uk ; ORCID: 0000-0002-7623-9128

S.M. shabaz.mohammed@chem.ox.ac.uk ; ORCID: 0000-0003-2640-9560

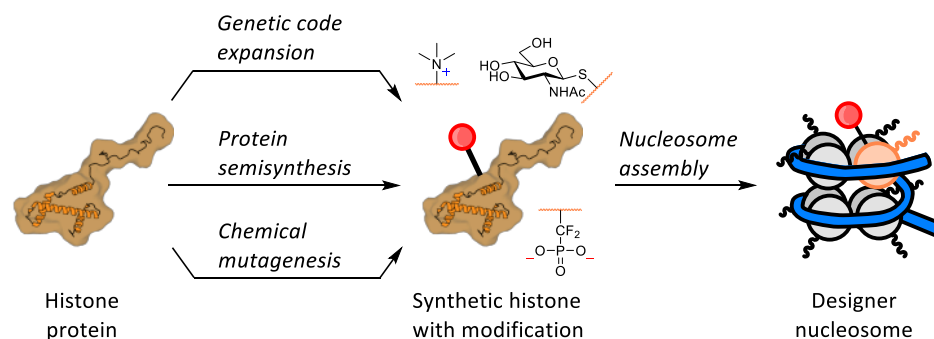
B.G.D. ben.davis@chem.ox.ac.uk ; ORCID: 0000-0002-5056-407X

* Corresponding author: Davis, Benjamin G. (ben.davis@chem.ox.ac.uk).

Abstract

Chromatin is the physiological template of genetic information in all eukaryotic cells, a highly organised complex of DNA and histone proteins central in regulating gene expression and genome organisation. A multitude of histone post-translational modifications (PTMs) have been discovered, providing a glance into the complex interplay of these epigenetic marks in cellular processes. In the last decade, synthetic and chemical biology techniques have emerged to study these modifications, including genetic code expansion, histone semisynthesis and post-translational chemical mutagenesis. These methods allow for the creation of histones carrying synthetic modifications which can in turn be assembled into designer nucleosomes. Their application in *in vitro* and *in vivo* studies is now beginning to have an important impact on chromatin biology. Efforts towards introducing multiple labile modifications in histones as well as expanding their use in cellular biology promise new powerful tools to study epigenetics.

Graphical abstract



Highlights

- Histones with specific PTMs can be synthesised and assembled in designer nucleosomes
- Codon suppression, semisynthesis or mutagenesis allow controlled insertion of PTMs
- Designer nucleosomes can probe stability and/or PTM specific interaction partners
- *In vivo* techniques show potential for locus-specific incorporation of PTM

Tags

Synthetic histones ; synthetic nucleosomes ; designer chromatin ; chemical modification ; histone post-translational modification ; genetic code expansion ; amber suppression ; protein semisynthesis ; native chemical ligation ; intein *trans* splicing ; split inteins ; chemical mutagenesis ; dehydroalanine ; tag-and-modify ; *in vivo* histone modification ; immobilised nucleosome pull-down ; mass spectrometry-based proteomics

Background and Motivation

Epigenetic mechanisms allow our cells to inherit changes in their gene expression and phenotype without changing the underlying DNA sequence. Epigenetic processes can happen at several layers of the nuclear organisation of the genome, from the 3D territorial positioning of interphase chromosomes in the nucleus down to the atomic level with histone post-translational modifications (PTMs) and nucleobase modifications (Figure 1.a in blue). [1,2] These epigenetic layers of regulation influence each other, thus understanding the role of histone post-translational modifications in this framework has been central to deciphering chromatin biology. [3,4]

Histone PTMs have a number of implications in regulatory tasks: specific modifications such as H3K14ac lead to transcriptional initiation [3], others such as H3K4me3 recruit specific transcription factors that have the required reader domain (i.e. plant homeodomain - PHD) [5]. Investigating histone PTMs poses many challenges: a plethora of PTMs have been discovered (Figure 1b), several distinct histone proteoforms are found in a single cell and histone tails can have multiple modifications in close proximity. Therefore, to study histone PTM access to large quantities of homogeneously modified histones (with site- and group- selectivity) is required – this is a difficult goal to attain with enzymatic techniques [6,7]. In order to address this ambitious challenge, synthetic and chemical biology techniques to generate ‘designer’ nucleosomes containing histones with precisely inserted modifications have emerged in the last decade [8-18].

Here, we review recent developments in chemical chromatin research; starting with identification processes of a given histone post-translational modification, moving to a comparison of methods for generating synthetic nucleosome constructs carrying the modification of interest (focusing on post-translational approaches) and closing with their use in the investigation of biological processes.

Searching for histone modifications using tandem mass spectrometry

Cellular chromatin consists of multiple nucleosome forms, with distinct histone proteoforms that can carry modifications in sometimes very low stoichiometry.[19] This large range of population distribution poses an analytical challenge for mass spectrometry (MS)-based proteomics, since both abundant marks and rare modifications have to be reliably identified – a prerequisite for knowing which modification to study with synthetic histones.

In a classical approach, when the sequence of histones has been previously well characterised, a bottom-up approach as depicted in Figure 1.c is chosen. Proteins of interest are digested with a

proteolytic enzyme such as trypsin and characterisation is performed on the subsequent pool of peptides. By bioinformatic comparison of the measured spectra with proteomic databases, peptides and their modifications are identified. In well-characterised samples, new modifications can be identified (unknown mass shift Δ mass on a given peptide). Accurate mass and peptide sequencing allow an empirical formula as well as chemical properties to be obtained – ultimately leading to the PTM structure. [20] Building on this well-established schematised workflow, broader application of proteomic methods in the analysis of histone modifications now include top-down proteomics, *de novo* identification of PTMs or quantitative experiments such as label-free quantification, multiple reaction monitoring (MRM) or isotopic tags. [19]

Mass spectrometry-based techniques have been useful in testing the efficacy of established chemical biology and chemical modification techniques and conversely these can enable mass spectrometry-based data acquisition. By probing the limits of the deleterious effects of cellularly-incorporated tags used in covalent retrieval methods (e.g. azidohomoalanine [21,22]) in histones on cell-cycle progression [23] their efficient use can be planned. Libraries of histone peptides carrying diverse post-translational modifications have allowed improved calibration of MS detection efficiencies and have allowed the effect upon wide-scale quantification experiments to be re-examined, leading in one example to a different biological conclusion. [24] Moreover, the development of selective enrichment techniques for modified peptides, such as TiO₂-mediated affinity chromatography for phosphopeptides further supported the identification of rare histone PTM in large proteomic screens.[25] These examples illustrate that both mass spectrometry-based proteomics and chemical biology techniques are not isolated tools used in histone PTM research, but are now progressing hand-in-hand.

The chemical biology toolkit

Once a modification has been identified and validated, several lines of investigation can be followed from biochemical assays to bioinformatic correlation analysis. One approach is to generate synthetic histones and assemble designer nucleosomes. Three main chemical biology strategies have been developed for this purpose that operate through a judicious mix of synthetic, co- and post-translational strategies: genetic code expansion, protein semi-synthesis and chemical modification / mutagenesis.

Genetic code expansion

Non-sense suppression mutagenesis is a powerful technology for introducing single unnatural amino acids (uAA) into cellular proteins. A stop codon, usually the constitutively rare amber stop codon (TAG), is re-attributed to the unnatural amino acid by incorporating both a cognate tRNA and an engineered aminoacyl tRNA synthetase (o-aaRS) into the host cell. The engineered synthetase can recognise the tRNA carrying the reattributed codon and load it with the uAA. Both need to be orthogonal to the host cell biosynthetic machinery, which ideally necessitates that the tRNA is not accepted by any constitutive synthetase and that the engineered synthetase only loads the orthogonal tRNA with the uAA. The cellular biosynthesis machinery is then able to incorporate the uAA for the amber codon into proteins being synthesised (Figure 2.a.). [26,27]

Engineering an aaRS-tRNA pair is thus a fine balancing act between creating orthogonality to the host cell and, at the same time, keeping the loaded tRNA structurally sufficiently close to constitutive tRNAs in the host cells so that the biosynthetic machinery (i.e. the ribosome) can recognise and use it. Often, aaRS-tRNA pairs from archaeal origins fulfil these criteria when used in prokaryotic expression systems. For the protein expression workhorse *Escherichia coli* orthogonal pairs based on PylRS-tRNA_{CUA}, TyrRS-tRNA_{CUA}, LeuRS-tRNA_{CUA} and TrpRS-tRNA_{CCA} are now routinely used.[28] Through engineering of corresponding synthetases, more than 160 structural analogues of amino acids bearing unnatural functional groups have been incorporated. At the same time the chemical diversity that can be reached by genetic code expansion remains dependent on the substrate recognition of the evolved synthetase and the ribosome, therefore the resulting chemical diversity typically remains structurally related to the original amino acids processed by those synthetases.[29]

Despite these limits, the PylRS-tRNA_{CUA} has been successfully used to site-specifically incorporate various naturally occurring PTMs into histones, in particular lysine PTM variant such as acyl-lysines (acetyl-[30], propionyl-[31], butyryl-[31] and crotonyllysine[32], 2-hydroxyisobutyryl-lysine [33], with special protocols for H4 [34]) or methylated lysines (methyl-Lys[35], dimethyl-Lys[36]), allowing the *in cellulo* and even the *in vivo* study of the effects of these modifications [37]. Moreover, precursors such as photoreactive lysine analogues (benzophenylalanine [38], photolysine [39], photocrotonyl-lysine [40], caged photolysine [41]) could also be incorporated, allowing their use in combination with post-translational chemical alteration (cross-linking, uncaging) and corresponding quantitative proteomics workflows leading to, for example, the identification of histone binding-factors in mammalian cells [39]. Recent results show that even more strategically challenging PTMs of other amino acids, such as phosphothreonine [42] and phosphoserine [43,44], can be incorporated into histones through the adaptation of other tRNA-synthetase pairs.

Current developments in genetic code expansion broadly diverge into two main directions of investigation, with potentially a strong impact on the study of histone PTMs. The first is to incorporate *in vivo* systems, described later in this review. The second aims at overcoming current limitations of synthetase-tRNA pairs. Screening technologies for the evolution of orthogonal aaRS have been designed [42,45] and biosynthetic re-engineering techniques of cells proposed, including deletion of termination factors and engineered ribosomes that recognise the amber codon with higher efficiency [28]. A particularly interesting innovation is the development of ribosomes recognising quadruplet codon (Figure 2a, right), allowing expansion beyond the restricted number of triplet codons and therefore potentially greater flexibility for the insertion of multiple uAA [46].

Protein semisynthesis

The discovery of native chemical ligation (NCL) [47] enabled a powerful unification of the world of shorter peptide synthesis with that of recombinant protein expression. [15,16] This can employ a polypeptide carrying a C-terminal thioester **1** with a recombinant protein/peptide containing an N-terminal cysteine **2** to generate a semisynthetic protein construct **4**, as described in Figure 2b. Suitable polypeptides carrying a C-terminal thioester can be generated by solid-phase peptide synthesis (SPPS) using specially designed safety-catch resins with linkers such as 3-thiopropionic acid or N-acylbenzimidazole.[6] Another possibility, referred to as expressed protein ligation (EPL), is to recombinantly express a construct of the protein of interest fused on its C-terminus to an intein domain. Inteins are proteins able to (self)splice themselves out of the fused construct to the target protein using a catalytic cysteine (or selenocysteine[48]) residue (*cis*-splicing). If this autoprocessing mechanism is performed in presence of intercepting thiols, then target proteins with a C-terminal thioester are generated.[6,16] NCL strategies have been used for the creation of synthetic histones for more than 15 years [8,49,50] and several different types of modifications have been incorporated, including pHis [51]. It has been the method of choice for histone ubiquitylation [52], a primary model to study *in vitro* protein-protein ligation chemistry. Ligation techniques including photo-cleavable thiol auxiliary **5** [52] and delta-mercaptolysine uAA **6** [53] have been subsequently developed to enable the traceless incorporation of whole protein modifications into nucleosomes (Figure 2.c).

One of the great strength of protein semisynthesis is the possibility to incorporate diverse modifications on several sites of the peptide, such as multiple glycosylations [54] or fully modified histone tails [55]. Elaborated schemes using several ligation steps can be used for a total chemical synthesis of histones; this allows for control over each residue in the protein chain. [55-59] In a

biological context, a large number of histone PTMs exert their full effect only in a concerted manner, for example by cooperative recruitment of transcription factors.[4] Semisynthetic strategies are thus particularly valuable to assess the multi-site PTM crosstalk that might constitute a so-called Histone Code.

The discovery of split inteins has further popularised protein semisynthesis, facilitating the assembly of protein constructs. The Muir group has recently published a technology based on ultrafast intein *trans* splicing in which the *N*-terminal part of the intein, Int^N, is fused to the *C*-terminus of the histone of interest and another protein construct containing the *C*-terminal part of the intein, Int^C, with a polypeptide tail carrying modifications is added. Upon association of Int^N and Int^C, the splicing reaction is triggered and the histone ligated with the polypeptide tail, creating a process that can be applied to *in cellulo* studies.[60] Further developments have seen improved inteins without extein dependence, allowing a splicing efficiency essentially independent of the construct context.[61,62] However, these splicing reactions are not yet quantitative and some drawbacks of inteins remain to be addressed, such the hydrolysis or thiolysis of inteins leading to *N*- and *C*-cleavage products. [61,63]

Both NCL and intein-based ligation strategies rely on solid-phase peptide synthesis (SPPS) to generate synthetic modifications or PTM-modified peptides and are thus subject to known SPPS limitations. These include the β -elimination of phosphorylated sites under basic conditions [64] and potentially extensive epimerisation of slow-coupling amino acids, such as glycosylated building blocks [65].

Chemical Mutagenesis

Altering protein residues by performing selective chemical protein modification reactions can allow at its extreme the creation of amino acids via that modification [66-68], a concept termed chemical mutagenesis. [68-70] Cysteine stands out among the proteinogenic amino acids by virtue of its unique chemistry, and its accessible pKa allowing generation of a more powerfully nucleophilic thiolate anion. [71] Moreover, its abundance in proteins (especially in histones) is low; in humans only H3 contains a cysteine (Cys110) that can be mutated to alanine without loss-of-function. [72] As such, use of cysteines is currently a central approach in the chemical modification of histones. In this section, we attempt to illustrate, using synthetic histone PTMs as examples, innovations in the field of chemical modification and how they have become more elaborate and effective as mimics (even approaching the goal of directly installing fully-native PTMs / chemical mutagenesis) over time.

Early modification studies targeted cysteine as a nucleophile by reacting it with haloalkyl reagent **8** to generate histones with methyl-lysine 'thialysine' analogues **9** (Figure 2.d). [73] Subsequently, reagent **10** for the generation of methyl-arginine analogues via conjugate addition was suggested. [74] Strategically similar cysteine-based chemistry for acetyl-lysine analogues **13** is more challenging; strain-promoted substitution reaction yielded the first promising results [75] and thiol-ene reactions using **12** (Figure 2d) prove reasonably efficient. [76] Bis-alkylation and elimination to dehydroalanine (Dha) allows the conversion of nucleophilic cysteinyl site to electrophilic Dha and has been exploited [77] through the addition of thiol nucleophiles to generate various PTM analogues, including thialysine variants, in histones. [72] Nevertheless, computational studies suggest that the thioether linkages in all such thialysines introduces structural distortions that may impact on their recognition (Box in Figure 2d-e). [78]

The generation of constitutionally native, modified lysine (and other) residues (chemical mutagenesis) via chemical modification could be envisaged through C–C bond formation. [69,70] Appropriate biorthogonal C_{sp^3} - C_{sp^3} bond formation reactions have recently been demonstrated using C^\bullet radical additions to Dha (Figure 2.e). [79,80] The procedure is performed in two steps: the first being the elimination (e.g. via bis-alkylation of cysteine) to Dha. Two main reagents for the formation of Dha on proteins are used: DBHDA [77] and MDBP, [81,82] allowing varied application to different target proteins and even regioselective modulation of elimination sites by reagent type (unpublished observations). The addition of a C-radical to Dha in the second step can be mediated by use of either $NaBH_4$ [79,83] or an elementary metal powder [79,80,83] with suitable radical precursors. Dha can also be generated from phosphoserine [80], phenylselenocysteine [84] or alkylselenocysteine [85] residues that can be incorporated by genetic code expansion, allowing useful flexibility in installation strategies and potentially multiplexing of epigenetic marks. Moreover, Dha formation on proteins containing selenocysteine can be performed site-selectively by exploiting differences in pK_a . [48]

Such post-translational chemical mutagenesis is a quite direct, economical and accessible method to synthetically incorporate constitutionally native PTMs into histones, under conditions that are sufficiently facile to be performed even at the histone sub-complex or nucleosome level. However, current mutagenesis protocols come at the cost of typically mixed configuration (D/L) at the α -carbon and the reaction currently does not yet efficiently allow for bond formation to aromatic residues. Nevertheless, promising approaches in peptidic substrates [86,87] suggest possible avenues in proteins.

Methodological diversity covers most PTMs

With more than 150 histone PTMs reported to date, techniques for generation of chemically-modified histones not only need to cover the broad chemical diversity in modifications, but also allow for a possible experimental high-throughput. Chemical mutagenesis experiments are based on a 'tag-and-modify' approach [88], where the addition reaction to a 'tag' such as Dha in a recombinantly expressed protein conceptually allows for subsequent combinatorial expansion via late-stage, divergent protein modification chemistry and hence for the incorporation of a broad range of functional groups. In strategies based on genetic code expansion for the incorporation of uAAs that demand development of tailored amino-acyl RNA synthetases only limited structural plasticity may be achieved, typically within quite confined limits. In more labour-intensive protein semisynthesis schemes requiring synthetic expertise, it can prove tough to match this broad accessibility and throughput despite the recent development of ingenious methods for the encoding (e.g. via DNA) of nucleosome libraries.[89] Yet both methods allow the circumvention of the current limitations encountered by chemical mutagenesis schemes, thereby increasing the chemical diversity of modifications that can be included in synthetic nucleosomes. As such, each method complements the other: non-sense codon suppression strategies are particularly useful for efficient incorporation of certain aromatic and lysine-type residues (often with ready transfer to cellular settings), protein-semisynthesis schemes allow for the incorporation of several modifications in the a single histone (an important feature when testing Histone Code hypotheses and/or recruitment of interaction partners based on modifications acting in concert) and chemical mutagenesis allows ready late stage diversification in a structurally broad manner (typically *in vitro*). In this spirit, suggestions for methods to introduce challenging modifications are presented in Table 1.

Whereas a range of techniques to chemically introduce post-translational modifications on histones have been described here separately, combinations of these three approaches are therefore most likely the way to achieve the diverse goals of chromatin chemical biology, including access in relevant contexts to many and large quantities of specifically and selectively modified histones and nucleosomes.

