

Crystal structure of Thermus thermophilus tRNA m1A58 methyltransferase and biophysical characterization of its interaction with tRNA.

Pierre Barraud, Béatrice Golinelli-Pimpaneau, Cédric Atmanene, Sarah Sanglier, Alain van Dorsselaer, Louis Droogmans, Frédéric Dardel, Carine Tisné

▶ **To cite this version:**

Pierre Barraud, Béatrice Golinelli-Pimpaneau, Cédric Atmanene, Sarah Sanglier, Alain van Dorsselaer, et al.. Crystal structure of Thermus thermophilus tRNA m1A58 methyltransferase and biophysical characterization of its interaction with tRNA.. Journal of Molecular Biology, 2008, 377 (2), pp.535-50. 10.1016/j.jmb.2008.01.041 . hal-00275721

HAL Id: hal-00275721
https://hal.science/hal-00275721

Submitted on 25 Apr 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Crystal structure of *Thermus thermophilus* tRNA m^1A_{58} methyltransferase and biophysical characterization of its interaction with tRNA.

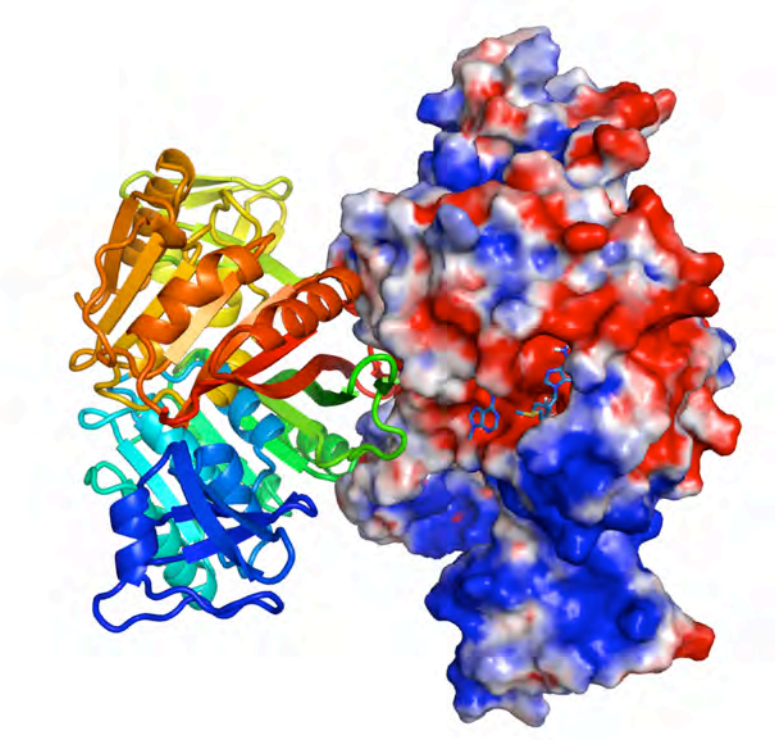
Pierre Barraud[1], Béatrice Golinelli-Pimpaneau[2], Cédric Atmanene[3], Sarah Sanglier[3], Alain Van Dorsselaer[3], Louis Droogmans[4], Frédéric Dardel[1] & Carine Tisné[1*]

[1] Laboratoire de Cristallographie et RMN biologiques, Université Paris-Descartes, CNRS UMR 8015, 4 avenue de l'Observatoire, 75006 Paris, France
[2] Laboratoire d'Enzymologie et Biochimie Structurales, CNRS Bâtiment 34, 1 avenue de la Terrasse, 91190 Gif-sur-Yvette, France
[3] Laboratoire de Spectrométrie de Masse Bio-Organique, IPHC-DSA, ULP, CNRS UMR7178, 25 rue Becquerel, 67087 Strasbourg, France
[4] Laboratoire de Microbiologie, Institut de Recherches Microbiologiques Wiame, Université Libre de Bruxelles, avenue E. Gryson 1, B-1070 Bruxelles, Belgium

*Address correspondence to: Dr. Carine Tisné, phone: +33 1 53 73 15 72, fax: +33 1 53 73 99 25, e-mail: carine.tisne@univ-paris5.fr

Short running title : Biophysical characterization of the interaction of *th*TrmI with tRNA

The abbreviations used are: AdoHcy, S-adenosyl-L-homocysteine; AdoMet, S-adenosyl-L-methionine; ESI-MS, Electrospray ionization mass spectrometry; MTases, methyltransferases; TrmI, m^1A_{58} tRNA methyltransferase; *th*TrmI, *Thermus thermophilus* TrmI; *myco*TrmI, *Mycobacterium tuberculosis* TrmI.



Abstract

Methyltransferases (MTases) from the TrmI family catalyse the S-adenosyl-L-methionine (AdoMet)-dependent N_1-methylation of tRNA adenosine 58. The crystal structure of *Thermus thermophilus* m^1A_{58} tRNA MTase (*th*TrmI) in complex with S-adenosyl-L-homocysteine was determined at 1.7 Å resolution. This structure is closely related to that of *Mycobacterium tuberculosis* TrmI (*myco*TrmI), and their comparison enabled us to enlighten two grooves in the TrmI structure that are large enough and electrostatically compatible to accommodate one tRNA per face of TrmI tetramer. We have then conducted a biophysical study based on electrospray ionization mass spectrometry, site-directed mutagenesis and molecular docking. First, we confirmed the tetrameric oligomerisation state of TrmI and we showed that this protein remains tetrameric upon tRNA binding with formation of complexes involving one to two molecules of tRNA per TrmI tetramer. Secondly, three key residues for the methylation reaction were identified, the universally conserved D170, and two conserved aromatic residues Y78 and Y194. We then used molecular docking to position a N_9-methyladenine in the active site of TrmI. The N_9-methyladenine snugly fits in the catalytic cleft where the side chain of D170 acts as a bidentate ligand binding the amino moiety of AdoMet and the exocyclic amino group of the adenosine. Y194 interacts with the N_9-methyladenine ring whereas Y78 can stabilize the sugar ring. From our results, we propose that the conserved residues that form the catalytic cavity (D170, Y78 and Y194) are essential to fashion an optimized shape of the catalytic pocket.

Keywords: TrmI, m1A58 methyltransferase, X-ray structure, noncovalent mass spectrometry, Protein-RNA interactions

Introduction

Functional tRNAs carry a number of chemically modified nucleosides that are formed enzymatically after transcription, during the tRNA maturation process. To date, more than 90 different modifications have been identified in tRNAs from various organisms.[1] Among these post-transcriptional nucleoside modifications, N_1-methyladenosine (m^1A) is found at a highly conserved A_{58} position in the TΨC loop of many tRNAs in the three domains of life (*Bacteria*, *Archaea* and *Eukarya*). This modification occurs infrequently in bacteria and, for instance, is absent in *E. coli* tRNAs. On the contrary, it is common in the tRNAs of most eukaryotes and archaea, and this modification seems to be essential for some organisms. Actually, in the yeast *Saccharomyces cerevisiae*, mutants defective in N_1-methylation of A_{58} are non-viable.[2] Likewise, in *Thermus thermophilus*, gene disruption studies have shown that m^1A_{58} modification is required for growth of this bacterium at high temperatures.[3]

S-Adenosyl-L-methionine (AdoMet) dependent methyltransferases (MTases) are involved in a wide variety of biological processes involving methylation of nucleic acids, proteins, phospholipids and small molecules using the ubiquitous methyl donor AdoMet. Recently, two essential genes of *S. cerevisiae,* GCD10 and GCD14 (renamed TRM6 and TRM61), were identified to encode the two types of subunits of the yeast m^1A_{58} MTase.[4] The recombinant enzyme behaves as an $\alpha_2\beta_2$ heterotetramer, Trm61p being responsible for AdoMet binding and presumably catalysis of the methyl transfer reaction,[2] whereas both types of subunits are essential for tRNA-binding.[5] Indeed, the purified recombinant Trm61p subunit, which binds AdoMet, cannot bind tRNA *in vitro* in the absence of Trm6p.[2] Interestingly, Trm61p was found to be closely related to a group of prokaryotic proteins which share not only the same AdoMet-binding site but also other highly conserved motifs. Therefore, the corresponding prokaryotic proteins were presumed to act as m^1A_{58} MTases.[6] This hypothesis was demonstrated experimentally for the bacterial orthologs in *Thermus thermophilus* and *Mycobacterium tuberculosis* and for the archaeal ortholog in *Pyrococcus abyssi*.[3,7-9] These proteins are α_4 homotetramers of Trm61p-like proteins, hereafter called TrmI. Interestingly, orthologs of the Trm6p protein could only be found in eukaryotes. It was shown that Trm6p, despite the absence of the characteristic MTases motifs, is structurally and evolutionary related to Trm61p suggesting that the eukaryotic m^1A_{58} MTases evolved by gene duplication and speciation to form a heteromultimeric protein, whereas their prokaryotic orthologs remained homomultimers.[6]

Despite the low level of sequence similarity between the various families of AdoMet-MTases, most of them contain a structurally highly conserved catalytic AdoMet-binding domain organized in a Rossman-like fold (For a review see [10]). Only one crystal structure of an m^1A_{58} tRNA MTase has been reported to date, that of the *M. tuberculosis* Rv2118c protein (*myco*TrmI) in complex with S-adenosyl-L-methionine.[7] It shows that the catalytic C-

terminal domain is very similar to that of other AdoMet-dependent MTases whereas the N-terminal domain which is mainly composed of β-sheets is not found in other MTases of known structure.

In the present work, we have first solved the crystal structure of the m^1A_{58} tRNA MTase from *T. thermophilus* (*th*TrmI) at 1.7 Å resolution in complex with S-Adenosyl-L-homocysteine (AdoHcy), the product after methyl transfer. We then focused our attention on the tRNA-binding properties of *th*TrmI using several biophysical techniques, including electrospray ionization mass spectrometry (ESI-MS), site-directed mutagenesis, steady-state kinetic assays, fluorescence spectroscopy and molecular docking. ESI-MS has demonstrated its particular suitability for the investigation of noncovalent complexes.[11-13] More particularly, so-called noncovalent ESI-MS has been extensively used to characterize supramolecular assemblies involving oligonucleotides in interaction with drugs and proteins (For reviews see [14-17]). As protein:RNA systems are partly driven by electrostatic-based interactions which are strongly enhanced in the gas phase, ESI-MS is well suited for their characterization.[18,19] Indeed, numerous studies of protein-RNA complexes have revealed a strong agreement between mass spectrometric gas-phase measurements and results obtained by solution phase techniques.[20-24] In our work, we used noncovalent ESI-MS: i) to unambiguously assess the oligomerisation state of *th*TrmI ; ii) to monitor the *th*TrmI oligomerisation state upon tRNA binding and iii) to determine the stoichiometry of the *th*TrmI-tRNA complexes. In addition, a mutagenesis study was performed to identify residues potentially crucial for the methylation reaction. We thus showed that the universally conserved D170 residue, together with two conserved aromatic residues, Y78 and Y194, which line the catalytic pocket are key residues for the enzymatic catalysis. Actually, D170A and Y78A variants are severely altered in their catalytic efficiency. Finally, we have used molecular docking to position a N_9-methyladenine in the active site of TrmI to further investigate the possible role of these residues.

Results and Discussion

Structure determination of *T. thermophilus* TrmI

Production and purification of *T. thermophilus* TrmI (*th*TrmI) have been described by Droogmans and coworkers.[3] However, the recombinant His_6-tagged protein exhibited a low solubility (< 0.5 mg/mL) unless high salt concentrations were added (0.2 M imidazole-HCl and 0.5 M KCl). This was a serious issue for crystallization assays, in which precipitating solutions must be added to concentrated protein. In order to improve solubility, the N-terminal His_6-tag of *th*TrmI was cleaved with thrombin. Indeed, after His_6-tag removal, *th*TrmI became highly soluble (> 10 mg/mL), even at moderate ionic strengths (150 mM). This type of behaviour has already been reported for some His-tagged proteins.[25] We then carried out crystallization assays on the complex between *th*TrmI, AdoHcy and *E. coli* $tRNA_i^{Met}$. The latter tRNA was selected because it is devoid of modification on A_{58} and is a substrate of the *T. thermophilus* enzyme (see paragraph on steady-state kynetic assays). Although the tRNA was present in all crystallization conditions, crystals only contained the free *th*TrmI protein in complex with AdoHcy. Table 1 summarizes the data collection and refinement statistics. The structure of *T. thermophilus* TrmI protein was then solved by molecular replacement with the structure of m^1A_{58} tRNA MTase from *M. tuberculosis* TrmI (PDB entry code 1I9G)[7] using PHASER[26] (Z score = 23.7). Actually, *th*TrmI shares 39% identical and 13% similar residues with *myco*TrmI. The model was then refined to 1.7 Å, to a final R_{free} factor of 20.9% (Table 1). The last four residues at the N-terminus were too poorly ordered to be included, and therefore the final model consisted of 251 residues per monomer. The asymmetric unit contained a dimer of *th*TrmI. However, the symmetry-related subunits interact extensively leading to a tetrameric organization of *th*TrmI (Figure 1(b)).

