Published in final edited form as:

Enzymes. 2017 ; 41: 51–88. doi:10.1016/bs.enz.2017.03.002.

The Evolution of Substrate Specificity by tRNA Modification Enzymes

Katherine M. McKenney, Mary Anne T. Rubio, and Juan D. Alfonzo¹

Department of Microbiology, Ohio State Biochemistry Program, The Center for RNA Biology, The Ohio State University, Columbus, OH, United States

Abstract

All types of nucleic acids in cells undergo naturally occurring chemical modifications, including DNA, rRNA, mRNA, snRNA, and most prominently tRNA. Over 100 different modifications have been described and every position in the purine and pyrimidine bases can be modified; often the sugar is also modified [1]. In tRNA, the function of modifications varies; some modulate global and/or local RNA structure, and others directly impact decoding and may be essential for viability. Whichever the case, the overall importance of modifications is highlighted by both their evolutionary conservation and the fact that organisms use a substantial portion of their genomes to encode modification enzymes, far exceeding what is needed for the de novo synthesis of the canonical nucleotides themselves [2]. Although some modifications occur at exactly the same nucleotide position in tRNAs from the three domains of life, many can be found at various positions in a particular tRNA and their location may vary between and within different tRNAs. With this wild array of chemical diversity and substrate specificities, one of the big challenges in the tRNA modification field has been to better understand at a molecular level the modes of substrate recognition by the different modification enzymes; in this realm RNA binding rests at the heart of the problem. This chapter will focus on several examples of modification enzymes where their mode of RNA binding is well understood; from these, we will try to draw general conclusions and highlight growing themes that may be applicable to the RNA modification field at large.

1. INTRODUCTION

The importance of RNA binding percolates all aspects of RNA processing and involves an amazing set of different RNA interaction domains; some are readily recognizable by their degree of sequence conservation, and others are more discreet and escape easy prediction [1]. In general, RNA-binding domains may be divided into two major groups: (1) those belonging to proteins that act alone (“professional RNA-binding proteins”); (2) those that are appended to other proteins, acting as binding modules that work in *cis* or *trans* to anchor other functionalities at the site of action. It is clear that a critical step for the evolution of extant life must have involved the appearance of polypeptides able to interact in some fashion with RNA, facilitating the transition into an RNA—protein world from an all-RNA world. Such developments had to occur after the emergence of templated peptide synthesis,

¹Corresponding author: alfonzo.1@osu.edu.

but not too much later. Although early RNA-binding proteins may have been act alone entities, the necessity for modularity must have been imminent [1]. Modularity implies that discreet protein domains, whose sole purpose was to bind RNA, had to fuse with catalytic domains [2,3]; these must have been an indispensable invention for the evolution of RNA processing and certainly for the appearance of RNA modification specificity.

2. THE PROBLEM OF tRNA RECOGNITION

All tRNAs in cells adopt a very similar 3D-folded structure, the characteristic L-shape (Fig. 1), whereby the tRNA can be divided into two functional domains: the codon recognition domain, formed by the coaxial stacking of the D-arm onto the anticodon arm, and the amino acid-accepting domain, formed by the coaxial stacking of the T Ψ C arm on top of the acceptor arm. These two domains come together as an L-shaped molecule whose structure is stabilized by base-pair interactions between the D-loop, the variable loop, and the aforementioned T Ψ C loop. Critical to this folding is a stable “elbow structure,” which is formed by specific base pairs between several nucleotides: residues 15 of the D-loop and 48 of the variable loop, G₁₈ of the D-loop, and Ψ ₅₅ of the T Ψ C loop and finally an additional Watson-Crick base pair between G₁₉ of the D-loop and the highly conserved C₅₆ (of T Ψ C). It has become apparent that this conserved structure (the elbow) has been used repeatedly for recognition by various tRNA-targeting enzymes such as a number of aminoacyl tRNA synthetases, ribonuclease P, and some tRNA transport factors [4]. In terms of modification, this conserved structure sets a critical baseline for how modification enzymes approach the recognition problem. A sure inference is the fact that if Ψ ₅₅ is important for proper tRNA folding, the enzyme that catalyzes its formation cannot be dependent on the L-shape structure for recognition. Predictably such enzymes learn to target tRNAs before the invention of the L-shape and inherently must have acted on proto-tRNAs before the elbow structure was adopted. Similarly, methylation of G₉ plays a critical role in tRNA folding [5] and had to be introduced early in the evolution of extant canonical L-shaped tRNAs.

Once the tRNA is folded, then other modifications can be added and the conserved tertiary features of a tRNA can be exploited for recognition. For example, many of the modifications that target the anticodon loop demand such a degree of specificity that mistakes, such as modification of neighboring nucleotides that are not the intended target, could prove detrimental to a cell's well-being. In general, modification enzymes that target the anticodon loop required a fully folded tRNA for activity and it is from the information found in such a structured molecule that exquisite specificity is achieved. This is especially true of anticodon deaminases, where deamination of the wrong nucleotide can potentially reassign the tRNA to a different codon, leading to a miscoding problem. It has been suggested then that tRNA modification enzymes can be divided into two major groups, based on how they recognize the substrate they target: (1) architecture-dependent and (2) architecture-independent enzymes; the former require a fully folded L-shape tRNA for activity, and the latter act earlier and are in general a prerequisite for folding [6]. In turn, many of the architecture-independent enzymes act at sites away from the anticodon arm and play critical roles as structural modulators, while architecture-dependent ones tend to target the anticodon loop, the modifications that they synthesize affecting only local structure and anticodon function. In the following pages, we will highlight how these structural constraints affect the mode of

substrate recognition by tRNA-modifying enzymes. We will focus on a few well-characterized examples of how modification enzymes have solved the tRNA-binding problem to discriminate between nearly identical substrates, while only targeting the desired ones. To illustrate this concept, we will focus on three different groups of modifications enzymes: (1) tRNA deaminases, (2) tRNA methyltransferases, and (3) the tRNA-specific pseudouridine synthases.

3. DEAMINASES THAT TARGET tRNA

Historically, the modification field has been divided into two groups: the editors and the modifiers. The term *modification* implies any biological change to a polynucleotide chain that alters its chemical composition beyond that of the canonical nucleotides (G, A, U (T), and C); such changes are programmed rather than stochastic and different from chemical mutagenesis [7]. Thus, RNA editing can be considered a well-defined category within the vast umbrella term of posttranscriptional modifications. Editing *stricto sensu* involves a programmed change of one canonical nucleotide to another (e.g., C to U) or a change of a canonical nucleotide that alters the meaning of the edited RNA beyond what is encoded by the genome (e.g., A-to-I deamination) [7]. Editing may also involve more extensive insertion and deletion of nucleotides as those occurring in *Physarum* and trypanosomatid mitochondria, but this will not be further discussed here.

Among the tRNA deaminases, some edit the first position of the anticodon (position 34), while others target additional positions away from the anticodon, for example, A-to-I deamination at position 37 in eukaryotes, or position 57 or 58 in Archaea [8–10]. The best studied among these enzymes are the adenosine deaminases acting on tRNA, or ADAT, where ADAT2/3 deaminates adenosine 34 (A_{34}) to inosine 34 (I_{34}). Inosine expands the decoding capacity of tRNAs allowing a single tRNA to decode multiple codons with the same amino acid. Other ADATs, for example, ADAT1, form inosine at position 37 of the anticodon loop of a single tRNA ($tRNA^{Ala}$) [11]. While I_{37} does not affect decoding specificity, it is important for translational accuracy. In addition, C-to-U changes at other positions have been described in organellar tRNAs from various organisms [12,13], but the enzymes involved have not been identified. Mechanistically these are inferred to occur by hydrolytic deamination. In this section, we will concentrate on the A-to-I enzymes, whose activities have been reconstituted in vitro and mechanisms of tRNA substrate recognition extensively characterized.

Anticodon A-to-I deaminases have been described in Bacteria (ADATa, TadA) and Eukarya (ADAT2/3) [14–18]. The bacterial enzyme is homodimeric in solution and targets a single tRNA, $tRNA^{Arg}_{ACG}$, the only A_{34} -containing tRNA in bacteria (Fig. 2). The crystal structure of ADATa was reported by several laboratories, which showed the general architecture of the active site and confirmed its similarity with other members of the deaminase superfamily [19–22]. Biochemical analysis demonstrated that this enzyme could efficiently catalyze deamination of minimal substrates representing only the anticodon arm of $tRNA^{Arg}$; a negligible difference was observed in the catalytic efficiency of this enzyme with shorter substrates when compared to a full-length tRNA [19,20]. This observation led to the proposal that ADATa was architecture independent. This may have been the result of its

specialization toward a single substrate, which then led to the evolution of a very restrictive RNA-binding domain limited to residues in the proximity of the active site [21]. This proposal was confirmed by the cocrystal structure of ADATa bound to an anticodon arm mimic, where the nucleotide to be deaminated (A₃₄) was replaced by the adenosine analog nebularine [20]. This was necessary to prevent the enzyme from turning over the substrate and enabled trapping of the substrate-bound enzyme. This cocrystal revealed that upon complex formation, the RNA substrate undergoes significant conformational changes. The anticodon nucleotides became splayed out, making them accessible for recognition. The RNA-binding cleft is part of a portion of the dimerization domain and makes sequence-specific contacts with the five anticodon loop nucleotides, including a direct minor groove interaction with the side chain of a conserved Lys106, which simultaneously forms hydrogen bonds with C₃₂ and A₃₈. Additionally, a conserved asparagine (Asn138) hydrogen bonds with A₃₈, while a series of positively charged residues (arginines and lysines) make general contacts with backbone phosphates in the anticodon stem. The overall anticodon loop structure poised for catalysis is finally stabilized by specific contacts with G₃₆ and U₃₃ [20]. The take-home lesson is that indeed, ADATa has built in its active site all that is needed for the specific recognition of a single tRNA, while excluding near cognate substrates.

The homologous enzyme in eukaryotes is ADAT2/3; unfortunately, there is no structural information available for this enzyme, but biochemical data support their heterodimeric structure in solution. The heterodimer is formed by two paralogous subunits encoded by the ADAT2 and ADAT3 genes. Both subunits bear the conserved signature H...CxxxC (where H stand for histidine, C for cysteine, and x for any amino acid) Zn²⁺-coordinating triad characteristic of members of the cytidine deaminase superfamily, but the conserved proton-shuttling glutamate required for catalysis has been naturally mutated through evolution in the ADAT3 sub-unit, while it is still present in the TbADAT2 subunit. This led to the suggestion that the ADAT3 subunit solely plays a structural role, a proposal not supported by studies with the *Trypanosoma brucei* enzyme, where ADAT3 coordinates the catalytic Zn²⁺ and in turn partakes in catalysis. Unlike the bacterial ADATa, ADAT2/3 targets seven to eight different tRNAs depending on the organism. Studies with the *T. brucei* enzyme revealed the presence of critical tRNA-binding residues at the C-terminus of TbADAT2, a domain composed of a stretch of positively charged amino acids containing arginines and lysines (the KR domain) [23]. Deletion of the KR domain leads to lack of deamination activity presumably due to the fact that tRNA binding is abrogated. However, the active site still contributes to binding and it is the combination of active-site residues and of the distal KR domain that partly explains substrate specificity. Currently, the contribution of the second subunit (TbADAT3) to binding has not been explored. Regardless, it is clear that unlike bacterial ADATa, eukaryotic ADATs may possess a bipartite-binding domain where active-site residues and more distal domains contribute to substrate recognition.

The described dichotomy between a homodimeric deaminase with a single substrate and a heterodimeric enzyme with multiple substrates sets an interesting scenario that has been suggested for how these enzymes evolved from single-substrate to multisubstrate specificity. Given their sequence conservation, it has been proposed that ADAT3 was the result of gene duplication from ADAT2. This event provided a second copy of an otherwise essential gene, which can now freely mutate by genetic drift. This must have served as the preexisting

condition that allowed the further rearrangement of RNA-binding domains away from the active site, which presumably also led to active-site relaxation. Once active-site binding becomes less critical, by virtue that the binding strength is provided by a distal domain, the enzyme was probably able to accommodate a wider variety of substrates, leading to the appearance of the extant eukaryotic tRNA deaminases. In this realm, gene duplication a purely neutral evolutionary event was a critical step in the advent of multisubstrate deaminases.

4. THE DIVERSITY OF RNA METHYLATIONS, METHYLTRANSFERASES, AND THEIR MODES OF tRNA BINDING

Methylation is one of the most abundant modifications found ubiquitously in a variety of RNAs including tRNA, rRNA, mRNA, tmRNA, snRNA, snoRNA, miRNA, and viral RNA [24–26]. While methylated nucleotides are involved in affecting functional aspects of certain RNAs, they are particularly important in maintaining structural elements as methyl groups can stabilize or modulate the local RNA structure. In either case, introducing methylations precisely in the proper location and substrate is essential. Consequently, RNA methyltransferases acting on multiple substrates have evolved different strategies to ensure that the correct substrate species and individual target nucleotide(s) are selected from a pool of potential choices. This portion of the chapter will focus primarily on tRNA methyltransferases, but will also touch on the methylation of other RNAs.

The majority of methyltransferases use *S*-adenosyl-L-methionine (SAM or AdoMet) as a methyl donor. Methyltransferases that utilize 5,10-methylenetetrahydrofolate or carboxy-*S*-adenosyl methionine as a methyl donor have been identified, but will not be extensively covered here [27,28]. SAM methyltransferases have been categorized into at least five classes (I–V) based on their conserved structural folds [5,29–31]. Most methyltransferases are class I type, harboring a Rossmann fold, but RNA methyltransferases in particular are predominantly confined to class I and class IV types [5,29]. Despite their highly conserved structural fold, the amino acid sequence as well as the mode of substrate binding can differ, often significantly, between enzymes within the same class. Over time, some methyltransferases have gained distinct auxiliary RNA-binding domains, supporting a common theme of modular evolution whereby catalytic domains are combined with different substrate-binding domains to achieve substrate specificity [32–35].

