Structural characterization of *B. subtilis* m^1^A^22^ tRNA methyltransferase TrmK: insights into tRNA recognition

Clément Dégut^1,†, Martine Roovers^2,†, Pierre Barraud^1, Franck Brachet^1, André Feller^4, Valéry Larue^1, Abdalla Al Refaii^4, Joël Caillet^3, Louis Droogmans^4,* and Carine Tisné^1,3,*

^1^Laboratoire de Cristallographie et RMN biologiques, CNRS, Université Paris Descartes, Sorbonne Paris Cité, 4 avenue de l’Observatoire, 75006 Paris, France, ^2^Institut de Recherches Labiris, B-1070 Bruxelles, Belgium, ^3^Laboratoire 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 ^4^Laboratoire 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 (m^1^A) 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 (*Bs*TrmK) responsible for the formation of m^1^A^22^ in tRNA. Here, we show that *Bs*TrmK displays a broad substrate specificity, and methylates seven out of eight tRNA isoacceptor families of *B. subtilis* bearing an A^22^. In addition to a non-Watson–Crick base-pair between the target A^22^ and a purine at position 13, the formation of m^1^A^22^ by *Bs*TrmK requires a full-length tRNA with intact tRNA elbow and anticodon stem. We solved the crystal structure of *Bs*TrmK 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 *Bst*RNASer to *Bs*TrmK in complex with its methyl-donor cofactor S-adenosyl-L-methionine (SAM). In this model, validated by methyltransferase activity assays on *Bs*TrmK mutants, both domains of *Bs*TrmK participate in tRNA binding. *Bs*TrmK recognises tRNA with very few structural changes in both partner, the non-Watson–Crick R^13^-A^22^ base-pair positioning the A^22^ 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 (m^1^A) found in tRNA. The formation of m^1^A 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 m^1^A^9^, m^1^A^14^ and m^1^A^58^ 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 m^1^A formation at the different positions in tRNAs. The m^1^A at position 9 (m^1^A^9^) 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 (m^1^G^9^) (7) and the Trm10 enzymes from human mitochondria and from several archaea display dual specificity, forming both m^1^A^9^ and m^1^G^9^ (5,6,8). In human mitochondrial tRNA^Lys^, m^1^A^9^ prevents the formation of an alternative structure of this tRNA by hampering a Watson–Crick base-pairing between A^9^ and U^64^ (9). The presence of m^1^A^14^ 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 m^1^A^58^ mod-
The presence of m1A22 in tRNA is rather scarce and its formation has been much less studied compared to m1A9 and m1A58. For organisms where tRNA sequences are available (1), m1A22 is found in tRNAs of *Bacillus subtilis* (tRNA^Ser^, tRNA^Gly^), both with large variable regions), of *Geobacillus stearothermophilus* (tRNA^Leu^ and tRNA^Lys^), of *Mycoplasma capricolum* (tRNA^Glu^, tRNA^Gly^, tRNA^His^, tRNA^Leu^, tRNA^Ser^ and tRNA^Arg^) and of *Mycoplasma mycoides* (tRNA^Ser^). Early studies showed that a SAM-dependent m1A22 MTase activity was present in B. subtilis extracts (20,21). More recently, the ygfN gene of *B. subtilis* (now trmK) was shown to encode the enzyme TrmK responsible for m1A22 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* m1A22 MTase TrmK (*B*\textsubscript{t}TrmK) with special focus on the enzyme-tRNA recognition. We investigated the properties that establish tRNAs as substrates of *B*\textsubscript{t}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*\textsubscript{t}TrmK recognition is involved in base-pairing with A22, the target site of methylation by *B*\textsubscript{t}TrmK. We also solved the crystal structure of *B*\textsubscript{t}TrmK and used NMR spectroscopy to gain insight into the recognition mode of *B. subtilis* tRNA^Ser^ (*B*\textsubscript{t}tRNA^Ser^) by *B*\textsubscript{t}TrmK. Based on the NMR data, we constructed a docking model of the *B*\textsubscript{t}TrmK/3\textsubscript{t}SAM/3\textsubscript{t}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*\textsubscript{t}TrmK for NMR and X-ray studies**