Overall structure of *th*TrmI

The *th*TrmI monomer is composed of two distinct domains connected by an extended linker containing a short α-helix (helix 2, Figure 1(a)). The large C-terminal domain (residues 63-255) adopts a typical class I MTase fold which consists of a central seven-stranded β sheet (β_1 to β_6 and β_8 – Figure 1(a)) flanked by α helices on both sides (α_A to α_E – Figure 1(a)).[10] The first five strands of the β sheet (β_1 to β_5 on Figure 1(a)) are parallel whereas the last two (β_6 to β_8) are antiparallel. The C-terminal domain contains the AdoMet binding site and the catalytic pocket. The smaller N-terminal domain (residues 5-62) is largely a β structure composed of six β-strands (β_A to β_F Figure 1(a)) and one small α helix (α_1).

The *th*TrmI tetramer is organised as a dimer of tight dimers. Two monomers interact via their extended β_6 strands which form a cross-subunit, antiparallel β-sheet (Figure 1(b), pale and dark blue). This dimer is further stabilized by ionic interactions between helices α_2 and α_3 from each subunit (Figure 1(b)). A pair of β_6 strands from each dimer packs on the related

β_6 strands of the other dimer to form a central anti-parallel β-sheet structure. The tetramer interaction is limited to this β-sheet region and stabilized by both ionic and hydrophobic interactions, mainly a salt bridge between E220 and R224, hydrophobic interactions between W226 and F245, and between H234 and F237 from one dimer to H242 from the other dimer. The salt bridge and the hydrophobic interactions between W226 and F245 have been also observed on the *myco*TrmI structure[7] and involve conserved residues.

In the tight dimer found in the asymmetric unit, the two cofactor-binding sites are occupied by AdoHcy molecules in identical conformations, similar to that of AdoMet in the *M. tuberculosis* TrmI structure. The binding cleft is formed by residues belonging to conserved motifs, named I, II, III, IV, (Figures 1(c), 1(d) and 7(a)).[27] The adenine ring of the AdoHcy moiety is bound by motifs II, III, and IV via water-mediated hydrogen bonds between the exocyclic amino group and residues E155 and E173, and between the N_7 position of the adenine and the backbone amide of E173 (Figure 1(c)). Adenine binding is also reinforced by hydrophobic contacts between its aromatic ring and residues L171, V177, K153, L154, Y124, A126 and A100. The ribose moiety of AdoHcy is bound by motif II, essentially by hydrogen bonds involving the 2' and 3' hydroxyl groups and residues E125 and H130. The amino-acid part of the cofactor is mainly bound by motif I. Interestingly, an aspartate side chain that could be involved in the methylation reaction, the carboxylate chain of D170 from motif IV, participates to a direct polar interaction with the ammonium group of AdoHcy. The GXGXGG pattern of motif I binds to the amine and the carboxylate through a tight network of water-mediated hydrogen bonds involving the backbone (Figure 1(c)).

Comparison with *myco*TrmI structure and tRNA recogition

Globally, the overall structure of *th*TrmI closely resembles that of *myco*TrmI. Superimposition of *T. thermophilus* and *M. tuberculosis* TrmI structures gives a root mean square deviation of 4.47 Å for Cα, the largest variation occurring in the β-strands (β_6, β_7, β_8) that protrude out of the structure. When comparing the active sites more accurately, that of *myco*TrmI is narrower by about 1 Å than that of *th*TrmI. Moreover, the protruding R249 in *myco*TrmI makes like a supplementary step at the exit of the pocket that does not exist in *th*TrmI as R249 is replaced by a shorter side chain residue (V240). As a consequence, the accessible surface of Y78 is twice larger in *th*TrmI than that of Y84 in *myco*TrmI. All these features seem to make the methyl donor less easily accessible in *myco*TrmI. Overall, in both structures, the adenine should enter deeply in the binding cleft to have access to the methyl donor therefore implicating a major deformation of the phosphodiester backbone around A_{58}. It is also highly interesting to compare electrostatic potentials of their molecular surfaces (Figure 2). These calculations uncover that both proteins have two linked grooves of positive electrostatic potentials indicated by large blue surfaces on Figure 2 that are large enough to accommodate double-helical RNA. The first groove that covers the β-sheet region forming

the tetrameric architecture is less positively charged in *th*TrmI than in *myco*TrmI; whereas the second groove that encompasses the N-terminal domain of one TrmI subunit and that looks like an open hand is more positively charged in *th*TrmI than in *myco*TrmI. Several residues with basic side chains are conserved among the TrmI protein family, namely K13 (R in some sequences), R15 (K), R72, R127, R217 and R229 (K) together with the histidines involved in the formation of the tetramer H234 and H242. R15 and R217 are only conserved in bacterial TrmI proteins. These grooves have the dimensions to accommodate the acceptor arm (groove 1, dimensions: 25 Å large and 45 Å long) and the anticodon one (groove 2, dimensions: 25 Å large and roughly 30 Å long) without steric clashes, enabling us to position the adenosine 58 near the catalytic pocket (manual docking, data not shown). The grey line in Figure 2(a) indicates a possible position of the helix axis of tRNA. Numerous clashes occur within the T-arm that undoubtedly undergoes huge conformational changes that probably involves the entire T-arm and not only a simple flip of the adenosine 58. These changes are difficult to predict and, for this reason, we could not pursue this docking. In conclusion, the conservation of two grooves of positively charged surfaces supports the idea that TrmI binds tRNA as a tetramer and that two molecules of tRNA can interact simultaneously with the TrmI tetramer. The fact that, in eukaryotes, the enzyme has evolved from an α_4 homotetramer to an $\alpha_2\beta_2$ heterotetramer is compatible with a two tRNA per TrmI tetramer stoichiometry, one tRNA being expected to interact with an $\alpha\beta$ subsystem.

***th*TrmI is a tetramer and binds up to two tRNA molecules as a tetramer.**

Close inspection of the crystal structures of *th*TrmI and *myco*TrmI[7] revealed that the tetramers are formed by two pairs of extensively interacting subunits stabilized by relatively small numbers of contacts between the two dimers. Given the architecture of TrmI, it was of interest to investigate the oligomerization state of *th*TrmI alone and upon tRNA binding, and to determine the binding stoichiometry of the *th*TrmI-tRNA complexes. Previously reported gel filtration experiments have indicated that both *myco*TrmI and *th*TrmI should form tetramer in solution.[3,9] However, in the case of *th*TrmI, due to an inherent limited precision of this technique, it is difficult to unambiguously state whether the protein is tetrameric or pentameric.[3] In this study, we used ESI-MS, which has been proven to be a valuable technique for the determination of oligomerization states of noncovalent assemblies[28-31] to first confirm the oligomerization state of *th*TrmI. Figure 3(a) presents the ESI mass spectrum obtained for *th*TrmI under non-denaturing conditions and carefully controlled operating conditions (see Materials and Methods): a single ion distribution is observed in the m/z 4000-5000 range. With a measured molecular weight (MW) of 115450 ± 5 Da, this distribution can be assigned to the +23 to +30 charge states of a *th*TrmI tetramer (theoretical MW = 115453 Da). Thus, ESI-MS allowed us to unambiguously assess the tetrameric nature of the protein in agreement with our crystallographic results.

The tRNA binding stoichiometry was first investigated by gel retardation assay. As shown on Figure 4, the free tRNA band progressively disappears upon addition of increasing amounts of *th*TrmI evidencing thus the existence of *th*TrmI-tRNA interactions. Moreover, the free tRNA band completely disappears at 1:2 *th*TrmI:tRNA ratio, substantiating the formation of complexes involving two molar equivalents of *th*TrmI monomer per tRNA. To further confirm gel retardation experiments, we also used the potentialities of noncovalent ESI-MS to investigate the oligomerization state of *th*TrmI upon tRNA binding and to determine the tRNA binding stoichiometry. Titration experiments, monitored by noncovalent ESI-MS, involving a fixed concentration of *th*TrmI and increasing amounts of $tRNA_i^{Met}$, revealed the presence of three ion distributions (Figure 3(b-d)). The first one in the m/z 4000-5000 region with a molecular mass of 115450 ± 5 Da is related to the *th*TrmI tetramer. For the second ion distribution in the m/z 5000-5900 region, the molecular mass of 140347 ± 6 Da corresponds to the $(thTrmI)_4$:$(tRNA)_1$ complex. Finally, a third ion distribution within the m/z 5900-6800 range with a molecular mass of 165455 ± 10 Da refers to the $(thTrmI)_4$:$(tRNA)_2$ complex. Even a ten-fold molar excess of tRNA per *th*TrmI tetramer does not lead to the detection of complexes with higher tRNA binding stoichiometries. Interestingly, upon increase of the tRNA concentration from 1.5 to 10 molar equivalents per *th*TrmI tetramer, the relative abundances of $(thTrmI)_4$:$(tRNA)_1$ and $(thTrmI)_4$:$(tRNA)_2$ complexes statistically increase, which is in favour of non-cooperative tRNA binding system and further suggests the presence of two independent tRNA binding sites.[32] The same experiments were also carried out in presence of AdoHcy (data not shown) and no effect either on the binding stoichiometry of the complex or on the binding affinity of the tRNA for *th*TrmI was observed.

As ESI-MS detects species in the gas phase of the mass spectrometer, control experiments are always needed to ensure that mass spectra faithfully reflect the behaviour in solution.[33] Thus, since electrostatic interactions are sensitive to the ionic strength of the medium,[34,35] experiments were carried out at different ammonium bicarbonate concentrations (Figure 5). Decreasing the ammonium bicarbonate concentration from 200 mM to 100 mM displaces the equilibrium towards the formation of $(thTrmI)_4$:$(tRNA)_1$ and $(TrmI)_4$:$(tRNA)_2$ complexes (Figure 5(a-b)). However, the signal intensities were dramatically lower when 100 mM ammonium bicarbonate was used rather than 200 mM buffer, mainly because of protein precipitation. Therefore, no ESI-MS analysis could be performed at buffer concentrations lower than 100 mM. While both $(thTrmI)_4$:$(tRNA)_1$ and $(thTrmI)_4$:$(tRNA)_2$ complexes are favoured at low ionic strengths, the use of higher salt concentrations destabilizes these assemblies (Figure 5(c-d)), leading even to a complete dissociation of *th*TrmI:tRNA complexes into *th*TrmI tetramer at 1250 mM ammonium bicarbonate (Figure 5(d)). The fact that mass spectra do reflect expected changes induced by modification of the solution conditions is a further evidence that allows us to definitely rule out the possibility that the TrmI-tRNA complexes result from an artefact of the technique. Moreover, as described in

other publications,[22,36] an additional ESI-MS control experiment was performed in strictly identical experimental and operating conditions with a non-substrate RNA as a negative control. The absence of any *th*TrmI-control RNA complex on ESI mass spectra even in presence of a five-fold molar excess of control RNA (Figure 6) indicates that the gas phase detection of the TrmI-tRNA complexes arises from a specific recognition in solution and not from any gas phase artefact.

In conclusion, noncovalent ESI-MS results clearly support that *th*TrmI remains tetrameric upon tRNA binding and that *th*TrmI binds up to two molecules of tRNA.

The D170A and Y78A variants of *th*TrmI are severely altered in their catalytic efficiency.

On the basis of multiple sequence alignments of the TrmI family members (Figure 7(a)) and on the comparison of the crystallographic structures of TrmI from *M. tuberculosis*[7] and *T. thermophilus*, we chose to mutate three conserved residues, D170, Y194 and Y78 in order to study their involvement in adenosine binding and/or in catalysis. Figure 7(b) presents the *th*TrmI residues that are conserved or semi-conserved in the catalytic pocket, except the conserved residues that bind to AdoMet since their role in the catalysis mechanism is known. Yet, the universally conserved AdoMet-binding D170 was selected as it forms the back of the catalytic pocket and may also bind the adenine 58 ring of the tRNA substrate (Figure 7(b)). Y194 and Y78 were mutated because they respectively form the left side and the floor of the adenosine-binding pocket in the 3D structure (Figure 7(b)) and are conserved as aromatic residues among m^1A_{58} tRNA MTases, belonging respectively to motifs V and X (Figure 7(a)). Therefore, these aromatic residues could be involved in the stabilization of the target adenine that needs to be flipped out of the tRNA structure to be methylated. P196 that is involved in making the shape of the catalytic cavity was not selected for mutation, because in *M. tuberculosis*, it is naturally replaced by an alanine. Therefore, we individually mutated Y78, D170 and Y194 to alanines. We expressed and purified the variant proteins as described in Materials and Methods section. We then determined the kinetic and RNA-binding parameters for wild-type and variant proteins (Table 2).

