protein structure communications

Acta Crystallographica Section F
Structural Biology and Crystallization Communications

ISSN 1744-3091

Structure of a class II TrmH tRNA-modifying enzyme from *Aquifex aeolicus*

Elizabeth Pleshe, John Truesdell and Robert T. Batey*

Department of Chemistry and Biochemistry, University of Colorado at Boulder, Campus Box 215, Boulder, CO 80309-0215, USA

Correspondence e-mail: robert.batey@colorado.edu

Received 16 May 2005
Accepted 19 July 2005
Online 30 July 2005

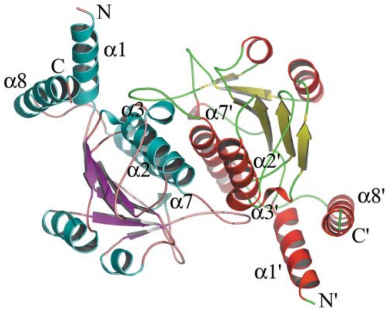
PDB Reference: *Aae*TrmH, 1zjr, r1zjrsf.

Biological RNAs contain a variety of post-transcriptional modifications that facilitate their efficient function in the cellular environment. One of the two most common forms of modification is methylation of the 2′-hydroxyl group of the ribose sugar, which is performed by a number of *S*-adenosylmethionine (SAM) dependent methyltransferases. In bacteria, many of these modifications in tRNA and rRNA are carried out by the α/β-knot superfamily of enzymes, whose SAM-binding pocket is created by a characteristic deep trefoil knot. TrmH, an enzyme found throughout all three kingdoms of life, modifies the universally conserved guanosine 18 position of tRNA. The crystal structure of TrmH from the thermophilic bacterium *Aquifex aeolicus* has been determined at 1.85 Å resolution using data collected from a synchrotron-radiation source. The protein reveals a fold typical of members of the SpoU clan of proteins, a subfamily of the α/β-knot superfamily, with α-helical extensions at the N- and C-termini that are likely to be involved in tRNA binding.

1. Introduction

RNAs involved in translation and post-transcriptional processing contain a large number of modified nucleosides (Limbach *et al.*, 1994). In tRNAs alone over 80 different modifications have been observed (Sprinzl *et al.*, 1998), many of which are important for tRNA function (Urbonavicius *et al.*, 2002; Nobles *et al.*, 2002). Modifications within the anticodon loop are necessary for proper codon recognition (Martinez Gimenez *et al.*, 1998), while those in the D- and T-loops facilitate folding and stabilization of the tRNA (Vermeulen *et al.*, 2005; Derrick & Horowitz, 1993). Ribosomal RNA modifications cluster around the catalytic center and other crucial regions of the ribosome, further underscoring their importance in translation (Decatur & Fournier, 2002). Bacteria also use RNA modifications to confer resistance to a number of different antibiotics. One of the most common modes of resistance in bacteria is through methylation of the 23S rRNA around the peptidyl exit channel, inhibiting the binding of macrolide-class antibiotics (Yonath, 2003; Steitz, 2005).

Of the almost 100 characterized RNA modifications, the two most common types are pseudouridine (ψ) and 2′-*O*-methylation of the ribose sugar (Bjork *et al.*, 1999). One group of enzymes responsible for 2′-*O*-methylation of both rRNA and tRNA is the α/β-knot superfamily of methyltransferases (MTases; Bateman *et al.*, 2004), which have been identified in all three kingdoms of life (Cavaille *et al.*, 1999). Like other methyltransferases, the α/β-knot superfamily requires SAM for methyl transfer, but does not use the classic Rossmann fold to form the SAM-binding domain (Martin & McMillan, 2002). Instead, they contain a unique α/β-fold involving the formation of a deep trefoil knot (Nureki *et al.*, 2002, 2004; Lim *et al.*, 2003). Outside the highly conserved α/β-knot core, each enzyme contains a smaller non-conserved region located at the N- and/or C-termini that often exhibits structural similarity to RNA-binding proteins. To date, the structures of a number of α/β-knot superfamily members have been solved, including *Streptomyces viridochromogenes* AviRb (Mosbacher *et al.*, 2005), *Thermus thermophilus* TrmH (Nureki *et al.*, 2004), *T. thermophilus* RrmA (Nureki *et al.*, 2002), *Haemophilus influenzae* TrmD (Ahn *et al.*, 2003), *H. influenzae* YibK



(Lim *et al.*, 2003), *Escherichia coli* RlmB (Michel *et al.*, 2002), *E. coli* TrmD (Elikins *et al.*, 2003) and *H. influenzae* YggJ (Forouhar *et al.*, 2003).