Recombinant *B*\textsubscript{t}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 His\textsubscript{6}-tagged recombinant protein. For structural studies, a variant of *B*\textsubscript{t}TrmK in which the two cysteine residues were replaced by serine ones (C35S and C152S) was used. Mutagenesis was performed using the QuickChange 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 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 His\textsubscript{6}-tag was removed from *B*\textsubscript{t}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*\textsubscript{t}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*\textsubscript{t}TrmK variant was purified as previously described (26). For the assignment procedure of NMR signals, samples of 3\textsuperscript{H}/13\textsuperscript{C}/15\textsuperscript{N}-labeled *B*\textsubscript{t}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\textsuperscript{H}/13\textsuperscript{C}/15\textsuperscript{N}-labeled *B*\textsubscript{t}TrmK concentrated around 0.2 mM in a 50 mM Tris–Cl buffer pH 8.0 containing 100 mM NaCl, 10 mM MgCl\textsubscript{2} and 2% glycerol.

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

*B*\textsubscript{t}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*\textsubscript{t}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\textsubscript{2} was then added. Fragments of *B*\textsubscript{t}tRNA^Ser^ were purchased from Dharmacon (GE Healthcare).
Analysis of \textit{B. subtilis} TrmK oligomerisation and subsequent MTase activity

Pure \textit{B. subtilis} 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 \textit{E. coli} tRNA using methyl-\textsuperscript{14}C SAM as methyl donor. The reaction mixture was 200 \mu l of total \textit{E. coli} tRNAs, 400 nCi of methyl-\textsuperscript{14}C SAM (50 mCi/mmol), the fraction containing \textit{B. subtilis} TrmK and H\textsubscript{2}O. Incubation was for 1 h at 37\textdegree 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 \textit{B. subtilis} tRNA genes in pUC18 plasmid

The cloning of the \textit{B. subtilis} genes of tRNA Cys, Gln, Gln, Tyr and Gly were performed as for the cloning of \textit{B. subtilis} tRNA\textsubscript{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 \textit{B. subtilis} TrmK using \textsuperscript{32}P-radiolabelled \textit{B. subtilis} tRNAs

The measurements of MTase activity were conducted on the wild-type (WT) \textit{B. subtilis} TrmK in presence of 10 mM DTT. Radioactive (\textsuperscript{32}P) 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). [\textsuperscript{\alpha-\textsuperscript{32}P}]-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\textdegree C in a 300 \mu l reaction mixture containing 0.2 mM SAM, 50 mM MOPS pH 6.5, 1 mM MgCl\textsubscript{2}, 10 mM DTT and various amounts of \textit{B. subtilis} 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 \textit{B. subtilis} TrmK

[\textsuperscript{Methyl-\textsuperscript{3}H}]-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 MgCl\textsubscript{2} at 37\textdegree C. Aliquots (300 \mu l) 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 \textit{K}_M and \textit{V}_\text{max} of \textit{B. subtilis} TrmK for tRNA\textsubscript{Ser}, the reaction mixtures contained \textit{B. subtilis} TrmK (100 nM), tRNA\textsubscript{Ser} (10–2000 nM) and \textsuperscript{3}H-AdoMet (90 \mu M). Initial rates (\textit{v}_i) for each substrate concentration (tRNA\textsubscript{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).

\[ \textit{v}_i = \frac{\textit{V}_\text{max}}{2 \textit{E}_0} (\textit{K}_M + \textit{S}_i + \textit{E}_0 - \sqrt{(\textit{K}_M + \textit{S}_i + \textit{E}_0)^2 - 4 \textit{S}_i \textit{E}_0}) \]  

Preparation of \textit{B. subtilis} TrmK mutants and MTase assays

All the TrmK mutants were generated using the QuikChange mutagenesis kit (Agilent) and purified by affinity chromatography on Ni\textsuperscript{2+} loaded Chelating Sepharose as described above. The MTase assays were performed in 50 mM MOPS-Na pH 6.5, 1 mM MgCl\textsubscript{2}, 10 mM DTT, 50 \mu g unfraccionated tRNA from the \textit{B. subtilis} (delta trmK) strain (22), 25 nCi [\textsuperscript{methyl-\textsuperscript{14}C}]-SAM (58 mCi/mmol, Perkin Elmer) and 0.2 \mu g enzyme in a total volume of 300 \mu l. Incubation was for 30 min at 37\textdegree 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 crystallography of \textit{B. subtilis} TrmK

Crystallisation of \textit{B. subtilis} TrmK was performed at 4\textdegree 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 \mu l were prepared by a Cybi-Disk robot mixing equal volume of protein and reservoir (100 \mu l). Crystals of \textit{B. subtilis} 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 \textit{sp}1610 (\textit{Streptococcus pneumoniae} ortholog...
of BsTrmK, 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 BsTrmK structure has been deposited to the Protein Data Bank under the accession code 6Q56.

