Published in final edited form as:

Biochemistry. 2019 February 05; 58(5): 336–345. doi:10.1021/acs.biochem.8b01047.

A family divided: Distinct structural and mechanistic features of the SpoU-TrmD (SPOUT) methyltransferase superfamily

Aiswarya Krishnamohan¹ and Jane E. Jackman^{1,*}

¹The Ohio State Biochemistry Program, Center for RNA Biology, and Department of Chemistry and Biochemistry, The Ohio State University, Columbus OH 43210, USA

Abstract

The SPOUT family of enzymes makes up the second largest of seven structurally distinct groups of methyltransferases and is named after two evolutionarily related RNA methyltransferases, SpoU and TrmD. A deep trefoil knotted domain in the tertiary structures of member enzymes defines the SPOUT family. For many years, formation of a homodimeric quaternary structure was thought to be a strict requirement for all SPOUT enzymes, critical for substrate binding and formation of the active site. However, recent structural characterization of two SPOUT members, Trm10 and Sfm1, revealed that they function as monomers without the requirement of this critical dimerization. This unusual monomeric form implies that these enzymes must exhibit a non-traditional substrate binding mode and active site architecture and may represent a new division in the SPOUT family with distinct properties removed from the dimeric enzymes. Here we discuss the mechanistic features of SPOUT enzymes with emphasis on the monomeric members and implications of this “novel” monomeric structure on cofactor and substrate binding.

For Table of contents use only

A family divided: Distinct structural and mechanistic features of the SpoU-TrmD (SPOUT) methyltransferase superfamily

*Corresponding author contact information: Department of Chemistry and Biochemistry The Ohio State University Columbus, OH 43210 Phone: 614-247-8097, Jackman.14@osu.edu.

I. Introduction

Methylation is one of the most common post-synthetic modifications found on biological macromolecules and these methylations play important roles in diverse biological processes from transcription to translation, as well as key roles in maintenance of DNA structure, function and genome integrity. The process of methylation itself is energetically expensive, as significant cellular resources are invested in production of the respective methyltransferase enzymes and biosynthesis of methyl donor cofactors, highlighting the importance of this process in biology. The methyl-donor used by most methyltransferases (MTases) to catalyze these methylations is the small molecule Sadenosylmethionine (SAM), which is converted to S-adenosyl homocysteine (SAH) as a consequence of the methyl-group transfer (Fig 1A)¹. The extremely favorable energetics (free energy change -17 kcal/mole) of this reaction likely explains the overwhelming preference for this cofactor over other methyl-group donors such as folate².

SAM-dependent MTases are classified into five evolutionarily distinct classes (classes I-V) based on structural and biochemical features that define each class. Among these five classes, the largest group identified to date is the class I MTases, characterized by a typical Rossmann-fold motif³. The second most populous group is the class IV MTases that are characterized by a catalytic domain with an unusual α/β fold forming a deep trefoil knot in the structure (Fig 1B)⁴. This class consists mostly of RNA MTases and are also known as SPOUT (SpoU-TrmD) MTases, named for the two tRNA MTases that were the founding members of the class: SpoU (renamed TrmH) that catalyzes 2'-O-methylation at G₁₈ (Gm18), and TrmD that catalyzes N1-methylation at G₃₇ (m¹G₃₇)⁴⁻⁷. The shared sequence motifs of these two enzymes enabled bioinformatics identification of other members, such as Nep1 and TrmY, which were later confirmed to belong to this class after structures revealed the characteristic knotted fold⁸⁻¹⁰. Multiple members have now been characterized and subsequently associated with a variety of substrate methylations to form methylated nucleotides such as m¹Ψ, m³Ψ, m¹G, m¹A and 2'-O methylation (Xm, where X = A, G, C or U) on diverse cellular RNAs (Table 1)¹¹⁻¹⁶. Interestingly, the SPOUT family is not restricted to RNA methylation, due to the recent identification of a family member that catalyzes mono- or di- methyl arginine in ribosomal proteins, highlighting the catalytic flexibility of the SPOUT fold¹⁷. Further, although the various 2'-O-MTases are thought to utilize generally similar catalytic mechanisms^{18,19}, the family also includes examples of enzymes that catalyze the same chemical modification by different mechanisms, as seen in the case of m¹G MTases, Trm10 and TrmD, indicating divergent catalytic strategies among evolutionarily related enzymes²⁰⁻²².

In addition to containing the topological knot in the core domain, all SPOUT enzymes were originally thought to be obligate homodimers, due to a requirement for contributions from both monomers to form the enzyme active site. Consequently, new members were further classified into either the TrmH or TrmD subfamily, based on whether the helices from the two subunits that form the dimer interface were perpendicular (TrmH) or antiparallel (TrmD). For both of these subfamilies, the RNA binding surface forms at the dimer interface, with the active site made up of residues from both subunits. Recently, however, structural characterization of two members – a tRNA m¹G₉/m¹A₉ methyltransferase

(Trm10) and a protein arginine methyltransferase (Sfm1), led to the discovery of monomeric members of the enzyme family^{17,23–26}. For these enzymes, formation of an analogous dimer appears to be prevented by the presence of a C-terminal helix that blocks the dimer interface, challenging the previously assumed strict requirement for dimerization^{17,23–26}.

Another notable feature of the SPOUT family is lack of significant sequence or structural conservation among the different members outside of the SPOUT core domain⁸. With the exception of a small group of minimalist members that only contain the SPOUT core, most enzymes have N or C terminal appended domains of varying lengths that are not significantly conserved (Fig 1C)^{8,18,27}. These appended domains clearly contribute significantly to the diversity of the SPOUT family, since the length and sequences of these domains may even differ among closely related homologs. These domains have been suggested to confer substrate specificity that is required for each enzyme⁸ although the molecular basis for the effects of the appended domains remains largely to be determined due to the relative dearth of enzyme-substrate complex structures among the SPOUT family. These developments underscore the structural flexibility possible within this family and suggest that the SPOUT fold, which is the unifying feature of these enzymes, may have been used through evolution as a scaffold to build enzymes with diverse structures, catalytic features and substrate specificities.

