2004

A third lineage with two-piece tmRNA

S. M. Sharkady

K. P. Williams

Nucleic Acids Research

Nucleic Acids Res. 2004; 32(15): 4531–4538. PMCID: PMC516066

Published online 2004 Aug 23. PMID: 15326226

doi: 10.1093/nar/gkh795

A third lineage with two-piece tmRNA

Stephen M. Sharkady and Kelly P. Williams*

Author information Article notes Copyright and License information Disclaimer

Associated Data

Supplementary Materials

ABSTRACT

tmRNA combines tRNA and mRNA properties and helps bacteria to cope with stalled ribosomes. Its termini normally pair in the tRNA domain, closing the mRNA portion into a looping domain. A striking variation is a two-piece form that effectively breaks open the mRNA domain loop, resulting from independent gene permutation events in alphaproteobacteria and cyanobacteria. Convergent evolution to a similar form in separate bacterial lineages suggests that loop-opening benefits tmRNA function. This argument is strengthened by the discovery of a third bacterial lineage with a loop-opened two-piece tmRNA. Whereas most betaproteobacteria have one-piece tmRNA, a permuted tmRNA gene was found for *Dechloromonas aromatica* and close relatives.

Correspondingly, two tmRNA pieces were identified, at approximately equal abundance and at a level one-fifteenth that of ribosomes, a 189 nt mRNA piece and a 65 nt aminoacylatable piece. Together these pieces were active with purified *Escherichia coli* translational components, but not alone. The proposed secondary structure combines common tmRNA features differently from the structures of other two-piece forms. The origin of the gene is unclear; horizontal transfer may be indicated by the similarity of the tRNA domain to that from a cyanobacterial two-piece tmRNA, but such transfer would not appear simple since the mRNA domain is most similar to that of other betaproteobacteria.

INTRODUCTION

Ribosomes are expected to stall at the 3′ end of defective mRNAs of a class termed nonstop mRNAs that lack an in-frame stop codon. Potential problems from stalling include sequestration of ribosomes and production of incomplete polypeptides. In bacteria these problems are solved by tmRNA, a specialized RNA with both tRNA-like and mRNA-like properties (1). The stalled ribosome can switch from the nonstop mRNA and resume translation on the reading frame in tmRNA; the ribosome is then freed by release at the tmRNA stop codon. A second result of tmRNA translation is the addition of a hydrophobic peptide tag onto the C-terminus of the polypeptide product of nonstop mRNA; this tag is a signal directing proteolysis of the entire tagged polypeptide.

Most tmRNAs are composed of a single RNA chain, whose termini base-pair, forming the tRNA domain and moreover sealing the mRNA region into a looping domain (2,3). Electron microscopy shows that this mRNA loop has a diameter approximately equal to that of the neck of the small ribosomal subunit (4). Translation of a small looped RNA raises topological issues, such as its mechanism of threading into the mRNA track through the decoding center and the continued path of the loop on the ribosome. It might be expected that strain would develop in this loop as it rachets around an unusual path during translation. The loop is studded with four pseudoknots in most bacteria that could open to help relieve such strain. A second possible solution to the strain problem is to permanently open the loop, as in the two-piece tmRNAs. Independent events of circular gene permutation have occurred in the alphaproteobacteria and cyanobacteria that are responsible for the effective break in the mRNA loop found in their two-piece tmRNAs (5,6). The establishment, twice, independently, of such complicated gene rearrangements implies that the two-piece form provides benefits to tmRNA function relative to the one-piece

form, perhaps solving topological problems as suggested above. Here, we extend this argument by showing that yet a third bacterial lineage has a permuted gene that produces two-piece tmRNA.

MATERIALS AND METHODS

Bacterial nucleic acids

Dechloromonas aromatica RCB, kindly supplied by John Coates and Romy Chakraborty (University of California, Berkeley, CA), was grown aerobically to an OD_{600} of 0.1 by shaking at 30°C in R2A medium (Gibco) supplemented with 20 mM lactate. The culture was split for the preparation of nucleic acids. Total cellular RNA (125 μg) was prepared from 100 ml culture using Tri-Reagent (MRC). Genomic DNA was prepared from 5 ml culture using a Puregene DNA isolation kit (Gentra). A 417 bp PCR product from the tmRNA gene was amplified from genomic DNA using primers 5′-ACTGATATCTGCTGCCAGT and 5′-CAATCGAGGCATCGGTCAT, sequenced directly and cloned into pCR2.1 Topo vector (Stratagene) with resequencing.

tmRNA sequences from *Dechloromonas agitata* and *Azospira oryzae* (originally called *Dechlorosoma suillum*) were obtained directly from purified PCR products that were amplified using primers 5′-GTAATACGACTCACTATAGGTTGTAATTCCGGCAATCT and 5′-TGACGACGGGGACCGAAGTCCCC, from genomic DNAs that were kind gifts from Laurie Achenbach (Southern Illinois University, Carbondale, IL).

In vitro transcripts

RNA standards corresponding to the acceptor and coding pieces of tmRNA from *D.aromatica* were transcribed with T7 RNA polymerase using PCR templates generated from the above clone. The PCR templates for coding piece standards were produced with primer 5′-GTAATACGACTCACTATAGGGGGTGTACTGGTCTCGA and either 5′-CGACGGGGACCGAAGTCCCC (186 nt RNA), 5-TGACGACGGGGACCGAAGTCCCC (189 nt RNA) or 5′-TGGACGTGACGACGGGGA (195 nt RNA). The PCR template for the acceptor piece standard was produced using primers 5′-GTAATACGACTCACTATAGGTTGTAATTCCGGCAATCT and 5′-TGG TGGAGGTGAGGGGAAT. A PCR template for the transcription of a standard for *D.aromatica* 5S RNA was created with primers 5′-

GTAATACGACTCACTATAGGCGTCTGGCGTCAATAGCCT GCTGG and 5′-GGAGTCTGGCGTTGACCTACTTTCGCGA using genomic DNA as a template.