Class I methyltransferases consist of alternating α -helices and parallel β -sheets characteristic of a typical Rossmann fold, apart from the presence of an additional (seventh) antiparallel β -strand [36]. Within this conserved fold, comparative sequence analyses have revealed a set of at least six conserved regions (motifs I–VI) encompassing the SAM-binding domain [29,36]. The first motif harbors a conserved nucleotide-binding site, GxGxG (x is any residue), which forms the first α -helix as it bends underneath SAM [29,30]. The other known conserved SAM-binding residue is an acidic residue at the end of β 2, which forms hydrogen bonds with both hydroxyls of the SAM ribose [30]. Recently, in the TlyA ribosomal RNA methyltransferase from *Mycobacterium tuberculosis* an additional novel SAM-binding motif, RxWV, was discovered [37], which affects the structure of a GxGxG

motif in a purposeful way [37], and both motifs are required for SAM binding [37]. Still, in general the SAM-binding region normally resides within the N-terminus of many methyltransferases and is assembled primarily from loops following strands 1, 2, and 3, whereas the substrate-binding region is generally located in the C-terminal portion of the β -sheet [29,38]. Beyond that, substrate binding for class I methyltransferases can vary greatly. There are several instances described in which a class I methyltransferase is fused to a known RNA-binding domain as discussed in more detail below [39,40].

Class IV methyltransferases are characterized by the SPOUT domain, which was first discovered by bioinformatics analyses based on the similarity to the SpoU (TrmH) and TrmD tRNA methyltransferases (Fig. 2) [41–46]. The SPOUT domain contains an unusual α/β fold distinguished by a deep trefoil knot formed by the folding of three β -sheets in the C-terminal portion of the SAM-binding motif [5,47,48]. Many SPOUT proteins have acquired additional N- and/or C-terminal extensions, which presumably participate in substrate binding [25,45,49]. Two SPOUT enzymes, TrmL (tRNA U_{m34}/C_{m34}) and RlmH (rRNA), are considered “minimalist” proteins where substrate binding is limited solely to the SPOUT domain and, in the case of TrmL, no other RNA-binding protein assists in binding [25,49–51]. TrmL acts as a homodimer to catalyze 2'-O-methylation at position 34 on two isoacceptors of tRNA^{Leu} (anticodons: CAA and UAA), and like many RNA-binding enzymes, it uses positively charged residues for substrate recognition [50,51]. Interestingly, only three other SPOUT tRNA methyltransferases are found in bacteria (TrmH, TrmD, and TrmJ); all methylate many more different substrates than TrmL [50]. Presumably, these enzymes acquired nucleic acid-binding domains as N- and C-terminal extensions that support recognition of multiple substrates, whereas the minimal SPOUT enzyme TrmL binds fewer substrates, making these extensions unnecessary. The nucleic acid-binding domains found within SPOUT as well as class I methyltransferases include PUA, THUMP, TRAM (OB-fold), and L30e domains (Fig. 3) [45].

The PUA domain (pseudouridine synthase and archeosine transglycosylase domain) is a highly conserved RNA-binding motif found in all domains of life. The PUA domain is made up of 64–96 amino acids and forms a compact pseudobarrel α/β fold (Fig. 3) [53–55]. In humans, NSUN6 is a class I methylase, which modifies m⁵C₇₂ of tRNA^{Cys} and tRNA^{Thr} and contains a PUA RNA-binding motif [53,56]. Interestingly, in vitro methylation by NSUN6 requires the 3'-CCA sequence, which is normally a prerequisite for aminoacylation [56]. A previous crystal structure of another PUA-containing enzyme, *Pyrococcus horikoshii* archaeosine tRNA-guanine transglycosylase, revealed that the PUA domain makes a direct contact with the 3'-CCA of tRNA, indicating a partially conserved mode of binding for enzymes containing this domain [56].

The putative THUMP RNA-binding domain (thiouridine synthase, methyltransferase, and pseudouridine synthase domain) was first discovered by in silico methods and thought to consist of 100–110 amino acids producing an α/β fold [57]. The crystal structure of the *Thermotoga maritima* Thil 4-thiouridine synthase, in complex with a truncated tRNA, confirmed the role of THUMP domains in RNA binding and revealed a potential molecular ruler mechanism to define substrate specificity [58]. Interestingly, like Thil, the RNA methyltransferases Trm11, RlmKL, Trm14/TrmN, and RlmN (a radical SAM enzyme) all

contain an N-terminal ferredoxin-like (NFLD) domain along with the THUMP domain, implying that THUMP and NFLD together could form a functional binding unit as is seen in Thil [58]. For example, Trm11 from Eukarya and Archaea is comprised of three domains: NFLD, THUMP, and a class I methyltransferase catalytic domain (Fig. 3).

The TRAM domain (Trm2 and MiaB domain) is conserved in Bacteria and Eukarya, but not widely found in Archaea [34]. Some enzymes have recruited the TRAM domain for RNA binding. This domain is predicted to form a simple β -barrel. In methyltransferases, TRAM domains have been identified at the N-terminus of Trm2-like tRNA methyltransferases, the N-terminus of RlmD (formerly RumA, m⁵U₁₉₃₉ *Escherichia coli* 23S rRNA), and the C-terminus of the 23S RNA-specific uracil O-methyltransferase, FtsJ, from *Halobacterium* (Fig. 3) [32,34]. The *EcRlmD* enzyme was the first to reveal the structure of the TRAM domain, showing a five-stranded antiparallel β -barrel fold with a Greek key topology [59]. This structure adopts an oligonucleotide binding (OB) fold, which is commonly found in DNA/RNA enzymes [59]. The putative SPOUT methyltransferase MT1 from *Methanobacterium thermoautotrophicum* also contains the OB-fold subdomain unusually located between the center of its β -fold architecture [60].

The L30e domain is an RNA-binding domain commonly found in ribosomal proteins (named for the ribosomal L30 protein) and forms a distinctive three-layer $\alpha/\beta/\alpha$ sandwich. There are three conserved residues within the L30 protein family, Gly26, Lys28, and Arg52, which are critical for RNA binding. The glycine residue, which lacks a bulky side chain, provides space for the RNA to bind, while the basic residues form hydrogen bonds with nucleotides in the RNA substrate [47]. The SPOUT 2'-O-methyltransferase, RlmB, is comprised of an N-terminal L30e domain extending from the catalytic subunit (Fig. 3) [60]. The RlmB family contains residues corresponding to the three RNA-binding residues in L30 and indeed RlmB has been shown to methylate Gm2251 of 23S rRNA in *E. coli* [60]. A putative SPOUT 2'-O-methyltransferase from *Thermus thermophilus* also appears to have an L30e domain at its N-terminus [61]. While the L30e domain is typically found in ribosomal proteins, methylation of 23S rRNA was not detected in intact ribosomes in *T. thermophilus*, and the putative methyltransferase displays no sequence similarity to L30; consequently, the target RNA is still unknown [61].

5. SEQUENCE vs STRUCTURAL DETERMINANTS FOR SPECIFICITY

There are two general strategies used by methyltransferases for identifying the correct substrate. They either survey structural features or recognize specific sequences within the RNA substrate. The majority of methyltransferases recognize local secondary structure or global tertiary folds, but some act based on a specific consensus sequence where recognition is usually in the context of local structural elements. RNA methyltransferases commonly detect structural features of RNA, as seen with the conserved consensus fold of m⁵U methyltransferases [62]. RlmD has three structural domains, an N-terminal TRAM domain, a central α/β domain, and a C-terminal class I methyltransferase domain [62]. RlmD strictly modifies 23S rRNA at a single position (U1939) to m⁵U and will not methylate it if the U position is mutated to C [59]. Early studies of another m⁵U methyltransferase, TrmA, suggested that it did not recognize a strict consensus sequence on tRNA and it could

efficiently modify short versions of the T-stem loop (TSL) [62]. However, a crystal structure of *EcTrmA* bound to the T-arm revealed a unique colinear base-stacking of G₅₃-A₅₈-G₅₇-C₅₆-U₅₅ in a conformation that is not normally seen in a typical unbound tRNA [62]. This presents the U₅₄-U₅₅-C₅₆ sequence (where the first U is the methylation target) into the active site [62]. This interaction has been proposed to be a conserved specificity determinant for RlmD and TrmA [62]. Outside of this consensus fold, other RNA-binding elements allow for discrimination between the different substrates (rRNA vs tRNA); for example, unlike RlmD, TrmA does not contain a TRAM domain [32,63].

As previously discussed for deaminases, tRNA methyltransferases are often categorized into two groups based on substrate recognition: those requiring the entire tRNA for activity and those that can act on shorter tRNA substrates. For example, the SPOUT Um₃₄/Cm₃₄ *EcTrmL* methyltransferase specifically recognizes the stem-loop structure of the anticodon stem-loop (ASL) (extended by two nucleotides). *EcTrmL* also recognizes the specific nucleotides A₃₆-A₃₇-A₃₈, as well as other determinants emphasized in the next section [50]. On the other hand, the *E. coli* TrmJ tRNA Xm₃₂ (X is any nucleotide) methyltransferase requires the full-length tRNA for activity [50]. Surprisingly, it detects specific regions within the D-loop rather than the anticodon loop where the nucleotide to be methylated is found [50]. In contrast to *EcTrmJ*, the archaeal TrmJ from *Sulfolobus acidocaldarius* (*SaTrmJ*) can act on small tRNA substrates, specifically tRNAs lacking the D- and T-loops [25]. Additionally, *SaTrmJ* can only methylate C₃₂, whereas bacterial *EcTrmJ* can modify all four nucleotides at position 32, although A and G are almost never found at that position [25]. Interestingly, the bacterial and archaeal TrmJ are SPOUT methyltransferases, whereas the eukaryotic tRNA Xm₃₂ enzyme (Trm7 with auxiliary protein Trm732) is a class I methyltransferase [25]. Likewise, the tRNA m¹G₃₇ methyltransferase, TrmD (SPOUT), does not require the L-shaped tertiary structure of tRNA, but instead recognizes the anticodon arm structure and also uses G₃₆-G₃₇ as a sequence determinant [50]. Unlike its bacterial counterpart, the eukaryotic m¹G₃₇ methyltransferase Trm5, an unrelated class I enzyme, surveys the entirety of the tRNA tertiary structure to identify the correct substrate. It is noteworthy that both TrmJ and Trm7 as well as TrmD and Trm5 are just a few examples of RNA modification enzymes that have undergone convergent evolution, whereas two unrelated enzymes (in this case SPOUT vs Rossmann-fold methyltransferases) have evolved independently yet perform the same catalytic reaction. This further highlights the functional significance of tRNA modifications in all domains of life.

6. PREEXISTING MODIFICATIONS AND THEIR REQUIREMENT FOR SPECIFICITY

Another aspect of specificity in RNA modification enzymes is the order of processing events and the role played by other modifications. Having a systematic sequence of processing events or addition of modifications can ensure the proper structure and function of RNA and allows for quality control. Of course, these events can be dictated by compartmentalization, perhaps best demonstrated by retrograde transport systems where some modifications require reimport into the nucleus.

Several methylations, in addition to other modifications, require the RNA to be processed before it can be introduced. This can be due to impediment via steric hindrance in the preprocessed substrate where processing can be necessary to form the proper sequence or structural recognition elements. An interesting case is the splicing of eukaryotic tRNA introns. The wybutosine (yW) biosynthesis pathway in eukaryotes requires the addition of m¹G₃₇ by Trm5 as the first of several steps [64,65]. In yeast, the tRNA must be spliced before the methylation reaction can take place; however, the splicing machinery is located on the cytoplasmic side of the outer mitochondrial membrane and it is in this context that intron removal takes place [48,65–69]. The Trm5 enzyme, however, is a nuclear enzyme [65]. Therefore, tRNA^{Phe} must be imported back into the nucleus after splicing in order to be methylated [65,66]. Afterward the tRNA is reexported to the cytoplasm to become further modified to yW. In contrast, some methylations are actually intron dependent. In yeast and humans, for example, the multisite m⁵C tRNA methyltransferase, Trm4, requires an intron for methylation at positions 34 and 40 [70,71]. Trm4 does not survey the overall L-shaped tRNA structure, but rather has been suggested to be dependent on specific nucleotide sequence and base pairs within the intron. However, Trm4-mediated m⁵C modification at positions 48 and 49 is unaffected by the presence or absence of an intron in yeast.

The majority of RNA methyltransferases modify canonical nucleotides, but they can also methylate more complex substrates like nucleotides with preexisting modifications. RlmH (SPOUT) from *E. coli* specifically methylates position 1915 of 23S rRNA exclusively if that position has a preexisting pseudouridine (Ψ) [31]. It is the only methylated pseudouridine found in bacterial RNA to date, although several have been found in Eukarya (m¹Ψ, Ψm, and m¹acp³Ψ) [31]. Modifications at positions outside of the target methylation site also influence methylation events. For example, *Ec*TrmL methylates pyrimidines but not purines at the wobble position and requires prior formation of i⁶A at position 37 [50]. The sequence A₃₆-A₃₇-A₃₈ is also important for *Ec*TrmL methylation, but it is unclear whether this sequence is necessary for recognition by the enzyme itself or because of the i⁶A₃₇ modification [50]. In any case, the addition of i⁶A₃₇ into an otherwise unmodified in vitro transcribed tRNA^{Leu} is sufficient to recruit TrmL to the tRNA [50]. A similar situation occurs in yeast where i⁶A₃₇ is a prerequisite for m³C₃₂ formation in tRNA^{ser} [72]. These results provide evidence for an anticodon loop circuitry where modifications within this loop are highly interrelated.