The SpoU family of MTases (Bateman *et al.*, 2004), a subfamily of the α/β-knot superfamily, includes TrmH, which catalyzes methyl-group transfer from SAM to the 2′-O position of guanosine 18, a universally conserved base in the D-loop of tRNA, forming Gm18 (Kumagai *et al.*, 1980; Persson *et al.*, 1997). Of all the modified nucleosides present in tRNA, only eight, Gm18 being one, are found in the same position and the same subpopulation of tRNA in organisms from all kingdoms, suggesting that they may have been present in the tRNA of the last common ancestor (Bjork *et al.*, 1999). Within fully folded tRNA, G18 is located in the highly conserved core of tRNA formed by base-pairing of the G18–ψ55 and G19–C56 residues between the D- and T-loops. Since 2′-*O*-methylation stabilizes the C3′-*endo* conformation of the ribose sugar, it is believed that it rigidifies the tRNA in a crucial region for establishing its global fold (Hori *et al.*, 2002). While early studies showed a lack of effect upon loss of G18 methylation (Persson *et al.*, 1997), new research has demonstrated a very necessary role for methylation at this position (Urbonavicius *et al.*, 2002). Deletion of the enzymes responsible for both G18 methylation and pseudouridylation of its pairing partner (ψ55) resulted in a reduced growth rate and in defects in the translation of certain codons. This was in part a result of decreased A-site selection and P-site slippage by $tRNA^{Tyr}$, resulting in an elevated level of frameshifting (Urbonavicius *et al.*, 2001, 2002). In addition, Gm18-deficient tRNA is unable to efficiently read the stop codon UAG, indicating a function beyond tRNA stabilization by altering its conformation and affecting its interaction with the ribosome (Urbonavicius *et al.*, 2002).

Two distinct classes of TrmH methyltransferases have been characterized based upon their specificity towards tRNA. Class I enzymes, usually found in thermophilic organisms, have no substrate specificity and methylate all tRNAs at the G18 position. These enzymes include the structurally characterized TrmH from *T. thermophilus* (*Tth*TrmH; Hori *et al.*, 1998, 2002). Class II enzymes, including *E. coli* (Persson *et al.*, 1997) and *A. aeolicus* TrmH (*Aae*TrmH; Hori *et al.*, 2003) and *Saccharomyces cerevisiae* Trm3 (Cavaille *et al.*, 1999), only methylate a subset of tRNAs *in vivo* (Hori *et al.*, 2003). The recognition of these enzymes as belonging to class II is based upon analysis of G18-methylation patterns *in vivo*. Both classes recognize a few common features of tRNA, primarily the conserved G18G19 dinucleotide step within the D-loop and the C11–G24 base pair of the D-stem (Hori *et al.*, 2002). Analysis of the kinetics of methylation by *Aae*TrmH revealed additional base pairs in the D-stem and two bases in the D-loop, Py15Py16, to be necessary for efficient methyl transfer (Hori *et al.*, 2003). Along with a larger D-loop and shorter D-stem, these features are common characteristics of those tRNA methylated by *Aae*TrmH *in vivo* [*Aae* $tRNA^{Leu}$ (CAG), $tRNA^{Pro}$ (GGG), $tRNA^{Gln}$ (TTG) and $tRNA^{Gly}$ (GCC); Hori *et al.*, 2003].

Although RNA methylation is known to be important, how substrate RNAs are recognized and specifically methylated is not well understood. Currently, there is little structural information on how these enzymes specifically interact with target RNAs or catalyze modification. Towards this goal, we have solved the first structure of a class II TrmH, *Aae*TrmH. Comparison of the structures of the two classes of TrmH illustrates an almost identical structure in both the catalytic domain and RNA-binding extension. Analysis of the electrostatic surface potential of the two classes reveals an additional basic region in *Aae*TrmH, which may harbor tRNA-recognition elements that give rise to the greater tRNA specificity for the class II enzymes.

Table 1
Crystallographic data statistics of *A. aeolicus* TrmH.

Values in parentheses correspond to the highest resolution shell.

Data collection	
Space group	$P6_122$
Unit-cell parameters (Å, °)	a = 66.658, b = 66.658, c = 194.884, $\alpha = \beta = 90$, $\gamma = 120$
Resolution (Å)	20–1.85 (1.92–1.85)
Wavelength (Å)	1.1271
Completeness (%)	95.3 (65.0)
Measured reflections	714432
Unique reflections	22968
Average redundancy	13.4
$\langle I\rangle/\langle\sigma(I)\rangle$	47.3 (2.9)
R_{merge}† (%)	5.8 (42.4)
Refinement	
No. of monomers per asymmetric unit	1
Resolution	20–1.85 (2.02–1.85)
No. of reflections	
Working	18851 [82.7%]
Test set	2061 [9.0%]
R_{xtal}‡	22.1 (32.3)
R_{free}	25.7 (37.9)
R.m.s.d. bonds (Å)	0.0053
R.m.s.d. angles (°)	1.17
Cross-validated Luzzati coordinate error (Å)	0.29
Cross-validated σ_A coordinate error (Å)	0.22
Average B factor (Å^2)	42.0
Ramachandran plot statistics (%)	
Residues in most favored regions	93.3
Residues in additional allowed regions	6.1
Residues in generously allowed regions	0.0
Residues in disallowed regions	0.6

† $R_{merge} = \sum|I - \langle I\rangle|/\sum I$, where I is the observed intensity and $\langle I\rangle$ is the average intensity of multiple measurements of symmetry-related reflections. ‡ $R_{xtal} = \sum||F_o| - |F_c||/\sum|F_o|$ for the working set. R_{free} is the same for the test set.

2. Methods and materials

2.1. Cloning and purification

The *trmH* gene (accession No. NP_214299) was amplified by PCR from genomic DNA of *A. aeolicus* VF5 using standard molecular-biological techniques (Sambrook *et al.*, 1989). The 5′ primer contained a sequence such that a TEV protease-cleavage site was added prior to the first initiator codon of the gene encoding *trmH*. The fragment was digested with *Nco*I and *Xba*I and cloned into the expression vector pThioHisB (Invitrogen), fusing TrmH to the C-terminus of thioredoxin. Ampicillin-resistant colonies were selected following transformation into *E. coli* strain DH5α and the result was plasmid-sequenced to verify its correctness.

