

Structural characterization of *B. subtilis* m¹A₂₂ tRNA methyltransferase TrmK: insights into tRNA recognition

Clément Dégut^{1,†}, Martine Roovers^{2,†}, Pierre Barraud^{1,3}, Franck Brachet¹, André Feller⁴, Valéry Larue¹, Abdalla Al Refai⁴, Joël Caillet³, Louis Droogmans^{4,*} and Carine Tisné^{1,3,*}

¹Laboratoire de Cristallographie et RMN biologiques, CNRS, Université Paris Descartes, Sorbonne Paris Cité, 4 avenue de l'Observatoire, 75006 Paris, France, ²Institut de Recherches Labiris, B-1070 Bruxelles, Belgium, ³Laboratoire d'Expression génétique microbienne, CNRS, Univ. Paris Diderot, Sorbonne Paris Cité, Institut de Biologie Physico-Chimique, IBPC, 13 rue Pierre et Marie Curie, 75005 Paris, France and ⁴Laboratoire de Microbiologie, Université libre de Bruxelles (ULB), 6041 Gosselies, Belgium

Received December 13, 2018; Revised March 04, 2019; Editorial Decision March 21, 2019; Accepted March 26, 2019

ABSTRACT

1-Methyladenosine (m¹A) is a modified nucleoside found at positions 9, 14, 22 and 58 of tRNAs, which arises from the transfer of a methyl group onto the N1-atom of adenosine. The *yqfN* gene of *Bacillus subtilis* encodes the methyltransferase TrmK (B_STrmK) responsible for the formation of m¹A₂₂ in tRNA. Here, we show that B_STrmK displays a broad substrate specificity, and methylates seven out of eight tRNA isoacceptor families of *B. subtilis* bearing an A₂₂. In addition to a non-Watson–Crick base-pair between the target A₂₂ and a purine at position 13, the formation of m¹A₂₂ by B_STrmK requires a full-length tRNA with intact tRNA elbow and anticodon stem. We solved the crystal structure of B_STrmK showing an N-terminal catalytic domain harbouring the typical Rossmann-like fold of Class-I methyltransferases and a C-terminal coiled-coil domain. We used NMR chemical shift mapping to drive the docking of B_StRNA^{Ser} to B_STrmK in complex with its methyl-donor cofactor S-adenosyl-L-methionine (SAM). In this model, validated by methyltransferase activity assays on B_STrmK mutants, both domains of B_STrmK participate in tRNA binding. B_STrmK recognises tRNA with very few structural changes in both partner, the non-Watson–Crick R₁₃–A₂₂ base-pair positioning the A₂₂ N1-atom close to the SAM methyl group.

INTRODUCTION

Transfer RNAs (tRNAs) contain numerous modified nucleosides formed post-transcriptionally by a variety of enzymes (1). Amongst nucleoside modifications, methylations are the most frequently occurring and position-wise diverse. Their formation is catalysed by methyltransferases (MTases) which most-commonly use S-adenosyl-L-methionine (SAM) as the methyl donor.

The tRNA core contains many different modified nucleosides, notably all the 1-methyladenosine (m¹A) found in tRNA. The formation of m¹A occurs by transfer of the SAM-methyl group onto the N1-atom of adenosine. It is found at positions 9, 14, 22 and 58 of tRNAs (1,2), and is also formed at position 57 as an intermediate in the biosynthesis of 1-methylinosine (3,4). The m¹A₉, m¹A₁₄ and m¹A₅₈ nucleosides are involved in non-Watson–Crick base-pairing that are important to assemble the tRNA elbow and to maintain the D- and the T-loops in close contact. Different enzymes are responsible for the m¹A formation at the different positions in tRNAs. The m¹A at position 9 (m¹A₉) in human mitochondrial and archaeal tRNAs is formed by the enzyme Trm10 which belongs to the SPOUT family of MTases (5,6). Remarkably, yeast Trm10 forms 1-methylguanosine at position 9 (m¹G₉) (7) and the Trm10 enzymes from human mitochondria and from several archaea display dual specificity, forming both m¹A₉ and m¹G₉. (5,6,8). In human mitochondrial tRNA^{Lys}, m¹A₉ prevents the formation of an alternative structure of this tRNA by hampering a Watson–Crick base-pairing between A₉ and U₆₄ (9). The presence of m¹A₁₄ has only been reported in a limited number of mammalian cytoplasmic tRNA^{Phe} (10) and the gene encoding the corresponding MTase is still unknown. In contrast, the m¹A₅₈ mod-

*To whom correspondence should be addressed. Tel: +33 1 58 41 50 13; Email: carine.tisne@cnrs.fr

Correspondence may also be addressed to Louis Droogmans. Tel: +32 2 650 99 22; Email: ldroogma@ulb.ac.be

†The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

ification is found in tRNAs from organisms belonging to all domains of life. The genes and corresponding enzymes responsible for this modification (Trm6-Trm61 in Eukarya and TrmI in Bacteria and Archaea) have been identified and structurally characterised in yeast (11), in human (12), in the bacteria *Thermus thermophilus* (13–17) and *Mycobacterium tuberculosis* (18), and in the archaeon *Pyrococcus abyssi* (4,19).

The presence of m^1A_{22} in tRNA is rather scarce and its formation has been much less studied compared to m^1A_9 and m^1A_{58} . For organisms where tRNA sequences are available (1), m^1A_{22} is found in tRNAs of *Bacillus subtilis* (tRNA^{Ser}, tRNA^{Tyr}, both with large variable regions), of *Geobacillus stearothermophilus* (tRNA^{Leu} and tRNA^{Tyr}), of *Mycoplasma capricolum* (tRNA^{Gln}, tRNA^{Cys}, tRNA^{Glu}, tRNA^{His}, tRNA^{Leu}, tRNA^{Ser} and tRNA^{Tyr}) and of *Mycoplasma mycoides* (tRNA^{Ser}). Early studies showed that a SAM-dependent m^1A_{22} MTase activity was present in *B. subtilis* extracts (20,21). More recently, the *yqfN* gene of *B. subtilis* (now *trmK*) was shown to encode the enzyme TrmK responsible for m^1A_{22} formation in tRNA (22). TrmK belongs to the COG2384 protein family and orthologs are found in Gram-positive and -negative bacteria, without any equivalent found in Eukarya or Archaea. A *B. subtilis* mutant in which the *trmK* gene has been inactivated showed no detectable phenotype, neither at growth nor at sporulation level (22). However in the bacteria *Streptococcus pneumoniae* and *M. mycoides*, the TrmK ortholog is essential for survival of the organism (23,24). TrmK is well conserved among pathogenic bacteria, such as *Vibrio cholerae*, *Listeria monocytogenes*, *Staphylococcus aureus*, *S. pneumoniae*, and displays high sequence identity across the family. There are currently three X-ray crystal structures available for COG2384 family members (Protein data bank (PDB): 3LEC, 3GNL, 3KR9) (25) with no further data published on these.

Here, we studied the relationship between the structure and the function of the *B. subtilis* m^1A_{22} MTase TrmK (B_s TrmK) with special focus on the enzyme-tRNA recognition. We investigated the properties that establish tRNAs as substrates of B_s TrmK, and we report how a single point mutation in tRNA can convert a non-substrate into a substrate. The identified nucleotide playing a central role in B_s TrmK substrate definition is involved in base-pairing with A_{22} , the target site of methylation by B_s TrmK. We also solved the crystal structure of B_s TrmK and used NMR spectroscopy to gain insight into the recognition mode of *B. subtilis* tRNA^{Ser} (B_s tRNA^{Ser}) by B_s TrmK. Based on the NMR data, we constructed a docking model of the B_s TrmK/SAM/ B_s tRNA^{Ser} complex, the validation of which was performed with additional biochemical data. This work provides a clear picture of the relationship between structure and activity for the TrmK protein family.

MATERIALS AND METHODS

Cloning, expression and purification of B_s TrmK for NMR and X-ray studies

Recombinant B_s TrmK was expressed and purified as previously described (22). The *trmK* gene was amplified by PCR and cloned into the pCRII blunt vector and then

transferred into the pET28b expression vector, allowing T7 expression of an N-terminal His₆-tagged recombinant protein. For structural studies, a variant of B_s TrmK in which the two cysteine residues were replaced by serine ones (C35S and C152S) was used. Mutagenesis was performed using the QuikChange site-directed mutagenesis kit (Agilent). The presence of the desired mutation in *trmK* was checked by sequencing. This variant was overexpressed in the *E. coli* (BL21(DE3) strain). The induction was performed by adding 1 mM isopropyl- β -D-thiogalactopyranoside (IPTG) after having grown the bacteria to an optical density of 0.6 at 600 nm. Cells were harvested after incubation at 18°C during 24 h, pelleted and frozen at -80°C until further use. The frozen cells were suspended in 20 ml of a 50 mM Tris/HEPES buffer pH 8.2 containing 500 mM NaCl, 5% glycerol and 1 mM of phenylmethanesulfonylfluoride (PMSF). The suspension was sonicated and the lysate was centrifuged for 30 min at 15 000 g. The resulting supernatant was applied to a 5 ml Nickel Sepharose column (HisTrap, GE Healthcare) previously equilibrated in a 50 mM Tris-Cl buffer pH 8.0 containing 500 mM NaCl and 5% glycerol (equilibration buffer). The resin was then washed with 30 ml of buffer and the protein was eluted with a gradient of the equilibration buffer supplemented with 500 mM imidazole pH 8.0. The N-terminal His₆-tag was removed from B_s TrmK by thrombin cleavage (25 U thrombin/mg of protein) performed overnight at 4°C. 250 μM PMSF and 1 mM EDTA were then added to the protein sample. For X-ray studies, an additional step of hydrophobic chromatography was performed. The protein in 1 M ammonium sulphate was applied to a Phenyl Sepharose column (HiLoad 26/10 SP Sepharose, GE Healthcare). B_s TrmK was eluted by a 1 to 0 M ammonium sulphate reverse gradient. The sample was then concentrated (Millipore Amicon, Molecular weight cut-off 10 kDa) and injected on a size exclusion chromatography column (Superdex-75 26/60, GE Healthcare) equilibrated in a 50 mM Tris-Cl buffer pH 8.0 containing 500 mM NaCl and 2% glycerol.

For NMR experiments, the (C35S, C152S) B_s TrmK variant was purified as previously described (26). For the assignment procedure of NMR signals, samples of $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ -labeled B_s TrmK at 0.7 mM were prepared in a 50 mM sodium phosphate buffer pH 7.0, 500 mM NaCl and 2% glycerol. NMR chemical shift mappings were performed with samples of $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ -labeled B_s TrmK concentrated around 0.2 mM in a 50 mM Tris-Cl buffer pH 8.0 containing 100 mM NaCl, 10 mM MgCl_2 and 2% glycerol.