Methylation can also act as a requirement for subsequent modifications at the same position or elsewhere in the RNA. The methylation of m¹A is found universally at position 58 of tRNA and is additionally present at several other positions within the T-loop depending on the organism [73]. In Archaea, certain tRNAs contain m¹A at position 58 or 57 as a mandatory step in the biosynthesis of m¹I₅₇, produced by adenosine-to-inosine (A-to-I) deamination. The methylation is catalyzed by the tetrameric archaeal TrmI enzyme and is strictly required for deamination to occur [10,73]. The orthologous m¹A₅₈ site-specific methyltransferase in yeast also forms a tetramer, but is comprised of two different evolutionary-related subunits (further discussed in the next section). A similar, yet unconventional, two-step pathway was recently discovered in *T. brucei* involving the TbADAT2/3 deaminase discussed previously. In addition to its role in the conserved A-to-I deamination at position 34, TbADAT2/3 also converts C-to-U at position 32 in

trypanosomes [16,74], but it requires a methylated substrate, m^3C_{32} , and the corresponding methyltransferase, TbTrm140, for this activity [75,76]. Remarkably, the methyltransferase TbTrm140 also requires the presence of TbADAT2/3 to form the methylated substrate prior to deamination [75,76]. Therefore, these enzymes are mutually dependent on each other for function [75,76].

In an analogous scenario and adding further complexity to the yW biosynthesis in tRNA^{Phe} highlighted earlier, it has been suggested that one of the enzymes in the pathway, Tyw1, depends on prior formation of Gm₃₄ and Cm₃₂ by the Trm7 enzyme (along with its associated partner proteins) [72,77]. Therefore, the entire proposed pathway consists of export of the intron-containing tRNA^{Phe} precursor from the nucleus to the cytoplasm for splicing and 2'-O-methylation by Trm7, followed by import back into the nucleus for m^1G_{37} and lastly reexport into the cytoplasm for the final steps to produce yW₃₇ (Fig. 4) [77].

Besides acting as an obligatory intermediate for additional modifications, methylations can positively or negatively influence other modifications and vice versa. The modification m^7G_{37} positively impacts Gm₁₈ and m^1G_{37} , while Ψ_{55} negatively affects the levels of Gm₁₈, $m^5S^2U_{54}$, and m^1A_{58} in *T. thermophilus* [72]. In trypanosomes, it has been proposed that upon mitochondrial import, editing of tRNA^{Trp} relies on the interplay between several modifications; it is 2'-O-methylated at position 32, thiolated at position 33 to form S²U₃₃, and edited from C₃₄ to U₃₄, while the unedited C₃₄ is methylated to Cm₃₄. Interestingly, downregulation of the thiolation activity leads to almost complete editing at position 34. Therefore, methylation may serve as a positive determinant for C₃₄ to U₃₄ editing, while thiolation negatively affects the editing reaction helping maintain two functional forms of tRNA^{Trp}, edited and unedited [78].

7. MULTISUBUNIT METHYLTRANSFERASES

One noteworthy aspect of RNA methyltransferases is their tendency to recruit additional partner proteins, which often play a part in substrate binding, catalytic efficiency, stability, or localization. Furthermore, oligomerization of the same subunit can also vary between methylation enzymes from different domains of life or even within the same domain affecting substrate specificity or catalysis. It has been proposed that the formation of homodimers could act as a possible step in the evolution of heterodimeric enzymes whereby (as discussed previously) a gene duplication event may arise and lead to functional specialization of each subunit over time [36,79]. In general, SPOUT methyltransferases form homodimers where each monomer partakes in substrate binding and the active site is present at the dimer interface. However, the SPOUT enzyme, Trm10, acts as a monomeric enzyme, suggesting that catalysis and substrate recognition for this enzyme may differ from canonical SPOUT methyltransferases [5,80,81]. Unlike the SPOUT methyltransferases, the class I methyltransferases are normally found as monomers, although there are few exceptions such as *Bacillus subtilis* TrmB, which will be discussed later in this section.

Some enzymes recruit additional proteins to mediate substrate binding. The yeast 2'-O-methyltransferase Trm7 interacts with an associated protein Trm732 to methylate C₃₂ and with another protein, Trm734 for methylation of X₃₄ on tRNA^{Phe}, tRNA^{Leu}, and tRNA^{Trp}

[77]. In this case, Trm7 is the catalytic subunit, which uses either Trm732 or Trm734 to direct substrate specificity. The adaptor protein Trm112 provides another example in which a methyltransferase requires an unrelated protein factor for activity. Trm112 interacts with multiple tRNA methyltransferases including Trm9 and Trm11 to form mcm⁵U₃₄ and m²G₁₀, respectively, in eukaryotic tRNAs [82]. Initial studies in yeast showed that deletion of Trm11 displayed a minimal growth defect, whereas Trm112 deletion resulted in a severe growth phenotype, indicating its importance in other cellular functions [82]. In fact, Trm112 associates with several different factors involved in other cellular activities; for instance, it provides structural stability for another methyltransferase, Bud32, which methylates G1575 of 18S rRNA in eukaryotes [83–85]. Therefore, Trm112 performs multiple functional roles in tRNA modification as well as other activities in the cell.

In yeast, the m¹A₅₈ is a common but nonuniversal modification as it appears in 21 of 32 sequenced tRNAs [86]. This methylation is catalyzed by a heterotetrameric enzyme composed of the subunits Trm6 (Gcd10p) and Trm61 (Gcd14p); mutation of the putative SAM-binding residues of Trm61 abolished activity, while in vitro neither Trm61 nor Trm6 could by themselves bind tRNA [86]. This suggested that Trm61 was responsible for catalysis, whereas Trm6 aided in tight binding to the tRNA substrate in the context of the other subunit. Mutational studies later demonstrated that residues in both Trm61 and Trm6 subunits participate in tRNA binding [87]. As mentioned earlier the orthologous bacterial/archaeal methyltransferase TrmI forms a tetramer, which may be conserved for its role in recognition, binding, or activity. However, the binding interface appears to diverge between prokaryotic TrmI and eukaryotic Trm6/Trm61, indicating that these enzymes evolved different means of substrate recognition and binding [87].

In eukaryotes, Trm8 and Trm82 form a complex which modifies position 46 of seven different tRNAs with m⁷G [88]. While Trm8 contains the putative catalytic site and has tRNA-binding ability, Trm82 does not engage in either catalysis or tRNA binding [79]. Instead, the formation of the Trm8-Trm82 complex causes Trm8 to undergo a considerable conformational change forming the correct architecture for catalysis [5]. Although both Trm8/Trm82 enzymes are conserved among Eukarya, Bacteria utilize a single subunit enzyme, TrmB, a homolog of Trm8. Interestingly, there are differences even among bacterial TrmB, where *E. coli* TrmB is a monomeric enzyme, whereas *B. subtilis* forms homodimers [36]. The dimerization of *Bs*TrmB is also unusual for a class I methyltransferase as its structural elements deviate from the expected Rossmann fold [36]. This finding supports the evolution of obligatory dimers observed for Trm8/Trm82 as well as a variety of other dimeric methyltransferases like those listed earlier.

Rather than enlisting other proteins, some methyltransferases, lacking intrinsic substrate specificity, utilize guide RNAs to direct catalysis. These enzymes form complexes with small C/D box guide RNAs that direct methylation to the target nucleotide. This mechanism is seen primarily in eukaryotic rRNA and small nuclear RNA as well as archaeal 2'-O-methylation of rRNA and in some cases tRNA [89]. The need for RNA-dependent methylation of tRNA is not conserved as 2'-O-methyltransferase of positions 32 and 34 of yeast tRNA^{Trp} and of position 18 from Bacteria (TrmH) and Eukarya (Trm3) does not

involve guide RNAs [73]. Methylation via guide RNAs is yet to be seen in Bacteria, suggesting that this system likely evolved after the Bacteria split from Archaea and Eukarya.

8. PSEUDOURIDINE SYNTHASES

Pseudouridine (Ψ) is the most common modified nucleotide found in RNAs from all three kingdoms of life [90–93]. In 1951, as the first modified nucleosides were discovered in RNA hydrolysates, pseudouridine was identified and became known as the fifth nucleoside, 5-ribosyl uracil [94–97]. Pseudouridine is a C5-glycoside isomer of uridine and the only C-C bond joining a nucleobase to a sugar known in nucleic acids [95,97]. Formation of pseudouridine is catalyzed by pseudouridine synthases (Pus or Ψ synthases), a class of RNA-modifying enzymes responsible for the posttranscriptional site-specific isomerization of uridine to pseudouridine in a wide array of RNA substrates including rRNA, mRNA, snoRNA, snRNA, 5S RNA, and tRNA [98–103]. Notably, these enzymes are unable to isomerize free nucleotides. Upon isomerization, there is no detectable release of free uridine and the reaction occurs without the need for ATP or any other cofactor [104,105]. Although a general isomerization mechanism is established, the specific chemical mechanism by which these enzymes catalyze the reaction has been a challenging problem [105–110]. A clear outcome of pseudouridylation is the formation of an extra hydrogen bond donor at the newly formed N1 position, imparting structural stability to the target RNA, which has in some cases proven important for translational efficiency and accuracy [111,112].

9. CLASSIFICATION OF RNA PSEUDOURIDINE SYNTHASES AND THE CONSERVED CATALYTIC CORE DOMAIN

The first tRNA Ψ synthases were identified from *Salmonella typhimurium* and *E. coli* and are the product of the *truA* gene [98,99,113]. Thereafter, the Ψ synthases have been classified into families based on limited sequence similarities and named using the bacterial nomenclature. The founding members of each family: TruA, TruB, RluA, RsuA, TruD are also present in Archaea and Eukarya (Fig. 1) [100,114,115]. The sixth and newest family, not found in Bacteria but found in Archaea and Eukarya, includes the human Pus10 (pseudouridine synthase) [116–118]. Thus, the original nomenclature was based on the target substrate. For instance, TruA and TruB modify tRNA, while RluA and RsuA act on large and small subunit ribosomal RNA, respectively [104,119]. However, there are unique exceptions where a pseudouridine synthase has multisubstrate RNA specificity; the RluA family of Ψ synthases, for example, catalyzes Ψ_{746} in the large ribosomal 23S rRNA and Ψ_{32} in tRNA^{Phe} [120]. Pseudouridine formation is catalyzed either by these independent proteins or by H/ACA snoRNA-dependent ribonucleoprotein particles [101,103,121–124]. The eukaryotic synthases (Pus), originally discovered in *Saccharomyces cerevisiae*, have been named numerically based on the order of their discovery, where Pus4p and Pus7p are the orthologs of *E. coli* TruB and TruD enzymes [125–129], respectively.

Crystal structures have been determined for proteins belonging to each of the six families and their structural diversity underscores the significant role that protein dynamics must play in RNA substrate recognition [104,119,130]. The underlying theme is that despite low level of primary structural similarity among Ψ synthases, all adopt exquisitely superimposable

Author Manuscript

folds with three-dimensional structural features that share a common catalytic core and a cleft comprised of a conserved set of active-site residues [104,110,114,119,130] (Fig. 5). In addition to a conserved catalytic aspartate, studies of the first four families (TruA, TruB, RsuA, and RluA) of Ψ synthases have revealed a second important residue: Tyr, in *E. coli* and archaeal Pus10; Asp in hPus10; and Glu in TruD [115,117,131–135]. This semiconserved residue is proposed to aid in the formation of a buried salt bridge along with the universal catalytic aspartate [104,130]. Moreover, the tyrosine residue may also help pack the substrate via the uracil base. In the fifth family of Ψ synthases, which includes TruD and *S. cerevisiae* Pus7, the tyrosine is not conserved, but mutational and enzymatic analyses indicate that a strictly conserved glutamate unique to this family likely participates as a general base [132].

Author Manuscript

Author Manuscript

The domain structure of Ψ synthases has been classified into three major conserved motifs I, II, and III [35,114]. Conserved among all Ψ synthases, the conserved catalytic core is found in Motif II and features an absolutely conserved aspartate residue. Having minimal sequence similarity to other synthases, TruA and its homologs appear to be the most distantly related [104]. The RluA and RsuA family are most closely related to each other as they share three conserved sequence motifs (motifs I, II, and III). Only motifs I and II were found in the TruB family. The structure of TruB revealed that motif II is part of the active site and that motif I plays an architectural role or reinforcing motif II, in agreement with site-directed mutagenesis studies of motif I residues [136–138]. Moreover, although TruB lacks motif III, the sequences of the TruB family of enzymes can be aligned with those from the RluA and RsuA families [139]. The TruD family, specific for pseudouridylation of position 13 in tRNA, has motifs I, II, and III and is remarkable because the crystal structure determination of TruD revealed a two-domain structure. TruD has a catalytic domain that differs in sequence, but is structurally very similar to the catalytic domain of other pseudouridine synthases and a second large domain with a novel α/β fold not found in any other protein [139–141]. Interestingly, all TruD homologs are present in all phyla and are all circularly permuted, meaning that they have spatially conserved structural elements with other synthases, but located at the opposite ends of the enzyme [142]. In other words, structural features found at the C-terminus of other synthases are at the N-terminus in TruD [142]. Therefore, in a parsimonious evolutionary scenario, it is proposed that TruD was the first to diverge from all other synthases [104,140].

Author Manuscript

The mechanism by which the active-site residues manage to disconnect, rotate, and reconnect the uracil base to the C5 carbon is still unclear [105–107,109,110]. One possibility is a nucleophilic attack by the aspartate on the sugar or the substrate uracil [105,107,110]. The chemical mechanism for this reaction would minimally involve three steps: (1) cleavage of the N-glycosidic linkage of the target uridine, (2) rotation of the uracil ring to position C5 of the pyrimidine near C1' of the ribosyl group of the RNA, and (3) formation of a new C1'–C5 carbon-carbon bond [105,107,110]. Traditional analysis of pseudouridine formation uses a tritium release assay; however, the quench-flow technique has been adapted to further study pseudouridine formation [143]. The results from studying three *E. coli* enzymes showed that single-turnover rate constants for pseudouridylation by TruB, RluA, and TruA were determined to be 0.5, 0.7, and 0.35 s⁻¹, respectively [143]. Thus, the reaction chemistry may be the rate-limiting step, since these rate constants were reported to be similar to the

k_{cat} tested under multiple-turnover conditions [143]. Although both the single- and multiple-turnover rate constants have been measured, the detailed mechanistic steps involved in catalysis still remain unsolved [119].

