Review

Emerging Technologies to Map the Protein Methylome

Scott M. Carlson and Or Gozani

Department of Biology, *Stanford University, 371 Serra Street, Stanford, CA 94305, USA*

Correspondence to Scott M. Carlson: *scottmc@stanford.eduz*
http://dx.doi.org/10.1016/j.jmb.2014.04.024
Edited by S. Khorasanizadeh

Abstract

Protein methylation plays an integral role in cellular signaling, most notably by modulating proteins bound at chromatin and increasingly through regulation of non-histone proteins. One central challenge in understanding how methylation acts in signaling is identifying and measuring protein methylation. This includes locus-specific modification of histones, on individual non-histone proteins, and globally across the proteome. Protein methylation has been studied traditionally using candidate approaches such as methylation-specific antibodies, mapping of post-translational modifications by mass spectrometry, and radioactive labeling to characterize methylation on target proteins. Recent developments have provided new approaches to identify methylated proteins, measure methylation levels, identify substrates of methyltransferase enzymes, and match methylated proteins to methyl-specific reader domains. Methyl-binding protein domains and improved antibodies with broad specificity for methylated proteins are being used to characterize the "protein methylome". They also have the potential to be used in high-throughput assays for inhibitor screens and drug development. These tools are often coupled to improvements in mass spectrometry to quickly identify methylated residues, as well as to protein microarrays, where they can be used to screen for methylated proteins. Finally, new chemical biology strategies are being used to probe the function of methyltransferases, demethylases, and methyl-binding "reader" domains. These tools create a "system-level" understanding of protein methylation and integrate protein methylation into broader signaling processes.
© 2014 Elsevier Ltd. All rights reserved.

Introduction

Post-translational methylation of arginine and lysine on histone proteins has well-established functions in epigenetics and chromatin regulation, and methylation of non-histone proteins has recently emerged as an important actor in broader biological regulation (Fig. 1, discussed in several recent reviews [1–3]). Arginine methylation has many critical functions such as in signal transduction [4], regulation of mRNA splicing and translation [5], and physiological processes such as glucose homeostasis [6] (reviewed in Refs. [7–9]). Non-histone lysine methylation regulates pathways including tumor suppression by p53 and signaling by NFκB [10,11] (reviewed in Ref. [12]). Until recently, most sites of protein methylation had been identified through candidate approaches, arising through detailed investigation of individual proteins.

Development of high-throughput approaches to identify and characterize methylated proteins has been ongoing for the past decade. In the last two years, many of these techniques have reached fruition, revealing thousands of sites of arginine methylation and hundreds of sites of lysine methylation across the human proteome. These discoveries raise important questions about the function and regulation of protein methylation. For example, which of the at least nine known protein arginine methyltransferases (PRMTs) in the human proteome are responsible for all these methylation events, or do additional enzymes remain to be discovered [3], and which lysine methyltransferases (KMTs) act to regulate non-histone methylation [13,14]? The SPOUT family of RNA methyltransferases has recently been shown to methylate arginine in yeast, perhaps this or other families are also acting in mammalian systems [15]. How is methylation regulated during dynamic biological processes, and how does

0022-2836/© 2014 Elsevier Ltd. All rights reserved.

Fig. 1. Methylation states of arginine and lysine.(a) Arginine residues are methylated by PRMTs at an N^G position to form MMA and then methylated at the same or the opposing nitrogen to form SDMA and ADMA. (b) Lysine residues aremethylated by KMTs at the ε-nitrogen to form mono-, di-, or tri-methyl lysine.

it interact with other modes of regulation such as kinase signaling, transcription factor activity, and mRNA splicing? Addressing these questions will require a combination of traditional biological approaches along with new system-level tools to investigate methylation across the entire proteome.

In this review, we will focus on emerging "system-level" tools to investigate protein methylation. We discuss approaches for proteome-wide identification of methylated proteins using target enrichment and tandem mass spectrometry [liquid chromatography and tandem mass spectrometry (LC-MS/MS)], affinity and high-throughput screening to identify PRMT and KMT substrates, and chemical biology approaches to investigate methylation signaling and discover PRMT and KMT substrates. These technologies have already enabled the first steps toward integrating methylation with broader signaling processes. We expect that future developments will make it possible to track methylation signaling dynamics directly and to examine how methylation contributes to signaling that regulates development and disease.

Introduction to Methods for Proteomic Analysis of Arginine and Lysine Methylation

Most strategies for proteomic analysis of post-translational modifications begin by enriching modified proteins from cell lysate or by digesting the lysate into short peptides and then enriching modified peptides. Such techniques are well established for phosphorylation, where selective resins are used to enrich phosphorylated peptides from complex mixtures [16]. Other modifications, notably phosphorylated tyrosine, lysine acetylation, and ubiquitinylation, have been

targeted effectively with "pan-specific" antibodies that bind modified peptides with little or no dependence on the local amino acid sequence [17–19]. Once modified peptides are enriched, they can be identified by LC-MS/MS [20] (Fig. 2). It has become routine to incorporate isotopic labels or other quantitative approaches to compare modifications across multiple biological conditions [21].

Similar techniques were first used to study arginine methylation more than a decade ago but only since 2012 have they improved to the point that analysis of the arginine methylation proteome could begin to become routine [22,23]. Arginine methylation is especially abundant, occurring on close to 1% of arginine residues in some types of mammalian cells [24]. Antibodies recognizing methylated arginine in the context of an Arg-Gly motif have been very effective for enriching modified proteins and peptides [22,25,26], and recent developments have extended this approach to thousands of arginine methylation sites [23]. In contrast, it has proven difficult to develop antibodies capable of enriching lysine methylation from crude protein extracts or digested peptides [26]. Only in the past year have several groups reported successful large-scale identification of lysine methylation using antibodies, while other groups have developed alternative approaches based on promiscuous methyl-lysine binding protein domains [23,27–30] (Table 1 summarizes the proteomic studies of protein methylation discussed here.) These techniques will make it possible to address system-level questions about the regulation and function of arginine and lysine methylation.

Global Approaches for Proteome-wide Analysis of Arginine Methylation

Foundations of arginine methylation proteomics using broad-specificity antibodies

The group of Stéphane Richard conducted early work to identify protein complexes containing symmetric or asymmetric di-methylated arginine [symmetric di-methyl arginine (SDMA) and asymmetric di-methyl arginine (ADMA)] using broadly specific antibodies in 2003 [22]. They used antibodies recognizing ADMA and SDMA embedded in Arg-Gly sequences frequently methylated by PRMT1 and PRMT5 (which generate ADMA and SDMA, respectively). This work identified over 200 proteins enriched by immunoprecipitation (IP) using antibodies against ADMA or SDMA but it did not directly identify any of their methylated residues. Unlike approaches that enrich modified peptides from digested protein, enriching modified proteins requires that modified residues be identified in a separate experiment or by high-coverage MS/MS of the precipitated proteins. This work suggested that arginine methylation is a widespread modification and formed a groundwork for proteomic identification of specific modified residues. Similar approaches have identified additional methylated proteins bound by the same antibodies, such as the PRMT1 substrate MRE11 [31].