For all these reasons, the SPOUT MTases constitute a biochemically and evolutionarily fascinating group of enzymes whose diversity is becoming increasingly appreciated. Here, we provide an overview of structural features and mechanisms of SPOUT enzymes with emphasis on the “novel” monomeric members, Trm10 and Sfm1, which may either be exceptions to the rule or the beginning of a new sub-family within the SPOUT class, distinct from the “conventional” dimeric enzymes. The implications of this novel monomeric form on cofactor and substrate binding are also considered.

II. Conserved SAM binding conformation:

The SAM binding site of SPOUT enzymes was initially predicted to be in the knotted domain with the cofactor bound in a bent conformation, through computational docking studies with three crystal structures available at the time²⁸. The location of the binding-site formed by the three connecting loops of the trefoil knot and the ~80° bend has now been confirmed through multiple co-crystal structures of SPOUT MTases with SAM and SAM-analogs. Strikingly, despite the local and quaternary structural variations described below, a common orientation of SAM has been observed in all of these crystal structures (Fig 1D).

The dimeric SPOUT enzymes bind 2 molecules of SAM per dimer, with the trefoil knotted fold of each subunit forming a cofactor binding location. Comparisons of the apoenzyme vs. SAM- or SAM analog-bound structures of different SPOUT members reveal some structural rearrangement in the presence of the cofactor^{9,17,23,24,29,30}. The extent of these structural changes differs, however, for different enzymes. For the minimalist SPOUT MTase, TrmL, no major structural changes are seen in the presence of the SAM analog, SAH, compared to the apoenzyme²⁹. Interestingly however, in the ligand-bound structure, SAH was found in two different conformations in the two binding pockets corresponding to each monomer. In

one, SAH adopted the bent conformation characteristic of SPOUT enzymes, with no observable interactions of the homocysteine moiety with any protein residues. In the second subunit, due to a buffer molecule present at the active site, SAH was seen in an altered conformation, with the carboxyl and amino groups of the homocysteine forming a larger interaction network with multiple protein residues. The buffer molecule was proposed to mimic some features of the bound substrate nucleotide, possibly supporting a model where rearrangement of the bound cofactor occurs upon substrate binding²⁹. Similar conformational differences between the cofactor-bound vs. free structures have also been reported in Nep1⁹. Other dimeric SPOUT members, such as RlmH and TrmD exhibit much more extreme changes to accommodate the SAM cofactor. The SAM-bound crystal structure of RlmH revealed that multiple protein residues surrounding the binding site were stabilized in the presence of SAM, including a wholesale rearrangement of the C-terminal loop into a 3₁₀-helix³¹. Similarly, in the case of TrmD, a large rearrangement of the cofactor-binding site appears to be necessary to accommodate SAM³². In fact, this accommodation was recently shown to trigger a series of complex internal movements in the knotted structure that is thought to signal local conformational changes and also inactivate the second SAM binding knot in the dimer³³. This process has been proposed to result in the half-of-the-sites model of TrmD catalysis, where only one active site is functional in the dimeric enzyme, and is supported by structural and kinetic studies^{33,34}.

Structural analyses of the two known monomeric enzymes, Trm10 and Sfm1, similarly revealed interesting structural rearrangements in the presence of the bound cofactor. Comparison of *Schizosaccharomyces pombe* Trm10 (SpTrm10) structures obtained in the presence and absence of the product molecule SAH reveal conformational rearrangements primarily in the loops surrounding SAH that involve rotation of side chains to face the bound cofactor. The largest structural change is seen for a lysine side chain, which moves ~13 Å to accommodate the cofactor. Similar structural alterations also seem likely in *Sulfolobus acidocaldarius* Trm10 (SaTrm10), as some protein side chains (specifically I182 and R188) in the apo-enzyme appear to block the cofactor when superimposed with the SAH-bound form, although these residues could not be modeled in the cofactor-bound structure²⁴. In the SAM-bound structure of the protein methyltransferase Sfm1, electron density was observed for two different conformations of the methionine moiety of SAM, which were deemed equally probable based on the extensive interaction network and roughly 50% occupancy seen in both cases³⁰. Interestingly, product molecule SAH and SAM-analog MTA (5' - (methylthio) adenosine) were only seen in a single conformation in their respective co-crystal structures^{17,30}. These observations could suggest that SAM itself may undergo conformational changes during the process of methylation. Nonetheless, overall structural changes to the protein in the presence and absence of cofactors were relatively minor.

Despite these enzyme-specific variations in structure that occur to accommodate the bound cofactor, a striking conformational similarity is seen when SAH molecules from multiple crystal structures of monomeric and dimeric enzymes (representative molecules shown in Fig 1D) are superimposed, indicating a relatively conserved cofactor binding mode in all SPOUT enzymes. This binding-mode is presumably imposed by the knotted structure of the SPOUT domain and not influenced by dimerization. Since co-crystal structures of monomeric enzymes with their RNA or protein substrates are not yet available, it is

uncertain whether substrate binding may influence other conformational rearrangements of the cofactor-binding site that differ from their dimeric counterparts. In any case, it is clear that despite deviating from the homodimeric quaternary structures, the monomeric enzymes still retain the general characteristics of a SPOUT cofactor-binding site.

III. Diverging catalytic strategies

The catalytic SPOUT domain has evolved to perform methylation on diverse substrates, not only in terms of substrate molecules (tRNAs, rRNAs and proteins) but also the atoms being methylated on the substrate residues. Consistent with this diversity, SPOUT enzymes employ varied catalytic strategies, despite their shared structural fold. Here, we discuss mechanistic features of the various methylations catalyzed by dimeric and monomeric SPOUT members.

IIIA: Dimeric SPOUT enzymes: conventional catalytic themes

The active site of dimeric SPOUT MTases lies at the dimer interface, with one subunit binding the methyl donor SAM and one or more catalytic residues contributed by the second subunit to form a catalytic center^{18,19}. Since methylation almost always occurs with concomitant removal of a proton from the target atom of the substrate, a general base is generally assumed to be involved in this proton-abstraction step. Protein residues that likely fill this catalytic role have been identified for most enzymes based on sequence conservation, severe catalytic defects observed upon alteration, and location of the residue at the active site in crystal structures. Discussed below are brief descriptions of proposed mechanisms for several representative dimeric SPOUT enzymes.