10. RNA BINDING: CONTRIBUTIONS FROM PERIPHERAL DOMAINS

The core Ψ synthase fold is embellished by divergent peripheral structural elements in some of the Ψ synthase enzymes [104,105]. These elements often provide additional independently folded domains that may confer unique substrate specificity. For example, RsuA, which catalyzes pseudouridine formation at position 516 in 16S rRNA important for ribosome assembly, possesses an appended domain at its N-terminus with similarity to the RNA-binding domain of ribosomal protein S4 (Fig. 5) [115,144,145]. Despite the conservation with S4, the N-terminal domain of RsuA and S4 binds rRNA at different locations [115,144]. A total of six Arg and two Lys residues on the surface of the N-terminal domain contribute the set of positively charged residues that bind rRNA in both proteins. Thus, it is the flexible linker that connects the central catalytic domain to this N-terminus that may provide the required flexibility to adjust the catalytic domain relative to a specific substrate [144]. The *Haemophilus influenza* RsuA also contains an N-terminal S4-like $\alpha 3/\beta 4$ domain followed by a catalytic domain [146]. Whereas the individual domains of *E. coli* and *H. influenza* RsuA are structurally similar, the relative spatial placement among the two structures differs greatly [146]. The *E. coli* structure displays an extended open conformation with no direct contacts between the domains, while the *H. influenza* RsuA is in a closed conformation with a large interface between the two domains [146]. Domain closure involves several basic and polar residues forming a putative RNA-binding cleft. It is proposed that this relative repositioning of the S4 and catalytic domains is used to modulate the shape and size of the rRNA-binding site in pseudouridine synthases possessing S4 domains (Fig. 5) [146].

Unique to the Pus10 family of synthases is the presence of an accessory N-terminal THUMP RNA-binding domain (Fig. 3), also found in tRNA methyltransferases, as described earlier, and in other RNA modification enzymes [57,117]. The N-terminal THUMP domain of Pus10 contributes to tRNA binding by a proposed induced-fit mechanism for pseudouridine formation. In archaeal Pus10a, in addition to the aforementioned catalytic aspartate, a forefinger loop and an Arg and a Tyr residue are essential for Ψ_{54} , but not for Ψ_{55} formation [118]. Since the forefinger loop is needed for both rRNA and tRNA Ψ synthase activities of RluA, but only for tRNA Ψ_{54} activity of Pus10, archaeal Pus10 proteins are proposed to use two distinct mechanisms for substrate uridine recognition, binding, and catalysis [134].

hPus10 is a crescent-shaped molecule with two domains, including the universally conserved Ψ synthase catalytic domain [117]. Superposition of the catalytic domains of Pus10 and other Ψ synthases shows that the full set of conserved Ψ synthase active-site residues are present in Pus10, indicating that it likely employs a catalytic mechanism similar to other Ψ synthases [116]. The Pus10 active site is located in a deep pocket of a basic cleft adjacent to flexible thumb and forefinger loops, which could provide further RNA substrate stabilization during binding [117]. Modeling studies support the role of the N-terminal domain as an accessory RNA-binding domain, with a cleft between the catalytic and

accessory domains large enough and electrostatically compatible to accommodate an RNA [117]. Homology models of archaeal Pus10 proteins based on the crystal structure of human Pus10 have revealed that subtle structural differences between all Pus10 proteins [118]. While most archaeal Pus10 produce both tRNA Ψ_{54} and Ψ_{55} , some produce only Ψ_{55} . The archaeal Pus10 synthase enzymes from *Methanocaldococcus jannaschii* and *Pyrococcus furiosus* can isomerize tRNA at two positions Ψ_{54} and Ψ_{55} with the two modifications occurring independently [134]. Furthermore, unlike bacterial TruB and yeast Pus4, archaeal Pus10 does not require a U54:A58 reverse-Hoogsteen base pair and a pyrimidine at position 56 [118]. Within the Pus10 family, another interesting observation is that the Pus10 gene is essential in *Haloferax volcanii*, whereas the orthologous tRNA Ψ_{55} synthase genes are nonessential in yeast (Pus4) and *E. coli* (TruB) [118,133,143].

Within the TruB family of Ψ synthases, *E. coli* TruB and yeast Pus4 are functional as monomers and specifically targets U₅₅ in the TSL of almost all tRNAs. For recognition and binding to its RNA substrate, the *E. coli* TruB PUA RNA-binding domain is present at the C-terminus, which differs from the S4 RNA-binding domain of RsuA found at its N-terminus. Although the RNA-binding domains of both proteins are present at antipodal termini, upon the superimposition of the central catalytic domain of RsuA and TruB structures, both PUA and S4 RNA-binding domains were located on the same structural side of the respective proteins [104,137,144].

As discussed earlier, Ψ synthases share common three-dimensional structural features and generally exist as monomeric proteins [104,105,147,148]. The one exception is the TruA family whose members are dimeric in solution with two positively charged RNA-binding clefts along its surface; each cleft contains the highly conserved aspartic acid at its center [122,148,149]. TruA utilizes its dimerization interface to form the key domain critical for binding via the conserved tRNA elbow. Furthermore, TruA is distinct from other Ψ synthases in its substrate specificity and modifies multiple tRNAs (for example, 17 tRNAs in *E. coli*) that contain uridines at positions 38, 39, and 40 [122,148,149]. The TruA enzymes are also more compact, in that they lack the extra peripheral domains at both the N-terminal and C-terminal regions. Since both the *T. thermophilus* and *E. coli* TruA enzymes consist of only the catalytic core domain in the monomeric structure, it is proposed that the other monomer in the active dimer of TruA probably plays a role similar to that of the PUA domain of TruB or the S4 domain of RsuA [122,148,149]. On the other hand, PusI (in the TruA family) modifies positions 38, 39, and 40 in the anticodon stem of tRNA but acts as a monomer [107,150–152]. The PusI monomer harbors a 10-stranded β -sheet formed by the interaction of the N-terminus and C-terminus domains. In the crystal structure of PusI, an unexpected topological relationship was revealed between both the N-terminal and C-terminal domains [59,107]. In the β_1 and β_2 strands of the $\beta\alpha\beta\beta\alpha\beta$ core topology of PusI, RNA binding is mediated by a classical RNA recognition motif, a well-characterized domain found in several RNA-binding proteins [39,153,154].

11. MULTISUBSTRATE SPECIFICITY OF PSEUDOURIDINE SYNTHASES

TruB is a single site-specific enzyme targeting position 55 in all *E. coli* tRNAs with the exception of the initiator tRNA^{Met} [112,155,156]. Therefore, this enzyme faces a different

challenge in discriminating against a single nonsubstrate tRNA^{iMet} in a sea of true substrates. This enzyme recognizes the reverse-Hoogsteen base pair between T₅₄ and A₅₈ found in a majority of tRNAs; surprisingly, this base pair is absent in initiator tRNA^{Met} [136,137]. In the cocrystal structure of TruB with tRNA, the enzyme specifically stacks a conserved histidine residue on the loop side of this pair, while disrupting stacking of the T-loop nucleotides in the folded tRNA [136,137]. The crystal structures of *E. coli* TruB bound to a 22-nucleotide TSL mini-helix portion of tRNA and an intact tRNA with 5-fluorouridine replacing U55 also revealed that this position is flipped out and bound to the enzymes' catalytic cleft [136,137]. In addition, three T-loop nucleotides were also flipped out of the helical stack, thereby disrupting the tertiary base-pairing interaction between the D- and T-loops and affecting the global fold of the tRNA. TruB uses its "thumb" to bind the major groove of the tRNA substrate [136,137]. Upon RNA binding, the thumb was observed to undergo structural reorganization [136,137,157–159].

The global tRNA structure is not a recognition element for the TruB Ψ synthase, as mutations in the tRNA that disrupt the T- and D-loop interaction have no effect on pseudouridylation [159]. However, sequence determinants for pseudouridylation at position 55 reside around the T-loop, whereby TSL mini-helices are efficient pseudouridylation substrates. Two substrate-free structures of TruB unveiled unusual protein flexibility involving hinge-bending movement in the central β-sheet of the catalytic module. Hinge bending is proposed to act as a clamp for positioning the substrate [159]. Moreover, the β-sheet-bending movement affects the shape the catalytic cleft and creates disordering of the active site. Based on the significant conformational change observed between the structures of RNA-bound and -free forms, a proposed model for tRNA recognition includes a combination of rigid docking followed by induced-fit binding [104,159].

Human Pus1 (hPus1), a member of the TruA family, pseudouridylates specific uridine residues in several noncoding RNAs, including the U2 spliceosomal RNA and steroid receptor activator RNA in addition to tRNAs [107,160,161]. Since it was not possible to obtain crystals of the full-length hPusI protein, a minimal system using a truncated protein composed of the complete catalytic core and of a minimal tRNA substrate showed comparable pseudouridylation activity to that seen with full-length tRNA [107,160,161]. This result supports the fact that the disordered N- and C-terminal extensions, which were removed in the truncated protein, are not essential for enzymatic activity, as was predicted a priori by the RONN server in silico [107,160,161]. The hPus1 enzyme has a fold similar to that found in bacterial Ψ synthases, with a central domain that contains the catalytic Asp146. It adopts a mixed α/β fold with an extended central antiparallel four-stranded N-terminal β-sheet that packs against a six-stranded C-terminal β-sheet decorated by helices and loops as found in other Ψ synthases [107,160,161].

Remarkably, it was the analyses of three separate crystal structures of hPus1 that facilitated the identification of conformational differences observed in flexible regions of the protein, which have been proposed to be important for substrate recognition by induced fit. A flexible hinge at the base of the sheet allows the enzyme to open and close around an electropositive active-site cleft [161]. Although hPusI has no N- or C-terminal peripheral domains, it has a positively charged electrostatic surface extending from the N-terminus of

the catalytic domain to the active site of the enzyme [161]. The N-terminal domain has been suggested as the RNA-binding site. Unique to hPusI are two long α -helices inserted at the C-terminus of the core domain [161]. These two C-terminal α -helices are proposed to specifically permit only a specific binding orientation of tRNA, which differs from that of TruA [104,161]. The structures of the catalytic core domain of hPusI also revealed that the C-terminus forms the walls of the RNA-binding surface by extending along the back and top of the central β -sheet of the catalytic core domain [104,161]. The walls of the cleft are formed by two structures: (1) the forefinger loop formed by N-terminal domain loops and (2) the thumb loop at the C-terminal domain loops. Docking of tRNA to hPusI in a productive orientation requires only minor conformational changes in the enzyme and tRNA [161]. The docked tRNA is bound by the electropositive surface of the protein—employing a completely different binding mode than that seen for the *E. coli* TruA homolog in complex with tRNA [161].

Among the six identified RNA Ψ synthase families, the RluA family of RNA Ψ synthases appears to be the most complex based on their difference in RNA substrate specificity [120,162]. The notable feature of the RluA enzyme is in its dual specificity for Ψ_{32} in several tRNAs as well as Ψ_{746} in the 23S rRNA [120,162]. RluA can pseudouridylylate shorter RNA substrates, including a tRNA ASL and also a stem-loop RNA that contains sequences flanking the target uridine at position 746 from 23S rRNA [163,164]. A number of nonrelated RNA-binding proteins use a “direct readout” to recognize RNA in a sequence-specific manner [165], reminiscent of what was discussed with bacterial ADATa in the deaminase section of this chapter. Given that the sequence $\Psi UXXAAA$ (X is any nucleotide) is found in all RluA substrates, it was expected that RluA Ψ synthase would also use a similar strategy for substrate recognition [120,155,162]. However, structural analyses revealed that RluA induces the reorganization of a 21-nucleotide ASL of tRNA^{Phe} [162]. The drastic structural change forces the anticodon loop to form a reverse-Hoogsteen base pair with three flipped-out nucleotides, instead of the canonical U-turn conformation [162]. Presumably, RluA indirectly reads out the RNA structure by probing RNA loops for their ability to adopt the reorganized fold [162]. This was further supported by results from structure-guided mutagenesis, sequence conservation, and cocrystal structure [162]. Upon association with RNA, the solvent-accessible surface of the free RluA enzyme becomes buried. Sharing the same structural context as the TruB thumb, the RluA thumb differs only in having a more compact thumb loop structure to interact with its substrate’s major groove [162]. It is proposed that the ASL RNA is positioned with its axis parallel to the β -sheet core of RluA [162].

The TruA active site only has a small region of positively charged residues, while PusI also has a larger positive binding surface from two extended α -helices dedicated to binding the substrate RNA. This would concur with the role of *E. coli* TruA in modifying ASLs from many tRNAs, while the hPusI has more substrate promiscuity in having a wider variety of RNAs as substrates. From the differences in their approach toward achieving substrate specificity, TruA and TruB are proposed to have evolved independently from their common ancestor [104,162]. TruB, having single-site specificity, uses accessory domains along with induced-fit conformational changes to actively sample the substrate sequence near the target site, while TruA utilizes dimerization to provide a scaffold for recognition of the conserved

elbow structure of tRNA, while maintaining multisite selection of its substrate by retaining flexibility in the ASL [148,158]. This has suggested a molecular ruler mechanism as a strategy for how TruA recognizes its target [148,149]. TruA binds the ASL through the initial docking of the tRNA elbow and the D stem, which differs from the TruB enzyme, which directly docks the TSL into the active site [104,149,158,162]. Consequently, TruB induces the thumb domain to clasp onto nucleotides near the target site and inserts His43 into the emptied space left behind by flipping out the base, whereas the double-stranded stem is always kept away from the active site [104,162]. In contrast to the extensive interactions between TruB and the neighboring region of the target site, TruA has more limited contact with the ASL near the target site and relies more on secondary structure recognition of the substrate to achieve U₅₅ site specificity [104,158,162].

