POST-TRANSLATIONAL MODIFICATIONS

TIMELINE

The winding path of protein methylation research: milestones and new frontiers

Jernej Murn and Yang Shi

Abstract | In 1959, while analysing the bacterial flagellar proteins, Ambler and Rees observed an unknown species of amino acid that they eventually identified as methylated lysine. Over half a century later, protein methylation is known to have a regulatory role in many essential cellular processes that range from gene transcription to signal transduction. However, the road to this now burgeoning research field was obstacle-ridden, not least because of the inconspicuous nature of the methyl mark itself. Here, we chronicle the milestone achievements and discuss the future of protein methylation research.

Post-translational modifications (PTMs) of proteins regulate almost every aspect of cell biology by altering the functional properties of proteins, commonly in a reversible manner and at a relatively low energetic cost. Protein methylation is widespread and is perhaps one of the most functionally versatile forms of PTMs, with the capacity to impinge on essentially any cellular process. Accordingly, deregulation of protein methylation has been associated with a plethora of human disorders. First reported in 1959 (REF. 1), protein methylation initially garnered a lot of attention, but it quickly became unpopular owing to the slow progress in understanding the physiological impact of this modification. Then, at the turn of the 21st century, fueled by a series of advances in molecular biology, the field of protein methylation flourished, and numerous discoveries were made that unveiled much of the biology of protein methylation that we know today.

In this Timeline article, we track the history of research into protein methylation, from its arduous beginnings until modern times. We point out several landmark discoveries that laid the foundations of the field, which include the discovery of the first methylated proteins, methyl-modifying enzymes (methyltransferases and demethylases; also known as methyl 'writers' and 'erasers', respectively), 'reader' proteins as effectors of methylation, the major biological processes that are regulated by this PTM, and the crosstalk between protein methylation and cellular metabolism. We discuss the historical reliance of the field — in particular, studies related to chromatin biology — on the availability of tools and technologies, and consider the prospect of novel quantitative proteomic approaches. We conclude with a summary of some of the key unanswered questions and challenges, and discuss potential directions that the field might take in the future.

Thrilling early discoveries

The story of protein methylation finds its deepest roots in studies of the metabolism of Lys, which Fred Sanger carried out as part of his Ph.D. thesis under the mentorship of Albert Neuberger, and which earned him a doctorate in 1943. In his work, the young Sanger observed that ε-*N*-methyl-lysine (in which the methyl group is attached to the side-chain amino group of Lys; herein referred to as methyl-lysine), but not α-*N*-methyl-lysine (in which the methyl group is attached to the terminal amino group of Lys), can readily substitute for Lys — an essential amino acid — in the diet of weaning rats[2]. Sanger reasoned that this was because only the ε-*N*-methyl compound could be converted into Lys or incorporated into proteins. He favoured conversion to Lys as the more likely of the two processes, in light of his own experiments showing that 'demethylation' of the ε-amino group of Lys readily occurs in kidney slices[3]. These findings were far ahead of their time and touched on several facets of an unsuspected regulatory mechanism, the inner workings of which did not become clear until the 21st century. However, Sanger made another far-reaching contribution to the field of protein methylation: the techniques of protein sequence determination he developed laid the foundation on which many of the early discoveries in the field were based.

The occurrence of protein methylation in living cells was first reported by Richard Ambler and Maurice Rees in 1959 (REF. 1). While analysing bacterial flagellar proteins (flagellins), Ambler observed that flagellin from *Salmonella typhimurium* contained a substantial amount of methyl-lysine, "an amino acid that has not been previously found to occur naturally" (REF. 1) (BOX 1). Two years later, Ambler and colleagues found that a gene separate to the gene encoding flagellin determined the presence or absence of methyl-lysine in flagellin, the first observation to suggest that methylation might occur post-translationally[4]. Given that several but not all lysines in flagellin were modified, Ambler further speculated on the existence of an enzyme that might methylate select Lys residues in a pre-formed protein. Together, the work of Sanger and Ambler set the methodological and intellectual base that instigated much of the progress in the field in the ensuing 30 years (FIG. 1).

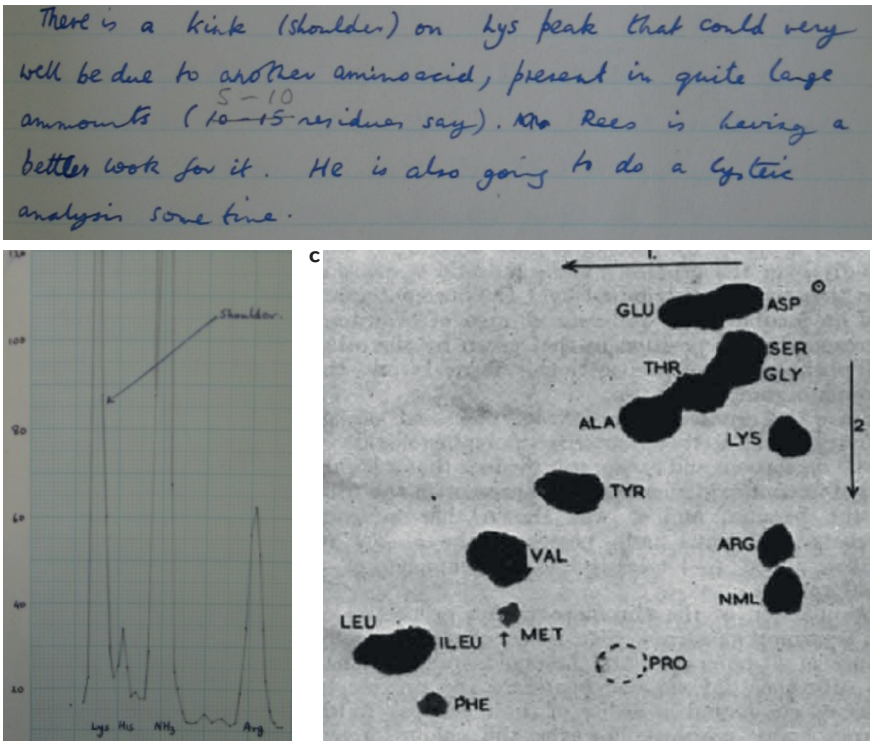
Following on from these studies, the detection of methyl-lysine in hydrolysates of mammalian histones in 1964 by Kenneth Murray was a discovery of particular significance[5]. In contrast to Ambler's flagellin, Murray observed consistently low and highly heterogeneous levels of the methylated residue in different histone fractions, which led him to speculate that histone methylation might be a highly specific process[5]. However, Murray was unable to tell whether

methylation occurred before or after histone synthesis. This issue was elegantly resolved by two independent studies, which showed that histones were methylated after synthesis of their polypeptide chains was completed. First, Allfrey and colleagues demonstrated that the translational inhibitor puromycin effectively blocked the incorporation of amino acids but had no effect on the incorporation of methyl groups into histones[6]. Consistent with this result, in 1965 Kim and Paik found that methyl-lysine could not be conjugated to tRNAs, which ruled out the possibility of its incorporation during translation[7].

Despite the intriguing insights that were gathered by the mid-1960s, the newborn field of protein methylation was missing a critical component in order to prosper — biological relevance. A major conceptual advance came with the demonstration in 1964 by Allfrey and colleagues that "acetylation of the histones can lower their effectiveness as inhibitors of RNA polymerase reaction" (REF. 6). This formed the basis for their visionary hypothesis that histone modifications (including methylation) could regulate gene transcription, a premise that sparked a lot of excitement across numerous scientific communities.

Box 1 | **Discovery of protein methylation**

Richard P. Ambler (1933–2013) was an English molecular biologist and a skilled practitioner of the techniques for protein sequence determination that were developed by, amongst others, his Ph.D. and postdoctoral adviser, Fred Sanger. Ambler employed protein sequencing to study the evolutionary relationships between organisms, predominantly bacteria. He is best known for having published, in 1963, the first sequence of a bacterial protein, and for his groundbreaking research into horizontal gene transfer, the primary reason for the spread of antibiotic resistance in bacteria. As a graduate student at the University of Cambridge, UK, Ambler worked together with Maurice W. Rees of the Agricultural Research Council's Virus Research Unit in Cambridge, UK, on the amino acid composition of flagellin in different strains of bacteria. Commenting on the amino acid analysis performed on 11 May 1958, Ambler wrote in his laboratory notebook: "There is a kink (shoulder) on Lys peak that could very well be due to another amino acid, present in quite large amounts (10–15 residues say [numbers corrected to 5–10]). Rees is having a better look for it. He is also going to do a cysteic analysis some time." (see the figure, part **a**). Ion-exchange chromatography with the indicated 'shoulder' (see the figure, part **b**) and two-dimensional paper chromatography of hydrolysates of *Salmonella typhimurium* flagellin (see the figure, part **c**) indicated the presence of a 'new amino acid', ε-*N*-methyl-lysine (NML), in bacterial flagellins. These analyses, together with an elemental analysis of the purified NML, provided the key evidence for the occurrence of protein methylation in living cells. Images in parts **a** and **b** courtesy of A. Ambler, UK. Part **c** reproduced with permission from REF. 1, Macmillan Publishers Limited.



Decades of stagnation

The apparent involvement of enzymes and their conservation in species as evolutionarily distant as bacteria and mammals suggested that protein methylation might serve important physiological functions. This seemed even more likely in light of the discovery of other PTMs at the time, which included protein phosphorylation in 1954 and acetylation in 1963, whose functions in the regulation of biological processes were just beginning to emerge[6,8–11]. However, much to the frustration of many aspiring investigators, no such physiological roles for protein methylation were documented until the late 1990s (FIG. 1). Given the complexity of protein methylation (BOX 2), this is perhaps not surprising from our current perspective.

With functional data lacking, the focus of the nascent field of protein methylation shifted to identifying the putative enzymes involved. Paik and Kim were particularly productive in this area in the 1960s and early 1970s. Together with other investigators, they defined the first methyltransferase activity that could transfer a methyl group from *S*-adenosylmethionine (SAM) to Lys, Arg, Asp or Glu residues of proteins[7,12–14]. These studies laid the groundwork for the identification of new methylated proteins as well as definition of different methylation states of Lys (mono-, di- or trimethyl) and Arg (N^G-monomethyl or N^GN^G-dimethyl, which could be symmetric or asymmetric; N^G refers to either of the two terminal nitrogen atoms of the guanidino group) (BOX 2), all of which were described by 1971 (REFS 15–18). In 1964, building on the early work of Sanger in the 1940s, Paik and Kim reported the identification of an ε-alkyllysinase, which removed methyl groups from free methyl-lysine[19]. They further reported a cellular activity — curiously enriched in the kidney — that appeared to remove methyl groups from histones, resulting in the production of formaldehyde[3,19,20]. However, the formation of the crucial reaction product, the demethylated histones, was not demonstrated. It also remained unclear whether this cellular activity was due to the ε-alkyllysinase or another factor. At around the same time, two other groups found that the turnover rate of methyl groups on histones was comparable to that of the histones themselves, which suggested that histone methylation was irreversible[21,22]. This latter view remained the dogma in the field until it was disproved three decades later.