2'-O-methylation—Multiple SPOUT MTases, including the founding member TrmH, catalyze 2'-O-methylation of RNA substrates. All identified 2'-O-MTases are homodimers and structurally similar to TrmH in terms of relative subunit orientation. Based on structural and mutational analysis of *T thermophiles* TrmH, an acid/base catalytic mechanism has been proposed for these enzymes³⁵. A highly conserved arginine (R41 in *T thermophilus*) is proposed to function as the general base and accepts the proton from the 2'-hydroxyl group thus facilitating methyl-group transfer (Fig 2). To perform this function, the pKa of R41 is thought to be lowered through an interaction with the phosphate group of the substrate nucleotide^{18,35}. Substitution of R41 to alanine leads to complete loss of activity, supporting its critical role³⁵. Analogous residues to the catalytic arginine and additional residues implicated in catalysis have been identified in other SPOUT 2'-O-MTases (TrmL, TrmJ, Trm56, Tsr, RlmB)¹⁸ suggesting a common mechanism and similar active site architecture for these enzymes.

m¹Ψ methylation—Two SPOUT enzymes have been identified that catalyze methylation at the N1-position of pseudouridine (m¹Ψ) – a tRNA MTase, TrmY, which catalyzes m¹Ψ₅₄³⁶ and an rRNA MTase, Nep1, which methylates 16S rRNA in bacteria at position 914 (*M jannaschii* numbering) and 18S rRNA in eukaryotes at position 1191 (*S. cerevisiae* numbering)^{37,38}. For Nep1, stabilization of developing negative charge on the substrate uridine ring by a conserved arginine is predicted to effect methylation³⁷ (Fig 3A). While one

of 4 highly conserved arginines (R88 in *S. cerevisiae*) has been proposed to fill this role based on its position in the crystal structure, mutations to any of these residues result in loss of substrate binding, making it difficult to fully assess this proposed catalytic role^{9,37}. Not much is known about the catalytic mechanism of TrmY, although a conserved aspartate (D157 in *M. jannaschii*) is only 5.5 Å from the bound cofactor in the SAH-bound crystal structure of *M. jannaschii* TrmY¹⁰ and has been suggested to function as a catalytic residue, by acting directly as a general base on the target N-1 atom (Fig 3B). Further biochemical studies are required to establish whether these proposed mechanisms are actually used by both groups of enzymes.

m³Ψ methylation—RlmH, a 23S rRNA MTase, catalyzes the N3-methylation of pseudouridine (m³Ψ) at position 1915, and is the smallest member of the SPOUT family with only 155 residues^{16,39}. The crystal structure of *E. coli* RlmH revealed a conserved tyrosine side chain (Y152 in *E. coli*) in an analogous position to the aspartate general base residue in TrmD, suggesting a role for this residue in accepting the N3-proton during catalysis³¹. Further, the enzyme was sensitive to loss of the hydroxyl group as seen by loss of activity in the Y152F variant, consistent with this hypothesis. Based on these observations it was proposed that a network of interactions between Y152 and two other highly conserved active site residues (R142 and E138 in *E. coli*) perturbs its pKa, which leads to its acceptance of the N3-proton from the substrate pseudouridine, thus facilitating methylation³¹ (Fig 4).

m¹G methylation—The mechanism of m¹G₃₇ (N1-methylation) formation in tRNAs by TrmD is one of the best studied mechanisms among all SPOUT members. Consistent with other SPOUT enzymes, the active site was determined through structural and biochemical analyses to lie at the dimer interface²⁰. Extensive kinetic analyses further helped identify the roles of several conserved active site residues in the reaction mechanism^{20,34}. A highly conserved aspartate residue (D169 in *E. coli*) functions as the general base to accept the N1-proton thus preparing N1 for the nucleophilic attack on the sulfonium center of SAM, a role that is strongly supported by kinetic and structural evidence^{20,21,32}. pH-rate analysis identified a single ionization required for catalysis which is very likely the deprotonation of N1, and this proton transfer is considered to be the rate-limiting step for the reaction (Fig 5A)²¹. The negative charge at O6 of substrate nucleotide G37 that results from this proton transfer is stabilized by Mg²⁺ that is required for catalysis, and stabilized by a conserved arginine residue (R154 in *E. coli*)²¹. The active site Mg²⁺-ion also appears to play a critical role in positioning the D169 side chain in the correct orientation for catalysis, making this enzyme the first example of a SPOUT member exhibiting a dependence on divalent metal ions for catalysis^{21,32}.

IIIB: Monomeric SPOUT enzymes

The monomeric state of the Trm10 and Sfm1 requires that the catalytic center be contained within a single subunit, which is a dramatic shift from the previously established mechanisms where catalytic residues are contributed by two different polypeptides. Discussed below are proposed catalytic mechanisms of the two known monomeric SPOUT MTases, Trm10 and Sfm1.

m¹G/m¹A methylation by Trm10 – alternative route to the same place—The crystal structures of two yeast homologs of Trm10 (from *S. pombe* and *S. cerevisiae*) were the first to be obtained for this family and confirmed the presence of the knotted SPOUT fold as predicted, but revealed a distinct monomeric structure that was confirmed by SAXS (Small Angle X-ray Scattering) analysis²³. Initially, based on strict conservation and location near the SAM cofactor, an aspartate residue (D210 in *S. cerevisiae*) was proposed to act as a general base, analogous to D169 in TrmD²³. However, subsequent biochemical studies showed that D210 alterations resulted in minimal effects on methylation and did not affect the pH dependence of the catalytic step, effectively ruling out this proposed role²². In fact, no other conserved active site aspartates/glutamates could be identified that are significantly required for catalysis and substantial loss of enzymatic activity was only observed with a triple variant in which three conserved carboxylate side-chains had been removed at the same time²². *S. cerevisiae* Trm10 (ScTrm10) also does not require any divalent metal ions for catalysis²². Thus, despite the fact that these two SPOUT enzymes catalyze the same chemical reaction, these data indicate that m¹G catalysis by Trm10 occurs via a mechanism that is fully distinct from the mechanism employed by TrmD.