Normal mode analysis of BsTrmK 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 BsTrmK 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 BsTrmK dialysis, were made into the cell containing BsTrmK (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 BsTrmK and Bs tRNASer and BsTrmK previously bound to SAH were obtained using 2H/13C/15N-labeled BsTrmK (0.14 mM) mixed with 4 equivalents of SAH and two equivalents of Bs tRNASer 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: Δ(H,N) = \[ \sqrt{[(\Delta 15N W_N)^2 + (\Delta 1H W_H)^2]} \], where Δ = δ (complex - δ free (difference of chemical shift between the complex and the free states)) and W_H = 1 and W_N = 1/5.

Molecular docking of BsTrmK/SAM/tRNASer
Bs tRNASer model was built using the T. thermophilus tRNASer crystal structure (PDB: ISER) as template except for the variable region that was not defined in this structure. To model the variable region of Bs tRNASer, the Human tRNASer crystal structure (PDB: 3A3A) was used. It was manually mutated to the Bs tRNASer sequence using Coot (29). Model geometry was then idealised with 300 iterations of RNAidealise from the Rosetta package, and subsequently minimised using the Rosetta force field through rna_minimise (35). Docking of B. subtilis tRNASer on BsTrmK/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 Bs tRNASer binding. Residues from BsTrmK with NMR chemical shift variations of their amide group superior to 0.03 ppm upon tRNASer 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 BsTrmK for methylation, was defined as active residue for the Bs tRNASer 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 tRNASer 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 BsTrmK 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 Bs tRNASer phosphate backbone between two models relative to the BsTrmK backbone. In other words, we compare the positioning of the Bs tRNASer on BsTrmK in the different models by superimposing the models along the BsTrmK 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 BsTrmK/SAM/Bs tRNASer complex are available as Supplementary data, TrmK_SAM_tRNASer.pdb.

RESULTS
BsTrmK 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 G10) could undergo m1A22 for-
m1A22 formation catalysed by B₈-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 tRNASer 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 B₈-TrmK (Figure 1). tRNA_Ser showed a two-fold decreased m1A22 formation (Figure 1G) compared to the other substrate tRNAs, and tRNA_Gly was not methylated at all by B₈-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.

B₈-TrmK requires a purine at position 13 of the tRNA to catalyse m1A22 formation

A comparison of the sequences of substrate (tRNA_Ser, tRNA_Gly, tRNA_Cys, tRNA_Glu, tRNA_Leu, tRNA_His and tRNA_Gln) and non-substrate tRNAs (tRNA_Gly) highlights two nucleotides, at positions 13 and 26, as potentially implicated in substrate discrimination by B₈-TrmK. Both positions are occupied by a purine (G or A) in all B₈-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 B₈-TrmK substrates. Both single-mutated tRNA_Gly (U13G and C26A) turned the previous non-substrate into a substrate for B₈-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 B₈-TrmK.

The purine at position 13, important for m1A22 formation catalysed by B₈-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 B₈-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 B₈-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 B₈-TrmK.

B₈-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 Ψ56 and G19 with C56, and a further stacking of the purine at position 57 between base-pairs G18-Ψ56 and G19-C56. To determine if B₈-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, U58G/C56G or U54C/U58G/C56G (Figure 2C). The triple-mutant of tRNA_His (U54C/U58G/C56G) showed very low substrate potential compared to WT tRNA_His, suggesting that B₈-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 B₈-tRNA_Ser. None of these transcripts were substrates for B₈-TrmK (data not shown).

Next, the importance of the anticodon stem-loop for tRNA recognition by B₈-TrmK was evaluated. In order to determine whether B₈-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 (U13G), an insertion of a G near position 36 (+G36A), and deletion of G37 (ΔG37) were produced. In the stem, a variant with three additional GC base pairs (C–G, G–C and C–G between A29–U31 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 B₈-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 B₈-TrmK substrates.

Altogether, these data show that B₈-TrmK displays a broad tRNA substrate specificity. B₈-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.