*Aae*TrmH was overexpressed in *E. coli* Rosetta(DE3)/pLysS cells (Novagen) in 2×YT medium at 310 K. To induce expression, 1 m*M* IPTG was added to the medium when the OD_{600} reached 0.35 and growth was allowed to continue at 310 K for 4 h prior to harvesting the cells by centrifugation. Cells were resuspended in 50 m*M* Tris–HCl pH 8.0, 300 m*M* NaCl and lysed by three rounds of freeze/thawing in liquid nitrogen; the viscosity of the resulting solution was reduced by adding 10 U DNaseI per millilitre of lysate and incubating on ice for 30 min. After centrifugation, the recombinant protein was in the pellet fraction as inclusion bodies, which were resuspended in 8 *M* urea, 50 m*M* sodium phosphate pH 8.0 and refolded by dialysis against a buffer containing 25 m*M* Tris–HCl pH 8.0, 125 m*M* NaCl and 10 m*M* β-mercaptoethanol overnight at 277 K. The soluble fraction was cleaved with TEV protease (Lucast *et al.*, 2001) for 2 d at room temperature to remove the thioredoxin tag and applied onto an SP-Sepharose column. Protein was eluted using a gradient of 0.1–1.5 *M* NaCl and fractions containing TrmH were pooled and applied onto a Sephadex-75 prep-grade column (Pharmacia) using a running buffer containing 20 m*M* Na MES pH 6.0, 100 m*M* NaCl, 5 m*M* Na

EDTA, 2 m*M* DTT and 0.1 m*M* sodium azide. The protein eluted with an apparent molecular weight of ∼54 000 Da, consistent with previous findings that this family of proteins always exists in the form of a homodimer. Fractions containing protein were pooled and analyzed by 15% SDS–PAGE; the resulting material was >99% pure as judged by Coomassie staining of the gel.

2.2. Crystallization and data collection

For crystallization, the protein was exchanged into a buffer containing 10 m*M* Na MES pH 6.0 and concentrated to 400 µ*M* monomer (9.8 mg ml^{-1}). The protein was crystallized by the hanging-drop method, in which 2 µl protein solution was mixed with 2 µl reservoir solution containing 50 m*M* sodium citrate pH 4.4, 200 m*M* ammonium sulfate and 7.5% PEG 1000 at 303 K. Crystals grew in 2–4 d to a maximal size of 100 × 100 × 200 µm. They were cryoprotected in a solution comprising the reservoir solution plus 30% glycerol for 1.5 h at room temperature and flash-frozen in liquid nitrogen prior to data collection.

Diffraction data were obtained on beamline 8.2.2 of the Lawrence Berkeley Advanced Light Source. Data collection was conducted on crystals at 100 K at a wavelength of 1.1271 Å using an ADSC Q315 2×2 CDC array detector. Data were indexed, integrated and scaled using the *HKL*2000 package (Otwinowski & Minor, 1997). Crystallographic data statistics are given in Table 1.

2.3. Structure solution and refinement

The structure of *Aae*TrmH (SpoU) was determined using diffraction amplitudes extending to 1.85 Å resolution and the molecular-replacement method with amino acids 85–242 from *E. coli* RlmB (PDB code 1gz0; Michel *et al.*, 2002) as the initial search model. Cross-rotational and translational searches performed in *CNS* (Brünger *et al.*, 1998) revealed a single clear solution that yielded an interpretable electron-density map. The structure was built using *O* (Jones *et al.*, 1991) and refined using iterative cycles of simulated-annealing/*B*-factor refinement in *CNS* and model building. Following the building of residues 6–202, for which clear electron density was observed, 174 waters, three glycerol molecules and a single sulfate ion were added to the model. The *R* factor and $R_{\rm free}$ values after the final round of simulated annealing and manual rebuilding are 22.1 and 25.7%, respectively. The root-mean-square deviation of the bonds is 0.005 Å and that for angles is 1.17°; analysis of the model using *PROCHECK* (Laskowski *et al.*, 1993) revealed that only one residue (Lys62) lies outside the allowed region of the Ramachandran plot and has an overall *G* factor (2.3 bandwidths from the mean) that is better than expected for a protein at this resolution.

3. Results and discussion

3.1. Purification and crystallization

SDS–PAGE analysis of the soluble, membrane-associated and insoluble fractions of the overexpression of the thioredoxin–TrmH fusion in *E. coli* using a number of temperatures, IPTG concentrations and media types revealed the protein to be consistently expressed as inclusion bodies. Following a standard inclusion-body preparation (Marston & Hartley, 1990), in which the insoluble fraction was detergent washed with buffer containing 1.0% Triton X-100, we were able to obtain a fraction that was >98% pure. The protein was refolded from a solution containing 8 *M* urea by dialysis, after which a majority of the protein was found in the soluble fraction, with only a small amount of *Aae*TrmH, along with a few minor impurities, reprecipitating. Like all structurally characterized members of the α/β-knot superfamily of proteins, the refolded protein displayed behavior on a gel-filtration column (data not shown) that was consistent with it being a dimer, indicating that it was properly refolded. Also, in a recent study the folding of the TrmH homolog YibK has been probed using intrinsic fluorescence and far-UV circular dichroism (Mallam & Jackson, 2005). Despite the presence of a deep topological knot in the protein, YibK efficiently refolded; this is consistent with our observation that TrmH can be refolded to yield a conformation that closely resembles that of other members of this family.

