Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Chemical Biology

Synthetic histone code

Wolfgang Fischle[1], Henning D Mootz[2] and Dirk Schwarzer[3]

Chromatin is the universal template of genetic information in all eukaryotic cells. This complex of DNA and histone proteins not only packages and organizes genomes but also regulates gene expression. A multitude of posttranslational histone modifications and their combinations are thought to constitute a code for directing distinct structural and functional states of chromatin. Methods of protein chemistry, including protein semisynthesis, amber suppression technology, and cysteine bioconjugation, have enabled the generation of so-called designer chromatin containing histones in defined and homogeneous modification states. Several of these approaches have matured from proof-of-concept studies into efficient tools and technologies for studying the biochemistry of chromatin regulation and for interrogating the histone code. We summarize pioneering experiments and recent developments in this exciting field of chemical biology.

Addresses
[1] Max Planck Institute for Biophysical Chemistry, 37077 Göttingen, Germany
[2] Institute of Biochemistry, University of Muenster, 48149 Muenster, Germany
[3] Interfaculty Institute of Biochemistry, University of Tübingen, 72076 Tübingen, Germany

Corresponding authors: Fischle, Wolfgang (wfischl@gwdg.de), Mootz, Henning D (henning.mootz@unimuenster.de) and Schwarzer, Dirk (dirk.schwarzer@uni-tuebingen.de)

Current Opinion in Chemical Biology 2015, **28**:131–140

This review comes from a themed issue on **Synthetic Biomolecules**

Edited by **Christian Hackenberger** and **Peng Chen**

http://dx.doi.org/10.1016/j.cbpa.2015.07.005

Chromatin and posttranslational histone modifications

The physiological template of genetic information in all eukaryotic cells is chromatin, the complex of DNA and histone proteins. As signaling platform, chromatin integrates a variety of internal and external cellular inputs. These direct distinct local and global structural and functional states of chromatin, thereby controlling gene expression [1].

In the repeating unit of chromatin, the nucleosome, DNA is wrapped around an octamer of histone proteins (two copies each of the core histones H2A, H2B, H3, and H4 and one copy of linker histone H1, Figure 1a) [1]. By various levels of folding, the originating primary chromatin fiber can be organized into various structural arrangements, including highly condensed mitotic chromosomes (Figure 1b). While the basic architecture of nucleosomes is the same throughout the genome, posttranslational modifications (PTMs) of histones are central means of increasing the biochemical divergence that regulates chromatin structure and function [2,3]. Histone PTMs are structurally diverse and include methylation, acetylation and ubiquitinylation of lysine, as well as phosphorylation of serine and threonine residues (Figure 1c). More than 150 histone modification sites have been identified in different experimental systems. Major sites of modification cluster within the unstructured regions of the N-terminal histone tails that vary between 10 and 35 amino acids in length. These are protruding out from the nucleosome core (Figure 1a).

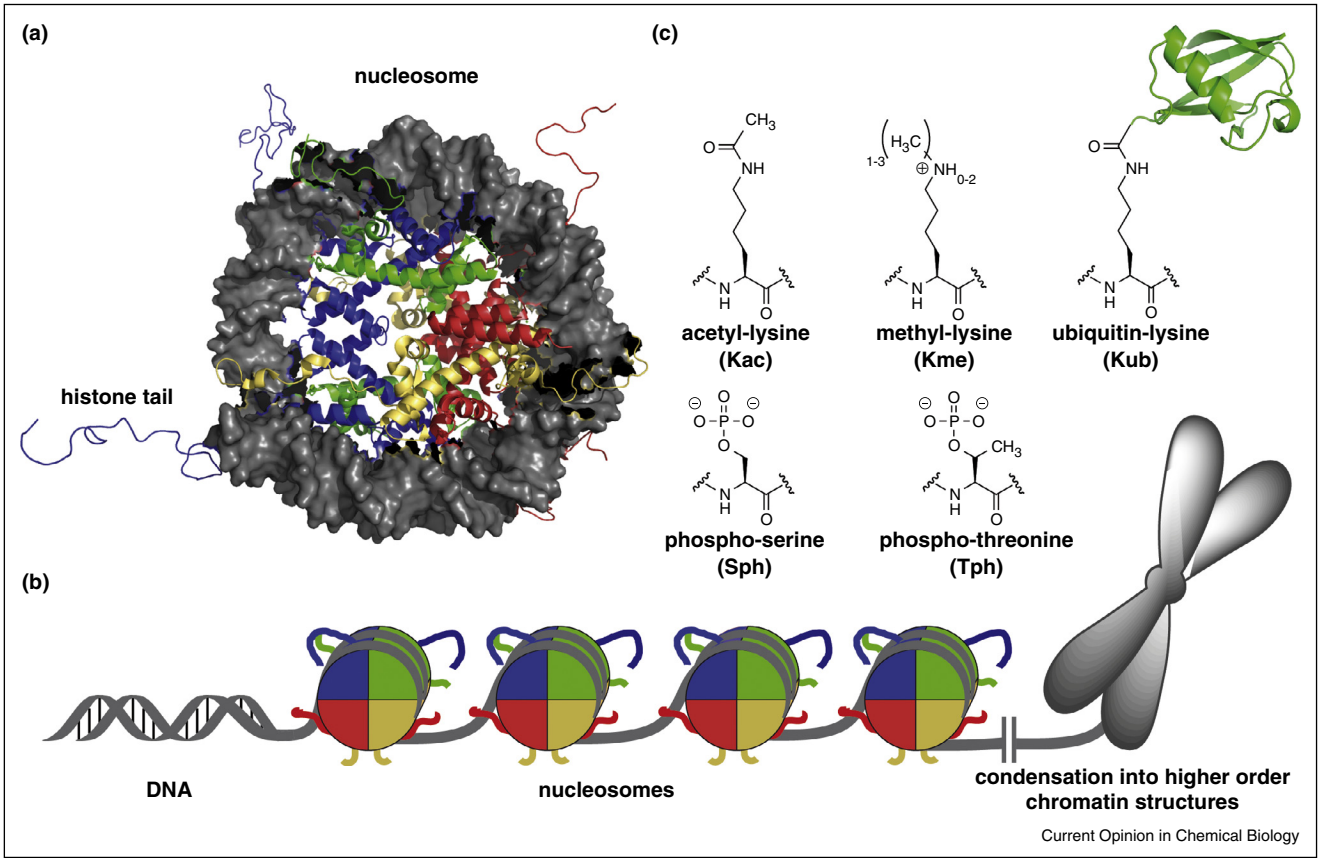
Research over the past years has focused on defining the distribution, biochemistry and cellular function of individual and combinations of histone marks. A 'histone code' hypothesis has been put forward that defines chromatin as dynamic programming platform, which integrates internal and external cellular signals [2,3]. Histone modifying enzymes are considered the writers of the histone code. Chromatin factors possessing PTM-binding domains serve as histone code readers and execute regulatory functions upon recruitment. Finally, dedicated PTM-removing enzymes are the erasers of histone PTMs [4–7].