For tmRNA translation experiments, the 189 nt coding piece, an acceptor piece differing from the above by the presence of one additional guanine nucleotide at the 5′ end and *Escherichia coli* tmRNA were transcribed. RNA transcripts were eluted from polyacrylamide–urea gel bands, ethanol precipitated and quantified spectrophotometrically.

Northern-blot analysis

An aliquot of 10 µg of total *D.aromatica* RNA or various quantities of mixed unlabeled standards were separated on polyacrylamide–urea gels and electrotransferred onto a nylon membrane. The blot was pre-hybridized 1 h at 50°C in QuikHyb solution (Stratagene), hybridized by incubating 1 h at 50°C with 5′-[^{32}P]-AACGTCGATCGCCTGACGAGTTGCCT (coding RNA probe), washed four times in 2× SSC and visualized by phosphorimaging. Successive hybridizations to 5′-[^{32}P]-CCCTGTCCAGAACGCCTCCAGATTGC (acceptor RNA probe) and 5′-[^{32}P]-GTTGACCTACTTTCGCGAGC (5S RNA probe) at 10-fold lower specific radioactivity than the other two probes were optional.

Primer extension analysis

The primer 5′-[^{32}P]-TGGTGGAGGTGAGGGGAAT (10 pmol) was annealed to 10 µg of total *D.aromatica* RNA by heating in 4.5 µl of 50 mM Tris–HCl (pH 8.3 at 42°C), 25 mM KCl for 2 min at 90°C then cooling to 30°C over 20 min. The hybridized primer was extended in 7.5 µl of 50 mM Tris–HCl (pH 8.3 at 42°C), 50 mM KCl, 10 mM dithiothreitol (DTT), 10 mM $MgCl_2$ and 200 µM of each dNTP with 1 U avian myeloblastosis virus reverse transcriptase (Seikagaku) for 1 h at 42°C. Reactions were terminated by the addition of 6 µl of 85% formamide, 20 mM EDTA, 0.1% bromophenol blue and 0.1% xylene cyanol. Dideoxy DNA sequencing employed the same primer, the cloned *D.aromatica* tmRNA gene as template and a kit from NEB. Reaction products were heated for 2 min at 80°C and resolved on a 10% polyacrylamide/8 M urea/TBE sequencing gel, which was dried and visualized by phosphorimaging.

In vitro tmRNA translation

Tightly coupled ribosomes were prepared (7) from *E.coli* KW1063,

the product of transducing the *ssrA::cat* allele into the strain CAN20-12E (8), deficient in the RNases I, II, D and BN. All additional factors were purified as six-histidine fusions of *E.coli* proteins (9). Reactions were carried out in 50 µl translation buffer [95 mM K-glutamate, 9 mM magnesium acetate, 8 mM putrescine, 5 mM NH_4Cl, 5 mM KPO_4 (pH 7.3), 1 mM spermidine, 1 mM DTT and 0.5 mM $CaCl_2$] containing 50 pmol ribosomes, 1 µg IF1, 2 µg IF2, 0.75 µg IF3, 50 pmol EF-G, 100 pmol EF-Tu, 50 pmol EF-Ts, 0.5 µg RF1, 0.5 µg RF3, 0.5 µg RRF, 0.2 µg creatine kinase (Sigma), 0.15 µg myokinase (Sigma), 0.054 µg nucleotide diphosphate kinase (Sigma), 3.5 µg AlaRS, 4.2 µg PheRS, 1 µg AsnRS, 0.4 µg AspRS, 0.65 µg GluRS, 0.02 µg TyrRS, 0.2 µg LeuRS, 0.2 µg GlnRS, 2 µg IleRS, 2.8 A_{260} units total *E.coli* tRNA (Roche), 1 mM ATP, 1 mM GTP, 0.1 mM [^{14}C]alanine (Amersham Pharmaciabiotech), and 0.1 mM each of phenylalanine, asparagine, aspartate, tyrosine, leucine, glutamine and isoleucine (9). Reactions additionally contained *E.coli* SmpB, and *in vitro* transcripts of *E.coli* tmRNA or the coding or acceptor pieces of *D.aromatica* tmRNA as specified. The tmRNAs were folded in translation buffer by heating for 2 min at 80°C and cooling to 20°C over 10 min. Reaction mixtures were incubated for 5 min at 37°C prior to initiation by the addition of 50 µg poly(U), incubated for 60 min at 37°C, terminated with 2 ml of 5% trichloroacetic acid, heated for 20 min at 90°C, incubated for 30 min on ice and filtered onto GF/C filters (Whatman) presoaked in 5% TCA. Filters were washed with 6 ml of 5% TCA then 2 ml ethanol, dried and counted in 7 ml BioSafe II scintillation cocktail (RPI).