B₈-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 B₈-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, B₈-TrmK eluted both as monomers and higher oligomeric species. The activity measurements of each eluted fraction showed that B₈-TrmK is functional as a monomer. The higher oligomeric states of B₈-TrmK are formed via disulphide bonds involving the two cysteines in
Figure 1. \( B. \text{TrmK} \) methylates a variety of \( B. \text{subtilis} \) tRNAs. Autoradiograms of chromatograms of P1 hydrolysates of \([\text{32P}]\text{ATP}\)-labelled T7-transcripts of \( B. \text{subtilis} \) tRNAs incubated during 15 minutes with SAM and increasing amounts of purified \( B. \text{TrmK} \) (from 0 to 600 nM). The sequences of the tRNAs tested as substrates are drawn above and below the autoradiograms. (A) tRNA\(^{\text{Ser}}\), (B) tRNA\(^{\text{His}}\), (C) tRNA\(^{\text{Tyr}}\), (D) tRNA\(^{\text{Cys}}\), (E) tRNA\(^{\text{Gln}}\), (F) tRNA\(^{\text{Glu}}\), (G) tRNA\(^{\text{Leu}}\), (H) tRNA\(^{\text{Gly}}\).

\( B. \text{TrmK} \) 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 \( B. \text{TrmK} \) functions as a monomer in solution.

Crystallisation and structure determination of \( B. \text{TrmK} \)

No crystals of \( B. \text{TrmK} \) 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 \( B. \text{TrmK} \) (Supplementary Figure S4). Therefore, all following structural studies were performed with this construct.

Unlike WT \( B. \text{TrmK} \), 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 \( R_{\text{free}} \) 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). \( B. \text{TrmK} \) 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 \( C_\alpha \) atoms).

The superimposition of the \( B. \text{TrmK} \) structure with its orthologs (PDB: 3KU1, 3GNL, 3KR9) gave an RMSD of 1.5 Å over around 220 \( C_\alpha \) atoms, consistent with the high sequence identity of these proteins.

\( B. \text{TrmK} \) consists of an N-terminal Class I MTase domain linked to a C-terminal coiled-coil domain

The \( B. \text{TrmK} \) 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 \( \beta \)-sheet flanked by six \( \alpha \)-helices. In the \( \beta \)-sheet, the five first \( \beta \)-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 \( B. \text{TrmK} \) identified, apart
Figure 2. \( \text{bsTrmK} \) MTase assays with \( B. \text{subtilis} \) tRNAs mutated at position 13, and with tRNA\(^{\text{His}} \) altered in the T-loop or in the anticodon stem-loop. \( [\alpha^{32}P] \text{ATP}\)-labelled T7-transcripts of mutated tRNA\(^{\text{Ser}} \) (A) or tRNA\(^{\text{His}} \) (B–D) of \( B. \text{subtilis} \) were incubated during 15 min with SAM and increasing amounts of purified \( \text{bsTrmK} \). P\(_1\) 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 \( \text{bsTrmK} \) concentration in the reaction mixture (A) for mutants of tRNA\(^{\text{Ser}} \) at position 13, (B) for mutants of tRNA\(^{\text{His}} \) at position 13, (C) for variants of tRNA\(^{\text{His}} \) in the T-loop, (D) for variants of tRNA\(^{\text{His}} \) in the anticodon stem-loop. (E) Summary on the tRNA\(^{\text{Ser}} \) structure of the determinants required for \( \text{bsTrmK} \) MTase activity: the length of the variable region (in green) is not important, the base-pair A\(_{22}\)–N\(_{13}\) (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 Ca), the m\(^2\)G\(_{63}\) tRNA MTase Trm14 (PDB: 3TM4, Z-score 17.4, RMSD 2.4 Å over 152 Ca), the m\(^1\)A\(_{58}\) tRNA MTase TrmI (PDB: 2PWY, Z-score 15.8, RMSD 2.4 Å over 151 Ca) and the m\(^1\)G\(_{37}\) tRNA MTase Trm5 (PDB: 2ZZM, Z-score 14.5, RMSD 2.9 Å over 147 Ca). These proteins are all Class-I type RFMs, confirming that \( \text{bsTrmK} \) 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 \( \text{bsTrmK} \) contains four \( \alpha \)-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 Ca) and the ribosomal protein L19 (PDB: 1S72, Z-score 6.1, RMSD 2.0 Å over 53 Ca). In these proteins, 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 Ca) and the ribosomal protein L19 (PDB: 1S72, Z-score 6.1, RMSD 2.0 Å over 53 Ca). In these proteins, the encompassed coiled-coil motif binds an RNA, suggesting that in \( \text{bsTrmK} \) the C-terminal domain could also play such a role. The N- and C-terminal domain of \( \text{bsTrmK} \) are held together by numerous contacts, and a normal mode analysis of the \( \text{bsTrmK} \) 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 \( \text{bsTrmK} \) and the cofactor product SAH was studied by ITC. The interaction is enthalpy-driven (\( \Delta H \approx -17 \text{ kcal/mol} \)) with a single binding site and a dissociation constant (\( K_D \)) of 1.7 μM (Supplementary Figure S7). Although SAM or SAH were present in crystallisation trials, crystals were of apo \( \text{bsTrmK} \), i.e. no ligand bound. The electron density is well-defined across the \( \text{bsTrmK} \) structure and particularly for residues constituting the SAM-binding site, even in the absence of the coiled-coil motif as flexible (Figure 3D).