Production of *B. subtilis* tRNA^{Ser}

B_s tRNA^{Ser} was produced *in vitro* by T7 transcription and also *in vivo* in *E. coli* as described previously (26). To ensure sample homogeneity, prior to purification, *in vitro* transcribed B_s tRNA^{Ser} was extensively dialyzed against water and diluted 10-fold in water, then refolded by heating the sample to 95°C and slowly cooled down to room temperature. 10 mM of MgCl_2 was then added. Fragments of B_s tRNA^{Ser} were purchased from Dharmacon (GE Healthcare).

Analysis of B_s TrmK oligomerisation and subsequent MTase activity

Pure B_s TrmK (5 mg) was loaded on a Superose P12 column equilibrated in a 50 mM Tris–Cl buffer pH 8.0 supplemented with 10 mM MgCl₂ and 500 mM NaCl at a flow rate of 24 ml/h. The same sample in the presence of 10 mM DTT was also similarly analysed. The calibration of the column was assessed by loading in the same buffer a mixture of ovalbumine and bovine serum albumine (BSA). The MTase activity of each peak was determined by measuring the amount of ¹⁴C transferred to total *E. coli* tRNA using methyl-¹⁴C SAM as methyl donor. The reaction mixture was 200 μl: 70 μg of total *E. coli* tRNAs, 400 nCi of methyl-¹⁴C SAM (50 mCi/mmol), the fraction containing B_s TrmK and H₂O. Incubation was for 1 h at 37°C. The reaction was stopped by phenol extraction, and tRNA was TCA precipitated and captured on a GF/C Glass microfiber filter (Whatman) for scintillation counting.

Cloning of *B. subtilis* tRNA genes in pUC18 plasmid

The cloning of the *B. subtilis* genes of tRNA^{Cys}, -^{Gln}, -^{Leu}, -^{His}, -^{Tyr} and -^{Gly} were performed as for the cloning of tRNA^{Ser} (22). The T7 promoter was added immediately upstream of each tRNA gene to promote high levels of transcription. An MvaI restriction site was added downstream of the genes in order to obtain the 3'-CCA extremity of the tRNA transcripts. tRNA sequences are presented in Supplementary Figure S1.

Measurements of the MTase activity of B_s TrmK using ³²P-radiolabelled *B. subtilis* tRNAs

The measurements of MTase activity were conducted on the wild-type (WT) B_s TrmK in presence of 10 mM DTT. Radioactive (³²P) *in vitro* transcripts were obtained as previously described (22) using MvaI digested plasmids containing the tRNA genes, or mutated variants, as templates. Mutated tRNA genes were obtained by site directed mutagenesis (Agilent). [α-³²P]-ATP (3000 Ci/mmol) was purchased from Perkin Elmer and T7 RNA polymerase from Promega. Radioactive transcripts were purified by 10% polyacrylamide gel electrophoresis. The transcripts (300 000 cpm) were incubated for 15 min at 37°C in a 300 μl reaction mixture containing 0.2 mM SAM, 50 mM MOPS pH 6.5, 1 mM MgCl₂, 10 mM DTT and various amounts of B_s TrmK. They were subsequently ethanol precipitated and digested with nuclease P1. The resulting nucleotides were analysed by thin layer chromatography (TLC) on cellulose plates followed by autoradiography as described previously (4). Plates were visualised by autoradiography.

Steady-state kinetic analysis of B_s TrmK

[Methyl-³H]-SAM (15 Ci/mmol, Perkin Elmer) was mixed with non-radioactive SAM (Sigma) to achieve a specific radioactivity of about 500 cpm/pmol. The methylation kinetic assays were performed in 50 mM Tris–Cl pH 8.0 with 5 mM MgCl₂ at 37°C. Aliquots (300 μl) were removed after different incubation times and transferred into 5 ml of 5% (w/v) trichloroacetic acid (TCA) at room temperature for

10 min in order to quench the reaction and to precipitate the tRNA. The precipitates were collected by filtration using GF/C filters (Whatman). The filters were washed with ethanol, dried, and the radioactivity was measured by liquid scintillation counting for 2 min, resulting in a counting error below 4%. Data were corrected by subtracting the background radioactivity determined from a control without enzyme.

For the determination of the K_M and V_{max} of B_s TrmK for tRNA^{Ser}, the reaction mixtures contained B_s TrmK (100 nM), tRNA^{Ser} (10–2000 nM) and ³H-AdoMet (90 μM). Initial rates (v_i) for each substrate concentration (tRNA^{Ser}) were determined from the slope of linear fittings of time course data points (5, 10, 15, 20, 30 min). Enzyme parameters were obtained by non-linear least square fitting using Equation (1) for Michaelis–Menten kinetics. Confidence limits of the parameters at 90% were estimated by Monte-Carlo sampling using the MC-Fit program (27).

$$v_i = \frac{V_{max}}{2E_0} \left(K_M + S_0 + E_0 - \sqrt{(K_M + S_0 + E_0)^2 - 4S_0E_0} \right) \quad (1)$$

Preparation of B_s TrmK mutants and MTase assays

All the TrmK mutants were generated using the QuikChange mutagenesis kit (Agilent) and purified by affinity chromatography on Ni²⁺ loaded Chelating Sepharose as described above. The MTase assays were performed in 50 mM MOPS-Na pH 6.5, 1 mM MgCl₂, 10 mM DTT, 50 μg unfractionated tRNA from the *B. subtilis* (delta *trmK*) strain (22), 25 nCi [methyl-¹⁴C]-SAM (58 mCi/mmol, Perkin Elmer) and 0.2 μg enzyme in a total volume of 300 μl. Incubation was for 30 min at 37°C. The reaction was stopped by phenol extraction and the tRNA was precipitated by transferring the aqueous phase in 5 ml of 5% TCA. The precipitates were collected by filtration using GF/C filters (Whatman). The filters were washed with ethanol, dried, and the radioactivity was measured by liquid scintillation counting. The experiments were performed in triplicates.

Crystallisation and X-ray structure resolution of B_s TrmK

Crystallisation of B_s TrmK was performed at 4°C using the sitting-drop vapor-diffusion method. Protein samples were prepared at 3 mg/ml in 50 mM Tris–Cl buffer (pH 8.0) containing 500 mM NaCl and 2% glycerol. Drops of 1.6 μl were prepared by a Cybi-Disk robot mixing equal volume of protein and reservoir (100 μl). Crystals of B_s TrmK were first obtained in condition A10 of the Hampton Research Crystal Screen kit. After optimisation, diffracting crystals were finally obtained in 0.2 M ammonium acetate, 0.1 M sodium acetate trihydrate pH 5, 30% (w/v) polyethylene glycol 4000 and 8% 2-methyl-2,4-pentanediol using microseeding from needle clusters. Crystals were harvested and flash-frozen in liquid nitrogen. Diffraction data were collected at beamline ID14-1 of the European Synchrotron Radiation Facility (Grenoble, France).

X-ray diffraction data were processed using XDS and scaled with SCALA. The structure was solved by molecular replacement with PHASER using an average model of the structures of sp1610 (*Streptococcus pneumoniae* ortholog

of B_s TrmK, PDB code 3KR9), LMOF2365.1472 (*Listeria monocytogenes*, PDB code 3GNL) and of SAG1203 (*Streptococcus agalactiae*, PDB code: 3LEC). In the resulting model, non-conserved residues were mutated into alanine and model building was first performed with ARP/WARP using the warpNtrace automated procedure. Restrained refinements of the structure were performed with the program phenix.refine in the Phenix suite (28). Model and map visualizations for manual reconstructions were performed with the program Coot (29). The B_s TrmK structure has been deposited to the Protein Data Bank under the accession code 6Q56.

Normal mode analysis of B_s TrmK was conducted with the program ProDy (30) and visualized using VMD (31).

Isothermal titration microcalorimetry (ITC)

ITC experiments were carried out at 18°C in a MicroCal[®] ITC₂₀₀ microcalorimeter with B_s TrmK extensively dialyzed against 50 mM Tris-Cl pH 8.0, 100 mM NaCl, 10 mM MgCl₂, 2% glycerol. Twenty four injections of 1.6 μl of *S*-adenosyl-L-homocysteine (SAH) at 800 μM, dissolved in the buffer of B_s TrmK dialysis, were made into the cell containing B_s TrmK (58 μM). Injections were spaced by 180 s. The ITC data were analysed with the software ORIGIN[®] using a single set of sites model.

NMR chemical shift mapping

NMR spectra were recorded at 18°C on Bruker spectrometers (either 950, 800 or 600 MHz) equipped with TCI-5 mm cryoprobes. Backbone assignments were performed as previously described (32).

Chemical shift mappings of the interaction between SAH and B_s TrmK and B_s tRNA^{Ser} and B_s TrmK previously bound to SAH were obtained using ²H/¹³C/¹⁵N-labeled B_s TrmK (0.14 mM) mixed with 4 equivalents of SAH and two equivalents of B_s tRNA^{Ser} in a 3 mm NMR tube. The NMR chemical shift mapping was conducted in a 50 mM Tris-Cl buffer pH 8.0 containing 100 mM NaCl, 10 mM MgCl₂ and 2% glycerol. For each mixture, one TROSY experiment (33) was recorded. TOPSPIN, mddNMR (34) and Sparky software were used to process and analyse NMR data. Chemical shift differences Δ(H,N) were derived from ¹H and ¹⁵N shift differences: $\Delta(H,N) = \sqrt{[(\Delta^{15}N W_N)^2 + (\Delta^1H W_H)^2]}$, where $\Delta = \delta_{\text{complex}} - \delta_{\text{free}}$ (difference of chemical shift between the complex and the free states) and $W_H = 1$ and $W_N = 1/5$.

Molecular docking of B_s TrmK/SAM/tRNA^{Ser}

B_s tRNA^{Ser} model was built using the *T. thermophilus* tRNA^{Ser} crystal structure (PDB: 1SER) as template except for the variable region that was not defined in this structure. To model the variable region of B_s tRNA^{Ser}, the Human tRNA^{Sec} crystal structure (PDB: 3A3A) was used. It was manually mutated to the B_s tRNA^{Ser} sequence using Coot (29). Model geometry was then idealised with 300 iterations of rna_idealise from the Rosetta package, and subsequently minimised using the Rosetta force field through rna_minimise (35). Docking of *B. subtilis* tRNA^{Ser}

on B_s TrmK/SAM structure was performed using HADDOCK 2.2 (36,37). HADDOCK program is particularly well-suited to deal with data from NMR chemical shift mapping through the possibility to introduce ambiguous restraints (i.e. restraints that are applied between residues and not between atoms). Restraints were generated based on NMR chemical shift mapping upon B_s tRNA^{Ser} binding. Residues from B_s TrmK with NMR chemical shift variations of their amide group superior to 0.03 ppm upon tRNA^{Ser} binding (8, 9, 13, 37, 46, 48, 49, 57, 67, 78, 80, 82, 83, 88, 93, 99, 101, 119, 123, 124, 144, 145, 147, 149) were defined as active residues whereas surrounding residues of the surface (3, 7, 10, 11, 14, 27, 28, 29, 30, 31, 32, 51, 52, 53, 54, 55, 58, 59, 60, 67, 75, 94, 95, 96, 97, 98, 10, 146, 150, 151, 153) were defined as passive. The A₂₂ nucleotide, the target of B_s TrmK for methylation, was defined as active residue for the B_s tRNA^{Ser} whereas, the nucleotides 12 to 23 of the D-arm were defined as passive ones. A 1.8 to 3.2 Å distance constraint was applied between the N1-atom of A₂₂ and the SAM-methyl group. The HADDOCK modeling program allows any of the atoms from each residue designated as an ambiguous interaction restraint to interact with the nucleotides 12–23 of the tRNA^{Ser} D-arm with the constraint that N1 of A₂₂ is placed near the methyl-group of SAM. Ten thousand structures of the complex were generated during the rigid-body energy minimization, and the 400 best solutions based on the intermolecular energy were used for the semiflexible refinement and were subsequently refined in the explicit solvent iteration. Semi-flexible zones (automatically determined by HADDOCK) and full-flexible ones defined on B_s TrmK loops (residues 3–10, 49–56, 93–99, 121–127, 146–153) were used. We set the Edesol weight to 0 for all stages of the docking as it is recommended for protein–nucleic acids docking.

