Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Chemical Biology

Studies of biochemical crosstalk in chromatin with semisynthetic histones

Calvin Jon Antolin Leonen, Esha Upadhyay and Champak Chatterjee

Reversible post-translational modifications of histone proteins in eukaryotic chromatin are closely tied to gene function and cellular development. Specific combinations of histone modifications, or *marks*, are implicated in distinct DNA-templated processes mediated by a range of chromatin-associated enzymes that install, erase and interpret the histone code. Mechanistic studies of the precise biochemical relationship between sets of marks and their effects on chromatin function are significantly complicated by the dynamic nature and heterogeneity of marks in cellular chromatin. Protein semisynthesis is a chemical technique that enables the piecewise assembly of uniformly and site-specifically modified histones in quantities sufficient for biophysical and biochemical analyses. Recent pioneering efforts in semisynthesis have yielded access to histones site-specifically modified by entire proteins, such as ubiquitin (Ub) and the small ubiquitin-like modifier (SUMO). Herein, we highlight key studies of biochemical crosstalk involving Ub and SUMO in chromatin that were enabled by histone semisynthesis.

Address
Department of Chemistry, University of Washington, Seattle, WA 98195, United States

Corresponding author: Chatterjee, Champak (chatterjee@chem.washington.edu)

Current Opinion in Chemical Biology 2018, **45**:27–34

This review comes from a themed issue on **Chemical genetics and epigenetics**

Edited by **Wolf Gang Fischle** and **Dirk Schwarzer**

https://doi.org/10.1016/j.cbpa.2018.02.005

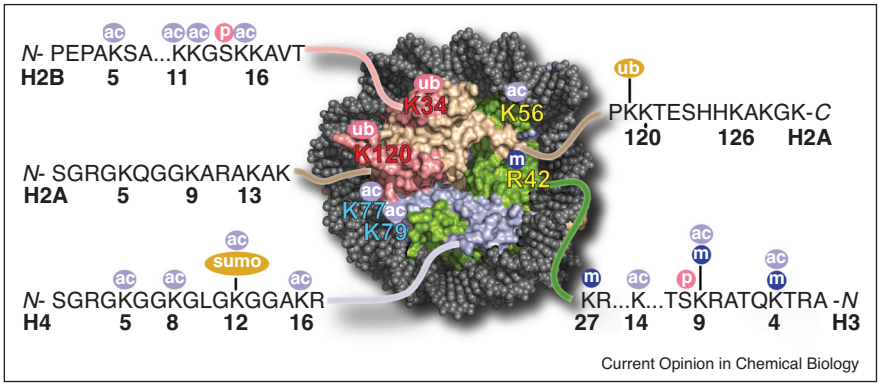
Introduction

The past decade has witnessed an unprecedented surge in chemical biology efforts directed toward studying the regulation of eukaryotic chromatin by reversible chemical modifications [1]. Chromatin is a massive and dynamic nucleoprotein complex that stores vast amounts of genetic material, about 3 billion base pairs of DNA in humans, within the minute cell nucleus. The fundamental building block of chromatin is the nucleosome core particle (NCP), which consists of ~147 bp of double-stranded DNA wrapped about 1.6 turns around a globular protein spool consisting of the core histones H2A, H2B, H3, and H4 [2,3•]. Extending outwards from the NCP, and hence exposed to the nucleosol, are the functional group rich histone *N*-termini or *tails*. Histone tails are the sites of a wide range of chemically distinct and reversible post-translational modifications (PTMs). These include varying degrees of arginine and lysine side-chain methylation (mono-methylation, di-methylation, and tri-methylation), serine, threonine, or tyrosine phosphorylation, and lysine acylation. The last ranges from modification by smaller acetyl to longer crotonyl and butyryl groups and even entire proteins such as ubiquitin (Ub) and SUMO (*s*mall *u*biquitin-like *mo*difier) (Figure 1) [4–6]. Histone PTMs or *marks* seldom exist in isolation and the diversity of modified histones revealed in numerous proteomic studies is staggering [7]. Despite the complex landscape of chromatin modifications, specific sets of marks may be correlated with distinct transcriptional states of their associated genes. This led to the hypothesis that combinatorial patterns of marks may constitute a *histone code* that regulates key DNA-templated processes such as transcription, replication and damage repair [8••]. This review will highlight important developments in protein chemistry that have enabled studies of the biochemical relationship, or *crosstalk*, between the Ub family of histone marks and histone methylation.