The electron density is well-defined across the \( \text{bsTrmK} \) structure and particularly for residues constituting the SAM-binding site, even in the absence of the coiled-coil motif as flexible (Figure 3D).

The C-terminal domain of \( \text{bsTrmK} \) is bound at the centre of the catalytic domain in a pocket with residues harbouring negative electrostatic potentials con-
Figure 3. Crystal structure of BsTrmK. SAM was modelled in the catalytic center of BsTrmK by superposing the BsTrmK structure with that of its ortholog from S. pneumoniae (PDB code 3KU1). It is represented as sticks. (A) Ribbon representation of BsTrmK, α-helices are coloured in red whereas β-strands are in yellow. (B) Location on the structure of BsTrmK 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 $-5kT/e$ (red) and $+5kT/e$ (blue). The figure was prepared with the APBS PyMOL plug-in and pdb2pqr webservice. (D) Normal mode analysis of BsTrmK 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 BsTrmK residue number.

Some post-transcriptional modifications were reported to stabilise and help fold tRNA structures. Therefore, we investigated if modifications in BsttRNA<sup>Ser</sup> could make this a better substrate for BsTrmK, compared to the unmodified BsttRNA. No homologue of TrmK is present in E. coli, enabling us to produce the tRNA<sup>Ser</sup> in this host, bearing all modifications formed by the E. coli enzymes i.e. Gm<sub>18</sub>, D<sub>19</sub>, T<sub>54</sub> and $\Psi_{55}$ (Figure 4A), but lacking $m^1$A<sub>22</sub>. The B. subtilis tRNA<sup>Ser</sup> was overexpressed and purified in E. coli as previously described (26). The catalytic efficiency of BsTrmK towards this tRNA<sup>Ser</sup> was then compared to that measured with an in vitro transcribed tRNA<sup>Ser</sup> lacking modifications...
Figure 4. Interaction of BsttRNA<sub>Ser</sub> with B<sub>s</sub>TrmK/SAH deciphered by NMR. (A) The presence of post-transcriptional modifications in B<sub>s</sub>TrmK substrate alters the MTase activity of B<sub>s</sub>TrmK. Sequence of B. subtilis tRNA<sub>Ser</sub> produced in vivo in E. coli showing post-transcriptional modifications incorporated by E. coli modifying enzymes. (B) Kinetic and enzymatic parameters of B<sub>s</sub>TrmK measured for BsttRNA<sub>Ser</sub> bearing (tRNA<sub>Ser</sub> in vivo) or not bearing (tRNA<sub>Ser</sub> in vitro) post-transcriptional modifications. Confidence intervals at 90% are indicated in parenthesis. An unmodified BsttRNA<sub>Ser</sub> was chosen to study the interaction of BsttRNA<sub>Ser</sub> with B<sub>s</sub>TrmK given that BsttRNA<sub>Ser</sub> prepared in vitro is a better substrate. (C) Selected regions from the superposition of three 2D <sup>1</sup>H–<sup>15</sup>N TROSY experiments, showing amide groups of B<sub>s</sub>TrmK, alone in black, with the SAH in red and with SAH and tRNA<sub>Ser</sub> in green. NMR chemical shift mapping of BsttRNA<sub>Ser</sub> binding reported onto the molecular surface of B<sub>s</sub>TrmK. Pink residues are those that disappeared upon addition of tRNA, residues that experience, upon tRNA binding, NMR chemical shift variations between 0.02 and 0.04 ppm are in orange and larger than 0.04 ppm are in red. Residues in grey are residues for which NMR chemical shift variations could not be measured either because it is a proline residue or a residue for which the NMR signal has disappeared upon SAH addition.
(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, $V_{\text{max}}/K_m$, 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 BstTrmK. The *in vitro* transcribed tRNA^Ser was thus chosen for structural investigations.