Application of chemical biology tools for understanding chromatin biology

The strikingly diverse chemical space of PTMs modulates the structure and function of histone proteins in a spatiotemporal manner, and thereby plays a fundamental role in chromatin regulation. This is governed through two different yet often coupled mechanisms.[3,4] The first involves the

modulation of nucleosome (and, thus, chromatin) structure due to the perturbation of histone-fold or histone-DNA interactions upon histone modification.[90] In the second mechanism, histone PTMs acts as molecular beacons for specific recruitment of 'reader' or 'effector' proteins to mediate downstream functions.[5,91] Earlier chromatin biology studies relied on the use of native nucleosome and chromatin templates, bearing undefined mixtures of histone marks, thus limiting the precise understanding of distinct histone marks in epigenetic regulation. The maturation of various chemical toolkits leading to the ready synthesis of defined histone proteins and designer nucleosomes, in conjugation with various biochemical and biophysical studies, has allowed us to overcome that limitation by being able to establish initial causal relationships between histone marks and chromatin dynamics and function.

Probing the modulation of chromatin structure

Charge-modulating histone PTMs such as acetylation, phosphorylation, and ADP-ribosylation can potentially alter electrostatic histone-DNA interactions, resulting in modulation of nucleosome structure and stability.[90] Indeed, acetylation of lysine residues is generally believed to inhibit stabilizing interactions between lysines and nucleosomal DNA leading to a 'relaxed' chromatin state.[92] A recent study using synthetic acetylated nucleosome revealed that modification of H3-Lys64 residue lowers the stability of the nucleosome due to reduced histone-DNA interaction.[93] The effect of histone phosphorylation on nucleosome structure has also been analysed; phosphorylation of H3-Thr118 leads to apparent formation of two suggested non-canonical nucleosome structures – so-called 'nucleosome duplex' and 'altosome complex' – whose exact physiological roles are yet to be determined [94] Elegant use of biomolecular NMR with synthetic phosphorylated H3 variants showed that the dynamics of the H3 tail increase upon charge-modulation, altering transient contacts to the DNA and enhancing its general modifiability.[95]

Similar to these charge-modulating histone PTMs, neutral modifications may also regulate chromatin architecture (Figure 3.a). Synthetic mimicking of GlcNAcylation of H2A-Thr101 appears to disrupt histone-fold interactions leading to destabilized nucleosome structure.[96] Acetylation and ubiquitination of H4-Lys91 regulates nucleosome structure by perturbing the dimer-tetramer interface of histone octamer [97,98] primarily via steric rather than electrostatic effects. Intriguingly, even relatively small modifications such as histone methylation can apparently influence chromatin structure. Casadio *et al.* used a semisynthetic approach to demonstrate that asymmetric dimethylation of H3-Arg42 residue lowers nucleosome stability and stimulates transcription.[99] This destabilization of the nucleosome was attributed to the change in histone-DNA interaction upon loss of potential hydrogen bond. Similarly, dimethylation of H3-Lys79 results in a subtle structural

reorganization leading to increased accessibility of the modified residue apparently through loss of a hydrogen bond.[100]

Probing interaction partners and downstream spreading of modifications

In addition to elucidating the structural effect of histone PTMs on chromatin dynamics, synthetic histone proteins and designer nucleosomes have also allowed identification of key interacting protein partners (Figure 3.b.). In seminal work, Bartke *et al.* used designer nucleosomes bearing modifications of nucleosomal DNA and H3 protein at Lys-4, Lys-9, and Lys-27 to affinity-enrich nucleosome-binding proteins from nuclear extracts.[101] The use of quantitative mass spectrometry allowed them to identify the preferential protein partners for individual as well as combinations of epigenetic marks in an unbiased manner. Rapid emergence of quantitative proteomics fueled by the advancement of mass spectrometers in terms of sensitivity, sequencing efficiency/speed, and mass accuracy coupled with new quantitative approaches has allowed this approach to now become a mainstay for identification of chromatin-binding proteins, over traditional western-blotting based methods.[19] This approach was further extended by using a defined oligo-nucleosome template for affinity-enrichment of binding protein partners.[102] A similar approach using synthetic GlcNAcylated nucleosome has allowed the identification of sub-units of FACT complex as interacting protein partners for nucleosomes GlcNAcylated at H2B-Ser112 [103] and led to a suggested updating of prior molecular mechanisms for GlcNAcylation-driven transcriptional activation that had been determined using only isolated GlcNAcylated-H2B protein. [104] The latter example highlights that more representative nucleosomal context may prove critical and that synthetic chemical biology should aim to build and use such higher order structures as preferred probes.

Towards *in vivo* chromatin chemical biology

In vitro chromatin biology experiments have provided valuable insights, yet these are reductionist approaches with the obvious possibility of consequent artefacts – i.e. key players might be missing from any reconstituted or probed system. Therefore, much effort is rightly focused on developing *in vivo* chromatin biology techniques [105].

Highly promising developments are occurring in the field of genetic code expansion. Non-sense suppression mutagenesis has been used for histones in mammalian cells providing exciting insights into lysine acetylation [37,41]. Given the recently claimed expansion of the technique to mouse as a model organism [106], more complex organismal chromatin biology might be expected to follow soon. Moreover, high affinity ultrafast split inteins now allow for the incorporation of specific modifications inside cells (Figure 3c).[60] Recently, this technique has been further (most

impressively) improved for locus-specific *in vivo* chromatin tailoring. Using the dCas9 targeting domains, Muir and co-workers were able to direct synthetic modification towards locations in a cellular genome (Figure 3d). [107] Such *in cellulo* host-guest-directed methods suggest modes of control that may even eventually allow use of other ligand-directed, site-selective methods; ligand-directed chemical protein acetylation can be performed on native nucleosomes *in vitro* [108] and could provide a good candidate small-molecule system for such *in vivo* transposition.

Outlook

In the last decade, synthetic chemical biology techniques have emerged that provide tools for the precise molecular investigation of chromatin biology. Whereas the *in vitro* generation of synthetic histones and designer nucleosomes can readily be achieved in complementary ways using genetic code expansion, protein semisynthesis or chemical mutagenesis, transposing these techniques to suitable *in vivo* applications remains one of the key challenges remaining in synthetic histone modification. Given the central role of dynamically- and diversely- altered higher order structures in epigenetics, the aim ultimately is to reveal, study and/or control epigenetic modification in its proper context both spatially and temporarily. It seems likely that for us to truly compete with the late-stage diversity generated by nature, then late-stage chemical tools and methods will prove important.

Conflicts of Interest

B.G.D. is editor-in-chief of Current Opinion in Chemical Biology. B.G.D. is inventor on patent WO 2009103941 filed by the University of Oxford assigned to Catalent Biologics that covers, in one claim, reaction of the carbon-carbon double bond of dehydroalanine residues reacted with an organohalide in the presence of elemental zinc. B.G.D. is a member of the Catalent Biologics Scientific Advisory Board. B.G.D. is a supplier of the DBHDA reagent via the Kerafast platform.

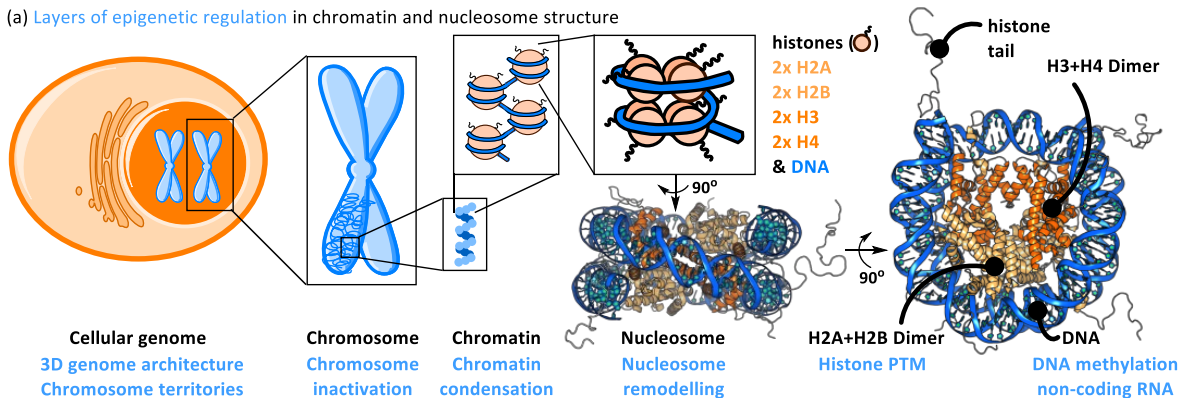
Acknowledgements

The authors thank Dr. Jitka Dadova for helpful discussions. This work was supported by the Felix Foundation (R.R) and EPSRC (S.N.).

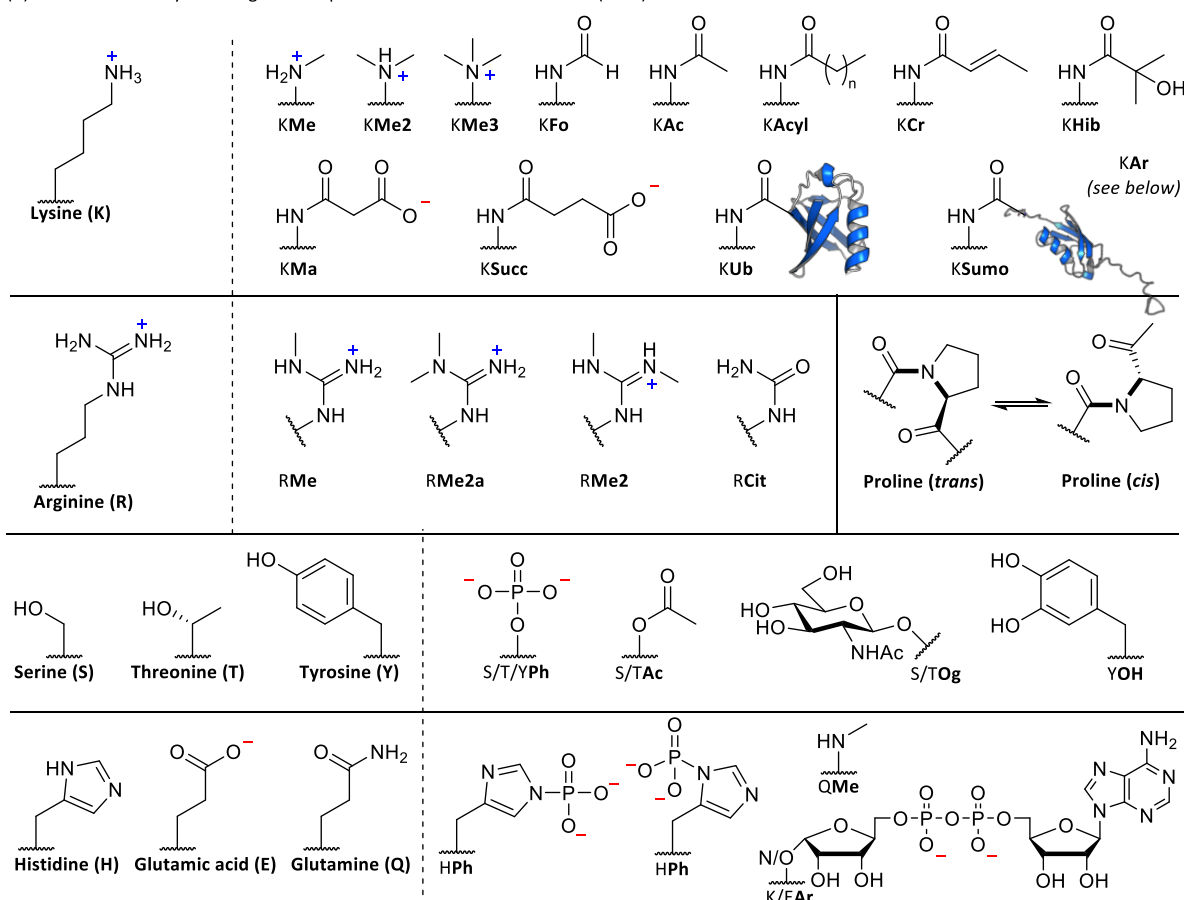
Figures

Figure 1

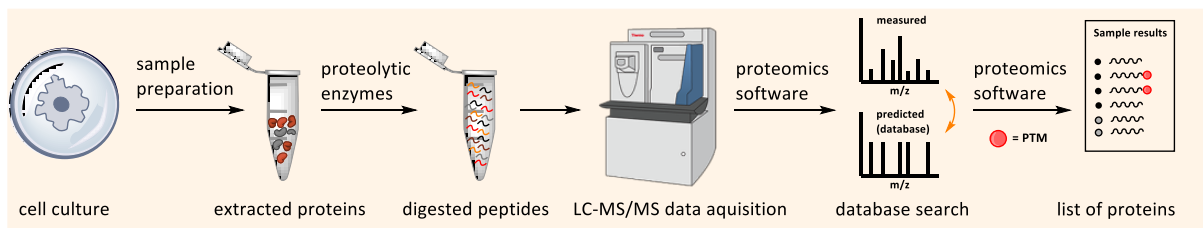
(a) Layers of epigenetic regulation in chromatin and nucleosome structure



(b) Selected naturally occurring histone post-translational modifications (PTM)



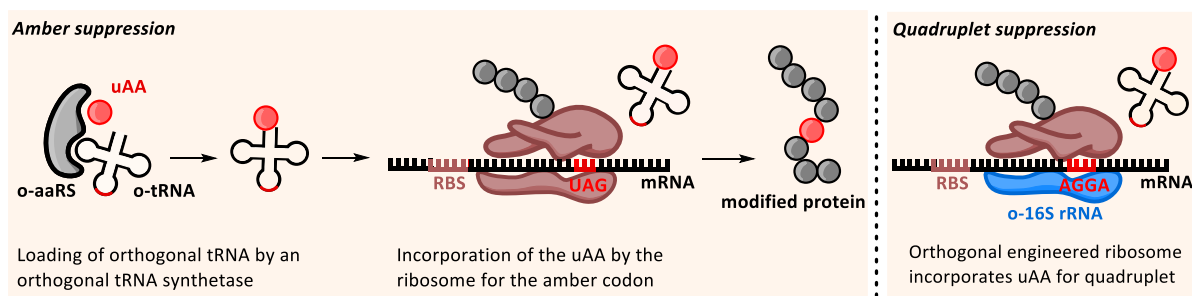
(c) Mass spectrometry-based proteomics for the identification of histone post-translational modifications



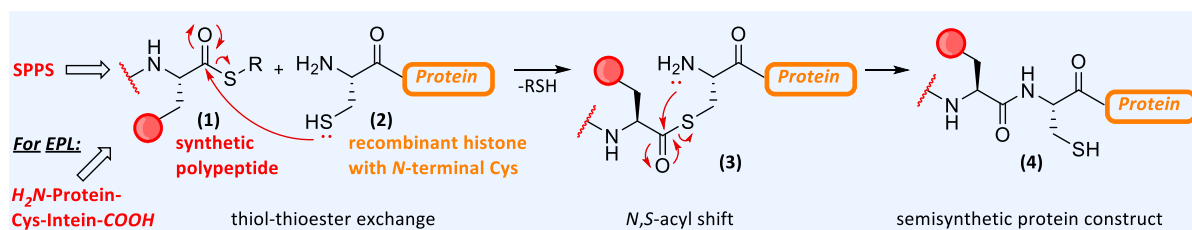
Post-translational modification of histones and their identification using mass-spectrometry based proteomics. **(a)** The packaging of cellular DNA into chromosomes and nucleosomes allows for several layers of epigenetic regulation. One of these is the post-translational modification of histone proteins, core components of the nucleosome. PDB: 1KX5 [109]. **(b)** Selected commonly observed histone post-translational modifications (PTMs), listed by amino acid site of modification, illustrate the broad range in chemistries found on histones. Adapted from ref [19]. **(c)** Histone PTMs can be identified by mass-spectrometry based proteomics. In this bottom-up proteomics workflow (based on digested proteolytic peptides) identification of known PTM is described. Ac – acetylation, Acyl - acylation, Ar – ADP-ribosylation, Cit – citrullination, Cr - crotonylation, Fo – formylation, Hib – 2-hydroxyisobutyrylation, Ma – malonylation, Me – methylation, Og – O-GlcNAcylation, OH – hydroxylation, Ph – phosphorylation, Succ – succinylation, Sumo – SUMOylation, Ub – ubiquitinylation.