Building on this idea, Ong and Mann conducted protein IP with an antibody recognizing methyl-arginine to enrich methylated proteins from lysate of HeLa S3 cells [26]. They used LC-MS/MS to identify 57 methylated arginine residues containing various methylation states of arginine across 31 proteins. Importantly, this work pioneered the use of *in vivo* heavy isotopic labeling to generate a distinctive mass shift for methylated residues. The authors prepared cells in media containing methionine labeled with deuterium and carbon-13 at the side-chain methyl group ($^{13}CD_3$). This methionine is converted by cells into *S*-adenosyl methionine (SAM) with isotopic label at the sulfonium methyl group, which is then transferred to proteins during the methylation reaction (Fig. 3). Changing the mass shift for methylation from 14 Da to 18 Da increases confidence in identification of methylated residues because it differentiates legitimate methylation sites from artifacts arising by chemical methylation, such as conversion of acidic residues to methyl esters due to methanol used during processing [32]. Ong and Mann also attempted to enrich methylated lysine but found identified peptides containing histone H3 lysine 27 and histone H4 lysine 20 from

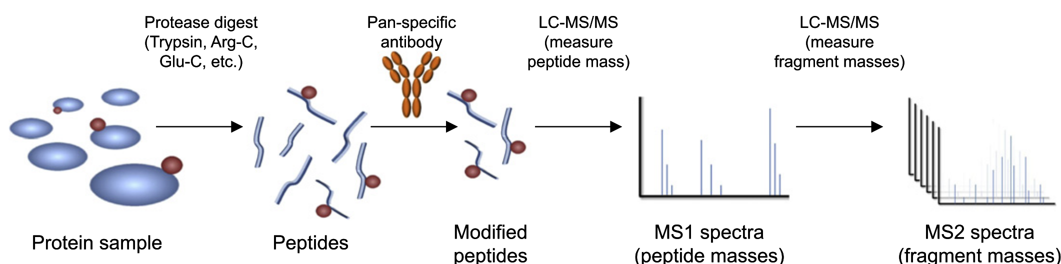

Fig. 2. Enriching and identify methylated peptides by peptide IP and LC-MS/MS.A protein sample is digested with enzymes such as trypsin, Arg-C, or Glu-C to generate a predictable mixture of peptides. Peptides containing modified residues are isolated by IP with a broad-specificity antibody and then identified using LC-MS/MS.

Table 1. Summary of proteome-wide studies identifying protein methylation

Study[a]	Year	Techniques	No. of proteins[b]	No. of sites[b]
Boisvert *et al.* [22] (Stéphane Richard)	2003	Protein IP for di-methyl-R	>200 enriched (R)	NA
Ong *et al.* [26] (Mattias Mann)	2004	Protein IP for methyl-R and K Isotopic labeling with heavy methionine in cell culture	31 (R) 2 (K)	57 (R) 2 (K)
Uhlmann *et al.* [25] (Oreste Acuto)	2012	Peptide IP for methyl-R, SCX, isoelectric focusing, HILIC	131 (R)	249 (R)
Bremang *et al.* [39] (Tiziana Bonaldi)	2013	Protein IP, heavy methionine labeling, fractionation by SDS-PAGE, isoelectric focusing	139	323 (R) 74 (K)
Fisk *et al.* [35] (Laurie Read)	2013	Protein IP for di-methyl-R followed by SCX and reverse phase chromatography (*T. brucei* mitochondria)	167 (R)	253 (R)
Lott *et al.* [36] (Laurie Read)	2013	Protein IP for di-methyl-R followed by SCX and reverse phase chromatography (*T. brucei*)	676 (R)	1332 (R)
Moore *et al.* [28] (Or Gozani)	2013	Methyl-K binding domain pull-down Heavy methionine labeling	313 enriched (K) 24 direct (K)	26 (K)
Liu *et al.* [30] (Shawn Shun-Cheng)	2013	Methyl-K binding domain pull-down Computational site prediction	109 enriched (K) 29 direct (K)	40 (K)
Cao *et al.* [27] (Benjamin Garcia)	2013	Peptide IP and SCX	413 (K)	540 (K)
Guo *et al.* [23] (Michael Comb)	2014	Peptide IP for methyl-R and K	>800 (R) 130 (K)	>1000 (R) 165 (K)
Sylvestersen *et al.* [37] (Michael Nielsen)	2014	Peptide IP for mono-methyl-R and SCX	494 (R)	1027 (R)

[a] Studies are listed by first author with corresponding author shown in parentheses.
[b] The number of modified proteins or sites is the number directly shown to be methylated, except when indicated otherwise (Boisvert *et al.*, Moore *et al.*, and Liu *et al.*). Methylation sites are indicated as arginine (R) or lysine (K) unless the original publication does not distinguish. Studies used mammalian cells or tissues unless indicated otherwise.

methyl-lysine IP. They noted that only a small fraction of proteins recovered by IP for methyl-lysine seemed to be were methylated and that they actually seemed to recover some amount of methylated arginine, indicating that available antibodies were limited in their selectivity.

Subsequent work focused on biochemical techniques to separate methylated proteins or peptides, optimizing mass spectrometry methods and improving antibodies and techniques for IP. Li and colleagues used 2-dimensional polyacrylamide gel electrophoresis (2D-PAGE) and Western blotting for methyl-arginine to select spots for identification by LC-MS/MS [33]. This approach identified a small number of methylated proteins, but the authors noted that many spots were identified as abundant proteins that are likely to co-migrate with proteins producing methyl-arginine signals in the Western blot. A similar strategy was used by Lin and colleagues to identify changes in protein methylation in LA29 rat fibroblasts expressing a temperature-sensitive form of the v-Src oncogene [34]. They used *in vitro* methylation of cell lysates from cells grown at different temperatures to identify differential methylation and 2D-PAGE with LC-MS/MS to identify likely methylated targets. As with earlier work using 2D-PAGE, this work identified a small number of likely methylated proteins. These were all very abundant proteins with roles in translation, metabolism, or cellular structure. Further improvements in antibody specificity and mass spectrometry techniques would be needed to achieve truly proteome-wide analysis of arginine methylation.

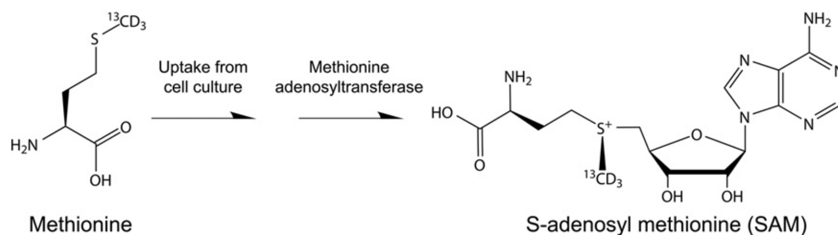

Fig. 3. *In vivo* labeling of methylated residues with heavy isotopes.Cells are grown in media containing methionine with a specific isotopic label. Labeled methionine is taken up from the media and converted to SAM by the enzyme methionine adenosyl transferase. The labeled sulfonium methyl group is finally used as a methyl donor by PRMTs and KMTs to produce methylated residues with a distinctive mass shift.

Global proteomic analysis of arginine methylation using antibodies and high-throughput LC-MS/MS

Several studies published since 2013 have improved antibodies and mass spectrometry techniques to extend proteomic analysis of arginine methylation to thousands of methylation sites. Read and colleagues used protein IP with several commercial antibodies to identify 167 arginine methylation sites in mitochondria of the parasite *Trypanosoma brucei* [35]. They separated material from each IP into 25 fractions using strong cation exchange (SCX). They increased the number of peptides identified by employing multiple strategies to fragment each peptide in the MS/MS analysis. In addition to the standard collision-induced dissociation, they fragmented peptides using electron transfer dissociation because of its utility in identifying post-translational modifications (reviewed by Afjehi-Sadat *et al.* [20]). They also increased coverage by performing additional analyses using the protease Glu-C instead of trypsin. Glu-C cleaves C-terminus to glutamic and aspartic acids depending on the buffer conditions, producing peptides that provide information complementary to those produced by trypsin. In subsequent work, the same group analyzed arginine methylation across the entire *T. brucei* proteome [36], reporting 1332 sites of arginine methylation. This work shows that the techniques for enrichment and analysis of arginine methylation have reached the point where deep coverage can be achieved. It will be exciting to see how directly these data apply to the human arginine methylome.