First, we investigated whether the mutations altered tRNA binding by the enzyme. Interaction with tRNA induces a quenching of the intrinsic tryptophan fluorescence of the protein of about 40 %. We therefore used fluorimetric titrations to determine the apparent dissociation constant between *th*TrmI variants and $tRNA_i^{Met}$ (Table 2). All mutants showed quite unchanged dissociation constants (K_d about 15 ± 3 nM), thereby indicating that they retained a native folding and similar RNA binding ability compared to the wild-type enzyme. We then analysed their catalytic properties by determining their kinetic parameters k_{cat} and K_M for both AdoMet and tRNA substrates. The overall results are summarized in Table 2. Interestingly, all mutants retain some catalytic activity, but are altered to various extents,

essentially in their catalytic turnover k_{cat}. The D170A mutant retains some activity, but its k_{cat} value is severely reduced, by a factor of about 300. D170A also shows an increased K_M for AdoMet substrate in agreement with the crystal structure. It seems therefore to be the most important residue for the catalysis of the reaction. The Y78A mutant leads to a roughly 20-fold decrease of the catalytic turn-over constant. Its binding constant for the methyl donor is unchanged compared to the wild-type protein. The Y194A mutant shows a smaller decrease of k_{cat}, by a factor of 3 (Table 2). The K_M value for AdoMet is increased suggesting that this residue is involved in cofactor binding. Yet, in the crystal structure, it is not involved in a direct interaction with AdoHcy in contrast to D170. The mutation of Y194 to an alanine probably slightly destabilizes the catalytic pocket or indirectly weakens AdoMet binding. In conclusion, D170A is the variant that is most largely altered in its catalytic efficiency followed by Y78A and then Y194A. And, therefore, D170, Y78 and Y194 are key residues in the catalytic mechanism.

The substrate adenine snugly fits in the active site pocket of *th*TrmI.

To investigate the positioning of the target A_{58} into the catalytic pocket of *T. thermophilus* and *M. tuberculosis* TrmI, we decided to locate a N_9-methyladenine in the catalytic pocket of these proteins by molecular docking. We docked a N_9-methyladenine and not an adenosine or an AMP molecule since it is impossible to predict the conformation of the sugar of this flipped-out nucleotide. Moreover Y78 that makes the exit of the catalytic pocket and which may interact with the ribose of the adenosine 58 can also flip to prevent from steric clashes with the A_{58} ribose. As the molecules of AdoHcy present in the active sites of *th*TrmI have identical conformations to the AdoMet co-factor present in *myco*TrmI, we modelled an AdoMet molecule in the catalytic pocket of *th*TrmI by superimposing the AdoMet from *myco*TrmI with AdoHcy in *th*TrmI. We used the program HADDOCK[37] that can take into account mutagenesis results by introducing data as Ambiguous Interaction Restraints (AIRs) to drive the docking process. An AIR is defined as an ambiguous distance between all residues shown to be involved in the interaction. Ambiguous intermolecular distance restraints were applied between D170, Y194, Y78 and the AdoMet cofactor on the one hand, and the N_9-methyladenine, on the other hand. A distance restraint between the N_1 atom of N_9-methyladenine and the methyl of AdoMet was also added. The model resulting from the docking with *th*TrmI is presented in Figure 8(a) and shows that the N_9-methyladenine fits snugly in the cavity where methylation takes place. The same result was obtained for *myco*TrmI (data not shown). Moreover, the methyl group introduced to substitute the ribose is pointing above Y78 (Figure 8(b)) enabling us to add a ribose without steric hindrance. Therefore, Y78 is likely to strongly interact with the ribose of the adenosine 58 probably participating to stabilization of the flipped-out conformation. The distance between the N_9-methyladenine N_1 and the AdoMet methyl group is 2.3 Å. This corresponds to a distance

between the N_9-methyladenine N_1 and the AdoMet sulphur atom of 4.1 Å. This is in keeping with distances observed in co-crystal structures of other endocyclic purine N-MTases (distances between 3.6 Å and 4.5 Å).[38] The docked model of the N_9-methyladenine ring suggests a binding mode involving essentially Van der Waals interactions with Y194, and a polar contact between the exocyclic amino group of the purine and the carboxylate of D170 (Figure 8(b)). Therefore, the side chain of D170 could act as a bidentate ligand binding the amino terminal moiety of AdoMet and the N_6 exocyclic amino group of the adenine ring substrate. Thus, the sulphur-methyl bond of AdoMet is almost coplanar with the N_9-methyladenine ring and is pointing to the same direction as the N_1 lone-pair orbital, which can be of significant importance for the chemical mechanism. We thus propose that D170 could minimize the transition state energy by positioning both the AdoMet and the target base in an optimal orientation for the methylation reaction. D170 could also deprotonate the N_6 group to activate the N_1 lone-pair orbital in order to attack the methyl group of AdoMet and to generate the exocyclic imino tautomer of N_1-methyladenosine as the initial product. The nucleophilic attack leads to the m^1A basic form which has to be subsequently protonated to give the m^1A cationic form, stable at physiological pH (pKa = 9.3).[39] As the deprotonation of N_6 of m^1A is easier than the deprotonation of N_6 of the non-methylated adenine ring (pKa = 16.7),[40] it seems more reasonable that the deprotonation and the nucleophilic attack occur in the same elementary step. The model perfectly situates D170 to serve that role, which is most definitely consistent with the kinetic results.

The chemical mechanism of adenine N_1 methylation of RNA had not yet been investigated in contrast to the N_1 methylation of guanine.[41] In the case of m^1G RNA methyltransferase (TrmD),[41] an aspartate residue was shown to be involved in the deprotonation step of N_1 of guanine. The two processes are likely to be different, as under physiological conditions, the N_1 of guanine is protonated but that of adenine is not. DNA methyltransferases have been extensively studied, but the m^1A modification is not encountered in DNA. The methylation of the N_6 position of adenine has been investigated in DNA (for structural studies see [42-44]) and more recently in rRNA.[45,46] Interestingly, the main catalytic residue in m^6A RNA or DNA MTases has been clearly identified by mutagenesis studies to be the D/N/S residue of the conserved pattern (D/N/S)PP(Y/F/W) which belongs to MTase motif IV (for reviews on DNA MTases see [47,48]). In known structures of m^6A DNA MTases (for example M•*Taq*I, PDB code 2ADM) or m^6A RNA MTases (ErmC', PDB code 1QAN) in complex with cofactor (AdoMet or AdoHcy), the position of the D/N/S residue of motif IV relative to the cofactor is closely similar to that of D170 residue in the *T. thermophilus* TrmI crystal structure. Additionally, two other residues have been identified to be important in catalysis in m^6A DNA methyltransferases. These residues are aromatic and stabilize the flipped-out target base and/or act as cation-π catalysts. The first one is the aromatic Y/F/W residue of motif IV, which is proposed to stabilize the cationic transition state by cation-π interaction.[49,50] And the

second one is the aromatic residue Y/F/W of motif VIII, which stabilizes the flipped base outside the DNA helix.[49,50] These aromatic residues are conserved in motifs IV and VIII in m^1A tRNA MTases (Figure 7(a), symbol **o**), but they are not part of the catalytic pocket in the TrmI structure. Strikingly, two other aromatic residues, *i.e.* Y78 and Y194, are located in the active site cavity of *th*TrmI (Figure 7(b)) at relative positions with respect to the AdoMet cofactor similar to those of the aromatic residues listed above for DNA MTases. These residues are also conserved as aromatic residues among m^1A_{58} tRNA MTases and belong respectively to motifs X and V (Figure 7(a)). The noticeable structural similarity between the catalytic pocket of m^1A tRNA MTases and m^6A DNA MTases suggests a similar role in catalysis for these residues. Y78 is thus the most probable candidate for the stabilization of the flipped-out adenosine given its position in the catalytic pocket and the dramatic decrease of k_{cat} of the Y78A mutant. Then, Y194 could act as a cation-π catalyst to stabilize a cationic transition state. These roles are also in agreement with our docking results.

Coordinates and structure factors for *T. thermophilus* TrmI protein in complex with AdoHcy have been deposited at the Protein Data Bank with accession code 2PWY.

Materials and Methods

Expression and purification of *th*TrmI for structural studies

Recombinant *T. thermophilus* TrmI protein (255 residues) was overexpressed and purified as previously described.[3] The N-terminal His_6-tag was removed by thrombin cleavage (25 units of thrombin per mg of protein) performed overnight at ambient temperature. The *th*TrmI protein was then further purified by gel filtration on a HiLoad 26/60 Superdex 200 prepgrade chromatography column (Amersham Biosciences) equilibrated in 50 mM Tris-HCl buffer (pH 8.0), 100 mM KCl. Fractions containing pure recombinant protein were pooled, concentrated to 8-10 mg/mL using Amicon® Ultra (Millipore) and stored at 4 °C.

Crystallization and crystal structure determination

Crystallization was performed at 19 °C using the sitting drop vapour diffusion method. Protein samples were prepared at 3 mg/mL in 50 mM Tris-HCl buffer (pH 8.0) containing 100 mM KCl, 2 mM AdoHcy and 1 equivalent of $tRNA_i^{Met}$. Drops of 1 µL were prepared by using a Cybi-Disk robot system that mixes equal volume of protein and reservoir solutions. Reservoir volumes of 100 µL were used. Crystals of *th*TrmI were obtained in 2 M ammonium sulphate, isopropanol 5 % (v/v) (condition number 5 of Hampton Research Crystal Screen II kit). The crystals were harvested, soaked in a cryoprotectant solution (2.2 M ammonium sulphate, isopropanol 5 % (v/v), glycerol 20 % (v/v)) and flash-frozen in liquid nitrogen before data collection. Diffraction data were collected at beam line ID14-1 of the European Synchrotron Radiation Facility (ESRF, Grenoble, France).

All crystallographic calculations were performed using the CCP4 suite version 6[51] as implemented in the graphical user interface.[52] X-ray diffraction data were processed using MOSFLM[53] and scaled with SCALA.[54] The structure of *th*TrmI was solved by molecular replacement using the program PHASER[26] and the structure of m^1A_{58} tRNA MTase from *M. tuberculosis* (PDB entry code 1I9G) as a model. In the model, non-conserved residues were truncated to alanine. Model building of *th*TrmI was first performed with ARP/wARP[55] using the warpNtrace automated procedure. Restrained refinements of the structure were performed with the program REFMAC5.[56] Model and maps visualisations for manual reconstruction were performed with the program COOT.[57] Solvent molecules were automatically added using ARP waters module implemented in REFMAC5. In the last stages of refinement, TLS parameters were refined[58] using one group for each domain of the protein, *i.e.* the large C-terminal domain and the small N-terminal domain.

We computed the electrostatic surface potential using APBS Tools, an interface for performing APBS electrostatics calculation[59] from within PyMOL (http://www.pymol.org).

Expression and purification of $tRNA_i^{Met}$

The *E. coli* $tRNA_i^{Met}$ was overexpressed from plasmid $pBSTtRNA_f^{Met}$ in *E. coli* JM101TR using a protocol derived from that of Meinnel and Blanquet.[60] Briefly, after phenol extraction of RNAs from bacteria, total tRNA was separated by gel filtration on a HiLoad 26/60 Superdex 75 prepgrade chromatography column (Amersham Biosciences) equilibrated in 25 mM Tris-HCl pH 7.0. The overexpressed $tRNA_i^{Met}$ was then separated from other tRNAs by an anion exchange step (Resource Q column) equilibrated in 25 mM Tris-HCl pH 7.0. tRNAs were eluted using a 350 mM to 550 mM NaCl gradient in the same buffer. The fractions containing the purified $tRNA_i^{Met}$ were pooled together, dialysed against 50 mM Tris-HCl, 100 mM KCl, pH 8.0, concentrated using Amicon® Ultra (Millipore) and stored at -20 °C.