Figure 2

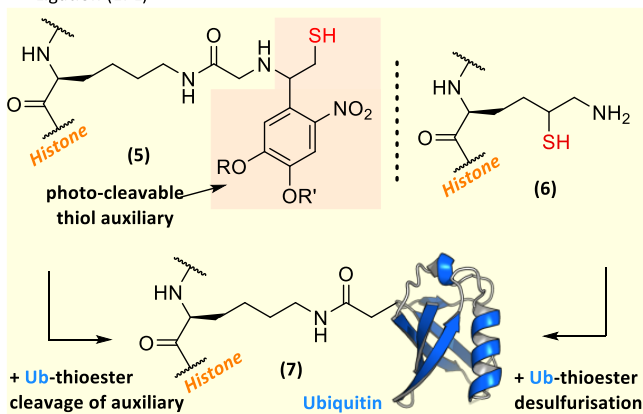
(a) Synthetic histones by genetic code expansion



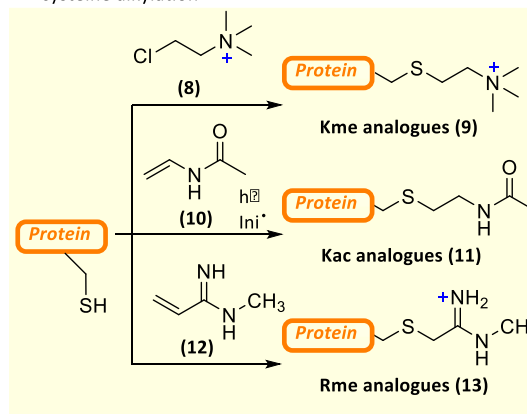
(b) Synthetic histones by protein semisynthesis using native chemical ligation (NCL)



(c) Synthetic histones by coupling to a full protein via Expressed Protein Ligation (EPL)

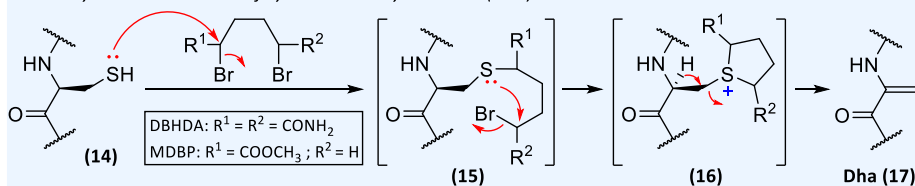


(d) Synthetic histones by chemical mutagenesis through cysteine alkylation

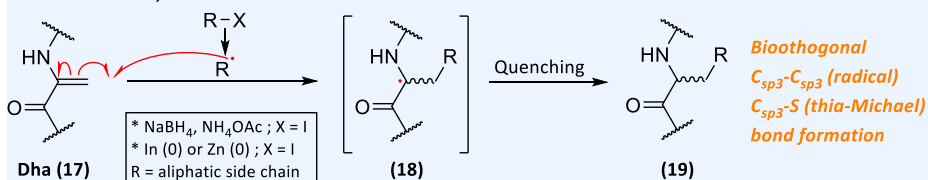


(e) Synthetic histones by dehydroalanine-mediated (Dha) chemical mutagenesis - 'tag-and-modify' approach

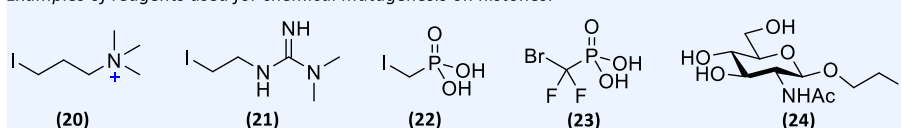
1. Bis-alkylation-elimination of cysteine to dehydroalanine (Dha)



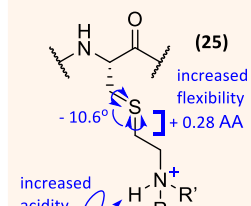
2. Addition to dehydroalanine



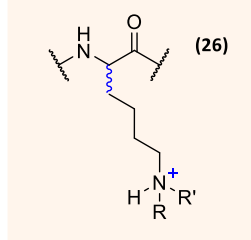
Examples of reagents used for chemical mutagenesis on histones:



Structural changes in sulfinated analogues



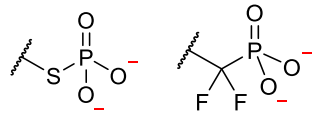
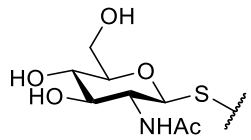
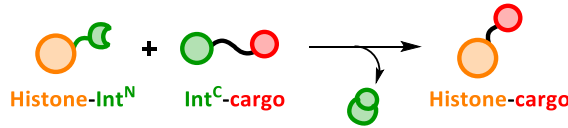
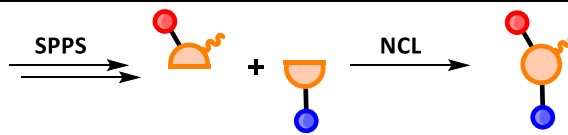
Loss of stereocenter in Dha-mediated mutagenesis



The chemical biology toolkit for synthetic histone modification. **(a)** Genetic code expansion. Amber codon (TAG) suppression is the most widely used technique for targeted genetic incorporation. Another possibility involves the use of an engineered ribosome with an orthogonal 16S ribosomal unit for quadruplet codon recognition. Adapted from [45]. **(b)** Mechanism of native chemical ligation (NCL) for protein semisynthesis. Using this technique, a protein with an *N*-terminal cysteine can be coupled to a synthetic peptide with a *C*-terminal thioester. **(c)** Expressed protein ligation can be used to couple proteins such as ubiquitin to peptides/proteins containing thiol auxiliaries **5** or the amino acid δ -mercaptolysine **6**. These approaches can be made traceless by UV-irradiation or desulfurisation. Adapted from [14]. **(d)** Historically, thioether-based PTM mimics were used by reacting a recombinant histone containing a cysteine at the desired position with thiol-alkylating reagents. These mimics however bring structural changes with them (frame below, adapted from [16]). **(e)** Many histone PTMs can be accessed using chemical mutagenesis, in a two step protocol. First, elimination of a recombinantly incorporated cysteine at the desired position of modification to dehydroalanine is performed using either the DBHDA or MDBP reagent. Unfortunately, this comes at the cost of mixed D/L configuration at the alpha-carbon. Second, addition to dehydroalanine is performed. Depending on the type of bond formation required, a radical C-C bond formation or thia-Michael reaction can be chosen; this representation emphasises the recently reported radical mechanism.

o-aaRS – orthogonal aminoacyl tRNA synthetase, uAA – unnatural amino acid, RBS – ribosomal binding site.

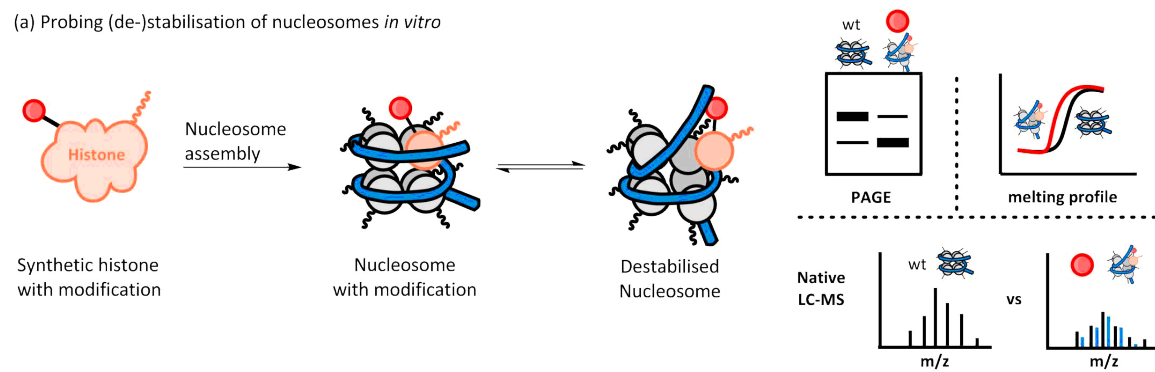
Table 1

Modification	Suggested incorporation technique	Scheme & notes
Phosphorylation	Chemical mutagenesis	 <ul style="list-style-type: none"> • non-hydrolysable, phosphatase-resistant • CF₂ is isoelectronic to natural phosphoester
<i>O</i> -GlcNAc glycosylation	Chemical mutagenesis	 <ul style="list-style-type: none"> • Same conformation as natural acetal linkage[110] • Resistant to hOGA glycolytic activity
Protein coupling	Native chemical ligation	<ul style="list-style-type: none"> • See Figure 2.c
Multiply modified histone tails	Intein trans-splicing or expressed protein ligation	
Multiply modified core histones	Histone total synthesis with NCL schemes	

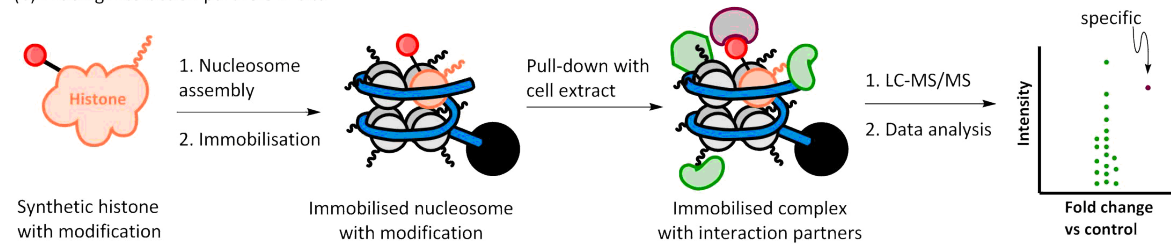
Challenging histone PTMs and illustrative current incorporation techniques.

Figure 3

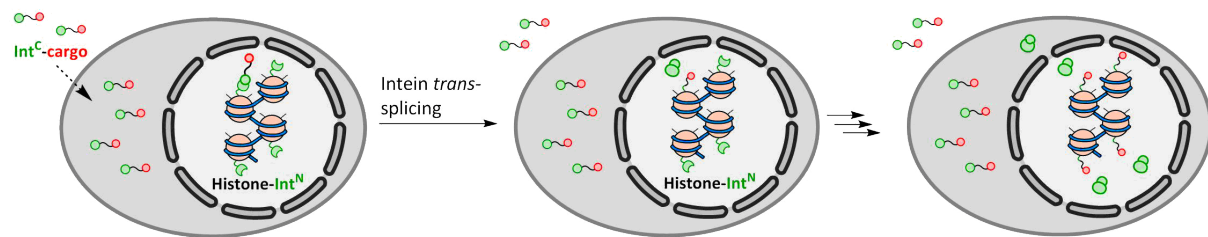
(a) Probing (de-)stabilisation of nucleosomes *in vitro*



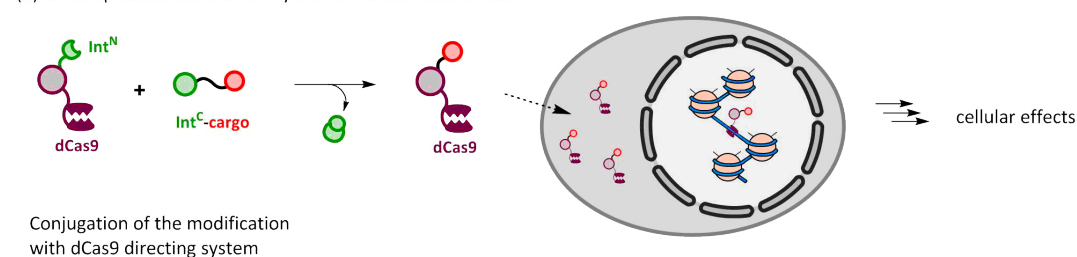
(b) Probing interaction partners *in-situ*



(c) Integrating modified nucleosomes *in vivo*



(d) Locus-specific insertion of a synthetic modification *in vivo*



Application of chemical biology tools for understanding chromatin biology. **(a)** Synthetic histone modifications can be used to probe their effects on nucleosome stability and chromatin architecture. **(b)** Synthetic histones can be assembled with other canonical histone proteins and biotinylated DNA to a designer nucleosome. These immobilized synthetic nucleosomes can be used for affinity enrichment of nucleosome-binding proteins from nuclear extract. Pooled proteins from sample and unmodified control are enzymatically digested and identified by LC-MS/MS analysis. The identified protein partners are quantified against each other using MS-based Label-free Quantification (LFQ). **(c)** Ultrafast split intein technology allows for the *in vivo* modification of chromatin. Using a pair with the Int^N genetically fused to histones and the synthetic Int^C exogenously added, protein semi-

synthesis can be performed *in vivo*, thus incorporating the desired modification. **(d)** Using split inteins, a synthetic construct of a histone bearing the target modification and the Cas9 directing system can be generated. Upon incorporation of the fusion construct into cells, the modified histone is delivered to the target location in the cellular genome, thus influencing epigenetic processes in its close environment. (Adapted from [105]).

References and recommended reading

- of special Interest
- of outstanding interest

1. Pombo A, Dillon N: **Three-dimensional genome architecture: players and mechanisms.** *Nature Reviews Molecular Cell Biology* 2015, **16**:245.
2. Fraser J, Williamson I, Bickmore WA, Dostie J: **An Overview of Genome Organization and How We Got There: from FISH to Hi-C.** *Microbiology and Molecular Biology Reviews* 2015, **79**:347-372.
3. Bannister AJ, Kouzarides T: **Regulation of chromatin by histone modifications.** *Cell Research* 2011, **21**:381.
4. Kouzarides T: **Chromatin Modifications and Their Function.** *Cell* 2007, **128**:693-705.
5. Musselman CA, Lalonde M-E, Côté J, Kutateladze TG: **Perceiving the epigenetic landscape through histone readers.** *Nature Structural & Molecular Biology* 2012, **19**:1218.
6. Dhall A, Chatterjee C: **Chemical approaches to understand the language of histone modifications.** *ACS Chem Biol* 2011, **6**:987-999.
7. Allis CD, Muir TW: **Spreading chromatin into chemical biology.** *Chembiochem* 2011, **12**:264-279.
8. Shogren-Knaak MA, Peterson CL: **Creating Designer Histones by Native Chemical Ligation.** In *Methods in Enzymology*. Edited by: Academic Press; 2003:62-76. Chromatin and Chromatin Remodeling Enzymes, Part A, vol 375.]
9. Schwarzer D: **Chemical tools in chromatin research.** *J Pept Sci* 2010, **16**:530-537.
10. Fierz B, Muir TW: **Chromatin as an expansive canvas for chemical biology.** *Nat Chem Biol* 2012, **8**:417-427.
11. Nikolov M, Fischle W: **Systematic analysis of histone modification readout.** *Mol Biosyst* 2013, **9**:182-194.
12. Fierz B: **Synthetic chromatin approaches to probe the writing and erasing of histone modifications.** *ChemMedChem* 2014, **9**:495-504.
13. Pick H, Kilic S, Fierz B: **Engineering chromatin states: chemical and synthetic biology approaches to investigate histone modification function.** *Biochim Biophys Acta* 2014, **1839**:644-656.
14. Fischle W, Mootz HD, Schwarzer D: **Synthetic histone code.** *Curr Opin Chem Biol* 2015, **28**:131-140.
15. Holt M, Muir T: **Application of the protein semisynthesis strategy to the generation of modified chromatin.** *Annu Rev Biochem* 2015, **84**:265-290.
16. Muller MM, Muir TW: **Histones: at the crossroads of peptide and protein chemistry.** *Chem Rev* 2015, **115**:2296-2349.

- This review comprehensively describes the broad scope of histone peptide chemistry as well as histone protein semisynthesis.

17. Chuh KN, Batt AR, Pratt MR: **Chemical Methods for Encoding and Decoding of Posttranslational Modifications**. *Cell Chem Biol* 2016, **23**:86-107.

18. Yang A, Cho K, Park H-S: **Chemical biology approaches for studying posttranslational modifications**. *RNA Biology* 2017:1-14.

19. Huang H, Lin S, Garcia BA, Zhao Y: **Quantitative Proteomic Analysis of Histone Modifications**. *Chemical Reviews* 2015, **115**:2376-2418.

- A comprehensive and still up-to-date review of mass spectrometry-based proteomics for the identification and discovery of histone PTM.

20. Dai L, Peng C, Montellier E, Lu Z, Chen Y, Ishii H, Debernardi A, Buchou T, Rousseaux S, Jin F, et al.: **Lysine 2-hydroxyisobutyrylation is a widely distributed active histone mark**. *Nat Chem Biol* 2014, **10**:365.

21. Howden AJM, Geoghegan V, Katsch K, Efstathiou G, Bhushan B, Boutureira O, Thomas B, Trudgian DC, Kessler BM, Dieterich DC, et al.: **QuaNCAT: quantitating proteome dynamics in primary cells**. *Nature Methods* 2013, **10**:343.

22. Dieterich DC, Link AJ, Graumann J, Tirrell DA, Schuman EM: **Selective identification of newly synthesized proteins in mammalian cells using bioorthogonal noncanonical amino acid tagging (BONCAT)**. *Proc Natl Acad Sci, U.S.A.* 2006, **103**:9482-9487.

23. Arnaudo AM, Link AJ, Garcia BA: **Bioorthogonal Chemistry for the Isolation and Study of Newly Synthesized Histones and Their Modifications**. *ACS Chem Biol* 2016, **11**:782-791.

24. Lin S, Wein S, Gonzales-Cope M, Otte GL, Yuan Z-F, Afjehi-Sadat L, Maile T, Berger SL, Rush J, Lill JR, et al.: **Stable-isotope-labeled Histone Peptide Library for Histone Post-translational Modification and Variant Quantification by Mass Spectrometry**. *Molecular & Cellular Proteomics* 2014, **13**:2450-2466.

- Using a library of synthetic histone peptides carrying defined PTMs, the effect of these PTMs on peptide detection and identification is investigated. The findings are food for thought on quantitative histone PTM experiments.

25. Kulej K, Avgousti DC, Weitzman MD, Garcia BA: **Characterization of histone post-translational modifications during virus infection using mass spectrometry-based proteomics**. *Methods* 2015, **90**:8-20.

26. Lang K, Chin JW: **Cellular Incorporation of Unnatural Amino Acids and Bioorthogonal Labeling of Proteins**. *Chemical Reviews* 2014, **114**:4764-4806.

27. Wang L, Brock A, Herberich B, Schultz PG: **Expanding the Genetic Code of *Escherichia coli***. *Science* 2001, **292**:498-500.

28. Chin JW: **Expanding and reprogramming the genetic code**. *Nature* 2017, **550**:53-60.

29. Dumas A, Lercher L, Spicer CD, Davis BG: **Designing logical codon reassignment - Expanding the chemistry in biology**. *Chemical Science* 2015, **6**:50-69.

- This review contains a list of all uAAs that have been successfully included by genetic code expansion (by 2014) which clearly shows the structural analogy of the uAA to the original amino acid of the synthetase.
30. Neumann H, Peak-Chew SY, Chin JW: **Genetically encoding N(epsilon)-acetyllysine in recombinant proteins.** *Nat Chem Biol* 2008, **4**:232-234.
 31. Gattner MJ, Vrabel M, Carell T: **Synthesis of epsilon-N-propionyl-, epsilon-N-butyryl-, and epsilon-N-crotonyl-lysine containing histone H3 using the pyrrolysine system.** *Chemical Communications* 2013, **49**:379-381.
 32. Kim CH, Kang M, Kim HJ, Chatterjee A, Schultz PG: **Site-Specific Incorporation of ε-N-Crotonyllysine into Histones.** *Angewandte Chemie International Edition* 2012, **51**:7246-7249.
 33. Knight WA, Cropp TA: **Genetic encoding of the post-translational modification 2-hydroxyisobutyryl-lysine.** *Organic & Biomolecular Chemistry* 2015, **13**:6479-6481.
 34. Wilkins BJ, Hahn LE, Heitmüller S, Frauendorf H, Valerius O, Braus GH, Neumann H: **Genetically Encoding Lysine Modifications on Histone H4.** *ACS Chemical Biology* 2015, **10**:939-944.
 35. Nguyen DP, Garcia Alai MM, Kapadnis PB, Neumann H, Chin JW: **Genetically Encoding Ne-Methyl-l-lysine in Recombinant Histones.** *Journal of the American Chemical Society* 2009, **131**:14194-14195.
 36. Nguyen DP, Alai MMG, Virdee S, Chin JW: **Genetically Directing ε-N, N-Dimethyl-l-Lysine in Recombinant Histones.** *Chemistry & Biology* 2010, **17**:1072-1076.
 37. Elsasser SJ, Ernst RJ, Walker OS, Chin JW: **Genetic code expansion in stable cell lines enables encoded chromatin modification.** *Nat Meth* 2016, **13**:158-164.
 - The authors could generate stable cell lines for amber suppression mutagenesis containing the PylRS/tRNA pair. This system was used to investigate the effects of H3 lysine acetylation *in vivo*.
 38. Chin JW, Martin AB, King DS, Wang L, Schultz PG: **Addition of a photocrosslinking amino acid to the genetic code of Escherichia coli.** *Proceedings of the National Academy of Sciences* 2002, **99**:11020-11024.
 39. Yang T, Li X-M, Bao X, Fung YME, Li XD: **Photo-lysine captures proteins that bind lysine post-translational modifications.** *Nat Chem Biol* 2015, **12**:70.
 40. Xie X, Li X-M, Qin F, Lin J, Zhang G, Zhao J, Bao X, Zhu R, Song H, Li XD, et al.: **Genetically Encoded Photoaffinity Histone Marks.** *Journal of the American Chemical Society* 2017, **139**:6522-6525.
 41. Wilkins BJ, Rall NA, Ostwal Y, Kruitwagen T, Hiragami-Hamada K, Winkler M, Barral Y, Fischle W, Neumann H: **A Cascade of Histone Modifications Induces Chromatin Condensation in Mitosis.** *Science* 2014, **343**:77-80.
 42. Zhang MS, Brunner SF, Huguenin-Dezot N, Liang AD, Schmied WH, Rogerson DT, Chin JW: **Biosynthesis and genetic encoding of phosphothreonine through parallel selection and deep sequencing.** *Nat Meth* 2017, **14**:729-736.
 - Through parallel positive selection and sequencing the authors present a scalable approach to the discovery of orthogonal synthetases.