Comb and colleagues produced rabbit monoclonal antibodies using peptide libraries containing monomethyl arginine (MMA) at the central position, MMA in the context of an RGG motif, or ADMA at four positions [23]. These antibodies were used for peptide IP to identify sites of MMA and ADMA in HCT116 lung adenocarcinoma cells, mouse brain tissue, and mouse embryos. In total, they identified over 1000 sites of MMA and ADMA. They also used label-free quantitation to compare the abundance of arginine methylation sites between mouse brain and embryo samples. In contrast to earlier studies, this analysis did not require additional fractionation or multi-dimensional chromatography. Nielsen and coworkers used the same antibodies for peptide IP of MMA from HEK 293T and U2OS cell [37]. By combining peptide IP with fractionation using SCX, they identified 1027 sites of MMA. They also used stable isotopic labeling by amino acids in cell culture (SILAC) [38] to examine changes in arginine methylation following transcriptional arrest by actinomycin D, identifying dynamic arginine methylation on many proteins involved in transcriptional regulation. To our knowledge, these are the only examples to-date of global quantitative analysis of arginine methylation. We expect that such experiments will make it possible to examine substrates of arginine methyltransferases, track dynamic methylation during signaling transduction, and measure how methylation changes through biological processes such as cellular differentiation.

Biochemical techniques for analyzing the methyl proteome

In addition to analyzing protein methylation using antibodies, at least two recent studies have used biochemical techniques to enrich methylated peptides for identification by LC-MS/MS. Acuto and colleagues identified 41 arginine methylation sites using peptide IP and compared these sites with several strategies to enrich arginine methylation based on its chemical properties [25]. They used SCX to enrich tryptic peptides containing arginine methylation based on the fact that cleavage of methyl-arginine by trypsin is inefficient. This causes methylated peptides to contain both an internal arginine and a C-terminal arginine or lysine and, thus, to carry a charge of at least +3 at pH less than 3. In comparison, about 70% of non-methyl tryptic peptides have +2 charge at the same pH. Using SCX, they identified 39 methylated residues, most of which were not previously reported. The authors next observed that most methylated peptides observed so far are basic, with isoelectric point (p*I*) of close to 11. They used isoelectric focusing to enrich and identify 66 methylated residues, again most of which were novel. They finally observed that most methylated peptides were highly hydrophilic and reasoned that hydrophilic interaction liquid chromatography (HILIC) would be able to separate methylated from non-methyl peptides. After fractionating cell lysate using HILIC, the authors identified 215 methylation sites, most of which had not been identified using the other techniques.

A similar study was conducted by Bonaldi and colleagues [39]. They used protein IP with five antibodies recognizing various methylation states of lysine and six recognizing methylation states of arginine. Protein IP was followed by separation into fractions using SDS-PAGE and analysis by LC-MS/MS. They compared methylation sites identified using this approach to sites identified by direct separation using SDS-PAGE and LC-MS/MS, as well as by isoelectric focusing of peptides from a tryptic digest. Importantly, they incorporated isotopic labeling using the method of Ong and Mann discussed above. These authors combined light and heavy cell lysates in a 1:1 mixture so that methylated peptides could be identified by the distinctive mass shift between their two isotopic forms. This work identified in total 74 sites of lysine methylation and 323 sites of arginine methylation. In contrast to earlier work by Acuto and colleagues, this study found that protein IP and fractionation by SDS-PAGE was the most effective strategy to

identify methylated residues (they used it to identify 54 and 254 methylation sites for lysine and arginine, respectively). This difference is likely due to the use of 11 distinct antibodies and deep protein sequencing by analyzing many molecular weight fractions separated by SDS-PAGE.

A key observation of these studies is that many arginine methylation sites are not contained within the Arg-Gly motif enriched in earlier work using peptide IP. This result suggested that better pan-specific antibodies could greatly increase coverage of the methyl-arginine proteome. It is also notable that there is little overlap among the methylation sites identified by the different enrichment strategies. It may be necessary to use multiple physical and chemical approaches to achieve thorough coverage of arginine methylation. Both of these studies separated their samples into many fractions for analysis using independent runs of the mass spectrometer. This strategy can greatly increase peptide coverage but also requires a large input of time and resources. Even with improvements in automated sample handling, high-throughput LC-MS/MS, and better bioinformatics tools, it is unclear whether such approaches can be used routinely to characterize the methyl proteome across biological contexts or to examine how dynamic methylation contributes to cellular behavior.

Proteomic Strategies for Non-histone Lysine Methylation

Over the past decade, many non-histone proteins have been found to be modified by mono-, di-, or tri-methylation of lysine, and these modifications have been shown to participate in a wide range of signaling pathways and regulatory processes [1,12]. Unlike arginine methylation, early efforts to develop broad-specificity antibodies recognizing the methylation states of lysine were largely unsuccessful [26]. This has made it difficult to establish a system-level understanding of lysine methylation, instead requiring methylated proteins to be identified and investigated individually. It has been only since 2013 that approaches have become available to probe lysine methylation across the proteome using antibodies and other methyl-lysine binding proteins [23,27–29]. We next discuss early efforts toward the methyl-lysine proteome and then a series of recent breakthroughs in antibodies and protein affinity tools that have made it possible to examine lysine methylation at a proteomic scale.

Lysine methylation proteomics using antibodies

An early effort to enrich lysine methylation for proteomic analysis was described by Ong and Mann along with their work discussed earlier on methylated arginine [3]. Using methyl-lysine antibodies available at the time, they were only able to identify one methylation site each on the histones H3 and H4. Further work was conducted over the next eight years to develop antibodies against methylation states of lysine, but they all suffered from some combination of low-affinity, poor selectivity for methylation and specificity for particular peptide sequences (e.g., Iwabata *et al.* [4] describe using an methyl-lysine antibody with 2D-PAGE and Western blotting and Levy *et al.* [5] used peptide arrays to evaluate commercially available methyl-lysine antibodies).

In 2013, Garcia and colleagues developed a series of polyclonal antibodies against mono-, di-, and tri-methylated lysine [27]. They used these antibodies for peptide IP, followed by separation into up to 11 fractions using SCX to increase coverage by LC-MS/MS. In total, they reported 323 sites of mono-methylation, 127 sites of di-methylation, and 102 sites of tri-methylation across a total of 413 proteins. To our knowledge, this is the most extensive single list of lysine methylation produced so far. Similar to much of the work on arginine methylation, the authors found it necessary to separate material from the peptide IP into many fractions for analysis by LC-MS/MS, suggesting that methylated peptides make up only a small part of the total. It is likely that the extensive fractionation was required because of limited selectivity of the antibodies for methylated relative to non-methylated peptides.

Independently, Comb and colleagues immunized rabbits with peptide libraries containing mono-, di-, or tri-methylated lysine to generate pan-specific methyl-lysine antibodies [23]. Using these antibodies for peptide IP identified 132 sites of mono-methylation, 35 sites of di-methylation, and 31 sites of tri-methylation from HCT116 lung adenocarcinoma cells (representing 111, 19, and 23 distinct proteins, respectively). Unlike the Garcia laboratory, the authors did not perform additional fractionation following peptide IP. This difference may account for the larger number of peptides identified in the earlier experiment. The authors did not find any specific sequence motif enriched by their methyl-lysine antibodies, suggesting that sequence specificity was not a limiting factor. Both this work and the study by Garcia and colleagues identified many new methylated residues but also did not identify many methylated lysine residues that have been discovered by other approaches. It may be that methylated species vary greatly between biological contexts or that these antibodies have limited ability to enrich low-abundance peptides from tryptic digest of the entire proteome.

These two studies have shown that it is possible to produce broad-specificity antibodies against methylated lysine and that these antibodies can be used to identify hundreds of distinct methylation sites. There is little overlap in the methylation sites identified by these two studies, perhaps indicating that the antibodies have only partly overlapping specificity. It may also be that lysine methylation changes dramatically between

the cell types used in these studies. Future work will help determine how thoroughly these experiments probe the methyl-lysine proteome. It will also be important to establish whether the depth of coverage, reproducibility, and ease of use will allow these techniques to be used for system-level investigation into cellular signaling through lysine methylation and how dynamic lysine methylation participates in biological regulation.

