The signal recognition particle (SRP) database

Christian Zwieb and Niels Larsen[1]
Department of Molecular Biology, The University of Texas Health Science Center, PO Box 2003, Tyler, TX 75710 and [1]Department of Microbiology, 131 Burrill Hall, 407 South Goodwin Avenue, University of Illinois, Urbana, IL 61801, USA

The SRP (Signal Recognition Particle) data collection, located at University of Texas Health Science Center at Tyler, consists of 47 aligned and annotated (see Table 1) SRP RNA sequences and three representative secondary structure models (from a bacterium, an archaeon and a eucaryote). Eventually, the database will contain protein sequences (there are currently only a few available), as well as related experimental data. The present release is an update of an earlier announcement (1), where 39 sequences were compared in order to refine the overall SRP RNA secondary structure and evaluate possible tertiary interactions.

The new entries include two *Mycoplasma* sequences, SRP RNAs from *Arabidopsis*, *Trypanosoma*, *Tetrahymena*, *Yarrowia*, and mouse B1 RNA (see Table 1). The SRP RNA from *Mycoplasma pneumoniae* (3) is the smallest known so far (79 bases). A *Saccharomyces cerevisiae* scRNA (4) may be the largest, but, since only a short portion near the 5· end can be aligned, we chose not to include it. The new sequences provide no additional evidence for the recently proposed helix 7 (1), but they all conform to, and in some cases support, the established minimal secondary structure models presented earlier (1).

The sequences can be obtained in formats used by the UWGCG sequence analysis package (5), GenBank (6) and EMBL (7). The alignment is available in the ALMA alignment editor format (2) or as files used by the AE2 alignment editor (see the Ribosomal Database Project announcement, this issue). We also offer the alignment as a printable PostScript version, which highlights base pairs and helices numbered using the previous convention (1).

We accept submission of SRP related data in any form. The submitter may request that the data may not be released until after a given date or upon notification. New sequences will be aligned and the updated alignment will be returned to the submitter.

The data are freely accessible by 'anonymous ftp': connect with ftp to 129.117.100.7, login with name 'anonymous' and with any password (preferably your e-mail address). A help file explains where to find the data. Hardcopies are also available by mailing a request to the first author.

ACKNOWLEDGMENTS

REFERENCES

1. Larsen, N. and Zwieb C. (1991) Nucleic acids Res. **19**, 209−215
2. Thirup, S. and Larsen, N. (1990) Proteins: Structure, Function and Genetics, **7**, 291−295
3. Simoneau, P. and Hu, P. (1992) J. Bact. **174**, 627−629.
4. Felici, F., Cesareni, G. and Hughes, J. M. (1989) Mol. Cell. Biol. **9**, 3260−3268
5. Devereux, J., Haeberli, P. and Smithies, O. (1984) Nucleic acids Res. **12**, 387−395
6. Burks, C., Cassidy, M., Cinkosky, M.J., Cumella, K.E., Gilna, P., Hayden, J.E-D., Keen, G.M., Kelley, T.A., Kelly, M., Kristofferson, D. and Ryals, J. (1991) Nucleic acids Res. **19**, Supplement, 2221−2225
7. Stoehr, P.J. and Cameron, G.N. (1991) Nucleic acids Res. **19**, Supplement, 2227−2230

Table 1. SRP RNA sequences in the SRP database

Domain	ID	Name	Source
Bacteria			
	MYC.PNE.[a]	Mycoplasma pneumoniae	DNA
	MYC.MYC.[a]	Mycoplasma mycoides	RNA
	THE.THE.	Thermus thermophilus	DNA
	LEG.PNE.	Legionella pneumophilia	DNA
	PSE.AER.	Pseudomonas aeruginosa	DNA
	ESC.COL.	Escherichia coli	DNA
	MIC.LYS.	Micrococcus lysodeikticus	DNA
	BAC.SUB.	Bacillus subtilis	RNA
Archaea			
	MET.VOL.	Methanococcus voltae	DNA
	MET.FER.	Methanothermus fervidus	DNA
	MET.THE.	Methanobacterium thermoautotrophicum	DNA
	MET.ACE.	Methanosarcina acetivorans	DNA
	HAL.HAL.	Halobacterium halobium	DNA
	ARC.FUL.	Archaeoglobus fulgidus	DNA
	SUL.SOL.	Sulfolobus solfataricus	DNA
	PYR.OCC.	Pyrodictium occultum	DNA
	THE.CEL.	Thermococcus celer	DNA
Eucarya			
	ZEA.MAY.A-G,H[b]	Zea mays	SRP
	TRI.AES.A-B,C[b]	Triticum aestivum	SRP
	ARA.THA.[a]	Arabidopsis thaliana	RNA
	CRY.MOR.[b]	Crysanthemum morifolium	RNA
	LYC.ESC.	Lycopersicon esculentum	RNA
	CIN.HYB.	Cineraria hybrida	RNA
	HOM.SAP.A-B	Homo sapiens	RNA
	CAN.SPE.	Canis species	SRP
	RAT.RAT.	Rattus rattus	RNA
	MUS.M.B1[a]	Mus musculus	RNA
	XEN.LAE.	Xenopus laevis	RNA
	DRO.MEL.	Drosophila melanogaster	RNA
	TRY.BRU.[a]	Trypanosoma brucei	RNA
	TET.THE.[a]	Tetrahymena thermophila	RNA
	TET.ROS.[a]	Tetrahymena rostrata	RNA
	SCH.POM.	Schizosaccharomyces pombe	SRP
	YAR.LIP.A,B[a]	Yarrowia lipolytica	SRP
	BEN.HIS.[b]	Benincasa hispida	RNA
	GYN.AUR.[b]	Gynura aurantiaca	RNA

DNA, sequence was determined from the gene; RNA, RNA was isolated from a crude mixture and either sequenced directly or from its cDNA; SRP, RNA was isolated from a purified particle and either sequenced directly or from its cDNA; [a] Recent addition to the database; [b] Incomplete sequence.