Electrospray ionization mass spectrometry

Noncovalent mass spectrometry experiments were performed on an electrospray time-of-flight mass spectrometer (LCT, Waters). Samples were continuously infused into the mass spectrometer with a syringe pump (Harvard Apparatus) at a flow rate of 6 µl/min. Mass spectra were acquired in the positive ion mode on the mass range 2000-8000 m/z after calibration with the multiply charged ions produced by horse heart myoglobin diluted to 2 µM in a 1:1 water/acetonitrile mixture (v/v) acidified with 1% (v/v) formic acid. Deconvoluted mass spectra were obtained using the Transform algorithm of MassLynx 4.0 software. Instrumental parameters were optimized to get the best compromise between ion desolvatation, ion transmission and preservation of noncovalent complexes during the ionization and desorption processes. This optimization concerned especially the pressure in the interface (Pi) and the accelerating voltage (Vc) which were set to 6.1 mbar and 120 V, respectively.

Purity and homogeneity of *th*TrmI were checked in denaturing conditions by diluting the protein to 5 µM in a 1:1 water/acetonitrile mixture acidified with 1% (v/v) formic acid. A good agreement was found between the measured molecular mass (28863.0 ± 0.3 Da) and the mass calculated from the amino acid sequence (28863.3 Da). For experiments performed in non-denaturing conditions, different ammonium bicarbonate concentrations ranging from 100 mM to 1250 mM were tested. A concentration of 200 mM ammonium bicarbonate was chosen as the best compromise to provide good quality ESI mass-spectra without extensive *th*TrmI:tRNA complexes dissociation (which takes place at high ionic strengths) or protein aggregation (occurring at low ionic strengths). Thus, this buffer concentration was used to study both *th*TrmI oligomerization state and its tRNA-binding properties. All experiments were carried out at 20 µM *th*TrmI (tetramer concentration) whereas *E. coli* $tRNA_i^{Met}$ concentration (molecular weight MW = 24888 Da) was increased up to 200 µM for titration experiments. To study the influence of AdoHcy on the *th*TrmI-tRNA complexation equilibrium, *th*TrmI and AdoHcy, both at 20 µM (tetramer concentration), were incubated in

the ammonium bicarbonate buffer in presence or in absence of tRNA (30 μM). Finally, a non-substrate control RNA of 59 nucleotides (MW = 19255 Da) was also used in a five-fold molar excess relative to the *th*TrmI tetramer concentration in order to investigate the specificity of the *th*TrmI-tRNA interactions. All solutions were incubated for 15 minutes at room temperature before analysis under non-denaturing conditions.

Non-denaturing agarose gel electrophoresis

Samples were prepared by mixing constant amount of *E. coli* $tRNA_i^{Met}$ (4 μg) with increasing amounts of *th*TrmI protein (from 0 to 1 molar equivalent of *th*TrmI tetramer) in loading buffer (Tris-HCl pH 8.0 50 mM, KCl 100 mM, glycerol 18 % (v/v)). Samples were kept at 4 °C for 12 hours and loaded on a 2 % (w/v) agarose gel. Migration was performed for one hour in Tris-Acetate buffer pH 8.0 at 4 °C under an electric field of 100 V. Bands were visualized by UV shadowing.

Site-directed Mutagenesis

Alanine site-directed mutagenesis was carried out using the Quickchange® kit and protocol (Stratagene). All oligonucleotides for mutagenesis were designed to have appropriate pairing stability (fusion temperature near 80 °C) and were between 27 to 32 nt long. All mutant genes were fully sequenced and contained only the desired substitutions.

Expression and purification of wild-type and mutant *th*TrmI proteins for kinetic studies

The wild type and mutant proteins were overexpressed and purified as for structural studies except that the thrombin cleavage was omitted. Protein concentrations were evaluated with the Bradford method. The proteins were then diluted in 50 mM Tris-HCl pH 8.0 to a final concentration of 0.25 mg/mL and stored at 4 °C.

Fluorescence titrations of *th*TrmI by $tRNA_i^{Met}$

Fluorescence measurements were performed at 20 °C on a JASCO spectrofluorimeter. Excitation and emission wavelengths were 280 nm and 343 nm respectively. The excitation and emission bandwidths were 5 nm and 10 nm respectively.

Fluorescence titrations experiments were performed by adding increasing concentrations of nucleic acid to a fixed amount of *th*TrmI protein (15 nM) in 50 mM Tris-HCl, 200 mM KCl, 0.5 % (w/v) PEG 8000, pH 8.0. Fluorescence intensities were corrected for dilution and were fitted using Eq. (1) assuming two equivalent tRNA binding sites per *th*TrmI tetramer (see Mass Spectrometry Results).

$$(1)\quad I = I_0 - \frac{I_0 - I_\infty}{2nN_t}\left(K_d + L_t + nN_t - \sqrt{\left(K_d + L_t + nN_t\right)^2 - 4L_t nN_t}\right)$$

where I_0, Fluorescence intensity without tRNA; I, fluorescence intensity at a given concentration of tRNA; I_∞, fluorescence intensity at the plateau; n, number of tRNA binding sites on the protein; N_t, total concentration of protein; L_t, total concentration of tRNA. Confidence limits on the parameters were estimated by Monte-Carlo sampling.[61]

Steady-state kinetic assay

[Methyl-^{3}H]-AdoMet (555 GBq/mmol, 20.5 MBq/mL, MP Biomedicals) was mixed with non-radioactive AdoMet (Sigma) to achieve a specific radioactivity of about 500 cpm/pmol (16.7 Bq/pmol). The methylation kinetic assays were performed in 50 mM Tris-HCl pH 8.0 at 60 °C. The tRNA and [methyl-^{3}H]-AdoMet were thermally equilibrated in 1.25 x concentrated reaction buffer (96 µL) at 60 °C for 5 min, and the reaction was started by adding the enzyme (24 µL). Aliquots (25 µL) were removed after different incubation times and transferred into 800 µL of 5 % (w/v) trichloroacetic acid (TCA) at 0 °C for 30 min in order to quench the reaction and to precipitate the tRNA. The precipitates were collected by filtration through GF/C filters (Whatman). The filters were washed with cold 5 % TCA, dried, and the radioactivity was measured by liquid scintillation counting for 2 min, resulting in a counting error below 4 %. Data were corrected by substracting the background radioactivity determined from a control without enzyme.

For the determination of AdoMet K_M and k_{cat}, the reaction mixtures contained MTase (wild type, 25 nM ; Y194A, 50 nM ; Y78A, 100 nM ; D170A, 500 nM), $tRNA_i^{Met}$ (20 µM) and ^{3}H-AdoMet (wild type, 0.5-8 µM ; Y194A, 0.5-8 µM ; Y78A, 0.5-8 µM ; D170A, 5-50 µM). For $tRNA_i^{Met}$ K_M and k_{cat}, the reaction mixtures contained MTase (wild type, 25 nM ; Y194A, 50 nM ; Y78A, 100 nM ; D170A, 500 nM), $tRNA_i^{Met}$ (wild type, 0.5-8 µM ; Y194A, 0.5-8 µM ; Y78A, 0.25-4 µM ; D170A, 0.5-8 µM) and ^{3}H-AdoMet (wild type, 20 µM ; Y194A, 50 µM ; Y78A, 20 µM ; D170A, 50 µM). Initial rates (v_i) for each substrate concentration ($tRNA_i^{Met}$) were determined from the slope of linear fits of time course data points (wild type, 9 min ; Y194A, 20 min ; Y78A, 40 min ; D170A, 40 min). Enzyme parameters were obtained by non-linear least square fitting using Eq. (2) of Michaelis-Menten kinetics (wild type and mutant Y194A) or Eq. (3) in cases where enzyme concentration was not negligible compared to substrate concentration (mutant Y78A and D170A). k_{cat} values were calculated for two catalytic sites per TrmI tetramer. Confidence limits on the parameters were estimated by Monte-Carlo sampling using the MC-Fit program.[61]

(2) $$v_i = \frac{V_{\max} S_0}{K_M + S_0}$$ (3) $$v_i = \frac{V_{\max}}{2E_0}\left(K_M + S_0 + E_0 - \sqrt{(K_M + S_0 + E_0)^2 - 4S_0E_0}\right)$$

Molecular Docking

Docking of the N_9-methyladenine ring into the active site of TrmI was performed using HADDOCK 1.3[37] (http://www.nmr.chem.uu.nl/haddock/) that makes use of CNS[62] as structure calculation software. HADDOCK allows to deal with ambiguous constraints, i.e. restraints can be applied between residues and not only between atoms. The protein structure used for the docking was the tetrameric *T. thermophilus* TrmI crystal structure in which the AdoHcy ligand was replaced with the AdoMet methyl donor. A 2 Å distance was used to define ambiguous restraints (AIRs) that were applied between the adenine and a set of neighbours composed of AdoMet cofactor and *th*TrmI residues identified by mutagenesis (*i.e.* Y78, D170 and Y194). One unambiguous restraint set to 2 Å was applied between the atoms that react in the methyl transfer reaction, *i.e.* the N_1 position of the N_9-methyladenine ring and the Cε atom of the AdoMet ligand. During the rigid body energy minimisation, 2000 structures were calculated and the 1000 best solutions based on the intermolecular energy were used for the semi-flexible refinement. The docking converged to a single solution. The same protocol was also applied to the *M. tuberculosis* TrmI.

REFERENCES

1. Rozenski, J., Crain, P. F. & McCloskey, J. A. (1999). The RNA Modification Database: 1999 update. *Nucl. Acids Res.* **27,** 196-197.
2. Anderson, J., Phan, L. & Hinnebusch, A. G. (2000). The Gcd10p/Gcd14p complex is the essential two-subunit tRNA(1-methyladenosine) methyltransferase of Saccharomyces cerevisiae. *Proc. Natl. Acad. Sci. USA* **97,** 5173-5178.
3. Droogmans, L., Roovers, M., Bujnicki, J. M., Tricot, C., Hartsch, T., Stalon, V. & Grosjean, H. (2003). Cloning and characterization of tRNA (m1A58) methyltransferase (TrmI) from Thermus thermophilus HB27, a protein required for cell growth at extreme temperatures. *Nucl. Acids Res.* **31,** 2148-2156.
4. Anderson, J., Phan, L., Cuesta, R., Carlson, B. A., Pak, M., Asano, K., Bjork, G. R., Tamame, M. & Hinnebusch, A. G. (1998). The essential Gcd10p-Gcd14p nuclear complex is required for 1-methyladenosine modification and maturation of initiator methionyl-tRNA. *Genes Dev.* **12,** 3650-3662.
5. Ozanick, S. G., Bujnicki, J. M., Sem, D. S. & Anderson, J. T. (2007). Conserved amino acids in each subunit of the heteroligomeric tRNA m1A58 Mtase from Saccharomyces cerevisiae contribute to tRNA binding. *Nucl. Acids Res.* **35,** 6808-6819.
6. Bujnicki, J. M. (2001). In silico analysis of the tRNA:m1A58 methyltransferase family: homology-based fold prediction and identification of new members from Eubacteria and Archaea. *FEBS Letters* **507,** 123-127.
7. Gupta, A., Kumar, P. H., Dineshkumar, T. K., Varshney, U. & Subramanya, H. S. (2001). Crystal structure of Rv2118c: an AdoMet-dependent methyltransferase from Mycobacterium tuberculosis H37Rv. *J. Mol. Biol.* **312,** 381-391.
8. Roovers, M., Wouters, J., Bujnicki, J. M., Tricot, C., Stalon, V., Grosjean, H. & Droogmans, L. (2004). A primordial RNA modification enzyme: the case of tRNA (m1A) methyltransferase. *Nucl. Acids Res.* **32,** 465-476.
9. Varshney, U., Ramesh, V., Madabushi, A., Gaur, R., Subramanya, H. S. & RajBhandary, U. L. (2004). Mycobacterium tuberculosis Rv2118c codes for a single-component homotetrameric m1A58 tRNA methyltransferase. *Nucl. Acids Res.* **32,** 1018-1027.
10. Schubert, H. L., Blumenthal, R. M. & Cheng, X. (2003). Many paths to methyltransfer: a chronicle of convergence. *Trends Biochem. Sci.* **28,** 329-335.
11. Loo, J. A. (1997). Studying noncovalent protein complexes by electrospray ionization mass spectrometry. *Mass Spectrom. Rev.* **16,** 1-23.
12. Loo, J. A. (2000). Electrospray ionization mass spectrometry: a technology for studying noncovalent macromolecular complexes. *Int. J. Mass Spectrom.* **200,** 175-186.
13. Heck, A. J. & Van Den Heuvel, R. H. (2004). Investigation of intact protein complexes by mass spectrometry. *Mass Spectrom. Rev.* **23,** 368-389.
14. Veenstra, T. D. (1999). Electrospray ionization mass spectrometry: a promising new technique in the study of protein/DNA noncovalent complexes. *Biochem. Biophys. Res. Commun.* **257,** 1-5.
15. Beck, J. L., Colgrave, M. L., Ralph, S. F. & Sheil, M. M. (2001). Electrospray ionization mass spectrometry of oligonucleotide complexes with drugs, metals, and proteins. *Mass Spectrom. Rev.* **20,** 61-87.
16. Hofstadler, S. A. & Griffey, R. H. (2001). Analysis of noncovalent complexes of DNA and RNA by mass spectrometry. *Chem. Rev.* **101,** 377-390.