Proteomics using methyl-binding domains

As an alternative to antibodies, two groups have recently used natural methyl-lysine binding domains to isolate and identify proteins modified by lysine methylation. This approach is based on the observation that some methyl-lysine binding domains, such as the triple MBT domains (3×MBT) of the protein L3MBTL1, recognize methylated lysine through a hydrophobic binding pocket that forms limited contacts with surrounding residues [40,41]. Taken out of their natural context, such domains can be used in much the same way as pan-specific antibodies [28].

We established that 3×MBT will bind to almost any protein or peptide containing mono- or di-methylated lysine (Fig. 4a). We then introduced a structure-based negative control by using a point mutant in the 3×MBT hydrophobic pocket to disrupt binding to methylated lysine [40] and used SILAC labeling to quantitatively compare proteins captured by native 3×MBT and the inactive point mutant (Fig. 4b). This work identified several hundred proteins that are consistently enriched by the MBT domain from nuclear extract of HEK 293T cells, but only directly identified methylated lysine on 26 of these. Importantly, only five of these methylation sites were previously reported, strongly suggesting that the majority of methylated residues remain to be identified. Many of the methylated proteins identified by peptide IP are strongly enriched in this experiment, but there is little overlap among the methylation sites identified directly. Advantages of this approach are that 3×MBT can be is easily produced in *Escherichia coli*, that it has well-established specificity for a broad range of peptide sequences, that protein methylation can be measured even when methylated residues are embedded in sequences difficult to identify by LC-MS/MS, and that the point mutant provides a rigorous negative control. The major disadvantage is that the domain only efficiently enriches methylated proteins; thus, additional experiments are often required to identify specific methylated residues.

In the same work, we used SILAC labeling to examine proteome-wide changes in lysine methylation after inhibiting the KMTs G9a and GLP (also called EHMT2 and EHMT1) [42]. In addition to being major methyltransferases producing di-methylation of histone H3 lysine 9 [43,44], these enzymes methylate numerous proteins *in vitro* and at least a fraction of

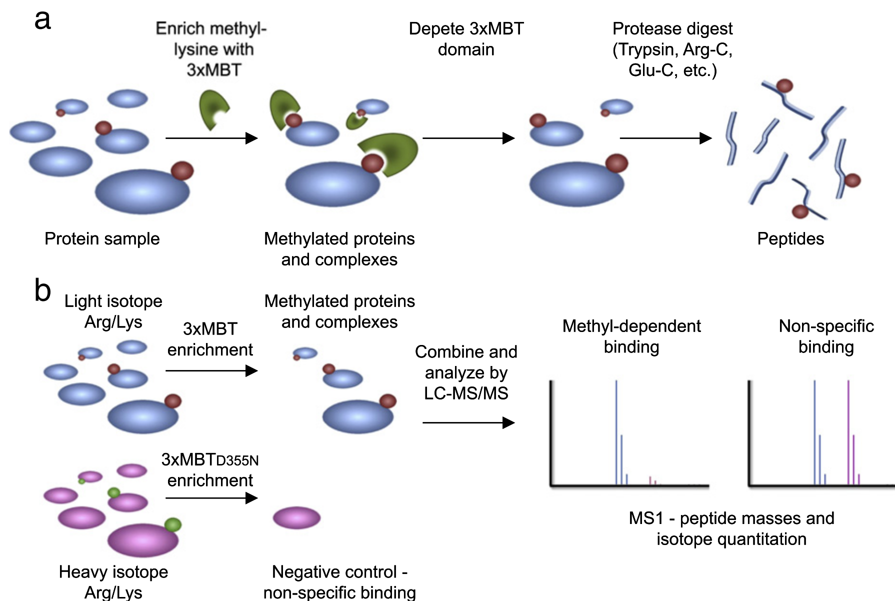

Fig. 4. Overview of enrichment and quantitative analysis of protein methylation using 3×MBT.(a) Methylated proteins are enriched from cell extract using immobilized 3×MBT [28,29]. The complex of domain with methyl-lysine is released from the beads, the domain is depleted to recover methylated proteins, and proteins are identified by LC-MS/MS. Only a subset of methylated residues is directly identified using this approach. (b) Cells are prepared in media containing isotopically labeled arginine and lysine [38]. The two extracts are incubated with immobilized 3×MBT or with the negative control 3×MBT$_{D355N}$. The two pull-downs are then combined and analyzed by LC-MS/MS to identify proteins bound in a methyl-dependent manner. The same approach can be used to compare methylation levels between different biological conditions.

these in their biological context [45,46]. We observed a strong decrease in methylation of reported G9a substrates WIZ and ACIN1, as well as several potential novel targets, but they we did not observe a decrease in bulk methylation of H3, most likely because the effect was buffered by methylation at other residues such as lysines 4, 27, and 36. The ability to monitor lysine methylation across the proteome in a quantitative manner opens the possibility of more sophisticated investigation into methyltransferase and demethylase substrates, as well as providing a route to detailed understanding of how methylation participates in dynamic signaling processes. However, these applications will greatly benefit from residue-specific information instead of total protein methylation.

Independently, Li and colleagues used a proteomic strategy to investigate methyl-specific binding targets of the chromodomain from the heterochromatin protein HP1β [30]. They used immobilized HP1β chromodomain as bait to capture methylated proteins from cell lysate, identifying a total *in vitro* interactome of 109 proteins in HEK 293T cells. They used peptide arrays to characterize the sequence specificity of HP1β chromodomain and then applied a bioinformatics approach to predict likely methylated lysine residues from the overall interactome. They tested specific predicted methylation sites using a targeted approach and confirmed methylation on 40 residues across 29 proteins, representing a 40% hit rate for predicted sites of methylation. It is notable that many proteins enriched by the 3×MBT domain are also enriched by HP1β chromodomain, but there is no overlap among the methylation sites directly identified. This may reflect the use of trypsin by our group to digest bound proteins, while the Li group used proteases Arg-C, Glu-C, chymotrypsin, and elastase. It is likely that multiple digestion and mass spectrometry strategies will continue to expand coverage of the lysine methylome.

An Integrative Perspective on the Methyl Proteome

Emerging proteomic strategies have identified thousands of methylated proteins, revealing methylation as a common post-translational modification. Understanding the biological functions of protein methylation will require better tools to analyze global methylation and methylation dynamics, as well as deeper investigation using traditional approaches of molecular biology. From a reductionist perspective, there are two key questions that need to be systematically addressed:

(1) Which methyltransferase and demethylase enzymes regulate each methylation event?
(2) How do specific methylation events regulate protein function and/or interactions with proteins, nucleic acids, or other biomolecules?

The rest of this review focuses on emerging proteome-wide and high-throughput strategies to address the first of these questions. We will examine strategies to identify non-histone targets of methyltransferase and demethylase enzymes. It will also be necessary to place these questions in a broader context. It will be important to understand how protein methylation is linked to upstream signaling pathways and regulatory processes and to discover how protein methylation acts to regulate downstream cellular decisions and biological processes. Addressing these questions will call for further development of system-level approaches, as well as experiments probing protein methylation along with other types of signal transduction such as phosphorylation.

Proteome-wide Approaches for Identifying Methyltransferase and Demethylase Substrates

Identifying proteins directly modified by any signaling enzyme is a challenging task. Fundamentally, it is difficult to know whether all the important substrates of an enzyme are expressed in any particular biological context and whether an enzyme/substrate interaction might depend on other factors such as interactions with scaffold proteins or prior post-translational modifications. In practice, it is necessary to show that a candidate substrate protein is directly modified *in vitro* and that the same modification site is regulated by the putative enzyme in cells [47].