Characterization of multiple members of the Trm10 subfamily from diverse species revealed another intriguing feature, in which either G₉ or A₉ nucleotides could be substrates for methylation. This type of flexibility had not been observed among enzymes that modify the nucleotide bases, and was particularly striking because of the predicted difference in pK_a between the target N1 atoms of guanine vs. adenine. In particular, since the guanine N1 is protonated at physiological pH, it presumably requires deprotonation before methylation, in contrast to adenine, where N1 is deprotonated at physiological pH, thus suggesting very different requirements for active site residues catalyzing both reactions. Trm10 enzymes have now been identified that are strictly m¹G₉-specific (fungal Trm10; human TRMT10A), strictly m¹A₉-forming (*S. acidocaldarius* Trm10), and even bifunctional, catalyzing both m¹G₉ and m¹A₉ (human TRMT10C and *T. kodakarensis* Trm10)^{14,15,40}. pH-rate analysis of the m¹G₉-specific ScTrm10 revealed a single basic ionization that is relevant for the enzyme-independent catalytic step, as is observed with TrmD, suggestive of a similar rate-determining N1-deprotonation of the target guanine G₉ that occurs without the participation of an obligate general base protein residue²². This hypothesis was further supported by analysis of the bifunctional m¹R₉ Trm10 from *Thermococcus kodakarensis* (TkTrm10), for which extensive mutagenesis also failed to identify any residue capable of serving as a general base²⁵, and for which pH-rate profiles of m¹A₉ methylation revealed pH-independent catalytic activity, consistent with the lack of a requirement to remove an N1-proton from A₉ (Krishnamohan, Dodbele and Jackman, unpublished) (Fig 5B). These observations suggest that although the N1-deprotonation is the critical ionization that drives methylation for both SPOUT m¹G methyltransferases TrmD and Trm10, the mechanism by which this is achieved in the two active sites differs markedly. While a single residue is responsible for this deprotonation in TrmD (Fig 5A), multiple residues at the active site collaborate to establish an active site environment that promotes this deprotonation in Trm10 (Fig 5B).

Arginine methylation by Sfm1 – introduction of protein methyltransferases

into the family: The catalytic mechanism used by Sfm1 enzymes has not yet been fully elucidated, but structural analyses identify active site residues that could participate in a plausible catalytic mechanism. Strikingly, the active site architecture of *S. cerevisiae* Sfm1 revealed four highly conserved residues (E9, W15, E19 and F180 in *S. cerevisiae*) analogous to catalytically important residues in other non-SPOUT type protein arginine MTases (PRMTs) in similar spatial orientations¹⁷. The two glutamate residues (E9 and E19) are structurally homologous to residues from the “double-E loop” in other PRMTs, which function as the catalytic center and are critical for catalysis (Fig 6)^{17,41–43}. By comparison with the roles of these glutamates in other PRMT active sites, these residues may interact with the guanidino group of substrate arginines and E19 would act as the N η -proton acceptor that facilitates methylation. These similarities in terms of active site architecture between unrelated PRMTs that catalyze the same reaction is a particularly intriguing example of convergent evolution and further demonstrates the catalytic flexibility of SPOUT enzymes.

IV. Quaternary structure and substrate recognition

The N and C-terminal extended domains (NTDs/CTDs) of SPOUT enzymes have generally been thought to aid in substrate recognition. These domains differ in both length and sequence among different family members (Fig 1C). Structural studies have further shown that these domains fold independent of the conserved SPOUT core domain, giving rise to the hypothesis that they allow for a wide variety of substrate specificities to be elaborated onto the catalytic core domain. For example, the rRNA MTase, RlmB, has an appended NTD that shares similarities with ribosomal proteins and therefore predicted to be involved in 23S rRNA recognition (Fig 1C)^{27,44}. tRNA MTases however, show much more variability in these domains, which may be a consequence of recognizing distinct positions on these highly structured small RNAs. Further, the distinct oligomeric states have clear implications on substrate recognition. In this section, differences in substrate recognition among SPOUT members are discussed. For clarity, we focus only on tRNA MTases to highlight important differences between various family members.

Dimeric minimalist SPOUT enzyme, TrmL:

The minimalist SPOUT member, TrmL (a Um₃₄/Cm₃₄ MTase), does not have any appended domains and is therefore restricted to only the SPOUT core to define its RNA binding properties (Fig 1C). Due to the absence of the extended domains in TrmL, requirement for a partner protein to compensate for this absence was initially proposed⁴⁵. However, it was later shown that *E. coli* TrmL alone was sufficient for activity on in vivo purified substrate tRNAs that contain certain modifications²⁹. Based on biochemical studies, a model in which the tRNA binding site spans both subunits of the TrmL dimer, and residues from the N-terminal half of one SPOUT core domain perform the function of the appended domains found in other SPOUT enzymes, has been proposed²⁹.

Dimeric enzymes with extended domains:

SPOUT tRNA MTases containing N-or C-terminal domains utilize these extensions in diverse ways to execute the required substrate methylation. The NTD and CTD of TrmH (a Gm₁₈ MTase) participate in non-specific binding to tRNAs but not in substrate vs. non-substrate discrimination²⁷. In the case of TrmD (an m¹G₃₇ MTase), residues from the CTD of one monomer and the NTD of the second monomer interact with the core of substrate tRNAs, helping to position the active site appropriately for methylation in the anticodon, thus contributing to regiospecificity during modification (Fig 7A)^{32,46}. TrmJ (an Xm₃₂ MTase) forms an asymmetric dimer, with the SPOUT domain and its associated CTD forming independent dimeric associations¹² connected by a flexible linker. In this case, a model in which both domains are responsible for overall binding with substrate specificity achieved by involvement of the flexible linker has been proposed¹².