Attempts to obtain co-crystals of unmodified BsttRNA^Ser with BstTrmK and SAH were unsuccessful, and we therefore turned to NMR spectroscopy to decipher the interaction of BstTrmK with BsttRNA in solution. No perturbation of BstTrmK resonances was observed for the couple BsttRNA^Ser/BstTrmK in the absence of SAH or for BstTrmK-SAHI in the presence of an RNA mimicking the D-arm, or the D-arm plus the anticodon stem of BsttRNA^Ser (data not shown). The NMR chemical shift mapping was performed by recording a $^{1\text{H}}$-15N TROSY experiment on 1H-15N-labeled BstTrmK bound to unlabeled SAH and unlabeled BsttRNA^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 BstTrmK. Binding of BsttRNA^Ser to BstTrmK-SAHI caused major perturbations on NMR signals of BstTrmK, 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 BsttRNA^Ser, for instance G49, G98, I83, I124, D149 and L156 in Figure 4C, outline the binding surface of BsttRNA^Ser (residues in red and orange, Figure 4C and D). Residues that do not show chemical shift variations upon binding to BsttRNA^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 BstTrmK 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 BstTrmK (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 BsttRNA^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 BstTrmK is dispensable for BstTrmK MTase activity, a BstTrmK variant lacking most of the C-terminal coiled-coil motif was designed. This construct of BstTrmK 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 BsttRNA^Ser upon binding of BstTrmK 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 BsttRNA^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 BstTrmK. Upon binding of BstTrmK, most of the peaks do not experience any chemical shift variation indicating that the global folding of BsttRNA^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 BstTrmK**

A model of the BstTrmK/SAM/BsttRNA^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 BstTrmK 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 BstTrmK. 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 BstTrmK with BsttRNA^Ser. The best model according to the HADDOCK scoring function is part of a cluster made of 102 structures with a mean RMSD$_\text{RNA}$ 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 BstTrmK 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 BstTrmK/SAM/BsttRNA^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 BstTrmK interact with BsttRNA^Ser, mainly through interactions with the phosphodiester backbone of the tRNA. The catalytic domain shares, to a large extend, a larger interaction surface with
Figure 5. Model of the \( B_\text{t} \) TrmK/SAM/\( B_\text{t} \) tRNA\(^\text{Ser}\) complex based on the NMR chemical shift mapping and validated by MTase activity of \( B_\text{t} \) 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 \( B_\text{t} \) tRNA\(^\text{Ser}\) binding on \( B_\text{t} \) TrmK, the D-arm is drawn in orange, the G\(_{13}\)-A\(_{22}\) base pair is in red sticks, the SAM on the \( B_\text{t} \) TrmK catalytic pocket is represented as sticks, (B) The base-pair G\(_{13}\)-A\(_{22}\) is in close proximity to the SAM-methyl donor in the model (C) MTase activity of \( B_\text{t} \) TrmK variants mutated in the tRNA binding interface (D) Positions of the mutations on \( B_\text{t} \) TrmK/SAM/\( B_\text{t} \) tRNA\(^\text{Ser}\) complex, the \( B_\text{t} \) tRNA\(^\text{Ser}\) is represented with the same color code as that used in Figure 2E, the tRNA elements crucial for \( B_\text{t} \) 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 _Bt_ 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 _Bt_ trnRNA^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 _Bt_ 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 _Bt_ TrmK MTase activity, and in the model, these two regions interact directly with _Bt_ TrmK. The model further suggests an explanation for helices abolished MTase activity. The base-pairing 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).

To further challenge the model for validation by experimental data, _Bt_ 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 _Bt_ 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 _Bt_ TrmK enzymatic activity, most probably by affecting tRNA binding given their position in _Bt_ 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 _Bt_ TrmK/SAM/tRNA^Ser_ complex is supported by MTase activity data performed on single-point mutants of _Bt_ TrmK. In the model, _Bt_ 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 _Bt_ TrmK.

**DISCUSSION**

NMR spectroscopy was a valuable method to get insight into _Bt_ TrmK/SAM/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. Crystalisation 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-tRNA 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 _Bt_ 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 _Bt_ TrmK and using the measured NMR data to guide the docking of _Bt_ trnRNA^Ser_ into the protein to build a model of the _Bt_ TrmK/SAM/btRNA^Ser_ complex. This model was further validated by measuring the MTase activity of _Bt_ TrmK variants with single-point mutations located at the proposed protein-tRNA interaction surface.

