Functional characterization of the YmcB and YqeV tRNA methylthiotransferases of Bacillus subtilis

Brian P. Anton 1,2,*, Susan P. Russell 3, Jason Vertrees 4, Simon Kasif 4,5, Elisabeth A. Raleigh 1, Patrick A. Limbach 3 and Richard J. Roberts 1

1New England Biolabs, Ipswich, MA 01938, 2Bioinformatics Program, Boston University, Boston, MA 02215, 3Department of Chemistry, University of Cincinnati, Cincinnati, OH 45221, 4Department of Biomedical Engineering and 5Center for Advanced Genomic Technology, Boston University, Boston, MA 02215, USA

Received December 8, 2009; Revised April 22, 2010; Accepted April 24, 2010

ABSTRACT
Methylthiotransferases (MTTases) are a closely related family of proteins that perform both radical-S-adenosylmethionine (SAM) mediated sulfur insertion and SAM-dependent methylation to modify nucleic acid or protein targets with a methyl thioether group (–SCH3). Members of two of the four known subgroups of MTTases have been characterized, typified by MiaB, which modifies N6-isopentenyladenosine (i6A) to 2-methylthio-N6-isopentenyladenosine (ms2i6A) in tRNA, and RimO, which modifies a specific aspartate residue in ribosomal protein S12. In this work, we have characterized the two MTTases encoded by Bacillus subtilis and find that, consistent with bioinformatic predictions, ymcB is required for ms2i6A formation (MiaB activity), and yqeV is required for modification of N6-threonylcarbamoyladenosine (t6A) to 2-methylthio-N6-threonylcarbamoyladenosine (ms2t6A) in tRNA. The enzyme responsible for the latter activity belongs to a third MTTase subgroup, no member of which has previously been characterized. We performed domain-swapping experiments between YmcB and YqeV to narrow down the protein domain(s) responsible for distinguishing i6A from t6A and found that the C-terminal TRAM domain, putatively involved with RNA binding, is likely not involved with this discrimination. Finally, we performed a computational analysis to identify candidate residues outside the TRAM domain that may be involved with substrate recognition. These residues represent interesting targets for further analysis.

INTRODUCTION
Transfer RNA molecules from all three domains of life undergo numerous, and often complex, post-transcriptional modifications in the course of maturation. Residue 37, which is 30-adjacent to the anticodon and almost invariably purine, is a frequent target of modification, the specific type of which appears to vary with the identity of residue 36 (the third residue of the anticodon) (1). Among the modified residues found exclusively at this position are N6-isopentenyladenosine (i6A) and N6-threonylcarbamoyladenosine (t6A), and their methylthiolated derivatives 2-methylthio-N6-isopentenyladenosine (ms2i6A) and 2-methylthio-N6-threonylcarbamoyladenosine (ms2t6A) (Figure 1). In Escherichia coli and Bacillus subtilis, i6A or ms2i6A occur in most tRNAs with A36 (reading UNN codons), while t6A or ms2t6A occur in most tRNAs with U36 (reading ANN codons) (1,2). Experimental evidence supports the hypothesis that these hydrophobic modifications stabilize the relatively weak A : U and U : A base pairs formed by the third base of the anticodons of these tRNAs with the first base of their complementary codons, thereby improving translational fidelity (1,3,4).

To the extent that the pathways are known, these modifications are constructed in a stepwise manner, and some of the enzymes involved have been discovered and characterized. Modification of adenosine to i6A is carried out by the gene products of miaA in E. coli and MOD5 in yeast, which transfer the /C2-isopentenyl group from dimethylallyl diphosphate, a mevalonic acid derivative, to A37 N6 (5–8). Mutations in miaA (originally called trpX) result in the accumulation of tRNA with strictly unmodified A37, indicating the formation of i6A is a prerequisite for any further modification (9). Formation of t6A from adenosine is less well characterized but is known to be an ATP-dependent reaction requiring threonine and

*To whom correspondence should be addressed. Tel.: +1 978 927 5054; Fax: +1 978 921 1350; Email: anton@neb.com

© The Author(s) 2010. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/2.5), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/nar/article-abstract/38/18/6195/1062111 by guest on 10 February 2018
carbonate (10). Recent genetic evidence shows that E. coli
yrdC and yeast SUA5 are required for this modification,
but the involvement of other proteins in the reaction
cannot yet be ruled out (11). Although not yet
demonstrated, it seems reasonable to expect that t^6A for-
mation is a requisite first step for further modification
based on analogy to the i^6A pathway.

Methylthiolation (–SCH3 addition) at C2 of the adeno-
sine ring forms ms2t^6A and ms^2t^6A from i^6A and t^6A,
respectively (Figure 1). The miaB gene product has been
shown to catalyze the formation of ms^2t^6A in both
E. coli and Thermotoga maritima (12–14). This reaction
was originally hypothesized to occur by sequential steps,
catalyzed by separate enzymes: sulfur insertion by a miaB activity, followed by
S-adenosylmethionine (SAM)-dependent methylation of a thiolated intermediate
by a miaC activity (15,16). This was seemingly borne out
by experiments using an E. coli rel met cys mutant (auxo-
trophic for methionine and cysteine yet capable of RNA
synthesis in their absence). Under conditions of methio-
nine starvation, such cells accumulated an uncharacterized
i^6A derivative that was capable of subsequent labeling
using 14C-SAM, presumably by methylation (15). It was
speculated that this uncharacterized nucleoside was the
intermediate s^2i^6A, but this has yet to be confirmed, and
this intermediate has not been observed elsewhere.
Furthermore, recent results with the closely related
enzyme RimO contradict this, suggesting instead that
methylation of the sulfur atom occurs on an enzyme-
bound FeS cluster prior to insertion (17). In any case,
once discovered, purified MiaB was shown to be a
methylthiotransferase (MTTase), responsible for both
thiolation and methylation of t^6A (18).

