The ribosome: A hot spot for the identification of new types of protein methyltransferases

Cellular physiology depends on the alteration of protein structures by covalent modification reactions. Using a combination of bioinformatic, genetic, biochemical, and mass spectrometric approaches, it has been possible to probe ribosomal proteins from the yeast *Saccharomyces cerevisiae* for posttranslationally methylated amino acid residues and for the enzymes that catalyze these modifications. These efforts have resulted in the identification and characterization of the first protein histidine methyltransferase, the first N-terminal protein methyltransferase, two unusual types of protein arginine methyltransferases, and a new type of cysteine methylation. Two of these enzymes may modify their substrates during ribosomal assembly because the final methylated histidine and arginine residues are buried deep within the ribosome with contacts only with RNA. Two of these modifications occur broadly in eukaryotes, including humans, whereas the others demonstrate a more limited phylogenetic range. Analysis of strains where the methyltransferase genes are deleted has given insight into the physiological roles of these modifications. These reactions described here add diversity to the modifications that generate the typical methylated lysine and arginine residues previously described in histones and other proteins.

**Regulation of biological function by protein methylation reactions**

It is more and more apparent just how much of the biological function depends upon post-translational modifications (1). The human genome encodes some 900 enzymes catalyzing just protein phosphorylation or ubiquitination (2, 3). However, it is now clear that methyl groups can stand beside phosphate groups and ubiquitin as major players in controlling the physiological functions of proteins. We are beginning to understand how the much greater diversity of protein methylation reactions can give rise to a greater diversity of function (1, 4–8). We are also learning the importance of cross-talk between protein and DNA methylation reactions (9) and among protein methylation, phosphorylation, acetylation, and ubiquitination reactions (10–12). Finally, we are discovering how alterations in protein methylation pathways can lead to human pathology, particularly in cancer (13–15).

The collection of over 60 human protein arginine and lysine methyltransferases that leave histone “marks” recognized by reader proteins to guide gene expression are perhaps the poster children for the importance of protein methyltransferases (16, 17). However, recent work has emphasized the importance of methylating nonhistone substrates at these residues, particularly ribosomal proteins, translation factors, and transcription factors in a wide variety of organisms (7, 8, 15, 18, 19). Furthermore, the methylation of lysine and arginine residues represents just the tip of the iceberg in protein methylation, and modifications have also been established at histidine, modified histidine (diphthamide), glutamate, glutamine, asparagine, L-isoaspartate, D-aspartate, cysteine, isoprenylcysteine, methionine, and N- and C-terminal residues (1, 4).

Recent work has demonstrated how methylated residues can be recognized by protein interaction domains for transcriptional control (13, 20), can be targets for ubiquitin-E3–linked degradation (21), can facilitate or block protein–protein interactions or enzyme–substrate interactions (6, 22), and can modulate interactions with RNA (7, 23, 24). An important recent discovery is that the three hydrogen atoms on methyl groups bound to positively charged nitrogen or sulfur atoms may themselves be able to serve as hydrogen bond donors, greatly expanding the possibilities for interactions (Fig. 1)(25, 26).

There are presently few systems where the full physiological role of protein methylation is understood. The question then arises of how useful is it to discover new types of protein methylation reactions when we do not fully understand the systems already described. Do such discovery efforts represent mere cataloging of modification enzymes and their substrates? The fact that the proteins are methylated by the products of genes that have often been conserved throughout the evolutionary development of organisms suggests that a full understanding of the biology of an organism needs to include an appreciation of these modifications and their functions. As we learn more from each new protein methylation system described, the range of functional roles also increases, and we are provided new targets for therapeutic intervention into human diseases (13, 27).