The formation of m^1^A22 by _Bt_ 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 m^1^A 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 _Bt_ 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 _Bt_ 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 m^1^A formation in tRNA does not require a base catalyst and agrees with (i) the previously proposed m^1^A1 methyltransferase mechanism for the bacterial m^1^A99 tRNA MTase Trm1 (2,14,17), and (ii) with data published on the dual-specific m^1^G9/m^1^A9 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 s²U₈, m⁶G₁₀, Gm₁₈, m¹A₂₂, m²G₁₆, m⁶G₆₆, m³C₄₈, m²U₅₄, m¹A₅₈, ψ₅₅. 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 archaepsine at G₁₅ (D-loop) bound to tRNVA₅ (45), and the human tRNA m¹A₅₈ MTase (Trm6-Trm61) in complex with its tRNA₃Lys₃ substrate (12). In the structure of TGT/tRNVA₅, 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/tRNA₃Lys₃, the D- and the T-arms are detached from each other to expose the A₅₈ 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 s²U₈, m⁶U₅₄ and ψ₅₅ 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 nucleotides located nearer the surface of tRNAs like A₉, no such conformational changes would be expected. According to a docking model of the Sulfolobus acidocaldarius m¹A₉ 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 m¹A modified bases, A₂₂ of B₉TrmK 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 B₉TrmK show that the variable hairpin is still formed upon binding of B₉TrmK, indicating that it is not folded in the λ-form. In the tRNA L-shape structure used for the docking, the A₂₂ 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 R₁₃–A₂₂ 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 R₁₃–A₂₂ pairing is supported by the fact that replacing pyrimidine 13 of a non-substrate tRNA by a purine renders the tRNA substrate of B₉TrmK. In the non-substrate tRNAGly, A₂₂ is involved in a Watson–Crick base pair with U₁₃, which hides the N1-atom of A₂₂ inside the AU base-pair and thus inside the tRNA structure. According to the model proposed in the present work, B₉TrmK can bind the tRNA with no to very little deformation of both partners, the G₁₃–A₂₂ base pair allowing the placement of the N1-atom of A₂₂ 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 B₉TrmK. These base-pairs are located apart from the G₁₃–A₂₂ 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 m¹A₂₂ in tRNAs