Phylogenetic analysis

Sequences were aligned for the mature tRNA domain (including the first 6 bp of helix P2, and repermuting permuted sequences) and tag reading frame (including 5 bp upstream of the resume codon) for all 246 known tmRNA sequences for which both sequence portions are complete and unique. The two alignments contained, respectively, 42 and 85 parsimony-informative positions, 40 and 31 of which were present in more than half of the sequences. A third alignment was a concatenation of the two above, but only for the 26 cyanobacterial and plastid sequences, with *E.coli* as an outgroup. The alignments were analyzed with PAUP* 4.0 b8, taking maximum-likelihood pairwise distances with the HKY85 substitution model (Table S.1), which were used to generate trees by either BIONJ neighbor-joining or UPGMA methods. Alignments were also analyzed using maximum parsimony as the optimality criterion. Heuristic searches using tree-bisection-

resection were performed with over 700 random sequence additions, abandoning when islands reached 3000 trees. The search attaining the tree of the shortest length (893 for the tRNA domain, 1639 for the reading frame and 413 for the cyanobacteria) was repeated, collecting up to 50 000 trees from which the 97% majority-rule cladogram was taken.

RESULTS

Permuted betaproteobacterial tmRNA gene sequences

D.aromatica RCB, best known for its ability to degrade benzene or perchlorate anaerobically, is a betaproteobacterium of the family Rhodocyclaceae (10). A BLAST search of genomic sequence data (www.jgi.doe.gov) identified only one candidate gene sequence for the *D.aromatica* tmRNA, based on the hallmark tRNA-like structure and peptide tag reading frame that it encodes (Figure 1). The sequence of this region was confirmed for a PCR product from genomic DNA. A likely strong promoter was identified, with −10 and −35 hexamers each matching five *E.coli* consensus bases and with the optimal 17 bp spacing. Downstream 303 bp from the predicted start site of transcription lies the stop codon for the convergently transcribed *tonB* gene, a probable limit to the extent of the tmRNA gene. The striking aspect of the *D.aromatica* gene is that it is circularly permuted relative to the standard gene form; that is, the portion encoding the aminoacylated 3′ tail of the tmRNA is positioned upstream of the portion encoding the mRNA domain.

Figure 1

Sequences of permuted betaproteobacterial tmRNA genes. Sequence data for *D.aromatica* (Dar) from DOE Joint Genome Institute

(www.jgi.doe.org) were confirmed in a PCR product from genomic DNA, and aligned with partial sequences for *D.agitata* (Dag) and *A.oryzae* (Aor). Uppercase sequence: segments coding for the mature tmRNA pieces (acceptor and coding) as identified here. Arrows: natural processing sites and some endpoints of RNA standards used in our experiments. Color-coding: proposed secondary structural features of RNA. Asterisks: Watson–Crick base-pair covariations among Dar, Dag and Aor. Corresponding segments from *G.violaceus* (Gvi) and *Azoarcus* sp. BH72 (Asp) genes are also aligned; mismatches between mature sequence of Dag and Gvi and between Aor and Asp are underlined.

All other available betaproteobacterial tmRNA sequences, from 13 genera including *Azoarcus* of the Rhodocyclaceae (B.Reinhold-Hurek, personal communication), had the standard gene form (11,12). In search of additional examples of the permuted gene, we examined two additional members of the Rhodocyclaceae, *D.agitata* and *A.oryzae*, that are closer relatives to *D.aromatica* (13,14). For both, a nearly complete gene sequence was amplified from genomic DNA, revealing a permuted tmRNA gene closely related to that of *D.aromatica* (Figure 1). Grossly similar permuted genes produce two-piece tmRNA in alphaproteobacteria and a subgroup of cyanobacteria (5,15).

D.aromatica tmRNA is in two pieces

For all tmRNAs, two RNA termini are determined by tRNA-like processing of a longer precursor, leaving the end sequences 5′p-GGGNNNN… and …N′N′N′N′UCCACCA-3′OH, where the underlined sequences base-pair to form the alanylatable acceptor stem and the CCA tail may be uncoded since it can be added post-transcriptionally by tRNA nucleotidyltransferase (16). These two processing sites lie at the extremities of the precursor for the standard one-piece tmRNA. With permuted gene structures, these sites are internal to the precursor, separated by an intervening segment that is removed by the processing (5). This leaves the mature tmRNA in two separate RNA chains, one containing the CCA tail (acceptor piece) and the other containing the tag reading frame (coding piece), although these chains remain noncovalently attached by base-pairing. Thus two of the four ends of two-piece tmRNAs are completely predictable. In contrast, the processes that produce the two remaining ends cannot necessarily be predicted and these ends must therefore be mapped.

We hypothesized that the *D.aromatica* mature tmRNA would be found in two pieces according to the above principles and tested this by northern blotting. Moreover, to make the blot quantitative and furnish especially appropriate size markers, we synthesized *in vitro* RNA standards of the same sequence and approximate length as our expectations for the two putative mature tmRNA pieces. We expected that the acceptor piece would be no longer than 65 nt, the distance from the putative transcription start site to the added CCA tail and synthesized a 63 nt RNA standard of the same sequence as this hypothetical maximum, but 2 nt shorter at the 5′ end. We suspected that a 10 bp stem–loop encoded at the other end of the gene would be part of the mature RNA and therefore that the 3′ end of the coding piece would map within the 19 bp segment between the stem–loop and the *tonB* stop codon. Accordingly two RNA standards were prepared, extending the putative coding piece to 186 or 195 nt. A fourth RNA standard mimicking the *D.aromatica* 5S ribosomal RNA was also prepared, for comparing the abundances of tmRNA and ribosomes.

Total cellular RNA was isolated from an aerobically grown *D.aromatica* culture and subjected to northern blotting, running the RNA standards described above as standard curves in outside lanes (Figure 2A). The blot was hybridized with probes for the *D.aromatica* coding and acceptor tmRNA pieces and 5S rRNA. The coding piece probe illuminated one band migrating between the 186 and 195 nt standards. The acceptor piece probe detected a band migrating just above the 63 nt standard. Separate hybridization with the individual probes demonstrated their specificity. A larger precursor tmRNA was not detected. The abundance of RNAs in 10 μg total cellular RNA was measured from the standard curves at 0.43 pmol coding RNA, 0.63 pmol acceptor RNA and 7.8 pmol 5S rRNA. These values yield a molar ratio of coding:acceptor pieces in *D.aromatica* of 0.68:1 and, in line with similar measurements for one-piece tmRNAs (5), a molar ratio of ribosomes:tmRNA (using the average for the two pieces) of 15:1.