The 400 structures were then sorted into clusters by calculating the RMSD_{RNA} that we define as the RMSD calculated on the B_s tRNA^{Ser} phosphate backbone between two models relative to the B_s TrmK backbone. In other words, we compare the positioning of the B_s tRNA^{Ser} on B_s TrmK in the different models by superimposing the models along the B_s TrmK backbone. To belong to the same cluster, two models must have a RMSD_{RNA} below 12 Å, this cut-off was selected to differentiate between major orientation differences but allows for small rocking move of the tRNA on the protein surface. HADDOCK score was then used to identify best solutions, to rank models, and to check consistency of each cluster (Supplementary Figure S12). The atomic coordinates of the top-scoring model of B_s TrmK/SAM/ B_s tRNA^{Ser} complex are available as Supplementary data, TrmK_SAM_tRNASer.pdb.

RESULTS

B_s TrmK displays a broad tRNA substrate specificity

Amongst the DNA sequences of all the tRNA genes of *B. subtilis*, eight different tRNA isoacceptor families bear an A at position 22. For only two of these tRNAs, i.e. tRNA^{Tyr} and tRNA^{Ser}, the modification pattern has been determined, and both carry the modified nucleoside m¹A₂₂ (1). To determine if the remaining *B. subtilis* tRNAs (tRNA^{Cys}, ^{-Gln}, ^{-Leu}, ^{-His}, ^{-Glu} and ^{-Gly}) could undergo m¹A₂₂ for-

mation catalysed by B_s TrmK, the gene of one representative of each of the tRNA isoacceptor families (Supplementary Figure S1) was cloned into the pUC18 vector downstream of a T7 promoter to allow *in vitro* transcription. These tRNAs present a large diversity of sequence, notably at the level of the variable region, e.g. tRNA^{Cys} with a short variable region and tRNA^{Ser} with a very long variable region. Transcripts were generated in the presence of α ³²P-ATP and used in an *in vitro* MTase assay with different amounts of purified recombinant B_s TrmK (Figure 1). tRNA^{Leu} showed a two-fold decreased m¹A₂₂ formation (Figure 1G) compared to the other substrate tRNAs, and tRNA^{Gly} was not methylated at all by B_s TrmK (Figure 1H). Therefore, all *B. subtilis* tRNAs with an A₂₂, except tRNA^{Gly}, could be methylated, with no discrimination between the length of the variable region, indicating that these tRNAs could also be N1-methylated on A₂₂ *in vivo*. This observation would mean that the occurrence of m¹A₂₂ is underestimated in tRNA modification databases (1,38), due to a lack of systematic mapping of tRNA modifications in *B. subtilis* and in Bacteria in general.

B_s TrmK requires a purine at position 13 of the tRNA to catalyse m¹A₂₂ formation

A comparison of the sequences of substrate (tRNA^{Ser}, -Tyr, -Cys, -Gln, -Leu, -His and -Glu) and non-substrate tRNAs (tRNA^{Gly}) highlights two nucleotides, at positions 13 and 26, as potentially implicated in substrate discrimination by B_s TrmK. Both positions are occupied by a purine (G or A) in all B_s TrmK substrates, and by U₁₃ and C₂₆ in the non-substrate. To mimic the substrates, nucleotides at these positions in the non-substrate tRNA^{Gly} were replaced by either G₁₃, A₂₆, or both (Supplementary Figure S2). Adenosine was chosen for position 26, as this is the most frequently appearing nucleoside at this position in B_s TrmK substrates. Both single-mutated tRNA^{Gly} (U₁₃G and C₂₆A) turned the previous non-substrate into a substrate for B_s TrmK, with U₁₃G increasing the amount of methylated A₂₂ the most (Supplementary Figure S2). Compared to the variants bearing a single mutation, the double-mutant (U₁₃G/C₂₆A) proved to be an even better substrate for B_s TrmK.

The purine at position 13, important for m¹A₂₂ formation catalysed by B_s TrmK, forms a non-Watson–Crick base pair, namely a Hoogsteen–Sugar base-pair, with the target nucleotide A₂₂ (as seen for instance in the structure of *T. thermophilus* tRNA^{Ser} (39), PDB: 1SER). To further investigate the importance of the nucleotide at position 13 for m¹A₂₂ formation, three variants of tRNA^{Ser} (G at position 13, long variable region) and tRNA^{His} (A at position 13, short variable region) were produced with varying nucleotides at position 13, i.e. A, U or C for tRNA^{Ser}, and G, U or C for tRNA^{His}, and tested as B_s TrmK substrates (Figure 2A and B). In agreement with the result obtained with tRNA^{Gly}, a pyrimidine at position 13 (U or C) reduced the substrate potential of tRNA^{Ser} and tRNA^{His} for B_s TrmK. However this reduction is more pronounced for tRNA^{Ser} indicating that other still unidentified elements are important for TrmK/tRNA recognition. Overall, the presence of purines at positions 13 and 26 renders tRNA a better substrate for B_s TrmK.

B_s TrmK requires the intact three-dimensional structure of tRNA to catalyse m¹A₂₂ formation

The tRNA molecule adopts a well-known L-shape 3D structure, in which the corner of the L, known as the tRNA elbow (40), is made up of tertiary interactions between the D- and T-loops (41). This network of interactions involves base-pairings of T₅₄ with A₅₈, G₁₈ with Ψ ₅₅ and G₁₉ with C₅₆, and a further stacking of the purine at position 57 between base-pairs G₁₈– Ψ ₅₅ and G₁₉–C₅₆. To determine if B_s TrmK recognises a canonical tRNA L-shape structure, these crucial interactions between the D- and T-loops were abolished by introducing the following mutations in tRNA^{His}: U₅₄C, U₅₅G/C₅₆G or U₅₄C/U₅₅G/C₅₆G (Figure 2C). The triple-mutant of tRNA^{His} (U₅₄C/U₅₅G/C₅₆G) showed very low substrate potential compared to WT tRNA^{His}, suggesting that B_s TrmK recognises a canonical L-shaped structure for tRNA binding. In a similar manner, we tested the substrate potential of an RNA transcript covering the D-arm sequence or the D-arm sequence plus the anticodon stem-loop of B_s tRNA^{Ser}. None of these transcripts were substrates for B_s TrmK (data not shown).

Next, the importance of the anticodon stem-loop for tRNA recognition by B_s TrmK was evaluated. In order to determine whether B_s TrmK interacts with tRNA anticodon loops or anticodon stems, five mutated tRNA^{His} constructs by altering either the loop or the stem were generated. In the loop, variants with a point mutation (U₃₃G), an insertion of a G near position 36 (+G_{36a}), and deletion of G₃₇ (Δ G₃₇) were produced. In the stem, a variant with three additional GC base pairs (C–G, G–C and C–G between A₂₉–U₄₁ and G₃₀–C₄₀) (ac stem +3 bp), and a variant of tRNA^{His} in which the entire anticodon stem-loop was removed (Δ (ac stem-loop), Supplementary Figure 1) were generated. None of the alterations in the anticodon-loop resulted in reduced m¹A₂₂ formation (Figure 2D). However, the Δ (ac stem-loop) construct showed a severe reduction in m¹A₂₂ formation (Figure 2D), suggesting that B_s TrmK interacts with the anticodon stem of tRNA. This interaction most probably occurs through non-nucleotide-specific interactions since the (ac stem +3bp) construct remains modified and since the sequences of the anticodon stems are not conserved across B_s TrmK substrates.

Altogether, these data show that B_s TrmK displays a broad tRNA substrate specificity. B_s TrmK methylates full-length tRNAs with a purine at position 13 and an intact tRNA elbow structure. It also requires an anticodon stem that is probably recognized in a non-sequence-specific manner.

B_s TrmK is active as a monomer

We previously showed that tetramerisation of bacterial m¹A₅₈ tRNA MTase (TrmI) was crucial for tRNA recognition (14,15). This prompted us to investigate the functional oligomeric state of B_s TrmK by size exclusion chromatography (SEC) by measuring the MTase activity of each eluted fraction along the chromatogram (Supplementary Figure S3). On the SEC column, B_s TrmK eluted both as monomers and higher oligomeric species. The activity measurements of each eluted fraction showed that B_s TrmK is functional as a monomer. The higher oligomeric states of B_s TrmK are formed via disulphide bonds involving the two cysteines in

