Biochimica et Biophysica Acta xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbagrm

Review

Engineering chromatin states: Chemical and synthetic biology approaches to investigate histone modification function☆

Horst Pick, Sinan Kilic, Beat Fierz

Fondation Sandoz Chair in Biophysical Chemistry of Macromolecules, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland

ARTICLE INFO

Article history:
Received 17 January 2014
Received in revised form 26 March 2014
Accepted 16 April 2014
Available online xxxx

Keywords:
Epigenetics
Histone PTM
EPL
NCL
Designer chromatin
Genetic code expansion
Synthetic biology

ABSTRACT

Patterns of histone post-translational modifications (PTMs) and DNA modifications establish a landscape of chromatin states with regulatory impact on gene expression, cell differentiation and development. These diverse modifications are read out by effector protein complexes, which ultimately determine their functional outcome by modulating the activity state of underlying genes. From genome-wide studies employing high-throughput ChIP-Seq methods as well as proteomic mass spectrometry studies, a large number of PTMs are known and their coexistence patterns and associations with genomic regions have been mapped in a large number of different cell types. Conversely, the molecular interplay between chromatin effector proteins and modified chromatin regions as well as their resulting biological output is less well understood on a molecular level. Within the last decade a host of chemical approaches has been developed with the goal to produce synthetic chromatin with a defined arrangement of PTMs. These methods now permit systematic functional studies of individual histone and DNA modifications, and additionally provide a discovery platform to identify further interacting nuclear proteins. Complementary chemical- and synthetic-biology methods have emerged to directly observe and modulate the modification landscape in living cells and to readily probe the effect of altered PTM patterns on biological processes. Herein, we review current methodologies allowing chemical and synthetic biological engineering of distinct chromatin states in vitro and in vivo with the aim of obtaining a molecular understanding of histone and DNA modification function. This article is part of a Special Issue entitled: Molecular mechanisms of histone modification function.

1. Introduction

Combinations of post-translational modifications of histone proteins (PTMs or 'marks') act as a synergistic signaling platform regulating chromatin function. Currently a large host of histone modifications is known, such as lysine acetylation (Kac), mono-, di- and trimethylation (Kme1/2/3), symmetric and asymmetric arginine methylation, arginine deimination, serine and threonine phosphorylation and glycosylation, ADP ribosylation, lysine ubiquitylation (Kub) as well as SUMOylation which can directly alter the physico-chemical properties of chromatin and function as recognition sites for chromatin effector proteins [1,2]. On the other hand, a different set of chemical modifications appears on the DNA template itself, not only foremost 5-methyl-cytosine in CpG dinucleotide [3], but also cytosine hydroxymethylation and higher oxidized forms [4]. The colocalization and spatial arrangement of combinations of histone modifications, together with specific non-histone chromatin proteins (effectors or regulators) form a chromatin state and are coupled to biological function. Fundamental chromatin states include heterochromatin, which is highly compacted, characterized by H3 K9 methylation and histone deacetylation and the presence of heterochromatin protein 1 (HP1) and where gene expression is silenced. Similarly, polycomb repressed regions are characterized by H3 K27 methylation and the presence of polycomb repressive complexes 1 and 2 (PRC1, 2) [5]. Conversely, active chromatin regions, such as transcriptionally active genes, promoters and enhancers exhibit various degrees of histone acetylation, methylation at H3 K4 and K36 and the presence of a large number of different effector proteins, including RNA polymerase II and general transcription factors [6]. DNA methylation patterns have important functions in long term gene regulation and epigenetic inheritance [7], whereas our understanding of the role of DNA hydroxymethylation is still limited [8,9].

Recognizing the exceedingly high complexity of possibly co-existing histone PTMs, the hypothesis of a 'histone code' was put forward to establish a causal link between PTM patterns and genome function [10]. Histone modifications can alter DNA accessibility and chromatin structure. Furthermore, they can serve as binding sites for chromatin effector proteins such as histone modifying enzymes. Many effectors contain one or several protein domains that specifically interact with histone modifications ('reader' domains). Examples include chromo-, tudor-, WD40 and malignant brain tumor MBT-domains

☆ This article is part of a Special Issue entitled: Molecular mechanisms of histone modification function.

http://dx.doi.org/10.1016/j.bbagrm.2014.04.016
1874-9399/

(which bind Kme1/2/3), bromodomains (BD, which recognize Kac), 14-3-3 domains (bind phosphorylated serines) and plant homeodomain PHD fingers (recognize unmodified or methylated lysines) [11]. Individual interactions between chromatin proteins and histones are fairly weak, exhibiting dissociation constants (K_d) in the micromolar range [12]. Therefore, a single PTM — reader domain interaction may not be sufficient to transduce a particular effector-mediated consequence. Indeed, a significant number of chromatin effectors contain multiple reader domains, or exist in higher order complexes containing multiple chromatin interaction motifs. The combinatorial action of several (low-affinity) reader domains (including other protein-protein, or protein-DNA interaction domains) then allows simultaneous recognition of multiple histone PTMs coexisting on a nucleosome in a multivalent fashion, with greatly increased affinity [12,13]. This presents an interesting mechanistic hypothesis how a chromatin state, characterized by combinations of PTMs, is specifically read-out by effectors and translated into a biological output. Since the development of this model [13], several examples of such interactions have been observed and studied [14–17].