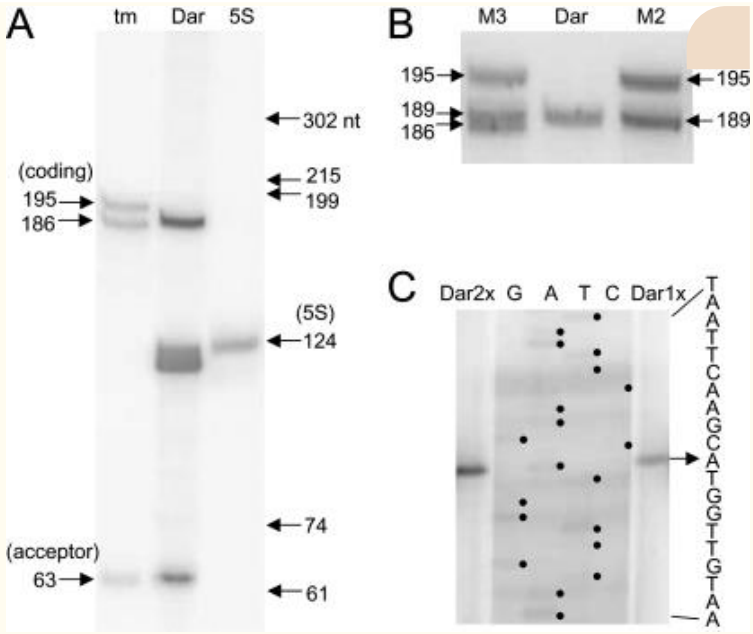

Figure 2

Primary structure of *D.aromatica* tmRNA. (**A**) Two pieces. A gel blot of 10 μg *D.aromatica* total RNA (Dar), 0.1 pmol each tmRNA standard (tm) or 1 pmol *D.aromatica* 5S rRNA standard (5S) was hybridized with probes specific for the 5S rRNA and coding and acceptor tmRNA pieces of *D.aromatica*, with separate use of the probes in other blots showed specificity for the middle, upper and lower bands, respectively. Additional lanes not shown continued upward the standard curves and included unrelated pre-radiolabeled RNA size markers whose positions are shown with arrows. (**B**) Size of the coding piece. Northern blot of 10 μg total *D.aromatica* RNA (Dar) or coding RNA standards of 189 and 195 nt (M2) or additionally the 186 nt standard (M3). (**C**) 5′ end of the acceptor piece. Primer extension analysis on 10 μg total *D.aromatica* RNA from *Dechloromonas* to determine the mature 5′ end of the acceptor RNA. Leftmost lane, two-fifths of reaction; rightmost lane, one-fifth of reaction. Contrast has been enhanced in place for the central lanes, containing the dideoxy sequencing products of the cloned *D.aromatica* tmRNA gene with the same primer.

Finer mapping of the tmRNA pieces

The size of the coding RNA was estimated at 189 nt from the preceding experiment. A new RNA standard corresponding precisely to a 189 nt coding RNA was prepared and the *D.aromatica* total RNA was subjected to extended electrophoresis adjacent to the three standards, with northern blotting (Figure 2B).

The coding RNA co-migrated with the 189 nt standard, a measurement that we estimate has ±1 nt accuracy considering the triphosphate 5′ end of the standard. With a priori knowledge of its 5′ end, this maps the 3′ end of the coding piece to a position downstream of the terminal stem–loop (Figure 1).

The experiment of Figure 2A mapped the 5′ end of the acceptor RNA to the vicinity of the predicted transcription start site, but slight truncation by post-transcriptional processing remained a possibility. Primer extension was used for better precision (Figure 2C). This mapped the end to the predicted transcription start site, a result compatible with activity of the promoter at the predicted transcription start site and no subsequent 5′-processing.

Proposed secondary structure for *D.aromatica* tmRNA

Phylogenetic analysis has been used to determine the secondary structure of standard one-piece tmRNA and both the alphaproteobacterial and cyanobacterial two-piece tmRNAs (5,15,17). The *D.aromatica* tmRNA sequence can be fitted into a secondary structure model whose features agree well with those proven for other tmRNAs (Figure 3). In this model, the tRNA domain and long protruding P2 pairing are as found in all tmRNAs. A long stem–loop within the intervening segment RNA would appear to promote folding of the tRNA domain in the precursor (Figure 1). At two positions, a pseudoknot abuts a helix, allowing mutual stabilization by coaxial stacking, as found at the corresponding positions in most one-piece tmRNAs. The first pseudoknot abuts P2. The second pseudoknot abuts helices at both ends, the P5 stem–loop (at the end of the tag reading frame) at one end and the 10 bp tetraloop-capped terminal stem–loop P8 at the other end.