43. Park H-S, Hohn MJ, Umehara T, Guo L-T, Osborne EM, Benner J, Noren CJ, Rinehart J, Söll D: **Expanding the Genetic Code of *Escherichia coli* with Phosphoserine.** *Science* 2011, **333**:1151-1154.
44. Lee S, Oh S, Yang A, Kim J, Söll D, Lee D, Park H-S: **A Facile Strategy for Selective Incorporation of Phosphoserine into Histones.** *Angewandte Chemie International Edition* 2013, **52**:5771-5775.
45. Des Soye BJ, Patel JR, Isaacs FJ, Jewett MC: **Repurposing the translation apparatus for synthetic biology.** *Current Opinion in Chemical Biology* 2015, **28**:83-90.
46. Sachdeva A, Wang K, Elliott T, Chin JW: **Concerted, Rapid, Quantitative, and Site-Specific Dual Labeling of Proteins.** *Journal of the American Chemical Society* 2014, **136**:7785-7788.
47. Dawson P, Muir T, Clark-Lewis I, Kent S: **Synthesis of proteins by native chemical ligation.** *Science* 1994, **266**:776-779.
48. Liu J, Chen Q, Rozovsky S: **Utilizing Selenocysteine for Expressed Protein Ligation and Bioconjugations.** *Journal of the American Chemical Society* 2017, **139**:3430-3437.
49. He S, Bauman D, Davis JS, Loyola A, Nishioka K, Gronlund JL, Reinberg D, Meng F, Kelleher N, McCafferty DG: **Facile synthesis of site-specifically acetylated and methylated histone proteins: reagents for evaluation of the histone code hypothesis.** *Proc Natl Acad Sci U S A* 2003, **100**:12033-12038.
50. Shogren-Knaak MA, Fry CJ, Peterson CL: **A Native Peptide Ligation Strategy for Deciphering Nucleosomal Histone Modifications.** *Journal of Biological Chemistry* 2003, **278**:15744-15748.
51. Kee J-M, Villani B, Carpenter LR, Muir TW: **Development of Stable Phosphohistidine Analogues.** *Journal of the American Chemical Society* 2010, **132**:14327-14329.
52. McGinty RK, Kim J, Chatterjee C, Roeder RG, Muir TW: **Chemically ubiquitylated histone H2B stimulates hDot1L-mediated intranucleosomal methylation.** *Nature* 2008, **453**:812-816.
53. Ajish Kumar KS, Haj-Yahya M, Olschewski D, Lashuel HA, Brik A: **Highly Efficient and Chemoselective Peptide Ubiquitylation.** *Angewandte Chemie International Edition* 2009, **48**:8090-8094.
54. Siman P, Brik A: **Chemical and semisynthesis of posttranslationally modified proteins.** *Org Biomol Chem* 2012, **10**:5684-5697.
55. Maity SK, Jbara M, Brik A: **Chemical and semisynthesis of modified histones.** *J Pept Sci* 2016, **22**:252-259.
56. Jbara M, Maity SK, Brik A: **Palladium in the Chemical Synthesis and Modification of Proteins.** *Angewandte Chemie International Edition* 2017, **56**:10644-10655.
57. Maity SK, Jbara M, Mann G, Kamnesky G, Brik A: **Total chemical synthesis of histones and their analogs, assisted by native chemical ligation and palladium complexes.** *Nat. Protocols* 2017, **12**:2293-2322.
58. Jbara M, Maity SK, Brik A: **Palladium in the Chemical Synthesis and Modification of Proteins.** *Angew Chem Int Ed Engl* 2017, **56**:10644-10655.

59. Bondalapati S, Jbara M, Brik A: **Expanding the chemical toolbox for the synthesis of large and uniquely modified proteins.** *Nat Chem* 2016, **8**:407-418.
60. David Y, Vila-Perello M, Verma S, Muir TW: **Chemical tagging and customizing of cellular chromatin states using ultrafast trans-splicing inteins.** *Nat Chem* 2015, **7**:394-402.
61. Stevens AJ, Sekar G, Shah NH, Mostafavi AZ, Cowburn D, Muir TW: **A promiscuous split intein with expanded protein engineering applications.** *Proceedings of the National Academy of Sciences* 2017, **114**:8538-8543.
62. Gramespacher JA, Stevens AJ, Thompson RE, Muir TW: **Improved Protein Splicing Using Embedded Split Inteins.** *Protein Sci* 2018, DOI: 10.1002/pro.3357
63. Zettler J, Schütz V, Mootz HD: **The naturally split Npu DnaE intein exhibits an extraordinarily high rate in the protein trans-splicing reaction.** *FEBS Letters* 2009, **583**:909-914.
64. Behrendt R, White P, Offer J: **Advances in Fmoc solid-phase peptide synthesis.** *Journal of Peptide Science* 2016, **22**:4-27.
65. Zhang Y, Muthana SM, Farnsworth D, Ludek O, Adams K, Barchi JJ, Gildersleeve JC: **Enhanced Epimerization of Glycosylated Amino Acids During Solid-Phase Peptide Synthesis.** *Journal of the American Chemical Society* 2012, **134**:6316-6325.
66. Zioudrou C, Wilchek M, Patchornik A: **Conversion of the L-Serine Residue to an L-Cysteine Residue in Peptides.** *Biochemistry* 1965, **4**:1811-1822.
67. Polgar L, Bender ML: **A New Enzyme Containing a Synthetically Formed Active Site. Thiol-Subtilisin1.** *Journal of the American Chemical Society* 1966, **88**:3153-3154.
68. Neet KE, Koshland DE: **The conversion of serine at the active site of subtilisin to cysteine: a "chemical mutation".** *Proceedings of the National Academy of Sciences of the United States of America* 1966, **56**:1606-1611.
69. Chalker JM, Davis BG: **Chemical mutagenesis: selective post-expression interconversion of protein amino acid residues.** *Current Opinion in Chemical Biology* 2010, **14**:781-789.
70. Wright TH, Vallée MRJ, Davis BG: **From Chemical Mutagenesis to Post-Expression Mutagenesis: A 50 Year Odyssey.** *Angewandte Chemie International Edition* 2016, **55**:5896-5903.
71. Gunnoo SB, Madder A: **Chemical Protein Modification through Cysteine.** *Chembiochem* 2016, **17**:529-553.
72. Chalker JM, Lercher L, Rose NR, Schofield CJ, Davis BG: **Conversion of cysteine into dehydroalanine enables access to synthetic histones bearing diverse post-translational modifications.** *Angew Chem Int Ed Engl* 2012, **51**:1835-1839.
73. Simon MD, Chu F, Racki LR, de la Cruz CC, Burlingame AL, Panning B, Narlikar GJ, Shokat KM: **The site-specific installation of methyl-lysine analogs into recombinant histones.** *Cell* 2007, **128**:1003-1012.

74. Le DD, Cortesi AT, Myers SA, Burlingame AL, Fujimori DG: **Site-specific and regiospecific installation of methylarginine analogues into recombinant histones and insights into effector protein binding.** *J Am Chem Soc* 2013, **135**:2879-2882.
75. Huang R, Holbert MA, Tarrant MK, Curtet S, Colquhoun DR, Dancy BM, Dancy BC, Hwang Y, Tang Y, Meeth K, et al.: **Site-Specific Introduction of an Acetyl-Lysine Mimic into Peptides and Proteins by Cysteine Alkylation.** *Journal of the American Chemical Society* 2010, **132**:9986-9987.
76. Li F, Allahverdi A, Yang R, Lua GB, Zhang X, Cao Y, Korolev N, Nordenskiöld L, Liu CF: **A direct method for site-specific protein acetylation.** *Angew Chem Int Ed Engl* 2011, **50**:9611-9614.
77. Chalker JM, Gunnoo SB, Boutureira O, Gerstberger SC, Fernández-González M, Bernardes GJL, Griffin L, Hailu H, Schofield CJ, Davis BG: **Methods for converting cysteine to dehydroalanine on peptides and proteins.** *Chemical Science* 2011, **2**:1666.
- This study introduces the DBHDA reagent for conversion of cysteine to dehydroalanine on native proteins.
78. Seeliger D, Soeroes S, Klingberg R, Schwarzer D, Grubmüller H, Fischle W: **Quantitative Assessment of Protein Interaction with Methyl-Lysine Analogues by Hybrid Computational and Experimental Approaches.** *ACS Chemical Biology* 2012, **7**:150-154.
- This study represents a rather unique comparison of experimental and computational data. Using a combination of molecular dynamics simulations, based on a state-of-the-art quantum mechanical parametrisation of alkylated cysteine analogues for methyllysine, the authors provide a framework and tools to evaluate the relevance of this synthetic construct. The model is subsequently evaluated experimentally. We predict such mixed computational and experimental approaches to gain an increasing role in Chemical Biology projects.
79. Wright TH, Bower BJ, Chalker JM, Bernardes GJ, Wiewiora R, Ng WL, Raj R, Faulkner S, Vallee MR, Phanumartwiwath A, et al.: **Posttranslational mutagenesis: A chemical strategy for exploring protein side-chain diversity.** *Science* 2016, **354**:aag1465.
- This seminal publication introduces the first bioorthogonal Csp3-Csp3 bond formation on native proteins, an approach termed chemical mutagenesis. This technique allows to alter residues of recombinantly expressed proteins to a broad range of natural and unnatural amino acids, including post-translational modifications.
80. Yang A, Ha S, Ahn J, Kim R, Kim S, Lee Y, Kim J, Söll D, Lee H-Y, Park H-S: **A chemical biology route to site-specific authentic protein modifications.** *Science* 2016, **354**:623-626.
81. Morrison PM, Foley PJ, Warriner SL, Webb ME: **Chemical generation and modification of peptides containing multiple dehydroalanines.** *Chem Commun (Camb)* 2015, **51**:13470-13473.
- This is the first report of the MDBP reagent for generation of Dha from cysteine. This reagent has faster reaction kinetics than DBHDA, has improved solubility and does not lead to overalkylation on lysine residues.
82. Dadová J, Wu K-J, Isenegger PG, Errey JC, Bernardes GJL, Chalker JM, Raich L, Rovira C, Davis BG: **Precise Probing of Residue Roles by Post-Translational β,γ -C,N Aza-Michael Mutagenesis in Enzyme Active Sites.** *ACS Central Science* 2017, **3**:1168-1173.
- The first reported use of MDBP as a reagent to generate Dha on proteins. Moreover, the first aza-Michael addition to Dha is described.

83. Wright TH, Davis BG: **Post-translational mutagenesis for installation of natural and unnatural amino acid side chains into recombinant proteins.** *Nat. Protocols* 2017, **12**:2243-2250.
84. Guo J, Wang J, Lee JS, Schultz PG: **Site-Specific Incorporation of Methyl- and Acetyl-Lysine Analogues into Recombinant Proteins.** *Angewandte Chemie International Edition* 2008, **47**:6399-6401.
85. Wang ZU, Wang Y-S, Pai P-J, Russell WK, Russell DH, Liu WR: **A Facile Method to Synthesize Histones with Posttranslational Modification Mimics.** *Biochemistry* 2012, **51**:5232-5234.
86. Key HM, Miller SJ: **Site- and Stereoselective Chemical Editing of Thiostrepton by Rh-Catalyzed Conjugate Arylation: New Analogues and Collateral Enantioselective Synthesis of Amino Acids.** *Journal of the American Chemical Society* 2017, **139**:15460-15466.
87. Chapman CJ, Matsuno A, Frost CG, Willis MC: **Site-selective modification of peptides using rhodium and palladium catalysis: complementary electrophilic and nucleophilic arylation.** *Chemical Communications* 2007:3903-3905.
88. Chalker JM, Bernardes GJL, Davis BG: **A “Tag-and-Modify” Approach to Site-Selective Protein Modification.** *Accounts of Chemical Research* 2011, **44**:730-741.
89. Nguyen UTT, Bittova L, Müller MM, Fierz B, David Y, Houck-Loomis B, Feng V, Dann GP, Muir TW: **Accelerated chromatin biochemistry using DNA-barcoded nucleosome libraries.** *Nature Methods* 2014, **11**:834.
90. Tessarz P, Kouzarides T: **Histone core modifications regulating nucleosome structure and dynamics.** *Nature Reviews Molecular Cell Biology* 2014, **15**:703.
91. Taverna SD, Li H, Ruthenburg AJ, Allis CD, Patel DJ: **How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers.** *Nature Structural & Molecular Biology* 2007, **14**:1025.
92. Wang Z, Zang C, Rosenfeld JA, Schones DE, Barski A, Cuddapah S, Cui K, Roh T-Y, Peng W, Zhang MQ, et al.: **Combinatorial patterns of histone acetylations and methylations in the human genome.** *Nature Genetics* 2008, **40**:897.
93. Di Cerbo V, Mohn F, Ryan DP, Montellier E, Kacem S, Tropberger P, Kallis E, Holzner M, Hoerner L, Feldmann A, et al.: **Acetylation of histone H3 at lysine 64 regulates nucleosome dynamics and facilitates transcription.** *eLife* 2014, **3**:e01632.
94. North JA, Šimon M, Ferdinand MB, Shoffner MA, Picking JW, Howard CJ, Mooney AM, van Noort J, Poirier MG, Ottesen JJ: **Histone H3 phosphorylation near the nucleosome dyad alters chromatin structure.** *Nucleic Acids Research* 2014, **42**:4922-4933.
95. Stützer A, Liokatis S, Kiesel A, Schwarzer D, Sprangers R, Söding J, Selenko P, Fischle W: **Modulations of DNA Contacts by Linker Histones and Post-translational Modifications Determine the Mobility and Modifiability of Nucleosomal H3 Tails.** *Molecular Cell* 2016, **61**:247-259.
96. Lercher L, Raj R, Patel NA, Price J, Mohammed S, Robinson CV, Schofield CJ, Davis BG: **Generation of a synthetic GlcNAcylated nucleosome reveals regulation of stability by H2A-Thr101 GlcNAcylation.** *Nature Communications* 2015, **6**:7978.

97. Yan Q, Dutt S, Xu R, Graves K, Juszczynski P, Manis JP, Shipp MA: **BBAP Monoubiquitylates Histone H4 at Lysine 91 and Selectively Modulates the DNA Damage Response.** *Molecular Cell* 2009, **36**:110-120.
98. Ye J, Ai X, Eugeni EE, Zhang L, Carpenter LR, Jelinek MA, Freitas MA, Parthun MR: **Histone H4 Lysine 91 Acetylation: A Core Domain Modification Associated with Chromatin Assembly.** *Molecular Cell* 2005, **18**:123-130.
99. Casadio F, Lu X, Pollock SB, LeRoy G, Garcia BA, Muir TW, Roeder RG, Allis CD: **H3R42me2a is a histone modification with positive transcriptional effects.** *Proceedings of the National Academy of Sciences* 2013, **110**:14894-14899.
100. Lu X, Simon MD, Chodaparambil JV, Hansen JC, Shokat KM, Luger K: **The effect of H3K79 dimethylation and H4K20 trimethylation on nucleosome and chromatin structure.** *Nat Struct Mol Biol* 2008, **15**:1122-1124.
101. Bartke T, Vermeulen M, Xhemalce B, Robson SC, Mann M, Kouzarides T: **Nucleosome-Interacting Proteins Regulated by DNA and Histone Methylation.** *Cell* 2010, **143**:470-484.
- This seminal paper combines the power of synthetic nucleosomes and quantitative proteomics for identification of nucleosome-binding protein partners in an unbiased manner.
102. Nikolov M, Stützer A, Mosch K, Krasauskas A, Soeroes S, Stark H, Urlaub H, Fischle W: **Chromatin Affinity Purification and Quantitative Mass Spectrometry Defining the Interactome of Histone Modification Patterns.** *Molecular & Cellular Proteomics* 2011, **10**.
103. Raj R, Lercher L, Mohammed S, Davis BG: **Synthetic Nucleosomes Reveal that GlcNAcylation Modulates Direct Interaction with the FACT Complex.** *Angewandte Chemie* 2016, **128**:9064-9068.
- Highlights the importance of 'proper' substrates i.e. nucleosome for studying epigenetic mechanisms as compared to peptides or proteins.
104. Fujiki R, Hashiba W, Sekine H, Yokoyama A, Chikanishi T, Ito S, Imai Y, Kim J, He HH, Igarashi K, et al.: **GlcNAcylation of histone H2B facilitates its monoubiquitination.** *Nature* 2011, **480**:557.
105. David Y, Muir TW: **Emerging Chemistry Strategies for Engineering Native Chromatin.** *Journal of the American Chemical Society* 2017, **139**:9090-9096.
106. Han S, Yang A, Lee S, Lee H-W, Park CB, Park H-S: **Expanding the genetic code of *Mus musculus*.** 2017, **8**:14568.
107. Liszczak GP, Brown ZZ, Kim SH, Oslund RC, David Y, Muir TW: **Genomic targeting of epigenetic probes using a chemically tailored Cas9 system.** *Proceedings of the National Academy of Sciences* 2017, **114**:681-686.
- By combining the split-intein technique with the dCas9 directing system, the authors were able to position a synthetic construct at a defined place in the genome.
108. Amamoto Y, Aoi Y, Nagashima N, Suto H, Yoshidome D, Arimura Y, Osakabe A, Kato D, Kurumizaka H, Kawashima SA, et al.: **Synthetic Posttranslational Modifications: Chemical Catalyst-Driven Regioselective Histone Acylation of Native Chromatin.** *J Am Chem Soc* 2017, **139**:7568-7576.
- First application of a ligand-directed catalyst-mediated site-specific acetylation of native nucleosomes. These findings could potentially bridge the world of small molecules with histone chemical biology.

109. Davey CA, Sargent DF, Luger K, Maeder AW, Richmond TJ: **Solvent Mediated Interactions in the Structure of the Nucleosome Core Particle at 1.9Å Resolution.** *Journal of Molecular Biology* 2002, **319**:1097-1113.
110. De Leon CA, Levine PM, Craven TW, Pratt MR: **The Sulfur-Linked Analogue of O-GlcNAc (S-GlcNAc) Is an Enzymatically Stable and Reasonable Structural Surrogate for O-GlcNAc at the Peptide and Protein Levels.** *Biochemistry* 2017, **56**:3507-3517.