Phylogenetic analysis of the MTTase family shows that
it consists of four clades, members of two of which have
been characterized (19,20). One clade includes MiaB and
its homologs, and is found exclusively in bacteria and eu-
karyotic organelles. The second characterized clade, also
exclusively bacterial, includes RimO, a MTTase that
modifies D88 of ribosomal protein S12 in
E. coli (19). A third bacterial clade, typified by B. subtilis
YqeV, and a fourth clade, found exclusively in archaea and eukaryotes
and typified by Methanocaldococcus jannaschii Mj0867,
remain uncharacterized. Given that the nucleoside
ms^2t^6A has been observed both in bacteria (including
B. subtilis (21) but not E. coli) and in archaea (22), it
seems reasonable to expect that members of both of
these clades are responsible for methylthiolating t^6A to
ms^2t^6A, with the phylogenetic distinction reflecting the
ancient split of the bacterial domain from archaea and
eukaryotes rather than a functional differentiation within
the protein family.

MTTases are members of the so-called ‘radical-SAM’
superfamily of proteins, which use a reducing equivalent
from a prosthetic [4Fe–4S]1+ cluster to cleave SAM,
generating methionine and a reactive 5’-dA radical (23).

In the case of MiaB and RimO, this radical facilitates the
difficult C–H to C–S bond conversion in the nucleoside or
amino acid substrate, respectively, by abstracting the
hydrogen atom and creating a reactive substrate radical
that is amenable to sulfur insertion. All MTTases share a
common tripartite domain structure, with the central
domain responsible for this radical-SAM chemistry. The
N-terminal domain has been shown in MiaB and RimO to
harbor a second FeS cluster (17,24), a feature common to
other radical-SAM proteins that catalyze sulfur insertion
reactions such as BioB and LipA (25); this cluster is
speculated to serve as the immediate sulfur donor in the
thiolation reaction. The C-terminal TRAM domain has
been shown to bind RNA in the context of other proteins (26), suggesting a similar function in tRNA-
modifying MTTases. Its presence in the protein-modifying
RimO is less obvious, but a recently published structure
suggests that in this particular protein it has adapted to
bind the protein substrate rather than RNA (27).

In this work, we identify and characterize the two
MTTases encoded in the genome of B. subtilis str. 168,
products of the ymcB and yqeV genes. We confirm the
earlier prediction that YqeV performs the novel t^6A
methylation function, and verify that YmcB, a
MiaB ortholog, methylthiolates i^6A. We then went on to
construct several chimeric proteins derived from these two
closely-related enzymes and show that, RNA-binding
function of the TRAM domain notwithstanding, the
ability of these two enzymes to discriminate between i^6A
and t^6A resides in the N-terminal and/or radical-SAM
domains, and not in the TRAM domain. Finally, we
were unable to complement the loss of yqeV in
B. subtilis with either the mj0867 gene or a mesophilic
ortholog in trans. We suggest that yqeV be renamed
mttB, for the second step in tRNA-methylthiolation,
reserving mttA for the t^6A modifying activity.

MATERIALS AND METHODS

Enzymes

All restriction enzymes and T4 DNA ligase were from
New England Biolabs (Ipswich, MA, USA). All PCR re-
actions were performed using Phusion DNA polymerase
(New England Biolabs) in the HF buffer, unless otherwise
noted.

Bacterial strains and media

Table 1 describes the bacterial strains and plasmids used in
this study. Bacillus subtilis strains BSF2608 and YQEVD
were derived from strain 168 by single-crossover
(Campbell type) insertion of the plasmid pMUTIN (28).
The insertion in BSF2608 is after ymcB nt 533, and that in

Figure 1. Schematic structures of the methylthiolated nucleic acid residues from B. subtilis. (A) t^6A when X = H and ms^2t^6A when
X = SCH3. (B) i^6A when Y = H and ms^2i^6A when Y = SCH3.
Table 1. Bacterial strains and plasmids used in this study