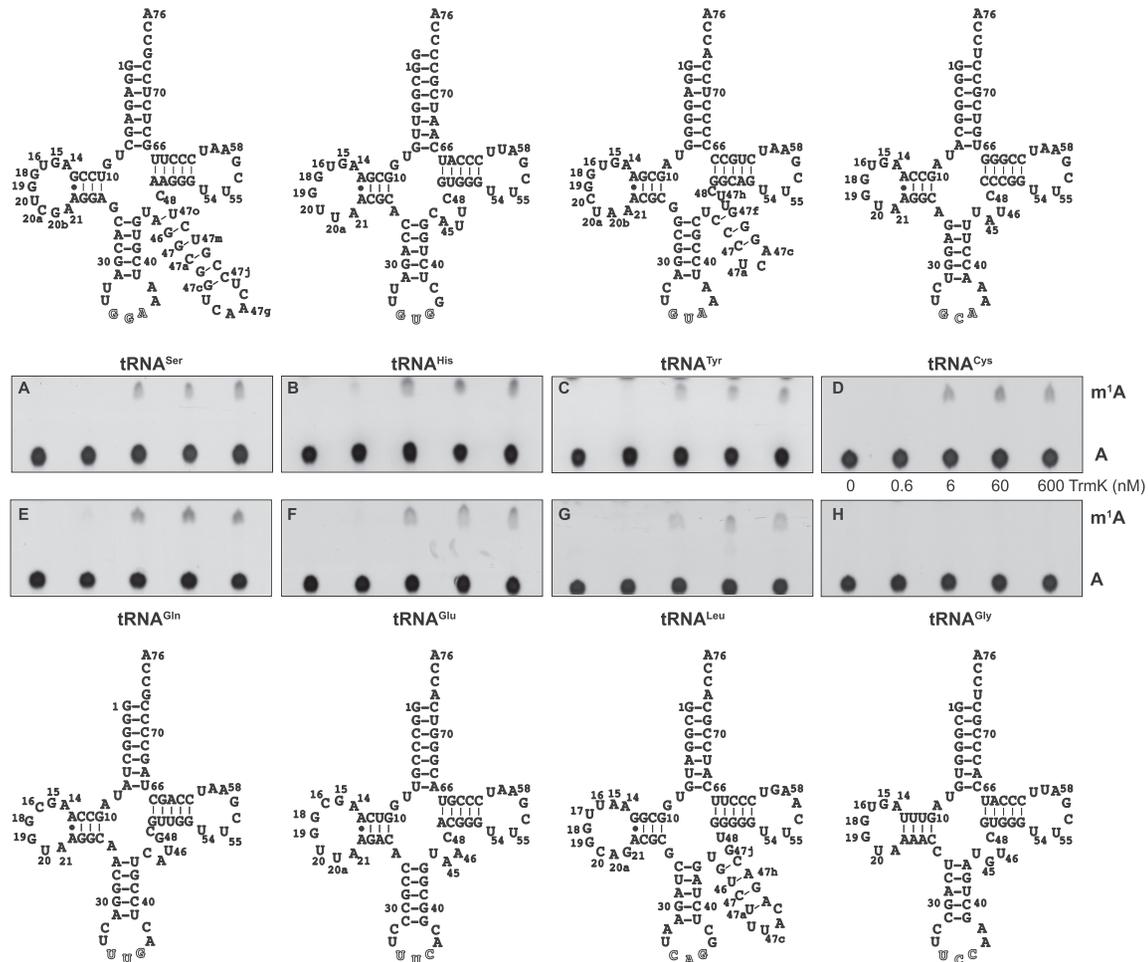


Figure 1. *B. subtilis* TrmK methylates a variety of *B. subtilis* tRNAs. Autoradiograms of chromatograms of P1 hydrolysates of [α -³²P]ATP-labelled T7-transcripts of *B. subtilis* tRNAs incubated during 15 minutes with SAM and increasing amounts of purified *B. subtilis* TrmK (from 0 to 600 nM). The sequences of the tRNAs tested as substrates are drawn above and below the autoradiograms. (A) tRNA^{Ser}, (B) tRNA^{His}, (C) tRNA^{Tyr}, (D) tRNA^{Cys}, (E) tRNA^{Gln}, (F) tRNA^{Glu}, (G) tRNA^{Leu}, (H) tRNA^{Gly}.

B. subtilis TrmK sequence at positions 35 and 152. Such bonds can be broken by addition of a reducing-agent, and addition of DTT to the MTase reaction buffer resulted in a dramatic increase of the enzymatic activity (Supplementary Figure S3), confirming that *B. subtilis* TrmK functions as a monomer in solution.

Crystallisation and structure determination of *B. subtilis* TrmK

No crystals of *B. subtilis* TrmK could be obtained either in the presence or absence of DTT. To facilitate crystallisation, the cysteines 35 and 152 were mutated into serine residues. The MTase activity of this mutant was equal to WT *B. subtilis* TrmK (Supplementary Figure S4). Therefore, all following structural studies were performed with this construct.

Unlike WT *B. subtilis* TrmK, we obtained crystals of this double-mutant, solved its crystal structure by molecular replacement and refined it to a resolution of 1.98 Å with a final R_{free} factor of 24% (Supplementary Table S1). Four molecules are present in the asymmetric unit (ASU) with a solvent percentage of 38.6% (Supplementary Figure S5A). *B. subtilis* TrmK displays an L-shaped structure (Figure 3) which allows two

2-fold symmetry related molecules to face each other in such a way that it hides the catalytic pocket of each protein. The four molecules in the ASU show little conformational difference between them (RMSD of 0.25 Å over all C α atoms). The superimposition of the *B. subtilis* TrmK structure with its orthologs (PDB: 3KU1, 3GNL, 3KR9) gave an RMSD of 1.5 Å over around 220 C α atoms, consistent with the high sequence identity of these proteins.

B. subtilis TrmK consists of an N-terminal Class I MTase domain linked to a C-terminal coiled-coil domain

The *B. subtilis* TrmK protein contains two domains (Figure 3A). The N-terminal catalytic domain, that harbours the binding site of the SAM cofactor, adopts a typical Rossmann-like fold of Class I MTases (RFM; Class-I MTases) with a twisted central β -sheet flanked by six α -helices. In the β -sheet, the five first β -strands are parallel and the last two are antiparallel. The catalytic core shows a high level of sequence conservation across COG2384 proteins (Figure 3B and Supplementary Figure S6). A DALI search (42) in the PDB with the catalytic domain of *B. subtilis* TrmK identified, apart

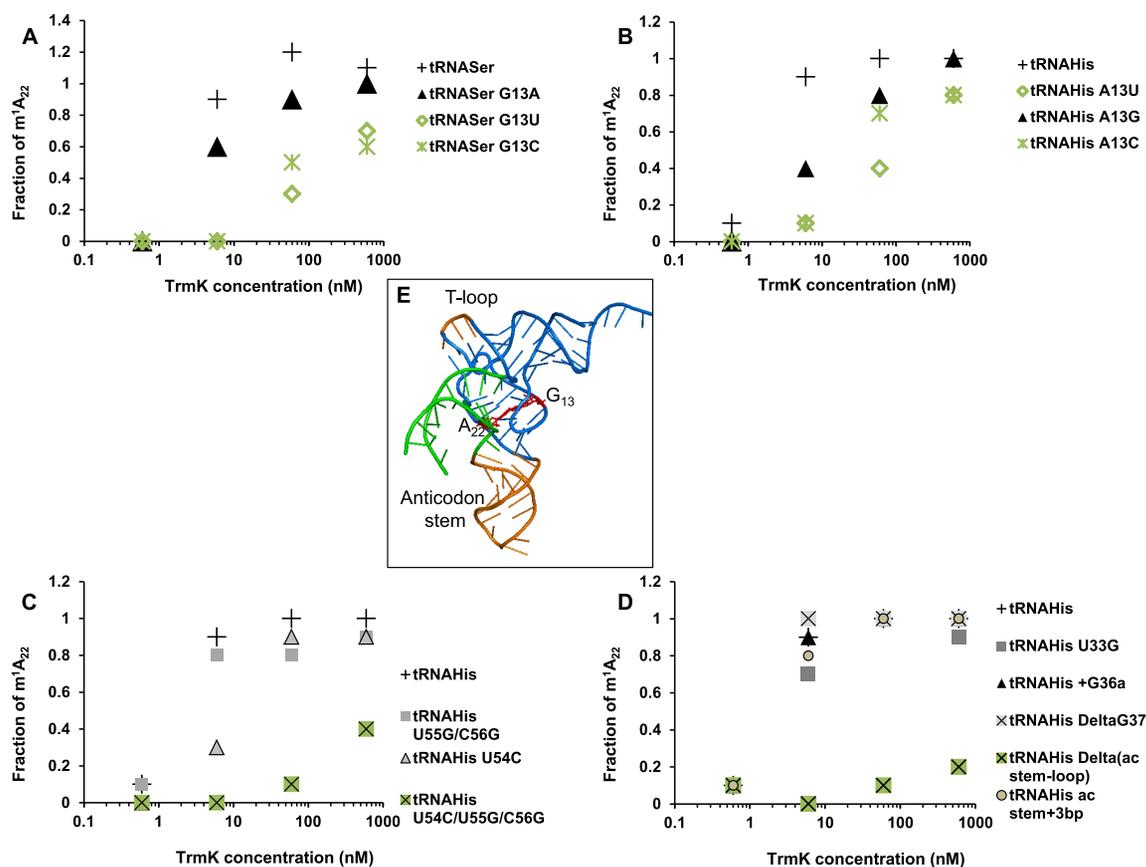


Figure 2. *Bs* TrmK MTase assays with *B. subtilis* tRNAs mutated at position 13, and with tRNA^{His} altered in the T-loop or in the anticodon stem-loop. [α^{32} P]ATP-labelled T7-transcripts of mutated tRNA^{Ser} (A) or tRNA^{His} (B–D) of *B. subtilis* were incubated during 15 min with SAM and increasing amounts of purified *Bs* TrmK. PI hydrolysates were separated by thin layer chromatography. After autoradiography the radioactivity of the spots corresponding to pA and pm¹A was measured by scintillation counting. The fraction of m¹A obtained was plotted versus the corresponding *Bs* TrmK concentration in the reaction mixture (A) for mutants of tRNA^{Ser} at position 13, (B) for mutants of tRNA^{His} at position 13, (C) for variants of tRNA^{His} in the T-loop, (D) for variants of tRNA^{His} in the anticodon stem-loop. (E) Summary on the tRNA^{Ser} structure of the determinants required for *Bs* TrmK MTase activity: the length of the variable region (in green) is not important, the base-pair A₂₂-N₁₃ (in red) involves a purine at position 13, an intact 3D structure is necessary notably at the level of the T-loop (in orange) and anticodon stem (in orange).

from COG2384 proteins, a ribosomal protein L11 MTase (PDB: 3GRZ, Z-score 19, RMSD 2.1 Å over 147 C α), the m²G₆ tRNA MTase Trm14 (PDB: 3TM4, Z score 17.4, RMSD 2.4 Å over 152 C α), the m¹A₅₈ tRNA MTase Trm1 (PDB: 2PWY, Z-score 15.8, RMSD 2.4 Å over 151 C α) and the m¹G₃₇ tRNA MTase Trm5 (PDB: 2ZZM, Z-score 14.5, RMSD 2.9 Å over 147 C α). These proteins are all Class-I type RFMs, confirming that *Bs* TrmK belongs to this class of proteins, but no obvious information for catalysis or substrate recognition could be drawn from these similarities. The C-terminal domain of *Bs* TrmK contains four α -helices, the two largest of which form a coiled-coil motif. The domain displays low sequence homology across COG2384 proteins (Figure 3B and Supplementary Figure S6), and a DALI search did not reveal any ideal protein-fold matches. However, two proteins with coiled-coil domain were identified with low Z-scores: the seryl-tRNA synthetase (PDB: 1SRY, Z-score 6.0, RMSD 1.1 Å over 50 C α) and the ribosomal protein L19 (PDB: 1S72, Z-score 6.1, RMSD 2.0 Å over 53 C α). In these proteins, the encompassed coiled-coil motif binds an RNA, suggesting that in *Bs* TrmK the C-terminal domain could also play such a role. The N- and

C-terminal domain of *Bs* TrmK are held together by numerous contacts, and a normal mode analysis of the *Bs* TrmK structure confirms a lack of free movement between the two domains by identifying only the very end of the coiled-coil motif as flexible (Figure 3D).