Despite the mechanistic insights gathered by the early 1970s, the continued lack of evidence for a biological role rendered protein methylation a rather unattractive subject of study. Isolated indirect links to biological processes began to emerge in the early 1990s from studies of carboxyl methylation (that is, methylation of terminal and side-chain carboxyl groups, which results in the formation of methyl esters) and Arg methylation of various proteins implicated in signalling pathways and RNA processing, although in several cases the specific roles of methylation were not determined[23–26].

The golden age

The long struggle to establish a definite functional role for protein methylation finally ended in the late 1990s owing to progress in genetics and molecular biology. The increased availability of genomic sequence information and the development of techniques to study chromatin biology were particularly critical to the rebirth of the field. Chromatin immunoprecipitation (ChIP) coupled with analysis of the precipitated DNA provided time- and location-resolved information about the *in vivo* associations of histones, their modifications and regulatory proteins with specific genomic regions[27,28]. Another major boost to the field came in 1996 with the identification of the first histone-modifying enzymes, an acetyltransferase and a deacetylase, by the groups of Allis and Schreiber, respectively[29,30]. Together, these advances not only led to a resurgence of interest in protein methylation but also swayed the field towards chromatin biology — in particular, histone methylation. However, the first evidence of a regulatory role for protein methylation came from the 'non-histone' world.

In 1996, a year after Klein and Houtz cloned the first gene that encodes a methyl-modifying enzyme, a ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) large subunit *N*-methyltransferase (LSMT; encoded by *RBCMT*) (FIG. 1)[31], the groups of Silver and Herschman isolated the gene for protein arginine *N*-methyltransferase 1 (PRMT1) from yeast and mammalian cells, respectively[32,33]. With the cloned genes in hand, the investigators were able to produce large amounts of defined methylating enzymes and study their properties. In 1998, the function of Arg methylation — catalysed by the yeast PRMT1 homologue Hmt1 — in the nuclear export of heterogeneous nuclear ribonucleoproteins (HNRNPs) was demonstrated, which provided the long sought-after evidence that protein methylation is indeed physiologically important[34] (FIG. 1).

Linking histone methylation to epigenetic control. More than three decades after its proposal, Allfrey's hypothesis about the impact of histone methylation on gene transcription could finally be tested. The first evidence that supported a role for

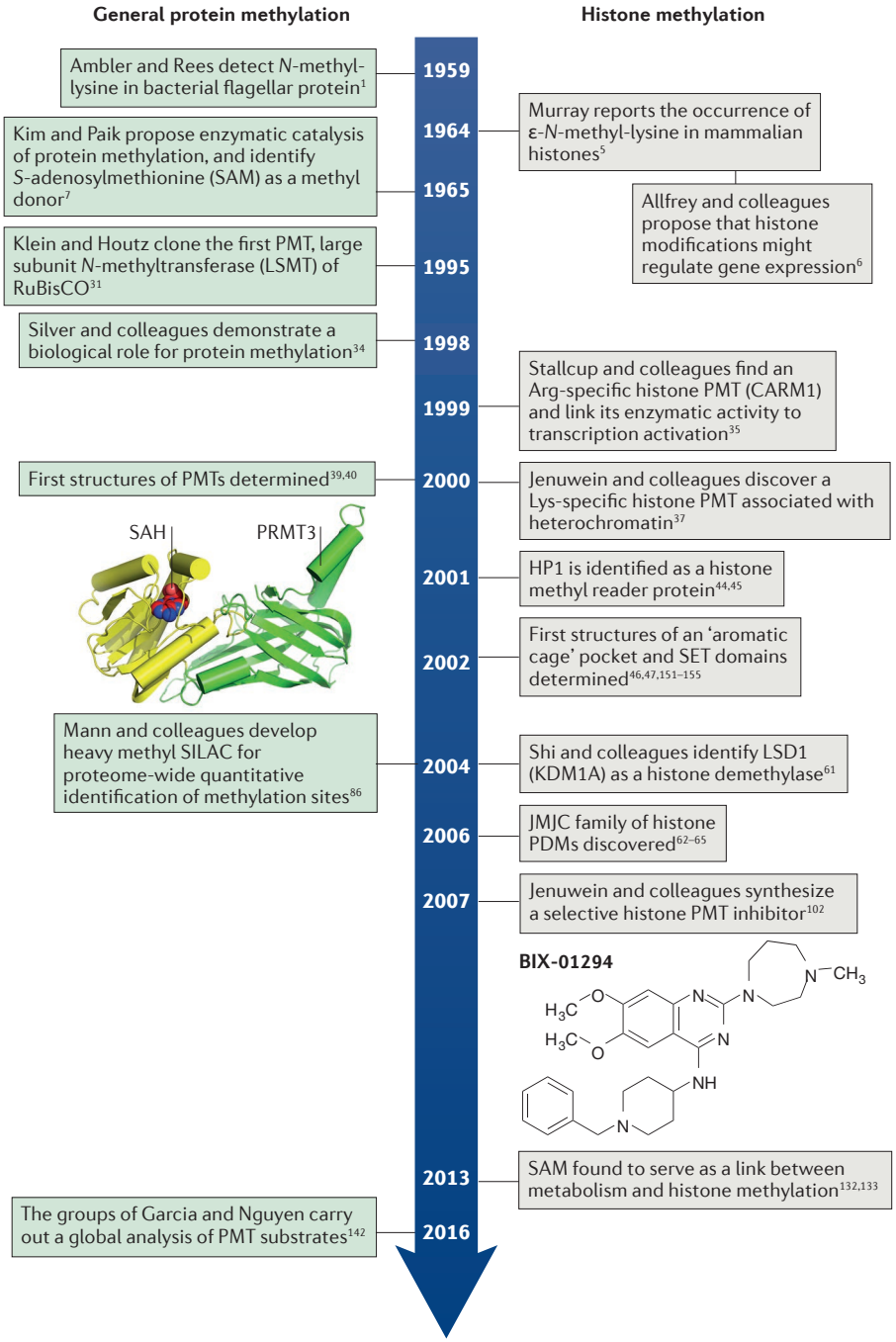


Figure 1 | **Milestones in protein methylation research.** Events listed on the right of the timeline are associated with histone methylation, whereas those on the left refer to more general studies of protein methylation (including non-histone proteins). The first structures of su(var)3-9, enhancer-of-zeste and trithorax (SET) domain-containing enzymes are described in REFS 152–156. CARM1, coactivator-associated arginine methyltransferase 1; HP1, heterochromatin protein 1; JMJC, Jumonji C-terminal domain; KDM1A, lysine-specific histone demethylase 1A; LSD1, lysine-specific demethylase 1; PDM, protein demethylase; PMT, protein methyltransferase; RuBisCO, ribulose-1,5-bisphosphate carboxylase/oxygenase; SAH, S-adenosylhomocysteine; SILAC, stable isotope labelling with amino acids in cell culture.

histone methylation in transcriptional regulation was provided in 1999 by Stallcup and co-workers[35]. Using a yeast two-hybrid screen and a sequence comparison with the PRMT enzymes known at that time, the team discovered an Arg-specific protein methyltransferase (PMT), coactivator-associated arginine methyltransferase 1 (CARM1; also known as PRMT4), which methylated histones *in vitro*. Markedly, mutation of conserved residues within the SAM-binding domain of CARM1 abolished its methyltransferase activity concomitant with a substantial drop in transcriptional activation[35].

A few months later in the same year, the team of Allis made the case for Lys methylation. Using biochemical approaches and microsequencing of *Tetrahymena thermophila*, yeast and mammalian histones, the investigators showed that Lys4 of H3 (H3K4) is a highly conserved site of methylation and suggested that it may facilitate transcription[36]. Then, in 2000, Jenuwein and colleagues observed sequence similarity between the heterochromatic protein SUV39H1 and the RuBisCO LSMTs. This similarity, together with elegant biochemical and gene ablation experiments led the group to identify SUV39H1 as the first Lys-specific histone PMT, and to link methylation of its target H3K9 to the establishment of transcriptionally less-permissive heterochromatin[37]. Of note, this report mapped the methyltransferase activity of SUV39H1 to the conserved su(var)3-9, enhancer-of-zeste and trithorax (SET) domain, which is the catalytic core of the largest clade of Lys-specific PMTs known to date[38].

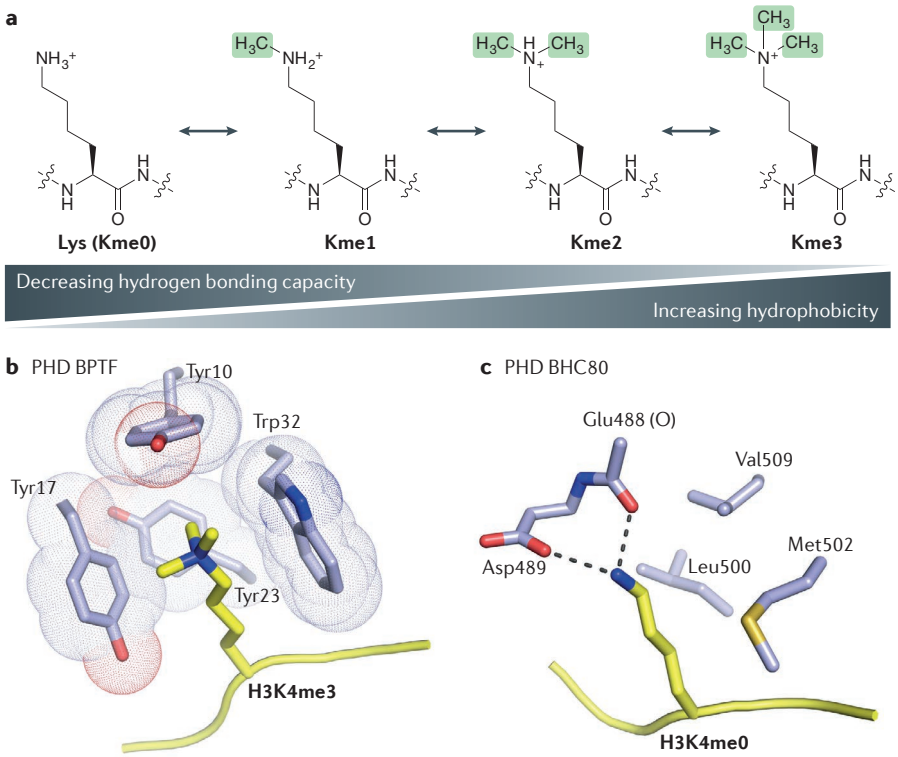
These studies set the stage for a flurry of discoveries of additional histone PMTs, their structures[39,40], their interacting proteins and their biological roles. In just over a decade, some fifty SET domain-containing proteins, many of which were validated as Lys-specific PMTs, were identified. This was accompanied by the discovery of nine Arg-specific seven-β-strand PMTs, designated PRMT1–PRMT9, as well as the H3K79-specific PMT DOT1-like protein (DOT1L). Altogether, these enzymes were capable of methylating all canonical Lys and Arg residues on histones, which include Arg2, Lys4, Arg8, Lys9, Arg17, Arg26, Lys27, Lys36 and Lys79 on H3, and Arg3 and Lys20 on H4 (reviewed in REF. 41). Notably, the location and the degree of methylation of a particular residue, determined by antibody-based and quantitative proteomic approaches as well as the co-occurrence of other PTMs, could often be associated with a particular transcriptional state or chromatin structure.