<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Genotype or description</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bacillus subtilis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>168</td>
<td>trpC2</td>
<td>(52)</td>
</tr>
<tr>
<td>BSF2608</td>
<td>168, ymcB::pMUTIN</td>
<td>(29)</td>
</tr>
<tr>
<td>YQEVD</td>
<td>168, yqeV::pMUTIN</td>
<td>(29)</td>
</tr>
<tr>
<td>B(124)</td>
<td>BSF2608 × pDM124c7 1</td>
<td>This work</td>
</tr>
<tr>
<td>V(124)</td>
<td>YQEVD × pDM124c7</td>
<td>This work</td>
</tr>
<tr>
<td>B(B)</td>
<td>BSF2608 × pDMymcB</td>
<td>This work</td>
</tr>
<tr>
<td>B(MM)</td>
<td>BSF2608 × pDMmmar</td>
<td>This work</td>
</tr>
<tr>
<td>V(V)</td>
<td>YQEVD × pDMqeV</td>
<td>This work</td>
</tr>
<tr>
<td>V(MJ)</td>
<td>YQEVD × pDMmjm0867</td>
<td>This work</td>
</tr>
<tr>
<td>V(MM)</td>
<td>YQEVD × pDMmmar</td>
<td>This work</td>
</tr>
<tr>
<td>B(B1V)</td>
<td>BSF2608 × pDM-B1V</td>
<td>This work</td>
</tr>
<tr>
<td>B(B2V)</td>
<td>BSF2608 × pDM-B2V</td>
<td>This work</td>
</tr>
<tr>
<td>B(V1B)</td>
<td>BSF2608 × pDM-V1B</td>
<td>This work</td>
</tr>
<tr>
<td>B(V2B)</td>
<td>BSF2608 × pDM-V2B</td>
<td>This work</td>
</tr>
<tr>
<td>V(B1V)</td>
<td>YQEVD × pDM-B1V</td>
<td>This work</td>
</tr>
<tr>
<td>V(B2V)</td>
<td>YQEVD × pDM-B2V</td>
<td>This work</td>
</tr>
<tr>
<td>V(V1B)</td>
<td>YQEVD × pDM-V1B</td>
<td>This work</td>
</tr>
<tr>
<td>V(V2B)</td>
<td>YQEVD × pDM-V2B</td>
<td>This work</td>
</tr>
<tr>
<td>Plasmids</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pDM124c7</td>
<td>E. coli/B. subtilis shuttle plasmid with the promoter region from the B. amylophilicaeens α-amylase gene followed by a multiple cloning site</td>
<td>J. Benner and D. Martin (unpublished)</td>
</tr>
<tr>
<td>pDMymcB</td>
<td>pDM124c7 with ymcB cloned between NdeI and XhoI</td>
<td>This work</td>
</tr>
<tr>
<td>pDMqeV</td>
<td>pDM124c7 with yqeV cloned between NdeI and XhoI</td>
<td>This work</td>
</tr>
<tr>
<td>pDMmjm0867</td>
<td>pDM124c7 with mjm0867 cloned between NdeI and XhoI</td>
<td>This work</td>
</tr>
<tr>
<td>pDMmmar</td>
<td>pDM124c7 with mmar cloned between NdeI and XhoI</td>
<td>This work</td>
</tr>
<tr>
<td>pDM-B1V</td>
<td>pDM124c7 with ymcB::yqeV chimera B1V between NdeI and XhoI</td>
<td>This work</td>
</tr>
<tr>
<td>pDM-B2V</td>
<td>pDM124c7 with ymcB::yqeV chimera B2V between NdeI and XhoI</td>
<td>This work</td>
</tr>
<tr>
<td>pDM-V1B</td>
<td>pDM124c7 with ymcB::yqeV chimera V1B between NdeI and XhoI</td>
<td>This work</td>
</tr>
<tr>
<td>pDM-V2B</td>
<td>pDM124c7 with ymcB::yqeV chimera V2B between NdeI and XhoI</td>
<td>This work</td>
</tr>
</tbody>
</table>

YQEVD is after yqeV nt 280 (29). Insertions were confirmed by PCR amplification (data not shown).

To make B. subtilis competent cells, 5 ml Rich medium supplemented with 3 mM MgSO4 was inoculated with a single colony and grown at 37°C with vigorous aeration to OD600 ~1.0. Then, 0.5 ml of this culture was used to inoculate 10 ml minimal medium (0.9× PC buffer, 2% glucose, 3 mM MgSO4, 2.5 mg/ml potassium aspartate, 11 μg/ml ferric ammonium citrate, 50 μg/ml phenylalanine, 50 μg/ml tryptophan), which was then grown 4 h at 37°C with vigorous aeration. Three milliliters of 50% glycerol was then added to the culture, which was frozen at −80°C in 1 ml aliquots until ready for use. The composition of 1× PC buffer was 0.1 M potassium phosphate, 3 mM trisodium citrate, pH adjusted to 7.5 with KOH.

Transformation of B. subtilis was accomplished by mixing 1.5 μg plasmid DNA with 200 μl competent cells, growing 37°C 1 h with vigorous aeration, and plating on Rich medium supplemented with 12.5 μg/ml chloramphenicol as required. Plasmids used for B. subtilis transformation were purified from E. coli NEBTurbo (New England Biolabs), a recA+ strain.

Plasmid construction

The shuttle plasmid used for complementation, pDM124c7 (J. Benner and D. Martin, unpublished data), contains a ColE1 origin of replication and ampicillin-resistance marker for propagation in E. coli and a chloramphenicol-resistance marker for selection in B. subtilis. Genes were cloned at the NdeI-XhoI sites, with expression in B. subtilis driven constitutively by the B. amylophilicaeens α-amylase promoter and expressed proteins targeted to the cytoplasm.