17. Rusconi, F., Guillonneau, F. & Praseuth, D. (2002). Contributions of mass spectrometry in the study of nucleic acid-binding proteins and of nucleic acid-protein interactions. *Mass Spectrom. Rev.* **21,** 305-348.
18. Robinson, C. V., Chung, E. W., Kragelund, B. B., Knudsen, J., Aplin, R. T., Poulsen, F. M. & Dobson, C. M. (1996). Probing the nature of noncovalent interactions by mass spectrometry. A study of protein-CoA ligand binding and assembly. *J. Am. Chem. Soc.* **118,** 8646-8653.
19. Rogniaux, H., Van Dorsselaer, A., Barth, P., Biellmann, J. F., Barbanton, J., van Zandt, M., Chevrier, B., Howard, E., Mitschler, A., Potier, N., Urzhumtseva, L., Moras, D. & Podjarny, A. (1999). Binding of aldose reductase inhibitors: Correlation of crystallographic and mass spectrometric studies. *J. Am. Soc. Mass Spectrom.* **10,** 635-647.
20. Mei, H. Y., Mack, D. P., Galan, A. A., Halim, N. S., Heldsinger, A., Loo, J. A., Moreland, D. W., Sannes-Lowery, K. A., Sharmeen, L., Truong, H. N. & Czarnik, A. W. (1997). Discovery of selective, small-molecule inhibitors of RNA complexes--I. The Tat protein/TAR RNA complexes required for HIV-1 transcription. *Bioorg. Med. Chem.* **5,** 1173-1184.
21. Liu, C., Tolic, L. P., Hofstadler, S. A., Harms, A. C., Smith, R. D., Kang, C. & Sinha, N. (1998). Probing RegA/RNA interactions using electrospray ionization-fourier transform ion cyclotron resonance-mass spectrometry. *Anal. Biochem.* **262,** 67-76.
22. Potier, N., Donald, L. J., Chernushevich, I., Ayed, A., Ens, W., Arrowsmith, C. H., Standing, K. G. & Duckworth, H. W. (1998). Study of a noncovalent trp repressor: DNA operator complex by electrospray ionization time-of-flight mass spectrometry. *Protein Sci.* **7,** 1388-1395.
23. Hagan, N. & Fabris, D. (2003). Direct mass spectrometric determination of the stoichiometry and binding affinity of the complexes between nucleocapsid protein and RNA stem-loop hairpins of the HIV-1 Psi-recognition element. *Biochemistry* **42,** 10736-10745.
24. Kilic, T., Sanglier, S., Van Dorsselaer, A. & Suck, D. (2006). Oligomerization behavior of the archaeal Sm2-type protein from Archaeoglobus fulgidus. *Protein Sci.* **15,** 2310-2317.
25. Hamilton, S., Odili, J., Pacifico, M. D., Wilson, G. D. & Kupsch, J. M. (2003). Effect of imidazole on the solubility of a His-tagged antibody fragment. *Hybridoma and Hybridomics* **22,** 347-355.
26. McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. (2005). Likelihood-enhanced fast translation functions. *Acta Crystallog. sect. D* **61,** 458-464.
27. Fauman, E. B., Blumenthal, R. M. & Cheng, X. (1999). Structure and evolution of AdoMet-dependent methyltransferases. In *S-adenosylmethionine-dependent methyltransferases: structures and functions* (Cheng, X. & Blumenthal, R. M., eds.), pp. 1-38. World Scientific Publishing.
28. Rostom, A. A. & Robinson, C. V. (1999). Detection of the intact GroEL chaperonin assembly by mass spectrometry. *J. Am. Chem. Soc.* **121,** 4718-4719.
29. van Berkel, W. J., van den Heuvel, R. H., Versluis, C. & Heck, A. J. (2000). Detection of intact megaDalton protein assemblies of vanillyl-alcohol oxidase by mass spectrometry. *Protein Sci.* **9,** 435-439.
30. Sanglier, S., Ramstrom, H., Haiech, J., Leize, E. & Van Dorsselaer, A. (2002). Electrospray ionization mass spectrometry analysis revealed a similar to 310 kDa noncovalent hexamer of HPr kinase/phosphatase from Bacillus subtilis. *Int. J. Mass Spectrom.* **219,** 681-696.

31. Sanglier, S., Leize, E., Van Dorsselaer, A. & Zal, F. (2003). Comparative ESI-MS study of approximately 2.2 MDa native hemocyanins from deep-sea and shore crabs: from protein oligomeric state to biotope. *J. Am. Soc. Mass Spectrom.* **14,** 419-429.
32. Rogniaux, H., Sanglier, S., Strupat, K., Azza, S., Roitel, O., Ball, V., Tritsch, D., Branlant, G. & Van Dorsselaer, A. (2001). Mass spectrometry as a novel approach to probe cooperativity in multimeric enzymatic systems. *Anal. Biochem.* **291,** 48-61.
33. Smith, R. D. & Light-Wahl, K. J. (1993). The observation of non-covalent interactions in solution by electrospray ionization mass spectrometry: Promise, pitfalls and prognosis. *Biol. Mass Spectrom.* **22,** 493-501.
34. Kapur, A., Beck, J. L., Brown, S. E., Dixon, N. E. & Sheil, M. M. (2002). Use of electrospray ionization mass spectrometry to study binding interactions between a replication terminator protein and DNA. *Protein Sci.* **11,** 147-157.
35. Gupta, R., Hamdan, S. M., Dixon, N. E., Sheil, M. M. & Beck, J. L. (2004). Application of electrospray ionization mass spectrometry to study the hydrophobic interaction between the epsilon and theta subunits of DNA polymerase III. *Protein Sci.* **13,** 2878-2887.
36. Cassiday, L. A., Lebruska, L. L., Benson, L. M., Naylor, S., Owen, W. G. & Maher, L. J., 3rd. (2002). Binding stoichiometry of an RNA aptamer and its transcription factor target. *Anal. Biochem.* **306,** 290-297.
37. Dominguez, C., Boelens, R. & Bonvin, A. M. (2003). HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.* **125,** 1731-1737.
38. McCarthy, A. A. & McCarthy, J. G. (2007). The structure of two N-methyltransferases from the caffeine biosynthetic pathway. *Plant Physiol.* **144,** 879-889.
39. Kettani, A., Gueron, M. & Leroy, J. L. (1997). Amino proton exchange processes in mononucleosides. *J. Am. Chem. Soc.* **119,** 1108-1115.
40. Stewart, R. & Harris, M. G. (1977). Amino Group Acidity in Nucleotide Bases. *Can. J. Chem.* **55,** 3807-3814.
41. Elkins, P. A., Watts, J. M., Zalacain, M., van Thiel, A., Vitazka, P. R., Redlak, M., Andraos-Selim, C., Rastinejad, F. & Holmes, W. M. (2003). Insights into catalysis by a knotted TrmD tRNA methyltransferase. *J. Mol. Biol.* **333,** 931-949.
42. Labahn, J., Granzin, J., Schluckebier, G., Robinson, D. P., Jack, W. E., Schildkraut, I. & Saenger, W. (1994). Three-dimensional structure of the adenine-specific DNA methyltransferase M.Taq I in complex with the cofactor S-adenosylmethionine. *Proc. Natl. Acad. Sci. USA* **91,** 10957-10961.
43. Tran, P. H., Korszun, Z. R., Cerritelli, S., Springhorn, S. S. & Lacks, S. A. (1998). Crystal structure of the DpnM DNA adenine methyltransferase from the DpnII restriction system of streptococcus pneumoniae bound to S-adenosylmethionine. *Structure* **6,** 1563-1575.
44. Goedecke, K., Pignot, M., Goody, R. S., Scheidig, A. J. & Weinhold, E. (2001). Structure of the N6-adenine DNA methyltransferase M.TaqI in complex with DNA and a cofactor analog. *Nat. Struct. Biol.* **8,** 121-125.
45. Schluckebier, G., Zhong, P., Stewart, K. D., Kavanaugh, T. J. & Abad-Zapatero, C. (1999). The 2.2 A structure of the rRNA methyltransferase ErmC' and its complexes with cofactor and cofactor analogs: implications for the reaction mechanism. *J. Mol. Biol.* **289,** 277-291.
46. Maravic, G., Feder, M., Pongor, S., Flogel, M. & Bujnicki, J. M. (2003). Mutational analysis defines the roles of conserved amino acid residues in the predicted catalytic pocket of the rRNA:m6A methyltransferase ErmC'. *J. Mol. Biol.* **332,** 99-109.

47. Bheemanaik, S., Reddy, Y. V. & Rao, D. N. (2006). Structure, function and mechanism of exocyclic DNA methyltransferases. *Biochem. J.* **399,** 177-190.
48. Jeltsch, A. (2002). Beyond Watson and Crick: DNA methylation and molecular enzymology of DNA methyltransferases. *Chembiochem* **3,** 274-293.
49. Pues, H., Bleimling, N., Holz, B., Wolcke, J. & Weinhold, E. (1999). Functional roles of the conserved aromatic amino acid residues at position 108 (motif IV) and position 196 (motif VIII) in base flipping and catalysis by the N6-adenine DNA methyltransferase from Thermus aquaticus. *Biochemistry* **38,** 1426-1434.
50. Roth, M., Helm-Kruse, S., Friedrich, T. & Jeltsch, A. (1998). Functional roles of conserved amino acid residues in DNA methyltransferases investigated by site-directed mutagenesis of the EcoRV adenine-N6-methyltransferase. *J. Biol. Chem.* **273,** 17333-17342.
51. Collaborative Computational Project Number 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D* **50,** 760-763.
52. Potterton, E., Briggs, P., Turkenburg, M. & Dodson, E. (2003). A graphical user interface to the CCP4 program suite. *Acta Crystallog. sect. D* **59,** 1131-1137.
53. Leslie, A. G. W. (1992). *Recent changes to the MOSFLM package for processing film and image plate data*. Joint CCP4+ ESF-EAMCB Newsletter on Protein Crystallography (Science and Engineering Research Council, D. L., Ed.), Daresbury Laboratory, Warrington, UK.
54. Evans, P. R. (1993). *Proceedings of the CCP4 Study Weekend: Data collection and processing*. Science and Engineering Research Council, Daresbury Laboratory, Warrington, United Kingdom.
55. Perrakis, A., Morris, R. & Lamzin, V. S. (1999). Automated protein model building combined with iterative structure refinement. *Nat. Struct. Biol.* **6,** 458-463.
56. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallog. sect. D* **53,** 240-255.
57. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallog. sect. D* **60,** 2126-2132.
58. Winn, M. D., Isupov, M. N. & Murshudov, G. N. (2001). Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallog. sect. D* **57,** 122-133.
59. Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. USA* **98,** 10037-10041.
60. Meinnel, T. & Blanquet, S. (1995). Maturation of pre-tRNA(fMet) by Escherichia coli RNase P is specified by a guanosine of the 5'-flanking sequence. *J. Biol. Chem.* **270,** 15908-15914.
61. Dardel, F. (1994). MC-Fit: using Monte-Carlo methods to get accurate confidence limits on enzyme parameters. *Comput. Appl. Biosci.* **10,** 273-275.
62. Brunger, 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). Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallog. sect. D* **54,** 905-921.
63. Schluckebier, G., O'Gara, M., Saenger, W. & Cheng, X. (1995). Universal catalytic domain structure of AdoMet-dependent methyltransferases. *J. Mol. Biol.* **247,** 16-20.

ACKNOWLEGMENTS

The authors would like to thank Franck Brachet for his skill in the set up of the pipetting robot. We acknowledge Dr Thibaut Crépin for his kind help in the synchrotron data collection on beamline ID14-1 at ESRF (Grenoble). We thank Nathalie Ulryck for help in site-directed mutagenesis. Pierre Barraud and Cédric Atmanene are supported by a studentship from the "Ministère de la recherche". We thank the CNRS for financial support. Louis Droogmans was supported by grants from the FRFC (Fonds pour la Recherche Fondamentale Collective) and IISN (Institut Interuniversitaire des Sciences Nucléaires).