**Yeast ribosomes and the discovery of novel methyltransferases**

In recent years, it has been recognized that proteins of the translational apparatus, including both ribosomal proteins and elongation factors, are major substrates for methylation reac-
tions (5, 7, 8, 19, 28). These proteins, and the enzymes that modify them, have been extensively studied from the intersection of two research directions. In the first place, the combination of bioinformatics and the biology of Saccharomyces cerevisiae have allowed an approach to identify methylated sites in yeast proteins (23, 24, 29, 30). Bioinformatic analyses have allowed for the identification of open reading frames for candidate methyltransferases from genomic DNA sequences both of the major seven–H9252-strand family and of the SET domain, SPOUT, and other structural families (31–35). Importantly, the ability of S. cerevisiae to transport S-adenosylmethionine across its plasma membrane (36) allows for radiolabeling of methylated proteins in intact cells and for their detection at sub-femtomole levels (see below). Significantly, the availability of gene-knockout strains of yeast allows for biochemical analyses of mutants to identify the methyltransferases responsible for each of the modifications. Second, mass spectrometric analyses, particularly top-down intact mass measurements, have allowed the location of methylated residues in the amino acid sequence of the modified proteins (23, 24, 37). The identification of the modifying enzymes can then allow for the study of the physiological effects of each methylation reaction by examining the biology of the mutant strains lacking the methyltransferase.

My attention to translation was initiated by the discovery in the Tam laboratory (38) that the large subunit ribosomal protein Rpl12ab in yeast contains at its Arg-66 residue a unique type of modification discovered a few years earlier in my laboratory (39). This arginine residue is modified by the addition of a methyl group not only increases the steric bulk at that position in the peptide but also decreases the nucleophilicity of the linked nitrogen or sulfur atom. Methylation can also provide additional carbon-based hydrogen bond donors (25, 26). These bonds can form when the carbon atom of the methyl group is bonded to a nitrogen atom bearing a net positive charge, as with the proline, histidine, and arginine methylated derivatives shown in the figure.

Figure 1. Structures and nomenclature of selected “less traveled” methylated amino acid residues. The replacement of a hydrogen atom with a methyl group not only increases the steric bulk at that position in the peptide but also decreases the nucleophilicity of the linked nitrogen or sulfur atom. Methylation can also provide additional carbon-based hydrogen bond donors (25, 26). These bonds can form when the carbon atom of the methyl group is bonded to a nitrogen atom bearing a net positive charge, as with the proline, histidine, and arginine methylated derivatives shown in the figure.

The replacement of a hydrogen atom with a methyl group not only increases the steric bulk at that position in the peptide but also decreases the nucleophilicity of the linked nitrogen or sulfur atom. Methylation can also provide additional carbon-based hydrogen bond donors (25, 26). These bonds can form when the carbon atom of the methyl group is bonded to a nitrogen atom bearing a net positive charge, as with the proline, histidine, and arginine methylated derivatives shown in the figure.
the large ribosomal subunit of yeast (43). This work had identified the S. cerevisiae proteins that were unmodified (with intact masses corresponding to the expected amino acid sequence), and the proteins that were modified by additional masses corresponding to methylation (14-Da increases) and/or acetylation (42-Da increases) (43). This study had already identified six large subunit ribosomal proteins as possible methyltransferases and functions by the approach described above. The result of this work was the identification of four additional protein lysine methyltransferases, three from the SET domain family (Rkm2, Rkm3, and Rkm4 (44, 45)) and one from the seven–β-strand family (Rkm5 (46)). Significantly, this work also identified four additional novel and/or unusual types of methyltransferases that are described below.