Genes encoding MTases were cloned by PCR amplification, with ymcB and yqeV from B. subtilis 168 genomic DNA, gene mjm0867 from M. jannaschii genomic DNA, and gene mmar from Methanococcus maripaludis genomic DNA, using primers described in Table 2. All four genes were digested at the NdeI and XhoI sites (underlined in Table 2) and inserted at the same sites in pDM124c7. Cloning steps were performed in E. coli, and constructs verified by DNA sequencing.

An internal NdeI site in mmar necessitated a second step in the cloning of that gene, as follows. The cloning steps above yielded a construct containing a truncated gene, with the 450 bp at the 5′ end of the gene missing. The missing fragment was amplified using primers mmar_F and mmar_I (Table 2), digested with NdeI and inserted at the NdeI site of the truncated construct, thus reconstructing the intact gene. Our strain of M. maripaludis is distinct from the four strains whose genome sequences have been published to date, so we use the generic name mmar to refer to the mjm0867 ortholog from our strain. The mmar gene sequence and a protein sequence alignment of Mmar with the four published M. maripaludis orthologs are shown in Supplementary Figure S1.

Plasmids containing ymcB::yqeV chimeric genes were constructed using the USER Friendly Cloning Kit (New England Biolabs).
England Biolabs). Two fragments (designated ‘N’ and ‘C’), which together encompassed the entirety of the plasmid template, were amplified from each of pDMycB and pDMyqeV using PfuCx Hotstart polymerase (Stratagene, Cedar Creek, TX, USA), with primers described in Table 2. Junctions between the two fragments were (i) within the ymcB or yqeV gene, as described in ‘Results’ section, and (ii) within the ampicillin-resistance marker. Following treatment with the USER enzyme, fragments were annealed such that amplicin-resistance marker. Following treatment with described in 'Results' section, and (ii) within the fragments were (i) within the USER enzyme, fragments were annealed such that junctions between the two primers described in Table 2. Junctions between the two fragments were (i) within the ymcB or yqeV gene, as described in ‘Results’ section, and (ii) within the ampicillin-resistance marker. Following treatment with the USER enzyme, fragments were annealed such that ampicillin-resistance marker. Following treatment with described in 'Results' section, and (ii) within the fragments were (i) within the USER enzyme, fragments were annealed such that primer (mmar_I), from the NdeI site (underlined) to the 3' end; for yqeVch and bla USER primers, the entire primer.

### tRNA purification

The tRNA was purified from 500 ml *B. subtilis* cultures grown at 37°C in rich medium and harvested at OD_{600} between 0.8 and 1.0. Cell pellets were washed with 4 ml TE (10 mM Tris–HCl pH 8.0, 1 mM EDTA), then incubated 37°C 3 h in 4 ml TE with 40 mg/ml lysozyme. Samples were then vigorously mixed with 16 ml TRI Reagent (Sigma-Aldrich, St Louis, MO, USA) and left 10 min at room temperature. Chloroform (3.2 ml) was added, and samples were left for 5 min at room temperature before centrifuging at 8000 r.p.m. for 10 min. Total RNA was precipitated from the aqueous phase with 0.3 M sodium acetate and 70% ethanol and collected by centrifugation at 12 000 r.p.m. for 20 min. tRNA was solubilized by vortexing pellets in 3.2 ml TE with 2 M LiCl and recentrifuging at 8000 r.p.m. for 10 min, then precipitated from the supernatant with 70% ethanol and resuspended in 450 μl TE.

The tRNA was cleaned for LC/MS by three rounds of precipitation with 0.3 M ammonium acetate and 70% ethanol, and finally resuspended in 450 μl TE. Typical yield was roughly 1 mg. Eighty micrograms of this was then purified through Nucleobond AX-R 80 ion exchange columns (Macherey-Nagel, Bethlehem, PA, USA), collecting 1.2 ml elute. tRNA was precipitated from the eluate with 45% isopropanol and resuspended in 40 μl H_{2}O. Typical yield was 65 μg tRNA.

### Liquid chromatography and ESI-MS analysis

Nucleosides were separated, at room temperature, on a Hitachi HPLC system (L-7100 pump) with UV detection at 260 nm (L-7400 UV detector). The column used was a Supelcosil LC-18-S (25 cm × 2.1 mm, 5 μm diameter particles, with a 2 cm × 2.1 mm guard column), which was run at a flow rate of 0.3 ml min⁻¹. The mobile phases used were (i) 5 mM ammonium acetate pH 5.3, and (ii) acetonitrile/water (40:60, v/v) with the gradient described by Pomerantz and McCloskey (30) with minor alterations. The column effluent was split, with one-third directed to a Thermo LTQ-XL (or LTQ-FT) mass spectrometer and two-thirds to the UV detector. Mass spectra were recorded in the positive ion mode from m/z 103 to 510. Electrospray and MS conditions were optimized using adenosine, introduced post-column. The capillary temperature used was 275°C, source voltage 3.7–5 kV, sheath gas flow 45
RESULTS
Identification of B. subtilis MTMase genes
A previous BLASTP search of the translated B. subtilis 168 genome using E. coli MiaB as the query revealed two putative MTMases, products of ymcB (BSU17010) and yqeV (BSU25430). Based on phylogenetic considerations, ymcB was predicted to be a true miaB ortholog, with its product responsible for the ms\(^{2t}\)tA modification observed in B. subtilis (31), while yqeV belonged to a novel subfamily predicted (19) to be responsible for the observed ms\(^{2t}\)tA modification (32). ymcB appears to be a part of bicistronic operon with the gene ymeA, with potential −35 and −10 elements upstream of ymcB and a putative rho-independent terminator downstream of ymcA (33). yqeV is the terminal gene of the heptacistronic dnaK operon, whose transcripts have been mapped and regulation described by Homuth and coworkers (who refer to yqeV as orf50) (34,35). It is expressed from both a heat-inducible promoter upstream of hrcA and from a constitutive promoter upstream of dnaJ. The tetracistronic constitutive transcript encodes, between dnaJ and yqeV, two other methyltransferases involved with modifying the translational machinery: yqeT, a probable homolog of the ribosomal protein L11 MTase, and yqeU, a probable homolog of the 16S rRNA MTase rsmE.