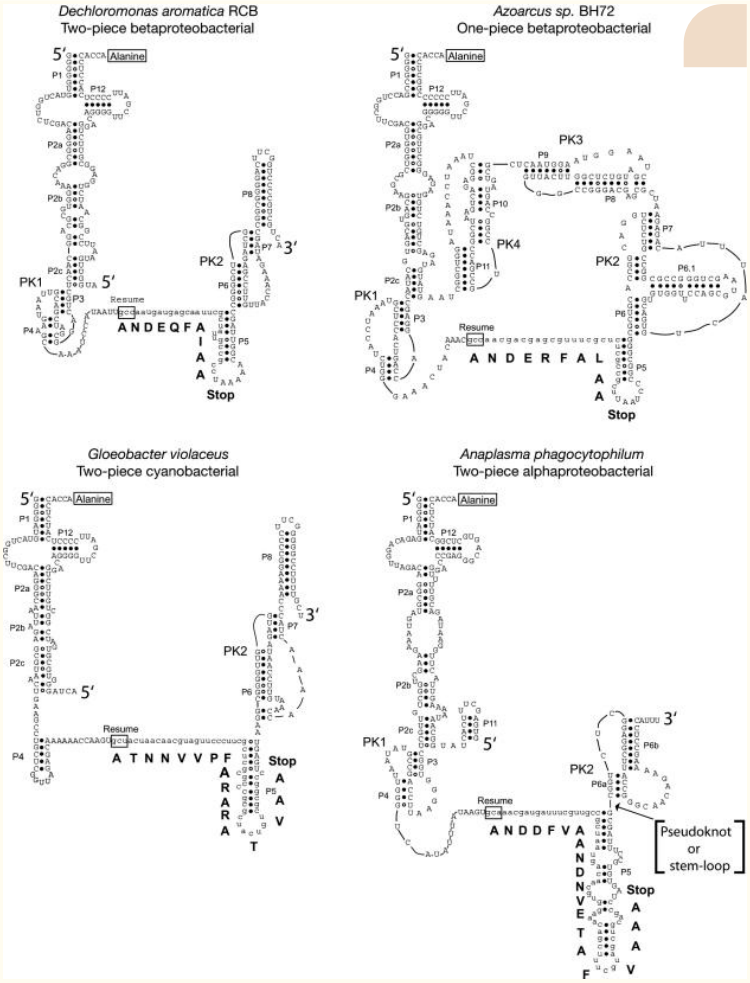

Figure 3

Proposed secondary structure of *D.aromatica* and other tmRNAs. Sequences from *G.violaceus* (22) and *Azoarcus sp*. BH72 (B. Reinhold-Hurek, personal communication) were fit into previously developed models for closely related tmRNAs (6,11). The encoded peptide tag sequences are shown. A site is marked where additional structures can be found in other alphaproteobacteria.

The partial *D.agitata* and *A.oryzae* tmRNA gene sequences provide support for much of the model, even though they do not extend fully through the P2 and P8 pairings (Figure 1). At least one Watson–Crick base-pair covariation is found in each of the following helices: the T stem (P12), a helix in the intervening segment, P4 in the first pseudoknot, P5 and both stems of the second pseudoknot. Stems P1 (acceptor) and P3 (in the first pseudoknot) receive no support from Watson–Crick base-pair covariation; moreover P3 would lose two Watson–Crick base-pairs in *D.agitata*. Nonetheless, P3 remains an attractive proposal due to

the opportunity for its stabilization by stacking on P2 and the possibility of purine–purine pairs maintaining it in *D.agitata*.

Activity of *D.aromatica* tmRNA

Based on an earlier assay (18), an *in vitro* tmRNA translation system has been described that uses purified *E.coli* translational components (19). In this assay, poly(U) serves as nonstop mRNA which does not itself direct the incorporation of alanine and each round of tmRNA usage at a ribosome stalled at the end of poly(U) incorporates five radiolabeled alanine molecules into acid-precipitable polypeptide (one alanine directly from the 3′ end of tmRNA and four additional alanines encoded in the tag reading frame). We used this system to assess the activity of *D.aromatica* two-piece tmRNA (Table 1). Neither piece alone was active, but the combined acceptor and coding pieces provided 8% of the activity of *E.coli* tmRNA and this depended on the tmRNA cofactor SmpB as does *E.coli* tmRNA activity (19). The reduction in activity may reflect imperfect interaction with the heterologous factors more than intrinsically low activity of the tmRNA.

Table 1.

***In vitro* translation of *D.aromatica* tmRNA**

tmRNA type	tmRNA (pmol)	SmpB (pmol)	Precipitable alanine (pmol)
None	0	60	2.4
Acceptor + coding	20	20	5.7
Acceptor + coding	40	40	8.7
Acceptor + coding	60	60	14.0
Acceptor + coding	60	0	3.2
Coding only	60	60	2.4
Acceptor only	60	60	2.7
E.coli	20	20	46.9
E.coli	40	40	66.4

E.coli	60	60	76.7
E.coli	60	0	3.1

DISCUSSION

It is shown here that the circularly permuted tmRNA gene of *D.aromatica*, like those in two other bacterial lineages, produces a two-piece tmRNA. This tmRNA is active *in vitro* with the *E.coli* translational machinery. Two of the four RNA ends could be predicted as the tRNA-type processing sites common to all tmRNAs. A third RNA end appears to be the unprocessed 5′ end of the primary transcript, as can be said for the two-piece tmRNAs of alphaproteobacteria. The process that produces the remaining RNA end is unclear; the G–C-rich, tetraloop-capped stem might be part of a rho-independent transcriptional terminator, although the downstream region contains few uracils.

