FEBS Letters 587 (2013) 1247–1257

journal homepage: www.FEBSLetters.org

Review

Deciphering post-translational modification codes

Adam P. Lothrop [a], Matthew P. Torres [b,*], Stephen M. Fuchs [a,*]

[a] Department of Biology, Tufts University, 200 Boston Ave. Suite 4700, Medford, MA 02155, United States
[b] School of Biology, Georgia Institute of Technology, Atlanta, GA 30332, United States

ARTICLE INFO

Article history:
Received 8 December 2012
Revised 20 January 2013
Accepted 23 January 2013
Available online 10 February 2013

Edited by Wilhelm Just

Keywords:
Post-translational modification
Histone
Phosphorylation
Epigenetic
Signaling
Methylation

ABSTRACT

Post-translational modifications (PTMs) occur on nearly all proteins. Many domains within proteins are modified on multiple amino acid sidechains by diverse enzymes to create a myriad of possible protein species. How these combinations of PTMs lead to distinct biological outcomes is only beginning to be understood. This manuscript highlights several examples of combinatorial PTMs in proteins, and describes recent technological developments, which are driving our ability to understand how PTM patterns may "code" for biological outcomes.

1. Patterns, signatures, and codes

Most proteins are post-translationally regulated in some manner by enzymes that directly alter the chemical makeup of the protein. These enzymes can be proteases, transferases (kinases, acetyltransferases, methyltransferases, glycosyltransferases, etc.), or enzymes that remove groups (phosphatases, deacetylases, glycosidases, etc.). In all, more than 400 discrete types of modifications can occur and, to date, more than 90 000 individual PTMs have been identified through biochemical and biophysical analysis [1]. PTMs are known to act alone and in combination to regulate nearly all aspects of protein function. Thus, deciphering how PTMs are coordinately regulated is of fundamental importance to our understanding of biology.

While many proteins are known to be heavily modified, combinatorial PTMs have perhaps been best studied in the context of histones, where, more than ten years ago, Strahl and Allis proposed that PTMs on the tails of histone proteins, alone or in combination, specify downstream events [2]. The original paper talks about a language of histone modifications, but today these concepts are generally referred to as the "histone code hypothesis" [2]. In the intervening years, several significant events have occurred. Tremendous technological advances have been made, allowing us to identify numerous histone PTMs, the enzymes responsible for transferring and removing many PTMs, and a host of protein domains that recognize specific histone PTMs. Improvements in mass spectrometry and proteomics techniques have also revolutionized the rate and detail with which PTMs are identified within the proteome. Consequently, there is great interest in identifying additional codes that modulate protein function [3–5]. Likewise, interdependences between PTMs within distant regions of the same protein, or on different proteins within complexes, are now commonplace. Thus, there is a nearly constant re-examination of how we define the interrelationships of multiple PTMs.

Among researchers interested in PTM biology, there is often debate over the nomenclature used to describe how the multitude of PTMs on a given protein regulates function. Currently, our capacity to detect PTMs far exceeds our ability to understand their biological function. This unfortunate, but important distinction is at the root of the controversy underlying the use of the term "code" to refer to patterns of PTMs that are "read" by the cell to drive biological outcomes. A general set of principles to decipher these codes has not yet emerged, and consequently many find the term "code" misleading. Whether they are called "codes" or not, we can clearly state that: (1) Many proteins have regions within their primary sequence that are targets for extensive, and often overlapping, modification by enzymes; (2) In many cases, these PTMs can recruit the binding or modulate the activity of other proteins; (3) Patterns of PTMs can be identified that correlate with differential biological states (e.g. normal or disease states, cell cycle stage, aging)

* Corresponding authors.
E-mail addresses: matthew.torres@biology.gatech.edu (M.P. Torres), stephen.fuchs@tufts.edu (S.M. Fuchs).