Table 1 summarizes the sites of yeast methylated ribosomal proteins that occur at positions other than those modified by well-established protein lysine and arginine methyltransferases and the enzymes responsible for their formation. The success of the work identifying these sites and enzymes depended upon the combined bioinformatic, genetic, biochemical, and mass spectrometric approaches described above. In the process, we used high-capacity, high-resolution chromatographic assays to detect radiolabeled methylated amino acid residues based on sulfonated polystyrene cation-exchange resins. Using in vivo and in vitro methyl labeling with high-specific activity, S-adenosyl[3H]methionine labeled with three tritium atoms per methyl group (giving a specific activity of about 100 cpm per femtomole of methyl groups), and our ability to detect less than 10 cpm by extended liquid scintillation counting, we can detect attomole levels of methyl groups in mg or larger amounts of protein (23). This extraordinary sensitivity allows us to detect modifications that would be missed by other commonly used approaches. Additionally, we can separate very similar methylated amino acid derivatives. The resolution of this chromatography is so high that we generally see a difference in elution position between the tritiated and hydrogen forms of an amino acid (47). These methods complement immunochemical methods that can sometimes match their sensitivity but can be much less specific, as well as top-down and bottom-up mass spectrometric methods coupled with fragmentation. Top-down mass spectrometric approaches, developed by us in collaboration with the Whitelegge laboratory at UCLA, have allowed for the detection of specific modifications at specific sites along the polypeptide chain (23, 24, 45, 48).

This work provided some surprises, including the identification of the first enzyme catalyzing the methylation of a histidine residue (24), the first enzyme catalyzing eukaryotic N-terminal methylation (48), the characterization of a brand-new type of protein arginine methyltransferase (23), and the discovery of a methylated cysteine residue in a zinc cluster (23). The current status of these methylation systems is described below.
Protein N-terminal methyltransferases

In addition to the common acetylation of the N-terminal amino group of proteins in a wide variety of organisms, a small number of proteins were established some years ago to be methylated at the α-amino nitrogen atom to give mono-, di-, and trimethylated derivatives (49). Full modification can fix the positive charge on the nitrogen atom and abolish its nucleophilic reactivity. In 1987, analysis of the N-terminal sequences known in eukaryotic species suggested the existence of an enzyme or enzymes that would recognize their common Xaa-Pro-Lys N-terminal sequences (49).

In 2004, a proteomic analysis of S. cerevisiae showed that the small subunit ribosomal protein Rps25A/Rps25B (N-terminal sequence PPK) contains an N-terminal dimethylproline residue (50). By 2006, our laboratory had found evidence for five added methyl groups on the large ribosomal subunit yeast protein Rpl12ab near the N terminus including modification at Lys-3 by a SET-domain methyltransferase designated Rkm2 and by a yet unidentified enzyme at Lys-10 (44). However, these results were called into question by reports on the modifications of the orthologs of Rpl12ab in Arabidopsis thaliana (51) and Schizosaccharomyces pombe (52); both reports suggested that our MS data were more consistent with the dimethylation of the N-terminal proline residue and the trimethylation of Lys-3 in line with the similar N-terminal modification of the Arabidopsis and S. pombe proteins. Further analysis in our laboratory (45) confirmed the N-terminal modification and the major methylation of Lys-3 as opposed to Lys-10 (53). The correction of our initial methylation assignment was particularly painful because we realized that the N-terminal sequence of Rpl12ab (PPK) should have alerted us to its modification by the prediction of an XPK methyltransferase made by us some 20 years earlier (49).

The methyltransferase responsible for the modification of the N terminus of Rpl12ab and Rps25A/Rps25B in S. cerevisiae was found by mass spectrometric analysis of these proteins purified from strains with deletions of genes encoding candidate methyltransferases. Loss of the YBR261C gene resulted in the absence of the N-terminal modification of both proteins, and the encoded protein was then designated Ntm1 (N-terminal methyltransferase 1) (48). Purified Ntm1 and its METTL11A human ortholog (now designated NTMT1) catalyzed the methylation of synthetic peptides with N-terminal proline, serine, or alanine residues followed by the proline–amino group of proteins in a wide variety of organisms, a small number of proteins were established some years ago to be methylated at the α-amino nitrogen atom to give mono-, di-, and trimethylated derivatives (49). Full modification can fix the positive charge on the nitrogen atom and abolish its nucleophilic reactivity. In 1987, analysis of the N-terminal sequences known in eukaryotic species suggested the existence of an enzyme or enzymes that would recognize their common Xaa-Pro-Lys N-terminal sequences (49).
be an ortholog of yeast Efm7 in the human proteome, so it appears that a distinct type of enzyme must exist with specificity only for the N terminus. It seems unlikely that this enzyme is NTMT1 because the N-terminal sequence of GKE is unlikely to be recognized given its XPK specificity.

