Colocation of Genes Encoding a tRNA-mRNA Hybrid and a Putative Signaling Peptide on Complementary Strands in the Genome of the Hyperthermophilic Bacterium Thermotoga maritima†

Clemente I. Montero, Derrick L. Lewis, Matthew R. Johnson, Shannon B. Conners, Elizabeth A. Nance, Jason D. Nichols, and Robert M. Kelly*

Department of Chemical and Biomolecular Engineering, North Carolina State University, Raleigh, North Carolina 27695-7905

Received 4 April 2006/Accepted 5 July 2006

In the genome of the hyperthermophilic bacterium Thermotoga maritima, TM0504 encodes a putative signaling peptide implicated in population density-dependent exopolysaccharide formation. Although not noted in the original genome annotation, TM0504 was found to colocate, on the opposite strand, with the gene encoding srrA, a hybrid of tRNA and mRNA (tmRNA), which is involved in a trans-translation process related to ribosome rescue and is ubiquitous in bacteria. Specific DNA probes were designed and used in real-time PCR assays to follow the separate transcriptional responses of the collocated open reading frames (ORFs) during transition from exponential to stationary phase, chloramphenicol challenge, and syntrophic coculture with Methanococcus jannaschii. TM0504 transcription did not vary under normal growth conditions. Transcription of the tmRNA gene, however, was significantly up-regulated during chloramphenicol challenge and in T. maritima bound in exopolysaccharide aggregates during methanogenic coculture. The significance of the colocation of ORFs encoding a putative signaling peptide and tmRNA in T. maritima is intriguing, since this overlapping arrangement (tmRNA associated with putative small ORFs) was found to be conserved in at least 181 bacterial genomes sequenced to date. Whether peptides related to TM0504 in other bacteria play a role in quorum sensing is not yet known, but their ubiquitous colocalization with respect to tmRNA merits further examination.

Although canonical models for bacterial signaling mechanisms have been established (4, 9), chemically diverse molecules that act as mediators for cell-to-cell communication continue to be discovered (42), concomitant with the expanding range of microorganisms found to be involved in inter- and intraspecies communication (18, 20, 36, 42). Recently, signaling has been connected with biologically extreme environments. For example, N-acetyl homoserine lactones have been reported in cultures of the haloalkaliphilic archaeon Natronococcus occultus (28). Furthermore, cell density-dependent, peptide-based signaling, tied to formation of exopolysaccharide-based cell aggregates, was identified in the hyperthermophilic bacterium Thermotoga maritima in syntrophic coculture with the hyperthermophilic archaean Methanococcus jannaschii (10, 11). Presumably, the formation of such aggregates allows M. jannaschii to make methane using the H2 produced by T. maritima cells as an auto-inhibitory by-product of carbohydrate metabolism (22). During syntrophic coculture of these two organisms, examination of genome-wide transcriptional profiles indicated increased transcript levels of a small hypothetical open reading frame (ORF), annotated as TM0504, in the T. maritima genome (25), in conjunction with up-regulation of adjacent ORFs annotated as putative ABC transporter permease and ATPase subunits (11). Subsequent reexamination of the genomic neighborhood of the TM0504 gene revealed that it was colocalized on the complementary strand within the srrA gene encoding a highly structured RNA having both tRNA and mRNA (tmRNA) domains and involved in “trans-translation” during ribosomal rescue (13, 44). The first domain in the tmRNA molecule consists of a region that mimics a tRNAAla, including an acceptor stem, a T-stem-loop, and a 3’ terminalCCA that can be aminoa cetlated by an alanyl tRNA synthase (15, 39, 41). The second domain consists of an internal ORF used for tagging nascent polypeptides. The tmRNA-like portion of tmRNA, charged with alanine, acts as a surrogate tRNA to accept the nascent polypeptide. The ribosome then switches templates from the stalled mRNA to the small ORF that is used as a peptide coding sequence to be incorporated into the nascent “stalled” peptide. The presence of a stop codon in the ORF included in tmRNA ensures translation termination and, hence, recycling of the ribosome (13, 44).