FIGURE LEGENDS

Figure 1: Crystal Structure of *T. thermophilus* TrmI.
(a) Overall structure of *th*TrmI monomer. The secondary structure elements are labelled following the nomenclature defined in Schluckebier *et al.*[63] The cofactor AdoHcy is shown in blue sticks. (b) Organization of *th*TrmI as a tetramer. Each monomer is drawn with a different colour. The two tight dimers are respectively in red and yellow, and in pale and dark blue. The tetramerization interface involved strands β_6 and β_7.
(c) The AdoHcy binding site. The ligand (blue) and the enzyme residues (gray) are shown as sticks. Water molecules are drawn as blue spheres. Polar contacts are represented with dashed lines. Strictly conserved residue among m^1A_{58} tRNA MTases are labelled with bold letters. Characteristic motifs from MTases are highlighted. (d) Density map (2Fo-Fc at 1σ) around the cofactor. The AdoHcy ligand (blue) and the enzyme residues (green) are shown as sticks. Water molecules are represented as red spheres.

Figure 2: Electrostatic surface potential of (a) *T. thermophilus* TrmI and (b) *M. tuberculosis* TrmI. Blue indicates positive charge and red negative one with the maximum color saturation corresponding to -3 kT (red) and +3 kT (blue). The figure was prepared using the APBS[59] PyMOL plugin.

Figure 3: Noncovalent ESI-MS analyses of *th*TrmI and *th*TrmI-tRNA assemblies. ESI mass spectra obtained under non-denaturing conditions (200 mM ammonium bicarbonate buffer pH 8.0; Vc = 120 V; Pi = 6.1 mbar) for (a) *th*TrmI tetramer (20 μM) alone and in presence of (b) 30 μM, (c) 100 μM and (d) 200 μM of *E. coli* $tRNA_i^{Met}$. Inserts correspond to deconvoluted spectra showing molecular weights on the x-scale. T (MW = 115450 ± 5 Da) corresponds to $(thTrmI)_4$, T+1tRNA (MW = 140347 ± 6 Da) corresponds to $(thTrmI)_4$:$(tRNA)_1$ complex and T+2tRNA (MW = 165455 ± 10 Da) corresponds to $(thTrmI)_4$:$(tRNA)_2$ complex. (*) correspond to tRNA signals.

Figure 4: Non-denaturing agarose gel electrophoresis
Constant amount of *E. coli* $tRNA_i^{Met}$ (4 μg) with increasing amounts of *th*TrmI protein (from 0 to 1 molar equivalent of *th*TrmI tetramer). Bands were visualized by UV shadowing.

Figure 5: Influence of the buffer ionic strength on *th*TrmI-tRNA complexes stabilities. ESI mass spectra obtained under non-denaturing conditions (ammonium bicarbonate pH 8.0; Vc = 120 V; Pi = 6.1 mbar) for *th*TrmI tetramer (20 μM) in presence of 100 μM of *E. coli* $tRNA_i^{Met}$. Ammonium bicarbonate concentration was set to (a) 100 mM, (b) 200 mM, (c) 400 mM and (d) 1250 mM. Inserts correspond to deconvoluted spectra showing molecular

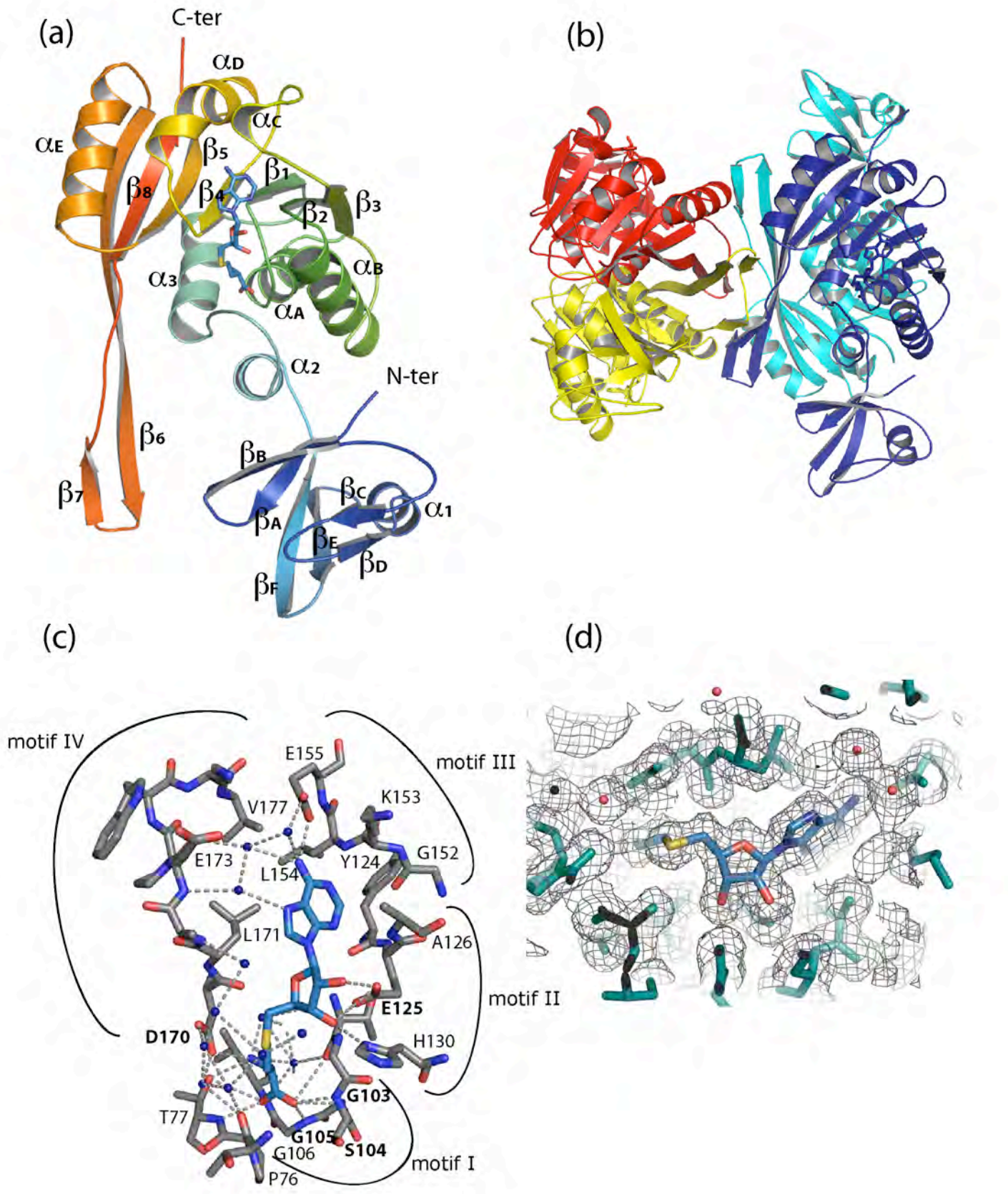
weights on the x-scale. T corresponds to (*th*TrmI)$_4$, T+1tRNA corresponds to (*th*TrmI)$_4$:(tRNA)$_1$ complex and T+2tRNA corresponds to (*th*TrmI)$_4$:(tRNA)$_2$ complex. (*) correspond to tRNA signals.

Figure 6: Specificity of *th*TrmI-tRNA complexes formation. ESI mass spectra obtained under non-denaturing conditions (200 mM ammonium bicarbonate buffer pH 8.0; Vc = 120 V; Pi = 6.1 mbar) for *th*TrmI tetramer (20 μM) in presence of (a) 100 μM *E. coli* $tRNA_i^{Met}$ (b) 100 μM non substrate control RNA. Inserts correspond to deconvoluted spectra showing molecular weights on the x-scale. T corresponds to (*th*TrmI)$_4$, T+1tRNA corresponds to (*th*TrmI)$_4$:(tRNA)$_1$ complex and T+2tRNA corresponds to (*th*TrmI)$_4$:(tRNA)$_2$ complex. (*) corresponds to RNA signal.

Figure 7: Sequence alignments of TrmI proteins and conserved residues in *th*TrmI catalytic pocket.
(a) Part of sequence alignment of *T. thermophilus*, *M. tuberculosis*, *P. abyssi*, *M. musculus*, *H. sapiens* and *S. cerevisiae* m^1A_{58} tRNA MTases. Residue numbers are those of *T. thermophilus* sequence. Highly conserved residues are shown on a dark blue background and residues with a similar physico-chemical character are on a pale blue background. The secondary structure of the *T. thermophilus* TrmI is shown below the sequences as arrows for β-strands and cylinder for α-helices. The names of the secondary structure elements refer to Figure 1(a). The sequence motifs typical of the AdoMet dependent MTase family are also indicated. Residues marked with a star (*) (i.e. Y78, D170 and Y194) are involved in the catalytic pocket formation and are those studied by mutagenesis in the present work. Residues marked with a circle (**o**) (i.e. W175 and F245) are conserved aromatic residues in motifs IV and VIII, belonging to the catalytic pocket in m^6A DNA MTases. (b) Conserved and semi-conserved residues around the catalytic pocket of *T. thermophilus* TrmI. The strictly conserved residues among the m^1A_{58} tRNA MTase family are shown as sticks on the *T. thermophilus* structure except the conserved residues that bind AdoMet. AdoMet was modeled from AdoHcy by adding a methyl group on the sulphur atom, the methyl is positionned by the sulfonium chirality. Y78, Y194 and D170 are indicated as red sticks.

Figure 8: Modeling a N_9-methyladenine in the catalytic pocket of *th*TrmI
(a) Average docked structure of the N_9-methyladenine ring inside the catalytic pocket of *th*TrmI. The enzyme surface is coloured according to the electrostatic surface potential. Positive potential is shown in blue and negative potential in red. The AdoMet ligand and the docked N_9-methyladenine ring are shown as sticks (carbon atoms in light blue, nitrogen atoms in dark blue, oxygen atoms in red, sulphur atom in yellow and hydrogen atoms in white). The Van der Waals surface of the N_9-methyladenine ring is shown as white dots.

(b) Distances between the average structure of the docked N_9-methyladenine ring and key moiety of the enzyme. The enzyme is shown in gray with the cartoon representation. The AdoMet ligand, the docked N_9-methyladenine ring and residues D170, Y78 and Y194 are shown as sticks. Distances are indicated with dashed lines and are given in Å.


(a)
C-ter
αD
αC
αE
β5
β1
β8
β4
β2
β3
αB
α3
αA
α2
N-ter
β6
βB
β7
βC
α1
βA
βE
βD
βF
(b)
(c)
motif IV
E155
motif III
V177
K153
E173
Y124
G152
L154
L171
A126
motif II
E125
D170
H130
G103
T77
G105
S104
G106
motif I
P76
(d)