Characterization of mutant phenotypes
Mutant strains with insertions in both ymcB and yqeV have been constructed (29), indicating that neither gene is essential under normal growth conditions. We obtained these mutant strains as well as the parental strain 168 from the B. subtilis sequencing consortium (http://bacillus.genome.ad.jp). To investigate the effect of these insertions mutants on tRNA modification, we purified and digested to nucleosides total tRNA from these insertion mutants on tRNA modification, we obtained these mutant strains as well as the parental strain 168, indicating the presence of a specific loss of the ms\(^{2i}\)tA peak (Figure 4B), and strain YQEvd (yqeV) shows specific loss of the ms\(^{2i}\)tA peak (Figure 4C). Both genes were separately subcloned into the plasmid pDM124c7 (Materials and methods’ section) and reintroduced in trans into the respective mutant strains to complement the disrupted alleles. In both cases, the intact plasmid-encoded gene rescued the modification-deficient phenotype, whereas the empty plasmid vector did not, indicating that ymcB and yqeV are required for ms\(^{2}\)tA and ms\(^{2}\)tA modification, respectively, confirming previous phylogenomic predictions (19). These results are summarized in Table 3, and complete LC/UV and LC/MS data for all samples are included as Supplementary Figure S2. Although we will continue to refer to these genes and their products here as ymcB and yqeV for clarity, we suggest that ymcB be renamed miaB and suggest the name yqeV.

UV chromatograms of nucleoside digests from the two mutant strains are shown in Figure 4. Strain B6208 (ymcB) shows specific loss of the ms\(^{2}\)tA peak (Figure 4B), and strain YQEvd (yqeV) shows specific loss of the ms\(^{2}\)tA peak (Figure 4C). Both genes were separately subcloned into the plasmid pDM124c7 (Materials and methods’ section) and reintroduced in trans into the respective mutant strains to complement the disrupted alleles. In both cases, the intact plasmid-encoded gene rescued the modification-deficient phenotype, whereas the empty plasmid vector did not, indicating that ymcB and yqeV are required for ms\(^{2}\)tA and ms\(^{2}\)tA modification, respectively, confirming previous phylogenomic predictions (19). These results are summarized in Table 3, and complete LC/UV and LC/MS data for all samples are included as Supplementary Figure S2. Although we will continue to refer to these genes and their products here as ymcB and yqeV for clarity, we suggest that ymcB be renamed miaB and suggest the name yqeV.

The ms\(^{2}\)tA- and ms\(^{2}\)tA-deficient B. subtilis strains identified above are, in addition, suitable for characterizing heterologous MTMases of unknown substrate specificity. We cloned mj0867 from M. jannaschii (a hyperthermophile), and the orthologous gene nmar from M. maripaludis (a mesophilic relative of M. jannaschii), into pDM124c7 and introduced them separately into the ms\(^{2}\)tA - strain YQEvd as above. However, despite being predicted to be ms\(^{2}\)tA MTMases, neither gene was able to rescue the modification-deficient phenotype (Table 3 and Supplementary Figure S2). In addition, nmar failed to
rescue the ms$^{2i6A}$ deficiency in strain BF2608. Both archaeal proteins appeared largely insoluble in the B. subtilis extracts (data not shown), suggesting that these proteins may be misfolded, and therefore inactive, in this heterologous context.

Construction and characterization of YmcB/YqeV chimeric proteins

We further used the availability of the modification-deficient strains to explore the determinants of substrate recognition in YmcB and YqeV by constructing chimeric proteins. Specifically, we sought to determine whether the TRAM domain, believed to bind RNA, also conferred the recognition of the N$^6$ moiety that differentiates the substrates of these two proteins. YmcB and YqeV share significant sequence similarity across all three domains (Figure 5). We wished to generate a breakpoint between the radical-SAM and TRAM domains at which the TRAM domains could be swapped. However, we noted significant sequence conservation in the region between these two domains as defined by Pfam (Figure 5), and it was thus unclear where the functionally required elements of one domain ended and the other began. Accordingly, we generated two alternative breakpoints for each construct, one close to the Pfam-defined boundary of the radical-SAM domain (breakpoint 1, Figure 5 red arrows) and the other close to the boundary of the TRAM domain (breakpoint 2, Figure 5 green arrows). Four constructs were generated in all, designated B1V, B2V, V1B and V2B, with the first letter ('B' for YmcB and 'V' for YqeV) describing the source of the N-terminal and radical-SAM domains, the second letter describing the source of the TRAM domain, and the number between them indicating the location of the breakpoint. Note that, due to the requirements of the cloning methodology, breakpoint 2 differs by three residues between B2V and V2B.