Proteomic tools to identify PRMT substrates have had a limited impact because there are only nine known human PRMTs [3]; it is easier and more reliable to test candidate substrate by screening against the entire set of PRMTs. In contrast, there are at least 55 human proteins containing the Su(var) 3-9, enhancer-of-zeste, trithorax (SET) domain, which is responsible for most known lysine methylation [13]. Many of these predicted KMTs have no known substrates, raising the possibilities that they lack enzymatic activity or that their substrates simply have not been discovered. In addition, some members of the seven-β-strand family of methyltransferases are able to methylate lysine [48], and more such enzymes are still being discovered [49,50]. Here we review proteome-wide strategies that have been used to identify substrates of the arginine and lysine protein methyltransferases.

Microarrays

Protein microarrays allow enzymes to be screened against thousands of potential substrate proteins independent of their expression in a biological system [51–54]. Modified proteins are detected by radiolabeling [55], by broad-specificity antibodies, or

by binding to methyl-protein recognition domains [28]. The major weaknesses of microarray approaches are that there is no guarantee that every protein is correctly folded, proteins are presented outside their natural biological context of protein interactions and other biomolecules, and the process of immobilizing them on a surface may block recognition of their substrate sites [56]. It is necessary to identify methylated residues individually using targeted follow-up experiments to ensure that *in vitro* substrates are bona fide substrates in their biological context [47].

In the context of protein methylation, high-density microarrays were first used by Bedford and colleagues to identify substrates of the arginine methyltransferases PRMT1 and CARM1 (also called PRMT4) [55]. This work identified five substrates for PRMT1 and two for CARM1. They confirmed the most prominent substrates by verifying that the original bacterial stocks from the array manufacturer expressed the correct protein and that each substrate was methylated in an independent *in vitro* labeling reaction. They then screened protein fragments and short peptides to identify the precise sites of methylation by CARM1 on poly(A)-binding protein 1. It is now a routine to verify specific modified residues using mass spectrometry [20], but site-directed mutagenesis provides strong evidence that all methylation sites on a substrate protein have been identified. They finally used labeled cells in culture with ^{3}H-methionine in the presence of inhibitors of protein synthesis to show that poly(A)-binding protein 1 is methylated *in vivo* [57], and they used co-IP to show that it interacts with CARM1.

Protein microarrays have recently emerged as a promising tool for identifying methyltransferase substrates. Levy *et al.* used the ProtoArray system (Life Technologies), consisting of ~9500 human proteins expressed as a fusion with glutathione *S*-transferase and immobilized on glass slides [56]. They treated arrays with SET domains from the enzymes SETD6 and SETD7 or with glutathione *S*-transferase as a negative control. Methylated proteins were detected by radiolabeling with tritium or by binding to fluorescent broad-specificity antibodies against methylated lysine. The authors identified 26 high-confidence *in vitro* substrates of SETD6 be both detection methods, validated six out of six tested using independent *in vitro* assays, and tested two, PLK1 and PAK4, in human cell culture. It remains to be determined how many of the lower confidence hits, or hits identified by a single detection method, are real *in vitro* substrates and how many are relevant in a physiological context.

The use of carefully validated pan-specific antibodies was critical in this work. Several of the commercial antibodies tested exhibited poor selectivity on a panel of methylated peptides, non-methylated peptides, and peptides carrying other modification. Recent work has shown that methyl-binding domains such as the L3MBTL1 3×MBT domain can be used in place of antibodies to probe lysine methylation on protein microarrays [28].

Interaction screens for candidate PRMT and KMT substrates

Many signaling molecules form interactions with their substrates that are stable enough to capture by co-IP [58,59]. Co-IP is a well-established approach for identifying kinase substrates and has recently been applied to both arginine methyltransferases and KMTs [9,50]. A major advantage of this approach is that it captures enzyme/substrate interactions that may require scaffold proteins or interactions through other biomolecules such as DNA or RNA. While this can be a useful approach to generate a relatively small pool of candidate substrates, its utility is limited because many enzyme/substrate interactions are transient and difficult to capture and because many enzymes interact with large protein complexes that are not direct substrates. Overall, proteomic identification of interacting proteins is often an effective way to begin identifying substrates for a methyltransferase.

Some examples of PRMTs that have been co-purified with their substrates include ribosomal protein S2 methylated by PRMT3 [60], DNA polymerase β by PRMT6 [61], estrogen receptor α by PRMT1 [62,63], and ribosomal protein S10 by PRMT5 [64]. The same approach has been used to identify the chaperone HSP90 as a direct substrate of SMYD2 [65]. In another recent example, Falnes and colleagues identified several novel KMTs from the seven-β-strand methyltransferase family and used tandem affinity purification to identify candidate substrates for these new enzymes. In particular, they identified the molecular chaperones VCP and HSP70 as substrates of the KMTs METTL21A and METTL21D [49,50]. It remains to be determined whether these KMTs have additional substrates that do not form stable complexes.

Using chemical genetics to identify methyltransferase substrates

A common challenge facing proteomic approaches is differentiating methylation by an enzyme of interest from background methylation by endogenous enzymes. It is therefore desirable to develop strategies that specifically label substrates of an individual enzyme. Chemical genetics, often called “bump-hole”, has been widely used to identify substrates for kinase phosphorylation. In this approach, a mutation is introduced into the enzyme of interest that expands its pocket for binding adenosine triphosphate (ATP); this allows the enzyme to use modified co-factors carrying bulky unnatural

substituents [66–69]. Replacing the terminal phosphate of ATP with a radiolabel or thiophosphate allows substrates of the enzyme to be uniquely tagged for detection or capture. This strategies was pioneered by the group of Kevan Shokat more than fifteen years ago, and it has recently been applied for proteome-wide identification of many kinase substrates [59,70–73].

In 2001, Gray and colleagues demonstrated the first example of this approach applied to protein methylation. They introduced a mutation in the yeast methyltransferase RMT1 so that it would be selectively inhibited using bulky analogs of *S*-adenosyl homocysteine (Fig. 5) [74] and would be able to utilize SAM substituted with a benzyl group at the adenosine N^6 position. In Binda *et al.*, the KMT SETDB1 was shown to utilize an analog of SAM with propargyl-substituted sulfonium (Fig. 5) [75]. This causes the enzyme to transfer the propargyl group in place of normal methylation. Click chemistry could then be used to conjugate alkyne-labeled substrates to a detection moiety such as a Flag epitope tag or a fluorophore [76,77]. This SAM analog was unstable and would not work with many other KMTs.

This issue was circumvented by Luo and colleagues, who found that mutations to the SAM binding pocket of arginine methyltransferase or KMT could allow them to accept SAM analogs with larger substituents at the sulfonium position (Fig. 5) [78–80]. This allows enzyme substrates to be labeled in cell lysate using *in vitro* reactions with only minimal background from endogenous protein methyltransferases. Using this approach, they engineered the KMTs G9a and GLP to accept bulky SAM analogs (Fig. 5) [81]. They expressed these enzymes in HEK 293T cells and used lysates from these cells for *in vitro* labeling with an azide-containing SAM analog. Labeled substrates were conjugated to a strained alkyne linked to biotin that allowed substrates to be enriched using immobilized streptavidin and identified by LC-MS/MS. This work identified 82 candidate substrates of G9a and 64 of GLP. It will be exciting to see how many of these substrates have important biological functions. The authors note that their substrates include only a subset of previously reported G9a and GLP targets, perhaps because some targets are present at low level in HEK 293T cells or because *in vitro* labeling does not recapitulate all the molecular interactions present in the cellular context. The Luo group has used the same strategy to identify novel substrates of the arginine methyltransferase PRMT3 [82]. This is a promising strategy that can be applied broadly to both PRMT and KMT enzymes. As with all proteomic methods, these studies will need to be paired with direct investigation to elucidate the physiologic importance of individual substrates.