Also in 2000, the amalgamation of diverse associations of histone methylation, acetylation and phosphorylation with transcription and chromatin structure kindled the hypothesis that distinct patterns of histone marks might encode a language or 'histone code' (REF. 42). Although there

Box 2 | **Biophysical properties and the readout of protein methylation**

Methylation stands out among other common post-translational modifications (PTMs) in terms of several biological and physicochemical properties. First, methylation can occur on the side chains of at least nine of the twenty common amino acid residues (Met, Cys, Lys, Arg, His, Gln, Asn, Glu and Asp)[26], although Lys and Arg are by far the most commonly methylated residues. Second, the methyl group is one of the smallest PTMs and contributes relatively little to the 'steric bulk' of the modified side chains. Third, methylation of Lys or Arg does not affect the overall charge of these residues, and their side chains remain positively charged even when methylated[149]. Fourth, a Lys residue can be methylated up to three times and an Arg residue up to two times. Each subsequent methylation event removes a proton from the ε-amino group of Lys or the guanidino group of Arg, thereby decreasing their hydrogen-bonding capacity and increasing their hydrophobicity (see the figure, part **a**)[149]. Fifth, the turnover of histone methylation is much slower compared with some other abundant PTMs, such as phosphorylation and acetylation[150,151]. The particularly slow kinetics of di- and trimethylation of histone Lys residues was proposed to contribute to epigenetic stability[150]. A lack of data on the kinetics of non-histone methylation, however, precludes extrapolation of these findings to non-histone proteins.

Methylated and unmodified Lys or Arg residues are recognized by specific 'reader' proteins, usually in a manner that is dependent on the flanking amino acid sequence. Methylated residues are typically recognized by 'aromatic cage' pockets, as exemplified by recognition of trimethylated Lys4 of histone 3 (H3K4me3) by the plant homeodomain (PHD) finger of bromodomain and PHD finger-containing transcription factor (BPTF; Protein Data Bank identifier (PDB ID): 2F6J) (see the figure, part **b**)[50]. The aromatic rings of two to four residues that line this aromatic cage stabilize a methylated side chain by cation–π interactions (the attraction force between a cation and the partially negatively charged π-surface of an aromatic ring), as well as by hydrophobic contacts[48]. By contrast, an unmodified Lys can be recognized by acidic residues through electrostatic and hydrogen-bonding interactions, as in the case of the recognition of H3K4me0 by the PHD finger of BRAF35–HDAC2 complex 80 kDa protein (BHC80; also known as PHF21A) that involves the formation of two hydrogen bonds (PDB ID: 2PUY) (see the figure, part **c**). The selectivity of readers for different Lys methylation states is typically achieved by tuning the ratio of aromatic versus acidic residues in the aromatic cage[149].



have been different views about this hypothesis[43], it is now generally accepted that histone modifications — including methylation — serve as signals for the recognition by effector or reader proteins, which impact chromatin structure and function (FIG. 2a, top panel). The first experimental support for this concept came in 2001 from a functional scrutiny of the SUV39H1-associated heterochromatin protein 1 (HP1) by the groups of Jenuwein and Kouzarides[44,45]. HP1 had been known to localize to heterochromatin and to facilitate gene silencing, but the reason why these processes were strictly dependent on SUV39H1 and the conserved chromodomain of HP1 were obscure. The investigators found that by methylating H3K9, SUV39H1 creates an epitope that is specifically recognized by the HP1 chromodomain, which allows the recruitment of HP1 to heterochromatin and, in turn, the HP1-dependent propagation of transcriptional silencing. The crystal structure of HP1 in a complex with trimethylated H3K9 (H3K9me3) was determined in 2002 and provided the first insight into the principles that underlie methyl-lysine recognition by the 'aromatic cage' pockets[46,47] (BOX 2).

A rapid flow of discoveries that followed unveiled the largest family of methyl readers, the plant homeodomain (PHD) finger-containing proteins, as well as numerous additional reader modules, which include Tudor domains, malignant brain tumour (MBT) repeats, ATRX–DNMT3–DNMT3L (ADD) domains, chromodomains, WD40 repeats and others (reviewed in REFS 48–50). Different readers, together with their associated effector complexes, were found to mediate different biological outputs through site-specific recognition of methylated Lys or Arg residues on a histone tail. For example, recognition of H3K4me3 at active promoters by TBP-associated factor 3 (TAF3), a component of the basal transcription factor TFIID, could explain the crosstalk between H3K4me3 and positive regulation of RNA polymerase II-mediated gene transcription[51].

Demethylases join the party. With reports on histone methylation pouring in during the early 2000s, several cases made it clear that mechanisms for the removal of methyl groups must also exist. For example, H3K9 methylation, which enables repression of the cyclin E gene during the G1 phase of the cell cycle, needs to be removed every time cells enter S phase and to be reinstated

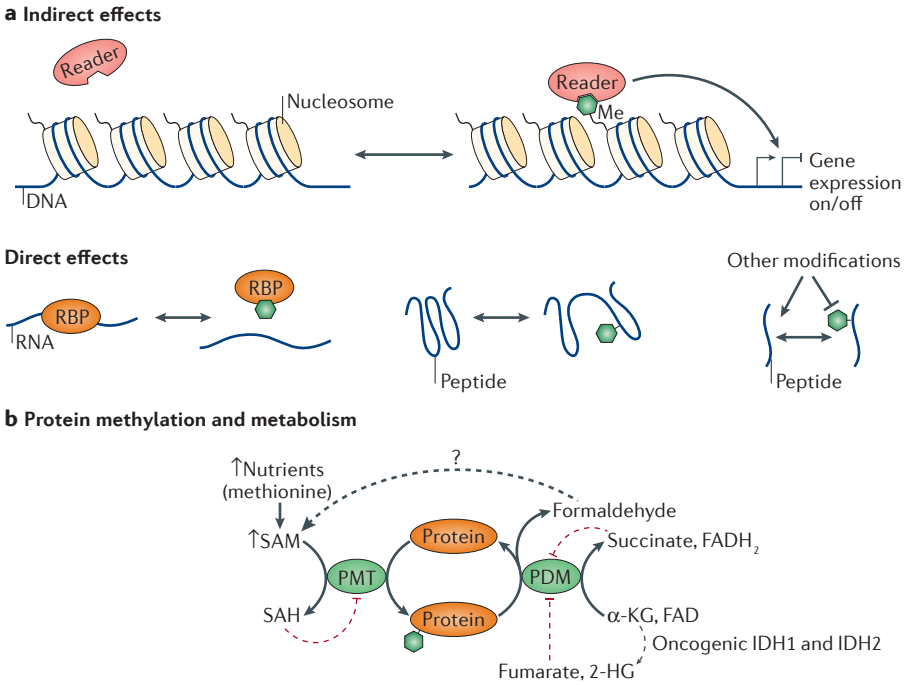


Figure 2 | **Interfaces between protein methylation and biological processes. a** | Regulatory potential of protein methylation. Protein methylation is thought to exert its biological function primarily indirectly, by recognition of methyl marks by effector proteins (also termed 'readers'; top panel). For example, readers that recognize histone methyl marks enact transcriptional changes or recruit other proteins to do so. Growing evidence also supports direct actions of protein methylation (bottom panel), which include modulation of the interaction of RNA-binding protein (RBPs) with RNA (bottom left panel); changes to protein structure (bottom centre panel), as is the case for dimethylation of Lys79 of histone 3 (H3K79me2) and H4K20me3, which regulate nucleosome surface and higher-order structure of chromatin, respectively[157]; and competition with other PTMs such as the acetylation of H3K9 and H3K27 (REFS 158,159) (bottom right panel). **b** | Metabolism can, in principle, affect protein methylation through changes in the availability of cofactors (which include *S*-adenosylmethionine (SAM), α-ketoglutarate (α-KG) and flavin adenine dinucleotide (FAD)) of protein methyltransferases (PMTs) and protein demethylases (PDMs), as well as through the generation of reaction products, which include *S*-adenosylhomocysteine (SAH), succinate and a reduced form of FAD, $FADH_2$, all of which are intermediate metabolites. In addition, fumarate, a metabolite of the citric acid cycle, has been shown to modulate PDM activity by acting as a competitor of α-KG. It has also been reported that oncogenic mutations in the genes that encode cytoplasmic and mitochondrial isocitrate dehydrogenases (IDHs; encoded by *IDH1* and *IDH2*, respectively) lead to the generation of 2-hydroxyglutarate (2-HG) from α-KG, which results in PDM inhibition. The question mark denotes speculation that the formaldehyde generated in a demethylation reaction could be converted back to SAM[53]. Dashed lines indicate effects for which the mechanism is incompletely understood.

soon afterwards[52]. How could such dynamic changes in methylation be accomplished? In 2002, Kouzarides and colleagues discussed three possible scenarios: histone exchange, histone tail clipping and active demethylation[53]. The first two possibilities were supported by some experimental evidence[54,55], whereas active demethylation was not. In the face of the longstanding dogma that histone methylation was an irreversible process, the existence of enzymes that would catalyse active removal of methyl groups seemed unlikely.