The interaction between *Bs* TrmK and the cofactor product SAH was studied by ITC. The interaction is enthalpy-driven ($\Delta H \approx -17$ kcal/mol) with a single binding site and a dissociation constant (K_D) of 1.7 μ M (Supplementary Figure S7). Although SAM or SAH were present in crystallisation trials, crystals were of apo *Bs* TrmK, *i.e.* no ligand bound. The electron density is well-defined across the *Bs* TrmK structure and particularly for residues constituting the SAM-binding site, even in the absence of the cofactor. Only two residues, K151 and Y153, near the active site, show poor electron density suggesting mobility (Supplementary Figure S5B). SAM was modelled into the active site of *Bs* TrmK, guided by the crystal structure of SAM-bound TrmK from *S. pneumoniae* (PDB: 3KU1).

In the *Bs* TrmK/SAM model, the SAM cofactor is bound at the centre of the catalytic domain in a pocket with residues harbouring negative electrostatic potentials con-

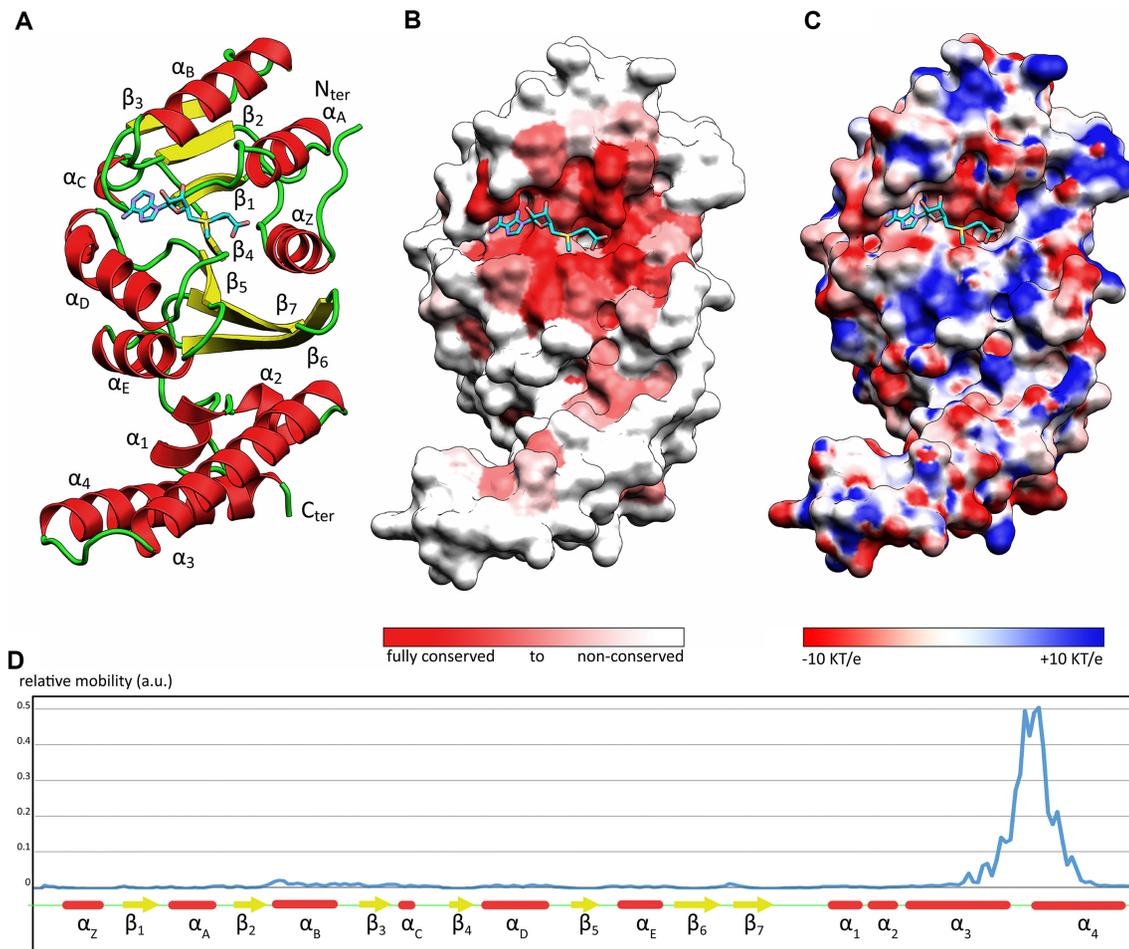


Figure 3. Crystal structure of B₅ TrmK. SAM was modelled in the catalytic center of B₅ TrmK by superposing the B₅ TrmK structure with that of its ortholog from *S. pneumoniae* (PDB code 3KU1). It is represented as sticks. (A) Ribbon representation of B₅ TrmK, α -helices are coloured in red whereas β -strands are in yellow. (B) Location on the structure of B₅ TrmK of conserved residues, the sequence of *B. subtilis* TrmK was aligned with those of the representative members of the COG2384 family, (C) Representation of electrostatic surface potentials of TrmK, Positive charges are in blue whereas negative ones are in red with the maximum color saturation corresponding to -5 kT/e (red) and $+5$ kT/e (blue). The figure was prepared with the APBS PyMOL plug-in (52) and pdb2pqr webserver (53). (D) Normal mode analysis of B₅ TrmK showing the mobility of the last part of the C-terminal domain. The mean mobility of residues in the six first non-trivial computed normal modes is plotted as a function of the B₅ TrmK residue number.

served amongst COG2384 proteins (Figure 3B). The methyl group is pointing towards a region of positively charged and well conserved residues, likely important for tRNA binding (Figure 3C). To verify the model, three well-conserved residues (Supplementary Figure S6) directly interacting with SAM, *i.e.* R9, D78 and M96 were mutated to alanine, and the activity of these B₅ TrmK variants were tested. Each point-mutation resulted in a dramatic loss of B₅ TrmK MTase activity, confirming the importance of these residues in SAM binding. Indeed, B₅ TrmK R9A is totally inactive, and B₅ TrmK M96A and B₅ TrmK D78A mutants have almost completely lost MTase activity (Supplementary Figure S8). To further characterise the cofactor binding site, an NMR chemical shift mapping (Supplementary Figure S9) of the SAH binding on B₅ TrmK was performed. The NMR chemical shift mapping monitors the binding of SAH on the signals of the ¹⁵N-labelled amide groups of B₅ TrmK that were previously assigned (32), and chemical shift variations and peak disappearances resulting from SAH interaction can subsequently be used to map the SAH-binding

site on B₅ TrmK structure. This NMR mapping identified, for instance, residues R9, L10, G77, D78, A94 and G95 as involved in SAH binding, in agreement with the modelled SAM in the B₅ TrmK structure (Supplementary Figure S9C).

NMR chemical shift mapping of the binding interface between B₅ TrmK and B₅ tRNA^{Ser}

Some post-transcriptional modifications were reported to stabilise and help fold tRNA structures. Therefore, we investigated if modifications in B₅ tRNA^{Ser} could make this a better substrate for B₅ TrmK, compared to the unmodified B₅ tRNA. No homologue of TrmK is present in *E. coli*, enabling us to produce the tRNA^{Ser} in this host, bearing all modifications formed by the *E. coli* enzymes *i.e.* Gm₁₈, D₁₉, T₅₄ and Ψ_{55} (Figure 4A), but lacking m¹A₂₂. The *B. subtilis* tRNA^{Ser} was overexpressed and purified in *E. coli* as previously described (26). The catalytic efficiency of B₅ TrmK towards this tRNA^{Ser} was then compared to that measured with an *in vitro* transcribed tRNA^{Ser} lacking modifications

(Figure 4B). This analysis showed that the tRNA^{Ser} produced in *E. coli* is less efficiently modified than the unmodified tRNA^{Ser}. The catalytic efficiencies, V_{\max}/K_M , for modified tRNA^{Ser} reached 44% of that obtained for the unmodified tRNA^{Ser}. The presence of post-transcriptional modifications in tRNA is likely to restrict m¹A₂₂ formation by B_STrmK. The *in vitro* transcribed tRNA^{Ser} was thus chosen for structural investigations.

Attempts to obtain co-crystals of unmodified B_StRNA^{Ser} with B_STrmK and SAH were unsuccessful, and we therefore turned to NMR spectroscopy to decipher the interaction of B_STrmK with B_StRNA in solution. No perturbation of B_STrmK resonances was observed for the couple B_StRNA^{Ser}/B_STrmK in the absence of SAH or for B_STrmK-SAH in the presence of an RNA mimicking the D-arm, or the D-arm plus the anticodon stem of B_StRNA^{Ser} (data not shown). The NMR chemical shift mapping was performed by recording a (¹H-¹⁵N) TROSY experiment on ²H-¹⁵N-¹³C-labeled B_STrmK bound to unlabeled SAH and unlabeled B_StRNA^{Ser} (Figure 4C, Supplementary Figure S10). The mapping monitors the binding of SAH and subsequently the binding of tRNA^{Ser} by recording NMR chemical shift variations of the amide groups of B_STrmK. Binding of B_StRNA^{Ser} to B_STrmK-SAH caused major perturbations on NMR signals of B_STrmK, *i.e.* disappearance of 2 peaks corresponding to residues V13 and L57 near the SAH binding pocket, and NMR chemical shift variations of 32 peaks of the N-terminal domain (Supplementary Figure S10B). The residues that experienced significant chemical shift variations upon binding of B_StRNA^{Ser}, for instance G49, G98, I83, I124, D149 and L156 in Figure 4C, outline the binding surface of B_StRNA^{Ser} (residues in red and orange, Figure 4C and D). Residues that do not show chemical shift variations upon binding to B_StRNA^{Ser} correspond to residues that do not take part to this interaction, *e.g.* Y167 (residues in blue, Figure 4C and D). Residues which disappear upon SAH addition are excluded from further monitoring (residues in grey, Figure 4C and D). The tRNA binding surface locates on a face of B_STrmK also covering the cofactor binding pocket. This surface contains a number of positively charged patches (Figure 3C) ideal for RNA binding.

No significant NMR chemical shift variations or disappearances of peaks could be observed within the C-terminal domain of B_STrmK (Figure 4D), indicating that this domain would not participate in tRNA binding. This is surprising as it is commonly thought that such domains in modification enzymes (*i.e.* domains surrounding the catalytic domain) are involved in RNA-binding, and other proteins with coiled-coil domains use this fold for RNA binding. Since the NMR chemical shift mapping uses the amide group resonances as probes, we cannot exclude that some residues with long side-chains like Arginine or Lysine could interact with B_StRNA^{Ser} in a way that would not be noticed by the backbone NMR chemical shift perturbation analysis. If this is the case, the binding surface of the C-terminal domain is limited to only a few residues, as it would otherwise have been detected by NMR. To investigate if the C-terminal domain of B_STrmK is dispensable for B_STrmK MTase activity, a B_STrmK variant lacking most of the C-

terminal coiled-coil motif was designed. This construct of B_STrmK is still well-folded as evidenced by 2D-NMR and able to bind SAH but exhibits no MTase activity (Supplementary Figure S11), suggesting that the coiled-coil motif of the C-terminal domain does, in fact, play a role in tRNA binding.