Highlights

- Histones with specific PTMs can be synthesised and assembled in designer nucleosomes
- Codon suppression, semisynthesis or mutagenesis allow controlled insertion of PTMs
- Designer nucleosomes can probe stability and/or PTM specific interaction partners
- *In vivo* techniques show potential for locus-specific incorporation of PTM

Synthetic post-translational modification of histones

Simon Nadal¹, Ritu Raj¹, Shabaz Mohammed¹ and Benjamin G. Davis^{1*}

Current Opinion in Chemical Biology

Addresses

¹Department of Chemistry, University of Oxford, Chemistry Research Laboratory, Mansfield Road, Oxford OX1 3TA, UK

S.N. simon.nadal@chem.ox.ac.uk ; ORCID: 0000-0002-0748-0284

R.R. ritu.raj@chem.ox.ac.uk ; ORCID: 0000-0002-7623-9128

S.M. shabaz.mohammed@chem.ox.ac.uk ; ORCID: 0000-0003-2640-9560

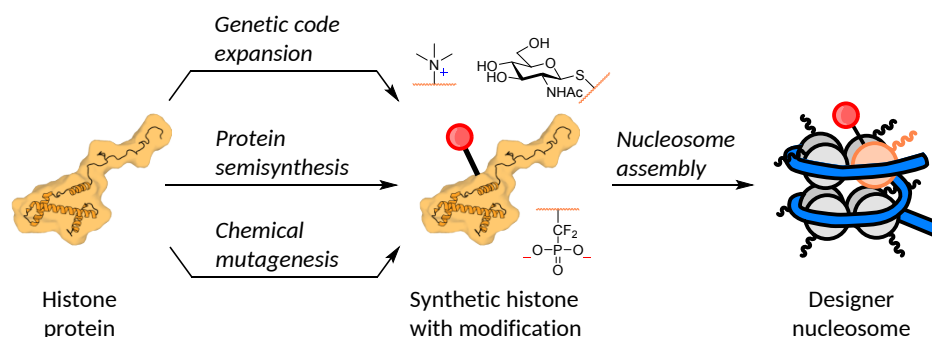
B.G.D. ben.davis@chem.ox.ac.uk ; ORCID: 0000-0002-5056-407X

* Corresponding author: Davis, Benjamin G. (ben.davis@chem.ox.ac.uk).

Abstract

Chromatin is the physiological template of genetic information in all eukaryotic cells, a highly organised complex of DNA and histone proteins central in regulating gene expression and genome organisation. A multitude of histone post-translational modifications (PTMs) have been discovered, providing a glance into the complex interplay of these epigenetic marks in cellular processes. In the last decade, synthetic and chemical biology techniques have emerged to study these modifications, including genetic code expansion, histone semisynthesis and post-translational chemical mutagenesis. These methods allow for the creation of histones carrying synthetic modifications which can in turn be assembled into designer nucleosomes. Their application in *in vitro* and *in vivo* studies is now beginning to have an important impact on chromatin biology. Efforts towards introducing multiple labile modifications in histones as well as expanding their use in cellular biology promise new powerful tools to study epigenetics.

Graphical abstract



Highlights

- Histones with specific PTMs can be synthesised and assembled in designer nucleosomes
- Codon suppression, semisynthesis or mutagenesis allow controlled insertion of PTMs
- Designer nucleosomes can probe stability and/or PTM specific interaction partners
- *In vivo* techniques show potential for locus-specific incorporation of PTM

Tags

Synthetic histones ; synthetic nucleosomes ; designer chromatin ; chemical modification ; histone post-translational modification ; genetic code expansion ; amber suppression ; protein semisynthesis ; native chemical ligation ; intein *trans* splicing ; split inteins ; chemical mutagenesis ; dehydroalanine ; tag-and-modify ; *in vivo* histone modification ; immobilised nucleosome pull-down ; mass spectrometry-based proteomics

Background and Motivation

Epigenetic mechanisms allow our cells to inherit changes in their gene expression and phenotype without changing the underlying DNA sequence. Epigenetic processes can happen at several layers of the nuclear organisation of the genome, from the 3D territorial positioning of interphase chromosomes in the nucleus down to the atomic level with histone post-translational modifications (PTMs) and nucleobase modifications (Figure 1.a in blue). [1,2] These epigenetic layers of regulation influence each other, thus understanding the role of histone post-translational modifications in this framework has been central to deciphering chromatin biology. [3,4]

Histone PTMs have a number of implications in regulatory tasks: specific modifications such as H3K14ac lead to transcriptional initiation [3], others such as H3K4me3 recruit specific transcription factors that have the required reader domain (i.e. plant homeodomain - PHD) [5]. Investigating histone PTMs poses many challenges: a plethora of PTMs have been discovered (Figure 1b), several distinct histone proteoforms are found in a single cell and histone tails can have multiple modifications in close proximity. Therefore, to study histone PTM access to large quantities of homogeneously modified histones (with site- and group- selectivity) is required – this is a difficult goal to attain with enzymatic techniques [6,7]. In order to address this ambitious challenge, synthetic and chemical biology techniques to generate ‘designer’ nucleosomes containing histones with precisely inserted modifications have emerged in the last decade [8-18].

Here, we review recent developments in chemical chromatin research; starting with identification processes of a given histone post-translational modification, moving to a comparison of methods for generating synthetic nucleosome constructs carrying the modification of interest (focusing on post-translational approaches) and closing with their use in the investigation of biological processes.

Searching for histone modifications using tandem mass spectrometry

Cellular chromatin consists of multiple nucleosome forms, with distinct histone proteoforms that can carry modifications in sometimes very low stoichiometry.[19] This large range of population distribution poses an analytical challenge for mass spectrometry (MS)-based proteomics, since both abundant marks and rare modifications have to be reliably identified – a prerequisite for knowing which modification to study with synthetic histones.

In a classical approach, when the sequence of histones has been previously well characterised, a bottom-up approach as depicted in Figure 1.c is chosen. Proteins of interest are digested with a

proteolytic enzyme such as trypsin and characterisation is performed on the subsequent pool of peptides. By bioinformatic comparison of the measured spectra with proteomic databases, peptides and their modifications are identified. In well-characterised samples, new modifications can be identified (unknown mass shift Δ mass on a given peptide). Accurate mass and peptide sequencing allow an empirical formula as well as chemical properties to be obtained – ultimately leading to the PTM structure. [20] Building on this well-established schematised workflow, broader application of proteomic methods in the analysis of histone modifications now include top-down proteomics, *de novo* identification of PTMs or quantitative experiments such as label-free quantification, multiple reaction monitoring (MRM) or isotopic tags. [19]

Mass spectrometry-based techniques have been useful in testing the efficacy of established chemical biology and chemical modification techniques and conversely these can enable mass spectrometry-based data acquisition. By probing the limits of the deleterious effects of cellularly-incorporated tags used in covalent retrieval methods (e.g. azidohomoalanine [21,22]) in histones on cell-cycle progression [23] their efficient use can be planned. Libraries of histone peptides carrying diverse post-translational modifications have allowed improved calibration of MS detection efficiencies and have allowed the effect upon wide-scale quantification experiments to be re-examined, leading in one example to a different biological conclusion. [24] Moreover, the development of selective enrichment techniques for modified peptides, such as TiO₂-mediated affinity chromatography for phosphopeptides further supported the identification of rare histone PTM in large proteomic screens.[25] These examples illustrate that both mass spectrometry-based proteomics and chemical biology techniques are not isolated tools used in histone PTM research, but are now progressing hand-in-hand.

The chemical biology toolkit

Once a modification has been identified and validated, several lines of investigation can be followed from biochemical assays to bioinformatic correlation analysis. One approach is to generate synthetic histones and assemble designer nucleosomes. Three main chemical biology strategies have been developed for this purpose that operate through a judicious mix of synthetic, co- and post-translational strategies: genetic code expansion, protein semi-synthesis and chemical modification / mutagenesis.

Genetic code expansion

Non-sense suppression mutagenesis is a powerful technology for introducing single unnatural amino acids (uAA) into cellular proteins. A stop codon, usually the constitutively rare amber stop codon (TAG), is re-attributed to the unnatural amino acid by incorporating both a cognate tRNA and an engineered aminoacyl tRNA synthetase (o-aaRS) into the host cell. The engineered synthetase can recognise the tRNA carrying the reattributed codon and load it with the uAA. Both need to be orthogonal to the host cell biosynthetic machinery, which ideally necessitates that the tRNA is not accepted by any constitutive synthetase and that the engineered synthetase only loads the orthogonal tRNA with the uAA. The cellular biosynthesis machinery is then able to incorporate the uAA for the amber codon into proteins being synthesised (Figure 2.a.). [26,27]

Engineering an aaRS-tRNA pair is thus a fine balancing act between creating orthogonality to the host cell and, at the same time, keeping the loaded tRNA structurally sufficiently close to constitutive tRNAs in the host cells so that the biosynthetic machinery (i.e. the ribosome) can recognise and use it. Often, aaRS-tRNA pairs from archaeal origins fulfil these criteria when used in prokaryotic expression systems. For the protein expression workhorse *Escherichia coli* orthogonal pairs based on PylRS-tRNA_{CUA}, TyrRS-tRNA_{CUA}, LeuRS-tRNA_{CUA} and TrpRS-tRNA_{CCA} are now routinely used.[28] Through engineering of corresponding synthetases, more than 160 structural analogues of amino acids bearing unnatural functional groups have been incorporated. At the same time the chemical diversity that can be reached by genetic code expansion remains dependent on the substrate recognition of the evolved synthetase and the ribosome, therefore the resulting chemical diversity typically remains structurally related to the original amino acids processed by those synthetases.[29]

Despite these limits, the PylRS-tRNA_{CUA} has been successfully used to site-specifically incorporate various naturally occurring PTMs into histones, in particular lysine PTM variant such as acyl-lysines (acetyl-[30], propionyl-[31], butyryl-[31] and crotonyllysine[32], 2-hydroxyisobutyryl-lysine [33], with special protocols for H4 [34]) or methylated lysines (methyl-Lys[35], dimethyl-Lys[36]), allowing the *in cellulo* and even the *in vivo* study of the effects of these modifications [37]. Moreover, precursors such as photoreactive lysine analogues (benzophenylalanine [38], photolysine [39], photocrotonyl-lysine [40], caged photolysine [41]) could also be incorporated, allowing their use in combination with post-translational chemical alteration (cross-linking, uncaging) and corresponding quantitative proteomics workflows leading to, for example, the identification of histone binding-factors in mammalian cells [39]. Recent results show that even more strategically challenging PTMs of other amino acids, such as phosphothreonine [42] and phosphoserine [43,44], can be incorporated into histones through the adaptation of other tRNA-synthetase pairs.

Current developments in genetic code expansion broadly diverge into two main directions of investigation, with potentially a strong impact on the study of histone PTMs. The first is to incorporate *in vivo* systems, described later in this review. The second aims at overcoming current limitations of synthetase-tRNA pairs. Screening technologies for the evolution of orthogonal aaRS have been designed [42,45] and biosynthetic re-engineering techniques of cells proposed, including deletion of termination factors and engineered ribosomes that recognise the amber codon with higher efficiency [28]. A particularly interesting innovation is the development of ribosomes recognising quadruplet codon (Figure 2a, right), allowing expansion beyond the restricted number of triplet codons and therefore potentially greater flexibility for the insertion of multiple uAA [46].

Protein semisynthesis

The discovery of native chemical ligation (NCL) [47] enabled a powerful unification of the world of shorter peptide synthesis with that of recombinant protein expression. [15,16] This can employ a polypeptide carrying a C-terminal thioester **1** with a recombinant protein/peptide containing an N-terminal cysteine **2** to generate a semisynthetic protein construct **4**, as described in Figure 2b. Suitable polypeptides carrying a C-terminal thioester can be generated by solid-phase peptide synthesis (SPPS) using specially designed safety-catch resins with linkers such as 3-thiopropionic acid or *N*-acylbenzimidazole.[6] Another possibility, referred to as expressed protein ligation (EPL), is to recombinantly express a construct of the protein of interest fused on its C-terminus to an intein domain. Inteins are proteins able to (self)splice themselves out of the fused construct to the target protein using a catalytic cysteine (or selenocysteine[48]) residue (*cis*-splicing). If this autoprocessing mechanism is performed in presence of intercepting thiols, then target proteins with a C-terminal thioester are generated.[6,16] NCL strategies have been used for the creation of synthetic histones for more than 15 years [8,49,50] and several different types of modifications have been incorporated, including pHis [51]. It has been the method of choice for histone ubiquitylation [52], a primary model to study *in vitro* protein-protein ligation chemistry. Ligation techniques including photo-cleavable thiol auxiliary **5** [52] and delta-mercaptolysine uAA **6** [53] have been subsequently developed to enable the traceless incorporation of whole protein modifications into nucleosomes (Figure 2.c).

One of the great strength of protein semisynthesis is the possibility to incorporate diverse modifications on several sites of the peptide, such as multiple glycosylations [54] or fully modified histone tails [55]. Elaborated schemes using several ligation steps can be used for a total chemical synthesis of histones; this allows for control over each residue in the protein chain. [55-59] In a

biological context, a large number of histone PTMs exert their full effect only in a concerted manner, for example by cooperative recruitment of transcription factors.[4] Semisynthetic strategies are thus particularly valuable to assess the multi-site PTM crosstalk that might constitute a so-called Histone Code.

The discovery of split inteins has further popularised protein semisynthesis, facilitating the assembly of protein constructs. The Muir group has recently published a technology based on ultrafast intein *trans* splicing in which the N-terminal part of the intein, Int^N, is fused to the C-terminus of the histone of interest and another protein construct containing the C-terminal part of the intein, Int^C, with a polypeptide tail carrying modifications is added. Upon association of Int^N and Int^C, the splicing reaction is triggered and the histone ligated with the polypeptide tail, creating a process that can be applied to *in cellulo* studies.[60] Further developments have seen improved inteins without extein dependence, allowing a splicing efficiency essentially independent of the construct context.[61,62] However, these splicing reactions are not yet quantitative and some drawbacks of inteins remain to be addressed, such the hydrolysis or thiolysis of inteins leading to N- and C-cleavage products. [61,63]

Both NCL and intein-based ligation strategies rely on solid-phase peptide synthesis (SPPS) to generate synthetic modifications or PTM-modified peptides and are thus subject to known SPPS limitations. These include the β -elimination of phosphorylated sites under basic conditions [64] and potentially extensive epimerisation of slow-coupling amino acids, such as glycosylated building blocks [65].

Chemical Mutagenesis

Altering protein residues by performing selective chemical protein modification reactions can allow at its extreme the creation of amino acids via that modification [66-68], a concept termed chemical mutagenesis. [68-70] Cysteine stands out among the proteinogenic amino acids by virtue of its unique chemistry, and its accessible pKa allowing generation of a more powerfully nucleophilic thiolate anion. [71] Moreover, its abundance in proteins (especially in histones) is low; in humans only H3 contains a cysteine (Cys110) that can be mutated to alanine without loss-of-function. [72] As such, use of cysteines is currently a central approach in the chemical modification of histones. In this section, we attempt to illustrate, using synthetic histone PTMs as examples, innovations in the field of chemical modification and how they have become more elaborate and effective as mimics (even approaching the goal of directly installing fully-native PTMs / chemical mutagenesis) over time.

Early modification studies targeted cysteine as a nucleophile by reacting it with haloalkyl reagent **8** to generate histones with methyl-lysine 'thialysine' analogues **9** (Figure 2.d). [73] Subsequently, reagent **10** for the generation of methyl-arginine analogues via conjugate addition was suggested. [74] Strategically similar cysteine-based chemistry for acetyl-lysine analogues **13** is more challenging; strain-promoted substitution reaction yielded the first promising results [75] and thiol-ene reactions using **12** (Figure 2d) prove reasonably efficient. [76] Bis-alkylation and elimination to dehydroalanine (Dha) allows the conversion of nucleophilic cysteinyl site to electrophilic Dha and has been exploited [77] through the addition of thiol nucleophiles to generate various PTM analogues, including thialysine variants, in histones. [72] Nevertheless, computational studies suggest that the thioether linkages in all such thialysines introduces structural distortions that may impact on their recognition (Box in Figure 2d-e). [78]

The generation of constitutionally native, modified lysine (and other) residues (chemical mutagenesis) via chemical modification could be envisaged through C-C bond formation. [69,70] Appropriate biorthogonal C_{sp3}-C_{sp3} bond formation reactions have recently been demonstrated using C• radical additions to Dha (Figure 2.e). [79,80] The procedure is performed in two steps: the first being the elimination (e.g. via bis-alkylation of cysteine) to Dha. Two main reagents for the formation of Dha on proteins are used: DBHDA [77] and MDBP, [81,82] allowing varied application to different target proteins and even regioselective modulation of elimination sites by reagent type (unpublished observations). The addition of a C-radical to Dha in the second step can be mediated by use of either NaBH₄ [79,83] or an elementary metal powder [79,80,83] with suitable radical precursors. Dha can also be generated from phosphoserine [80], phenylselenocysteine [84] or alkylselenocysteine [85] residues that can be incorporated by genetic code expansion, allowing useful flexibility in installation strategies and potentially multiplexing of epigenetic marks. Moreover, Dha formation on proteins containing selenocysteine can be performed site-selectively by exploiting differences in pK_a. [48]

Such post-translational chemical mutagenesis is a quite direct, economical and accessible method to synthetically incorporate constitutionally native PTMs into histones, under conditions that are sufficiently facile to be performed even at the histone sub-complex or nucleosome level. However, current mutagenesis protocols come at the cost of typically mixed configuration (D/L) at the α -carbon and the reaction currently does not yet efficiently allow for bond formation to aromatic residues. Nevertheless, promising approaches in peptidic substrates [86,87] suggest possible avenues in proteins.

Methodological diversity covers most PTMs

With more than 150 histone PTMs reported to date, techniques for generation of chemically-modified histones not only need to cover the broad chemical diversity in modifications, but also allow for a possible experimental high-throughput. Chemical mutagenesis experiments are based on a 'tag-and-modify' approach [88], where the addition reaction to a 'tag' such as Dha in a recombinantly expressed protein conceptually allows for subsequent combinatorial expansion via late-stage, divergent protein modification chemistry and hence for the incorporation of a broad range of functional groups. In strategies based on genetic code expansion for the incorporation of uAAs that demand development of tailored amino-acyl RNA synthetases only limited structural plasticity may be achieved, typically within quite confined limits. In more labour-intensive protein semisynthesis schemes requiring synthetic expertise, it can prove tough to match this broad accessibility and throughput despite the recent development of ingenious methods for the encoding (e.g. via DNA) of nucleosome libraries.[89] Yet both methods allow the circumvention of the current limitations encountered by chemical mutagenesis schemes, thereby increasing the chemical diversity of modifications that can be included in synthetic nucleosomes. As such, each method complements the other: non-sense codon suppression strategies are particularly useful for efficient incorporation of certain aromatic and lysine-type residues (often with ready transfer to cellular settings), protein-semisynthesis schemes allow for the incorporation of several modifications in the a single histone (an important feature when testing Histone Code hypotheses and/or recruitment of interaction partners based on modifications acting in concert) and chemical mutagenesis allows ready late stage diversification in a structurally broad manner (typically *in vitro*). In this spirit, suggestions for methods to introduce challenging modifications are presented in Table 1.

