446–447 Nucleic Acids Research, 2003, Vol. 31, No. 1
DOI: 10.1093/nar/gkg018

tmRDB (tmRNA database)

Christian Zwieb*, Jan Gorodkin[1], Bjarne Knudsen[2], Jody Burks and Jacek Wower[3]

Department of Molecular Biology, The University of Texas Health Science Center at Tyler, 11937 US Highway 271, Tyler, TX 75708-3154, USA, [1]Division of Genetics, The Royal Veterinary and Agricultural University, Institute of Animal Science and Animal Health, Groennegaardsvej 3, DK-1870 Frederiksberg C, Denmark, [2]Bioinformatics Research Center, University of Aarhus, Høegh Guldbergsgade 10, DK-8000 Aarhus C, Denmark and [3]Department of Animal and Dairy Sciences, Program in Cellular and Molecular Biosciences, Auburn University, Auburn, AL 36849-5415, USA

Received August 24, 2002; Accepted September 20, 2002

ABSTRACT

Maintained at the University of Texas Health Science Center at Tyler, Texas, the tmRNA database (tmRDB) is accessible at the URL http://psyche.uthct.edu/dbs/tmRDB/tmRDB.html with mirror sites located at Auburn University, Auburn, Alabama (http://www.ag.auburn.edu/mirror/tmRDB/) and the Bioinformatics Research Center, Aarhus, Denmark (http://www.bioinf.au.dk/tmRDB/). The tmRDB collects and distributes information relevant to the study of tmRNA. In *trans*-translation, this molecule combines properties of tRNA and mRNA and binds several proteins to form the tmRNP. Related RNPs are likely to be functional in all bacteria. In this release of tmRDB, 186 new entries from 10 bacterial groups for a total of 274 tmRNA sequences have been added. Lists of the tmRNAs and the corresponding tmRNA-encoded tag-peptides are presented in alphabetical and phylogenetic order. The tmRNA sequences are aligned manually, assisted by computational tools, to determine base pairs supported by comparative sequence analysis. The tmRNA alignment, available in a variety of formats, provides the basis for the secondary and tertiary structure of each tmRNA molecule. Three-dimensional models of the tmRNAs and their associated proteins in PDB format give evidence for the recent progress that has been made in the understanding of tmRNP structure and function.

tmRDB TABLE OF CONTENTS

tmRNA FUNCTION

Trans-translation is a process characterized by the movement of a ribosome from the 3′-end of damaged mRNA to the resume codon of tmRNA. This step results in the covalent attachment of a tmRNA-encoded tag-peptide to the incompletely-translated polypeptide and is followed by the proteolytic destruction of the protein. The ribosome is released to initiate another round of translation. Several proteins have been shown to bind to tmRNA, including elongation factor Tu, protein SmpB and ribosomal protein S1. For reviews, see (1,2). Recent studies highlight some circumstances that trigger tmRNA binding to ribosomes and the transfer of the truncated protein from tRNA to tmRNA (3–5).

tmRDB DESCRIPTION

The tmRDB provides aligned, annotated, and phylogenetically ordered sequences related to tmRNA structure and function. New and updated tmRNAs were obtained by using a representative subset of sequences as input for BLASTN to search GenBank (6), the various genome sequencing projects accessible via the World Wide Web, and the tmRNA Website (7).

186 new tmRNA sequences (a total of 274) were from the following species (ordered alphabetically): *Acinetobacter ADP1, Actinobacillus pleuropneumoniae, Agrobacterium tumefaciens, Alcaligenes faecalis, Azotobacter vinelandii, Bacillus halodurans, B.stearothermophilus, Bacteroides forsythus, B.fragilis, Bolidomonas* species plastid, *Bordetella parapertussis, Brucella melitensis, Buchnera* species, *Burkholderia cepacia, B.fungorum, B.mallei, B.pseudomallei, Carboxydothermus hydrogenoformans, Carnobacterium piscicola* (partial sequence), *Chlamydophila abortus, Chloroflexus aurantiacus, Chromobacterium violaceum, Chroococcidiopsis* species,