Monomeric SPOUT tRNA MTase, Trm10 (m¹R₉ MTase):

The positively charged surface observed in crystal structures of three homologs of Trm10 (SpTrm10, ScTrm10 and SaTrm10), led to prediction of the tRNA binding mode of this monomeric enzyme^{23,24}. Based on biochemical studies, the NTD of SpTrm10 and ScTrm10 appear to participate in overall tRNA binding, while some residues in the SPOUT domain may ensure substrate specificity, although the identity of these has not been determined²³. A computational docking model of *E coli* initiator tRNA with full length SaTrm10 predicted that the overall shape of the enzyme complements the tertiary structure of the substrate tRNA (Fig 7B)²⁴. Based on the tRNA-like shape of the full-length enzyme, it is tempting to speculate that these enzymes use the tertiary structure of substrate tRNAs to identify the position of the substrate R₉ nucleotide. Since the appended domains contribute to formation of the tRNA-like shape, this could be one way in which these domains also contribute to the regiospecificity of the m¹R₉ modification. Interestingly, the tRNA binding surface identified in both fungal and archaeal Trm10 corresponds to helices that form the dimer interface of traditional SPOUT enzymes, implying that Trm10 replaces the second subunit with the substrate tRNA itself (Fig 7B)²⁴. Further, this model could also help explain the results of a previous biochemical study which showed that ScTrm10 is capable of methylating all tRNAs with short (so called Type I) variable loops *in vitro*, but does not methylate tRNAs with extended (Type II) variable loops⁴⁷. Considering that the position of the extended variable loop in the overall tertiary structure of a tRNA substrate would lie at the predicted enzyme:tRNA interface, it seems likely that its presence may block access to the substrate nucleotide located in the tRNA core. However, given these predictions entirely rely on computational models and crystal structures determined in the absence of RNA substrate, further structural and biochemical studies are required to obtain an accurate representation of the monomeric SPOUT binding mode.

V. Conclusions

Characterization of a broad variety of SPOUT family members has revealed significant catalytic flexibility of these enzymes and their ability to act on a wide variety of substrates. These observations underscore an evolutionary scenario in which the SPOUT core domain could have provided a scaffold for development of diverse catalytic strategies and substrate

specificities through acquisition of necessary catalytic residues at key locations in the structure. Further diversity appears to have been achieved through addition of appended domains (NTDs and CTDs) and by varying the relative orientations of the core domains as they are assembled into the homodimeric structures adopted by some enzymes. With the discovery of the monomeric enzymes, yet another way has been revealed to achieve diversity in function and mechanism. Remarkably, the two monomeric SPOUT enzymes identified so far have unique capabilities in terms of substrate specificity (acting in multiple-nucleotide-specific base methylation and protein methylation) that distinguish them from the rest of the family. Whether the monomeric forms of Trm10 and Sfm1 are a result of developing new substrate specificities or in contrast drove acquisition of new specificity is an exciting question that remains to be answered. In any case, it is clear that the monomeric enzymes represent an intriguing new addition to the SPOUT family to which new members may yet be added through future studies of uncharacterized members.

ACKNOWLEDGMENT

Supported by NIGMS of the National Institutes of Health under award number R01GM130135 to J.E.J.

REFERENCES

- (1). Chiang PK, Gordon RK, Tal J, Zeng GC, Doctor BP, Pardhasaradhi K, and McCann PP (1996) S-Adenosylmethionine and methylation. *FASEB J* 10, 471–480. [PubMed: 8647346]
- (2). Cantoni GL (1975) Biological Methylation: Selected Aspects. *Annu. Rev. Biochem.* 44, 435–451. [PubMed: 1094914]
- (3). Schubert HL, Blumenthal RM, and Cheng X (2003, 6) Many paths to methyltransfer: A chronicle of convergence. *Trends Biochem. Sci.* 28, 329–335. [PubMed: 12826405]
- (4). Anantharaman V, Koonin EV, and Aravind L (2002) SPOUT: a class of methyltransferases that includes spoU and trmD RNA methylase superfamilies, and novel superfamilies of predicted prokaryotic RNA methylases. *J Mol. Microbiol. Biotechnol.* 4, 71–75. [PubMed: 11763972]
- (5). Persson BC, Jäger G, and Gustafsson C (1997) The spoU gene of *Escherichia coli*, the fourth gene of the spoT operon, is essential for tRNA (Gm18) 2'-O-methyltransferase activity. *Nucleic Acids Res.* 25, 4093–4097. [PubMed: 9321663]
- (6). Cavaillé J, Chetouani F, and Bachellerie JP (1999) The yeast *Saccharomyces cerevisiae* YDL112w ORF encodes the putative 2'-O-ribose methyltransferase catalyzing the formation of Gm18 in tRNAs. *RNA* 5, 66–81. [PubMed: 9917067]
- (7). Björk GR, Wikström PM, and Byström AS (1989) Prevention of translational frameshifting by the modified nucleoside 1-methylguanosine. *Science.* 244, 986–989. [PubMed: 2471265]
- (8). Tkaczuk KL, Dunin-Horkawicz S, Purta E, and Bujnicki JM (2007) Structural and evolutionary bioinformatics of the SPOUT superfamily of methyltransferases. *BMC Bioinformatics* 8, 73. [PubMed: 17338813]
- (9). Taylor AB, Meyer B, Leal BZ, Kötter P, Schirf V, Demeler B, Hart PJ, Entian KD, and Wöhnert J (2008) The crystal structure of Nep1 reveals an extended SPOUT-class methyltransferase fold and a pre-organized SAM-binding site. *Nucleic Acids Res.* 36, 1542–1554. [PubMed: 18208838]
- (10). Chen H-Y, and Yuan YA (2010) Crystal Structure of Mj1640/DUF358 Protein Reveals a Putative SPOUT-Class RNA Methyltransferase. *J Mo. Cell Bio.* 2, 366–374.
- (11). Purta E, van Vliet F, Tkaczuk KL, Dunin-Horkawicz S, Mori H, Droogmans L, and Bujnicki JM (2006) The yfhQ gene of *Escherichia coli* encodes a tRNA: Cm32/Um32 methyltransferase. *BMC Mol. Biol.* 7, 23. [PubMed: 16848900]
- (12). Liu R-J, Long T, Zhou M, Zhou X-L, and Wang E-D (2015) tRNA recognition by a bacterial tRNA Xm32 modification enzyme from the SPOUT methyltransferase superfamily. *Nucleic Acids Res.* 43, 7489–7503. [PubMed: 26202969]