Lastly, the structure of B_StRNA^{Ser} upon binding of B_STrmK was investigated by NMR chemical shift mapping with a ¹⁵N-labeled tRNA (Supplementary Figure S12). The overlaps encountered in the NMR spectra did not allow us to assign all imino groups of B_StRNA^{Ser}, notably no assignment is available in the D-stem. However, since the assignment of the variable region is nearly complete, the titration data confirm that this region does not take part in the binding of B_STrmK. Upon binding of B_STrmK, most of the peaks do not experience any chemical shift variation indicating that the global folding of B_StRNA^{Ser} is conserved. Disappearances of peaks are observed at the top of the anticodon stem and at the junction between the acceptor and the T-stem for imino groups that are assigned.

tRNA is bound *via* interactions with both domains of B_STrmK

A model of the B_STrmK/SAM/B_StRNA^{Ser} complex was generated with the high-ambiguity-driven biomolecular docking program HADDOCK (36) using the data obtained from NMR as ambiguous interaction restraints (see Materials and Methods). Input data for HADDOCK calculations used B_STrmK residues with significant chemical shift perturbations as residue-level ambiguous interaction restraints, and further included the nucleotides 12 to 23 in the D-arm of tRNA^{Ser}, assuming that nucleotides around A₂₂ could interact with B_STrmK. A distance restraint between the N1-atom of A₂₂ and the methyl of SAM was also added.

Six clusters of models were extracted from HADDOCK calculations (Supplementary Figure S13), each exhibiting a very different relative orientation of the protein on the tRNA (Supplementary Figure S14). This reflects the feature that we do not have any restraint for the interaction of the C-terminal domain of B_STrmK with B_StRNA^{Ser}. The best model according to the HADDOCK scoring function is part of a cluster made of 102 structures with a mean RMSD_{RNA} of 4.2 Å containing most of the best-scoring structures. The 5 other clusters present best-scoring structures with much lower scores. In clusters 5 and 6, there is no interaction between the C-terminal domain of B_STrmK and the tRNA, which is not compatible with our data. In cluster 2, the contacts are located at the level of the anticodon loop, in cluster 3 at the level of the variable stem and in cluster 4 at the level of the D-loop and the T-stem. According to our data, clusters 2 and 3 are not valid. Cluster 4 is only represented by five structures and harbours a much lower binding energy than the best-scoring structure of cluster 1. The top-scoring model of the cluster 1 was thus chosen as the best model of the B_STrmK/SAM/B_StRNA^{Ser} complex (the atomic coordinates of the model are available as Supplementary data, TrmK_SAM.tRNASer.pdb).

In this model (Figure 5A), both domains of B_STrmK interact with B_StRNA^{Ser}, mainly through interactions with the phosphodiester backbone of the tRNA. The catalytic domain shares, to a large extent, a larger interaction surface

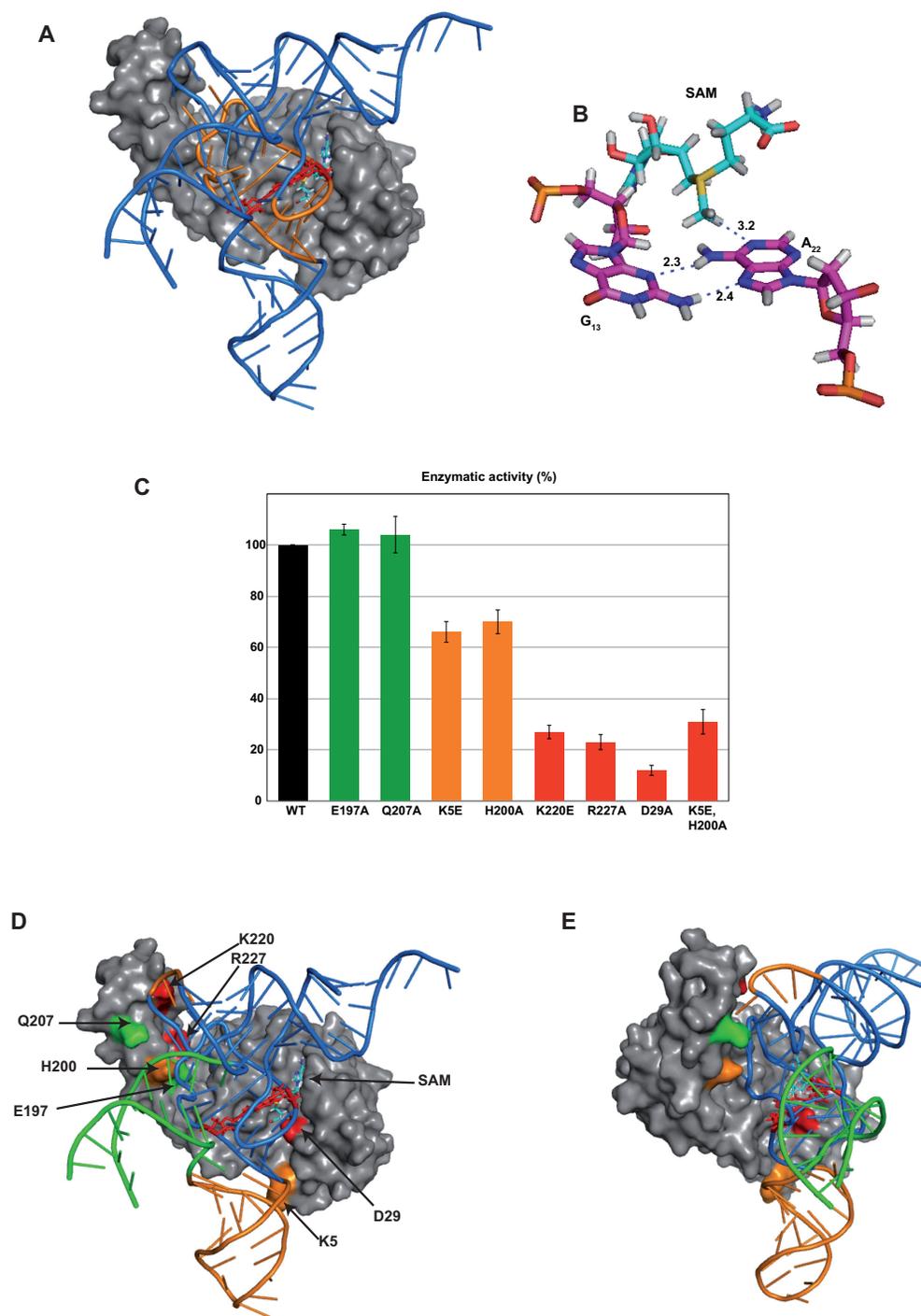


Figure 5. Model of the B_s TrmK/SAM/ B_s tRNA^{Ser} complex based on the NMR chemical shift mapping and validated by MTase activity of B_s TrmK variants mutated at the binding interface with tRNA. (A) Top-scoring structure from HADDOCK docking guided by the NMR chemical shift mapping of the B_s tRNA^{Ser} binding on B_s TrmK, the D-arm is drawn in orange, the G₁₃-A₂₂ base pair is in red sticks, the SAM on the B_s TrmK catalytic pocket is represented as sticks, (B) The base-pair G₁₃-A₂₂ is in close proximity to the SAM-methyl donor in the model (C) MTase activity of B_s TrmK variants mutated in the tRNA binding interface (D) Positions of the mutations on B_s TrmK/SAM/ B_s tRNA^{Ser} complex, the B_s tRNA^{Ser} is represented with the same color code as that used in Figure 2E, the tRNA elements crucial for B_s TrmK activity are in red and orange (E) Profile view of the model.

with the tRNA, compared to that of the C-terminal domain. Almost all nucleotides of the D-arm (G₁₃, A₁₄, G₁₅, U₁₆, G₁₉, G₂₃ and G₂₄) and two nucleotides of the anticodon stem (U₃₉ and G₄₁) interact with the catalytic domain of *B_s*TrmK whereas its C-terminal domain only interacts with nucleotides U₂₀ and C₅₆. Indeed, R227 and H200 are involved in the binding of U₂₀ and a surface cluster made by K220, Q216, N217 and Q207 interacts with C₅₆ in the T-loop of *B_s*tRNA^{Ser}. The few interaction points with the tRNA and the nature of the residues involved in interaction (long side-chains) could explain why no interaction was detected by NMR for this domain. In the model, the interaction between the coiled-coil motif of the C-terminal domain and the D- and T-loops helps maintain the position of A₂₂ near the catalytic pocket, explaining why removal of these helices abolished MTase activity. The base-pairing of A₂₂ with G₁₃ allows for positioning the N1-atom of A₂₂ in close proximity to the methyl group of SAM (Figure 5B).

In the proposed model, the variable region of the tRNA molecule does not interact with *B_s*TrmK, in agreement with the fact that the length of this region does not affect substrate potential. The anticodon stem and the tRNA elbow were shown to be important for *B_s*TrmK MTase activity, and in the model, these two regions interact directly with *B_s*TrmK. The model further suggests an explanation for the preference of a purine at position 13, a pyrimidine at this position would form hydrogen bonds to the Watson-Crick face of A₂₂, which would render the N1-atom non-accessible.

To further challenge the model for validation by experimental data, *B_s*TrmK residues in the binding interface were mutated to glutamate or alanine residues and the mutant proteins were tested for MTase activity (Figure 5C). Residues were selected according to the model as: (i) two N-terminal domain residues that are predicted to interact with the tRNA phosphodiester backbone in the anticodon stem (K5E) and in the D-stem (D29A), (ii) C-terminal domain residues predicted to interact with tRNA (H200A, Q207A, K220E, R227A), (iii) A residue not expected to be involved in tRNA binding (E197A) was added as a negative control. All mutant proteins are still folded (Supplementary Figure S15) and showed reduced MTase activity compared to WT *B_s*TrmK, except for the Q207A variant and the E197A one bearing a mutation of a residue outside of the tRNA binding surface. These results indicate that the mutated residues, except Q207 and E197, are important for *B_s*TrmK enzymatic activity, most probably by affecting tRNA binding given their position in *B_s*TrmK structure (Figure 5D and E). In cluster 4, among the mutated residues, only H200 and Q207 make contact with the tRNA, which comforts us in the choice of the best-scoring structure of cluster 1 as the best model.

Overall, the obtained model for the *B_s*TrmK/SAM/tRNA^{Ser} complex is supported by MTase activity data performed on single-point mutants of *B_s*TrmK. In the model, *B_s*TrmK recognises the overall L-shaped structure of tRNA, rather than specific nucleotides, via extensive interactions with the phosphodiester backbone of nucleotides in the D-loop, and point-interactions with the phosphodiester backbone of the T-loop and anticodon stem. This recognition pattern explains how multiple

tRNAs, with a variety of sequences, can be accommodated as substrates for *B_s*TrmK.