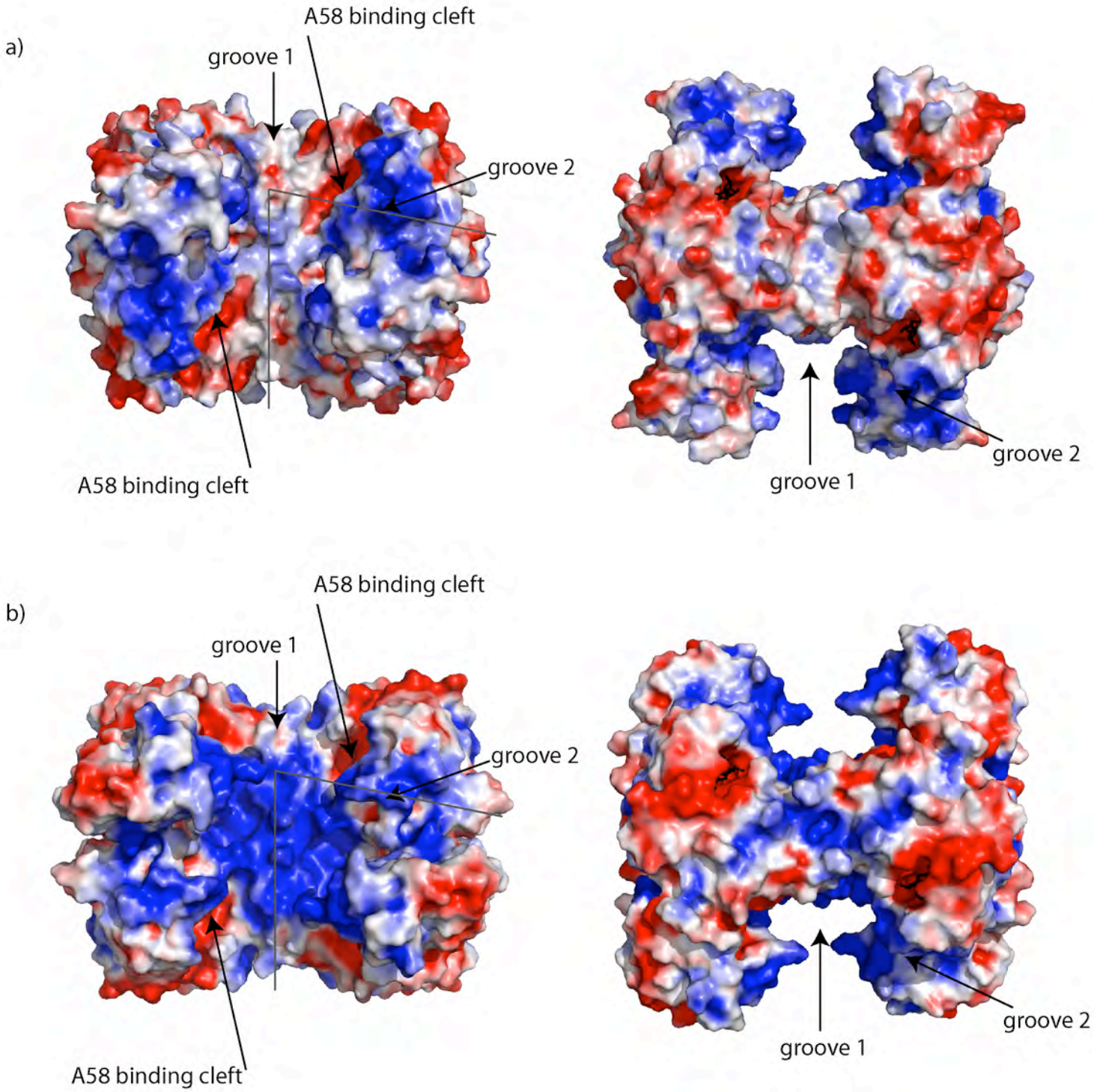
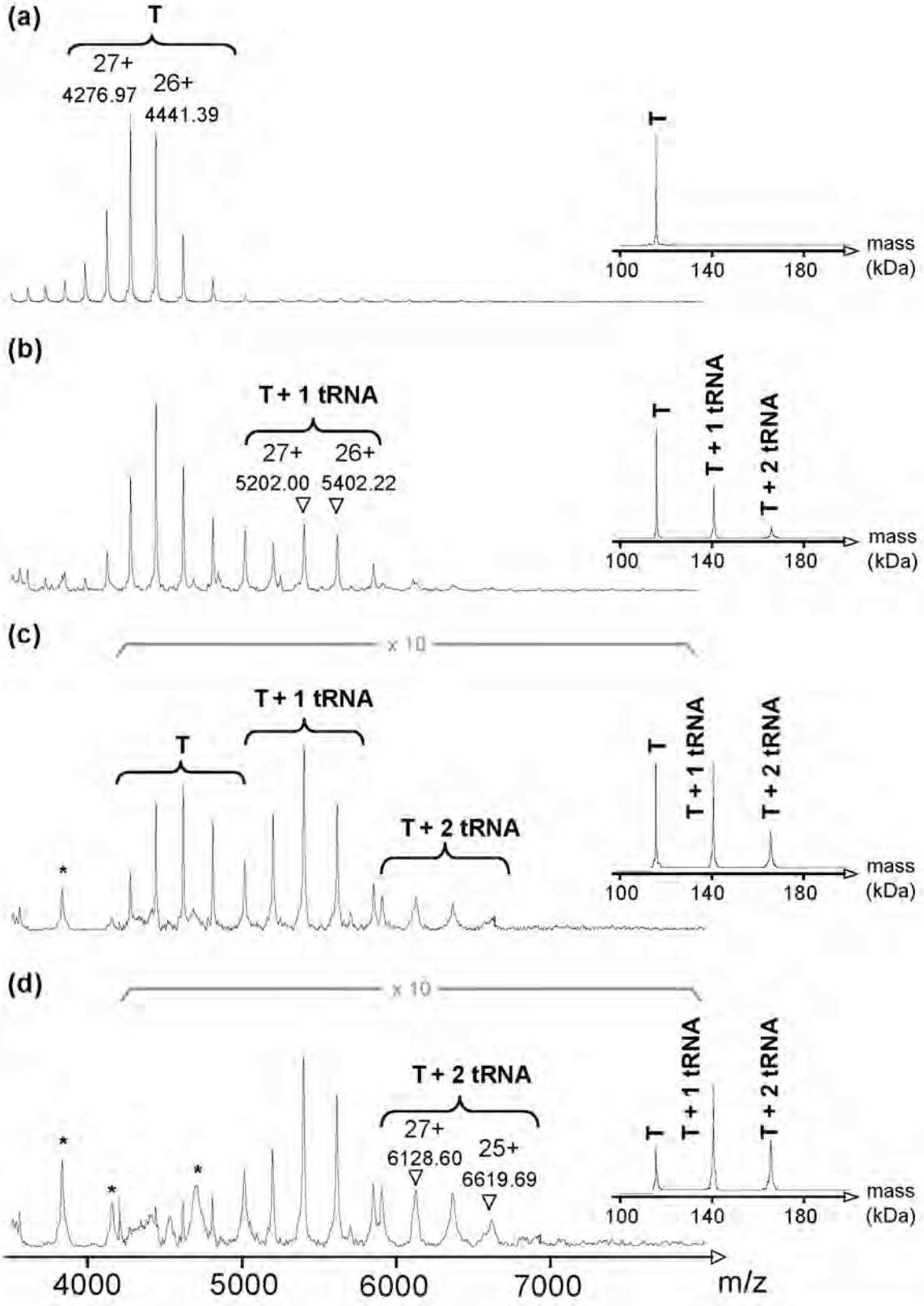

a)
A58 binding cleft
groove 1
groove 2
A58 binding cleft
groove 2
groove 1
b)
A58 binding cleft
groove 1
groove 2
A58 binding cleft
groove 2
groove 1

Figure 2



T : 115450 ± 5 Da **T + 1 tRNA** : 140347 ± 6 Da **T + 2 tRNA** : 165455 ± 10 Da

* : tRNA

Figure 3

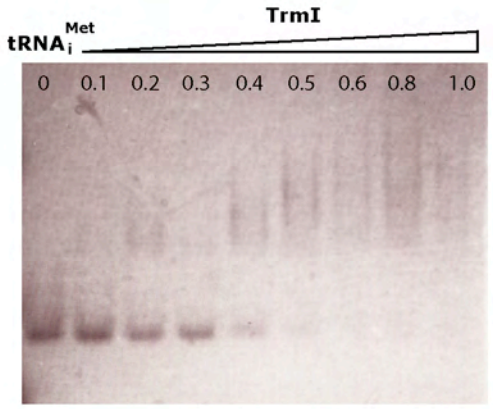


Figure 4

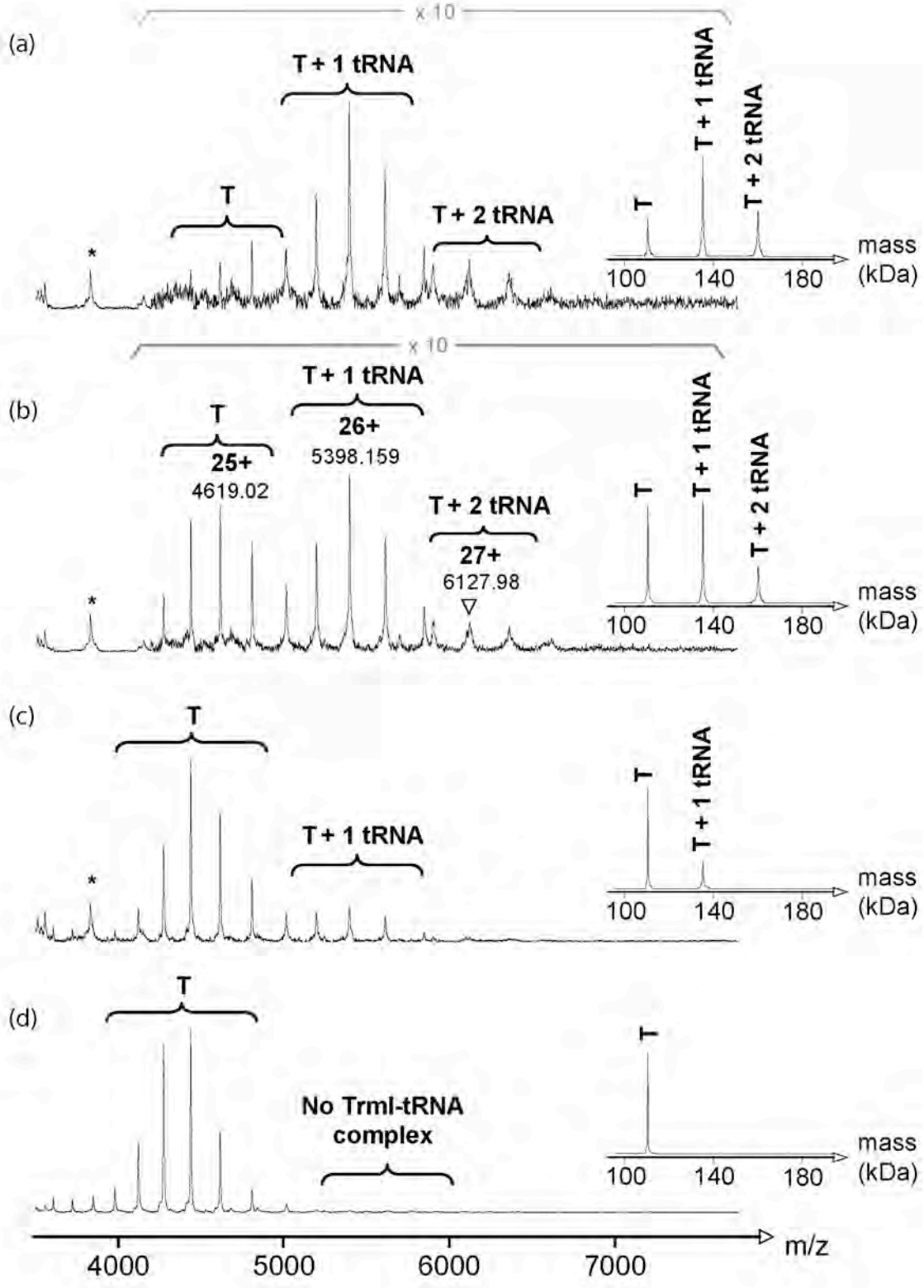

(a)
x 10
T + 1 tRNA
T
T + 2 tRNA
*
T
T + 1 tRNA
T + 2 tRNA
100
140
180
mass
(kDa)
(b)
x 10
T + 1 tRNA
26+
5398.159
T
25+
4619.02
T + 2 tRNA
27+
6127.98
*
T
T + 1 tRNA
T + 2 tRNA
100
140
180
mass
(kDa)
(c)
T
T + 1 tRNA
*
T
T + 1 tRNA
100
140
180
mass
(kDa)
(d)
T
No TrmI-tRNA
complex
T
100
140
180
mass
(kDa)
4000
5000
6000
7000
m/z