**Protein histidine methyltransferases**

Prior to 2010, methylation of protein histidine residues at either the \( \pi \)- or the \( \tau \)-nitrogen atom in the imidazole ring was known for a small group of animal proteins, but nothing was known of the enzymes that catalyze their formation (24). One of the proteins identified from intact MS as possibly methylated in *S. cerevisiae* was the cytoplasmic large subunit ribosomal protein Rpl3 (43). In 2010, we were able to show that the candidate methyltransferase gene *YIL110W* encoded an enzyme that was responsible for the \( \tau \)-methylation of histidine 243 in the “tryptophan finger” of the Rpl3 (24). The methyltransferase, now designated Hpm1 (histidine protein methyltransferase 1) is a member of the seven–\( \beta \)-strand family of enzymes.

Hsp-243 of Rpl3 is buried deep within the 25S RNA near the A-site and peptidyltransferase center (73); the methylated N-3 atom contacts guanine-878 and adenine-876 of the 25S rRNA (7, 74). Yeast hpm1 null cells have a pronounced deficiency of 60S subunits reflecting a significant defect in early rRNA processing with the accumulation of 35S and 23S intermediate RNA species (74). Using a dual-luciferase reporter system, we have also been able to show significant (\( >2 \)-fold) misincorporation of amino acids in cells deficient in Hpm1 (75). Hpm1-deficient cells also demonstrated increased resistance to the ribosome-binding antibiotics cycloheximide and verrucarin A, suggesting structural changes in the ribosome that accompanied the abnormal ribosome biogenesis (76). We thus proposed that Hpm1 plays a role in the orchestration of the early assembly of the large ribosomal subunit leading to a structure that gives faithful protein production.

We had hoped that the identification of Hpm1 in yeast would allow us to then identify the enzyme or enzymes responsible for the methylation of histidine residues previously established in other proteins, including the \( \tau \)-methylation of actin at His-73 in almost all eukaryotes (77). However, we noted that actin from *S. cerevisiae* was the cytoplasmic large subunit ribosomal protein Rpl3 (43). In 2010, we were able to show that the candidate methyltransferase gene *YIL110W* encoded an enzyme that was responsible for the \( \tau \)-methylation of histidine 243 in the “tryptophan finger” of the Rpl3 (24). The methyltransferase, now designated Hpm1 (histidine protein methyltransferase 1) is a member of the seven–\( \beta \)-strand family of enzymes.

**An unusual protein arginine methyltransferase that modifies the internal nitrogen atom of the guanidino group**

In mammalian cells, almost all of the extensive protein arginine methylation reactions are catalyzed by members of a sequence-related family of nine gene products, designated PRMT1 to PRMT9 (40). These cytoplasmic and nuclear enzymes have been shown to complement protein lysine methyltransferases and indeed perhaps even protein kinases in modulating transcriptional activation/repression and controlling mRNA splicing, DNA repair, the cell cycle, and signaling pathways (13, 15, 40, 82). A mitochondrial seven–\( \beta \)-strand methyltransferase from an unrelated family, designed NDUFAD7, modifies a subunit of complex I in the mammalian electron transport chain (83, 84). All of these enzymes modify only the terminal \( \omega \)-nitrogen atoms of the arginine residue.

In *S. cerevisiae*, there are two established orthologs of the PRMT1–9 family, Rmt1 (mammalian PRMT1 ortholog; Ref. 85) and Hsl7 (mammalian PRMT5 ortholog; Ref. 86). These enzymes also modify only the terminal \( \omega \)-nitrogen atoms of the arginine side chain. However, as discussed above, a novel methylated arginine residue was found in a yeast extract where the bridging \( \delta \)-nitrogen atom was monomethylated (Fig. 1) (39) in a reaction catalyzed by the Rmt2 methyletransferase (31). Interestingly, Rmt2 does not share any sequence similarities with the PRMT/Rmt1/Hsl7 family outside of the \( S \)-adenosylmethylamine-binding motifs, but it does share some similarity with the mammalian small molecule guanidinoacetate methyltransferase where a bridging guanidino nitrogen atom is also modified to form creatine (31).