*To whom correspondence should be addressed. Tel: +1 903 877 7689; Fax: +1 903 877 5731; Email: zwieb@uthct.edu

Clostridium botulinum, C.perfringens, C.thermocellum, Colwellia species, *Comamonas testosteroni, Corynebacterium glutamicum, Coxiella burnetii, Cytophaga hutchinsonii, Dehalococcoides ethenogenes, Desulfitobacterium hafniense, Ehrlichia chaffeensis, Enterococcus durans, Erwinia carotovora, E.chrysanthemi, Fremyella diplosiphon, Haemophilus somnus, Heliobacillus mobilis, Hydrogenophaga palleronii, Lactobacillus delbrueckii, L.gallinarum, L.gasseri, L.helveticus, L.plantarum, L.sakei, Lactococcus garvieae, L.lactis, L.plantarum, L.raffinolactis, Leuconostoc lactis, L.mesenteroides, L.pseudomesenteroides, Listeria grayi, L.innocua, L.ivanovii, L.monocytogenes, L.seeligeri, L.welshimeri, Magnetospirillum magnetotacticum, Mannheimia haemolytica, Mesorhizobium loti, Mesostigma viride* chloroplast, *Methylobacillus glycogenes, Methylococcus capsulatus, Microbulbifer degradans, Mycobacterium marinum, M.smegmatis, Mycoplasma pulmonis, Nitrosomonas cryotolerans, Oenococcus oeni, Oscillatoria* species, *Pavlova lutheri* chloroplast (partial sequence), *Pediococcus pentosaceus, Photobacterium phosphoreum, Plectonema boryanum, Prevotella intermedia, Prochlorococcus marinus, Providencia rettgeri, Pseudomonas fluorescens, P.syringae, Ralstonia eutropha, R.metallidurans, R.pickettii, R.solanacearum, Rhizobium leguminosarum, Rhodobacter sphaeroides, Rhodospirillum rubrum, Rickettsia conorii, R.typhi, Ruminococcus albus, Salmonella enterica, S.typhi, Shigella dysenteriae, S.flexneri, Silicibacter pomeroyi, Sphingomonas aromaticivorans, Spiroplasma kunkelii, Staphylococcus saprophyticus, S.xylosus, Streptococcus mitis, S.suis* (partial sequence), *S.thermophilus, S.uberis, Tannerella forsythensis, Thermoanaerobacter tengcongensis, Thermobifida fusca, Treponema denticola, Trichodesmium erythraeum, Tropheryma whipple, Variovorax paradoxus, Wolbachia* species, *Xanthomonas axonpodis, Xanthomonas campestris*, and *Yersinia enterocolitica*. In addition, 43 sequences were derived from uncultured samples.

Because of the significant increase in the number of sequences, the tmRDB nomenclature for the abbreviated species names were revised to include the first four letters of the genus and species name, respectively, with modifiers to indicate organelles (chloroplast, cyanelle, or mitochondrion), a partial sequence, or ambiguous residues.

Using BioEdit, available at http://www.mbio.ncsu.edu/BioEdit/bioedit.html, the previously described criteria (8) were applied to incorporate the new sequences into the previous tmRNA alignment (9). Automated procedures were avoided as they tended to maximize base pairing and ignore idiosyncratic properties. The alignment was submitted iteratively to RNAdbtool (10), accessible at http://www.bioinf.au.dk/rnadbtool, to check base pairing consistency, phylogenetic support, and discover helical extensions. The tRNA- and mRNA-like portions, the four pseudoknots, and a total of 17 helices were confirmed, although certain features were absent in some tmRNAs (11). The secondary structures of the nine organelle tmRNAs were improved, and two residues were found to be inserted in the tmRNA of *Dehalococcoides ethenogenes* which might form a base pair in the conserved tRNA-like portion. However, it is currently unknown if this tmRNA is functional.