Interestingly, upon juxtaposition of the RNA-bound structure RluA with that of TruB, the cores of the two enzymes superimpose well, yet the position of the RNA substrate in their active site differs by as much as a 60-degree angle [104,162]. Presumably, the difference may be accounted for by the presence of the peripheral PUA domain in RluA, and lack of a forefinger loop in TruB, in contrast to RluA [104,162]. Thus, the forefinger loop in combination with an appended binding domain can drastically change how the enzyme binds the substrate [104,162]. Furthermore, structural comparison of RluA with the distantly related TruA shows a strikingly similar conformation including both forefinger and thumb loops despite little sequence similarity [104,162]. Juxtaposition of the RNA-bound RluA with the structure-free RsuA shows that RsuA possesses a forefinger structure, but not a thumb, as opposed to TruB which lacks a forefinger loop, but does have a thumb loop [104,162]. The use of such structural protrusions (forefinger and thumb loop) for recognition had already been proposed in the study of RluC and RluD, which together with RluB and RluE are part of the bacterial RluA family of synthases [104,105,112,166]. However, members of the bacterial RluA family have different RNA specificities in comparison to their yeast counterparts. Four members of the RluA family are present in both *E. coli* (RluA, RluC, RluD, and TruC) and *S. cerevisiae* (Pus5p, Pus6p, Rib2/Pus8p, and Pus9p) [104,105,112,166]. Three of the four *E. coli* enzymes modify 23S ribosomal RNA: RluC, RluD, and RluA (which also modifies position 32 in tRNAs) [105,112]. In notable contrast, only one of the four *S. cerevisiae* enzymes, Pus5p, modifies mitochondrial 21S rRNA [129]. In conclusion, despite a degree of domain conservation between this group of enzymes, their variability in domain arrangement makes it difficult to determine their intended targets a priori.

12. CONCLUSION

RNA modification enzymes act on a great diversity of substrates, and despite much research in recent years, still little is known about their mechanism for substrate recognition and specifically as to what their RNA-binding domains look like. In the particular case of tRNA, the matured substrate is highly structured and a general theme is that some modifications act early contributing to tRNA folding; others act after the near-native structure has formed and are thus limited to recognizing salient features of the 3D L-shape tRNA. Therefore, although all fully folded tRNAs are very similar in the 3D structure, modification enzymes as a family survey a much larger structural landscape depending on how early or late they act in the

folding pathway. In this chapter, we have attempted to illustrate well-known examples of how tRNA-modifying enzymes solve the binding problem. We have also highlighted the exquisite specificity that is needed to only target a subset of substrates in a vast pool of very similar nonsubstrates. A general principle that comes from this is that through evolution RNA-binding domains have been appended to the ends of modification enzymes and in doing so have provided a degree of active-site relaxation. By contrast, another family of tRNA-interacting enzymes, the aminoacyl tRNA synthetases, use a more limited set of features for tRNA recognition. For instance, studies using in vitro amino-acylation revealed the acceptor stem as a key identity element. Studies with portions of tRNA^{Ala} and yeast AlaRS showed that the first three base pairs of the acceptor stem were important for recognition [167–169]. In addition, the anticodon loop often bears identity elements for the selection of the correct tRNA by the cognate synthetase, and as such, it directly serves as a link between the protein synthesis and the genetic code [167,169,170]. Thus, in general, synthetases primarily recognize features of the anticodon loops and acceptor stems of tRNAs [171–179]. In the case of modification enzymes, to modify tRNAs at different stages of folding, active-site flexibility was necessary to accommodate a diversity of substrates. For instance, with deaminases, as discussed earlier, it may have also allowed a given enzyme to perform more than one reaction, as seen with the *T. brucei* ADAT2/3 enzyme, which can catalyze both A-to-I and C-to-U deamination. Likewise, it may have been that the modularity imparted by appended binding domains also provided the driving force for the evolution of multisubstrate specificity. Given the bewildering number of modified substrates, the modification field is indeed a fertile ground to further explore the limits of RNA binding and the basis for substrate specificity by this large family of enzymes.

ACKNOWLEDGMENTS

We thank NIH Grant GM084065 to J.D.A. and a Graduate Fellowship from The Ohio State University Center for RNA Biology to K.M. for support.

REFERENCES

- [1]. Lunde BM, Moore C, Varani G, RNA-binding proteins: modular design for efficient function, *Nat. Rev. Mol. Cell Biol* 8 (6) (2007) 479–490. [PubMed: 17473849]
- [2]. Blanc V, Navaratnam N, Henderson JO, Anant S, Kennedy S, Jarmuz A, Scott J, Davidson NO, Identification of GRY-RBP as an apolipoprotein B RNA-binding protein that interacts with both apobec-1 and apobec-1 complementation factor to modulate C to U editing, *J. Biol. Chem* 276 (13) (2001) 10272–10283. [PubMed: 11134005]
- [3]. Mehta A, Kinter MT, Sherman NE, Driscoll DM, Molecular cloning of apobec-1 complementation factor, a novel RNA-binding protein involved in the editing of apolipoprotein B mRNA, *Mol. Cell. Biol* 20 (5) (2000) 1846–1854. [PubMed: 10669759]
- [4]. Zhang J, Ferré-D'Amaré AR, The tRNA elbow in structure, recognition and evolution, *Life (Basel, Switzerland)* 6 (1) (2016) 3.
- [5]. Swinehart WE, Jackman JE, Diversity in mechanism and function of tRNA methyltransferases, *RNA Biol.* 12 (4) (2015) 398–411. [PubMed: 25626150]
- [6]. Grosjean H, Edqvist J, Stråby KB, Giegé R, Enzymatic formation of modified nucleosides in tRNA: dependence on tRNA architecture, *J. Mol. Biol* 255 (1) (1996) 67–85. [PubMed: 8568876]
- [7]. Gray MW, Diversity and evolution of mitochondrial RNA editing systems, *IUBMB Life* 55 (4–5) (2003) 227–233. [PubMed: 12880203]

- [8]. Paris Z, Fleming IMC, Alfonzo JD, Determinants of tRNA editing and modification: avoiding conundrums, affecting function, *Semin. Cell Dev. Biol* 23 (3) (2012) 269–274. [PubMed: 22024020]
- [9]. Randau L, Stanley BJ, Kohlway A, Mechta S, Xiong Y, Söll D, A cytidine deaminase edits C to U in transfer RNAs in Archaea, *Science* 324 (5927) (2009) 657–659. [PubMed: 19407206]
- [10]. Grosjean H, Constantinesco F, Foiret D, Benachenhou N, A novel enzymatic pathway leading to 1-methylinosine modification in *Haloferax volcanii* tRNA, *Nucleic Acids Res.* 23 (21) (1995) 4312–4319. [PubMed: 7501451]
- [11]. Gerber A, Grosjean H, Melcher T, Keller W, Tad1p, a yeast tRNA-specific adenosine deaminase, is related to the mammalian pre-mRNA editing enzymes ADAR1 and ADAR2, *EMBO J.* 17 (16) (1998) 4780–4789. [PubMed: 9707437]
- [12]. Janke A, Pääbo S, Editing of a tRNA anticodon in marsupial mitochondria changes its codon recognition, *Nucleic Acids Res.* 21 (7) (1993) 1523–1525. [PubMed: 8479901]
- [13]. Alfonzo JD, Blanc V, Estévez AM, Rubio MAT, Simpson L, C to U editing of the anticodon of imported mitochondrial tRNA(Trp) allows decoding of the UGA stop codon in *Leishmania tarentolae*, *EMBO J.* 18 (24) (1999) 7056–7062. [PubMed: 10601027]
- [14]. Wolf J, Gerber AP, Keller W, TadA, an essential tRNA-specific adenosine deaminase from *Escherichia coli*, *EMBO J.* 21 (14) (2002) 3841–3851. [PubMed: 12110595]
- [15]. Gerber AP, Keller W, An adenosine deaminase that generates inosine at the wobble position of tRNAs, *Science* 286 (5442) (1999) 1146–1149. [PubMed: 10550050]
- [16]. Rubio MAT, Pastar I, Gaston KW, Ragone FL, Janzen CJ, Cross GAM, Papavasiliou FN, Alfonzo JD, An adenosine-to-inosine tRNA-editing enzyme that can perform C-to-U deamination of DNA, *Proc. Natl. Acad. Sci. U.S.A* 104 (19) (2007) 7821–7826. [PubMed: 17483465]
- [17]. Torres AG, Piñeyro D, Rodríguez-Escribà M, Camacho N, Reina O, Saint-Léger A, Filonava L, Batlle E, Ribas De Pouplana L, Inosine modifications in human tRNAs are incorporated at the precursor tRNA level, *Nucleic Acids Res.* 43 (10) (2015) 5145–5157. [PubMed: 25916855]
- [18]. Zhou W, Karcher D, Bock R, Identification of enzymes for adenosine-to-inosine editing and discovery of cytidine-to-uridine editing in nucleus-encoded transfer RNAs of *Arabidopsis*, *Plant Physiol.* 166 (4) (2014) 1985–1997. [PubMed: 25315605]
- [19]. Kim J, Malashkevich V, Roday S, Lisbin M, Schramm VL, Almo SC, Structural and kinetic characterization of *Escherichia coli* TadA, the wobble-specific tRNA deaminase, *Biochemistry* 45 (20) (2006) 6407–6416. [PubMed: 16700551]
- [20]. Losey HC, Ruthenburg AJ, Verdine GL, Crystal structure of *Staphylococcus aureus* tRNA adenosine deaminase TadA in complex with RNA, *Nat. Struct. Mol. Biol* 13 (2) (2006) 153–159. [PubMed: 16415880]
- [21]. Elias Y, Huang RH, Biochemical and structural studies of A-to-I editing by tRNA: A34 deaminases at the wobble position of transfer RNA, *Biochemistry* 44 (36) (2005) 12057–12065. [PubMed: 16142903]
- [22]. Kuratani M, Ishii R, Bessho Y, Fukunaga R, Sengoku T, Shirouzu M, Sekine S, Yokoyama S, Crystal structure of tRNA adenosine deaminase (TadA) from *Aquifex aeolicus*, *J. Biol. Chem* 280 (16) (2005) 16002–16008. [PubMed: 15677468]
- [23]. Ragone FL, Spears JL, Wohlgamuth-Benedum JM, Kreel N, Papavasiliou FN, Alfonzo JD, The C-terminal end of the *Trypanosoma brucei* editing deaminase plays a critical role in tRNA binding, *RNA* 17 (7) (2011) 1296–1306. [PubMed: 21602302]
- [24]. Motorin Y, Helm M, RNA nucleotide methylation, *Wiley Interdiscip. Rev. RNA* 2 (5) (2011) 611–631. [PubMed: 21823225]
- [25]. Liu RJ, Long T, Zhou M, Zhou XL, Wang ED, tRNA recognition by a bacterial tRNA Xm32 modification enzyme from the SPOUT methyltransferase superfamily, *Nucleic Acids Res.* 43 (15) (2015) 7489–7503. [PubMed: 26202969]
- [26]. Hori H, Methylated nucleosides in tRNA and tRNA methyltransferases, *Front. Genet* 5 (5) (2014) 144. [PubMed: 24904644]
- [27]. Yamagami R, Yamashita K, Nishimasu H, Tomikawa C, Ochi A, Iwashita C, Hirata A, Ishitani R, Nureki O, Hori H, The tRNA recognition mechanism of folate/FAD-dependent tRNA methyltransferase (TrmFO), *J. Biol. Chem* 287 (51) (2012) 42480–42494. [PubMed: 23095745]