Colocation of TM0504 and tmRNA in the T. maritima genome may not be coincidental. Examination of more than 200 bacterial genomes completed to date revealed that genes encoding small putative peptides are commonly colocalated on the opposite strand to the tmRNA gene. Since the double-stranded probe on the cDNA microarray used previously (11) to detect changes in transcript levels of TM0504 would not have differentiated between the peptide gene and tmRNA transcription, a strategy was developed to separately track changes in transcript levels of each of these two colocalated ORFs in T. maritima for several growth conditions. These results were examined with respect to the functional roles of these molecules in T. maritima and other bacteria.

* Corresponding author. Mailing address: Department of Chemical and Biomolecular Engineering, North Carolina State University, Raleigh, NC 27695-7905. Phone: (919) 515-6396. Fax: (919) 515-3465. E-mail: rmkelly@eos.ncsu.edu.
† Supplemental material for this article may be found at http://jb.asm.org/.
Materials and Methods

Cultivation protocols. T. maritima MSB8 was grown anaerobically at 80°C in batch-pure and methanogenic coculture (11) and in continuous pure culture, as described previously (31, 35). For antibiotic challenge experiments, a chemostat culture was exposed to 100 mg/ml of chloramphenicol; the thermostability of this antibiotic at 80°C for periods of time relevant to this experiment was previously found to be negligible (C. I. Montero, unpublished data). For RNA sampling, 350 ml of each culture was harvested into centrifuge bottles and cooled rapidly to 0°C by immersion in ice water, followed by immediate centrifugation of the cells at 10,000 × g for 15 min at 4°C. RNA extraction was then performed following established protocols (7). RNA/cDNA obtained previously during synchronous coculture of T. maritima with M. jannaschii (11) was reexamined.

Real-time PCR protocols. For real-time-PCR analysis, a gene-specific strategy was used involving a strand tag for proper identification of transcripts from the overlapping tmRNA and TM0504 genes. All RNA samples used were pretreated with DNase I (amplification grade) (Invitrogen), following manufacturer’s recommendations. The concentration and purity of RNA samples were monitored using denaturing agarose-formaldehyde electrophoresis gels and standard spectrophotometric measurements. For cDNA generation and real-time PCR, a two-step, gene-specific strategy was adopted. Superscript III (Invitrogen) was used for the generation of cDNA from 2 μg of RNA sample, following the manufacturer’s protocols. The generation of cDNA was performed using a single 40-mer antisense primer per gene. The polyacrylamide gel electrophoresis-purified antisense primer was then designed to be complementary to the cDNA generated and paired with the terminal CCA sequence. This CCA sequence in tmRNA is associated with the aminosylation of the molecule. (B) Genomic neighborhood of TM0504 locus in T. maritima.

Results and Discussion

Relative transcript levels of tmRNA and TM0504 ORFs in T. maritima. The fact that a chemically synthesized, truncated version of the peptide encoded in TM0504 triggered exopolysaccharide production in T. maritima suggested that this peptide plays a role in the molecular ecology of this bacterium (11). Following initial work focusing on the increase in transcript levels corresponding to the TM0504 ORF, it became apparent that TM0504 was located in reverse orientation and complementary to the tmRNA gene in the T. maritima genome, as shown in Fig. 1 (25). The unusual organization of these two genes included portions of the coding and template strands of both the tmRNA and TM0504. The fact that a chemically synthesized, truncated version of the peptide encoded in TM0504 triggered exopolysaccharide production in T. maritima suggested that this peptide plays a role in the molecular ecology of this bacterium (11). Following initial work focusing on the increase in transcript levels corresponding to the TM0504 ORF, it became apparent that TM0504 was located in reverse orientation and complementary to the tmRNA gene in the T. maritima genome, as shown in Fig. 1 (25). The unusual organization of these two genes included portions of the coding and template strands of both the peptide and tmRNA. Therefore, transcripts for either tmRNA or TM0504 could have hybridized to this probe, such that the transcriptional response to specific conditions would have been confounded. In order to resolve the specific transcription of these genes, a strategy was developed in which real-time PCR was adapted to independently estimate the transcript levels of the overlapping genes. This approach was then used to interrogate RNA samples taken for T. maritima subjected to chlor-
amphenicol challenge at 100 µg/ml in chemostat culture, and as well as to reinterrogate samples from previous experiments with syntrophic coculture with *M. jannaschii* (10, 11) (Table 1).