- (13). Somme J, Van Laer B, Roovers M, Steyaert J, Versées W, and Droogmans L (2014) Characterization of two homologous 2'-O-methyltransferases showing different specificities for their tRNA substrates. *RNA* 20, 1257–1271. [PubMed: 24951554]
- (14). Jackman JE, Montange RK, Malik HS, and Phizicky EM (2003) Identification of the yeast gene encoding the tRNA m1G methyltransferase responsible for modification at position 9. *RNA* 9, 574–585. [PubMed: 12702816]
- (15). Kempenaers M, Roovers M, Oudjama Y, Tkaczuk KL, Bujnicki JM, and Droogmans L (2010) New archaeal methyltransferases forming 1-methyladenosine or 1-methyladenosine and 1-methylguanosine at position 9 of tRNA. *Nucleic Acids Res.* 38, 6533–6543. [PubMed: 20525789]
- (16). Purta E, Kaminska KH, Kasprzak JM, Bujnicki JM, and Douthwaite S (2008) YbeA is the m3 methyltransferase RlmH that targets nucleotide 1915 in 23S rRNA. *RNA* 14, 2234–2244. [PubMed: 18755835]
- (17). Lv F, Zhang T, Zhou Z, Gao S, Wong CC, Zhou J-Q, and Ding J (2015) Structural basis for Sfm1 functioning as a protein arginine methyltransferase. *Cell Discov.* 1, 15037. [PubMed: 27462434]
- (18). Hori H (2017) Transfer RNA methyltransferases with a SpoU-TrmD (SPOUT) fold and their modified nucleosides in tRNA. *Biomolecules* 7, 23.
- (19). Swinehart WE, and Jackman JE (2015) Diversity in mechanism and function of tRNA methyltransferases. *RNA Biol.* 12, 398–411. [PubMed: 25626150]
- (20). Elkins PA, Watts JM, Zalacain M, van Thiel A, Vitazka PR, Redlak M, Andraos-Selim C, Rastinejad F, and Holmes WM (2003) Insights into catalysis by a knotted TrmD tRNA methyltransferase. *J Mo. Biol.* 333, 931–949.
- (21). Sakaguchi R, Lahoud G, Christian T, Gamper H, and Hou Y-M (2014) A divalent metal ion-dependent N(1)-methyl transfer to G37-tRNA. *Chem. Biol.* 21, 1351–1360. [PubMed: 25219964]
- (22). Krishnamohan A, and Jackman JE (2017) Mechanistic features of the atypical tRNA m1G9 SPOUT methyltransferase, Trm10. *Nucleic Acids Res.* 45, 9019–9029. [PubMed: 28911116]
- (23). Shao Z, Yan W, Peng J, Zuo X, Zou Y, Li F, Gong D, Ma R, Wu J, Shi Y, Zhang Z, Teng M, Li X, and Gong Q (2014) Crystal structure of tRNA m1G9 methyltransferase Trm10: insight into the catalytic mechanism and recognition of tRNA substrate. *Nucleic Acids Res.* 42, 509–525. [PubMed: 24081582]
- (24). Van Laer B, Roovers M, Wauters L, Kasprzak JM, Dyzma M, Deyaert E, Kumar Singh R, Feller A, Bujnicki JM, Droogmans L, and Versées W (2016) Structural and functional insights into tRNA binding and adenosine N1-methylation by an archaeal Trm10 homologue. *Nucleic Acids Res.* 44, 940–953. [PubMed: 26673726]
- (25). Singh RK, Feller A, Roovers M, Elder DVAN, Wauters L, Droogmans L, and Versées W (2018) Structural and biochemical analysis of the dual-specificity Trm10 enzyme from *Thermococcus kodakaraensis* prompts reconsideration of its catalytic mechanism. *RNA* 24, 1080–1092. [PubMed: 29848639]
- (26). Oerum S, Roovers M, Rambo RP, Kopec J, Bailey HJ, Fitzpatrick F, Newman JA, Newman WG, Amberger A, Zschocke J, Droogmans L, Oppermann U, and Yue WW (2018) Structural insight into the human mitochondrial tRNA purine N1-methyltransferase and ribonuclease P complexes. *J Biol. Chem.* 293, 12862–12876. [PubMed: 29880640]
- (27). Ochi A, Makabe K, Yamagami R, Hirata A, Sakaguchi R, Hou YM, Watanabe K, Nureki O, Kuwajima K, and Hori H (2013) The catalytic domain of topological knot tRNA methyltransferase (TrmH) discriminates between substrate tRNA and nonsubstrate tRNA via an induced-fit process. *J Biol. Chem.* 288, 25562–25574. [PubMed: 23867454]
- (28). Kurowski MA, Sasin JM, Feder M, Debski J, and Bujnicki JM (2003) Characterization of the cofactor-binding site in the SPOUT-fold methyltransferases by computational docking of S-adenosylmethionine to three crystal structures. *BMC Bioinformatics* 4, 9. [PubMed: 12689347]
- (29). Liu RJ, Zhou M, Fang ZP, Wang M, Zhou XL, and Wang ED (2013) The tRNA recognition mechanism of the minimalist SPOUT methyltransferase, TrmL. *Nucleic Acids Res.* 41, 7828–7842. [PubMed: 23804755]
- (30). Wang C, Zeng J, and Xie W (2017) A flexible cofactor-binding loop in the novel arginine methyltransferase Sfm1. *FEBS Lett.* 591, 433–441. [PubMed: 27990635]