Origin of the gene

If the tmRNA gene of *Dechloromonas/Azospira* is the product of a third independent case of permutation, it would support the idea that the two-piece structure provides benefits for tmRNA function. Alternatively, the gene may have transferred horizontally from one of the other two known lineages with a permuted tmRNA gene. Sequence comparisons suggest a more complicated and specific variation on the latter hypothesis, namely, that the permuted tRNA domain portion from cyanobacteria fused with the segment containing the tag reading frame from the endogenous betaproteobacterial gene. By far, the best match to the combined segments encoding the tRNA domain is to that from the cyanobacterium *Gloeobacter violaceus*, whose tmRNA gene is also permuted; 57 out of 59 bp in these segments are identical between *G.violaceus* and *D.agitata* (Figure 1). However, homology to the *G.violaceus* gene is not detectable beyond the tRNA domain region. Instead another portion of the gene best matches that of the closest known beta-proteobacterial relative whose tmRNA gene sequence is available, *Azoarcus* sp. BH72; 35 out of 38 bp in a block containing the reading frame region are identical between *Azoarcus* sp. and *A.oryzae* (Figure 1). These two observations suggest the hypothesis that a DNA segment encoding the tRNA domain region was horizontally transferred from a *Gloeobacter*-like cyanobacterium into a *Dechloromonas/Azospira* ancestor where it fused appropriately with the endogenous tmRNA gene. According to this dual-origin hypothesis, the

Dechloromonas/Azospira tmRNA gene does not represent a third independent gene permutation event. The intervening segment between the two portions of the tRNA domain differs between the *Dechloromonas/Azospira* and *G.violaceus* sequences, but this is not a problem for the hypothesis; the intervening segment is hypervariable in length and sequence within both the cyanobacterial and alphaproteobacterial clades.

We investigated the dual-origin hypothesis further by separate phylogenetic analysis of the mature tRNA domain and reading frame sequence blocks, for all available unique tmRNA sequences. The clustering of the tRNA domain sequences of *G.violaceus* and *D.agitata* and of the reading frame regions of *Azoarcus* sp. and *A.oryzae*, suggested by sequence similarity as noted above, was reproduced in the cladograms generated by parsimony analysis (Supplementary Material), and again in trees generated with distance methods. Three points must be made about this phylogenetic analysis. First, it is biased toward the dual-origin hypothesis because the sequence blocks chosen for study were those identified by sequence similarity that led to the hypothesis; this is somewhat justified by the facts that these blocks correspond to functional subregions of tmRNA and that they are the most reliably alignable portions within tmRNA sequences. Second, although sequence clustering in the cladograms largely respects the bacterial phyla determined by other methods, there are many instances of phylum-splitting, such that the trees are only of intermediate reliability. This is most likely due to the small number of phylogenetically informative characters present in either of the sequence blocks. Third, a major character was omitted from the phylogenetic analysis, that of gene configuration (permuted versus standard), which means that the relationship to the *G.violaceus* gene was underevaluated.

In consideration of the dual-origin hypothesis, it should be mentioned that the putative gene transfer donor may itself have gained its tmRNA gene by horizontal transfer. *G.violaceus* appears to be the deepest-branching line of descent in the cyanobacteria/plastid phylum, by multiple criteria (20) including small subunit (SSU) rRNA analysis (21). In a 90% majority-rule cladogram of cyanobacterial/plastidial SSU rRNA sequences, the remainder of the phylum radiates in six deep branches from the next node after that for the *Gloeobacter* line (20). tmRNA sequences are known from six of these seven lines (Figure 4). The tree for tmRNA sequences would not appear to match the SSU rRNA tree. The cyanobacterial one-piece tmRNAs are much like

other bacterial one-piece tmRNAs, yet have distinguishing features such as a doubled pseudoknot and a branched P5 stem that unify them as homologs (15). Plastid tmRNAs have diverged sufficiently to have lost these and other features but are nonetheless tied to cyanobacterial tmRNAs by the tag sequences they encode (12). The two-piece cyanobacterial tmRNAs likewise appear monophyletic. Although the *G.violaceus* sequence diverges somewhat from the two-piece tmRNAs of the *Synechococcus* group, its primary alignment with the *Synechococcus* WH8102 sequence has 70% identity with only five gaps, it fits easily into the secondary structure model developed for the *Synechococcus* group (6) (Figure 3), and it clusters with the two-piece cyanobacterial tmRNAs in parsimony analysis (Figure 4). The key to the relationship between the one-piece and two-piece cyanobacterial tmRNAs lies in the *Synechococcus* group, where both forms are found sharing particularly high sequence similarity; the two-piece *Synechococcus* WH8102 tmRNA has 95 and 76% sequence identity to the one-piece *Synechococcus* PCC6301 tmRNA, in the tRNA domain and tag reading frame, respectively (15). One form therefore appears to have arisen from the other within the *Synechococcus* group, but in which direction? The possibility that two-piece tmRNA was ancestral in cyanobacteria and that the one-piece form arose in a member of the *Synechococcus* group that became an ancestor to most other cyanobacteria, disagrees with much of the shape of the SSU rRNA tree. Instead by comparison with most other bacterial phyla the one-piece form is a more plausible ancestral state for the cyanobacteria, as borne out in the tmRNA cladogram (Figure 4). The idea that the two-piece cyanobacterial tmRNA arose from an event within the *Synechococcus* group agrees with most of the SSU rRNA tree, except for the appearance of the homologous form in *G.violaceus*. Thus by its tmRNA sequence, *G.violaceus* does not appear to be primitive among the cyanobacteria, as it is for so many other characteristics, but lies instead at a tip of the tree. The mismatch with the SSU rRNA tree can be resolved by proposing that the two-piece tmRNA entered *G.violaceus* by horizontal gene transfer from a member of the *Synechococcus* group.