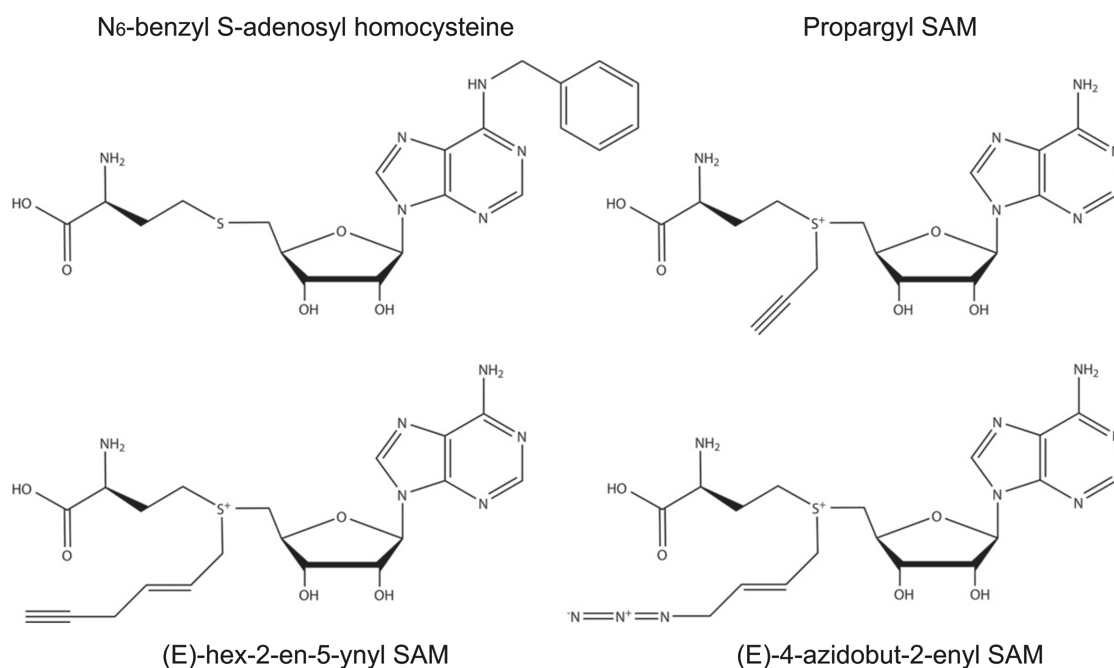

Fig. 5. Examples of SAM and *S*-adenosyl homocysteine analogs used for chemical biology. N^6-Benzyl *S*-adenosyl homocysteine can be used to selectively inhibit methyltransferase enzymes engineered with an expanded ATP binding pocket [74]. Propargyl SAM can be utilized directly by some KMTs to generate alkynlated protein substrates [75]. Larger alkyne or azide substituents at the sulfonium position can only be used efficiently by engineered methyltransferase enzymes and have improved stability in solution relative to propargyl SAM [78,81]. In each case, Click chemistry can be used to conjugate proteins labeled with alkyne or azide to moieties for enrichment (such as biotin) or a fluorescent label.

In an exciting extension of this approach, Luo and colleagues have engineered the human methionine adenosyl transferase (the enzyme responsible for synthesizing SAM from methionine) so that it can accept bulky methionine analogs introduced during cell culture. These analogs are processed into the corresponding analog of SAM in living cells. They showed that when appropriately engineered G9a or GLP are expressed in cells along with engineered methionine adenosyl transferase, the bulky alkyl group is transferred to histone H3. To track genome-wide activity of G9a and GLP at chromatin, the authors conjugated modified H3 to a cleavage biotin probe and chromatin capture followed by high-throughput sequencing. This allowed them to compare and contrast G9a and GLP activity across the genome in a way that was previously impossible. Such an approach may be a powerful tool to label and identify non-histone methyltransferases in their natural biological context.

Conclusions

Emerging proteomic tools have shown that post-translational methylation of arginine and lysine is common in species ranging from yeast to human. Non-histone methylation is likely involved in a vast range of signaling processes. Only in the last year have system-level tools to probe and measure protein methylation entered widespread use. Pan-specific antibodies have enabled high-throughput proteomic identification of arginine methylation, while tools to study lysine methylation are rapidly improving, such as broad-specificity antibodies and pan-specific methyl-lysine binding domains. Approaches adapted from work on other signaling processes, especially phosphorylation, have provided tools to map methylation signaling networks using proteomics, protein microarrays, and chemical biology. All these tools have allowed us to establish a new perspective on the broad role of protein methylation in biological regulation. They will be critical as we seek to develop an integrated understanding of how methylation interacts with other signaling processes, contributes to physiologic homeostasis, and acts in development of disease.

Acknowledgements

S.M.C. was supported in part by postdoctoral fellowship 123711-PF-13-093-01-TBE from the American Cancer Society Illinois Division. This work was supported in part by the National Institutes of Health grant R01 GM079641 to O.G.

Conflict of Interest Statement: O.G. is a co-founder of EpiCypher, Inc. Stanford University has submitted a provisional US patent describing the use of the 3×MBT domain as an affinity reagent.

Received 11 March 2014;
Received in revised form 25 April 2014;
Accepted 28 April 2014
Available online 5 May 2014

Keywords:
lysine methylation;
arginine methylation;
methyltransferase;
signaling;
proteomics

Abbreviations used:
ADMA, asymmetric di-methyl arginine; HILIC, hydrophilic interaction liquid chromatography; IP, immunoprecipitation; KMT, lysine methyltransferase; LC-MS/MS, liquid chromatography and tandem mass spectrometry; MMA, mono-methyl arginine; 2D-PAGE, 2-dimensional polyacrylamide gel electrophoresis; PRMT, protein arginine methyltransferase; SAM, S-adenosyl methionine; SCX, strong cation exchange; SDMA, symmetric di-methyl arginine; SET, Su(var)3-9, enhancer-of-zeste, trithorax; SILAC, stable isotopic labeling by amino acids in cell culture.

References

[1] Zhang X, Wen H, Shi X. Lysine methylation: beyond histones. Acta Biochim Biophys Sin (Shanghai) 2012;44:14–27.
[2] Di Lorenzo A, Bedford MT. Histone arginine methylation. FEBS Lett 2011;585:2024–31.
[3] Yang Y, Bedford MT. Protein arginine methyltransferases and cancer. Nat Rev Cancer 2013;13:37–50.
[4] Mowen KA, Tang J, Zhu W, Schurter BT, Shuai K, Herschman HR, et al. Arginine methylation of STAT1 modulates IFNalpha/beta-induced transcription. Cell 2001;104:731–41.
[5] Boisvert FM, Cote J, Boulanger MC, Cleroux P, Bachand F, Autexier C, et al. Symmetrical dimethylarginine methylation is required for the localization of SMN in Cajal bodies and pre-mRNA splicing. J Cell Biol 2002;159:957–69.
[6] Iwasaki H, Yada T. Protein arginine methylation regulates insulin signaling in L6 skeletal muscle cells. Biochem Biophys Res Commun 2007;364:1015–21.
[7] Bedford MT, Clarke SG. Protein arginine methylation in mammals: who, what, and why. Mol Cell 2009;33:1–13.
[8] Bedford MT, Richard S. Arginine methylation an emerging regulator of protein function. Mol Cell 2005;18:263–72.
[9] Han HS, Jung CY, Yoon YS, Choi S, Choi D, Kang G, et al. Arginine methylation of CRTC2 is critical in the transcriptional control of hepatic glucose metabolism. Sci Signaling 2014;7: ra19.
[10] Shi X, Kachirskaia I, Yamaguchi H, West LE, Wen H, Wang EW, et al. Modulation of p53 function by SET8-mediated methylation at lysine 382. Mol Cell 2007;27:636–46.