Initial microcrystals of *Aae*TrmH were obtained under a single condition (25% PEG 4000, 0.1 *M* sodium acetate pH 4.6, 0.2 *M* ammonium sulfate) at room temperature in the Magic 50 sparse-matrix screen (Jancarik & Kim, 1991). Optimization of these conditions showed that this protein would crystallize at a pH lower than 6.0, but diffraction-quality crystals could only be obtained in the pH range 4.0–4.5. The structure reveals that one of the principal crystal lattice contacts occurs between six acidic residues: Glu185, Glu186 and Glu187 in one dimer and Asp31′, Glu128′ and Glu135′ in the neighbor. As two of these (Glu187 and Glu135′) form a direct hydrogen bond (Fig. 1), it is likely that one or more of these residues needed to be protonated to form the extended lattice. Crystallization trials with *Aae*TrmH were also performed in the presence of the natural cofactor *S*-adenosylmethionine as well as *S*-adenosylhomocysteine and a substrate mimic, 5′-guanosine monophosphate. Diffraction-quality crystals could not be obtained in the presence of any of these compounds and thus only the unliganded form of this protein has been solved.

3.2. Overall structure

*Aae*TrmH, like other SpoU rRNA/tRNA methylase family members, consists of a catalytic core domain of six parallel β-strands (β3-β2-β1-β5-β4-β6) flanked by four α-helices (α4–α6 on one side of the sheet and α2–α7 on the other) and a peripheral extension

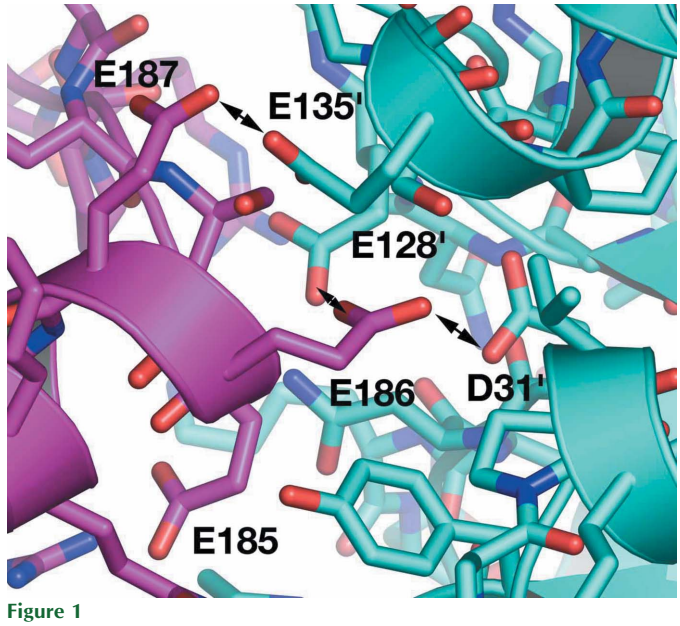


Figure 1
Acidic residues at the interface between proteins in the crystal lattice. The arrows indicate residues in close proximity: Glu187 and Glu135′ are within hydrogen-bonding distance (2.5 Å) and Glu186 is within 4.4–4.6 Å distance from Asp31′ and Glu128′.

consisting of two α-helices (α1 and α8; Fig. 2*a*). The catalytic core of the enzyme (residues 26–175) is similar to the Rossmann fold, with the exception of a C-terminal knot formed by the threading of the β6 strand through a loop connecting β4 to β5. In addition to the characteristic left-handed deep trefoil knot that defines the α/β-knot superfamily (Michel *et al.*, 2002; Lim *et al.*, 2003), *Aae*TrmH forms a homodimer, with interactions between α2 and α7 of each monomer and a long flexible loop region connecting β6 and α7 (residues 144–157) forming much of the dimer interface. At the N- and C-termini, a structure specific for the TrmH subfamily of proteins is formed through packing of α1 and the α8 helices.

Superposition of the catalytic core domains of the *Tth*TrmH and *Aae*TrmH variants showed significant similarities between the two structures, with an r.m.s. deviation of 1.095 Å between their backbone atoms (total of 155 C^{α} atoms). Substantial differences in the two structures only occur in two regions of the protein. The most loosely organized part of the protein, as judged by *B* factors, lies in the loop connecting the β2 and β3 strands, for which the backbone was built, but most of the side chains are too disordered to place in the model. Interestingly, this region is also the most disordered part of the knotted domains of *Tth*TrmH and YibK and is altogether absent in TrmD, suggesting that this may be a loosely organized region of the fold or that it only becomes structured in the presence of its RNA target. The other region of the protein that differs between the two known structures of TrmH consists of two loops (β4–α5 and β5–α6) directly adjacent to the SAM-binding pocket (see below).