- (31). Koh CS, Madireddy R, Beane TJ, Zamore PD, and Korostelev AA (2017) Small methyltransferase RlmH assembles a composite active site to methylate a ribosomal pseudouridine. *Sci. Rep.* 7, 969. [PubMed: 28428565]
- (32). Ito T, Masuda I, Yoshida K, Goto-Ito S, Sekine S, Suh SW, Hou Y-M, and Yokoyama S (2015) Structural basis for methyl-donor-dependent and sequence-specific binding to tRNA substrates by knotted methyltransferase TrmD. *Proc. Natl. Acad. Sci. U. S. A* 112, E4197–4205. [PubMed: 26183229]
- (33). Christian T, Sakaguchi R, Perlinska AP, Lahoud G, Ito T, Taylor EA, Yokoyama S, Sulkowska JI, and Hou Y-M (2016) Methyl transfer by substrate signaling from a knotted protein fold. *Nat Struct Mol Biol* 23, 941–948. [PubMed: 27571175]
- (34). Christian T, Lahoud G, Liu C, and Hou Y-M (2010) Control of catalytic cycle by a pair of analogous tRNA modification enzymes. *J Mol. Biol.* 400, 204–217. [PubMed: 20452364]
- (35). Nureki O, Watanabe K, Fukai S, Ishii R, Endo Y, Hori H, and Yokoyama S (2004) Deep knot structure for construction of active site and cofactor binding site of tRNA modification enzyme. *Structure* 12, 593–602. [PubMed: 15062082]
- (36). Chatterjee K, Blaby IK, Thiaville PC, Majumder M, Grosjean H, Adam Yuan Y, Gupta R, and De Crecy-Lagard V (2012) The archaeal C0G1901/DUF358 SPOUT-methyltransferase members, together with pseudouridine synthase Pus10, catalyze the formation of 1-methylpseudouridine at position 54 of tRNA. *RNA* 18, 421–433. [PubMed: 22274953]
- (37). Thomas SR, Keller CA, Szyk A, Cannon JR, and LaRonde-LeBlanc NA (2011) Structural insight into the functional mechanism of Nep1/Emg1 N1-specific pseudouridine methyltransferase in ribosome biogenesis. *Nucleic Acids Res.* 39, 2445–2457. [PubMed: 21087996]
- (38). Wurm JP, Meyer B, Bahr U, Held M, Frolov O, Köster P, Engels JW, Heckel A, Karas M, Entian KD, and Wöhnert J (2010) The ribosome assembly factor Nep1 responsible for Bowen-Conradi syndrome is a pseudouridine-N1-specific methyltransferase. *Nucleic Acids Res.* 38, 2387–2398. [PubMed: 20047967]
- (39). Ero R, Peil L, Liiv A, and Remme J (2008) Identification of pseudouridine methyltransferase in *Escherichia coli*. *RNA* 14, 2223–2233. [PubMed: 18755836]
- (40). Vilardo E, Nachbagauer C, Buzet A, Taschner A, Holzmann J, and Rossmann W (2012) A subcomplex of human mitochondrial RNase P is a bifunctional methyltransferase—extensive moonlighting in mitochondrial tRNA biogenesis. *Nucleic Acids Res.* 40, 11583–11593. [PubMed: 23042678]
- (41). Yue Y, Chu Y, and Guo H (2015) Computational Study of Symmetric Methylation on Histone Arginine Catalyzed by Protein Arginine Methyltransferase PRMT5 through QM/MM MD and Free Energy Simulations. *Molecules* 20, 10032–10046. [PubMed: 26035101]
- (42). Zhang R, Li X, Liang Z, Zhu K, Lu J, Kong X, Ouyang S, Li L, Zheng YG, and Luo C (2013) Theoretical Insights into Catalytic Mechanism of Protein Arginine Methyltransferase 1. *PLoS One* (Jeltsch A, Ed.) 8, e72424. [PubMed: 23977297]
- (43). Zhang X, Zhou L, and Cheng X (2000) Crystal structure of the conserved core of protein arginine methyltransferase PRMT3. *EMBO J* 19, 3509–3519. [PubMed: 10899106]
- (44). Michel G, Sauv V, Larocque R, Li Y, Matte A, and Cygler M (2002) The structure of the RlmB 23S rRNA methyltransferase reveals a new methyltransferase fold with a unique knot. *Structure* 10, 1303–1315. [PubMed: 12377117]
- (45). Lim K, Zhang H, Tempczyk A, Krajewski W, Bonander N, Toedt J, Howard A, Eisenstein E, and Herzberg O (2003) Structure of the YibK methyltransferase from *Haemophilus influenzae* (HI0766): A cofactor bound at a site formed by a knot. *Proteins Struct. Funct. Genet.* 51, 56–67.
- (46). Christian T, and Hou Y-M (2007) Distinct determinants of tRNA recognition by the TrmD and Trm5 methyl transferases. *J Mol. Biol.* 373, 623–632. [PubMed: 17868690]
- (47). Swinehart WE, Henderson JC, and Jackman JE (2013) Unexpected expansion of tRNA substrate recognition by the yeast m1G9 methyltransferase Trm10. *RNA* 19, 1137–1146. [PubMed: 23793893]
- (48). Bystrm AS, and Bjrk GR (1982) Chromosomal location and cloning of the gene (trmD) responsible for the synthesis of tRNA (m1G) methyltransferase in *Escherichia coli* K-12. *MGG Moi. Gen. Genet.* 188, 440–446.

- (49). Byström AS, Hjalmarsson KJ, Wikström PM, and Björk GR (1983) The nucleotide sequence of an *Escherichia coli* operon containing genes for the tRNA(m¹G)methyltransferase, the ribosomal proteins S16 and L19 and a 21-K polypeptide. *EMBO J* 2, 899–905. [PubMed: 6357787]
- (50). Hjalmarsson KJ, Bystrom AS, and Bjork GR (1983) Purification and characterization of transfer RNA (guanine-I)methyltransferase from *Escherichia coli*. *J Biol. Chem.* 258, 1343–1351. [PubMed: 6337136]
- (51). Benitez-Paez A, Villarroya M, Douthwaite S, Gabaldon T, and Armengod ME (2010) YibK is the 2'-O-methyltransferase TrmL that modifies the wobble nucleotide in *Escherichia coli* tRNA^{Leu} isoacceptors. *RNA* 16, 2131–2143. [PubMed: 20855540]
- (52). Kuratani M, Bessho Y, Nishimoto M, Grosjean H, and Yokoyama S (2008) Crystal Structure and Mutational Study of a Unique SpoU Family Archaeal Methylase that Forms 2'-O-Methylcytidine at Position 56 of tRNA. *J Mot. Biol.* 375, 1064–1075.
- (53). Renalier M-H, Joseph N, Gaspin C, Thebault P, and Mougin A (2005) The Cm56 tRNA modification in archaea is catalyzed either by a specific 2'-O-methylase, or a C/D sRNP. *RNA* 11, 1051–1063. [PubMed: 15987815]
- (54). Dunstan MS, Hang PC, Zelinskaya NV, Honek JF, and Conn GL (2009) Structure of the Thiostrepton Resistance Methyltransferase-S-Adenosyl-I-methionine Complex and Its Interaction with Ribosomal RNA. *J Biol. Chem.* 284, 17013–17020. [PubMed: 19369248]
- (55). Thompson J, Schmidt F, and Cundliffe E (1982) Site of action of a ribosomal RNA methylase conferring resistance to thiostrepton. *J Biol. Chem.* 257, 7915–7917. [PubMed: 6806287]
- (56). Lovgren JM, and Wikstrom PM (2001) The rlmB Gene Is Essential for Formation of Gm2251 in 23S rRNA but Not for Ribosome Maturation in *Escherichia coli*. *J Bacteriol.* 183, 6957–6960.
- (57). Young BD, Weiss DI, Zurita-Lopez CI, Webb KJ, Clarke SG, and McBride AE (2012) Identification of methylated proteins in the yeast small ribosomal subunit: A role for SPOUT methyltransferases in protein arginine methylation. *Biochemistry* 51, 5091–5104. [PubMed: 22650761]