Whereas a range of techniques to chemically introduce post-translational modifications on histones have been described here separately, combinations of these three approaches are therefore most likely the way to achieve the diverse goals of chromatin chemical biology, including access in relevant contexts to many and large quantities of specifically and selectively modified histones and nucleosomes.

Application of chemical biology tools for understanding chromatin biology

The strikingly diverse chemical space of PTMs modulates the structure and function of histone proteins in a spatiotemporal manner, and thereby plays a fundamental role in chromatin regulation. This is governed through two different yet often coupled mechanisms.[3,4] The first involves the

modulation of nucleosome (and, thus, chromatin) structure due to the perturbation of histone-fold or histone-DNA interactions upon histone modification.[90] In the second mechanism, histone PTMs acts as molecular beacons for specific recruitment of 'reader' or 'effector' proteins to mediate downstream functions.[5,91] Earlier chromatin biology studies relied on the use of native nucleosome and chromatin templates, bearing undefined mixtures of histone marks, thus limiting the precise understanding of distinct histone marks in epigenetic regulation. The maturation of various chemical toolkits leading to the ready synthesis of defined histone proteins and designer nucleosomes, in conjugation with various biochemical and biophysical studies, has allowed us to overcome that limitation by being able to establish initial causal relationships between histone marks and chromatin dynamics and function.

Probing the modulation of chromatin structure

Charge-modulating histone PTMs such as acetylation, phosphorylation, and ADP-ribosylation can potentially alter electrostatic histone-DNA interactions, resulting in modulation of nucleosome structure and stability.[90] Indeed, acetylation of lysine residues is generally believed to inhibit stabilizing interactions between lysines and nucleosomal DNA leading to a 'relaxed' chromatin state.[92] A recent study using synthetic acetylated nucleosome revealed that modification of H3-Lys64 residue lowers the stability of the nucleosome due to reduced histone-DNA interaction.[93] The effect of histone phosphorylation on nucleosome structure has also been analysed; phosphorylation of H3-Thr118 leads to apparent formation of two suggested non-canonical nucleosome structures - so-called 'nucleosome duplex' and 'altosome complex' - whose exact physiological roles are yet to be determined [94] Elegant use of biomolecular NMR with synthetic phosphorylated H3 variants showed that the dynamics of the H3 tail increase upon charge-modulation, altering transient contacts to the DNA and enhancing its general modifiability.[95]

Similar to these charge-modulating histone PTMs, neutral modifications may also regulate chromatin architecture (Figure 3.a). Synthetic mimicking of GlcNAcylation of H2A-Thr101 appears to disrupt histone-fold interactions leading to destabilized nucleosome structure.[96] Acetylation and ubiquitination of H4-Lys91 regulates nucleosome structure by perturbing the dimer-tetramer interface of histone octamer [97,98] primarily via steric rather than electrostatic effects. Intriguingly, even relatively small modifications such as histone methylation can apparently influence chromatin structure. Casadio *et al.* used a semisynthetic approach to demonstrate that asymmetric dimethylation of H3-Arg42 residue lowers nucleosome stability and stimulates transcription.[99] This destabilization of the nucleosome was attributed to the change in histone-DNA interaction upon loss of potential hydrogen bond. Similarly, dimethylation of H3-Lys79 results in a subtle structural

reorganization leading to increased accessibility of the modified residue apparently through loss of a hydrogen bond.[100]

Probing interaction partners and downstream spreading of modifications

In addition to elucidating the structural effect of histone PTMs on chromatin dynamics, synthetic histone proteins and designer nucleosomes have also allowed identification of key interacting protein partners (Figure 3.b.). In seminal work, Bartke *et al.* used designer nucleosomes bearing modifications of nucleosomal DNA and H3 protein at Lys-4, Lys-9, and Lys-27 to affinity-enrich nucleosome-binding proteins from nuclear extracts.[101] The use of quantitative mass spectrometry allowed them to identify the preferential protein partners for individual as well as combinations of epigenetic marks in an unbiased manner. Rapid emergence of quantitative proteomics fueled by the advancement of mass spectrometers in terms of sensitivity, sequencing efficiency/speed, and mass accuracy coupled with new quantitative approaches has allowed this approach to now become a mainstay for identification of chromatin-binding proteins, over traditional western-blotting based methods.[19] This approach was further extended by using a defined oligo-nucleosome template for affinity-enrichment of binding protein partners.[102] A similar approach using synthetic GlcNAcylated nucleosome has allowed the identification of sub-units of FACT complex as interacting protein partners for nucleosomes GlcNAcylated at H2B-Ser112 [103] and led to a suggested updating of prior molecular mechanisms for GlcNAcylation-driven transcriptional activation that had been determined using only isolated GlcNAcylated-H2B protein. [104] The latter example highlights that more representative nucleosomal context may prove critical and that synthetic chemical biology should aim to build and use such higher order structures as preferred probes.

Towards *in vivo* chromatin chemical biology

In vitro chromatin biology experiments have provided valuable insights, yet these are reductionist approaches with the obvious possibility of consequent artefacts – i.e. key players might be missing from any reconstituted or probed system. Therefore, much effort is rightly focused on developing *in vivo* chromatin biology techniques [105].

Highly promising developments are occurring in the field of genetic code expansion. Non-sense suppression mutagenesis has been used for histones in mammalian cells providing exciting insights into lysine acetylation [37,41]. Given the recently claimed expansion of the technique to mouse as a model organism [106], more complex organismal chromatin biology might be expected to follow soon. Moreover, high affinity ultrafast split inteins now allow for the incorporation of specific modifications inside cells (Figure 3c).[60] Recently, this technique has been further (most

impressively) improved for locus-specific *in vivo* chromatin tailoring. Using the dCas9 targeting domains, Muir and co-workers were able to direct synthetic modification towards locations in a cellular genome (Figure 3d). [107] Such *in cellulo* host-guest-directed methods suggest modes of control that may even eventually allow use of other ligand-directed, site-selective methods; ligand-directed chemical protein acetylation can be performed on native nucleosomes *in vitro* [108] and could provide a good candidate small-molecule system for such *in vivo* transposition.

Outlook

In the last decade, synthetic chemical biology techniques have emerged that provide tools for the precise molecular investigation of chromatin biology. Whereas the *in vitro* generation of synthetic histones and designer nucleosomes can readily be achieved in complementary ways using genetic code expansion, protein semisynthesis or chemical mutagenesis, transposing these techniques to suitable *in vivo* applications remains one of the key challenges remaining in synthetic histone modification. Given the central role of dynamically- and diversely- altered higher order structures in epigenetics, the aim ultimately is to reveal, study and/or control epigenetic modification in its proper context both spatially and temporarily. It seems likely that for us to truly compete with the late-stage diversity generated by nature, then late-stage chemical tools and methods will prove important.

Conflicts of Interest

B.G.D. is editor-in-chief of Current Opinion in Chemical Biology. B.G.D. is inventor on patent WO 2009103941 filed by the University of Oxford assigned to Catalent Biologics that covers, in one claim, reaction of the carbon-carbon double bond of dehydroalanine residues reacted with an organohalide in the presence of elemental zinc. B.G.D. is a member of the Catalent Biologics Scientific Advisory Board. B.G.D. is a supplier of the DBHDA reagent via the Kerafast platform.

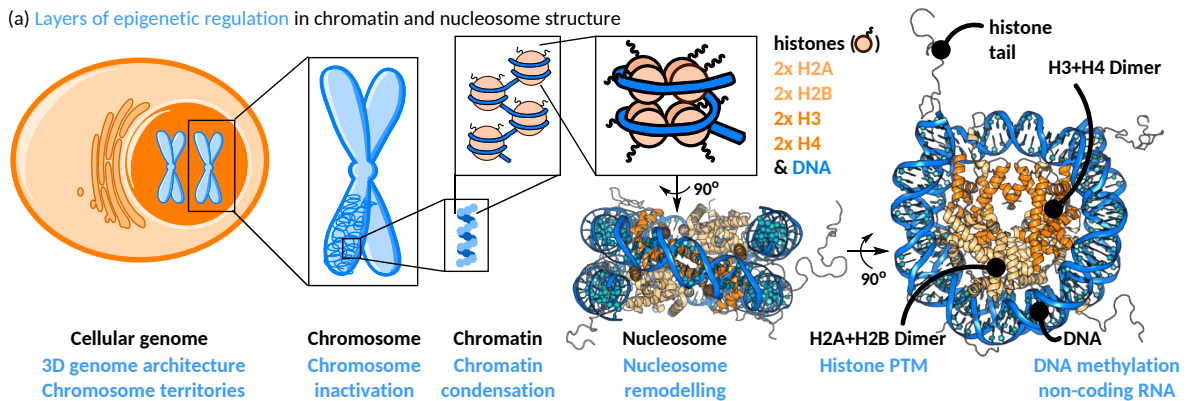
Acknowledgements

The authors thank Dr. Jitka Dadova for helpful discussions. This work was supported by the Felix Foundation (R.R) and EPSRC (S.N.).

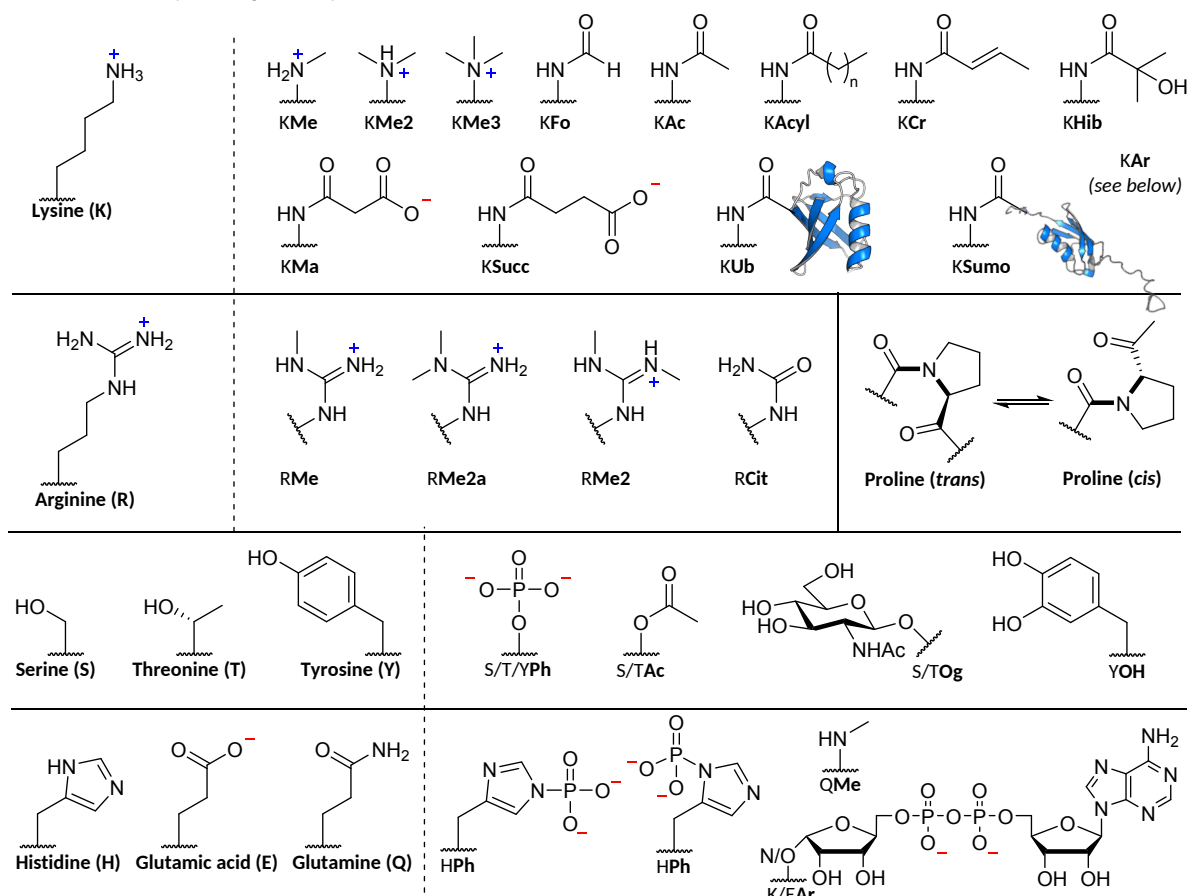
Figures

Figure 1

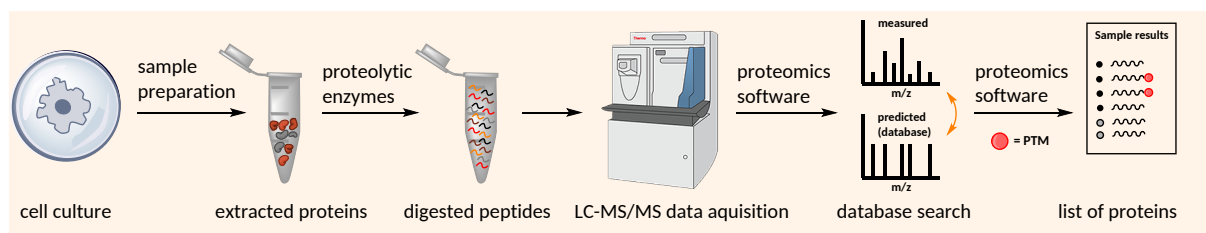
(a) Layers of epigenetic regulation in chromatin and nucleosome structure



(b) Selected naturally occurring histone post-translational modifications (PTM)



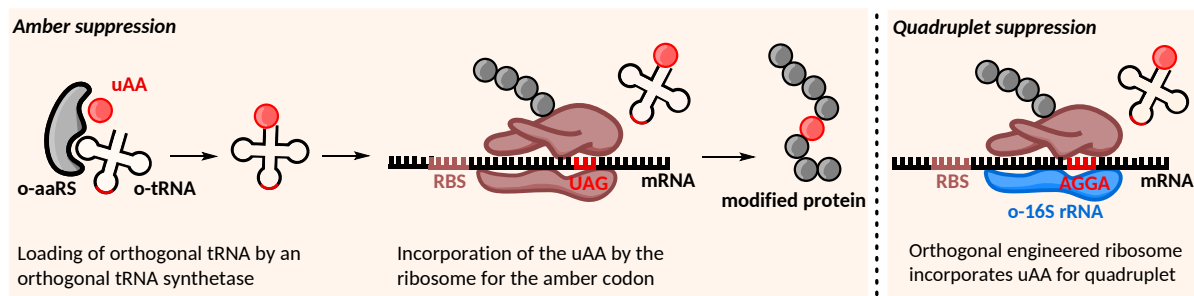
(c) Mass spectrometry-based proteomics for the identification of histone post-translational modifications



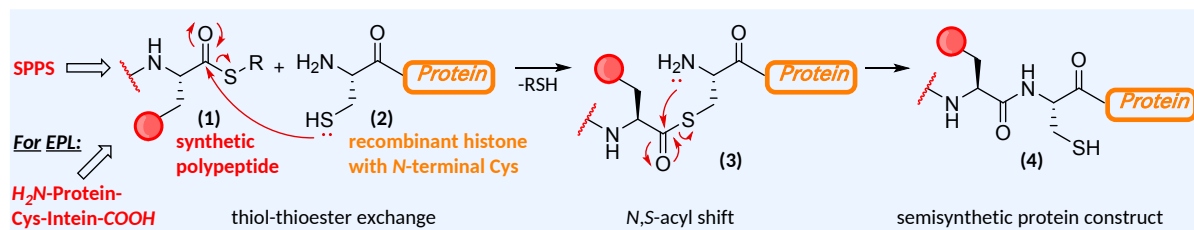
Post-translational modification of histones and their identification using mass-spectrometry based proteomics. **(a)** The packaging of cellular DNA into chromosomes and nucleosomes allows for several layers of epigenetic regulation. One of these is the post-translational modification of histone proteins, core components of the nucleosome. PDB: 1KX5 [109]. **(b)** Selected commonly observed histone post-translational modifications (PTMs), listed by amino acid site of modification, illustrate the broad range in chemistries found on histones. Adapted from ref [19]. **(c)** Histone PTMs can be identified by mass-spectrometry based proteomics. In this bottom-up proteomics workflow (based on digested proteolytic peptides) identification of known PTM is described. Ac - acetylation, Acyl - acylation, Ar - ADP-ribosylation, Cit - citrullination, Cr - crotonylation, Fo - formylation, Hib - 2-hydroxyisobutyrylation, Ma - malonylation, Me - methylation, Og - O-GlcNAcylation, OH - hydroxylation, Ph - phosphorylation, Succ - succinylation, Sumo - SUMOylation, Ub - ubiquitinylation.