- [28]. Kim J, Xiao H, Bonanno JB, Kalyanaraman C, Brown S, Tang X, Al-Obaidi NF, Patskovsky Y, Babbitt PC, Jacobson MP, et al., Structure-guided discovery of the metabolite carboxy-SAM that modulates tRNA function, *Nature* 498 (7452) (2013) 123–126. [PubMed: 23676670]
- [29]. Kozbial PZ, Mushegian AR, Natural history of S-adenosylmethionine-binding proteins, *BMC Struct. Biol* 5 (2005) 19. [PubMed: 16225687]
- [30]. Schubert HL, Blumenthal RM, Cheng X, Many paths to methyltransfer: a chronicle of convergence, *Trends Biochem. Sci* 28 (6) (2003) 329–335. [PubMed: 12826405]
- [31]. Ero R, Peil L, Liiv A, Remme J, Identification of pseudouridine methyltransferase in *Escherichia coli*, *RNA* 14 (10) (2008) 2223–2233. [PubMed: 18755836]
- [32]. Urbonavicius J, Jäger G, Björk GR, Amino acid residues of the *Escherichia coli* tRNA(m⁵U54)methyltransferase (TrmA) critical for stability, covalent binding of tRNA and enzymatic activity, *Nucleic Acids Res.* 35 (10) (2007) 3297–3305. [PubMed: 17459887]
- [33]. McKenney KM, Alfonzo JD, From prebiotics to probiotics: the evolution and functions of tRNA modifications, *Life (Basel, Switzerland)* 6 (1) (2016) 13.
- [34]. Anantharaman V, Koonin EV, Aravind L, TRAM, a predicted RNA-binding domain, common to tRNA uracil methylation and adenine thiolation enzymes, *FEMS Microbiol. Lett* 197 (2) (2001) 215–221. [PubMed: 11313137]
- [35]. Aravind L, Koonin EV, Novel predicted RNA-binding domains associated with the translation machinery, *J. Mol. Evol* 48 (3) (1999) 291–302. [PubMed: 10093218]
- [36]. Zegers I, Gigot D, van Vliet F, Tricot C, Aymerich S, Bujnicki JM, Kosinski J, Droogmans L, Crystal structure of *Bacillus subtilis* TrmB, the tRNA (m⁷G46) methyltransferase, *Nucleic Acids Res.* 34 (6) (2006) 1925–1934. [PubMed: 16600901]
- [37]. Witek MA, Kuiper EG, Minten E, Crispell EK, Conn GL, A novel motif for S-adenosyl-L-methionine binding by the ribosomal RNA methyltransferase TlyA from *Mycobacterium tuberculosis*, *J. Biol. Chem* 292 (5) (2017) 1977–1987. [PubMed: 28031456]
- [38]. Martin JL, McMillan FM, SAM (dependent) I AM: the S-adenosylmethionine-dependent methyltransferase fold, *Curr. Opin. Struct. Biol* 12 (6) (2002) 783–793. [PubMed: 12504684]
- [39]. Pastore C, Topalidou I, Forouhar F, Yan AC, Levy M, Hunt JF, Crystal structure and RNA binding properties of the RNA recognition motif (RRM) and AlkB domains in human AlkB homolog 8 (ABH8), an enzyme catalyzing tRNA hypermodification, *J. Biol. Chem* 287 (3) (2012) 2130–2143. [PubMed: 22065580]
- [40]. Towns WL, Begley TJ, Transfer RNA methyltransferases and their corresponding modifications in budding yeast and humans: activities, predilections, and potential roles in human health, *DNA Cell Biol.* 31 (4) (2012) 434–454. [PubMed: 22191691]
- [41]. Anantharaman V, Koonin EV, Aravind L, 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 (1) (2002) 71–75. [PubMed: 11763972]
- [42]. Ahn HJ, Kim H-W, Yoon H-J, Lee II B, Suh SW, Yang JK, Crystal structure of tRNA(m¹G37)methyltransferase: insights into tRNA recognition, *EMBO J.* 22 (11) (2003) 2593–2603. [PubMed: 12773376]
- [43]. Lim K, Zhang H, Tempczyk A, Krajewski W, Bonander N, Toedt J, Howard A, Eisenstein E, Herzberg O, Structure of the YibK methyltransferase from *Haemophilus influenzae* (HI0766): a cofactor bound at a site formed by a knot, *Proteins* 51 (1) (2003) 56–67. [PubMed: 12596263]
- [44]. Elkins PA, Watts JM, Zalacain M, Van Thiel A, Vitazka PR, Redlak M, Andraos-Selim C, Rastinejad F, Holmes WM, Insights into catalysis by a knotted TrmD tRNA methyltransferase, *J. Mol. Biol* 333 (5) (2003) 931–949. [PubMed: 14583191]
- [45]. Tkaczuk K, Dunin-Horkawicz S, Purta E, Bujnicki JM, Structural and evolutionary bioinformatics of the SPOUT superfamily of methyltransferases, *BMC Bioinformatics* 8 (2007) 73. [PubMed: 17338813]
- [46]. Nureki O, Watanabe K, Fukai S, Ishii R, Endo Y, Hori H, Yokoyama S, Deep knot structure for construction of active site and cofactor binding site of tRNA modification enzyme, *Structure* 12 (4) (2004) 593–602. [PubMed: 15062082]

- [47]. Michel G, Sauvé V, Larocque R, Li Y, Matte A, Cygler M, The structure of the RlmB 23S rRNA methyltransferase reveals a new methyltransferase fold with a unique knot, *Structure* 10 (10) (2002) 1303–1315. [PubMed: 12377117]
- [48]. Christian T, Hou Y-M, Distinct determinants of tRNA recognition by the TrmD and Trm5 methyltransferases, *J. Mol. Biol.* 373 (3) (2007) 623–632. [PubMed: 17868690]
- [49]. Van Laer B, Roovers M, Wauters L, Kasprzak JM, Dyzma M, Deyaert E, Singh RK, Feller A, Bujnicki JM, Droogmans L, et al., Structural and functional insights into tRNA binding and adenosine N1-methylation by an archaeal Trm10 homologue, *Nucleic Acids Res.* 44 (2) (2016) 940–953. [PubMed: 26673726]
- [50]. Zhou M, Long T, Fang Z-P, Zhou X-L, Liu R-J, Wang E-D, Identification of determinants for tRNA substrate recognition by *Escherichia coli* C/U34 2'-O-methyltransferase, *RNA Biol.* 12 (8) (2015) 900–911. [PubMed: 26106808]
- [51]. Liu RJ, Zhou M, Fang ZP, Wang M, Zhou XL, Wang ED, The tRNA recognition mechanism of the minimalist SPOUT methyltransferase TrmL, *Nucleic Acids Res.* 41 (16) (2013) 7828–7842. [PubMed: 23804755]
- [52]. Humphrey W, Dalke A, Schulten K, VMD: visual molecular dynamics, *J. Mol. Graph* 14 (1) (1996) 33–38. [PubMed: 8744570]
- [53]. Cerrudo CS, Ghiringhelli PD, Gomez DE, Protein universe containing a PUA RNA-binding domain, *FEBS J.* 281 (1) (2014) 74–87. [PubMed: 24393395]
- [54]. Pérez-Arellano I, Gallego J, Cervera J, The PUA domain—a structural and functional overview, *FEBS J.* 274 (19) (2007) 4972–4984. [PubMed: 17803682]
- [55]. Larsen LHG, Rasmussen A, Giessing AMB, Jøgl G, Kirpekar F, Identification and characterization of the *Thermus thermophilus* 5-methylcytidine (m5C) methyltransferase modifying 23 S ribosomal RNA (rRNA) base C1942, *J. Biol. Chem* 287 (33) (2012) 27593–27600. [PubMed: 22711535]
- [56]. Haag S, Warda AS, Kretschmer J, Günnigmann MA, Höbartner C, Bohnsack MT, NSUN6 is a human RNA methyltransferase that catalyzes formation of m5C72 in specific tRNAs, *RNA* 21 (9) (2015) 1532–1543. [PubMed: 26160102]
- [57]. Aravind L, Koonin EV, THUMP—a predicted RNA-binding domain shared by 4-thiouridine, pseudouridine synthases and RNA methylases, *Trends Biochem. Sci* 26 (4) (2001) 215–217. [PubMed: 11295541]
- [58]. Neumann P, Lakomek K, Naumann PT, Erwin WM, Lauhon CT, Ficner R, Crystal structure of a 4-thiouridine synthetase-RNA complex reveals specificity of tRNA U8 modification, *Nucleic Acids Res.* 42 (10) (2014) 6673–6685. [PubMed: 24705700]
- [59]. Lee TT, Agarwalla S, Stroud RM, Crystal structure of RumA, an iron-sulfur cluster containing *E. coli* ribosomal RNA 5-methyluridine methyltransferase, *Structure* 12 (3) (2004) 397–407. [PubMed: 15016356]
- [60]. Leulliot N, Bohnsack MT, Graille M, Tollervey D, Van Tilbeurgh H, The yeast ribosome synthesis factor Emg1 is a novel member of the superfamily of alpha/beta knot fold methyltransferases, *Nucleic Acids Res.* 36 (2) (2008) 629–639. [PubMed: 18063569]
- [61]. Nureki O, Shirouzu M, Hashimoto K, Ishitani R, Terada T, Tamakoshi M, Oshima T, Chijimatsu M, Takio K, Vassylyev DG, et al., An enzyme with a deep trefoil knot for the active-site architecture, *Acta Crystallogr. D Biol. Crystallogr* 58 (7) (2002) 1129–1137. [PubMed: 12077432]
- [62]. Alian A, Lee TT, Griner SL, Stroud RM, Finer-Moore J, Structure of a TrmA-RNA complex: a consensus RNA fold contributes to substrate selectivity and catalysis in m5U methyltransferases, *Proc. Natl. Acad. Sci. U.S.A* 105 (19) (2008) 6876–6881. [PubMed: 18451029]
- [63]. Walbott H, Leulliot N, Grosjean H, Golinelli-Pimpaneau B, The crystal structure of *Pyrococcus abyssi* tRNA (uracil-54, C5)-methyltransferase provides insights into its tRNA specificity, *Nucleic Acids Res.* 36 (15) (2008) 4929–4940. [PubMed: 18653523]
- [64]. Sample PJ, Ko ený L, Paris Z, Gaston KW, Rubio MAT, Fleming IMC, Hinger S, Horáková E, Limbach PA, Lukeš J, et al., A common tRNA modification at an unusual location: the discovery of wyosine biosynthesis in mitochondria, *Nucleic Acids Res.* 43 (8) (2015) 4262–4273. [PubMed: 25845597]

- [65]. Ohira T, Suzuki T, Retrograde nuclear import of tRNA precursors is required for modified base biogenesis in yeast, *Proc. Natl. Acad. Sci. U.S.A* 108 (26) (2011) 10502–10507. [PubMed: 21670254]
- [66]. Hopper AK, Transfer RNA post-transcriptional processing, turnover, and subcellular dynamics in the yeast *Saccharomyces cerevisiae*, *Genetics* 194 (1) (2013) 43–67. [PubMed: 23633143]
- [67]. Huh W-K, Falvo JV, Gerke LC, Carroll AS, Howson RW, Weissman JS, O'Shea EK, Global analysis of protein localization in budding yeast, *Nature* 425 (6959) (2003) 686–691. [PubMed: 14562095]
- [68]. Yoshihisa T, Yunoki-Esaki K, Ohshima C, Tanaka N, Endo T, Possibility of cytoplasmic pre-tRNA splicing: the yeast tRNA splicing endonuclease mainly localizes on the mitochondria, *Mol. Biol. Cell* 14 (8) (2003) 3266–3279. [PubMed: 12925762]
- [69]. Goto-Ito S, Ito T, Kuratani M, Bessho Y, Yokoyama S, Tertiary structure checkpoint at anticodon loop modification in tRNA functional maturation, *Nat. Struct. Mol. Biol* 16 (10) (2009) 1109–1115. [PubMed: 19749755]
- [70]. Motorin Y, Grosjean H, Multisite-specific tRNA: m5C-methyltransferase (Trm4) in yeast *Saccharomyces cerevisiae*: identification of the gene and substrate specificity of the enzyme, *RNA* 5 (8) (1999) 1105–1118. [PubMed: 10445884]
- [71]. Brzezicha B, Schmidt M, Makalowska I, Jarmolowski A, Piekowska J, Szwejkowska-Kuliska Z, Identification of human tRNA: m5C methyltransferase catalysing intron-dependent m5C formation in the first position of the anticodon of the pre-tRNA(CAA)Leu, *Nucleic Acids Res.* 34 (20) (2006) 6034–6043. [PubMed: 17071714]
- [72]. Arimbasseri AG, Iben J, Wei F, Rijal K, Tomizawa K, Hafner M, Maraia RJ, Evolving specificity of tRNA 3-methyl-cytidine-32 (m3C32) modification: a subset of tRNAs^{Ser} requires N6-isopentenylolation of A37, *RNA* 22 (9) (2016) 1400–1410. [PubMed: 27354703]
- [73]. Roovers M, A primordial RNA modification enzyme: the case of tRNA (m1A) methyltransferase, *Nucleic Acids Res.* 32 (2) (2004) 465–476. [PubMed: 14739239]
- [74]. Rubio MAT, Ragone FL, Gaston KW, Ibba M, Alfonzo JD, C to U editing stimulates A to I editing in the anticodon loop of a cytoplasmic threonyl tRNA in *Trypanosoma brucei*, *J. Biol. Chem* 281 (1) (2006) 115–120. [PubMed: 16269406]
- [75]. Rubio MAT, Gaston KW, McKenney KM, Fleming IMC, Paris Z, Limbach PA, Alfonzo JD, Editing and methylation at a single site by functionally interdependent activities, *Nature* 542 (7642) (2017) 494–497. [PubMed: 28230119]
- [76]. Yewdell WT, Chaudhuri J, Molecular biology: RNA editing packs a one-two punch, *Nature* 542 (7642) (2017) 420–421. [PubMed: 28230135]
- [77]. Guy MP, Podyma BM, Preston MA, Shaheen HH, Krivos KL, Limbach PA, Hopper AK, Phizicky EM, Yeast Trm7 interacts with distinct proteins for critical modifications of the tRNA^{Phe} anticodon loop, *RNA* 18 (10) (2012) 1921–1933. [PubMed: 22912484]
- [78]. Rubio M, Alfonzo JD, Editing and modification in trypanosomatids: the reshaping of non-coding RNAs, *Top. Curr. Genet* 12 (2005) 71–86.
- [79]. Guy MP, Phizicky EM, Two-subunit enzymes involved in eukaryotic post-transcriptional tRNA modification, *RNA Biol.* 11 (12) (2014) 1608–1618. [PubMed: 25625329]
- [80]. Shao Z, Yan W, Peng J, Zuo X, Zou Y, Li F, Gong D, Ma R, Wu J, Shi Y, et al., Crystal structure of tRNA m1G9 methyltransferase Trm10: insight into the catalytic mechanism and recognition of tRNA substrate, *Nucleic Acids Res.* 42 (1) (2014) 509–525. [PubMed: 24081582]
- [81]. Kurowski MA, Sasin JM, Feder M, Debski J, Bujnicki JM, Characterization of the cofactor-binding site in the SPOUT-fold methyltransferases by computational docking of S-adenosylmethionine to three crystal structures, *BMC Bioinformatics* 4 (2) (2003) 9. [PubMed: 12689347]
- [82]. Purushothaman SK, Bujnicki JM, Grosjean H, Lapeyre B, Trm11p and Trm112p are both required for the formation of 2-methylguanosine at position 10 in yeast tRNA, *Mol. Cell. Biol* 25 (11) (2005) 4359–4370. [PubMed: 15899842]
- [83]. Sardana R, Johnson AW, The methyltransferase adaptor protein Trm112 is involved in biogenesis of both ribosomal subunits, *Mol. Biol. Cell* 23 (21) (2012) 4313–4322. [PubMed: 22956767]