In 2004, our attention was drawn to a known but uncharacterized component of several co-repressor complexes, KIAA0601, a conserved protein that shares extensive sequence similarity with flavin adenine dinucleotide (FAD)-dependent amine oxidases[56–60]. Using a number of biochemical and *in vivo* assays, we identified KIAA0601 as a Lys-specific demethylase — hence we termed it LSD1 (also known as KDM1A) — which has substrate specificity for H3K4me1 and H3K4me2 (REF. 61). The enzymatic activity of LSD1 provided definitive proof that protein methylation is a reversible process that is catalysed by demethylases (FIG. 1). A year later, LSD1 was shown to also possess H3K9me1 and H3K9me2 demethylase activity[62], but the biochemical mechanism has remained incompletely understood. The discovery of LSD1 thus dismissed the decades-old

dogma that histone methylation was static and irreversible, and revealed the dynamic nature of the regulation of histone methylation.

As recombinant LSD1 was not reactive towards lysines at other positions on H3 and could not demethylate H3K4me3 (REF. 61), we speculated that additional histone protein demethylases (PDMs) must exist. In 2006, several groups, beginning with Zhang and colleagues, discovered a class of α-ketoglutarate (α-KG)-dependent Jumonji C-terminal domain (JMJC)-containing PDMs that used an alternative mechanism for demethylation[63–66] (FIG. 1; see also FIG. 2b). In contrast to LSD1, which oxidized the ε-amino group of Lys, the JMJC PDMs were found to oxidize the attached methyl groups, a mechanism that allowed for the demethylation of mono-, di- and trimethyl-lysine. The numerous JMJC PDMs discovered since then were found to target all key methylated histone lysines, with the exception of H3K79 (reviewed in REF. 67).

Probing methylation of non-histone proteins. The expansion in the number of histone methylation studies in the 2000s largely overshadowed the research of other methylated proteins. This bias was heavily influenced by a relatively slow development of quantitative proteomic versus genomic tools and the simple fact that most proteins are orders of magnitude less abundant than histones. Nonetheless, the realization that many histone methyl modifiers and readers also targeted non-histone proteins had begun to illuminate the vast regulatory scope of protein methylation.

A pioneering report in 1998 on methylation-controlled HNRNP shuttling indicated that protein methylation might be involved in RNA biology[34]. Indeed, within a few years, methylation of individual RNA-binding proteins (RBPs) had been associated with regulation of RNP assembly and localization, pre-mRNA splicing, mRNA stability and small RNA biogenesis[68–74]. In contrast to histones, for which methylation of lysines has drawn most attention, Arg has emerged as the predominant methylated residue in RBPs. Methylarginines were found to facilitate the recognition of RBPs by Tudor domain-containing readers, or were proposed to directly interfere with RNA binding[24,75,76]. By the early 2000s, most known Arg-methylated proteins were found to be associated with RNA metabolism[77], and overrepresentation of RBPs among the Arg-methylated proteins was confirmed in later proteome-wide studies (see below).

In 2004, Reinberg and colleagues reported that a SET domain-containing protein 7 (SETD7)-dependent methylation of a specific Lys in the TP53 tumour suppressor protein promoted its nuclear stability, transcriptional activity and cellular apoptosis[78]. Remarkably, follow-up studies showed that methylation of four lysines and three arginines dictated the activity of TP53, either as an activator or as a repressor, in a site-specific manner[79–82]. In 2007, the team of Berger found that TP53, as the first non-histone protein, could also undergo enzymatic demethylation. The investigators found that by demethylating K370me2 of TP53, LSD1 removed the recognition site for the methyl reader p53-binding protein 1 (53BP1), thereby repressing TP53 function in the DNA damage response[83]. The reversibility, site-specific effects and readout of TP53 methylation suggested that this PTM could provide similar regulatory complexity for histones and non-histone proteins (FIG. 2a).

The dawn of protein methylomes

The few-dozen methylated proteins known by the early 2000s were mostly nuclear proteins that participate in gene transcription and RNA metabolism[23,77]. However, methylated substrates such as myosin[84], cytochrome *c*[85], and myelin basic protein[86], together with the apparent promiscuity of a growing number of histone PMTs, had many wondering whether protein methylation might serve a broader regulatory role. However, the question could not be adequately addressed with conventional biochemical approaches alone. Once again, the field required a technological breakthrough.

In 2004, Mann and colleagues described a revolutionary mass spectrometry-based approach, called heavy methyl stable isotope labelling with amino acids in cell culture (heavy methyl SILAC), for the quantitative identification of methylation sites *in vivo*[87] (FIG. 1). The method relied on labelling of methylation events with a 'heavy' methyl group comprising carbon-13 and three deuterium atoms, $^{13}CD_3$, which was achieved through the endogenous conversion of the supplied [$^{13}CD_3$] methionine to the sole biological methyl donor, [$^{13}CD_3$]*S*-adenosylmethionine. Heavy-methyl labelling *in vivo* followed by antibody-based enrichment of methylated peptides allowed for proteome-wide identification of methylation events by mass spectrometry. Using a methylarginine-specific antibody, these investigators detected 59 methylation sites in 33 different proteins, which considerably expanded the repertoire of known methylation events[87].

The capability to reliably detect methylation events provided an opportunity to obtain a global view of protein methylation. However, the difficulty of generating methyl-specific — in particular, methyl-lysine-specific — antibodies prevented a more efficient enrichment of methylated proteins and in-depth analysis of methylproteomes by mass spectrometry[87,88]. Intense efforts over the next decade to tackle this challenge led to the development of 'pan-methyl' antibodies with improved affinity[89,90], as well as alternative enrichment strategies, including the use of natural methyl-binding domains (chromodomains or MBT repeats)[91,92] and chemical labelling methods[93–96]. To date, these efforts — in combination with high-resolution mass spectrometry — have identified, in total, over 16,000 unique Lys and Arg methylation events in more than 5,500 human proteins (FIG. 3).

One of the more intriguing predictions that emerged from these proteomic analyses was that protein methylation, which is kinetically slow (at least for histones), might have important regulatory roles in some of the most dynamic cellular processes, including signal transduction, which is typically driven by phosphorylation cascades (reviewed in REF. 97; see also BOX 2). Indeed, Lys and Arg methylation of components of the major signalling pathways, such as the mitogen-activated protein kinase (MAPK), WNT, nuclear factor-κB (NF-κB) and Janus kinase (JAK)–signal transducer and activator of transcription (STAT) pathways, was shown to either promote or impede phosphorylation to affect the strength and duration of signalling[98–102].

The dramatic increase in the number of known methylproteomes established protein methylation as a globally important PTM that has regulatory roles in most essential cellular processes, which range from gene transcription and RNA processing to protein translation and cell signalling.

Drugging protein methylation

Functional and structural characterization of methyl writers, readers and erasers in the 2000s paved the way for probing protein methylation using small-molecule inhibitors. Of primary interest was the potential for therapeutic targeting of methyl-modifying enzymes — in particular, Lys-specific histone PMTs and PDMs that

are associated with chronic conditions such as cancer and inflammation. In contrast to the wide and shallow methyl-binding pockets of most readers[50], writers and erasers were found to contain at least two domains that could be readily targeted for enzymatic inhibition: the substrate channel and the cofactor-binding site.

The first selective methylation inhibitor, BIX-01294, was reported in 2007 (REF. 103) (FIG. 1). As a substrate-competitive inhibitor, BIX-01294 exhibited selectivity for euchromatic histone-lysine *N*-methyltransferase 2 (EHMT2; also known as G9A) and the closely related EHMT1 (also known as GLP1), both of which methylate H3K9 and several non-histone substrates[104,105]. Although cytotoxic and with limited affinity, BIX-01294 served as the lead compound for the development of highly selective and potent inhibitors of EHMT1 and EHMT2; some of these compounds have shown potential for treatment of EHMT2-overexpressing human cancers[103,106]. Therapeutically relevant substrate competitors were also designed to target SET and MYND domain-containing protein 2 (SMYD2), an oncogenic PMT that, through methylation, functionally represses TP53 and retinoblastoma protein (RB)[79,107,108]. In 2016, an Arg channel-blocking PRMT5 inhibitor, GSK3326595, entered a phase I clinical trial for the treatment of solid tumours and non-Hodgkin lymphoma, which marks a clinical transition for substrate-competitive PMT inhibitors[109,110].

PMT inhibitors that compete for binding with the cofactor SAM have made the most progress towards therapeutic use in humans. The addiction of leukaemic cells that harbour oncogenic genome rearrangements involving myeloid/lymphoid or mixed-lineage leukaemia protein 1 (*MLL1*, also known as *KMT2A*) to the enzymatic activity of DOT1L, the only known human H3K79 PMT, has spurred the development of selective DOT1L inhibitors[111]. Beginning in 2011 with a SAM-competing compound, EPZ004777, highly selective DOT1L inhibitors with picomolar and nanomolar potencies have been generated and are currently being evaluated in cancer clinical trials[106,112,113]. SAM-competing inhibitors found further use in targeting enhancer of zeste homologue 2 (EZH2), the H3K27 methylating component of the Polycomb repressive complex 2 (PRC2) that is frequently amplified in solid tumours and presents with gain-of-function mutations in B cell lymphomas[114–118]. Since the report of the first selective EZH2 inhibitor in 2012, several derivative EZH2-specific compounds have moved forwards to early phase clinical trials[106,119,120]. Recent discoveries of allosteric inhibitors of PRMT3 suggest that allosteric binding sites could also be exploited for the design of potent and selective PMT inhibitors[121,122].

Among the demethylases, LSD1 has been targeted with the greatest success. As demonstrated in 2006, the amine oxidase activity of LSD1 could be inhibited by non-selective monoamine oxidase (MAO) inhibitors, in particular tranylcypromine (TCP), a US Food and Drug Administration (FDA)-approved drug for the treatment of psychological disorders[123]. Development of LSD1-selective TCP derivatives, all of which form a covalent adduct with FAD in the cofactor-binding site, was stimulated by observations that inhibition of LSD1 could promote the differentiation of acute myeloid leukaemia cells[124–128]. In 2014, LSD1 inhibitors, alone and in combination with all-*trans* retinoic acid or histone deacetylase inhibitors, entered clinical trials for the treatment of haematological malignancies[129].