We attempted to rescue the modification defect in both BSF2608 and YQEvd with each of the four chimeric constructs as was done for ymcB, yqeV and mj0867. tRNA was purified and the modified nucleosides examined for each of these eight strains (four chimeras in two backgrounds), and we observed successful rescue in two of them: in strains B(B1V) and B(B2V), chimeras B1V and

Figure 3. LC/MS spectra of the adenosine derivatives t$^6$A, ms$^{2i6A}$, i$^6$A and ms$^{2i6A}$ from digested wild-type B. subtilis 168 total tRNA. (A) Selected ion chromatograms of molecular and base ions. (B) Mass spectra.
B2V rescued the loss of ms²t⁶A modification (Table 3 and Supplementary Figure S2). These results are consistent with differential recognition of the t⁶ or i⁶ substituent residing in either the N-terminal or radical-SAM domain of these MTTases rather than in the TRAM domain. The inability of V1B and V2B to rescue the ms²t⁶A defect (or the ms²i⁶A defect) is likely due to misfolding of the proteins. Indeed, all four chimeric proteins appeared largely insoluble in the cell extracts, which was not the case with the two native proteins (data not shown). However, the B1V and B2V proteins were significantly more highly expressed than V1B and V2B, and enough protein may have folded correctly in these cases to observe enzymatic activity.

In order to pinpoint residues within the N-terminal and/or radical-SAM domains that may be responsible for substrate recognition, we performed a computational analysis that is described more fully in the Supplementary Data. Briefly, we examined a collection of sequences from the three characterized MTTase subfamilies, MiaB, YqeV and RimO, to identify residues that are conserved within each subfamily but differ between them. Such residues are candidates for discriminating the three types of substrates specific for each subfamily. Of the 13 highest scoring residues in our analysis, shown in Supplementary Figure S3, 6 are closely proximal to the 6 invariant cysteines involved with coordinating the FeS clusters, suggesting they may be of structural importance, and 2 more are within the TRAM domain. The remaining five, all within the radical-SAM domain yet distal in the primary sequence to the FeS cluster motif, represent strong candidates for substrate interaction; these are residues 195, 231, 232, 297 and 328, using numbering from B. subtilis YqeV. Note that the highest scoring of these, residue 328, is technically outside of the Pfam-defined radical-SAM domain, but is just upstream of breakpoint 1 in our chimera construction, and therefore segregated with the radical-SAM domain in all of the chimeras we constructed.

Relative modification levels

A semi-quantitative investigation of the LC/UV data revealed apparent differences in the relative amounts of

Figure 4. UV chromatograms of B. subtilis total tRNA nucleoside digests from the wild-type and two mutant strains, focused on later eluting species. Peaks representing t⁶A, ms²t⁶A, t⁶A and ms²i⁶A are labeled, with parentheses indicating loss of the expected peak. A non-nucleoside peak eluting close to ms²t⁶A is indicated by an asterisk. (A) Wild-type strain 168, (B) strain BSF2608 and (C) strain YQEvd.
the modifications noted in Table 3 for the 16 strains listed there. We first calculated the peak area ratios of t 6A, ms 2i6A, i 6A and ms 2i6A to m 5U, a modified nucleoside found at position U54 of all sequenced B. subtilis tRNA species (http://modomics.genesilico.pl), and whose modification is independent of those studied in this work; these ratios can therefore be interpreted as the fractions of all tRNA molecules bearing t 6A, ms 2t6A, i 6A or ms 2i6A, respectively. We then calculated the ratios of ms 2i6A and ms 2t6A to (ms 2i6A+i 6A) and (ms 2t6A+t 6A), respectively, that is, the fractions of i 6A- and t 6A-containing residues that are methylthiolated. This information is summarized in Supplementary Table S2, and the original peak areas from which the ratios were calculated can be found in Supplementary Table S1.

The fraction of t 6A residues that are methylthiolated is roughly similar to the wild-type across ms 2t6A+ strains at ~10%, regardless of whether YqeV is expressed from the chromosome or the plasmid (Supplementary Table S2). The wild-type strain 168 exhibits roughly equal amounts of t 6A and ms 2t6A (50% methylthiolation), but in other ms 2t6A strains this fraction ranges as high as 70% in strain B(B), where wild-type YmcB is expressed from the plasmid, to as low as 8% for strain B(B1V) (Supplementary Table S2). These more extreme values are likely caused by promiscuous modification of non-cognate tRNA by overexpressed YmcB in the case of B(B), and misfolding of a substantial fraction of chimeric protein in the case of B(B1V), as described earlier.

A final observation from the LC/UV and LC/MS data is consistent with the UV results except where indicated (footnote c). UV peak indicative of this modified nucleoside. MS results were consistent with the UV results except where indicated (footnote c).