Homodimer formation by TrmH and its homologs is mediated by a number of interface contacts between the α7 helix and the β6–α7 loop of each monomer. This results in a sizable amount of buried surface area at the interface (3049 $Å^2$ in total, with 1525 $Å^2$ contributed from each monomer); dimerization buries 14.5% of the total surface area of each monomer. Each monomer is related to the other *via* a *C*2 symmetry axis running down the interface (Fig. 2*b*), similar to all other members of the α/β-knot superfamily. The majority of the buried surface involves two of the three conserved motifs found within the SpoU family, motif I (residues 28–45) and motif III (residues 144–165), where motif I interacts principally with motif III of the other monomer. This interface is mediated through van der Waals packing of a number of highly conserved hydrophobic residues in the two conserved motifs. A smaller part of the interface

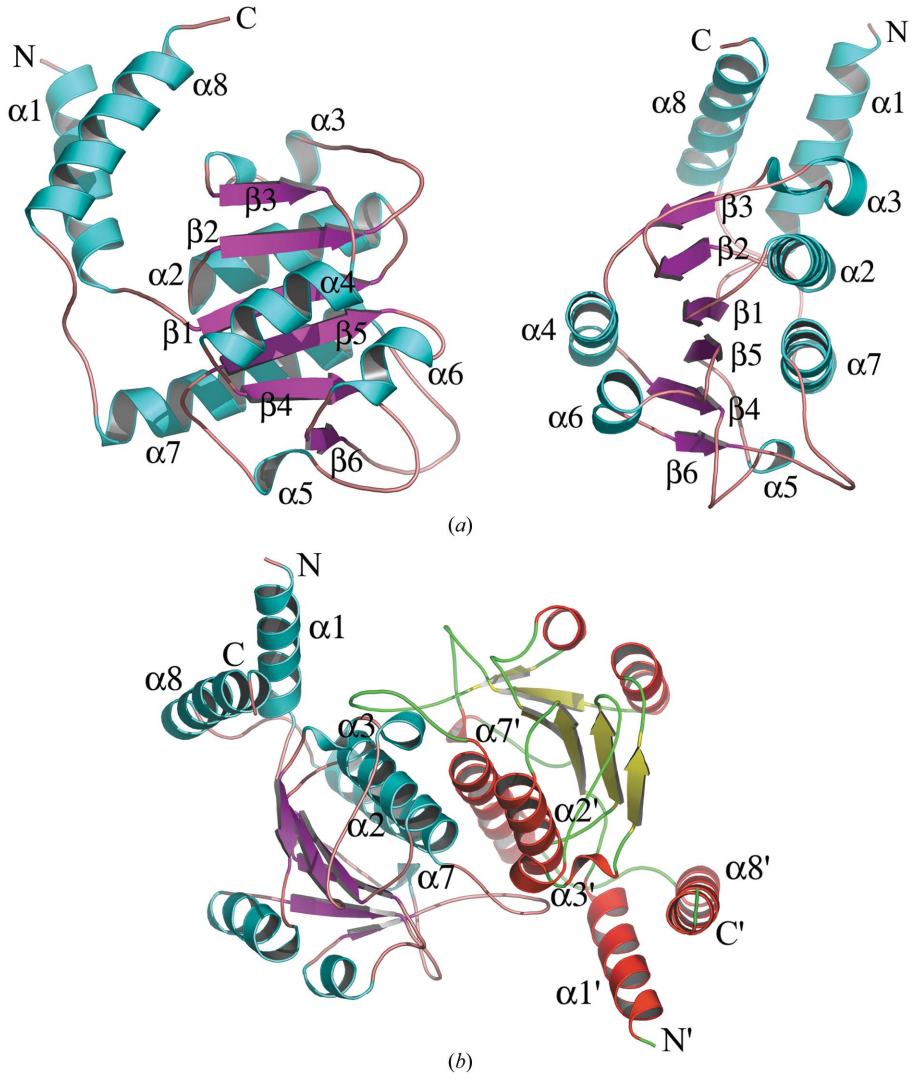


Figure 2
(*a*) Overall structure of the TrmH monomer from *A. aeolicus*. α-Helices are shown in cyan, β-sheets in magenta and loops in pink. The SAM-binding domain common to all members of the α/β-knot superfamily corresponds to the structure from β1–α7, while the RNA-binding extension consists of α1 and α8. The left panel shows a side view of the protein and the right shows a 90° clockwise rotation of the monomer with respect to the perspective on the left, emphasizing the central β-sheet flanked by α-helices on each side. (*b*) The biological dimer, with one monomer in cyan and magenta and the second represented in red and yellow.

comprises interactions between the C-terminal portion of α7 and α5 of the adjacent monomer, which is primarily mediated through electrostatic and polar interactions (Gln169–Arg113′, Tyr166–Pro147′ and Glu173–Tyr149′).

3.3. Active site

Although *Aae*TrmH was not able to be crystallized with SAM or an analog in the active site, several observations based upon a comparison to *Tth*TrmH are instructive. Superposition of the unliganded *Aae*TrmH and the SAM-bound *Tth*TrmH clearly illustrate that the binding pocket for the two enzymes is organized in the same fashion, with SAM fitting perfectly into a deep crevice formed between the β4–α5, β5–α6 and β6–α7 loops part of conserved sequence motifs II (residues 124–132) and III (Fig. 3*a*). However, there are several differences between the two structures that reflect changes in the protein structure upon cofactor binding. The first is a substantial change in the conformation of the β5–α6 loop, which is pushed out owing to interactions with the ribose sugar and methionine of the SAM cofactor in the bound form. While this conformational change has been noted in comparison of unliganded and liganded forms of *Tth*TrmH (Nureki *et al.*, 2004), similar changes are not observed in YibK or AviRB, suggesting that it only occurs in TrmH. The other change in the binding pocket is a movement of the side chain of Leu105, which forms van der Waals contacts with the adenine nucleobase in SAM in the bound form. Within the SAM-binding pocket of *Aae*TrmH, two very well ordered glycerol molecules were observed (Fig. 3*b*), forming extensive hydrogen-bonding interactions with each other and the protein. These molecules occupy the same location as the ribose sugar and methionine side chain in the SAM-bound form of *Tth*TrmH.