Selective inhibition of JMJC PDMs has proven more challenging owing to the structural similarity of these enzymes. However, the α-KG-competitive selective inhibitors of lysine demethylase 5B (KDM5B; also known as PLU1) and the KDM6 family members KDM6B (also known as JMJD3) and KDM6A (also known as UTX) suggest that in the future the JMJC group of PDMs could be exploited pharmacologically[129–131].

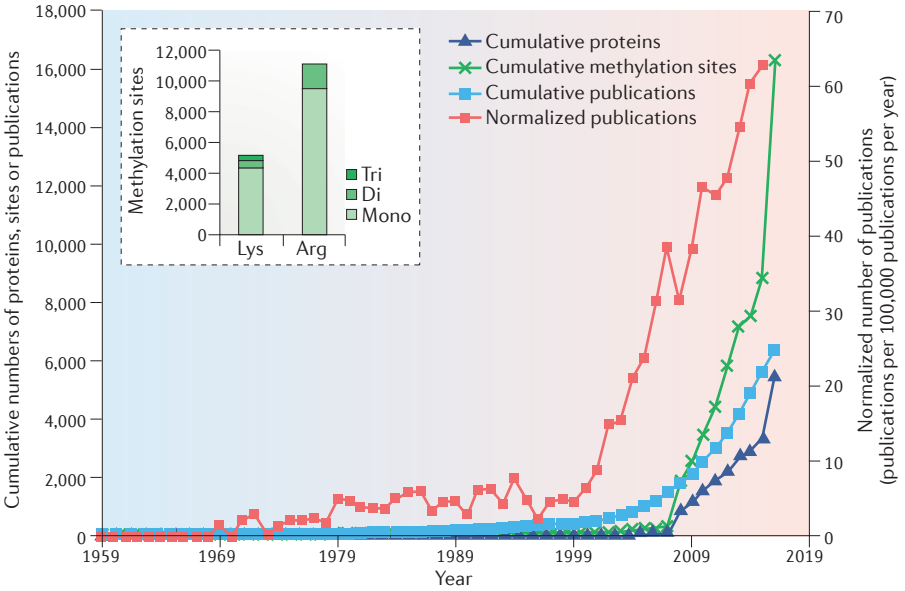


Figure 3 | **Protein methylation in numbers.** The graph summarizes the growing number of publications and known methylated proteins. Cumulative (light blue line) and normalized (red line) numbers of publications about protein methylation were obtained from PubMed. The normalized values show the number of relevant papers per 100,000 papers published per year. Note that the normalized value for the year 2016 is missing. Cumulative numbers of known methylated human proteins (cumulative proteins) or methylated side chains of any residues within these proteins (cumulative methylation sites) were obtained from the PhosphoSitePlus database, and span the period from 1959 until October 2016. Only proteins with determined methylation sites were included in the analysis. The right *y*-axis indicates the normalized values, whereas the left *y*-axis corresponds to cumulative values. Note the increase in publication numbers following the discoveries, between 1998 and 2000, of the first functional roles of protein methylation and the first histone post-translational modifications[34,35,37]. The surge in the number of known methylated sites and proteins shortly after the description of heavy methyl stable isotope labelling with amino acids in cell culture (heavy methyl SILAC) in 2004 (REF. 87) is apparent. See also FIG. 1. The inset shows the number of known methylated Lys and Arg residues categorized by the degree of methylation into mono-, di- and (for Lys only) trimethyl modifications. Recent large-scale, mostly quantitative methylproteomic studies are included in these analyses, although their lists may not be exhaustive.

Crosstalk with metabolism

It is intriguing to note that essentially all PMTs and PDMs discovered to date utilize cofactors that are intermediary metabolites in core metabolic pathways, including SAM, FAD and α-KG (FIG. 2b). This inspired

a tantalizing hypothesis, which proposed that changes in metabolism might induce adaptive responses through changes in the methylation of proteins, especially histones[132]. Recent evidence suggests that this may indeed be the case.

In 2013, the teams of Daley and Cantley found that a reduced production of SAM in embryonic stem cells, induced by restricted metabolism of Thr, substantially decreased the global H3K4me3 level, which led to slowed growth and a loss of pluripotency[133]. In line with these findings, Cravatt and colleagues reported that overabundant nicotinamide *N*-methyltransferase (NNMT) in human cancers caused overconsumption of SAM, which resulted in hypomethylation of H3K4, H3K9, H3K27 and H4K20 (REF. 134) (FIG. 1). Interestingly, histone PMTs with lower affinity for SAM were found to be particularly sensitive to excessive NNMT levels. Finally, in 2015, Locasale and colleagues showed that a crosstalk between SAM metabolism and histone methylation also occurs *in vivo*[135]. The investigators found that the amount of dietary methionine could dictate the level of H3K4me3 methylation by modulating SAM availability. The induced changes in H3K4me3 led to altered gene transcription and provided feedback regulation that decreased the consumption of SAM by alternative pathways.

Circumstantial evidence obtained since 2012 also supports a potential 'sensing' of the metabolic state by PDMs. For example, Thompson and colleagues found that experimental reduction of α-KG levels in embryonic stem cells resulted in hypermethylation of multiple histone Lys residues, presumably by reducing the enzymatic activity of JMJC PDMs[136] (FIG. 2b). These researchers also found further support for such a connection in the cancer-associated mutations in the genes of the cytoplasmic and mitochondrial isocitrate dehydrogenases, *IDH1* and *IDH2*, respectively[137]. The mutant IDH enzymes convert their normal product, α-KG, to 2-hydroxyglutarate (2-HG)[138,139], and the latter was found to inhibit the activity of sensitive PDMs (FIG. 2b), including KDM4C, leading to a block in cell differentiation[137]. In an analogous manner, Nakao and colleagues proposed that LSD1 could sense the metabolic state by modulation of its demethylating activity depending on the availability of FAD (with FAD promoting demethylation)[140] (FIG. 2b). Together, these studies established that fluctuating concentrations of metabolically regulated cofactors of PTMs and PDMs could be sufficiently dynamic to affect the activities of these enzymes, thereby making protein methylation sensitive to metabolic alterations.

The reaction products *S*-adenosylhomocysteine (SAH) and succinate — intermediary metabolites that competitively inhibit PMTs and JMJC PDMs, respectively — also participate in the crosstalk between protein methylation and metabolism[141] (FIG. 2b). In 2012, it was found that excessive accumulation of succinate or fumarate, another metabolic competitor of α-KG, in human cancers broadly inhibited α-KG-dependent PDMs, resulting in potentially oncogenic hypermethylation of histones[142]. It remains to be seen how these and other metabolism-induced changes in protein methylation contribute to metabolic adaptations, pathology and transgenerational inheritance.

Future prospects

One important task for the future will be to define the proteome-wide substrate specificity of individual PMTs and PDMs. Most of these enzymes are known chiefly for their histone methyl-modifying activities; however, a growing number of reports, including a recent PMT-specific proteome-wide study[143], suggest the existence of a myriad of additional, non-histone substrates. The definition of the complete targeting repertoires of individual PMTs and PDMs will be especially crucial if these enzymes are to serve as drug targets for therapy in humans.

Furthermore, it is unclear whether the dominance of Arg- versus Lys-methylated sites and the low overall proportion of residues with a higher degree of methylation (di- and trimethyl) (FIG. 3) reflect the true methylation state *in vivo* or an experimental bias. The latter could, at least in part, result from differences in the quality of the antibodies that are used for the enrichment of methylated species before mass spectrometry analysis, or a less efficient enzymatic digestion at residues with a higher degree of methylation[144]. Methyl-specific antibodies with improved affinity might, for example, be raised against a library of degenerate peptides that contain a methylated residue flanked by 1–3 more or less random amino acids. A similar strategy using a library of much longer peptides recently yielded encouraging results, although the length of the flanking sequences may have reduced the methyl-specificity[90]. The latest advances in mass spectrometry-based approaches that enable ultra-deep analyses of proteomes inspire hope that it might be possible to eliminate the need for enrichment of methylated peptides altogether (J. Olsen, personal communication).

Beyond the identification of methylation events, it will be critical to investigate their regulatory roles. The challenge should be tackled from a systems perspective, using quantitative proteomics to study dynamic changes (for example, time- or stimulus-dependent changes) in methylproteomes. This will allow for methylation networks to be inferred, which might exist to connect different proteins into functionally coherent modules. In addition, knowledge of methylation stoichiometry, that is, the proportion of molecules methylated at a particular site, will inform about the dynamic range (and thus the extent) of regulation by methylation. The few known 'static' values indicate substantial differences between different methylation sites. For example, up to 80% of all available H4K20 sites are dimethylated, and H3K9me1 is at around 30%, whereas the level of H3K4me3 is below 0.01%. Similar methylation stoichiometry ranges may be in place for non-histone proteins (REF. 145; and O. Gozani, personal communication).

At present, little is known about the readout of protein methylation beyond histone methylation, mainly because few methyl sites on non-histone proteins have had their cognate readers determined. As Mann and colleagues demonstrated, quantitative mass spectrometry can be a powerful asset in the designation of readers for specific methyl sites, and it would be of interest if mass spectrometry-based approaches with a higher throughput could be developed[51,146].

Curiously, a substantial number of methylated sites appear to be buried deep within nucleic acid-binding proteins where they are unlikely to be accessible to methyl reader proteins[76]. In RBPs, many of these buried sites are located in the immediate vicinity of the bound RNA, which suggests that protein methylation could directly affect protein–RNA interactions[24,76,147] (FIG. 2a, bottom left panel). A recent large-scale analysis revealed that RNA-binding sites are hotspots for protein methylation[148]. Strikingly, Lys methylation and, to a lesser extent, Arg methylation were found to be more than 16 times more common at the RNA-binding sites compared with non-RNA-binding regions within RBPs[148]. Future studies are warranted to explore the

significance of these methylation events in more detail.

Finally, several fundamental questions about protein methylation remain open. For example, how are PMTs and PDMs specifically recruited to their target residues? Are these enzymes constitutively active or is their activity regulated in the cell, and if so, how? How is the slow kinetics of histone and, potentially, non-histone protein methylation coupled to dynamic cellular responses? Is the observation from 1973 that mitochondria contain the highest histone Lys 'demethylating' activity — particularly in the kidney, of all organs[20] — supported by modern biochemistry? If so, what is the explanation for this enrichment? What is the fate of the highly reactive formaldehyde that is released during enzymatic demethylation? Are there PDMs dedicated to the removal of methyl groups from arginines, H3K79 and buried methylated residues? Does histone methylation directly affect transcription?