In addition to thermodynamic considerations, the average dwell time at a particular chromatin locus is a critical parameter for effector action. Large scale chromatin compartments, such as heterochromatin, are highly stable while the individual factors are in rapid exchange with soluble proteins [18]. This allows a fast response to stimuli, e.g. for chromatin remodeling in DNA damage repair [19]. Such malleability can be understood to be a result of many weak effectors — chromatin interactions, associated with fast dissociation rate constants. In this context, multivalent interactions are an attractive way to establish selectivity while retaining kinetically dynamic interactions, as they mainly result in an increase of the local concentration of factors at its target chromatin region thereby accelerating the binding kinetics [13]. In turn, local competition for binding sites can still take place and, through processes such as facilitated dissociation [20,21], rapid chromatin factor exchange can take place. A crucial factor for a deeper understanding of such interactions is thus the local concentration of both factors and PTMs in the nucleus, and the resulting interaction thermodynamics and kinetics. While the degree of modification is mostly unknown at given chromatin loci, recent findings demonstrated that nucleosomes are often not homogenously modified, but that many histone PTMs exist in an asymmetric fashion: One copy of a particular histone might carry one modification whereas the other copy is unmodified or modified differently [22]. This might have important consequences for the downstream readout by multivalent effectors.

Methylated DNA, on the other hand, recruits its own set of associated proteins, for example methyl CpG binding protein 2 (MeCP2) and the SET- and RING-associated (SRA) domain in the ubiquitin-like, containing PHD and RING finger domains 1 (UHRF1) complex. Contrariwise, other DNA binding proteins are expelled from their target sequences by methylation resulting in a cell specific biological output [23]. In summary, patterns of coexisting chromatin modifications and associated effectors define distinct chromatin states, which correlate with the expression levels of underlying genes, splicing activity, and replication and repair processes [24,25]. Chromatin states can persist over cell generations and are involved in the regulation of cell differentiation and lineage commitment. Therefore, the combination of the chemical modifications in DNA and histones, the associated effector proteins and the chromatin structural states can be considered to contribute to epigenetic inheritance [26].

Currently the knowledge on the location and amounts as well as the combinatorial complexity and dynamics of histone modifications in cell populations is rapidly expanding through genome wide investigations employing ChIP-Seq methodologies [27] and mass spectrometry (MS) based investigations [28]. Moreover, low cell number and time-dependent ChIP methodologies are emerging, enabling ChIP analysis on small sample sizes, minimizing information loss through ensemble averaging and allowing kinetic investigations [29–32]. However the molecular mechanism of PTMs function is poorly understood, information on the single cell dynamics of modifications is often lacking and the biological downstream effects of the epigenetic PTM landscape remain elusive.

In recent years, synthetic chemistry and synthetic biology approaches have gained an increasingly important role in elucidating basic molecular mechanisms of chromatin regulation: Synthetically produced designer chromatin, for instance, in combination with biochemical and biophysical methods allows quantitative measurements under exactly defined experimental conditions. Such approaches have revealed how histone marks locally regulate chromatin structure and function, and enabled comprehensive analyses of effector binding to mono- and multivalent histone marks thus testing the basic tenets of the histone code hypothesis. Likewise, protein engineering and synthetic biology methods result in novel tools for reading out and re-writing endogenous histone marks in living cells. These strategies pave new ways to quantitatively monitor changes in histone modification dynamics, e.g., under normal and disease conditions and allow to study the output of synthetically introduced histone modifications on biological functions. In this review we cover synthetic chemical and biological methods to study and to rewrite the histone language in vitro and in cells to complement discovery based approaches.

2. Engineering chromatin states in vitro

Due to the combinatorial complexity of histone modifications it has been exceedingly complicated to sort out the detailed effects of each PTM separately. Synthetic histone tail peptides carrying distinct modifications have proven useful to study the functional role of individual reader domains in chromatin-associated proteins. However, in the majority of cases chromatin readers function in tandem or in the context of homo- or hetero-oligomeric effector complexes. In such cases, reconstituted nucleosomes and chromatin fibers with defined patterns of histone PTMs and DNA modifications are required for the establishment of direct causal relationships between the specific type of modification and downstream effects.

Several molecular components are needed to engineer defined chromatin states in vitro (Fig. 1): *Firstly*, histone proteins have to be produced, carrying desired site-specific modifications. The isolation of histones from biological material yields proteins with heterogeneous PTM patterns, whereas enzymatic modification of histones in vitro often does not result in site-specifically or homogenously modified histones. A solution to this problem is provided by chemical synthesis and protein engineering approaches. *Secondly*, suitable DNA templates are needed, tailored for the specific experiments allowing defined localization of nucleosomes by the use of nucleosome positioning sequences (NPS) and potentially containing modified bases and reactive sites for the controlled assembly of heterogeneous chromatin arrays. *Thirdly*, methodologies to assemble, connect, and purify reconstituted chromatin systems are required. In the following, we will discuss the currently available chemical methods and engineering approaches to produce designer chromatin substrates and highlight recent applications for solving biological questions.

2.1. Native chemical ligation and expressed protein ligation approaches

Several chemical methods have been developed to site-specifically introduce PTMs or mimics thereof into full-length histones, of which arguably native chemical ligation (NCL) [33] as well as expressed protein ligation (EPL) [34] exhibit the broadest scope. The efficient chemical synthesis of peptides usually breaks down after 50–70 amino acids [35], falling short of most but the smallest proteins. Efficient ligation of individually synthesized peptide fragments however can circumvent these synthetic limitations, and thereby enable the straightforward introduction of unnatural or modified amino acids into protein. Following on a number of peptide ligation

Please cite this article as: H. Pick, et al., Engineering chromatin states: Chemical and synthetic biology approaches to investigate histone modification function, Biochim. Biophys. Acta (2014), http://dx.doi.org/10.1016/j.bbagrm.2014.04.016