DISCUSSION

NMR spectroscopy was a valuable method to get insight into *B_s*TrmK/SAM/*B_s*tRNA^{Ser} interactions

X-ray crystallography is the method of choice for obtaining high-resolution structure of tRNA modification enzymes in complex with their tRNA substrates, but the crystallisation of these complexes remains a highly challenging task. Crystallisation is often limited by the flexibility of the tRNAs, the low binding affinity of the partners and the salt concentration needed in many crystallisation assays that weakens the often-electrostatic protein-RNA interaction. Heteronuclear NMR spectroscopy can be a valuable alternative to X-ray crystallography in the absence of diffracting crystals. This method can provide information on the binding interface between the enzyme and the cognate tRNA. In this work, NMR proved a powerful tool for deciphering the tRNA binding patch of *B_s*TrmK. To our knowledge, this is the first time that NMR has been used to obtain information on the binding interface between a modification enzyme and its full-length tRNA substrate. The strategy consisted in solving the crystal structure of *B_s*TrmK and using the measured NMR data to guide the docking of *B_s*tRNA^{Ser} into the protein to build a model of the *B_s*TrmK/SAM/*B_s*tRNA^{Ser} complex. This model was further validated by measuring the MTase activity of *B_s*TrmK variants with single-point mutations located at the proposed protein-RNA interaction surface.

The formation of m¹A₂₂ by *B_s*TrmK does not require a general base catalyst

In the methylation reaction mechanism, the N1-atom of adenine can act as a nucleophile either alone or assisted by a general base catalyst. From a chemical point of view, the N1-position of adenine is a powerful nucleophile which is easily methylated by an electrophile such as methylmethanesulfonate (43). On the other hand, SAM is a natural electrophile in the cell and it was even proposed that some m¹A formation in the cell can occur by direct reaction with SAM for RNAs that have accessible N1-atom of adenosines, especially in environments rich in SAM (43). For *B_s*TrmK, the only residue that is conserved across COG2384 proteins and that could act as a general base catalyst in the reaction mechanism is the aspartate D29. Mutation of D29 to alanine reduced the activity of *B_s*TrmK, but did not fully deactivate the enzyme (Figure 5A), indicating that D29 is not absolutely required for reaction. This result adds to the growing lines of evidence that m¹A formation in tRNA does not require a base catalyst and agrees with (i) the previously proposed m¹A methylation mechanism for the bacterial m¹A₅₈ tRNA MTase TrmI (2,14,17), and (ii) with data published on the dual-specific m¹G₉/m¹A₉ tRNA MTase Trm10 from *Thermococcus kodakaraensis* (8,44).

Enzymes modifying the tRNA core use different strategies to get access to the target nucleoside

Considering the published tRNA sequences (1,38), the tRNA core contains many different modifications, such as s^4U_8 , m^2G_{10} , Gm_{18} , m^1A_{22} , m^2G_{26} , m^7G_{46} , m^5C_{48} , m^5U_{54} , m^1A_{58} , Ψ_{55} . All these modified nucleosides are involved in base-pairing or tertiary interactions that help shape and lock the tRNA elbow structure. Two crystal structures of tRNA-core-modifying enzymes bound to a full-length tRNA are available: the tRNA-guanine transglycosylase (TGT) that catalyses the formation of archaeosine at G_{15} (D-loop) bound to tRNA^{Val} (45), and the human tRNA m^1A_{58} MTase (Trm6-Trm61) in complex with its tRNA^{Lys}₃ substrate (12). In the structure of TGT/tRNA^{Val}, the bound tRNA shows an alternative conformation named the λ -form, which is drastically different from the canonical L-shape. In the λ -form, all D-arm secondary base-pairs and canonical tertiary interactions are disrupted and the helical structure is reorganized such that the D-arm nucleotides 9 to 22 are unpaired and protrude from the tRNA. A D-variable helix formed by the base-pairing between residues 23–48, 24–47 and 25–46 is stacked on the anticodon stem to form a double-helical structure. In the structure of Trm6-Trm61/tRNA^{Lys}₃, the D- and the T-arms are detached from each other to expose the A_{58} N1-atom for methylation, but the L-shape of the tRNA is otherwise maintained. The binding is stabilized by the formation of numerous hydrogen bonds with the C56 nucleobase and the sugar-phosphate backbone. For the s^4U_8 , m^5U_{54} and Ψ_{55} formations, crystal structures of enzymes responsible for each of these modifications in complex with a fragment of tRNA are available. These structures show a flip of the nucleoside to be modified into the catalytic pocket of the enzyme that is stabilized by numerous hydrogen bonds between the protein and the sugar-phosphate moieties of the nucleotides surrounding the target base (46–48).

In all these complexes, the nucleoside to be modified is buried in full-length tRNA structure and is base-paired. To get access to the nucleoside, the enzyme has to both disrupt the base-pairing and partially unfold the tRNA. For nucleosides located nearer the surface of tRNAs like A_9 , no such conformational changes would be expected. According to a docking model of the *Sulfolobus acidocaldarius* m^1A_9 MTase and tRNA, the canonical L-shape of the tRNA is nearly perfectly retained, with only a small flip of the D-stem (49). Like other m^1A modified bases, A_{22} of B_s TrmK substrates takes part in non-Watson–Crick base-pairing by binding to either a G or an A at position 13. The NMR data obtained on the B_s tRNA^{Ser} show that the variable hairpin is still formed upon binding of B_s TrmK indicating that it is not folded in the λ -form. In the tRNA L-shape structure used for the docking, the A_{22} N1-atom, when base-paired with a G or an A, is not buried deep, and is accessible for modification without need for much structural rearrangement of the tRNA molecule. Therefore, the target atom seems to be rendered accessible by the specific R_{13} – A_{22} base-pairing that is very stable (50), and allows the placement of the N1-atom of the adenine 22 accessible in the major groove of the tRNA. The major demonstration of the crucial role of the R_{13} – A_{22} pairing is supported by the

fact that replacing pyrimidine 13 of a non-substrate tRNA by a purine renders the tRNA substrate of B_s TrmK. In the non-substrate tRNA^{Gly}, A_{22} is involved in a Watson–Crick base pair with U_{13} , which hides the N1-atom of A_{22} inside the AU base-pair and thus inside the tRNA structure. According to the model proposed in the present work, B_s TrmK can bind the tRNA with no to very little deformation of both partners, the G_{13} – A_{22} base pair allowing the placement of the N1-atom of A_{22} in close contact with the methyl group of SAM in the active site of the enzyme. However, we cannot definitely rule out that an induced-fit process occurs after binding to stabilize the complex. Disappearance of peaks in the NMR spectra for residues at the top of the anticodon stem and at the junction between the acceptor and the T-stem are observed upon binding of B_s TrmK. These base-pairs are located apart from the G_{13} – A_{22} pair. For instance, a bending at the interface between the acceptor and the T-stems could explain that the imino groups of these base-pairs become more exchangeable with the solvent and disappear from the NMR spectra. This bending would likely allow the acceptor stem to interact with TrmK. Further experiments with a co-crystal between TrmK and a tRNA would be needed to test this hypothesis.

Future investigations are needed to establish the real occurrence of m^1A_{22} in tRNAs

The presence of m^1A_{22} is not frequently reported in tRNA database and has to date only been identified in some bacteria. In our opinion, this may largely be due to the lack of systematic modification mapping in tRNAs. *In vitro*, B_s TrmK modifies any provided tRNA bearing a purine at position 13, indicating that many tRNAs carrying an A at position 22 could be modified in *B. subtilis* *in vivo*. The seemingly low occurrence of this modification therefore reflects either the low availability of *B. subtilis* tRNA sequences for which the complete identification of modifications has not been performed or that B_s TrmK specificity is different *in vivo* like previously observed for the yeast m^1G_9 forming enzyme Trm10 (51). Systematic mapping of *B. subtilis* tRNA modifications could in the future give an answer to this question and confirm whether the specificity observed in this study *in vitro* is retained *in vivo*.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank the TGIR-RMN-THC FR3050 at Gif-sur-Yvette for NMR measurement time and Nelly Morellet for helpful discussion and technical advice. The X-ray data were collected on the ID14-eh1 beamline at the European Synchrotron Radiation Facility (ESRF, Grenoble, France). We are grateful to the Local Contact at the ESRF for assisting during data collection. Abdalla Al Refaii was supported by a Khaled Al-As'ad fellowship delivered by the Université libre de Bruxelles. The Fonds D. et A. Van Buuren and the Fonds J. Brachet (ULB) are gratefully acknowledged for financial support. We also thank Kevin Englebert, Jessica

Berlier and Marie Morel for taking part to the early stages of this work.

FUNDING

CNRS, Paris Descartes University, Labex Dynamo Funding for open access charge: CNRS.

Conflict of interest statement. None declared.