Table 3. Modified nucleosides observed in B. subtilis strains

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype^a</th>
<th>ymcB</th>
<th>yqeV</th>
<th>t 6A</th>
<th>ms 2t6A</th>
<th>t 6A</th>
<th>ms 2i6A</th>
</tr>
</thead>
<tbody>
<tr>
<td>168</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>BSF2608</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B(124)</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>B(B)</td>
<td>+</td>
<td></td>
<td></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>B(MM)</td>
<td>(?)^p</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>B(B1V)</td>
<td>(?)^p</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>B(B2V)</td>
<td>(?)^p</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>V(V1B)</td>
<td>(?)^p</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>V(B2V)</td>
<td>(?)^p</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>YQEVD</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>V(124)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>V(V)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>V(V2B)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>V(V1B)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>V(B1V)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

^aIntact genes indicated by ‘plus’; insertion mutants indicated by ‘minus’; question marks indicate the presence of a heterologous gene that may or may not complement the gene in question; superscript ‘P’ indicates plasmid-encoded copy.

^bPhenotypes indicate the presence (plus) or absence (minus) of an LC/UV peak indicative of this modified nucleoside. MS results were consistent with the UV results except where indicated (footnote c).

^cWeak MS signal detected in these samples (Supplementary Figure S2).

DISCUSSION

Regulation of MTase activity

Early studies of ms 2i6A in bacteria revealed that the methylthiol group, in modulating codon–anticodon interactions, acts as an environmental sensor, mediating cellular responses to changing growth conditions. In E. coli, most i 6A residues are methylthiolated to ms 2i6A at all stages of growth under normal conditions (31,39). However, i 6A is undermodified under conditions of iron limitation in E. coli and other Proteobacteria (40–42). An effect of the loss of ms 2t6A and tRNA 3⁹Phc is to specifically reduce the efficiency at which these tRNAs read the regulatory leader sequences preceding the tryptophan and phenylalanine biosynthetic operons, respectively. The slower rate at which these sequences are read by ribosomes disfavors the formation of attenuator structures that prematurely terminate transcription, with the end result being specific up-regulation of the trp and phe operons in response to iron depletion.

The same effect of iron-limitation has been observed in B. subtilis (43), but in addition, the degree of ms 2 modification in this differentiating bacterium is growth-phase dependent. During exponential growth, ms 2i6A levels are

modification pathways. In three of the four strains where no ms 2t6A peak could be detected by LC/UV, trace amounts of the diagnostic MH + and BH + ions could still be detected by LC/MS (Table 3 and Supplementary Figure S2). One of these three was the original insertion mutant strain BSF2608, suggesting that this trace modification could be due either to incomplete inactivation of ymcB by the insertion, or to trace promiscuous activity by the fully active YqeV in these strains. Levels of this nucleotide are low enough that stochastic variation between samples could explain the absence of LC/MS-detectable amounts in the fourth strain.

Other tRNA modifications present

In addition to enabling identification of t 6A, ms 2t6A, i 6A and ms 2i6A, Figure 2 represents one of the most comprehensive censuses of tRNA modification in B. subtilis to date, with a total of 23 modified nucleosides identified as present in physiologically relevant amounts. The spectrum of modified nucleosides in Figure 2 is largely consistent with previous studies, including those of Vold and coworkers (36,37), and the collection of sequenced tRNA species in the MODOMICS database (2). All nucleosides positively identified in at least one of those studies can be identified in Figure 2 with the exception of s 4U and cmnm 5s2U. In addition, we identified k 2C, which was previously identified only in a more specialized analysis (38). We detected three nucleosides (m 5G, m 7A and m 5i 6A) that are typical of rRNA and therefore may be indicative of trace rRNA contamination in our tRNA preparation. One other nucleoside, ms 2A, is present in trace amounts in all samples and is assumed to be a breakdown product created during tRNA purification. We could identify no modified nucleosides besides those observed in previous studies.
Breakpoints 1 (red) and 2 (green) used to construct chimeric proteins (see text). Defined by Pfam are boxed in pink (UPF0004; PF00919), cyan (radical-SAM; PF04055) and brown (TRAM; PF01938). Arrows indicate the relative activity of behavior of YmcB and YqeV suggests there are additional levels of regulation at work.

Confirmation of the functions encoded by the ymcB and yqeV MTTase genes should facilitate better understanding of the precise regulatory networks that coordinate their activities. The ymcB gene appears to be co-transcribed with ymcA, a master regulator of biofilm formation that is required for pellicle formation and colony differentiation (33). This gene pair is extensively conserved within the Bacilli, but the functional relationship between the two genes remains unclear. The fact that both genes are associated with cell differentiation is nonetheless intriguing. Based on published results, these genes do not appear to be part of any well-characterized regulon, particularly those associated with sporulation (σ^E, σ^I), Spo0A or iron utilization (Fur) (45-49). Given the condition-dependent pattern of ms^2i6A modification described earlier, this is somewhat surprising and suggests that some regulatory element remains to be discovered. Transcription of yqeV appears to be dual regulated: a background level of transcript (along with other modification-related genes and dnaJ) expressed from a vegetative promoter, supplemented with expression from a second, longer dnaK transcript governed by a CIRCE-regulated σ^E promoter. Again the functional connection between the sRNA-related genes and the translation-related modification genes is unclear, but this gene order is again well conserved among both Bacilli and Clostridia, so there appears to be some selective pressure to maintain this arrangement.