The only substrate presently known for Rmt2 is the arginine 66 side chain in the yeast ribosomal protein Rpl12ab (38). Rmt2 is located in both in the nucleus and the cytoplasm of *S. cerevisiae* and co-purifies with several nuclear pore components (87). Importantly, a genome-wide transcription study of yeast cells deleted in the *RMT2* gene revealed down-regulation of a gene encoding the type II myosin heavy chain (87). In a high-throughput study, an Rmt2-GFP fusion protein was shown to relocate to the nucleus after DNA damage in yeast cells induced by hydroxyurea or methyl methanesulfonate (88). These results suggest the possibility that Rmt2 has additional protein substrates in yeast, although it is also possible that methylation on the ribosomal protein is in some way coupled to DNA transcription or repair.

BLAST searches reveal clear orthologs of Rmt2 in fungi in both the ascomycetes and basidomycetes families as well as in a variety of higher and lower plants. The only significant matches in animal cells are in flatworms such as *Schistosoma hematothium*, where a protein exists with \( 32 \% \) identity over 341 residues. The human ortholog of *S. cerevisiae* Rpl12ab, the 60 S ribosomal protein L12, is 70% identical in amino acid sequence over the entire polypeptide length to the yeast protein, and the QNRQ sequence flanking the Arg-66 residue is identical in humans and yeast. However, there is no evidence for methylation of this residue in mammals. It is unclear why a modification of a highly conserved protein would itself not be highly conserved through nature.
Spout family methyltransferases are not just for RNA substrates—a new type of protein arginine methyltransferase

Analysis of the methylated proteins of the *S. cerevisiae* ribosome yielded a further surprise. ω-Monomethylation was detected at arginine 146 in the Rps3 protein. However, mutations in none of the known three members of the yeast protein arginine methyltransferase family (Rmt1, Rmt2, and Hsl7) resulted in the loss of the modification (23). By searching for this modification in the collection of candidate methyltransferase deletion mutants, it was found that Rps3 methylation depended upon the expression of the *YOR121C* gene (23). This gene, now designated *SFM1* for spout family methyltransferase 1, encodes a member of the SPOUT family of methyltransferases, whose members previously were only known to modify RNA species (23).

Crystal structure analyses of Sfm1 showed that although the overall structure is similar to that of SPOUT family RNA methyltransferases, the active site has similarities to other protein arginine methyltransferases (89, 90). A negatively-charged surface adjacent to the active site was proposed to prevent RNA substrate binding and to enhance the binding of the basic region surrounding the Arg-146 substrate (KLRAARAKAMK). The presence in Sfm1 of a C-terminal domain in addition to the SPOUT family of methyltransferases, whose members previously were only known to modify RNA species (23).

Crystal structure analyses of Sfm1 showed that although the overall structure is similar to that of SPOUT family RNA methyltransferases, the active site has similarities to other protein arginine methyltransferases (89, 90). A negatively-charged surface adjacent to the active site was proposed to prevent RNA substrate binding and to enhance the binding of the basic region surrounding the Arg-146 substrate (KLRAARAKAMK). The presence in Sfm1 of a C-terminal domain in addition to the SPOUT family of methyltransferases, whose members previously were only known to modify RNA species (23).

Significantly, the modification site in Rps3 interacts with the 18S rRNA (7, 23). This result suggests that the modification may occur during the assembly of the ribosomal small subunit in the nucleolus.

It appears that the phylogenetic distribution of Sfm1 is limited to the ascomycetes group of fungi. The yeast Rps3 ribosomal small subunit substrate of Sfm1 is highly conserved in the human ribosome in its S3 protein (66% identical amino acid sequence), and the yeast methylation site (AARAK) is partially conserved in the human ortholog (GQRAK). However, there is no evidence for methylation at or near this site in the human ribosomal protein, although it has been reported that PRMT1 can modify human S3 at alternative arginine residues (91).