REFERENCES

- Boccalletto, P., Machnicka, M.A., Purta, E., Piatkowski, P., Baginski, B., Wirecki, T.K., de Crecy-Lagard, V., Ross, R., Limbach, P.A., Kotter, A. *et al.* (2018) MODOMICS: a database of RNA modification pathways. 2017 update. *Nucleic Acids Res.*, **46**, D303–D307.
- Oerum, S., Degut, C., Barraud, P. and Tisne, C. (2017) m1A Post-Transcriptional modification in tRNAs. *Biomolecules*, **7**, E20.
- Grosjean, H., Constantinesco, F., Foiret, D. and Benachenhou, N. (1995) A novel enzymatic pathway leading to 1-methylinosine modification in *Haloflex volcanii* tRNA. *Nucleic Acids Res.*, **23**, 4312–4319.
- Roovers, M., Wouters, J., Bujnicki, J.M., Tricot, C., Stalon, V., Grosjean, H. and Droogmans, L. (2004) A primordial RNA modification enzyme: the case of tRNA (m1A) methyltransferase. *Nucleic Acids Res.*, **32**, 465–476.
- Kempnaers, M., Roovers, M., Oudjama, Y., Tkaczuk, K.L., Bujnicki, J.M. and Droogmans, L. (2010) New archaeal methyltransferases forming 1-methyladenosine or 1-methyladenosine and 1-methylguanosine at position 9 of tRNA. *Nucleic Acids Res.*, **38**, 6533–6543.
- Vilardo, E., Nachbagauer, C., Buzet, A., Taschner, A., Holzmann, J. and Rossmanith, W. (2012) A subcomplex of human mitochondrial RNase P is a bifunctional methyltransferase–extensive moonlighting in mitochondrial tRNA biogenesis. *Nucleic Acids Res.*, **40**, 11583–11593.
- Jackman, J.E., Montange, R.K., Malik, H.S. and Phizicky, E.M. (2003) Identification of the yeast gene encoding the tRNA m1G methyltransferase responsible for modification at position 9. *RNA*, **9**, 574–585.
- Singh, R.K., Feller, A., Roovers, M., Van Elder, D., Wauters, L., Droogmans, L. and Versees, W. (2018) Structural and biochemical analysis of the dual-specificity Trm10 enzyme from *Thermococcus kodakaraensis* prompts reconsideration of its catalytic mechanism. *RNA*, **24**, 1080–1092.
- Helm, M., Giege, R. and Florentz, C. (1999) A Watson–Crick base-pair-disrupting methyl group (m1A9) is sufficient for cloverleaf folding of human mitochondrial tRNA^{Lys}. *Biochemistry*, **38**, 13338–13346.
- Sprinzl, M. and Vassilenko, K.S. (2005) Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res.*, **33**, D139–D140.
- Wang, M., Zhu, Y., Wang, C., Fan, X., Jiang, X., Ebrahimi, M., Qiao, Z., Niu, L., Teng, M. and Li, X. (2016) Crystal structure of the two-subunit tRNA m(1)A58 methyltransferase TRM6-TRM61 from *Saccharomyces cerevisiae*. *Sci. Rep.*, **6**, 32562.
- Finer-Moore, J., Czudnochowski, N., O’Connell, J.D. 3rd, Wang, A.L. and Stroud, R.M. (2015) Crystal structure of the human tRNA m(1)A58 Methyltransferase-tRNA(3)(Lys) complex: refolding of substrate tRNA allows access to the methylation target. *J. Mol. Biol.*, **427**, 3862–3876.
- Droogmans, L., Roovers, M., Bujnicki, J.M., Tricot, C., Hartsch, T., Stalon, V. and Grosjean, H. (2003) Cloning and characterization of tRNA (m1A58) methyltransferase (TrmI) from *Thermus thermophilus* HB27, a protein required for cell growth at extreme temperatures. *Nucleic Acids Res.*, **31**, 2148–2156.
- Barraud, P., Golinelli-Pimpaneau, B., Atmanene, C., Sanglier, S., Van Dorselaer, A., Droogmans, L., Dardel, F. and Tisne, C. (2008) Crystal structure of *Thermus thermophilus* tRNA m1A58 methyltransferase and biophysical characterization of its interaction with tRNA. *J. Mol. Biol.*, **377**, 535–550.
- Guelorget, A., Barraud, P., Tisne, C. and Golinelli-Pimpaneau, B. (2011) Structural comparison of tRNA m(1)A58 methyltransferases revealed different molecular strategies to maintain their oligomeric architecture under extreme conditions. *BMC Struct. Biol.*, **11**, 48.
- Takuma, H., Ushio, N., Minoji, M., Kazayama, A., Shigi, N., Hirata, A., Tomikawa, C., Ochi, A. and Hori, H. (2015) Substrate tRNA recognition mechanism of eubacterial tRNA (m1A58) methyltransferase (TrmI). *J. Biol. Chem.*, **290**, 5912–5925.
- Degut, C., Ponchon, L., Folly-Klan, M., Barraud, P. and Tisne, C. (2016) The m1A(58) modification in eubacterial tRNA: An overview of tRNA recognition and mechanism of catalysis by TrmI. *Biophys. Chem.*, **210**, 27–34.
- Varshney, U., Ramesh, V., Madabushi, A., Gaur, R., Subramanya, H.S. and RajBhandary, U.L. (2004) Mycobacterium tuberculosis Rv2118c codes for a single-component homotetrameric m1A58 tRNA methyltransferase. *Nucleic Acids Res.*, **32**, 1018–1027.
- Guelorget, A., Roovers, M., Guérineau, V., Barbey, C., Li, X. and Golinelli-Pimpaneau, B. (2010) Insights into the hyperthermostability and unusual region-specificity of archaeal *Pyrococcus abyssi* tRNA m1A57/58 methyltransferase. *Nucleic Acids Res.*, **38**, 6206–6218.
- Raettig, R., Kersten, H., Weissenbach, J. and Dirheimer, G. (1977) Methylation of an adenosine in the D-loop of specific transfer RNAs from yeast by a procaryotic tRNA (adenine-1) methyltransferase. *Nucleic Acids Res.*, **4**, 1769–1782.
- Kersten, H., Raettig, R., Weissenbach, J. and Dirheimer, G. (1978) Recognition of individual procaryotic and eucaryotic transfer-ribonucleic acids by *B. subtilis* adenine-1-methyltransferase specific for the dihydrouridine loop. *Nucleic Acids Res.*, **5**, 3033–3042.
- Roovers, M., Kaminska, K.H., Tkaczuk, K.L., Gigot, D., Droogmans, L. and Bujnicki, J.M. (2008) The YqfN protein of *Bacillus subtilis* is the tRNA: m1A22 methyltransferase (TrmK). *Nucleic Acids Res.*, **36**, 3252–3262.
- Thanassi, J.A., Hartman-Neumann, S.L., Dougherty, T.J., Dougherty, B.A. and Pucci, M.J. (2002) Identification of 113 conserved essential genes using a high-throughput gene disruption system in *Streptococcus pneumoniae*. *Nucleic Acids Res.*, **30**, 3152–3162.
- Hutchison, C.A. 3rd, Chuang, R.Y., Noskov, V.N., Assad-Garcia, N., Deerinck, T.J., Ellisman, M.H., Gill, J., Kannan, K., Karas, B.J., Ma, L. *et al.* (2016) Design and synthesis of a minimal bacterial genome. *Science*, **351**, aad6253.
- Ta, H.M. and Kim, K.K. (2010) Crystal structure of *Streptococcus pneumoniae* Sp1610, a putative tRNA methyltransferase, in complex with S-adenosyl-L-methionine. *Protein Sci.*, **19**, 617–624.
- Degut, C., Monod, A., Brachet, F., Crepin, T. and Tisne, C. (2016) In Vitro/In vivo production of tRNA for X-Ray studies. *Methods Mol. Biol.*, **1320**, 37–57.
- Dardel, F. (1994) MC-Fit: using Monte-Carlo methods to get accurate confidence limits on enzyme parameters. *Comput. Appl. Biosci.*, **10**, 273–275.
- Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W. *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 213–221.
- Emsley, P., Lohkamp, B., Scott, W.G. and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 486–501.
- Bakan, A., Meireles, L.M. and Bahar, I. (2011) ProDy: protein dynamics inferred from theory and experiments. *Bioinformatics*, **27**, 1575–1577.
- Humphrey, W., Dalke, A. and Schulten, K. (1996) VMD: visual molecular dynamics. *J. Mol. Graph.*, **14**, 33–38.
- Degut, C., Barraud, P., Larue, V. and Tisne, C. (2016) Backbone resonance assignments of the m1A22 tRNA methyltransferase TrmK from *Bacillus subtilis*. *Biomol. NMR Assign.*, **10**, 253–257.
- Pervushin, K.V., Wider, G. and Wuthrich, K. (1998) Single Transition-to-single transition polarization transfer (ST2-PT) in [15N,1H]-TROSY. *J. Biomol. NMR*, **12**, 345–348.
- Orekhov, V.Y. and Jaravine, V.A. (2011) Analysis of non-uniformly sampled spectra with multi-dimensional decomposition. *Prog. Nucl. Magn. Reson. Spectrosc.*, **59**, 271–292.
- Das, R., Karanicolas, J. and Baker, D. (2010) Atomic accuracy in predicting and designing noncanonical RNA structure. *Nat. Methods*, **7**, 291–294.

36. Dominguez, C., Boelens, R. and Bonvin, A.M. (2003) HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.*, **125**, 1731–1737.
37. de Vries, S.J., van Dijk, A.D., Krzeminski, M., van Dijk, M., Thureau, A., Hsu, V., Wassenaar, T. and Bonvin, A.M. (2007) HADDOCK versus HADDOCK: new features and performance of HADDOCK2.0 on the CAPRI targets. *Proteins*, **69**, 726–733.
38. Juhling, F., Morl, M., Hartmann, R.K., Sprinzl, M., Stadler, P.F. and Putz, J. (2009) tRNAdb 2009: compilation of tRNA sequences and tRNA genes. *Nucleic Acids Res.*, **37**, D159–D162.
39. Biou, V., Yaremchuk, A., Tukalo, M. and Cusack, S. (1994) The 2.9 Å crystal structure of *T. thermophilus* seryl-tRNA synthetase complexed with tRNA(Ser). *Science*, **263**, 1404–1410.
40. Zhang, J. and Ferre-D'Amare, A.R. (2016) The tRNA elbow in structure, recognition and evolution. *Life (Basel)*, **6**, E3.
41. Giege, R., Juhling, F., Putz, J., Stadler, P., Sauter, C. and Florentz, C. (2012) Structure of transfer RNAs: similarity and variability. *Wiley Interdiscip. Rev. RNA*, **3**, 37–61.
42. Holm, L. and Sander, C. (1998) Touring protein fold space with Dali/FSSP. *Nucleic Acids Res.*, **26**, 316–319.
43. Reichle, V.F., Kaiser, S., Heiss, M., Hagelskamp, F., Borland, K. and Kellner, S. (2019) Surpassing limits of static RNA modification analysis with dynamic NAIL-MS. *Methods*, **156**, 91–101.
44. Krishnamohan, A., Dodbele, S. and Jackman, J.E. (2019) Insights into catalytic and tRNA recognition mechanism of the Dual-Specific tRNA methyltransferase from *Thermococcus kodakarensis*. *Genes (Basel)*, **10**, E100.
45. Ishitani, R., Nureki, O., Nameki, N., Okada, N., Nishimura, S. and Yokoyama, S. (2003) Alternative tertiary structure of tRNA for recognition by a posttranscriptional modification enzyme. *Cell*, **113**, 383–394.
46. Alian, A., Lee, T.T., Griner, S.L., Stroud, R.M. and Finer-Moore, J. (2008) Structure of a TrmA-RNA complex: a consensus RNA fold contributes to substrate selectivity and catalysis in m5U methyltransferases. *Proc. Natl. Acad. Sci. U.S.A.*, **105**, 6876–6881.
47. Hoang, C. and Ferre-D'Amare, A.R. (2001) Cocystal structure of a tRNA Psi55 pseudouridine synthase: nucleotide flipping by an RNA-modifying enzyme. *Cell*, **107**, 929–939.
48. Pan, H., Agarwalla, S., Moustakas, D.T., Finer-Moore, J. and Stroud, R.M. (2003) Structure of tRNA pseudouridine synthase TruB and its RNA complex: RNA recognition through a combination of rigid docking and induced fit. *Proc. Natl. Acad. Sci. U.S.A.*, **100**, 12648–12653.
49. Van Laer, B., Roovers, M., Wauters, L., Kasprzak, J.M., Dyzma, M., Deyaert, E., Kumar Singh, R., Feller, A., Bujnicki, J.M., Droogmans, L. et al. (2016) Structural and functional insights into tRNA binding and adenosine N1-methylation by an archaeal Trm10 homologue. *Nucleic Acids Res.*, **44**, 940–953.
50. Mladek, A., Sharma, P., Mitra, A., Bhattacharyya, D., Sponer, J. and Sponer, J.E. (2009) Trans Hoogsteen/sugar edge base pairing in RNA. Structures, energies, and stabilities from quantum chemical calculations. *J. Phys. Chem. B*, **113**, 1743–1755.
51. Swinehart, W.E., Henderson, J.C. and Jackman, J.E. (2013) Unexpected expansion of tRNA substrate recognition by the yeast m1G9 methyltransferase Trm10. *RNA*, **19**, 1137–1146.
52. Baker, N.A., Sept, D., Joseph, S., Holst, M.J. and McCammon, J.A. (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U.S.A.*, **98**, 10037–10041.
53. Dolinsky, T.J., Nielsen, J.E., McCammon, J.A. and Baker, N.A. (2004) PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.*, **32**, W665–W667.