0014-5793/$36.00
http://dx.doi.org/10.1016/j.febslet.2013.01.047

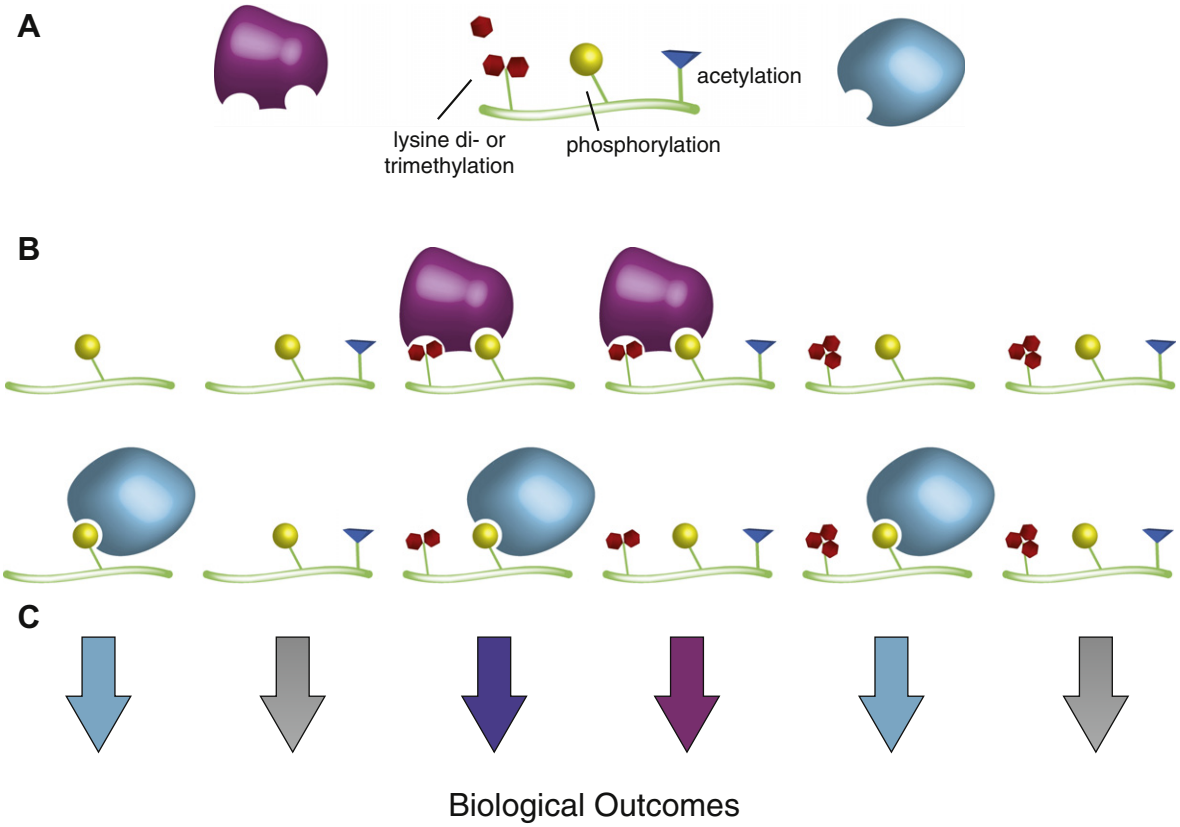

Fig. 1. Combinatorial PTMs can code for complex biological outcomes. (A) Modifications such as methylation (red), phosphorylation (yellow), or acetylation (blue) are commonly recognized by proteins with PTM-recognition domains (purple and cyan). Modifications such as lysine methylation can occur up to three times on a single residue resulting in PTMs with distinct activity. (B) Neighboring PTMs have differing effects on the ability of proteins to recognize a phosphorlyation site. For example, the purple protein requires dimethylation of the lysine, but is occluded by trimethylysine and uninfluenced by the neighboring acetylation. In contrast, the cyan protein can be blocked by aceylation but is uneffected by methylation. (C) The combinatorial PTMs setup a "code," that determines which protein–protein interactions lead to distinct biological outcomes.

(Fig. 1). This review focuses on methods developed to understand the complex biology of PTMs, and the growing evidence demonstrating that the interactions of modifications that exist across a landscape of proteins act concomitantly to orchestrate complex biological outcomes.

2. Combinatorial PTMs coordinate protein–protein interactions – lessons from histone tails

DNA is packaged around two copies each of four histones (H2A, H2B, H3, and H4) to form nucleosomes, the basic unit of chromatin structure. Nucleosomes play pivotal roles in compacting the genome and protecting it from damage. However, packaging of DNA into chromatin is repressive towards DNA-templated processes such as transcription [6]. Eukaryotic cells balance the needs to copy and read, but simultaneously protect, genetic information through a complex network of PTMs primarily directed toward the N- and C-terminal tails of the four histones. Histone PTMs can alter the charge of histones (e.g. lysine acetylation) and recruit specific binding domains (e.g. acetylation, methylation, phosphorylation) associated with proteins such as chromatin remodelers, transcriptional coactivators/repressors, and DNA repair proteins. Histone PTMs have gained prominence since the mid 1990s when the Allis and Schreiber groups demonstrated that histone-modifying enzymes have direct roles in regulating gene expression [7,8]. The sequencing of the human genome, development of chromatin immunoprecipitation (ChIP), and next-generation sequencing technologies have now made histone PTMs the best studied of all cellular modifications. Just as there are a large number of PTMs on the histone tails, there are also numerous protein domains that recognize and bind to particular PTMs on these tails. For example, PTM-recognition domains such as PHD (Plant homeodomain) fingers, chromodomains, and Tudor domains all recognize methylated lysine residues, whereas bromodomains and 14-3-3 domains recognize acetylysine and phosphoserine/threonine respectively [9].

Most PTM-recognition domains recognize a particular modification within a defined amino acid sequence, indicating that neighboring amino acid sequence is important in the context of substrate recognition by these proteins. Because histone tails are rich in PTMs, the presence of nearby modifications influences the ability of protein factors to recognize a particular PTM. For example, phosphorylation of Ser10 on histone H3 (H3S10) negatively influences HP1 (heterochromatin protein 1) recognition of methylation on neighboring Lys9 (H3K9) [10]. This phenomenon is referred to as the "phospho-methyl" switch, as H3S10 phosphorylation acts as a switch to prevent the binding of HP1 to chromatin in mitosis [10]. Recently, investigators uncovered an exception to this finding, where the tandem Tudor domain of UHRF1 binds to H3K9 methylation irrespective of the H3S10 phosphorylation state [11]. This finding suggests that multiple PTMs can act in concert to carefully orchestrate the binding of numerous factors to the same primary modification. Similar examples have been observed elsewhere on the H3 tail, where modification at either Arg2 or Thr3 can impact recognition of either neighboring Lys4 methylation or the free N-terminus of histone H3 [12].

The example of the histone tails demonstrates that the influence of neighboring PTMs on recognition likely extends far beyond the few examples we have currently identified. Several groups re-