Mechanisms of chromatin regulation by histone marks

Histone marks influence chromatin structure and function by two mutually non-exclusive mechanisms. In some instances, a mark may directly change local chromatin structure, facilitating or denying access to numerous chromatin-modifying enzymes. Secondly, a mark may serve to recruit chromatin-associated proteins that deposit (*writers*), remove (*erasers*), or bind (*readers*) specific sets of marks [9]. Indeed, a body of literature exists for certain privileged histone marks, such as H4 Lys16 acetylation (H4K16ac), which is associated with actively transcribed genes [10] and an open euchromatin structure [11••]. The absence of H4K16ac in chromatin is strongly associated with transcriptional repression, heterochromatin formation, and the appearance of repressive marks such as

Figure 1



Histone tail marks. Scheme showing the diversity of marks that have been accessed by chemical protein synthesis and semisynthesis. Chemical groups are indicated as ac, acetyl; m, methyl; p, phosphoryl; sumo, small ubiquitin-like modifier; ub, ubiquitin. The globular core of the nucleosome core particle is shown with histones colored as H2A (gold), H2B (red), H3 (green) and H4 (blue) and double-stranded DNA (gray). PDB code 1KX5.

H3K9me3 [12,13] and H3K27me3 [14,15]. Further complexity in the writing and execution of the histone code arises from the crosstalk between marks, whereby one mark may lead to the addition or removal of others on the same histone or on a different histone [16]. Beyond histone-centric crosstalk alone are the relationships between marks and DNA modifying enzymes [17] or long non-coding RNAs [18], which are also important regulators of cellular outcomes.

Histone semisynthesis to unravel a complex code

The empirical association between histone marks and gene function at different loci is a critical and important step toward understanding chromatin regulation, but it provides little clarity at the molecular level. Cellular signaling is inherently complex and immediate changes in chromatin prior to or following the installation of specific marks are not readily captured in cell-based experiments. Elucidating the precise biochemical crosstalk between sets of marks in chromatin is particularly challenging due to their complex pattern, varying abundance and dynamic nature in cells. Indeed, a molecular understanding of the roles for specific marks requires uniformly and site-specifically modified NCPs, in order to establish a direct line of causality between a mark and its functions. Therefore, the last decade has seen the development of a large number of chemical biology techniques to access homogeneously and site-specifically modified histones [1]. Protein semisynthesis is one early and key technique that has enabled detailed mechanistic studies of crosstalk in well-defined NCP substrates [19•].

The enzymatic modification of recombinant histones is a valuable strategy for generating substrates used in mechanistic studies, but it requires knowledge of the site-specific writer and its isolation in an active form. Semisynthesis has several advantages including (1) the ease of scalability, (2) chemospecificity, and (3) the ability to install multiple different marks in close proximity to each other. The latter may be particularly challenging when the natural order of histone modification is unknown, leading to the inhibition of a desired enzymatic activity by partially modified histones.

Native chemical ligation (NCL) is a widely utilized semisynthetic strategy that overcomes inherently limited yields from the solid-phase peptide synthesis of proteins longer than ~50 amino acids, such as the histones [20••]. NCL permits chemoselective amide bond formation between two polypeptide fragments, by the incorporation of a C-terminal α-thioester in one fragment and N-terminal Cys residue in the other (Figure 2). An initial reaction between the fragments links them by a thioester, which spontaneously rearranges to the native backbone amide bond at neutral-to-alkaline pH. The fragment with N-terminal Cys may be synthetic or recombinant in nature depending upon its desired PTM state. Expressed protein ligation (EPL) further extends the scope of NCL by employing an intein, a single-turnover enzyme that undertakes protein splicing in bacteria, to generate the C-terminal α-thioester [21•]. Thus, either N-terminally or C-terminally modified histone tails are readily accessible by protein semisynthesis employing two asymmetrical fragments, with the smaller synthetically accessible fragment containing any desired marks (Figure 2). The total chemical synthesis of histones by successive NCL steps has also rendered marks in the interior of histones, such as acetylation at H3K56 [22] and phosphorylation at H2AY57, H3Y41 and T45, accessible in good quantities for mechanistic studies [23].