Translational inhibiting antibiotics, such as kanamycin, gentamicin (1), and chloramphenicol (38), cause read-through of stop codons and increased frequency of ribosome stalling (38). These effects presumably trigger a tmRNA response. Upon chloramphenicol challenge (100 µg/ml), *T. maritima* growing in continuous culture demonstrated significant up-regulation of the tmRNA gene (40-fold at 5 min and 23-fold at 30 min after dosing versus before challenge), presumably reflecting the deleterious effect of the antibiotic on translational processes (Table 1). A relatively modest change (twofold) was noted for TM0504. Previously, TM0504 was identified by cDNA microarray analysis as among the most highly up-regulated genes (13-fold) in *T. maritima* growing in stationary phase (38).

These effects presumably trigger a tmRNA response. Upon chloramphenicol challenge (100 µg/ml), *T. maritima* growing in continuous culture demonstrated significant up-regulation of the tmRNA gene (40-fold at 5 min and 23-fold at 30 min after dosing versus before challenge), presumably reflecting the deleterious effect of the antibiotic on translational processes (Table 1). A relatively modest change (twofold) was noted for TM0504. Previously, TM0504 was identified by cDNA microarray analysis as among the most highly up-regulated genes (13-fold) in *T. maritima* growing in stationary phase (38).

The significant down-regulation of TM0504 in the coculture aggregate compared to pure culture could relate to the fact that cells no longer participate in population-based behaviors (i.e., exopolysaccharide formation) once they are constrained to aggregates. Under normal growth conditions in either pure or coculture, no significant change in either TM0504 or tmRNA transcription was noted for transition from mid-exponential to early stationary phase (data not shown).

**Identification of small ORFs encoded within the tmRNA gene locus in bacteria.** As shown in Fig. 1, the TM0504 ORF encodes a peptide situated in the complementary strand of tmRNA. Whether this arrangement is common to other tmRNA genes was determined for 200 representative bacterial species with completely sequenced genomes. Given that tmRNA is a highly conserved molecule with a relatively high GC content, a lower frequency of start and stop codons was expected within the tmRNA ORF. However, small ORFs could be identified on the opposite strand of the tmRNA molecule in all bacterial genomes examined. Therefore, analysis focused on ORFs encoded in the 3′ end “antisense” strand of the tmRNA gene (155 genomes) and on ORFs located on the 5′ end on the complementary strand of tmRNAs that had a circular gene permutation, commonly called two-piece tmRNAs (3, 34) (26 genomes). From this group, 127 had a conformation similar to that observed in TM0504 in which the ORF starts exactly two bases after the 3′ end encoding putative peptides of 20 amino acids or more. It is noteworthy that these putative peptides have not been annotated in the majority of the 200 bacterial genome sequences examined. In fact, in only three other cases, in addition to *T. maritima*, the predicted

<table>
<thead>
<tr>
<th>Culture and assay conditions for:</th>
<th>TM0504</th>
<th>tmRNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Culture before CHL challenge</td>
<td>29.1 ± 0.4</td>
<td>28.4 ± 0.3</td>
</tr>
<tr>
<td>Culture 5 min after CHL challenge</td>
<td>28.5 ± 1.0</td>
<td>22.9 ± 0.1</td>
</tr>
<tr>
<td>Culture 30 min after CHL challenge</td>
<td>28.0 ± 0.9</td>
<td>23.8 ± 0.2</td>
</tr>
<tr>
<td>Culture 5 min after CHL challenge vs. before</td>
<td>1.0</td>
<td>40.3</td>
</tr>
<tr>
<td>Culture 30 min after CHL challenge vs. before</td>
<td>2.0</td>
<td>22.8</td>
</tr>
</tbody>
</table>

**Coculture effect**

- Pure culture (mid-log phase): 30.3 ± 0.4
- Coculture (mid-log phase): 35.5 ± 0.1
- Coculture vs. pure culture: −30.6

**Value for ORF**

- *C*<sub>T</sub>, cycle threshold.