[11] Levy D, Kuo AJ, Chang Y, Schaefer U, Kitson C, Cheung P, et al. Lysine methylation of the NF-κB subunit RelA by SETD6 couples activity of the histone methyltransferase GLP at chromatin to tonic repression of NF-κB signaling. Nat Immunol 2011;12:29–36.
[12] Huang J, Berger SL. The emerging field of dynamic lysine methylation of non-histone proteins. Curr Opin Genet Dev 2008;18:152–8.
[13] Petrossian TC, Clarke SG. Uncovering the human methyltransferasome. Mol Cell Proteomics 2011;10 [M110.000976].
[14] Petrossian TC, Clarke SG. Multiple motif scanning to identify methyltransferases from the yeast proteome. Mol Cell Proteomics 2009;8:1516–26.
[15] Young BD, Weiss DI, Zurita-Lopez CI, Webb KJ, Clarke SG, McBride AE. Identification of methylated proteins in the yeast small ribosomal subunit: a role for SPOUT methyltransferases in protein arginine methylation. Biochemistry 2012;51:5091–104.
[16] Ficarro SB, McCleland ML, Stukenberg PT, Burke DJ, Ross MM, Shabanowitz J, et al. Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae*. Nat Biotechnol 2002;20:301–5.
[17] Choudhary C, Kumar C, Gnad F, Nielsen M, Rehman M, Walther T, et al. Lysine acetylation targets protein complexes and co-regulates major cellular functions. Science 2009;325:834–40.
[18] Zhang Y, Wolf-Yadlin A, Ross P, Pappin D, Rush J, Lauffenburger D, et al. Time-resolved mass spectrometry of tyrosine phosphorylation sites in the epidermal growth factor receptor signaling network reveals dynamic modules. Mol Cell Proteomics 2005;4:1240–50.
[19] Xu G, Paige J, Jaffrey S. Global analysis of lysine ubiquitination by ubiquitin remnant immunoaffinity profiling. Nat Biotechnol 2010;28:868–73.
[20] Afjehi-Sadat L, Garcia BA. Comprehending dynamic protein methylation with mass spectrometry. Curr Opin Chem Biol 2013;17:12–9.
[21] Altelaar AF, Munoz J, Heck AJ. Next-generation proteomics: towards an integrative view of proteome dynamics. Nat Rev Genet 2013;14:35–48.
[22] Boisvert FM, Côté J, Boulanger MC, Richard S. A proteomic analysis of arginine-methylated protein complexes. Mol Cell Proteomics 2003;2:1319–30.
[23] Guo A, Gu H, Zhou J, Mulhern D, Wang Y, Lee KA, et al. Immunoaffinity enrichment and mass spectrometry analysis of protein methylation. Mol Cell Proteomics 2014;13:372–87.
[24] Bulau P, Zakrzewicz D, Kitowska K, Wardega B, Kreuder J, Eickelberg O. Quantitative assessment of arginine methylation in free versus protein-incorporated amino acids *in vitro* and *in vivo* using protein hydrolysis and high-performance liquid chromatography. Biotechniques 2006;40:305–10.
[25] Uhlmann T, Geoghegan VL, Thomas B, Ridlova G, Trudgian DC, Acuto O. A method for large-scale identification of protein arginine methylation. Mol Cell Proteomics 2012;11:1489–99.
[26] Ong SE, Mittler G, Mann M. Identifying and quantifying *in vivo* methylation sites by heavy methyl SILAC. Nat Methods 2004;1:119–26.
[27] Cao XJ, Arnaudo AM, Garcia BA. Large-scale global identification of protein lysine methylation *in vivo*. Epigenetics 2013;8:477–85.
[28] Moore KE, Carlson SM, Camp ND, Cheung P, James RG, Chua KF, et al. A general molecular affinity strategy for global detection and proteomic analysis of lysine methylation. Mol Cell 2013;50:444–56.
[29] Carlson SM, Moore KE, Green EM, Martin GM, Gozani O. Proteome-wide enrichment of proteins modified by lysine methylation. Nat Protoc 2014;9:37–50.
[30] Liu H, Galka M, Mori E, Liu X, Lin YF, Wei R, et al. A method for systematic mapping of protein lysine methylation identifies functions for HP1β in DNA damage response. Mol Cell 2013;50:723–35.
[31] Boisvert FM, Déry U, Masson JY, Richard S. Arginine methylation of MRE11 by PRMT1 is required for DNA damage checkpoint control. Genes Dev 2005;19:671–6.
[32] Jung SY, Li Y, Wang Y, Chen Y, Zhao Y, Qin J. Complications in the assignment of 14 and 28 Da mass shift detected by mass spectrometry as *in vivo* methylation from endogenous proteins. Anal Chem 2008;80:1721–9.
[33] Hung CJ, Lee YJ, Chen DH, Li C. Proteomic analysis of methylarginine-containing proteins in HeLa cells by two-dimensional gel electrophoresis and immunoblotting with a methylarginine-specific antibody. Protein J 2009;28:139–47.
[34] Chiou YY, Fu SL, Lin WJ, Lin CH. Proteomics analysis of *in vitro* protein methylation during Src-induced transformation. Electrophoresis 2012;33:451–61.
[35] Fisk JC, Li J, Wang H, Aletta JM, Qu J, Read LK. Proteomic analysis reveals diverse classes of arginine methylproteins in mitochondria of trypanosomes. Mol Cell Proteomics 2013;12:302–11.
[36] Lott K, Li J, Fisk JC, Wang H, Aletta JM, Qu J, et al. Global proteomic analysis in trypanosomes reveals unique proteins and conserved cellular processes impacted by arginine methylation. J Proteomics 2013;91:210–25.
[37] Sylvestersen KB, Horn H, Jungmichel S, Jensen LJ, Nielsen ML. Proteomic analysis of arginine methylation sites in human cells reveals dynamic regulation during transcriptional arrest. Mol Cell Proteomics 2014 [Epub ahead of print].
[38] Ong S, Blagoev B, Kratchmarova I, Kristensen D, Steen H, Pandey A, et al. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics 2002;1:376–86.
[39] Bremang M, Cuomo A, Agresta AM, Stugiewicz M, Spadotto V, Bonaldi T. Mass spectrometry-based identification and characterisation of lysine and arginine methylation in the human proteome. Mol BioSyst 2013;9:2231–47.
[40] Min J, Allali-Hassani A, Nady N, Qi C, Ouyang H, Liu Y, et al. L3MBTL1 recognition of mono- and dimethylated histones. Nat Struct Mol Biol 2007;14:1229–30.
[41] Nady N, Krichevsky L, Zhong N, Duan S, Tempel W, Amaya MF, et al. Histone recognition by human malignant brain tumor domains. J Mol Biol 2012;423:702–18.
[42] Vedadi M, Barsyte-Lovejoy D, Liu F, Rival-Gervier S, Allali-Hassani A, Labrie V, et al. A chemical probe selectively inhibits G9a and GLP methyltransferase activity in cells. Nat Chem Biol 2011;7:566–74.
[43] Tachibana M, Sugimoto K, Fukushima T, Shinkai Y. Set domain-containing protein, G9a, is a novel lysine-preferring mammalian histone methyltransferase with hyperactivity and specific selectivity to lysines 9 and 27 of histone H3. J Biol Chem 2001;276:25309–17.
[44] Shinkai Y, Tachibana M. H3K9 methyltransferase G9a and the related molecule GLP. Genes Dev 2011;25:781–8.
[45] Rathert P, Dhayalan A, Murakami M, Zhang X, Tamas R, Jurkowska R, et al. Protein lysine methyltransferase G9a acts on non-histone targets. Nat Chem Biol 2008;4:344–6.
[46] Sampath SC, Marazzi I, Yap KL, Krutchinsky AN, Mecklenbräuker I, Viale A, et al. Methylation of a histone

mimic within the histone methyltransferase G9a regulates protein complex assembly. Mol Cell 2007;27:596–608.

[47] Berwick D, Tavaré J. Identifying protein kinase substrates: hunting for the organ-grinder's monkeys. Trends Biochem Sci 2004;29:227–32.