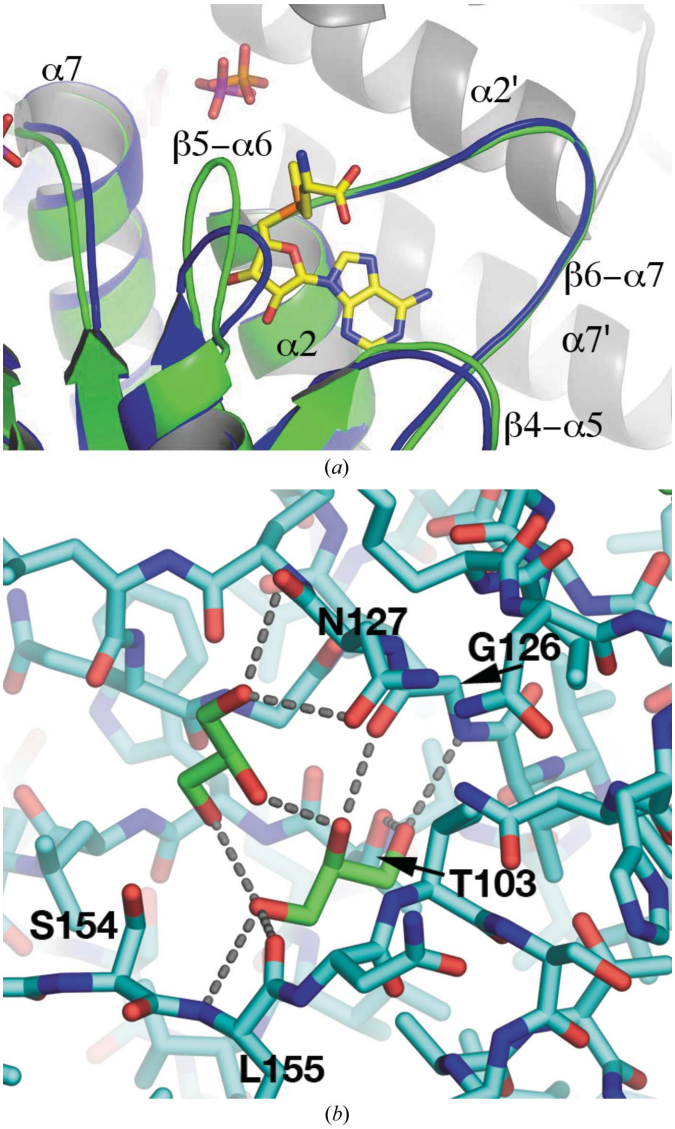


Figure 3
(*a*) Superposition of the active site of the unliganded *Aae*TrmH (cyan) and the SAM-bound *Tth*TrmH (green, with cofactor in yellow), with the other *A. aeolicus* monomer shown in gray for perspective. The position of a bound sulfate in the *A. aeolicus* structure (magenta) is nearly identical to that in the *T. thermophilus* structure (orange). (*b*) Binding of two glycerol molecules within the active site of TrmH. Each glycerol forms two hydrogen bonds with the other glycerol, along with a number of hydrogen bonds to neighboring protein residues.

Outside the SAM-binding pocket, a number of conserved residues are found that have been implicated in the mechanism by which TrmH methylates the 2′-hydroxyl group. Most of the amino-acid side chains suggested to play a role in substrate binding and the reaction mechanism are found in almost identical positions in the *Aae*TrmH and *Tth*TrmH structures. The only substantial difference is that of Glu154, which helps to buttress Arg44′ (supplied by the other monomer) at the active site; in our structure this side chain was disordered, which is likely to be a consequence of the absence of hydrogen-bonding interactions with the SAM cofactor. A key similarity in the two structures was the presence of a highly ordered sulfate ion adjacent to the active site, which has been proposed as the binding site for the G18 phosphate of tRNA (Fig. 3*a*; Nureki *et al.*, 2004). While the position of this ion is nearly identical in the two structures, it interacts with different residues in *Aae*TrmH (Asn35, His37′, Asn38, Asn156) and *Tth*TrmH (Lys32, His34, Asn35 and Asn152). Thus, our structure is entirely consistent with the available biochemical and structural data regarding the critical components of the active site and their likely arrangement during catalysis.

3.4. RNA-binding extension

Each member of the α/β-knot superfamily uses a unique and specialized RNA-binding extension that lies outside the conserved SAM-binding domain, with the exception of YibK. This is most clear for RlmB, a ribosomal RNA 2′-*O*-methyltransferase whose RNA-binding domain displays significant structural homology to the ribosomal proteins L7ae and L30 (Michel *et al.*, 2002). Limited proteolysis of *Tth*TrmH showed tRNA-binding activity resides within the first ten amino acids at the N-terminus and the last 20 amino acids at the C-terminus (Hori *et al.*, 2002). These residues in *Aae*TrmH form two α-helices (α1 and α8; Fig. 2*a*) that pack against each other and against the SAM-binding domain through interactions with the α2–β2 and α3–β3 loops. Superposition of the SAM-binding domains of *Aae*TrmH and *Tth*TrmH reveal that the α1/α8 extension is oriented slightly differently, suggesting that its orientation is somewhat flexible with respect to the SAM-binding domain prior to binding tRNA.