* TABLE 1. Relative transcript levels of ORFs encoding TM0504 peptide and tmRNA

* Chemostat culture. CHL, chloramphenicol.
* Serum bottles; no sparging.

---

**FIG. 2.** Size distribution among putative tmRNA colocated peptides in 200 selected bacterial genome sequences. tmRNA sequences were retrieved from publicly available databases. Vector NTI was used to analyze the strand complementary to tmRNA to identify ORFs at the 3′ end encoding putative peptides of 20 amino acids or more.
peptide annotated as a hypothetical protein: LP2510 in Lactobacillus plantarum WCFS1 (14), TC0287 in Chlamydia muridarum Nigg (32), and CpB0109 in Chlamydia pneumoniae TW-183 (unpublished). This may relate to default sizes used for small ORF identification during genome annotation. It is also of note that tmRNA itself has not been annotated in most microbial genome sequences determined to date.

Analysis of tmRNA gene neighborhoods in bacterial genomes. In the case of T. maritima, TM0504 is found upstream of ORFs encoding an ABC transporter lacking a substrate-binding protein, suggesting a possible export pathway for the signaling peptide (11). Figure 1B shows the gene neighborhood containing TM0504 and tmRNA. TM0503 and TM0502 are annotated as permease subunits of a putative oligopeptide transporter system, while TM0500 and TM0501 are putative ABC transporter ATPases. Unlike other ABC transporter operons of the oligopeptide/dipeptide family found in T. maritima, which include associated substrate binding proteins that in many cases have recently been shown to bind various sugars (24), this transporter showed little response to changes in medium carbohydrate composition (5).

Genes adjacent to tmRNA in T. maritima and selected bacteria were examined in order to elucidate possible functions that might be associated with putative peptides. Not unexpectedly, smpB genes, encoding small binding proteins involved in tmRNA function, were found within the tmRNA genomic neighborhood (±5 ORFs from the tmRNA/peptide locus) in 63 genomes (Fig. 3). However, given the size of our sample, the notion of a higher frequency of gene rearrangements and gene integration in the vicinity of the tmRNA gene in Enterobacteriaceae seems to prevail in other bacterial groups (43). Transport and transmembrane proteins were also found in higher than expected proportions, based upon analysis of the average proportion of functional categories within genomes available at The Institute for Genomic Research through the Comprehensive Microbial Resource (29). As a group, hypothetical proteins were found in higher than expected numbers; it remains to be seen if these share common functional properties.

Colocalization of tmRNA with small peptides. The possibility that a biologically active peptide is encoded on and independently transcribed from the strand opposite tmRNA in most bacterial genomes is intriguing. In T. maritima, tmRNA and TM0504 appear to be transcribed independently; within this locus apparently exists the capability for both quorum sensing and ribosome rescue. While this phenomenon has yet to be described in other microorganisms, a previous study utilizing Northern analysis reported transcription of an ORF corresponding to a putative small peptide colocated with the tmRNA in Bradyrhizobium japonicum (6). This was believed to result from contamination of the RNA sample, given that the bacteria were cultivated with plant tissue. Efforts to locate the translated peptide in B. japonicum were unsuccessful. In Salmonella enterica serovar Typhimurium, insertion mutagenesis of the 3’ end of the tmRNA molecule has been implicated

FIG. 3. Functional assignment of the genes adjacent to tmRNA in representative bacterial genomes. Functional category assignments for five genes upstream and five genes downstream of the tmRNA were tabulated for 200 bacterial genomes. Functional categories were designated as follows: A, Smgb; B, cellular processes/cell division; C, DNA metabolism; D, protein synthesis; E, amino acids and nucleotide synthesis; F, hypothetical proteins; G, cellular processes; H, two-component systems and regulatory proteins; I, protein fate; J, transporters and transmembrane proteins; K, energy metabolism and cell envelope; L, Small peptides with orientations similar to TM0504; M, other functions, including stress and intermediate metabolism.