Cysteine S-methylation

Analysis by top-down MS of the proteins of the small ribosomal subunit of *S. cerevisiae* revealed that Rps27a was methylated at cysteine 39 (23). However, it appears that this modification may be fundamentally different from the ones described above in that the methylation reaction may be nonenzymatic. The crystal structure of the yeast small subunit showed that the sulfur atom of Cys-39 is displaced from a tetrahedral position of a four-cysteine zinc cluster presumably due to its methylation (23). This cluster is similar to that seen in bacterial DNA repair enzymes that demethylate spontaneously alkylated DNA, leading to the hypothesis that Rps27a may also be involved in DNA repair and/or resistance to alkylating agents (92, 93). Here, the methylated protein may simply represent a dead-end product of the repair reaction.

Methylation of cysteine residues may also occur as intermediates involved in methyl transfer in other types of enzymes, including radical S-adenosylmethionine rRNA methyltransferases (94) and methyl-coenzyme M reductases (95).

Recent work has provided evidence for methyltransferase-dependent modifications of cysteine residues that may be of regulatory significance. The NleE protein of pathogenic *E. coli* is a methyltransferase with a unique structural fold that can modify the human TAB2 and TAB3 proteins involved in NF-κB signaling to blunt the inflammatory response against the bacterium (96–98). Interestingly, the modified cysteine residue in these proteins is also coordinated to a zinc ion, although the significance of this ligation is unclear. Biochemical studies have now identified an additional NleE substrate as the ZRANB3 protein that functions in the remodeling of stalled replication forks resulting from DNA damage (97). Finally, preliminary evidence for the enzymatic methylation of a potential cysteine residue in yeast trehalose-6-phosphate synthetase has been published (99), but more work is needed to show the nature of the modified amino acid and the identity of the putative methyltransferase.

Concluding notes

Of the methyltransferases reviewed here, only the enzyme that modifies the N-terminal α-amino group of XPK substrates has a both a widespread distribution in eukaryotes and the ability to modify a wide variety of proteins. One other enzyme, the Hpm1 protein histidine methyltransferase, has a similar broad distribution in nature but may be specific for its one ribosomal protein substrate. Both the unusual Rmt2 and Rps3 protein arginine methyltransferases appear to have a much more limited phylogenetic distribution and again may only modify their one ribosomal protein substrate. Of course, it is possible that further studies may identify additional substrates for these proteins. If this is case, the observed phenotypes of the methyltransferase mutants may reflect methylation changes occurring on nonribosomal proteins. Additional work with mutants of ribosomal proteins themselves incapable of methylation may clarify this situation.

It is unclear why ribosomal proteins are the targets of such a variety of methyltransferases. In yeast, some two-thirds of all known methyltransferases act on components of the translational system, including 10 ribosomal protein methyltransferases (Rkm1–5, Rmt1–2, Sfm1, Hpm1, and Ntm1), 10 release and elongation factor protein methyltransferases (Mtlq1–2, Efml–1–7, and Dph5), 13 rRNA methyltransferases, 16 tRNA methyltransferases, and 2 mRNA methyltransferases. Because null mutants of all of the protein methyltransferase genes are viable in yeast, their collective role may be fine-tuning the protein biosynthetic apparatus for optimum activity. Further work will be needed to assess the possibility that subtypes of ribosomes exist with proteins with distinct patterns of methylation, perhaps in response to changing environmental conditions. A recent study has shown that the Hpm1 and Efml7 methyltransferases described here are themselves modified by multiple potential regulatory phosphorylation and acetylation reactions (100).

Acknowledgments—I thank Drs. Catherine Clarke and Jonathan Lowenson for their helpful comments on the paper.
Ribosomal protein methylation

References


Ribosomal protein methylation