Deciphering the complex cross-talk between histone PTMs, chromatin factors and gene regulation is a major research challenge [8]. While the accessibility of 'designer chromatin' composed of histones with defined and homogeneous modification states has been a central bottleneck in chromatin research, novel chemical approaches are providing powerful tools enabling important discoveries of the rules of the histone code [4–7,9–12].

Tools for designer chromatin

Chromatin research has benefitted strongly from peptide chemistry. Especially, the availability of building blocks of modified amino acids has enabled the synthesis of histone peptides with defined modification patterns. Such histone-derived peptides have been instrumental for many key discoveries on the readers, writers and eraser of histone PTMs. However, histone-peptides can only recapitulate a minor fraction of the complex chromatin structure (Figure 1). Consequently, major efforts have been untertaken to site-specifically introduce

Figure 1



Histones, nucleosomes and chromatin. **(a)** Structure of the nucleosome core particle. Histones are color-coded as follows: H2A: yellow, H2B: red, H3: blue, H4: green. The illustration was generated from pdb file: 1KX5. **(b)** Packaging of DNA into nucleosomes and chromosomes. Intermediate packaging stages are not illustrated. **(c)** Selected posttranslational modifications of histones including acetylation, methylation and ubiquitination of lysine as well as phosphorylation of serine and threonine residues.

modification marks into full-length histones, nucleosomes, arrays of nucleosomes and even into chromatin of living cells. These approaches draw from the technological repertoire of modern protein chemistry [4–7].

Selective modification of cysteine residues

The most widely used strategy for generating modified histones is the selective conversion of cysteine residues into mimics of modified amino acids. The chemical properties of the cysteine thiol group are unique among the proteinogenic amino acids. Selective alkylation of this soft nucleophile has been used in traditional biochemistry and this approach has recently experienced a renaissance for site-specific introduction of mimics of PTMs. Since natural cysteine residues are absent in core histones with exception of a single residue in H3, this method is particularly attractive for chromatin chemistry. In a pioneering report site-directed mutagenesis was used to introduce cysteine at sites of methylated lysine residues in otherwise Cys-free mutant histone H3 as well as histone H4 [13]. Upon treatment with 2-chloro-ethyl-methylammonium or bromo-ethyl-methylammonium compounds (**1**) the respective methyl-thialysine (**2**) residues were generated in high yields (Figure 2a). Downstream experiments showed that these so-called methyl-lysine analogs (MLAs) mimicked properties of methylated lysine in peptide and nucleosomal context by recruiting chromatin readers.

In a similar fashion methyl-arginine analogs (**4**) were generated [14]. In this case α,β-unsaturated amidines (**3**) served as alkylation agents. Extending this approach to acetyl-thialysine (**6**) turned out much more challenging [15]. The treatment of cysteine with 2-bromoethyl-acetylamine (**5**) or with more reactive N-acetyl-aziridine, were either poor yielding or led to undesired side-products. Alkylation with methylthiocarbonyl-aziridine (**7**) converted cysteine residues into methylthiocarbonyl-thialysine (**8**), which mimicked some aspects of acetylated lysine [15]. Efficient formation of acetyl-thiaLys (**6**) was finally achieved with N-vinyl-acetamidine (**9**), however not by direct conjugation but via the radical thiol-ene