FIG. 4. Alignment of selected TM0504-like peptides. Alignments were performed using the Align X/ClustalW tool of Vector NTI. As shown in the alignment, the amino terminal end of the putative peptides is conserved among phylogenetically diverse groups of bacteria (for a complete listing of all bacterial genomes analyzed, see Tables S1 and S2 in the supplemental material). T. neapolitana, Thermotoga neapolitana; N. gonorrhoeae, Neisseria gonorrhoeae; P. syringae, Pseudomonas syringae; M. tuberculosis, Mycobacterium tuberculosis; S. coelicolor, Streptomyces coelicolor; T. thermophiles, Thermus thermophilus; C. tepidum, Chlororibium tepidum; B. subtilis, Bacillus subtilis; O. iheyensis, Oceanobacillus iheyensis; L. acidophilus, Lactobacillus acidophilus; L. lactis, Lactococcus lactis; S. pneumoniae, Streptococcus pneumoniae.
in a reduction of virulence and inhibition of expression of genes including a toxin-related (cypA) gene. The interpretation of these results is complicated by the insertion of a prophage in the ssrA region. This observation was attributed to a regulatory role of tmRNA in which alteration of DNA interactions or translation associated with this gene directly affects virulence (12). An alignment of TM0504 with tmRNA-colocated putative peptides encoded in selected bacterial genomes is shown in Fig. 4, and a comprehensive listing of all tmRNA-colocated putative peptides in bacterial genomes examined here is provided in Tables S1 and S2 in the supplemental material. It is interesting that beyond the highly conserved regions overlapping the 3’ end of the tmRNA, there is little, if any, sequence similarity within the peptide coding regions of distantly related species. If, in fact, these peptides are implicated in ecologically important functions akin to what was observed for T. maritima, perhaps their unique sequences may trigger responses that are highly specific to particular species.

Conclusions. The colocation and apparently independent regulation of an unusual locus of overlapping genes in the tmRNA gene neighborhood in T. maritima are intriguing, given that similar constructs occur in most bacterial genomes sequenced to date. For T. maritima, sterically limitations likely restrict both transcripts from being read simultaneously within the 400-bp region, although this has not yet been examined. Examples of overlapping genes and, more precisely, the independent regulation of those genes in bacteria are relatively rare but have been reported. The pic/set locus in Shigella flexneri and Escherichia coli, corresponding to a mucinase and enterotoxin arrangement, is one such example of differentially transcribed overlapping genes (2). The repEA and repEB genes are another such example and are differentially transcribed during T4 bacteriophage DNA replication (40). It has been suggested that overlapping genes may compress genome size in species subject to reductive evolution (33). T. maritima is a free-living bacterium that would seem to have frequent opportunity to expand its genome through exchange of genetic elements with colocated microorganisms, including archaeal hyperthermophiles (21, 25–27). This feature of overlapping genes at the tmRNA locus might, in fact, have primitive origins; a similar arrangement occurs in the hyperthermophilic bacterium Aquifex aeolicus (unpublished observation). It is also possible that this feature has been conserved across several bacterial phyla as a possible means for protection against mutation events (16, 33). Here, the constraints associated with a functional t-alanyl-acceptor region implicate a locus highly restrictive for modifications. Based on the alignments provided at the tmRNA database, this seems to be the case (45). More work is needed to determine the significance of this genetic colocation, including testing for evidence that specific microbial behaviors arise through peptide-induced regulation.

ACKNOWLEDGMENTS

This work was supported in part by grants from the NASA Exobiology, Energy Biosciences of the U.S. Department of Energy, and NSF Biotechnology Programs.

The authors also thank Fred Fuller from the Veterinary Medicine School at North Carolina State University for his suggestions for the design of the real-time PCR primers.

REFERENCES