- [84]. Figaro S, Wacheul L, Schillewaert S, Graille M, Huvelle E, Mongeard R, Zorbas C, Lafontaine DLJ, Heurgue-Hamard V, Trm112 is required for Bud23-mediated methylation of the 18S rRNA at position G1575, *Mol. Cell. Biol* 32 (12) (2012) 2254–2267. [PubMed: 22493060]
- [85]. Ltoquart J, Huvelle E, Wacheul L, Bourgeois G, Zorbas C, Graille M, Heurg-Hamard V, Lafontaine DLJ, Structural and functional studies of Bud23-Trm112 reveal 18S rRNA N7-G1575 methylation occurs on late 40S precursor ribosomes, *Proc. Natl. Acad. Sci. U.S.A* 111 (51) (2014) e5518–e5526. [PubMed: 25489090]
- [86]. Anderson J, Phan L, Hinnebusch AG, The Gcd10p/Gcd14p complex is the essential two-subunit tRNA(1-methyladenosine) methyltransferase of *Saccharomyces cerevisiae*, *Proc. Natl. Acad. Sci. U.S.A* 97 (10) (2000) 5173–5178. [PubMed: 10779558]
- [87]. Ozanick SG, Bujnicki JM, Sem DS, Anderson JT, Conserved amino acids in each subunit of the heteroligomeric tRNA m1A58 Mtase from *Saccharomyces cerevisiae* contribute to tRNA binding, *Nucleic Acids Res.* 35 (20) (2007) 6808–6819. [PubMed: 17932071]
- [88]. Alexandrov A, Martzen MR, Phizicky EM, Two proteins that form a complex are required for 7-methylguanosine modification of yeast tRNA, *RNA* 8 (10) (2002) 1253–1266. [PubMed: 12403464]
- [89]. Somme J, Van Laer B, Roovers M, Steyaert J, Verses W, Droogmans L, Characterization of two homologous 2'-O-methyltransferases showing different specificities for their tRNA substrates, *RNA* 20 (8) (2014) 1257–1271. [PubMed: 24951554]
- [90]. Dunin-Horkawicz S, Czerwoniec A, Gajda MJ, Feder M, Grosjean H, Bujnicki JM, MODOMICS: a database of RNA modification pathways, *Nucleic Acids Res.* 34 (Database issue) (2006) d145–d149. [PubMed: 16381833]
- [91]. Cantara WA, Crain PF, Rozenski J, McCloskey JA, Harris KA, Zhang X, Vendeix FAP, Fabris D, Agris PF, The RNA modification database, RNAMDB: 2011 update, *Nucleic Acids Res.* 39 (Database issue) (2011) d195–d201. [PubMed: 21071406]
- [92]. Machnicka MA, Milanowska K, Osman OO, Purta E, Kurkowska M, Olchowik A, Januszewski W, Kalinowski S, Dunin-Horkawicz S, Rother KM, et al., MODOMICS: a database of RNA modification pathways—2013 update, *Nucleic Acids Res.* 41 (Database issue) (2013) d262–d267. [PubMed: 23118484]
- [93]. Grosjean H, Benne R, Modification and Editing of RNA, American Society of Microbiology, Washington, DC, 1998.
- [94]. Cohn WE, Volkin E, Nucleoside-5'-phosphates from ribonucleic acid, *Nature* 167 (4247) (1951) 483–484.
- [95]. Davis FF, Allen FW, Ribonucleic acids from yeast which contain a fifth nucleotide, *J. Biol. Chem* 227 (2) (1957) 907–915. [PubMed: 13463012]
- [96]. Yu CT, Allen FW, Studies on an isomer of uridine isolated from ribonucleic acids, *Biochim. Biophys. Acta* 32 (1959) 393–406. [PubMed: 13846687]
- [97]. Cohn WE, 5-Ribosyl uracil, a carbon-carbon ribofuranosyl nucleoside in ribonucleic acids, *Biochim. Biophys. Acta* 32 (1959) 569–571. [PubMed: 13811055]
- [98]. Cortese R, Kammen HO, Spengler SJ, Ames BN, Biosynthesis of pseudouridine in transfer ribonucleic acid, *J. Biol. Chem* 249 (4) (1974) 1103–1108. [PubMed: 4592259]
- [99]. Ciampi MS, Arena F, Cortese R, Biosynthesis of pseudouridine in the in vitro transcribed tRNA^{Tyr} precursor, *FEBS Lett.* 77 (1) (1977) 75–82. [PubMed: 323061]
- [100]. Ofengand J, Ribosomal RNA pseudouridines and pseudouridine synthases, *FEBS Lett.* 514 (1) (2002) 17–25. [PubMed: 11904174]
- [101]. Ge J, Yu Y-T, RNA pseudouridylation: new insights into an old modification, *Trends Biochem. Sci* 38 (4) (2013) 210–218. [PubMed: 23391857]
- [102]. Pomerantz SC, McCloskey JA, Detection of the common RNA nucleoside pseudouridine in mixtures of oligonucleotides by mass spectrometry, *Anal. Chem* 77 (15) (2005) 4687–4697. [PubMed: 16053277]
- [103]. Yu Y-T, Meier UT, RNA-guided isomerization of uridine to pseudouridine— pseudouridylation, *RNA Biol.* 11 (12) (2014) 1483–1494. [PubMed: 25590339]
- [104]. Hamma T, Ferr-D' Amar AR, Pseudouridine synthases, *Chem. Biol* 13 (11) (2006) 1125–1135. [PubMed: 17113994]

- [105]. Spenkuch F, Motorin Y, Helm M, Pseudouridine: still mysterious, but never a fake (uridine)! *RNA Biol.* 11 (12) (2015) 1540–1554.
- [106]. McDonald MK, Miracco EJ, Chen J, Xie Y, Mueller EG, The handling of the mechanistic probe 5-fluorouridine by the pseudouridine synthase TruA and its consistency with the handling of the same probe by the pseudouridine synthases TruB and RluA, *Biochemistry* 50 (3) (2011) 426–436. [PubMed: 21142053]
- [107]. Czudnochowski N, Ashley GW, Santi DV, Alian A, Finer-Moore J, Stroud RM, The mechanism of pseudouridine synthases from a covalent complex with RNA, and alternate specificity for U2605 versus U2604 between close homologs, *Nucleic Acids Res.* 42 (3) (2014) 2037–2048. [PubMed: 24214967]
- [108]. Silva RG, Schramm VL, Uridine phosphorylase from *Trypanosoma cruzi*: kinetic and chemical mechanisms, *Biochemistry* 50 (42) (2011) 9158–9166. [PubMed: 21932786]
- [109]. Veerareddygarri GR, Singh SK, Mueller EG, The pseudouridine synthases proceed through a glycol intermediate, *J. Am. Chem. Soc.* 138 (25) (2016) 7852–7855. [PubMed: 27292228]
- [110]. Mueller EG, Chips off the old block, *Nat. Struct. Biol.* 9 (5) (2002) 320–322. [PubMed: 11976723]
- [111]. Ofengand J, Malhotra A, Remme J, Gutsell NS, Del Campo M, Jean-Charles S, Peil L, Kaya Y, Pseudouridines and pseudouridine synthases of the ribosome, *Cold Spring Harb. Symp. Quant. Biol.* 66 (2001) 147–159. [PubMed: 12762017]
- [112]. Gutsell NS, Deutscher MP, Ofengand J, The pseudouridine synthase RluD is required for normal ribosome assembly and function in *Escherichia coli*, *RNA* 11 (7) (2005) 1141–1152. [PubMed: 15928344]
- [113]. Kammen HO, Marvel CC, Hardy L, Penhoet EE, Purification, structure, and properties of *Escherichia coli* tRNA pseudouridine synthase I, *J. Biol. Chem.* 263 (5) (1988) 2255–2263. [PubMed: 3276686]
- [114]. Koonin EV, Pseudouridine synthases: four families of enzymes containing a putative uridine-binding motif also conserved in dUTPases and dCTP deaminases, *Nucleic Acids Res.* 24 (12) (1996) 2411–2415. [PubMed: 8710514]
- [115]. Conrad J, Niu L, Rudd K, Lane BG, Ofengand J, 16S ribosomal RNA pseudouridine synthase RsuA of *Escherichia coli*: deletion, mutation of the conserved Asp102 residue, and sequence comparison among all other pseudouridine synthases, *RNA* 5 (6) (1999) 751–763. [PubMed: 10376875]
- [116]. Roovers M, Hale C, Tricot C, Terns MP, Terns RM, Grosjean H, Droogmans L, Formation of the conserved pseudouridine at position 55 in archaeal tRNA, *Nucleic Acids Res.* 34 (15) (2006) 4293–4301. [PubMed: 16920741]
- [117]. McCleverty CJ, Hornsby M, Spraggon G, Kreusch A, Crystal structure of human Pus10, a novel pseudouridine synthase, *J. Mol. Biol.* 373 (5) (2007) 1243–1254. [PubMed: 17900615]
- [118]. Gurha P, Gupta R, Archaeal Pus10 proteins can produce both pseudouridine 54 and 55 in tRNA, *RNA* 14 (12) (2008) 2521–2527. [PubMed: 18952823]
- [119]. Spenkuch F, Motorin Y, Helm M, Pseudouridine: still mysterious, but never a fake (uridine)! *RNA Biol.* 11 (12) (2014) 1540–1554. [PubMed: 25616362]
- [120]. Raychaudhuri S, Niu L, Conrad J, Lane BG, Ofengand J, Functional effect of deletion and mutation of the *Escherichia coli* ribosomal RNA and tRNA pseudouridine synthase RluA, *J. Biol. Chem.* 274 (27) (1999) 18880–18886. [PubMed: 10383384]
- [121]. Hamma T, Ferré-D'Amaré AR, The box H/ACA ribonucleoprotein complex: interplay of RNA and protein structures in post-transcriptional RNA modification, *J. Biol. Chem.* 285 (2) (2010) 805–809. [PubMed: 19917616]
- [122]. Hur S, Stroud RM, Finer-Moore J, Substrate recognition by RNA 5-methyluridine methyltransferases and pseudouridine synthases: a structural perspective, *J. Biol. Chem.* 281 (51) (2006) 38969–38973. [PubMed: 17085441]
- [123]. Kiss T, Fayet-Lebaron E, Jány BE, Box H/ACA small ribonucleoproteins, *Mol. Cell* 37 (5) (2010) 597–606. [PubMed: 20227365]

- [124]. Decatur WA, Liang X, Piekna-Przybylska D, Fournier MJ, Identifying effects of snoRNA-guided modifications on the synthesis and function of the yeast ribosome, *Methods Enzymol.* 425 (2007) 283–316. [PubMed: 17673089]
- [125]. Samuelsson T, Olsson M, Transfer RNA pseudouridine synthases in *Saccharomyces cerevisiae*, *J. Biol. Chem* 265 (15) (1990) 8782–8787. [PubMed: 2187870]
- [126]. Ma X, Zhao X, Yu Y-T, Pseudouridylation (Psi) of U2 snRNA in *S. cerevisiae* is catalyzed by an RNA-independent mechanism, *EMBO J.* 22 (8) (2003) 1889–1897. [PubMed: 12682021]
- [127]. Becker HF, Motorin Y, Planta RJ, Grosjean H, The yeast gene YNL292w encodes a pseudouridine synthase (Pus4) catalyzing the formation of psi55 in both mitochondrial and cytoplasmic tRNAs, *Nucleic Acids Res.* 25 (22) (1997) 4493–4499. [PubMed: 9358157]
- [128]. Hellmuth K, Grosjean H, Motorin Y, Deinert K, Hurt E, Simos G, Cloning and characterization of the *Schizosaccharomyces pombe* tRNA:pseudouridine synthase Pus1p, *Nucleic Acids Res.* 28 (23) (2000) 4604–4610. [PubMed: 11095668]
- [129]. Urban A, Behm-Ansmant I, Branlant C, Motorin Y, RNA sequence and two-dimensional structure features required for efficient substrate modification by the *Saccharomyces cerevisiae* RNA:{Psi}-synthase Pus7p, *J. Biol. Chem* 284 (9) (2009) 5845–5858. [PubMed: 19114708]
- [130]. Ferré-D'Amaré AR, RNA-modifying enzymes, *Curr. Opin. Struct. Biol* 13 (1) (2003) 49–55. [PubMed: 12581659]
- [131]. Phannachet K, Elias Y, Huang RH, Dissecting the roles of a strictly conserved tyrosine in substrate recognition and catalysis by pseudouridine 55 synthase, *Biochemistry* 44 (47) (2005) 15488–15494. [PubMed: 16300397]
- [132]. Chan CM, Huang RH, Enzymatic characterization and mutational studies of TruD—the fifth family of pseudouridine synthases, *Arch. Biochem. Biophys* 489 (1–2) (2009) 15–19. [PubMed: 19664587]
- [133]. Chatterjee K, Blaby IK, Thiaville PC, Majumder M, Grosjean H, Yuan YA, Gupta R, de Crécy-Lagard V, The archaeal COG1901/DUF358 SPOUT-methyltransferase members, together with pseudouridine synthase Pus10, catalyze the formation of 1-methylpseudouridine at position 54 of tRNA, *RNA* 18 (3) (2012) 421–433. [PubMed: 22274953]
- [134]. Joardar A, Jana S, Fitzek E, Gurha P, Majumder M, Chatterjee K, Geisler M, Gupta R, Role of forefinger and thumb loops in production of Ψ54 and Ψ55 in tRNAs by archaeal Pus10, *RNA* 19 (9) (2013) 1279–1294. [PubMed: 23898217]
- [135]. Ramamurthy V, Swann SL, Paulson JL, Spedaliere CJ, Mueller EG, Critical aspartic acid residues in pseudouridine synthases, *J. Biol. Chem* 274 (32) (1999) 22225–22230. [PubMed: 10428788]
- [136]. Hoang C, Ferré-D'Amaré AR, Cocystal structure of a tRNA Ψ55 pseudouridine synthase: nucleotide flipping by an RNA-modifying enzyme, *Cell* 107 (7) (2001) 929–939. [PubMed: 11779468]
- [137]. Hoang C, Hamilton CS, Mueller EG, Ferré-D'Amaré AR, Precursor complex structure of pseudouridine synthase TruB suggests coupling of active site perturbations to an RNA-sequestering peripheral protein domain, *Protein Sci.* 14 (8) (2005) 2201–2206. [PubMed: 15987897]
- [138]. Spedaliere CJ, Hamilton CS, Mueller EG, Functional importance of motif I of pseudouridine synthases: mutagenesis of aligned lysine and proline residues, *Biochemistry* 39 (31) (2000) 9459–9465. [PubMed: 10924141]
- [139]. Kaya Y, Del Campo M, Ofengand J, Malhotra A, Crystal structure of TruD, a novel pseudouridine synthase with a new protein fold, *J. Biol. Chem* 279 (18) (2004) 18107–18110. [PubMed: 14999002]
- [140]. Kaya Y, Ofengand J, A novel unanticipated type of pseudouridine synthase with homologs in bacteria, archaea, and eukarya, *RNA* 9 (6) (2003) 711–721. [PubMed: 12756329]
- [141]. Ericsson UB, Nordlund P, Hallberg BM, X-ray structure of tRNA pseudouridine synthase TruD reveals an inserted domain with a novel fold, *FEBS Lett.* 565 (1–3) (2004) 59–64. [PubMed: 15135053]