Biochemical analysis of *Aae*TrmH and *Tth*TrmH revealed that class II enzymes recognize additional elements in the tRNA, giving rise to their preferences for certain tRNAs. It has been suggested that the additional sequences at the C-terminus of the class II enzymes may be responsible for this specificity. In our structure, however, this additional sequence (approximately ten amino acids) is disordered and thus does not provide a structural basis for the distinction between the class I and II enzymes. However, an analysis of the electrostatic potential along the surfaces of the two enzymes reveals a significant difference. Both enzymes display a bipartite distribution of

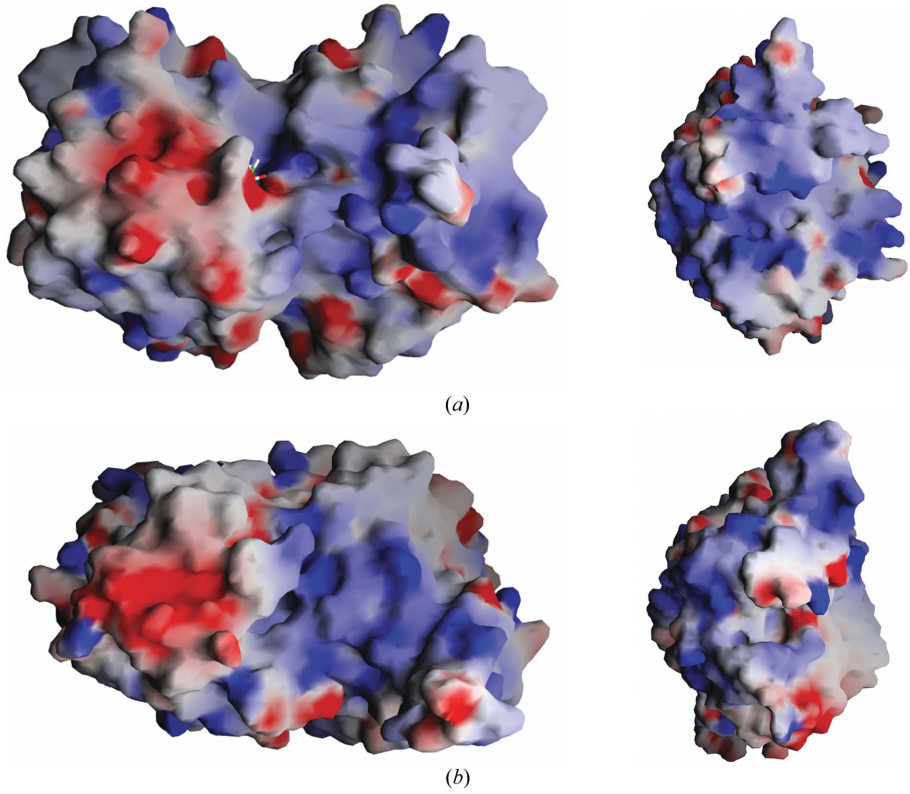


Figure 4
Electrostatic potential along the surface of the TrmH homodimers, with $-10kT$ shown in red and $+10kT$ in blue. (*a*) On the left is the *Aae*TrmH displayed looking at one of the active sites in the dimer, showing one of the basic patches directly adjacent the SAM-binding pocket. The right shows the second site of significant positive electrostatic potential as seen from the back side of the $\alpha1/\alpha8$ extension. (*b*) The same perspectives shown for *Tth*TrmH.

acidic and basic residues on the faces that contain the SAM-binding pocket. The side that contains the cofactor-binding pocket is predominantly acidic in both enzymes (Figs. 4*a* and 4*b*), while the opposite side, formed by the $\alpha1/\alpha8$ extension, is very basic, suggesting it contacts negatively charged phosphate backbone of the tRNA. On the opposite side of the $\alpha1/\alpha8$ extension, however, the *Aae*TrmH protein contains another very basic region (Fig. 4*a*) that is not present in the *Tth*TrmH (Fig. 4*b*). This region may harbor residues involved in the additional specificity of the class II enzymes. As it is likely that the tRNA must be unfolded for TrmH to gain access to the 2′-hydroxyl group of G18, as occurs in other tRNA-modifying enzymes (Ishitani *et al.*, 2003), the tRNA could conceivably contact both regions of the protein.

This work was supported through grants from the Research Corporation and the W. M. Keck foundation. The authors wish to thank Steve Edwards for maintaining the X-ray crystallography facility, Anne Gooding and Elaine Podell for collecting data at the Advanced Light Source and Sunny Gilbert, Rebecca Montange and Colby Stoddard for critically reading this manuscript.