It is clear that we still have much to learn about the molecular details of protein methylation. As defects in protein methylation networks have been linked to human pathologies, an in-depth understanding of the biology of this PTM will be essential to harness the full potential of protein methylation for the development of therapies to treat disease in humans.

Conclusion

The evolution of the field of protein methylation reflects its heavy and continuous reliance on technological breakthroughs. Most milestone discoveries in the field were based on one or more of the three key inventions, starting with the methods for protein sequence determination in the 1950s, continuing with the renaissance in genetics and molecular biology approaches in the 1990s, and ending with the advent of modern proteomic technologies about a decade later (FIG. 1). These quantum leaps have transformed our view of protein methylation from it being an obscure epiphenomenon into what we now consider a powerful regulator of essentially every cellular process. Technological advances are sure to continue driving the evolution of this exciting field in the future.

Jernej Murn and Yang Shi are at the Department of Cell Biology, Harvard Medical School, and the Division of Newborn Medicine, Boston Children's Hospital, 300 Longwood Avenue, Boston, Massachusetts 02115, USA.

murn.jernej@gmail.com;
yang_shi@hms.harvard.edu

doi:10.1038/nrm.2017.35
Published online 17 May 2017

1. Ambler, R. P. & Rees, M. W. Epsilon-N-methyl-lysine in bacterial flagellar protein. *Nature* **184**, 56–57 (1959).
2. Neuberger, A. & Sanger, F. The availability of in-acetyl-d-lysine and in-methyl-dl-lysine for growth. *Biochem. J.* **38**, 125–129 (1944).
3. Neuberger, A. & Sanger, F. The metabolism of lysine. *Biochem. J.* **38**, 119–125 (1944).
4. Stocker, B. & McDonough, M. A gene determining presence or absence of ε-N-methyl-lysine in *Salmonella* flagellar protein. *Nature* **189**, 556–558 (1961).
5. Murray, K. The occurrence of epsilon-N-methyl lysine in histones. *Biochemistry* **3**, 10–15 (1964).
6. Allfrey, V. G., Faulkner, R. & Mirsky, A. E. Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *Proc. Natl Acad. Sci. USA* **51**, 786–794 (1964).
7. Kim, S. & Paik, W. K. Studies on the origin of epsilon-N-methyl-L-lysine in protein. *J. Biol. Chem.* **240**, 4629–4634 (1965).
8. Fischer, E. H., Graves, D. J., Crittenden, E. R. & Krebs, E. G. Structure of the site phosphorylated in the phosphorylase b to a reaction. *J. Biol. Chem.* **234**, 1698–1704 (1959).
9. Phillips, D. M. The presence of acetyl groups of histones. *Biochem. J.* **87**, 258–263 (1963).
10. Walsh, D. A., Perkins, J. P. & Krebs, E. G. An adenosine 3′,5′-monophosphate-dependant protein kinase from rabbit skeletal muscle. *J. Biol. Chem.* **243**, 3763–3765 (1968).
11. Burnett, G. & Kennedy, E. P. The enzymatic phosphorylation of proteins. *J. Biol. Chem.* **211**, 969–980 (1954).
12. Paik, W. K. & Kim, S. Protein methylase I. Purification and properties of the enzyme. *J. Biol. Chem.* **243**, 2108–2114 (1968).
13. Paik, W. K. & Kim, S. Solubilization and partial purification of protein methylase 3 from calf thymus nuclei. *J. Biol. Chem.* **245**, 6010–6015 (1970).
14. Liss, M. & Edelstein, L. M. Evidence for the enzymatic methylation of crystalline ovalbumin preparations. *Biochem. Biophys. Res. Commun.* **26**, 497–504 (1967).
15. Hempel, K. & Lange, H. W. Nε-methylated lysine in histones from chicken erythrocytes. *Hoppe-Seyler's Z. Physiol. Chem.* **349**, 603–607 (in German) (1968).
16. Baldwin, G. S. & Carnegie, P. R. Specific enzymic methylation of an arginine in the experimental allergic encephalomyelitis protein from human myelin. *Science* **171**, 579–581 (1971).
17. Brostoff, S. & Eylar, E. H. Localization of methylated arginine in the A1 protein from myelin. *Proc. Natl Acad. Sci. USA* **68**, 765–769 (1971).
18. Kakimoto, Y. & Akazawa, S. Isolation and identification of N^G,N^G- and N^G,N'^G-dimethyl-arginine, N^ε-mono-, di-, and trimethyllysine, and glucosylgalactosyl- and galactosyl-δ-hydroxylysine from human urine. *J. Biol. Chem.* **245**, 5751–5758 (1970).
19. Kim, S., Benoiton, L. & Paik, W. K. Epsilon-alkyllysinase, purification and properties of the enzyme. *J. Biol. Chem.* **239**, 3790–3796 (1964).
20. Paik, W. K. & Kim, S. Enzymatic demethylation of calf thymus histones. *Biochem. Biophys. Res. Commun.* **51**, 781–788 (1973).
21. Duerre, J. A. & Lee, C. T. *In vivo* methylation and turnover of rat brain histones. *J. Neurochem.* **23**, 541–547 (1974).
22. Byvoet, P., Shepherd, G. R., Hardin, J. M. & Noland, B. J. The distribution and turnover of labeled methyl groups in histone fractions of cultured mammalian cells. *Arch. Biochem. Biophys.* **148**, 558–567 (1972).
23. Stallcup, M. R. Role of protein methylation in chromatin remodeling and transcriptional regulation. *Oncogene* **20**, 3014–3020 (2001).
24. Calnan, B. J., Tidor, B., Biancalana, S., Hudson, D. & Frankel, A. D. Arginine-mediated RNA recognition: the arginine fork. *Science* **252**, 1167–1171 (1991).
25. Najbauer, J., Johnson, B. A., Young, A. L. & Aswad, D. W. Peptides with sequences similar to glycine, arginine-rich motifs in proteins interacting with RNA are efficiently recognized by methyltransferase(s) modifying arginine in numerous proteins. *J. Biol. Chem.* **268**, 10501–10509 (1993).
26. Clarke, S. Protein methylation. *Curr. Opin. Cell Biol.* **5**, 977–983 (1993).
27. Orlando, V. Mapping chromosomal proteins *in vivo* by formaldehyde-crosslinked-chromatin immunoprecipitation. *Trends Biochem. Sci.* **25**, 99–104 (2000).
28. Solomon, M. J., Larsen, P. L. & Varshavsky, A. Mapping protein-DNA interactions *in vivo* with formaldehyde: evidence that histone H4 is retained on a highly transcribed gene. *Cell* **53**, 937–947 (1988).
29. Brownell, J. E. *et al.* Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* **84**, 843–851 (1996).
30. Taunton, J., Hassig, C. A. & Schreiber, S. L. A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. *Science* **272**, 408–411 (1996).
31. Klein, R. R. & Houtz, R. L. Cloning and developmental expression of pea ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit N-methyltransferase. *Plant Mol. Biol.* **27**, 249–261 (1995).
32. Henry, M. F. & Silver, P. A. A novel methyltransferase (Hmt1p) modifies poly(A) + -RNA-binding proteins. *Mol. Cell. Biol.* **16**, 3668–3678 (1996).
33. Lin, W. J., Gary, J. D., Yang, M. C., Clarke, S. & Herschman, H. R. The mammalian immediate-early TIS21 protein and the leukemia-associated BTG1 protein interact with a protein-arginine N-methyltransferase. *J. Biol. Chem.* **271**, 15034–15044 (1996).
34. Shen, E. C. *et al.* Arginine methylation facilitates the nuclear export of hnRNP proteins. *Genes Dev.* **12**, 679–691 (1998).
35. Chen, D. *et al.* Regulation of transcription by a protein methyltransferase. *Science* **284**, 2174–2177 (1999).
36. Strahl, B. D., Ohba, R., Cook, R. G. & Allis, C. D. Methylation of histone H3 at lysine 4 is highly conserved and correlates with transcriptionally active nuclei in *Tetrahymena*. *Proc. Natl Acad. Sci. USA* **96**, 14967–14972 (1999).
37. Rea, S. *et al.* Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* **406**, 593–599 (2000).
38. Herz, H. M., Garruss, A. & Shilatifard, A. SET for life: biochemical activities and biological functions of SET domain-containing proteins. *Trends Biochem. Sci.* **38**, 621–639 (2013).
39. Zhang, X., Zhou, L. & Cheng, X. Crystal structure of the conserved core of protein arginine methyltransferase PRMT3. *EMBO J.* **19**, 3509–3519 (2000).
40. Weiss, V. H. *et al.* The structure and oligomerization of the yeast arginine methyltransferase, *Hmt1*. *Nat. Struct. Biol.* **7**, 1165–1171 (2000).
41. Greer, E. L. & Shi, Y. Histone methylation: a dynamic mark in health, disease and inheritance. *Nat. Rev. Genet.* **13**, 343–357 (2012).
42. Strahl, B. D. & Allis, C. D. The language of covalent histone modifications. *Nature* **403**, 41–45 (2000).
43. Lee, J. S., Smith, E. & Shilatifard, A. The language of histone crosstalk. *Cell* **142**, 682–685 (2010).
44. Lachner, M., O'Carroll, D., Rea, S., Mechtler, K. & Jenuwein, T. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410**, 116–120 (2001).
45. Bannister, A. J. *et al.* Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410**, 120–124 (2001).
46. Nielsen, P. R. *et al.* Structure of the HP1 chromodomain bound to histone H3 methylated at lysine 9. *Nature* **416**, 103–107 (2002).
47. Jacobs, S. A. & Khorasanizadeh, S. Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. *Science* **295**, 2080–2083 (2002).
48. Taverna, S. D., Li, H., Ruthenburg, A. J., Allis, C. D. & Patel, D. J. How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat. Struct. Mol. Biol.* **14**, 1025–1040 (2007).
49. Musselman, C. A., Lalonde, M. E., Cote, J. & Kutateladze, T. G. Perceiving the epigenetic landscape through histone readers. *Nat. Struct. Mol. Biol.* **19**, 1218–1227 (2012).
50. Patel, D. J. & Wang, Z. Readout of epigenetic modifications. *Annu. Rev. Biochem.* **82**, 81–118 (2013).
51. Vermeulen, M. *et al.* Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell* **131**, 58–69 (2007).
52. Nielsen, S. J. *et al.* Rb targets histone H3 methylation and HP1 to promoters. *Nature* **412**, 561–565 (2001).
53. Bannister, A. J., Schneider, R. & Kouzarides, T. Histone methylation: dynamic or static? *Cell* **109**, 801–806 (2002).