Substrate recognition

The substrate elements recognized by MTTases are not completely defined. However, the absence of observed ms^3A in miaA mutants indicates that the i^6A group is absolutely required for MiaB modification. Given the strong similarity among all MTTases, it seems reasonable to expect that YqeV should correspondingly require the t^A group for modification; however, no t^A^- phenotype has been created in an organism harboring an YqeV-type MTTase to test this. Our results with YmcB/YqeV chimeric proteins, while incomplete, nonetheless suggest that recognition of the N^6 substituent lies in either the N-terminal or radical-SAM domain, not in the C-terminal TRAM domain. The TRAM domain may function as a non-specific RNA clamp by which the MTTase holds its substrate, such that TRAM domains function as a non-specific RNA clamp by which the MTTase holds its substrate, such that TRAM domains are completely interchangeable, or alternatively, TRAM may confer additional substrate specificity, perhaps at the RNA sequence level. We have shown that the B1V and B2V chimeras form ms^2i6A, but without further characterization of individual tRNA species, it cannot be assumed that these chimeras modify the same set of tRNAs as YmcB.

Bacillus subtilis YmcB and YqeV both appear relatively restrictive in their additional substrate requirements. Of the 7 i^6A-modified tRNA species, only tRNA^Phe^ and tRNA^lys^ are methylthiolated (2), and of the 9 t^A-modified species, only tRNA^lys^ is methylthiolated (2,21). Pierrel and coworkers have shown that E. coli MibA can modify a 17 base RNA oligonucleotide (N^6)-related genes and the translation-related modification genes is unclear, but this gene order is again well conserved among both Bacilli and Clostridia, so there appears to be some selective pressure to maintain this arrangement.
of these regions are provided as Supplementary Figure S4.) In the case of i^A-bearing tRNAs, the only sequence feature outside the anticodon that is unique to the two methylthiolated species is the A31:Ψ39 bp at the end of the stem-loop. Ψ:A base pairs at the corresponding position of a synthetic stem-loop oligonucleotide have been shown to have a stabilizing effect on the structure (50). However, tRNA1^{TP}, which has a Ψ31:A39 bp, is not methylthiolated, suggesting that if this base pair has any significance with respect to YmcB recognition, the important element may be Ψ39 itself (and/or A31) rather than the stem structure it confers. In the case of i^A-bearing tRNAs, there is only a single methylthiolated example, tRNA1^{Lys}. It again contains an A31:Ψ39 bp, but tRNA1^{Thr} also contains this base pair yet is not methylthiolated. Indeed, the sequences of tRNA1^{Lys} and tRNA1^{Thr} are nearly identical in this region outside the anticodon, except for the U27:A39 bp, which is unique to tRNA1^{Lys}.

Escherichia coli MiaB appears much less restrictive in its modification than B. subtilis YmcB: of 10 i^A-modified tRNA species, nine are methylthiolated. In addition, mutant tRNAs that have been altered to become MiaA substrates, such as tRNA2^{Glu} C36A, tRNA3^{Lys} U36A and tRNA4^{Met} A35U U36A (51), acquire ms2i6A37, indicating they too are MiaA substrates when isopentenylated. It is unclear to what degree the difference in the suite of methylthiolated tRNA species between these two organisms is due to changes in the proteins and to what degree it is due to changes in tRNA, as both enzyme and substrate coevolve. In B. subtilis, ms2iA modification has come to be associated with sporation, a process that has no analog in E. coli. We speculate that this additional role may have influenced these changes, but alternative explanations are certainly possible. In any case, proper assessment of any differences in substrate specificity between E. coli MiaB and B. subtilis YmcB will require complementation of one with the other and full analysis of the modified tRNA species. This work provides a first step towards an understanding of the process of substrate recognition by MTases, and further progress in this area may shed light on multiple key biological processes.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors are indebted to Alessandro Galizzi, Tsutomu Sato and the National BioResource Project (NIG, Japan): B. subtilis for strains; to Jack Benner, Deana Martin, Marta Meda and Agnieszka Sekowska for plasmids; and to Antoine Danchin for additional help in obtaining materials. The authors also thank Marion Sibley for maintaining bacterial strains, and Irena Nikcevic and Balasubrahmanyam Addepalli for technical advice. Special thanks to Melanie Berkmen for protocols and advice related to B. subtilis. Finally, we thank Don Comb for support.

FUNDING

National Science Foundation (CHE 0602413 cofunded by the MPS/CHE and BIO/IDBR Divisions and by the MPS Office of Multidisciplinary Activities, and CHE 0910751 cofunded by the MPS/CHE and BIO/MBD Divisions, to P.A.L.); National Institutes of Health (NIH R019900 to P.A.L.); and National Institutes of Health (NIH 1R01GM092602-01 to S.K., J.V. and R.J.R.). Funding for open access charge: New England Biolabs.

Conflict of interest statement. None declared.

REFERENCES