Figure 2

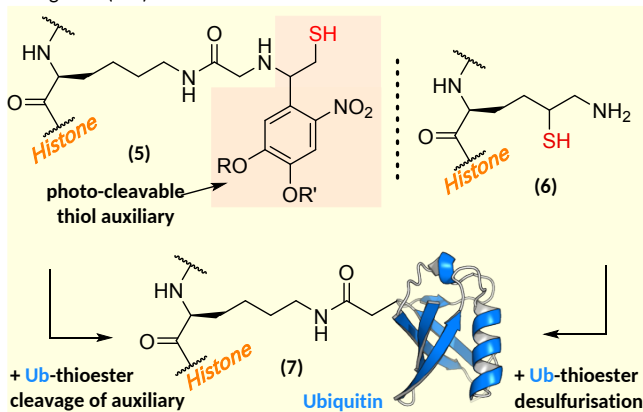
(a) Synthetic histones by genetic code expansion



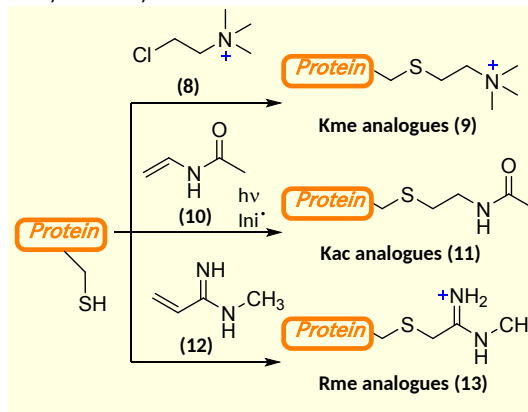
(b) Synthetic histones by protein semisynthesis using native chemical ligation (NCL)



(c) Synthetic histones by coupling to a full protein via Expressed Protein Ligation (EPL)

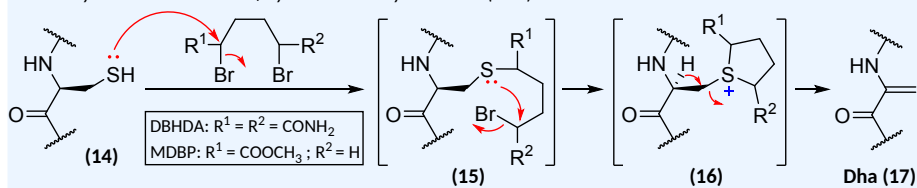


(d) Synthetic histones by chemical mutagenesis through cysteine alkylation

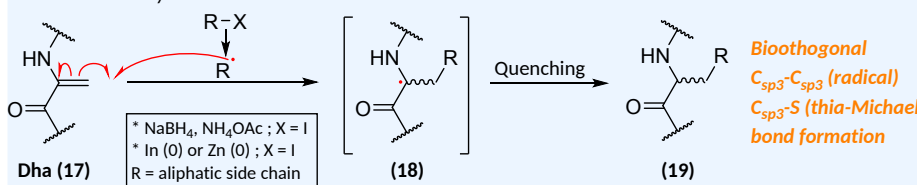


(e) Synthetic histones by dehydroalanine-mediated (Dha) chemical mutagenesis - 'tag-and-modify' approach

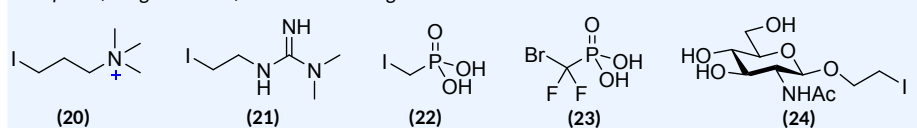
1. Bis-alkylation-elimination of cysteine to dehydroalanine (Dha)



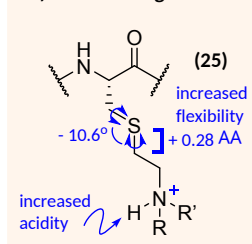
2. Addition to dehydroalanine



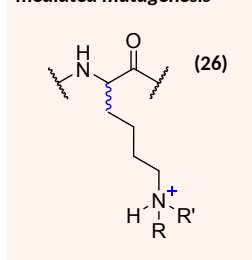
Examples of reagents used for chemical mutagenesis on histones:



Structural changes in sulfinated analogues



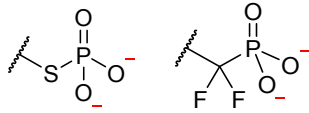
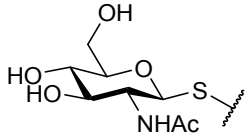
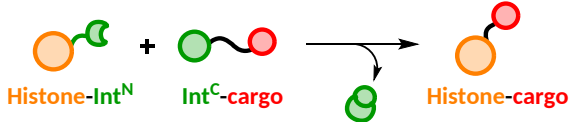
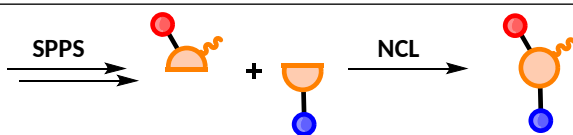
Loss of stereocenter in Dha-mediated mutagenesis



The chemical biology toolkit for synthetic histone modification. **(a)** Genetic code expansion. Amber codon (TAG) suppression is the most widely used technique for targeted genetic incorporation. Another possibility involves the use of an engineered ribosome with an orthogonal 16S ribosomal unit for quadruplet codon recognition. Adapted from [45]. **(b)** Mechanism of native chemical ligation (NCL) for protein semisynthesis. Using this technique, a protein with an *N*-terminal cysteine can be coupled to a synthetic peptide with a C-terminal thioester. **(c)** Expressed protein ligation can be used to couple proteins such as ubiquitin to peptides/proteins containing thiol auxiliaries **5** or the amino acid δ -mercaptolysine **6**. These approaches can be made traceless by UV-irradiation or desulfurisation. Adapted from [14]. **(d)** Historically, thioether-based PTM mimics were used by reacting a recombinant histone containing a cysteine at the desired position with thiol-alkylating reagents. These mimics however bring structural changes with them (frame below, adapted from [16]). **(e)** Many histone PTMs can be accessed using chemical mutagenesis, in a two step protocol. First, elimination of a recombinantly incorporated cysteine at the desired position of modification to dehydroalanine is performed using either the DBHDA or MDBP reagent. Unfortunately, this comes at the cost of mixed D/L configuration at the alpha-carbon. Second, addition to dehydroalanine is performed. Depending on the type of bond formation required, a radical C-C bond formation or thia-Michael reaction can be chosen; this representation emphasises the recently reported radical mechanism.

o-aaRS – orthogonal aminoacyl tRNA synthetase, uAA – unnatural amino acid, RBS – ribosomal binding site.

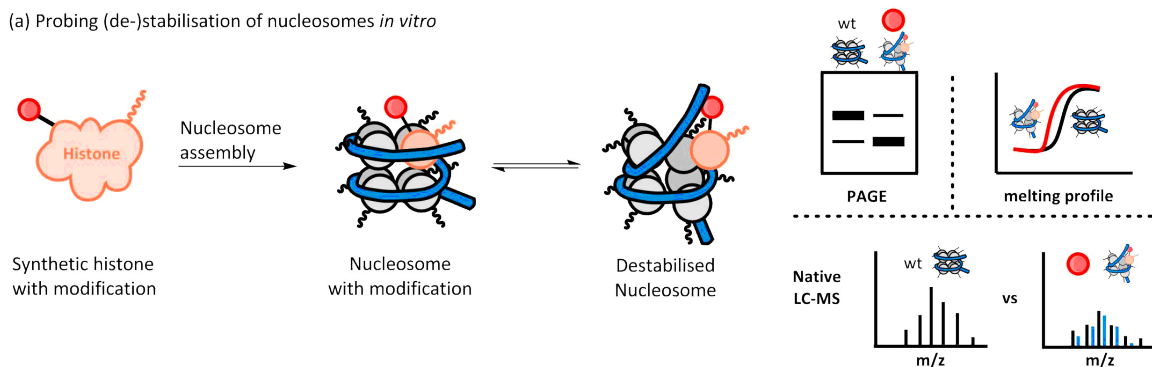
Table 1

Modification	Suggested incorporation technique	Scheme & notes
Phosphorylation	Chemical mutagenesis	 <ul style="list-style-type: none"> • non-hydrolysable, phosphatase-resistant • CF₂ is isoelectronic to natural phosphoester
O-GlcNAc glycosylation	Chemical mutagenesis	 <ul style="list-style-type: none"> • Same conformation as natural acetal linkage[110] • Resistant to hOGA glycolytic activity
Protein coupling	Native chemical ligation	<ul style="list-style-type: none"> • See Figure 2.c
Multiply modified histone tails	Intein trans-splicing or expressed protein ligation	
Multiply modified core histones	Histone total synthesis with NCL schemes	

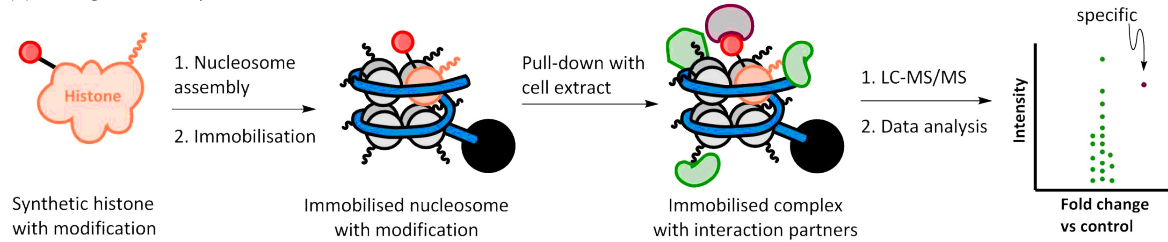
Challenging histone PTMs and illustrative current incorporation techniques.

Figure 3

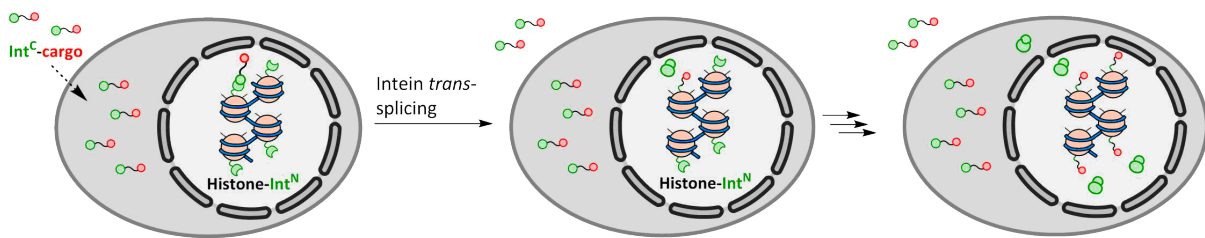
(a) Probing (de-)stabilisation of nucleosomes *in vitro*



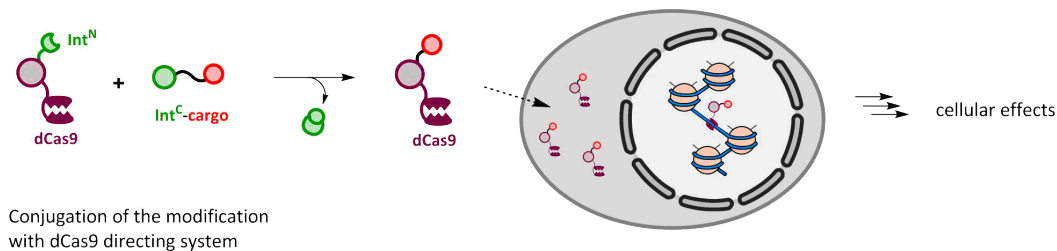
(b) Probing interaction partners *in-situ*



(c) Integrating modified nucleosomes *in vivo*



(d) Locus-specific insertion of a synthetic modification *in vivo*



Application of chemical biology tools for understanding chromatin biology. **(a)** Synthetic histone modifications can be used to probe their effects on nucleosome stability and chromatin architecture. **(b)** Synthetic histones can be assembled with other canonical histone proteins and biotinylated DNA to a designer nucleosome. These immobilized synthetic nucleosomes can be used for affinity enrichment of nucleosome-binding proteins from nuclear extract. Pooled proteins from sample and unmodified control are enzymatically digested and identified by LC-MS/MS analysis. The identified protein partners are quantified against each other using MS-based Label-free Quantification (LFQ). **(c)** Ultrafast split intein technology allows for the *in vivo* modification of chromatin. Using a pair with the Int^N genetically fused to histones and the synthetic Int^C exogenously added, protein semi-

synthesis can be performed *in vivo*, thus incorporating the desired modification. **(d)** Using split inteins, a synthetic construct of a histone bearing the target modification and the Cas9 directing system can be generated. Upon incorporation of the fusion construct into cells, the modified histone is delivered to the target location in the cellular genome, thus influencing epigenetic processes in its close environment. (Adapted from [105]).

References and recommended reading

- of special Interest
- of outstanding interest

1. Pombo A, Dillon N: **Three-dimensional genome architecture: players and mechanisms.** *Nature Reviews Molecular Cell Biology* 2015, **16**:245.
2. Fraser J, Williamson I, Bickmore WA, Dostie J: **An Overview of Genome Organization and How We Got There: from FISH to Hi-C.** *Microbiology and Molecular Biology Reviews* 2015, **79**:347-372.
3. Bannister AJ, Kouzarides T: **Regulation of chromatin by histone modifications.** *Cell Research* 2011, **21**:381.
4. Kouzarides T: **Chromatin Modifications and Their Function.** *Cell* 2007, **128**:693-705.
5. Musselman CA, Lalonde M-E, Côté J, Kutateladze TG: **Perceiving the epigenetic landscape through histone readers.** *Nature Structural & Molecular Biology* 2012, **19**:1218.
6. Dhall A, Chatterjee C: **Chemical approaches to understand the language of histone modifications.** *ACS Chem Biol* 2011, **6**:987-999.
7. Allis CD, Muir TW: **Spreading chromatin into chemical biology.** *ChemBiochem* 2011, **12**:264-279.
8. Shogren-Knaak MA, Peterson CL: **Creating Designer Histones by Native Chemical Ligation.** In *Methods in Enzymology*. Edited by: Academic Press; 2003:62-76. Chromatin and Chromatin Remodeling Enzymes, Part A, vol 375.]
9. Schwarzer D: **Chemical tools in chromatin research.** *J Pept Sci* 2010, **16**:530-537.
10. Fierz B, Muir TW: **Chromatin as an expansive canvas for chemical biology.** *Nat Chem Biol* 2012, **8**:417-427.
11. Nikolov M, Fischle W: **Systematic analysis of histone modification readout.** *Mol Biosyst* 2013, **9**:182-194.
12. Fierz B: **Synthetic chromatin approaches to probe the writing and erasing of histone modifications.** *ChemMedChem* 2014, **9**:495-504.
13. Pick H, Kilic S, Fierz B: **Engineering chromatin states: chemical and synthetic biology approaches to investigate histone modification function.** *Biochim Biophys Acta* 2014, **1839**:644-656.
14. Fischle W, Mootz HD, Schwarzer D: **Synthetic histone code.** *Curr Opin Chem Biol* 2015, **28**:131-140.
15. Holt M, Muir T: **Application of the protein semisynthesis strategy to the generation of modified chromatin.** *Annu Rev Biochem* 2015, **84**:265-290.
16. Muller MM, Muir TW: **Histones: at the crossroads of peptide and protein chemistry.** *Chem Rev* 2015, **115**:2296-2349.

- This review comprehensively describes the broad scope of histone peptide chemistry as well as histone protein semisynthesis.
17. Chuh KN, Batt AR, Pratt MR: **Chemical Methods for Encoding and Decoding of Posttranslational Modifications**. *Cell Chem Biol* 2016, **23**:86-107.
 18. Yang A, Cho K, Park H-S: **Chemical biology approaches for studying posttranslational modifications**. *RNA Biology* 2017:1-14.
 19. Huang H, Lin S, Garcia BA, Zhao Y: **Quantitative Proteomic Analysis of Histone Modifications**. *Chemical Reviews* 2015, **115**:2376-2418.
 - A comprehensive and still up-to-date review of mass spectrometry-based proteomics for the identification and discovery of histone PTM.
 20. Dai L, Peng C, Montellier E, Lu Z, Chen Y, Ishii H, Debernardi A, Buchou T, Rousseaux S, Jin F, et al.: **Lysine 2-hydroxyisobutyrylation is a widely distributed active histone mark**. *Nat Chem Biol* 2014, **10**:365.
 21. Howden AJM, Geoghegan V, Katsch K, Efstathiou G, Bhushan B, Boutureira O, Thomas B, Trudgian DC, Kessler BM, Dieterich DC, et al.: **QuaNCAT: quantitating proteome dynamics in primary cells**. *Nature Methods* 2013, **10**:343.
 22. Dieterich DC, Link AJ, Graumann J, Tirrell DA, Schuman EM: **Selective identification of newly synthesized proteins in mammalian cells using bioorthogonal noncanonical amino acid tagging (BONCAT)**. *Proc Natl Acad Sci, U.S.A.* 2006, **103**:9482-9487.
 23. Arnaudo AM, Link AJ, Garcia BA: **Bioorthogonal Chemistry for the Isolation and Study of Newly Synthesized Histones and Their Modifications**. *ACS Chem Biol* 2016, **11**:782-791.
 24. Lin S, Wein S, Gonzales-Cope M, Otte GL, Yuan Z-F, Afjehi-Sadat L, Maile T, Berger SL, Rush J, Lill JR, et al.: **Stable-isotope-labeled Histone Peptide Library for Histone Post-translational Modification and Variant Quantification by Mass Spectrometry**. *Molecular & Cellular Proteomics* 2014, **13**:2450-2466.
 - Using a library of synthetic histone peptides carrying defined PTMs, the effect of these PTMs on peptide detection and identification is investigated. The findings are food for thought on quantitative histone PTM experiments.
 25. Kulej K, Avgousti DC, Weitzman MD, Garcia BA: **Characterization of histone post-translational modifications during virus infection using mass spectrometry-based proteomics**. *Methods* 2015, **90**:8-20.
 26. Lang K, Chin JW: **Cellular Incorporation of Unnatural Amino Acids and Bioorthogonal Labeling of Proteins**. *Chemical Reviews* 2014, **114**:4764-4806.
 27. Wang L, Brock A, Herberich B, Schultz PG: **Expanding the Genetic Code of *Escherichia coli***. *Science* 2001, **292**:498-500.
 28. Chin JW: **Expanding and reprogramming the genetic code**. *Nature* 2017, **550**:53-60.
 29. Dumas A, Lercher L, Spicer CD, Davis BG: **Designing logical codon reassignment - Expanding the chemistry in biology**. *Chemical Science* 2015, **6**:50-69.

- This review contains a list of all uAAs that have been successfully included by genetic code expansion (by 2014) which clearly shows the structural analogy of the uAA to the original amino acid of the synthetase.
30. Neumann H, Peak-Chew SY, Chin JW: **Genetically encoding N(epsilon)-acetyllysine in recombinant proteins.** *Nat Chem Biol* 2008, **4**:232-234.
 31. Gattner MJ, Vrabel M, Carell T: **Synthesis of epsilon-N-propionyl-, epsilon-N-butyryl-, and epsilon-N-crotonyl-lysine containing histone H3 using the pyrrolysine system.** *Chemical Communications* 2013, **49**:379-381.
 32. Kim CH, Kang M, Kim HJ, Chatterjee A, Schultz PG: **Site-Specific Incorporation of ε-N-Crotonyllysine into Histones.** *Angewandte Chemie International Edition* 2012, **51**:7246-7249.
 33. Knight WA, Cropp TA: **Genetic encoding of the post-translational modification 2-hydroxyisobutyryl-lysine.** *Organic & Biomolecular Chemistry* 2015, **13**:6479-6481.
 34. Wilkins BJ, Hahn LE, Heitmüller S, Frauendorf H, Valerius O, Braus GH, Neumann H: **Genetically Encoding Lysine Modifications on Histone H4.** *ACS Chemical Biology* 2015, **10**:939-944.
 35. Nguyen DP, Garcia Alai MM, Kapadnis PB, Neumann H, Chin JW: **Genetically Encoding Ne-Methyl-l-lysine in Recombinant Histones.** *Journal of the American Chemical Society* 2009, **131**:14194-14195.
 36. Nguyen DP, Alai MMG, Virdee S, Chin JW: **Genetically Directing ε-N, N-Dimethyl-l-Lysine in Recombinant Histones.** *Chemistry & Biology* 2010, **17**:1072-1076.
 37. Elsasser SJ, Ernst RJ, Walker OS, Chin JW: **Genetic code expansion in stable cell lines enables encoded chromatin modification.** *Nat Meth* 2016, **13**:158-164.
 - The authors could generate stable cell lines for amber suppression mutagenesis containing the PylRS/tRNA pair. This system was used to investigate the effects of H3 lysine acetylation *in vivo*.
 38. Chin JW, Martin AB, King DS, Wang L, Schultz PG: **Addition of a photocrosslinking amino acid to the genetic code of Escherichia coli.** *Proceedings of the National Academy of Sciences* 2002, **99**:11020-11024.
 39. Yang T, Li X-M, Bao X, Fung YME, Li XD: **Photo-lysine captures proteins that bind lysine post-translational modifications.** *Nat Chem Biol* 2015, **12**:70.
 40. Xie X, Li X-M, Qin F, Lin J, Zhang G, Zhao J, Bao X, Zhu R, Song H, Li XD, et al.: **Genetically Encoded Photoaffinity Histone Marks.** *Journal of the American Chemical Society* 2017, **139**:6522-6525.
 41. Wilkins BJ, Rall NA, Ostwal Y, Kruitwagen T, Hiragami-Hamada K, Winkler M, Barral Y, Fischle W, Neumann H: **A Cascade of Histone Modifications Induces Chromatin Condensation in Mitosis.** *Science* 2014, **343**:77-80.
 42. Zhang MS, Brunner SF, Huguenin-Dezot N, Liang AD, Schmied WH, Rogerson DT, Chin JW: **Biosynthesis and genetic encoding of phosphothreonine through parallel selection and deep sequencing.** *Nat Meth* 2017, **14**:729-736.
 - Through parallel positive selection and sequencing the authors present a scalable approach to the discovery of orthogonal synthetases.