The tmRDB also provides alignments of predicted tmRNA encoded tag peptides, several tmRNA secondary structure models, and tentative three-dimensional models in PDB format generated interactively with ERNA-3D (12) followed by refinement using VCMD (13), as well as information about the tmRNA-associated proteins alanyl-tRNA synthetase, elongation factor Tu, SmpB, and ribosomal protein S1.

ACCESS

The data are freely accessible for research purposes at http://psyche.uthct.edu/dbs/tmRDB/tmRDB.html, http://www.ag.auburn.edu/mirror/tmRDB/ and http://www.bioinf.au.dk/tmRDB/. Suggestions should be directed to zwieb@uthct.edu. J.G. and B.K. can be reached at the email addresses gorodkin@bioinf.kvl.dk and bk@bioinf.au.dk, respectively. J.B. can be contacted at jodymburks@hotmail.com, and J.W. at jwower@acesag.auburn.edu. This article should be cited in research projects assisted by the use of the tmRDB.

ACKNOWLEDGEMENTS

We thank Florian Müller for ERNA-3D and Ingolf Sommer for VCMD software. J.G. is supported by the Danish Technical Research Council and the Ministry of Food, Agriculture and Fisheries. This work was supported by NIH grant GM-58267 to J.W.

REFERENCES

1. Karzai,A.W., Roche,E.D. and Sauer,R.T. (2000) The SsrA-SmpB system for protein tagging, directed degradation and ribosome rescue. *Nature Struct. Biol.,* **7**, 449–455.
2. Wower,J., Zwieb,C. and Wower,I.K. (2000) tmRNA: a case of structural and functional mimicry on and off the ribosome? In Garrett,R.A., Douthwaite,S.R., Liljas,A., Matheson,A.T., Moore,P.B. and Noller,H.F. (eds), *The Ribosome: Structure, Function, Antibiotics, and Cellular Interactions*. ASM Press, Washington, D.C., pp. 397–405.
3. Collier,J., Binet,E. and Bouloc,P. (2002) Competition between SsrA tagging and translational termination at weak stop codons in *Escherichia coli*. *Mol. Microbiol.*, **45**, 745–754.
4. Hayes,C.S., Bose,B. and Sauer,R.T. (2002) Proline residues at the C-terminus of nascent chains induce SsrA-tagging during translation termination. *J. Biol. Chem.*, **277**, 33825–33832.
5. Hayes,C.S., Bose,B. and Sauer,R.T. (2002) Stop codons preceded by rare arginine codons are efficient determinants of SsrA tagging in *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **99**, 3440–3445.
6. Altschul,S.F., Madden,T.L., Schäffer,A.A., Zhang,J., Zhang,Z., Miller,W. and Lipman,D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, **25**, 3389–3402.
7. Williams,K. (2002) The tmRNA Website: invasion by an intron. *Nucleic Acids Res.*, **30**, 179–182.
8. Larsen,N. and Zwieb,C. (1991) SRP-RNA sequence alignment and secondary structure. *Nucleic Acids Res.*, **19**, 209–215.
9. Knudsen,B., Wower,J., Zwieb,C., and Gorodkin,J. (2001) tmRDB (tmRNA database). *Nucleic Acids Res.,* **29**, 171–172.
10. Gorodkin,J., Zwieb,C. and Knudsen,B. (2001) Semi automated update and clean up of structural RNA alignment databases. *Bioinformatics*, **17**, 642–645.
11. Zwieb,C., Wower,I. and Wower,J. (1999) Comparative sequence analysis of tmRNA. *Nucleic Acids Res.*, **27**, 2063–2071.
12. Müller,F., Döring,T., Erdemir,T., Greuer,B., Jünke,N., Osswald,M., Rinke-Appel,L., Stade,K., Thamm,S. and Brimacombe,R. (1995) Getting closer to an understanding of the three-dimensional structure of ribosomal RNA. *Biochem. Cell Biol.*, **73**, 767–773.
13. Sommer,I. and Brimacombe,R. (2001) Methods for refining interactively established models of ribosomal RNA towards physico-chemically plausible structure. *J. Comput. Chem.*, **22**, 407–417.