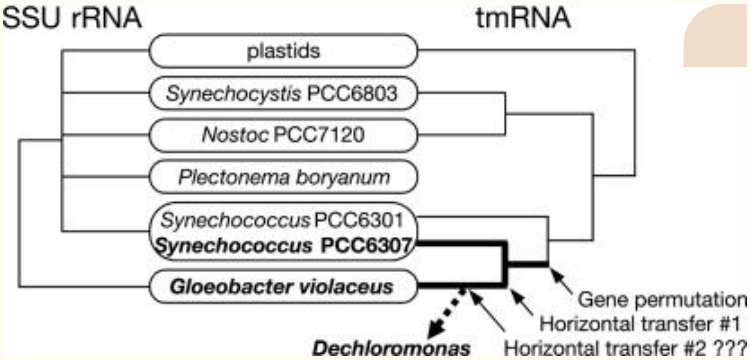

Figure 4

Relationships among cyanobacteria. The SSU rRNA cladogram shows the first two nodes from 90%-rule maximum-likelihood analysis (20) with each of the main lines for which tmRNA sequences are available terminating in boxes labeled with representative species, those with two-piece tmRNA in boldface type. The cladogram from parsimony analysis of tmRNA sequences has the two-piece lineage drawn in thick lines and does not include the *Plectonema* group because no complete sequences are available. Sequence similarity between *G.violaceus* and *Dechloromonas* tRNA domains (Figure 3) suggests the possibility of horizontal transfer into a betaproteobacterium (dashed arrow).

Horizontal transfer of this gene at least once among the cyanobacteria lends support to the idea that it was transferred again into the *Dechloromonas/Azospira* line. Explaining the homology with the betaproteobacterial tag reading frame then becomes difficult. Precise fusion of tmRNA genes from two sources is problematic; sequence homology would not seem sufficient for homologous DNA recombination or any RNA-mediated process. Although we have presented some evidence for a dual-origin hypothesis, in our opinion the gene fusion problem is so severe that we consider it more likely that the similarity to the *G.violaceus* gene is spurious.

A new form for two-piece tmRNA

We derive confidence in the secondary structure model proposed for *D.aromatica* tmRNA from its matching of structures commonly found in other tmRNAs, the opportunities it presents for mutual stabilization of several structures by coaxial stacking and phylogenetic support.

The model has a new form for two-piece tmRNA, whether or not

its gene represents a third historical permutation event (Figure 3). Most tmRNAs including the alphaproteobacterial two-piece tmRNAs have a pseudoknot upstream of the tag reading frame and abutting P2; this pseudoknot would appear well-formed in *D.aromatica*. This would distinguish *D.aromatica* tmRNA from the two-piece cyanobacterial tmRNAs, which have a stem–loop replacing this pseudoknot. It may appear that a potential pairing of the cyanobacterial sequence between P2 and the stem–loop with the loop producing a pseudoknot, but this is not supported by base-pair covariation, and chemical probing showed that these potential pairing partners were single-stranded (6); moreover the potential pseudoknot would not abut P2.

The *D.aromatica* tmRNA is distinguished from the alphaproteobacterial tmRNAs by its termination in a stem–loop rather than a pseudoknot and by the absence of the 5′ stem–loop found in the alphaproteobacterial acceptor piece. A further distinction is that *D.aromatica* has a normal tRNA domain, with a GG dinucleotide in the region corresponding to the D loop and a match to the UUCRANY pre-modification sequence of the T loop, where alphaproteobacteria have an AG dinucleotide and either GGCRGWN or UGCRAUW, respectively (5).

CONCLUSION

The introduction presents the argument that the opening of the mRNA loop in two-piece tmRNAs solves topological problems that arise during translation when the loop is closed. This is supported by the independent evolution of a similar two-piece form in the alphaproteobacteria and cyanobacteria. The newly discovered two-piece tmRNA in a betaproteobacterial lineage may represent yet a third gene permutation event, lending further support to the argument. Alternatively, as suggested by the similarity with a portion of a cyanobacterial two-piece tmRNA, it may have arisen wholly or partially by horizontal transfer of the cyanobacterial gene. In this case there would be no third gene permutation event, yet the replacement of an endogenous one-piece tmRNA with an exogenous two-piece gene would likewise support the idea that the two-piece tmRNA improves tmRNA function in some way. The *D.aromatica* tmRNA is a new structural variation on the two-piece form, although the similarities among the two-piece tmRNAs are probably more striking than their variations. They all effectively break the mRNA loop of one-piece tmRNAs at approximately the same position and have fewer pseudoknots than their one-piece ancestors, in keeping with the notion of two alternative solutions to a topological problem, either pseudo-knots in a looped domain or

an opened loop.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at NAR Online.

[Supplementary Material]

Click here to view.

ACKNOWLEDGEMENTS

We thank John Coates and Romy Chakraborty (University of California, Berkeley, CA) for the *D.aromatica* strain, Laurie Achenbach (Southern Illinois University, Carbondale, IL) for genomic DNAs, Barbara Reinhold-Hurek (University of Bremen) and her collaborators in the Department of Genetics at University of Bielefeld for kindly sharing the *Azoarcus* sp. BH72 tmRNA sequence, Keith Obye for preparing ribosomes and Takuya Ueda (University of Tokyo) for translation factor clones. This work was supported by an NIH grant.