43. Park H-S, Hohn MJ, Umehara T, Guo L-T, Osborne EM, Benner J, Noren CJ, Rinehart J, Söll D: **Expanding the Genetic Code of *Escherichia coli* with Phosphoserine.** *Science* 2011, **333**:1151-1154.
44. Lee S, Oh S, Yang A, Kim J, Söll D, Lee D, Park H-S: **A Facile Strategy for Selective Incorporation of Phosphoserine into Histones.** *Angewandte Chemie International Edition* 2013, **52**:5771-5775.
45. Des Soye BJ, Patel JR, Isaacs FJ, Jewett MC: **Repurposing the translation apparatus for synthetic biology.** *Current Opinion in Chemical Biology* 2015, **28**:83-90.
46. Sachdeva A, Wang K, Elliott T, Chin JW: **Concerted, Rapid, Quantitative, and Site-Specific Dual Labeling of Proteins.** *Journal of the American Chemical Society* 2014, **136**:7785-7788.
47. Dawson P, Muir T, Clark-Lewis I, Kent S: **Synthesis of proteins by native chemical ligation.** *Science* 1994, **266**:776-779.
48. Liu J, Chen Q, Rozovsky S: **Utilizing Selenocysteine for Expressed Protein Ligation and Bioconjugations.** *Journal of the American Chemical Society* 2017, **139**:3430-3437.
49. He S, Bauman D, Davis JS, Loyola A, Nishioka K, Gronlund JL, Reinberg D, Meng F, Kelleher N, McCafferty DG: **Facile synthesis of site-specifically acetylated and methylated histone proteins: reagents for evaluation of the histone code hypothesis.** *Proc Natl Acad Sci U S A* 2003, **100**:12033-12038.
50. Shogren-Knaak MA, Fry CJ, Peterson CL: **A Native Peptide Ligation Strategy for Deciphering Nucleosomal Histone Modifications.** *Journal of Biological Chemistry* 2003, **278**:15744-15748.
51. Kee J-M, Villani B, Carpenter LR, Muir TW: **Development of Stable Phosphohistidine Analogues.** *Journal of the American Chemical Society* 2010, **132**:14327-14329.
52. McGinty RK, Kim J, Chatterjee C, Roeder RG, Muir TW: **Chemically ubiquitylated histone H2B stimulates hDot1L-mediated intranucleosomal methylation.** *Nature* 2008, **453**:812-816.
53. Ajish Kumar KS, Haj-Yahya M, Olschewski D, Lashuel HA, Brik A: **Highly Efficient and Chemoselective Peptide Ubiquitylation.** *Angewandte Chemie International Edition* 2009, **48**:8090-8094.
54. Siman P, Brik A: **Chemical and semisynthesis of posttranslationally modified proteins.** *Org Biomol Chem* 2012, **10**:5684-5697.
55. Maity SK, Jbara M, Brik A: **Chemical and semisynthesis of modified histones.** *J Pept Sci* 2016, **22**:252-259.
56. Jbara M, Maity SK, Brik A: **Palladium in the Chemical Synthesis and Modification of Proteins.** *Angewandte Chemie International Edition* 2017, **56**:10644-10655.
57. Maity SK, Jbara M, Mann G, Kamnesky G, Brik A: **Total chemical synthesis of histones and their analogs, assisted by native chemical ligation and palladium complexes.** *Nat. Protocols* 2017, **12**:2293-2322.
58. Jbara M, Maity SK, Brik A: **Palladium in the Chemical Synthesis and Modification of Proteins.** *Angew Chem Int Ed Engl* 2017, **56**:10644-10655.

59. Bondalapati S, Jbara M, Brik A: **Expanding the chemical toolbox for the synthesis of large and uniquely modified proteins.** *Nat Chem* 2016, **8**:407-418.
60. David Y, Vila-Perello M, Verma S, Muir TW: **Chemical tagging and customizing of cellular chromatin states using ultrafast trans-splicing inteins.** *Nat Chem* 2015, **7**:394-402.
61. Stevens AJ, Sekar G, Shah NH, Mostafavi AZ, Cowburn D, Muir TW: **A promiscuous split intein with expanded protein engineering applications.** *Proceedings of the National Academy of Sciences* 2017, **114**:8538-8543.
62. Gramespacher JA, Stevens AJ, Thompson RE, Muir TW: **Improved Protein Splicing Using Embedded Split Inteins.** *Protein Sci* 2018, DOI: 10.1002/pro.3357
63. Zettler J, Schütz V, Mootz HD: **The naturally split Npu DnaE intein exhibits an extraordinarily high rate in the protein trans-splicing reaction.** *FEBS Letters* 2009, **583**:909-914.
64. Behrendt R, White P, Offer J: **Advances in Fmoc solid-phase peptide synthesis.** *Journal of Peptide Science* 2016, **22**:4-27.
65. Zhang Y, Muthana SM, Farnsworth D, Ludek O, Adams K, Barchi JJ, Gildersleeve JC: **Enhanced Epimerization of Glycosylated Amino Acids During Solid-Phase Peptide Synthesis.** *Journal of the American Chemical Society* 2012, **134**:6316-6325.
66. Zioudrou C, Wilchek M, Patchornik A: **Conversion of the L-Serine Residue to an L-Cysteine Residue in Peptides.** *Biochemistry* 1965, **4**:1811-1822.
67. Polgar L, Bender ML: **A New Enzyme Containing a Synthetically Formed Active Site. Thiol-Subtilisin1.** *Journal of the American Chemical Society* 1966, **88**:3153-3154.
68. Neet KE, Koshland DE: **The conversion of serine at the active site of subtilisin to cysteine: a "chemical mutation".** *Proceedings of the National Academy of Sciences of the United States of America* 1966, **56**:1606-1611.
69. Chalker JM, Davis BG: **Chemical mutagenesis: selective post-expression interconversion of protein amino acid residues.** *Current Opinion in Chemical Biology* 2010, **14**:781-789.
70. Wright TH, Vallée MRJ, Davis BG: **From Chemical Mutagenesis to Post-Expression Mutagenesis: A 50 Year Odyssey.** *Angewandte Chemie International Edition* 2016, **55**:5896-5903.
71. Gunnoo SB, Madder A: **Chemical Protein Modification through Cysteine.** *Chembiochem* 2016, **17**:529-553.
72. Chalker JM, Lercher L, Rose NR, Schofield CJ, Davis BG: **Conversion of cysteine into dehydroalanine enables access to synthetic histones bearing diverse post-translational modifications.** *Angew Chem Int Ed Engl* 2012, **51**:1835-1839.
73. Simon MD, Chu F, Racki LR, de la Cruz CC, Burlingame AL, Panning B, Narlikar GJ, Shokat KM: **The site-specific installation of methyl-lysine analogs into recombinant histones.** *Cell* 2007, **128**:1003-1012.

74. Le DD, Cortesi AT, Myers SA, Burlingame AL, Fujimori DG: **Site-specific and regiospecific installation of methylarginine analogues into recombinant histones and insights into effector protein binding.** *J Am Chem Soc* 2013, **135**:2879-2882.
75. Huang R, Holbert MA, Tarrant MK, Curtet S, Colquhoun DR, Dancy BM, Dancy BC, Hwang Y, Tang Y, Meeth K, et al.: **Site-Specific Introduction of an Acetyl-Lysine Mimic into Peptides and Proteins by Cysteine Alkylation.** *Journal of the American Chemical Society* 2010, **132**:9986-9987.
76. Li F, Allahverdi A, Yang R, Lua GB, Zhang X, Cao Y, Korolev N, Nordenskiöld L, Liu CF: **A direct method for site-specific protein acetylation.** *Angew Chem Int Ed Engl* 2011, **50**:9611-9614.
77. Chalker JM, Gunnoo SB, Boutureira O, Gerstberger SC, Fernández-González M, Bernardes GJL, Griffin L, Hailu H, Schofield CJ, Davis BG: **Methods for converting cysteine to dehydroalanine on peptides and proteins.** *Chemical Science* 2011, **2**:1666.
- This study introduces the DBHDA reagent for conversion of cysteine to dehydroalanine on native proteins.
78. Seeliger D, Soeroes S, Klingberg R, Schwarzer D, Grubmüller H, Fischle W: **Quantitative Assessment of Protein Interaction with Methyl-Lysine Analogues by Hybrid Computational and Experimental Approaches.** *ACS Chemical Biology* 2012, **7**:150-154.
- This study represents a rather unique comparison of experimental and computational data. Using a combination of molecular dynamics simulations, based on a state-of-the-art quantum mechanical parametrisation of alkylated cysteine analogues for methyllysine, the authors provide a framework and tools to evaluate the relevance of this synthetic construct. The model is subsequently evaluated experimentally. We predict such mixed computational and experimental approaches to gain an increasing role in Chemical Biology projects.
79. Wright TH, Bower BJ, Chalker JM, Bernardes GJ, Wiewiöra R, Ng WL, Raj R, Faulkner S, Vallee MR, Phanumartwath A, et al.: **Posttranslational mutagenesis: A chemical strategy for exploring protein side-chain diversity.** *Science* 2016, **354**:aag1465.
- This seminal publication introduces the first bioorthogonal Csp3-Csp3 bond formation on native proteins, an approach termed chemical mutagenesis. This technique allows to alter residues of recombinantly expressed proteins to a broad range of natural and unnatural amino acids, including post-translational modifications.
80. Yang A, Ha S, Ahn J, Kim R, Kim S, Lee Y, Kim J, Söll D, Lee H-Y, Park H-S: **A chemical biology route to site-specific authentic protein modifications.** *Science* 2016, **354**:623-626.
81. Morrison PM, Foley PJ, Warriner SL, Webb ME: **Chemical generation and modification of peptides containing multiple dehydroalanines.** *Chem Commun (Camb)* 2015, **51**:13470-13473.
- This is the first report of the MDBP reagent for generation of Dha from cysteine. This reagent has faster reaction kinetics than DBHDA, has improved solubility and does not lead to overalkylation on lysine residues.
82. Dadová J, Wu K-J, Isenegger PG, Errey JC, Bernardes GJL, Chalker JM, Raich L, Rovira C, Davis BG: **Precise Probing of Residue Roles by Post-Translational β,γ -C,N Aza-Michael Mutagenesis in Enzyme Active Sites.** *ACS Central Science* 2017, **3**:1168-1173.
- The first reported use of MDBP as a reagent to generate Dha on proteins. Moreover, the first aza-Michael addition to Dha is described.

83. Wright TH, Davis BG: **Post-translational mutagenesis for installation of natural and unnatural amino acid side chains into recombinant proteins.** *Nat. Protocols* 2017, **12**:2243-2250.
84. Guo J, Wang J, Lee JS, Schultz PG: **Site-Specific Incorporation of Methyl- and Acetyl-Lysine Analogues into Recombinant Proteins.** *Angewandte Chemie International Edition* 2008, **47**:6399-6401.
85. Wang ZU, Wang Y-S, Pai P-J, Russell WK, Russell DH, Liu WR: **A Facile Method to Synthesize Histones with Posttranslational Modification Mimics.** *Biochemistry* 2012, **51**:5232-5234.
86. Key HM, Miller SJ: **Site- and Stereoselective Chemical Editing of Thiostrepton by Rh-Catalyzed Conjugate Arylation: New Analogues and Collateral Enantioselective Synthesis of Amino Acids.** *Journal of the American Chemical Society* 2017, **139**:15460-15466.
87. Chapman CJ, Matsuno A, Frost CG, Willis MC: **Site-selective modification of peptides using rhodium and palladium catalysis: complementary electrophilic and nucleophilic arylation.** *Chemical Communications* 2007:3903-3905.
88. Chalker JM, Bernardes GJL, Davis BG: **A "Tag-and-Modify" Approach to Site-Selective Protein Modification.** *Accounts of Chemical Research* 2011, **44**:730-741.
89. Nguyen UTT, Bittova L, Müller MM, Fierz B, David Y, Houck-Loomis B, Feng V, Dann GP, Muir TW: **Accelerated chromatin biochemistry using DNA-barcoded nucleosome libraries.** *Nature Methods* 2014, **11**:834.
90. Tessarz P, Kouzarides T: **Histone core modifications regulating nucleosome structure and dynamics.** *Nature Reviews Molecular Cell Biology* 2014, **15**:703.
91. Taverna SD, Li H, Ruthenburg AJ, Allis CD, Patel DJ: **How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers.** *Nature Structural & Molecular Biology* 2007, **14**:1025.
92. Wang Z, Zang C, Rosenfeld JA, Schones DE, Barski A, Cuddapah S, Cui K, Roh T-Y, Peng W, Zhang MQ, et al.: **Combinatorial patterns of histone acetylations and methylations in the human genome.** *Nature Genetics* 2008, **40**:897.
93. Di Cerbo V, Mohn F, Ryan DP, Montellier E, Kacem S, Tropberger P, Kallis E, Holzner M, Hoerner L, Feldmann A, et al.: **Acetylation of histone H3 at lysine 64 regulates nucleosome dynamics and facilitates transcription.** *eLife* 2014, **3**:e01632.
94. North JA, Šimon M, Ferdinand MB, Shoffner MA, Picking JW, Howard CJ, Mooney AM, van Noort J, Poirier MG, Ottesen JJ: **Histone H3 phosphorylation near the nucleosome dyad alters chromatin structure.** *Nucleic Acids Research* 2014, **42**:4922-4933.
95. Stützer A, Liokatis S, Kiesel A, Schwarzer D, Sprangers R, Söding J, Selenko P, Fischle W: **Modulations of DNA Contacts by Linker Histones and Post-translational Modifications Determine the Mobility and Modifiability of Nucleosomal H3 Tails.** *Molecular Cell* 2016, **61**:247-259.
96. Lercher L, Raj R, Patel NA, Price J, Mohammed S, Robinson CV, Schofield CJ, Davis BG: **Generation of a synthetic GlcNAcylated nucleosome reveals regulation of stability by H2A-Thr101 GlcNAcylation.** *Nature Communications* 2015, **6**:7978.

97. Yan Q, Dutt S, Xu R, Graves K, Juszczynski P, Manis JP, Shipp MA: **BBAP Monoubiquitylates Histone H4 at Lysine 91 and Selectively Modulates the DNA Damage Response.** *Molecular Cell* 2009, **36**:110-120.
98. Ye J, Ai X, Eugeni EE, Zhang L, Carpenter LR, Jelinek MA, Freitas MA, Parthun MR: **Histone H4 Lysine 91 Acetylation: A Core Domain Modification Associated with Chromatin Assembly.** *Molecular Cell* 2005, **18**:123-130.
99. Casadio F, Lu X, Pollock SB, LeRoy G, Garcia BA, Muir TW, Roeder RG, Allis CD: **H3R42me2a is a histone modification with positive transcriptional effects.** *Proceedings of the National Academy of Sciences* 2013, **110**:14894-14899.
100. Lu X, Simon MD, Chodaparambil JV, Hansen JC, Shokat KM, Luger K: **The effect of H3K79 dimethylation and H4K20 trimethylation on nucleosome and chromatin structure.** *Nat Struct Mol Biol* 2008, **15**:1122-1124.
101. Bartke T, Vermeulen M, Xhemalce B, Robson SC, Mann M, Kouzarides T: **Nucleosome-Interacting Proteins Regulated by DNA and Histone Methylation.** *Cell* 2010, **143**:470-484.
- This seminal paper combines the power of synthetic nucleosomes and quantitative proteomics for identification of nucleosome-binding protein partners in an unbiased manner.
102. Nikolov M, Stützer A, Mosch K, Krasauskas A, Soeroes S, Stark H, Urlaub H, Fischle W: **Chromatin Affinity Purification and Quantitative Mass Spectrometry Defining the Interactome of Histone Modification Patterns.** *Molecular & Cellular Proteomics* 2011, **10**.
103. Raj R, Lercher L, Mohammed S, Davis BG: **Synthetic Nucleosomes Reveal that GlcNAcylation Modulates Direct Interaction with the FACT Complex.** *Angewandte Chemie* 2016, **128**:9064-9068.
- Highlights the importance of 'proper' substrates i.e. nucleosome for studying epigenetic mechanisms as compared to peptides or proteins.
104. Fujiki R, Hashiba W, Sekine H, Yokoyama A, Chikanishi T, Ito S, Imai Y, Kim J, He HH, Igarashi K, et al.: **GlcNAcylation of histone H2B facilitates its monoubiquitination.** *Nature* 2011, **480**:557.
105. David Y, Muir TW: **Emerging Chemistry Strategies for Engineering Native Chromatin.** *Journal of the American Chemical Society* 2017, **139**:9090-9096.
106. Han S, Yang A, Lee S, Lee H-W, Park CB, Park H-S: **Expanding the genetic code of *Mus musculus*.** 2017, **8**:14568.
107. Liszczak GP, Brown ZZ, Kim SH, Oslund RC, David Y, Muir TW: **Genomic targeting of epigenetic probes using a chemically tailored Cas9 system.** *Proceedings of the National Academy of Sciences* 2017, **114**:681-686.
- By combining the split-intein technique with the dCas9 directing system, the authors were able to position a synthetic construct at a defined place in the genome.
108. Amamoto Y, Aoi Y, Nagashima N, Suto H, Yoshidome D, Arimura Y, Osakabe A, Kato D, Kurumizaka H, Kawashima SA, et al.: **Synthetic Posttranslational Modifications: Chemical Catalyst-Driven Regioselective Histone Acylation of Native Chromatin.** *J Am Chem Soc* 2017, **139**:7568-7576.
- First application of a ligand-directed catalyst-mediated site-specific acetylation of native nucleosomes. These findings could potentially bridge the world of small molecules with histone chemical biology.

109. Davey CA, Sargent DF, Luger K, Maeder AW, Richmond TJ: **Solvent Mediated Interactions in the Structure of the Nucleosome Core Particle at 1.9Å Resolution.** *Journal of Molecular Biology* 2002, **319**:1097-1113.
110. De Leon CA, Levine PM, Craven TW, Pratt MR: **The Sulfur-Linked Analogue of O-GlcNAc (S-GlcNAc) Is an Enzymatically Stable and Reasonable Structural Surrogate for O-GlcNAc at the Peptide and Protein Levels.** *Biochemistry* 2017, **56**:3507-3517.