Figure 1: Shared features of SPOUT methyltransferases.

(A) Conversion of S-adenosyl methionine (SAM) to S-adenosyl homocysteine (SAH) during methylation by SPOUT enzymes. (B) SPOUT domain of *E. coli* TrmD (from PDB 1P9P) showing the characteristic topological trefoil knot found in all SPOUT enzymes. (C) Schematic cartoon of select SPOUT enzymes showing differences in N/C-terminal extension domains (adapted from ^{18,27}). The α/β refers to the general arrangement of α -helices and β -sheets observed in crystal structures. The RNA binding motif of RlmB is predicted based on similarity to ribosomal proteins. (D) The SAM/SAH cofactors from multiple crystal structures of SPOUT methyltransferases were aligned using PyMol. SAH from crystal structures of SaTrm10 (5A7Z, green), ScTrm10 (4JWJ, blue), *E. coli* TrmL (4JAL, red), *S. cerevisiae* Nep1 (3OII, orange), *S. cerevisiae* Sfm1 (5H5E, teal), *E. coli* TrmD (1P9P, purple) and SAM from *E. coli* RlmH (5TWJ, black) are pictured.

Figure 2: Putative mechanism of TrmH-catalyzed 2'-O-methylation.

A conserved arginine, whose pKa is likely altered through interaction with the phosphate of the substrate nucleotide, acts as the general base for abstraction of the 2'-OH proton, facilitating nucleophilic attack³⁵.

Figure 3: Two predicted mechanisms for m¹Ψ catalysis.

(A) A conserved arginine in the Nep1 active site may aid in stabilizing the negative charge of the uridine ring, which in turn prepares N1 for nucleophilic attack³⁷ (B) A highly conserved aspartate in the TrmY active site potentially acts as a general base for abstracting the N1 proton¹⁰.

Figure 4: Hypothetical mechanism for m³Ψ catalysis by RlmH.

An arginine in the active site of RlmH forms a salt bridge with an active site glutamate, thus influencing the pK_a of a tyrosine residue to act as a general base for N3 proton abstraction³¹.

Figure 5: Differences in m^1G catalysis by two SPOUT members, TrmD and Trm10.

(A) Mechanism of m^1G_{37} formation by TrmD. An aspartate general base catalyzes N1 proton abstraction, and the resultant negative charge on O6 is stabilized through a Mg^{2+} -ion^{20,21,32} (B) Predicted mechanism of m^1R_9 catalysis by Trm10. Trm10 does not utilize a general base for N1 proton abstraction but may use a collaborative active site that lowers the pKa of the N1 proton and enables transfer to bulk solvent, possibly through a water network. No metal ions are required for methylation by Trm10²².

Figure 6: Predicted mechanism of arginine methylation by Sfm1.

Two conserved active site glutamates position the guanido group, one of which acts as the general base to accept the N η -proton¹⁷.

Figure 7: Substrate tRNA-binding modes of TrmD and Trm10.

(A) Co-crystal structure of TrmD with tRNA illustrates the binding of tRNA at the dimer interface. The two monomers forming the dimer are indicated in light and dark purple. Residues from the CTD of one monomer and the NTD from the second monomer interact with the core of the tRNA (indicated by the circle) and contribute to positioning the substrate G₃₇ (orange) in the anticodon loop at the active site (PDB: 4YVI)^{32,46} (B) Computational docking model of SaTrm10 with *E. coli* initiator tRNA predicts that the shape of full-length enzyme recognizes the L-shape of the tRNA which may help specifically identify the substrate core nucleotide R₉ (orange)²⁴.

Table 1:

Diversity in methylations catalyzed by SPOUT MTases

SPOUT MTase	Sub-family	Substrate	Modification
TrmH (SpoU)	TrmH	tRNA	Gm ₁₈ ⁵
TrmD	TrmD	tRNA	m ¹ G ₃₇ ⁴⁸⁻⁵⁰
TrmY	TrmH ¹⁰	tRNA	m ¹ Ψ ₅₄ ³⁶
TrmL	TrmH ^{29,45}	tRNA	Um ₃₄ /Cm ₃₄ ⁵¹
TrmJ	TrmH ¹²	tRNA	Xm ₃₂ ¹¹
Trm56	TrmH ⁵²	tRNA	Cm ₅₆ ⁵³
Trm10	Monomeric ²³⁻²⁵	tRNA	m ¹ G ₉ /m ¹ A ₉ ^{14,15}
NepI	TrmH ⁹	16S rRNA 18S rRNA	^a m ¹ Ψ ₉₁₄ ³⁸ ^b m ¹ Ψ ₁₁₉₁ ³⁸
RlmH	TrmD ^{8,31}	23SrRNA	^c m ³ Ψ ₁₉₁₅ ¹⁶
Tsr	TrmH ⁵⁴	23SrRNA	^c Am ₁₀₆₇ ⁵⁵
RlmB	TrmH ⁴⁴	23SrRNA	^c Gm ₂₂₅₁ ⁵⁶
Sfml	Monomeric ¹⁷	Ribosomal protein S3	ω-methylation of ^b R146 ⁵⁷

Numbering corresponds to;

^a – *Methanococcus jannaschii*;^b – *Saccharomyces cerevisiae*;^c – *Escherichia coli*