REFERENCES

1. Keiler K.C., Waller,P.R. and Sauer,R.T. (1996) Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. Science, 271, 990–993. [PubMed] [Google Scholar]

2. Komine Y., Kitabatake,M., Yokogawa,T., Nishikawa,K. and Inokuchi,H. (1994) A tRNA-like structure is present in 10Sa RNA, a small stable RNA from *Escherichia coli*. Proc. Natl Acad. Sci. USA, 91, 9223–9227. [PMC free article] [PubMed] [Google Scholar]

3. Ushida C., Himeno,H., Watanabe,T. and Muto,A. (1994) tRNA-like structures in 10Sa RNAs of *Mycoplasma capricolum* and *Bacillus subtilis*. Nucleic Acids Res., 22, 3392–3396. [PMC free article] [PubMed] [Google Scholar]

4. Valle M., Gillet,R., Kaur,S., Henne,A., Ramakrishnan,V. and Frank,J. (2003) Visualizing tmRNA entry into a stalled ribosome. Science, 300, 127–130. [PubMed] [Google Scholar]

5. Keiler K.C., Shapiro,L. and Williams,K.P. (2000) tmRNAs that encode proteolysis-inducing tags are found in all known bacterial genomes: a two-piece tmRNA functions in *Caulobacter*. Proc. Natl

Acad. Sci. USA, 97, 7778–7783. [PMC free article] [PubMed] [Google Scholar]

6. Gaudin C., Zhou,X., Williams,K.P. and Felden,B. (2002) Two-piece tmRNA in cyanobacteria and its structural analysis. Nucleic Acids Res., 30, 2018–2024. [PMC free article] [PubMed] [Google Scholar]

7. Spedding G. (1990) Isolation and analysis of ribosomes from prokaryotes, eukaryotes, and organelles. In Spedding,G. (ed.), Ribosomes and Protein Synthesis: A Practical Approach. Oxford University Press, NY, pp. 1–29. [Google Scholar]

8. Reuven N.B. and Deutscher,M.P. (1993) Multiple exoribonucleases are required for the 3′ processing of *Escherichia coli* tRNA precursors *in vivo*. FASEB J., 7, 143–148. [PubMed] [Google Scholar]

9. Shimizu Y., Inoue,A., Tomari,Y., Suzuki,T., Yokogawa,T., Nishikawa,K. and Ueda,T. (2001) Cell-free translation reconstituted with purified components. Nat. Biotechnol., 19, 751–755. [PubMed] [Google Scholar]

10. Coates J.D., Chakraborty,R., Lack,J.G., O'Connor,S.M., Cole,K.A., Bender,K.S. and Achenbach,L.A. (2001) Anaerobic benzene oxidation scoupled to nitrate reduction in pure culture by two strains of *Dechloromonas*. Nature, 411, 1039–1043. [PubMed] [Google Scholar]

11. Felden B., Massire,C., Westhof,E., Atkins,J.F. and Gesteland,R.F. (2001) Phylogenetic analysis of tmRNA genes within a bacterial subgroup reveals a specific structural signature. Nucleic Acids Res., 29, 1602–1607. [PMC free article] [PubMed] [Google Scholar]

12. Gueneau de Novoa P. and Williams,K.P. (2004) The tmRNA website: reductive evolution of tmRNA in plastids and other endosymbionts. Nucleic Acids Res., 32, D104–D108. [PMC free article] [PubMed] [Google Scholar]

13. Achenbach L.A., Michaelidou,U., Bruce,R.A., Fryman,J. and Coates,J.D. (2001) *Dechloromonas agitata gen. nov., sp. nov*. and *Dechlorosoma suillum gen. nov., sp. nov*., two novel environmentally dominant (per)chlorate-reducing bacteria and their phylogenetic position. Int. J. Syst. Evol. Microbiol., 51, 527–533. [PubMed] [Google Scholar]

14. Tan Z. and Reinhold-Hurek,B. (2003) *Dechlorosoma suillum* Achenbach *et al*. 2001 is a later subjective synonym of *Azospira*

oryzae Reinhold-Hurek and Hurek 2000. Int. J. Syst. Evol. Microbiol., 53, 1139–1142. [PubMed] [Google Scholar]

15. Williams K.P. (2002) Descent of a split RNA. Nucleic Acids Res., 30, 2025–2030. [PMC free article] [PubMed] [Google Scholar]

16. Mörl M. and Marchfelder,A. (2001) The final cut. The importance of tRNA 3′-processing. EMBO Rep., 2, 17–20. [PMC free article] [PubMed] [Google Scholar]

17. Williams K.P. and Bartel,D.P. (1996) Phylogenetic analysis of tmRNA secondary structure. RNA, 2, 1306–1310. [PMC free article] [PubMed] [Google Scholar]

18. Himeno H., Sato,M., Tadaki,T., Fukushima,M., Ushida,C. and Muto,A. (1997) *In vitro* trans translation mediated by alanine-charged 10Sa RNA. J. Mol. Biol., 268, 803–808. [PubMed] [Google Scholar]

19. Shimizu Y. and Ueda,T. (2002) The role of SmpB protein in trans-translation. FEBS Lett., 514, 74–77. [PubMed] [Google Scholar]

20. Turner S., Pryer,K.M., Miao,V.P. and Palmer,J.D. (1999) Investigating deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence analysis. J. Eukaryot. Microbiol., 46, 327–338. [PubMed] [Google Scholar]

21. Nelissen B., Van de Peer,Y., Wilmotte,A. and De Wachter,R. (1995) An early origin of plastids within the cyanobacterial divergence is suggested by evolutionary trees based on complete 16S rRNA sequences. Mol. Biol. Evol., 12, 1166–1173. [PubMed] [Google Scholar]

22. Nakamura Y., Kaneko,T., Sato,S., Mimuro,M., Miyashita,H., Tsuchiya,T., Sasamoto,S., Watanabe,A., Kawashima,K., Kishida,Y. *et al*. (2003) Complete genome structure of *Gloeobacter violaceus* PCC 7421, a cyanobacterium that lacks thylakoids. DNA Res., 10, 137–145. [PubMed] [Google Scholar]

A third lineage with two-piece tmRNA