[48] Ng HH, Feng Q, Wang H, Erdjument-Bromage H, Tempst P, Zhang Y, et al. Lysine methylation within the globular domain of histone H3 by Dot1 is important for telomeric silencing and Sir protein association. Genes Dev 2002;16:1518–27.

[49] Kernstock S, Davydova E, Jakobsson M, Moen A, Pettersen S, Mælandsmo GM, et al. Lysine methylation of VCP by a member of a novel human protein methyltransferase family. Nat Commun 2012;3:1038.

[50] Jakobsson ME, Moen A, Bousset L, Egge-Jacobsen W, Kernstock S, Melki R, et al. Identification and characterization of a novel human methyltransferase modulating Hsp70 protein function through lysine methylation. J Biol Chem 2013;288:27752–63.

[51] Zhu H, Klemic JF, Chang S, Bertone P, Casamayor A, Klemic KG, et al. Analysis of yeast protein kinases using protein chips. Nat Genet 2000;26:283–9.

[52] Zhu H, Bilgin M, Bangham R, Hall D, Casamayor A, Bertone P, et al. Global analysis of protein activities using proteome chips. Science 2001;293:2101–5.

[53] Meng L, Michaud GA, Merkel JS, Zhou F, Huang J, Mattoon DR, et al. Protein kinase substrate identification on functional protein arrays. BMC Biotechnol 2008;8:22.

[54] Del Rincón SV, Rogers J, Widschwendter M, Sun D, Sieburg HB, Spruck C. Development and validation of a method for profiling post-translational modification activities using protein microarrays. PLoS One 2010;5:e11332.

[55] Lee J, Bedford MT. PABP1 identified as an arginine methyltransferase substrate using high-density protein arrays. EMBO Rep 2002;3:268–73.

[56] Levy D, Liu CL, Yang Z, Newman AM, Alizadeh AA, Utz PJ, et al. A proteomic approach for the identification of novel lysine methyltransferase substrates. Epigenetics Chromatin 2011;4:19.

[57] Liu Q, Dreyfuss G. *in vivo* and *in vitro* arginine methylation of RNA-binding proteins. Mol Cell Biol 1995;15:2800–8.

[58] Sharrocks AD, Yang SH, Galanis A. Docking domains and substrate-specificity determination for MAP kinases. Trends Biochem Sci 2000;25:448–53.

[59] Al-Ayoubi AM, Zheng H, Liu Y, Bai T, Eblen ST. Mitogen-activated protein kinase phosphorylation of splicing factor 45 (SPF45) regulates SPF45 alternative splicing site utilization, proliferation, and cell adhesion. Mol Cell Biol 2012;32:2880–93.

[60] Swiercz R, Person MD, Bedford MT. Ribosomal protein S2 is a substrate for mammalian PRMT3 (protein arginine methyltransferase 3). Biochem J 2005;386:85–91.

[61] El-Andaloussi N, Valovka T, Toueille M, Steinacher R, Focke F, Gehrig P, et al. Arginine methylation regulates DNA polymerase beta. Mol Cell 2006;22:51–62.

[62] Koh SS, Li H, Lee YH, Widelitz RB, Chuong CM, Stallcup MR. Synergistic coactivator function by coactivator-associated arginine methyltransferase (CARM) 1 and beta-catenin with two different classes of DNA-binding transcriptional activators. J Biol Chem 2002;277:26031–5.

[63] Le Romancer M, Treilleux I, Leconte N, Robin-Lespinasse Y, Sentis S, Bouchekioua-Bouzaghou K, et al. Regulation of estrogen rapid signaling through arginine methylation by PRMT1. Mol Cell 2008;31:212–21.

[64] Ren J, Wang Y, Liang Y, Zhang Y, Bao S, Xu Z. Methylation of ribosomal protein S10 by protein-arginine methyltransferase 5 regulates ribosome biogenesis. J Biol Chem 2010;285:12695–705.

[65] Abu-Farha M, Lanouette S, Elisma F, Tremblay V, Butson J, Figeys D, et al. Proteomic analyses of the SMYD family interactomes identify HSP90 as a novel target for SMYD2. J Mol Cell Biol 2011;3:301–8.

[66] Shah K, Liu Y, Deirmengian C, Shokat K. Engineering unnatural nucleotide specificity for Rous sarcoma virus tyrosine kinase to uniquely label its direct substrates. Proc Natl Acad Sci U S A 1997;94:3565–70.

[67] Bishop A, Shah K, Liu Y, Witucki L, Kung C, Shokat K. Design of allele-specific inhibitors to probe protein kinase signaling. Curr Biol 1998;8:257–66.

[68] Liu Y, Shah K, Yang F, Witucki L, Shokat K. Engineering Src family protein kinases with unnatural nucleotide specificity. Chem Biol 1998;5:91–101.

[69] Allen J, Lazerwith S, Shokat K. Bio-orthogonal affinity purification of direct kinase substrates. J Am Chem Soc 2005;127:5288–9.

[70] Blethrow J, Glavy J, Morgan D, Shokat K. Covalent capture of kinase-specific phosphopeptides reveals Cdk1-cyclin B substrates. Proc Natl Acad Sci U S A 2008;105:1442–7.

[71] Carlson SM, Chouinard CR, Labadorf A, Lam CJ, Schmelzle K, Fraenkel E, et al. Large-scale discovery of ERK2 substrates identifies ERK-mediated transcriptional regulation by ETV3. Sci Signaling 2011;4:rs11.

[72] Banko MR, Allen JJ, Schaffer BE, Wilker EW, Tsou P, White JL, et al. Chemical genetic screen for AMPKα2 substrates uncovers a network of proteins involved in mitosis. Mol Cell 2011;44:878–92.

[73] Berchowitz LE, Gajadhar AS, van Werven FJ, De Rosa AA, Samoylova ML, Brar GA, et al. A developmentally regulated translational control pathway establishes the meiotic chromosome segregation pattern. Genes Dev 2013;27:2147–63.

[74] Lin Q, Jiang F, Schultz PG, Gray NS. Design of allele-specific protein methyltransferase inhibitors. J Am Chem Soc 2001;123:11608–13.

[75] Binda O, Boyce M, Rush JS, Palaniappan KK, Bertozzi CR, Gozani O. A chemical method for labeling lysine methyltransferase substrates. ChemBioChem 2011;12:330–4.

[76] Kolb HC, Finn MG, Sharpless KB. Click chemistry: diverse chemical function from a few good reactions. Angew Chem Int Ed Engl 2001;40:2004–21.

[77] Best MD. Click chemistry and bioorthogonal reactions: unprecedented selectivity in the labeling of biological molecules. Biochemistry 2009;48:6571–84.

[78] Islam K, Zheng W, Yu H, Deng H, Luo M. Expanding cofactor repertoire of protein lysine methyltransferase for substrate labeling. ACS Chem Biol 2011;6:679–84.

[79] Wang R, Ibáñez G, Islam K, Zheng W, Blum G, Sengelaub C, et al. Formulating a fluorogenic assay to evaluate *S*-adenosyl-L-methionine analogues as protein methyltransferase cofactors. Mol BioSyst 2011;7:2970–81.

[80] Wang R, Zheng W, Yu H, Deng H, Luo M. Labeling substrates of protein arginine methyltransferase with engineered enzymes and matched *S*-adenosyl-L-methionine analogues. J Am Chem Soc 2011;133:7648–51.

[81] Islam K, Bothwell I, Chen Y, Sengelaub CA, Wang R, Deng H, et al. Bioorthogonal profiling of protein methylation (BPPM) using azido derivative of *S*-adenosyl-L-methionine. J Am Chem Soc 2012;134:5909–15.

[82] Guo H, Wang R, Zheng W, Chen Y, Blum G, Deng H, et al. Profiling substrates of protein arginine *N*-methyltransferase 3 with *S*-adenosyl-L-methionine analogues. ACS Chem Biol 2014;9:476–84.