References

Ahn, H. J., Kim, H. W., Yoon, H. J., Lee, B., Suh, S. W. & Yang, J. K. (2003). *EMBO J.* **22**, 2593–2603.
Bateman, A., Coin, L., Durbin, R., Finn, R. D., Hollich, V., Griffiths-Jones, S., Khanna, A., Marshall, M., Moxon, S., Sonnhammer, E. L., Studholme, D. J., Yeats, C. & Eddy, S. R. (2004). *Nucleic Acids Res.* **32**, D138–D141.
Bjork, G. R., Durand, J. M. B., Hagervall, T. G., Leipuviene, R., Lundgren, H. K., Nilsson, K., Chen, P., Qian, Q. & Urbonavicius, J. (1999). *FEBS Lett.* **452**, 47–51.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Cavaille, D., Chetouani, F. & Bachellerie, J. P. (1999). *RNA*, **5**, 66–81.
Decatur, W. A. & Fournier, M. J. (2002). *Trends Biochem. Sci.* **27**, 344–351.
Derrick, W. B. & Horowitz, J. (1993). *Nucleic Acids Res.* **21**, 4948–4953.
Elikins, P. A., Watts, J. M., Zalacain, M., van Thiel, A., Vitazka, P. R., Redlak, M., Andraos-Selim, C., Rastinejad, F. & Holmes, W. M. (2003). *J. Mol. Biol.* **333**, 931–949.
Forouhar, F., Shen, J., Xiao, R., Acton, T. B., Montelione, G. T. & Tong, L. (2003). *Proteins*, **53**, 329–332.
Hori, H., Kubota, S., Watanabe, K., Kim, J. M., Ogasawara, T., Sawasaki, T. & Endo, Y. (2003). *J. Biol. Chem.* **278**, 25081–25090.
Hori, H., Suzuki, T., Sugawara, K., Inoue, Y., Shibata, T., Kuramitsu, S., Yokoyama, S., Oshima, T. & Watanabe, K. (2002). *Genes Cells*, **7**, 259–272.
Hori, H., Yamazaki, N., Matsumoto, T., Watanabe, Y., Ueda, T., Nishikawa, K., Kumagai, I. & Watanabe, K. (1998). *J. Biol. Chem.* **273**, 25721–25727.
Ishitani, R., Nureki, N., Okada, N., Nishimura, S. & Yokoyama, S. (2003). *Cell*, **113**, 383–394.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Jones, T. A., Zou, J. Y. & Kjeldgaard, S. W. (1991). *Acta Cryst.* A**47**, 110–119.
Kumagai, I., Watanabe, K. & Oshima, T. (1980). *Proc. Natl Acad. Sci. USA*, **77**, 1922–1926.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Lim, K., Tempczyk, A., Krajewski, W., Bonander, N., Toedt, J., Howard, A., Eisenstein, E. & Herzberg, O. (2003). *Proteins*, **51**, 56–67.
Limbach, P. A., Crain, P. F. & McCloskey, J. A. (1994). *Nucleic Acids Res.* **22**, 2183–2196.
Lucast, L. J., Batey, R. T. & Doudna, J. A. (2001). *Biotechniques*, **30**, 544–546.
Mallam, A. L. & Jackson, S. E. (2005). *J. Mol. Biol.* **346**, 1409–1421.
Marston, F. A. O. & Hartley, D. L. (1990). *Methods Enzymol.* **182**, 264–276.

Martin, J. L. & McMillan, F. M. (2002). *Curr. Opin. Struct. Biol.* **12**, 783–793.
Martinez Gimenez, J. A., Saez, G. T. & Seisdedos, R. F. (1998). *J. Theor. Biol.* **194**, 485–490.
Michel, G., Sauve, V., Larocque, R., Li, Y., Matte, A. & Cygler, M. (2002). *Structure*, **10**, 1303–1315.
Mosbacher, T. G., Bechthold, A. & Schulz, G. E. (2005). *J. Mol. Biol.* **345**, 535–545.
Nobles, K. N., Yarian, C. S., Liu, G., Guenther, R. H. & Agris, P. F. (2002). *Nucleic Acids Res.* **30**, 4751–4760.
Nureki, O., Shirouzu, M., Hashimoto, K., Ishitani, R., Terada, T., Tamakoshi, M., Oshima, T., Chijimatsu, M., Takio, K., Vassylyev, D., Shibata, T., Inoue, Y., Kuramitsu, S. & Yokoyama, S. (2002). *Acta Cryst.* D**58**, 1129–1137.
Nureki, O., Watanabe, K., Fukai, S., Ishii, R., Endo, Y., Hori, H. & Yokoyama, S. (2004). *Structure*, **12**, 593–602.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Persson, B. C., Jager, G. & Gustafsson, C. (1997). *Nucleic Acids Res.* **25**, 4093–4097.
Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, 2nd ed. Cold Spring Harbor, NY, USA: Cold Spring Harbor Press.
Sprinzl, M., Horn, C., Brown, M., Ioudovitch, A. & Steinberg, S. (1998). *Nucleic Acids Res.* **26**, 148–153.
Steitz, T. A. (2005). *FEBS Lett.* **579**, 955–958.
Urbonavicius, J., Durand, J. M. B. & Bjork, G. R. (2002). *J. Bacteriol.* **184**, 5348–5357.
Urbonavicius, J., Qian, Q., Durand, J. M., Hagervall, T. G. & Bjork, G. R. (2001). *EMBO J.* **20**, 4863–4873.
Vermeulen, A., McCallum, S. A. & Pardi, A. (2005). *Biochemistry*, **44**, 6024–6033.
Yonath, A. (2003). *Chembiochem*, **4**, 1008–1017.