54. Ahmad, K. & Henikoff, S. The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Mol. Cell* **9**, 1191–1200 (2002).
55. Allis, C. D., Bowen, J. K., Abraham, G. N., Glover, C. V. & Gorovsky, M. A. Proteolytic processing of histone H3 in chromatin: a physiologically regulated event in *Tetrahymena* micronuclei. *Cell* **20**, 55–64 (1980).
56. Humphrey, G. W. *et al.* Stable histone deacetylase complexes distinguished by the presence of SANT domain proteins CoREST/kiaa0071 and Mta-L1. *J. Biol. Chem.* **276**, 6817–6824 (2001).
57. Shi, Y. *et al.* Coordinated histone modifications mediated by a CtBP co-repressor complex. *Nature* **422**, 735–738 (2003).
58. Tong, J. K., Hassig, C. A., Schnitzler, G. R., Kingston, R. E. & Schreiber, S. L. Chromatin deacetylation by an ATP-dependent nucleosome remodelling complex. *Nature* **395**, 917–921 (1998).
59. You, A., Tong, J. K., Grozinger, C. M. & Schreiber, S. L. CoREST is an integral component of the CoREST-human histone deacetylase complex. *Proc. Natl Acad. Sci. USA* **98**, 1454–1458 (2001).
60. Hakimi, M. A. *et al.* A core-BRAF35 complex containing histone deacetylase mediates repression of neuronal-specific genes. *Proc. Natl Acad. Sci. USA* **99**, 7420–7425 (2002).
61. Shi, Y. *et al.* Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **119**, 941–953 (2004).
62. Metzger, E. *et al.* LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. *Nature* **437**, 436–439 (2005).
63. Tsukada, Y. *et al.* Histone demethylation by a family of JmjC domain-containing proteins. *Nature* **439**, 811–816 (2006).
64. Whetstine, J. R. *et al.* Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. *Cell* **125**, 467–481 (2006).
65. Fodor, B. D. *et al.* Jmjd2b antagonizes H3K9 trimethylation at pericentric heterochromatin in mammalian cells. *Genes Dev.* **20**, 1557–1562 (2006).
66. Cloos, P. A. *et al.* The putative oncogene GASC1 demethylates tri- and dimethylated lysine 9 on histone H3. *Nature* **442**, 307–311 (2006).
67. Kooistra, S. M. & Helin, K. Molecular mechanisms and potential functions of histone demethylases. *Nat. Rev. Mol. Cell Biol.* **13**, 297–311 (2012).
68. Brahms, H., Meheus, L., de Brabandere, V., Fischer, U. & Luhrmann, R. Symmetrical dimethylation of arginine residues in spliceosomal Sm protein B/B' and the Sm-like protein LSm4, and their interaction with the SMN protein. *RNA* **7**, 1531–1542 (2001).
69. Friesen, W. J., Massenet, S., Paushkin, S., Wyce, A. & Dreyfuss, G. SMN, the product of the spinal muscular atrophy gene, binds preferentially to dimethylarginine-containing protein targets. *Mol. Cell* **7**, 1111–1117 (2001).
70. Yu, M. C. *et al.* Arginine methyltransferase affects interactions and recruitment of mRNA processing and export factors. *Genes Dev.* **18**, 2024–2035 (2004).
71. Li, H. *et al.* Lipopolysaccharide-induced methylation of HuR, an mRNA-stabilizing protein, by CARM1. Coactivator-associated arginine methyltransferase. *J. Biol. Chem.* **277**, 44623–44630 (2002).
72. Nishida, K. M. *et al.* Functional involvement of Tudor and dPRMT5 in the piRNA processing pathway in *Drosophila* germlines. *EMBO J.* **28**, 3820–3831 (2009).
73. Reuter, M. *et al.* Loss of the Mili-interacting Tudor domain-containing protein-1 activates transposons and alters the Mili-associated small RNA profile. *Nat. Struct. Mol. Biol.* **16**, 639–646 (2009).
74. Vagin, V. V. *et al.* Proteomic analysis of murine Piwi proteins reveals a role for arginine methylation in specifying interaction with Tudor family members. *Genes Dev.* **23**, 1749–1762 (2009).
75. Chen, C., Nott, T. J., Jin, J. & Pawson, T. Deciphering arginine methylation: Tudor tells the tale. *Nat. Rev. Mol. Cell Biol.* **12**, 629–642 (2011).
76. Clarke, S. G. Protein methylation at the surface and buried deep: thinking outside the histone box. *Trends Biochem. Sci.* **38**, 243–252 (2013).
77. Pahlich, S., Zakaryan, R. P. & Gehring, H. Protein arginine methylation: cellular functions and methods of analysis. *Biochim. Biophys. Acta* **1764**, 1890–1903 (2006).
78. Chuikov, S. *et al.* Regulation of p53 activity through lysine methylation. *Nature* **432**, 353–360 (2004).
79. Huang, J. *et al.* Repression of p53 activity by Smyd2-mediated methylation. *Nature* **444**, 629–632 (2006).
80. Shi, X. *et al.* Modulation of p53 function by SET8-mediated methylation at lysine 382. *Mol. Cell* **27**, 636–646 (2007).
81. Huang, J. *et al.* G9a and Glp methylate lysine 373 in the tumor suppressor p53. *J. Biol. Chem.* **285**, 9636–9641 (2010).
82. Jansson, M. *et al.* Arginine methylation regulates the p53 response. *Nat. Cell Biol.* **10**, 1431–1439 (2008).
83. Metzger, E. *et al.* LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. *Nature* **437**, 436–439 (2005).
84. Huszar, G. & Elzinga, M. Epsilon-N-methyl lysine in myosin. *Nature* **223**, 834–835 (1969).
85. DeLange, R. J., Glazer, A. N. & Smith, E. L. Presence and location of an unusual amino acid, epsilon-N-trimethyllysine, in cytochrome *c* of wheat germ and *Neurospora*. *J. Biol. Chem.* **244**, 1385–1388 (1969).
86. Pollack, B. P. *et al.* The human homologue of the yeast proteins Skb1 and Hsl7p interacts with Jak kinases and contains protein methyltransferase activity. *J. Biol. Chem.* **274**, 31531–31542 (1999).
87. Ong, S. E., Mittler, G. & Mann, M. Identifying and quantifying *in vivo* methylation sites by heavy methyl SILAC. *Nat. Methods* **1**, 119–126 (2004).
88. Levy, D. *et al.* A proteomic approach for the identification of novel lysine methyltransferase substrates. *Epigenetics Chromatin* **4**, 19 (2011).
89. Cao, X. J., Arnaudo, A. M. & Garcia, B. A. Large-scale global identification of protein lysine methylation *in vivo*. *Epigenetics* **8**, 477–485 (2013).
90. Guo, A. *et al.* Immunoaffinity enrichment and mass spectrometry analysis of protein methylation. *Mol. Cell. Proteomics* **13**, 372–387 (2014).
91. Liu, H. *et al.* A method for systematic mapping of protein lysine methylation identifies functions for HP1beta in DNA damage response. *Mol. Cell* **50**, 723–735 (2013).
92. Moore, K. E. *et al.* A general molecular affinity strategy for global detection and proteomic analysis of lysine methylation. *Mol. Cell* **50**, 444–456 (2013).
93. Islam, K. *et al.* Bioorthogonal profiling of protein methylation using azido derivative of S-adenosyl-L-methionine. *J. Am. Chem. Soc.* **134**, 5909–5915 (2012).
94. Islam, K. *et al.* Defining efficient enzyme-cofactor pairs for bioorthogonal profiling of protein methylation. *Proc. Natl Acad. Sci. USA* **110**, 16778–16783 (2013).
95. Binda, O. *et al.* A chemical method for labeling lysine methyltransferase substrates. *Chembiochem* **12**, 330–334 (2011).
96. Peters, W. *et al.* Enzymatic site-specific functionalization of protein methyltransferase substrates with alkynes for click labeling. *Angew. Chem. Int. Ed.* **49**, 5170–5173 (2010).
97. Biggar, K. K. & Li, S. S. Non-histone protein methylation as a regulator of cellular signalling and function. *Nat. Rev. Mol. Cell Biol.* **16**, 5–17 (2015).
98. Andreu-Perez, P. *et al.* Protein arginine methyltransferase 5 regulates ERK1/2 signal transduction amplitude and cell fate through CRAF. *Sci. Signal.* **4**, ra58 (2011).
99. Mazur, P. K. *et al.* SMYD3 links lysine methylation of MAP3K2 to Ras-driven cancer. *Nature* **510**, 283–287 (2014).
100. Bikkavilli, R. K. *et al.* Dishevelled3 is a novel arginine methyl transferase substrate. *Sci. Rep.* **2**, 805 (2012).
101. Kim, E. *et al.* Phosphorylation of EZH2 activates STAT3 signaling via STAT3 methylation and promotes tumorigenicity of glioblastoma stem-like cells. *Cancer Cell* **23**, 839–852 (2013).
102. Levy, D. *et al.* Lysine methylation of the NF-kappaB subunit RelA by SETD6 couples activity of the histone methyltransferase GLP at chromatin to tonic repression of NF-kappaB signaling. *Nat. Immunol.* **12**, 29–36 (2011).
103. Kubicek, S. *et al.* Reversal of H3K9me2 by a small-molecule inhibitor for the G9a histone methyltransferase. *Mol. Cell* **25**, 473–481 (2007).
104. Tachibana, M. *et al.* Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9. *Genes Dev.* **19**, 815–826 (2005).
105. Rathert, P. *et al.* Protein lysine methyltransferase G9a acts on non-histone targets. *Nat. Chem. Biol.* **4**, 344–346 (2008).
106. Kaniskan, H. U., Konze, K. D. & Jin, J. Selective inhibitors of protein methyltransferases. *J. Med. Chem.* **58**, 1596–1629 (2015).
107. Ferguson, A. D. *et al.* Structural basis of substrate methylation and inhibition of SMYD2. *Structure* **19**, 1262–1273 (2011).
108. Saddic, L. A. *et al.* Methylation of the retinoblastoma tumor suppressor by SMYD2. *J. Biol. Chem.* **285**, 37733–37740 (2010).
109. Castillo-Aguilera, O., Depreux, P., Halby, L., Arimondo, P. B. & Goossens, L. DNA methylation targeting: the DNMT/HMT crosstalk challenge. *Biomolecules* **7**, E3 (2017).
110. Chan-Penebre, E. *et al.* A selective inhibitor of PRMT5 with *in vivo* and *in vitro* potency in MCL models. *Nat. Chem. Biol.* **11**, 432–437 (2015).
111. Okada, Y. *et al.* hDOT1L links histone methylation to leukemogenesis. *Cell* **121**, 167–178 (2005).
112. Daigle, S. R. *et al.* Selective killing of mixed lineage leukemia cells by a potent small-molecule DOT1L inhibitor. *Cancer Cell* **20**, 53–65 (2011).
113. Daigle, S. R. *et al.* Potent inhibition of DOT1L as treatment of MLL-fusion leukemia. *Blood* **122**, 1017–1025 (2013).
114. Cao, R. *et al.* Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* **298**, 1039–1043 (2002).
115. Morin, R. D. *et al.* Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. *Nat. Genet.* **42**, 181–185 (2010).
116. Kleer, C. G. *et al.* EZH2 is a marker of aggressive breast cancer and promotes neoplastic transformation of breast epithelial cells. *Proc. Natl Acad. Sci. USA* **100**, 11606–11611 (2003).
117. Varambally, S. *et al.* The polycomb group protein EZH2 is involved in progression of prostate cancer. *Nature* **419**, 624–629 (2002).
118. Sneeringer, C. J. *et al.* Coordinated activities of wild-type plus mutant EZH2 drive tumor-associated hypertrimethylation of lysine 27 on histone H3 (H3K27) in human B-cell lymphomas. *Proc. Natl Acad. Sci. USA* **107**, 20980–20985 (2010).
119. Knutson, S. K. *et al.* A selective inhibitor of EZH2 blocks H3K27 methylation and kills mutant lymphoma cells. *Nat. Chem. Biol.* **8**, 890–896 (2012).
120. Knutson, S. K. *et al.* Selective inhibition of EZH2 by EPZ-6438 leads to potent antitumor activity in EZH2-mutant non-Hodgkin lymphoma. *Mol. Cancer Ther.* **13**, 842–854 (2014).
121. Liu, F. *et al.* Exploiting an allosteric binding site of PRMT3 yields potent and selective inhibitors. *J. Med. Chem.* **56**, 2110–2124 (2013).
122. Kaniskan, H. U. *et al.* A potent, selective and cell-active allosteric inhibitor of protein arginine methyltransferase 3 (PRMT3). *Angew. Chem. Int. Ed.* **54**, 5166–5170 (2015).
123. Lee, M. G., Wynder, C., Schmidt, D. M., McCafferty, D. G. & Shiekhattar, R. Histone H3 lysine 4 demethylation is a target of nonselective antidepressive medications. *Chem. Biol.* **13**, 563–567 (2006).
124. Schmidt, D. M. & McCafferty, D. G. *trans*-2-phenylcyclopropylamine is a mechanism-based inactivator of the histone demethylase LSD1. *Biochemistry* **46**, 4408–4416 (2007).
125. Binda, C. *et al.* Biochemical, structural, and biological evaluation of tranylcypromine derivatives as inhibitors of histone demethylases LSD1 and LSD2. *J. Am. Chem. Soc.* **132**, 6827–6833 (2010).
126. Yang, M. *et al.* Structural basis for the inhibition of the LSD1 histone demethylase by the antidepressant *trans*-2-phenylcyclopropylamine. *Biochemistry* **46**, 8058–8065 (2007).
127. Schenk, T. *et al.* Inhibition of the LSD1 (KDM1A) demethylase reactivates the all-*trans*-retinoic acid differentiation pathway in acute myeloid leukemia. *Nat. Med.* **18**, 605–611 (2012).
128. Harris, W. J. *et al.* The histone demethylase KDM1A sustains the oncogenic potential of MLL-AF9 leukemia stem cells. *Cancer Cell* **21**, 473–487 (2012).
129. Morera, L., Lubbert, M. & Jung, M. Targeting histone methyltransferases and demethylases in clinical trials for cancer therapy. *Clin. Epigenetics* **8**, 57 (2016).
130. Kruidenier, L. *et al.* A selective jumonji H3K27 demethylase inhibitor modulates the proinflammatory macrophage response. *Nature* **488**, 404–408 (2012).
131. Hancock, R. L., Dunne, K., Walport, L. J., Flashman, E. & Kawamura, A. Epigenetic regulation by histone demethylases in hypoxia. *Epigenomics* **7**, 791–811 (2015).