The presence of m¹A₂₂ 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, B₉TrmK 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 B₉TrmK specificity is different in vivo like previously observed for the yeast m¹G₉ 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. Brachel (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
1. Boccacietto,P., Machnicka,M.A., Purta,E., Pietakowski,P., Baginski,B., Wirecki,T.K., de Crecy-Lagarde,V., Ross,R., Limbach.P., Kotter,A. et al. (2018) MODOMICS: a database of RNA modification pathways. 2017 update. Nucleic Acids Res., 46, D303–D307.
2. Oerum,S., Degut,C., Barraud,P. and Tsine,C. (2017) m1A Post-Transcriptional modification in tRNAs. Biomolecules, 7, E20.
3. Grosjean,H., Constantinesco,F., Foiret,D. and Benachouen,N. (1995) A novel enzymatic pathway leading to 1-methylinosine modification in Haloferax volcanii tRNA. Nucleic Acids Res., 23, 4312–4319.
4. Roovers,M., Wouters,J., Bujnicki,J.M., Tricot,C., Stalon,V. and Grosjean,H. (2004) A primordial RNA modification enzyme: the case of tRNA (m1A) methyltransferase. Nucleic Acids Res., 32, 465–476.
5. Kempenaers,M., Pourboiton,M., Oudjama,Y., Tkaczuk,K.L., Bujnicki,J.M. and Grosjean,H. (2010) New archaeal methyltransferases forming 1-methyladenosine or 1-methyladenosine and 1-methylguanosine at position 9 of tRNA. Nucleic Acids Res., 38, 6533–6543.
6. Vilardo,E., Nachbagauer,C., Buzet,A., Tsch treffen,A., Holzmann,J. and Rossmann,W. (2012) A subcomplex of human mitochondrial RNA P is a bifunctional methyltransferase–extensive moonlighting in mitochondrial RNA biogenesis. Nucleic Acids Res., 40, 11583–11593.
7. 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.
8. Singh,R.K., Bella,J., Roovers,M., Van Elder,D., Wauters,L., Roovers,M. and Verschae,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.
9. Helm,M., Giege,R. and Florentz,C. (1999) A Watson–Crick base-pair-dissociating methyl group (m1A9) is sufficient for cloverleaf folding of human mitochondrial tRNAs. Biosychemistry, 38, 13338–13346.
10. Sprinzl,M. and Vassilenko,K.S. (2005) Compilation of tRNA (m1A58) methyltransferase (TrmI) from Thermus thermophilus HB27, a protein required for cell growth at extreme temperatures. Nucleic Acids Res., 31, 2148–2156.
11. Barraud,P., Golinelli-Pimpanbe,A., Atmane,C., Sanglier,S., Van Dorsselaer,A., Roovers,M., Dardel,F. and Tsine,C. (2008) Crystal structure of Thermus thermophilus tRNA (m1A58) methyltransferase and biophysical characterization of its interaction with tRNA. J. Mol. Biol., 377, 535–550.
12. Guelorget,A., Barraud,P., Tsine,C. and Golinelli-Pimpanbe,A. (2011) Structural comparison of tRNA (m1A58) methyltransferases revealed different molecular strategies to maintain their oligomeric architecture under extreme conditions. BMC Struct. Biol., 11, 48.
13. Takuma,H., Ushio,N., Minoji,M., Kazayama,A., Shigi,N., Hirata,A., Tomikawa,C., Ochi,A. and Horii,H. (2015) Substrate tRNA recognition mechanism of eubacterial tRNA (m1A58) methyltransferase (TrmI). J. Biol. Chem., 290, 5912–5925.
14. Degut,C., Poncho,C., Bujnicki,J.M., Barraud,P. and Tsine,C. (2016) The m1A58 modification in eubacterial tRNA: An overview of tRNA recognition and mechanism of catalysis by TrmI. Biophys. Chem., 210, 27–34.
15. 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.
16. Guoelget,A., Roovers,M., Guerineau,V., Babay,C., Li,X. and Golinelli-Pimpanbe,A. (2010) Insights into the hyperthermostability and unusual region-specificity of archaeal Pyrococcus abyssi tRNA (m1A57/58) methyltransferase. Nucleic Acids Res., 38, 6206–6218.
17. Raettig,R., Kersten,H., Weissenbach,J. and Dirheimer,G. (1978) Identification 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.
18. Roovers,M., Kaminska,K.H., Tkaczuk,K.L., Gigot,D., Dardel,F. and Barraud,P. (2008) The YqN protein of Bacillus subtilis is the tRNA: m1A22 methyltransferase (TrmK). Nucleic Acids Res., 36, 3252–3262.
19. 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.
20. Hutchinson,C.A. 3rd, Chuang,R.Y., Noskov,V.N., Assad-Garcia,N., Dougherty,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.
21. Th,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.
22. Degut,C., Monod,A., Brachet,F., Crepin,T. and Tsine,C. (2016) In Vitro In vivo production of tRNA for X-ray studies. Methods Mol. Biol., 1320, 37–57.
23. Dardel,F. (1994) MC-Fit: using Monte-Carlo methods to get accurate molecular dynamics inferred from theory and experiments. J. Mol. Biol., 230, 37–57.
24. 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.
25. Emsley,P., Lohkamp,B., Scott,W.G. and Cowtan,K. (2010) Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr., 66, 486–501.
26. Golas,K. and Bork,P. (2011) ProDiv: protein dynamics inferred from theory and experiments. Bioinformatics, 27, 1575–1577.
27. Humphrey,W., Dalke,A. and Schulten,K. (1996) VMD: visual molecular dynamics. J. Mol. Graph., 14, 33–38.
28. Degut,C., Barraud,P., Larue,V. and Tsine,C. (2016) Backbone resonance assignments of the m1A22 tRNA methyltransferase TrmK from Bacillus subtilis. Biomol. NMR Assign., 10, 253–257.
29. Pervushin,K.V., Wider,G. and Wuthrich,K. (1998) Single Transition-to-single transition polarization transfer (ST2-PT) in [15N,1H]-TROSY. J. Biol. Chem., 273, 495–499.
30. 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.
31. 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., Morf, 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 A 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. Ailin, 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) Co-crystal 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., Dzyma, 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 archael Trm10 homologue. _Nucleic Acids Res._, 44, 940–953.

50. Mladenik, 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.