Figure 5

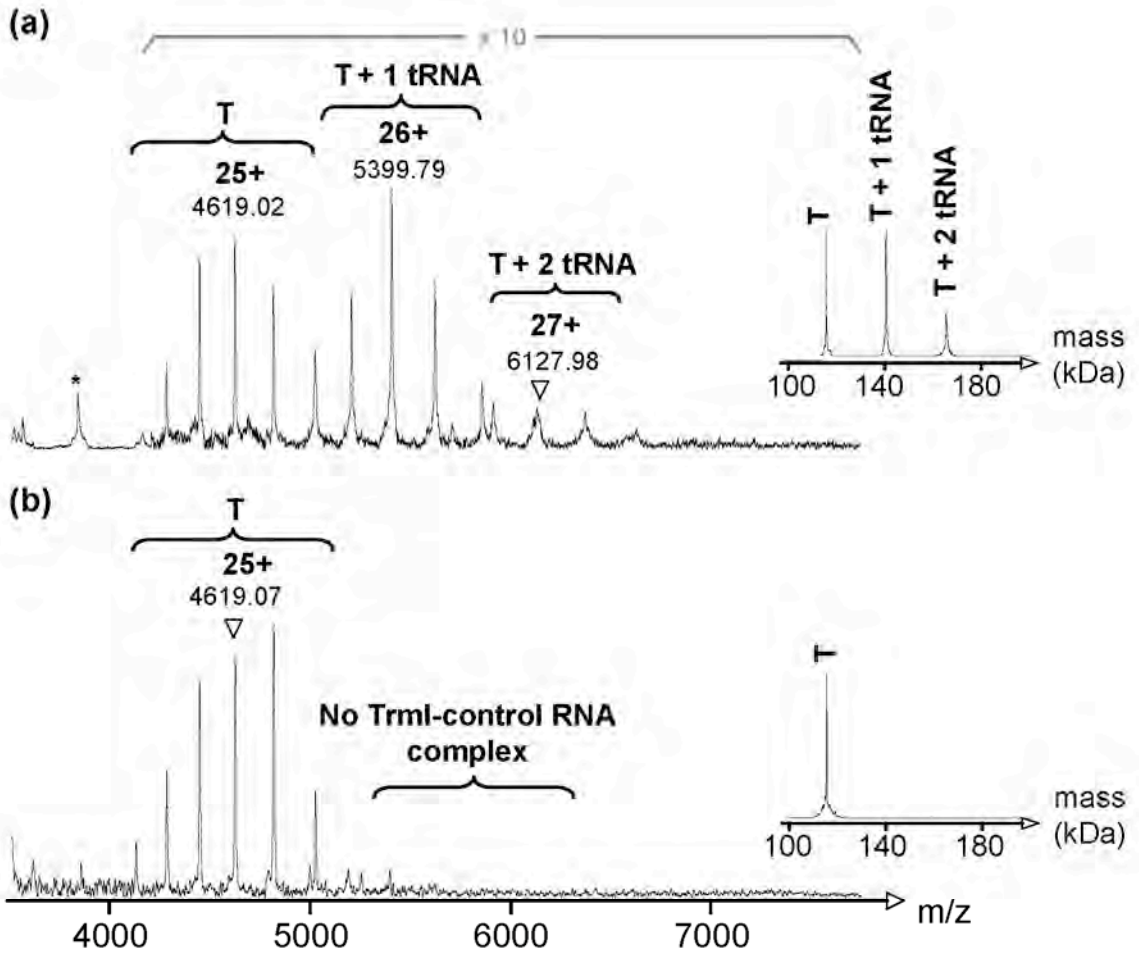

(a)
x 10
T
25+
4619.02
T + 1 tRNA
26+
5399.79
T + 2 tRNA
27+
6127.98
*
T
T + 1 tRNA
T + 2 tRNA
100
140
180
mass
(kDa)
(b)
T
25+
4619.07
No TrmI-control RNA
complex
T
100
140
180
mass
(kDa)
4000
5000
6000
7000
m/z

Figure 6

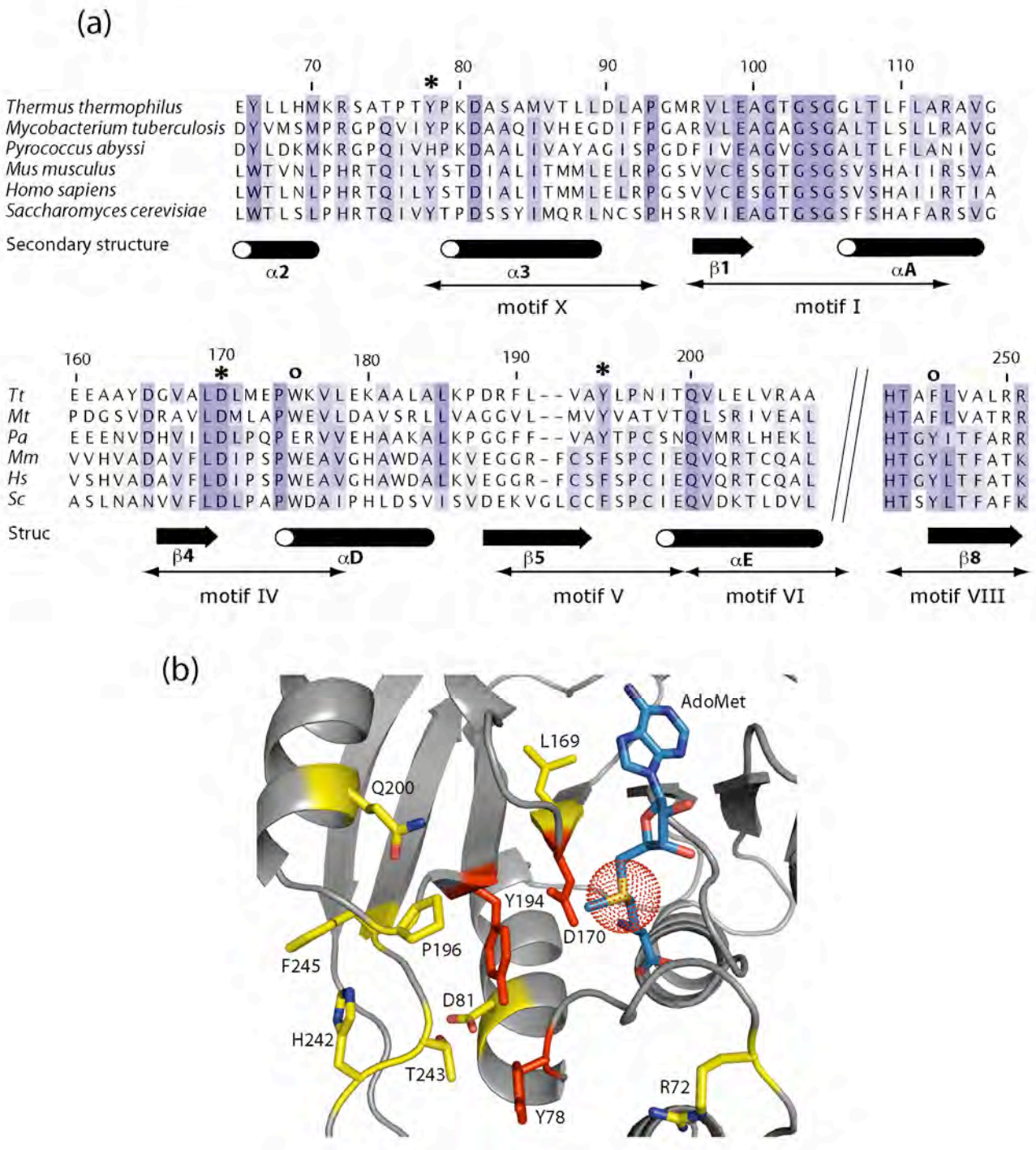

(a)
70
80
90
100
110
Thermus thermophilus
Mycobacterium tuberculosis
Pyrococcus abyssi
Mus musculus
Homo sapiens
Saccharomyces cerevisiae
Secondary structure
α2
α3
β1
αA
motif X
motif I
160
170
180
190
200
250
Tt
Mt
Pa
Mm
Hs
Sc
Struc
β4
αD
β5
αE
β8
motif IV
motif V
motif VI
motif VIII
(b)
AdoMet
L169
Q200
Y194
D170
P196
F245
D81
H242
T243
R72
Y78

Figure 7

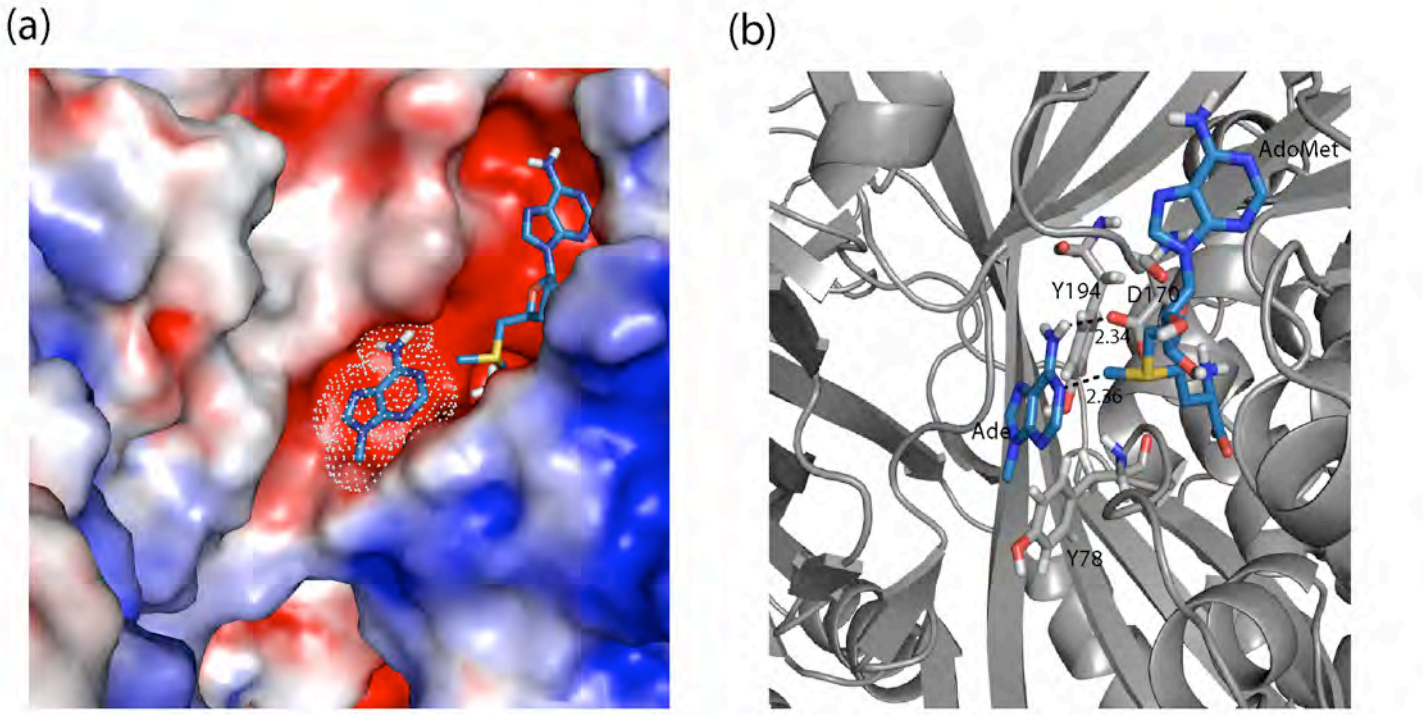


Figure 9

Table 1 *Data collection and refinement statistics for the determination of TrmI structure.*
Values in parenthesis are for the outer resolution shell.

A. *Data collection statistics*		B. *Refinement statistics*	
Wavelength (Å)	0.9340	Resolution limits (Å)	33.67-1.70 (1.74-1.70)
Space group	$C222_1$	Number of reflections	64480 (4709)
Unit-cell parameters (Å)		*R*-factor (%)	18.0 (24.7)
a	90.4406	*R*free (%)	20.9 (27.9)
b	96.7946	Number of atoms	4311
c	140.5837	Protein	3823
No. of measured reflections	499190 (71441)	Ligands	67
No. of unique reflections	67981 (9830)	Waters	421
Multiplicity	7.3 (7.3)	Mean B-factors (Å^2)	15.92
Resolution limits (Å)	48.39-1.70 (1.79-1.70)	RMSD (bounds) (Å)	0.009
Data completeness (%)	100.0 (100.0)	RMSD (angles) (°)	1.265
Mean $I/\sigma(I)$	13.6 (2.9)		
R_{merge}	10.6 (67.6)		

Table 2 *Kinetic and thermodynamic parameters of wild-type and mutant TrmI enzymes.*
The parameters are indicated with a confidence interval of 90 %.

	AdoMet		$tRNA_i^{Met}$			
Mutants	k_{cat} (s^{-1} x 10^{-3})	K_M (µM)	k_{cat} (s^{-1} x 10^{-3})	K_M (µM)	K_D (nM)	Relative k_{cat}
WT	37 (30 - 44)	2.1 (1.4-3.0)	36 (28 - 45)	2.7 (1.7-4.8)	13 (5 - 24)	100.0
Y194A	6 (5 – 11)	5.7 (3.9-8.7)	10 (9 - 13)	2.2 (1.9-3.3)	18 (10 - 37)	27.4
Y78A	1.4 (1.1 – 1.7)	1.9 (1.2-2.7)	1.6 (1.5 – 1.8)	0.16 (0.11-0.23)	16 (7 - 33)	4.1
D170A	0.13 (0.11 - 0.14)	5.4 (3.7-8.0)	0.11 (0.10 - 0.12)	0.17 (0.07-0.39)	14 (6 - 25)	0.3

Figure 1 (Supplementary Material): Determination of the catalytic parameters of wild type *th*TrmI enzyme.

(a) Time course measurements of incorporated ^{3}H-Methyl in $tRNA_i^{Met}$ for different AdoMet concentrations (0.5, 1, 2, 4, 8 µM). (b) Michaelis-Menten fit of initial rates corresponding to several AdoMet concentrations.