132. Teperino, R., Schoonjans, K. & Auwerx, J. Histone methyl transferases and demethylases; can they link metabolism and transcription? *Cell Metab.* **12**, 321–327 (2010).
133. Shyh-Chang, N. *et al.* Influence of threonine metabolism on S-adenosylmethionine and histone methylation. *Science* **339**, 222–226 (2013).
134. Ulanovskaya, O. A., Zuhl, A. M. & Cravatt, B. F. NNMT promotes epigenetic remodeling in cancer by creating a metabolic methylation sink. *Nat. Chem. Biol.* **9**, 300–306 (2013).
135. Mentch, S. J. *et al.* Histone methylation dynamics and gene regulation occur through the sensing of one-carbon metabolism. *Cell Metab.* **22**, 861–873 (2015).
136. Carey, B. W., Finley, L. W., Cross, J. R., Allis, C. D. & Thompson, C. B. Intracellular alpha-ketoglutarate maintains the pluripotency of embryonic stem cells. *Nature* **518**, 413–416 (2015).
137. Lu, C. *et al.* IDH mutation impairs histone demethylation and results in a block to cell differentiation. *Nature* **483**, 474–478 (2012).
138. Dang, L. *et al.* Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature* **462**, 739–744 (2009).
139. Ward, P. S. *et al.* The common feature of leukemia-associated IDH1 and IDH2 mutations is a neomorphic enzyme activity converting alpha-ketoglutarate to 2-hydroxyglutarate. *Cancer Cell* **17**, 225–234 (2010).
140. Hino, S. *et al.* FAD-dependent lysine-specific demethylase-1 regulates cellular energy expenditure. *Nat. Commun.* **3**, 758 (2012).
141. Luka, Z., Mudd, S. H. & Wagner, C. Glycine N-methyltransferase and regulation of S-adenosylmethionine levels. *J. Biol. Chem.* **284**, 22507–22511 (2009).
142. Xiao, M. *et al.* Inhibition of alpha-KG-dependent histone and DNA demethylases by fumarate and succinate that are accumulated in mutations of FH and SDH tumor suppressors. *Genes Dev.* **26**, 1326–1338 (2012).
143. Olsen, J. B. *et al.* Quantitative profiling of the activity of protein lysine methyltransferase SMYD2 using SILAC-based proteomics. *Mol. Cell. Proteomics* **15**, 892–905 (2016).
144. Zee, B. M. & Garcia, B. A. Discovery of lysine post-translational modifications through mass spectrometric detection. *Essays Biochem.* **52**, 147–163 (2012).
145. Ostareck-Lederer, A. *et al.* Asymmetric arginine dimethylation of heterogeneous nuclear ribonucleoprotein K by protein-arginine methyltransferase 1 inhibits its interaction with c-Src. *J. Biol. Chem.* **281**, 11115–11125 (2006).
146. Vermeulen, M. *et al.* Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell* **142**, 967–980 (2010).
147. Hu, S. B. *et al.* Protein arginine methyltransferase CARM1 attenuates the paraspeckle-mediated nuclear retention of mRNAs containing IRAlus. *Genes Dev.* **29**, 630–645 (2015).
148. Castello, A. *et al.* Comprehensive identification of RNA-binding domains in human cells. *Mol. Cell* **63**, 696–710 (2016).
149. Beaver, J. E. & Waters, M. L. Molecular recognition of Lys and Arg methylation. *ACS Chem. Biol.* **11**, 643–653 (2016).
150. Barth, T. K. & Imhof, A. Fast signals and slow marks: the dynamics of histone modifications. *Trends Biochem. Sci.* **35**, 618–626 (2010).
151. Zee, B. M., Levin, R. S., Dimaggio, P. A. & Garcia, B. A. Global turnover of histone post-translational modifications and variants in human cells. *Epigenetics Chromatin* **3**, 22 (2010).
152. Jacobs, S. A. *et al.* The active site of the SET domain is constructed on a knot. *Nat. Struct. Biol.* **9**, 833–838 (2002).
153. Min, J., Zhang, X., Cheng, X., Grewal, S. I. & Xu, R. M. Structure of the SET domain histone lysine methyltransferase Clr4. *Nat. Struct. Biol.* **9**, 828–832 (2002).
154. Trievel, R. C., Beach, B. M., Dirk, L. M., Houtz, R. L. & Hurley, J. H. Structure and catalytic mechanism of a SET domain protein methyltransferase. *Cell* **111**, 91–103 (2002).
155. Wilson, J. R. *et al.* Crystal structure and functional analysis of the histone methyltransferase SET7/9. *Cell* **111**, 105–115 (2002).
156. Zhang, X. *et al.* Structure of the *Neurospora* SET domain protein DIM-5, a histone H3 lysine methyltransferase. *Cell* **111**, 117–127 (2002).
157. Lu, X. *et al.* The effect of H3K79 dimethylation and H4K20 trimethylation on nucleosome and chromatin structure. *Nat. Struct. Mol. Biol.* **15**, 1122–1124 (2008).
158. Pasini, D. *et al.* Characterization of an antagonistic switch between histone H3 lysine 27 methylation and acetylation in the transcriptional regulation of Polycomb group target genes. *Nucleic Acids Res.* **38**, 4958–4969 (2010).
159. Tie, F. *et al.* CBP-mediated acetylation of histone H3 lysine 27 antagonizes *Drosophila* Polycomb silencing. *Development* **136**, 3131–3141 (2009).

Acknowledgements

The authors thank L. Sawyer for help with retrieving information and materials relating to Richard Ambler's research; P. Hornbeck for help with the analyses of known methylated proteins; O. Gozani, J. Olsen and B. Fischer for discussions; and M. Teplova for help with the figures. Work in the laboratory of Y.S. is supported by the US National Institutes of Health (CA118487, GM117264 and MH096066) and Boston Children's Hospital. Y.S. is an American Cancer Society Research Professor. The authors apologize to colleagues whose work could not be cited owing to space limitations.

Competing interests statement

The authors declare competing interests: see Web version for details.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

DATABASES

Phosphositeplus: http://www.phosphosite.org
Pubmed: https://www.ncbi.nlm.nih.gov/pubmed/
RCSB Protein Data Bank: http://www.rcsb.org/pdb/home/home.do

ALL LINKS ARE ACTIVE IN THE ONLINE PDF