- [142]. Hoang C, Ferre-D'Amare AR, Crystal structure of the highly divergent pseudouridine synthase TruD reveals a circular permutation of a conserved fold, *RNA* 10 (7) (2004) 1026–1033. [PubMed: 15208439]
- [143]. Wright JR, Keffer-Wilkes LC, Dobing SR, Kothe U, Pre-steady-state kinetic analysis of the three *Escherichia coli* pseudouridine synthases TruB, TruA, and RluA reveals uniformly slow catalysis, *RNA* 17 (12) (2011) 2074–2084. [PubMed: 21998096]
- [144]. Sivaraman J, Sauvé V, Larocque R, Stura EA, Schrag JD, Cygler M, Matte A, Structure of the 16S rRNA pseudouridine synthase RsuA bound to uracil and UMP, *Nat. Struct. Biol* 9 (5) (2002) 353–358. [PubMed: 11953756]
- [145]. Staker BL, Korber P, Bardwell JC, Saper MA, Structure of Hsp15 reveals a novel RNA-binding motif, *EMBO J.* 19 (4) (2000) 749–757. [PubMed: 10675344]
- [146]. Matte A, Louie GV, Sivaraman J, Cygler M, Burley SK, Structure of the pseudouridine synthase RsuA from *Haemophilus influenzae*, *Acta Crystallogr. F Struct. Biol. Cryst. Commun* 61 (Pt 4) (2005) 350–354.
- [147]. Lecointe F, Simos G, Sauer A, Hurt EC, Motorin Y, Grosjean H, Characterization of yeast protein Deg1 as pseudouridine synthase (Pus3) catalyzing the formation of psi 38 and psi 39 in tRNA anticodon loop, *J. Biol. Chem* 273 (3) (1998) 1316–1323. [PubMed: 9430663]
- [148]. Hur S, Stroud RM, How U38, 39, and 40 of many tRNAs become the targets for pseudouridylation by TruA, *Mol. Cell* 26 (2) (2007) 189–203. [PubMed: 17466622]
- [149]. Dong X, Bessho Y, Shibata R, Nishimoto M, Shirouzu M, Kuramitsu S, Yokoyama S, Crystal structure of tRNA pseudouridine synthase TruA from *Thermus thermophilus* HB8, *RNA Biol.* 3 (3) (2006) 115–122. [PubMed: 17114947]
- [150]. Chen J, Patton JR, Cloning and characterization of a mammalian pseudouridine synthase, *RNA* 5 (3) (1999) 409–419. [PubMed: 10094309]
- [151]. Arluison V, Buckle M, Grosjean H, Pseudouridine synthetase Pus1 of *Saccharomyces cerevisiae*: kinetic characterisation, tRNA structural requirement and real-time analysis of its complex with tRNA, *J. Mol. Biol* 289 (3) (1999) 491–502. [PubMed: 10356324]
- [152]. Grosshans H, Lecointe F, Grosjean H, Hurt E, Simos G, Pus1p-dependent tRNA pseudouridylation becomes essential when tRNA biogenesis is compromised in yeast, *J. Biol. Chem* 276 (49) (2001) 46333–46339. [PubMed: 11571299]
- [153]. Daubner GM, Cléry A, Allain FH-T, RRM-RNA recognition: NMR or crystallography...and new findings, *Curr. Opin. Struct. Biol* 23 (1) (2013) 100–108. [PubMed: 23253355]
- [154]. Allain FH, Howe PW, Neuhaus D, Varani G, Structural basis of the RNA-binding specificity of human U1A protein, *EMBO J.* 16 (18) (1997) 5764–5772. [PubMed: 9312034]
- [155]. Nurse K, Wrzesinski J, Bakin A, Lane BG, Ofengand J, Purification, cloning, and properties of the tRNA psi 55 synthase from *Escherichia coli*, *RNA* 1 (1) (1995) 102–112. [PubMed: 7489483]
- [156]. Kinghorn SM, O'Byrne CP, Booth IR, Stansfield I, Physiological analysis of the role of truB in *Escherichia coli*: a role for tRNA modification in extreme temperature resistance, *Microbiology* 148 (11) (2002) 3511–3520. [PubMed: 12427942]
- [157]. Phannachet K, Huang RH, Conformational change of pseudouridine 55 synthase upon its association with RNA substrate, *Nucleic Acids Res.* 32 (4) (2004) 1422–1429. [PubMed: 14990747]
- [158]. Pan H, Agarwalla S, Moustakas DT, Finer-Moore J, Stroud RM, Structure of tRNA pseudouridine synthase TruB and its RNA complex: RNA recognition through a combination of rigid docking and induced fit, *Proc. Natl. Acad. Sci. U.S.A* 100 (22) (2003) 12648–12653. [PubMed: 14566049]
- [159]. Chaudhuri BN, Chan S, Perry LJ, Yeates TO, Crystal structure of the apo forms of psi 55 tRNA pseudouridine synthase from *Mycobacterium tuberculosis*: a hinge at the base of the catalytic cleft, *J. Biol. Chem* 279 (23) (2004) 24585–24591. [PubMed: 15028724]
- [160]. Huang L, Pookanjanatavip M, Gu X, Santi DV, A conserved aspartate of tRNA pseudouridine synthase is essential for activity and a probable nucleophilic catalyst, *Biochemistry* 37 (1) (1998) 344–351. [PubMed: 9425056]
- [161]. Czudnochowski N, Wang AL, Finer-Moore J, Stroud RM, In human pseudouridine synthase 1 (hPus1), a C-terminal helical insert blocks tRNA from binding in the same orientation as in the

- Pus1 bacterial homologue TruA, consistent with their different target selectivities, *J. Mol. Biol.* 425 (20) (2013) 3875–3887. [PubMed: 23707380]
- [162]. Hoang C, Chen J, Vizthum CA, Kandel JM, Hamilton CS, Mueller EG, Ferré-D'Amaré AR, Crystal structure of pseudouridine synthase RluA: indirect sequence readout through protein-induced RNA structure, *Mol. Cell* 24 (4) (2006) 535–545. [PubMed: 17188032]
- [163]. Cabello-Villegas J, Nikonowicz EP, Solution structure of psi32-modified anticodon stem-loop of *Escherichia coli* tRNA^{Phe}, *Nucleic Acids Res.* 33 (22) (2005) 6961–6971. [PubMed: 16377777]
- [164]. Hamilton CS, Greco TM, Vizthum CA, Ginter JM, Johnston MV, Mueller EG, Mechanistic investigations of the pseudouridine synthase RluA using RNA containing 5-fluorouridine, *Biochemistry* 45 (39) (2006) 12029–12038. [PubMed: 17002302]
- [165]. Williamson JR, Induced fit in RNA-protein recognition, *Nat. Struct. Biol* 7 (10) (2000) 834–837. [PubMed: 11017187]
- [166]. Del Campo M, Kaya Y, Ofengand J, Identification and site of action of the remaining four putative pseudouridine synthases in *Escherichia coli*, *RNA* 7 (11) (2001) 1603–1615. [PubMed: 11720289]
- [167]. Giegé R, Sissler M, Florentz C, Universal rules and idiosyncratic features in tRNA identity, *Nucleic Acids Res.* 26 (22) (1998) 5017–5035. [PubMed: 9801296]
- [168]. Chambers RW, On the recognition of tRNA by its aminoacyl-tRNA ligase, *Prog. Nucleic Acid Res. Mol. Biol* 11 (1971) 489–525. [PubMed: 4934251]
- [169]. Giegé R, The early history of tRNA recognition by aminoacyl-tRNA synthetases, *J. Biosci* 31 (4) (2006) 477–488. [PubMed: 17206068]
- [170]. Kisselev LL, The role of the anticodon in recognition of tRNA by aminoacyl-tRNA synthetases, *Prog. Nucleic Acid Res. Mol. Biol* 32 (1985) 237–266. [PubMed: 3911276]
- [171]. Saks ME, Sampson JR, Abelson JN, The transfer RNA identity problem: a search for rules, *Science* 263 (5144) (1994) 191–197. [PubMed: 7506844]
- [172]. Perona JJ, Hadd A, Structural diversity and protein engineering of the aminoacyl-tRNA synthetases, *Biochemistry* 51 (2012) 8705–8729. [PubMed: 23075299]
- [173]. Yaremchuk A, Kriklivyi I, Tukalo M, Cusack S, Class I tyrosyl-tRNA synthetase has a class II mode of cognate tRNA recognition, *EMBO J.* 21 (14) (2002) 3829–3840. [PubMed: 12110594]
- [174]. Cavarelli J, Moras D, Recognition of tRNAs by aminoacyl-tRNA synthetases, *FASEB J.* 7 (1) (1993) 79–86. [PubMed: 8422978]
- [175]. Carter CW Jr., Cognition, mechanism, and evolutionary relationships in aminoacyl-tRNA synthetases, *Annu. Rev. Biochem* 62 (1) (1993) 715–748. [PubMed: 8352600]
- [176]. Ibba M, Soll D, Aminoacyl-tRNA synthesis, *Annu. Rev. Biochem* 69 (2000) 617–650. [PubMed: 10966471]
- [177]. Carter CW, Wolfenden R, tRNA acceptor-stem and anticodon bases embed separate features of amino acid chemistry, *RNA Biol.* 13 (2) (2016) 145–151. [PubMed: 26595350]
- [178]. Bullwinkle TJ, Ibba M, Emergence and evolution, *Top. Curr. Chem* 344 (2014) 43–87. [PubMed: 23478877]
- [179]. Beuning PJ, Musier-Forsyth K, Transfer RNA recognition by aminoacyl-tRNA synthetases, *Biopolymers* 52 (1) (1999) 1–28. [PubMed: 10737860]

Fig. 1. Characteristic secondary cloverleaf structure of tRNA with relevant stem loops indicated and color-coded (*left*). L-shaped tertiary structure following the same color-coding with the amino acid-accepting arm, codon recognition arm, and the elbow indicated (*right*). The specific base pairs that are needed to form a stable L-shaped structure are numbered and shown as *bonds* in *colors* according to their position in the tRNA (base pairs include nucleotides 15 and 48, 18 and 55, and 19 and 56).

Fig. 2.

A schematic view of the modified nucleotide positions within the cloverleaf secondary structure of tRNA and the respective bacterial modification enzymes for pseudourine (Y, *left*), methylation (M, *middle*), and A-to-I deamination (I, *right*).

Fig. 3.

Topological diagrams of common RNA-binding domains (*left*) and representative crystal structures of RNA methyltransferases containing these domains (*right*). In the topological diagrams the α -helices are shown as *cylinders* in *blue* and β -sheets as *arrows* in *red*. α -Helices that are not well conserved are shown in *gray*. Representative crystal structures are depicted in *cartoon form* with the corresponding domain shown with β -sheets in *yellow* and α -helices in *purple* and *dark blue*. All enzymes are shown here in monomeric form and the protein structure outside of the featured domain is shown in *gray cartoon*. These models were created using the Visual Molecular Dynamics (VMD) program [52].

Fig. 4.

Processing pathway of tRNA^{Phe} in *Saccharomyces cerevisiae*. tRNA^{Phe} is transcribed in the nucleus with 5'- and 3'-extensions (*orange*) and an intron (*red*). The extensions are removed and a 3'-CCA sequence is added in the nucleus. The tRNA is exported to the cytoplasm and undergoes splicing by the splicing endonuclease (SEN) protein complex located to the outer mitochondrial membrane facing the cytoplasm. The spliced tRNA is 2'-O-methylated by Trm7/Trm732 for Cm₃₂ and Trm7/Trm734 for Gm₃₄. The tRNA is then imported back into the nucleus where it is methylated to m¹G₃₇ by Trm5, initiating the first step of the yW pathway. The tRNA is reexported where it undergoes further biosynthesis steps to form yW when the presence of Cm₃₂ and Gm₃₄ serve as key determinants.

Fig. 5.

Topological diagram of the S4 RNA-binding domain (*top left*) and representative crystal structure of RsuA containing an S4 domain (*top right*). The protein structure outside of the featured domain is shown in *gray cartoon*. Topological diagram of the core domain of pseudouridine synthases based off of the *E. coli* enzyme TruA (*bottom left*). The core of pseudouridine synthases is comprised of two domains each with a $\beta\alpha\beta\beta\alpha\beta$ arrangement forming an alternating eight-stranded β -sheet. The arrangement, but not necessarily the connectivity shown here, is conserved among pseudouridine synthases. Representative crystal structure of human Pus1 (*bottom right*). These models were created using the Visual Molecular Dynamics (VMD) program [52].