© 1997 Oxford University Press Nucleic Acids Research, 1997, Vol. 25, No. 1 107–108

# The Signal Recognition Particle Database (SRPDB)

**Christian Zwieb* and Niels Larsen[1]**

Department of Molecular Biology, The University of Texas Health Science Center, PO Box 2003, Tyler, TX 75710, USA and [1]Department of Microbiology, Giltner Hall, Michigan State University, East Lansing, MI 48824, USA

Received October 8, 1996; Accepted October 9, 1996

## ABSTRACT

**The SRPDB (Signal Recognition Particle Database) offers aligned SRP RNA and SRP protein sequences, phylogenetically ordered and annotated. This release adds three SRP RNA sequences (totaling 96 SRP RNA sequences) and 11 SRP protein sequences (a total of 39 protein sequences from SRP9, SRP14, SRP19, SRP21, SRP54, SRP68 or SRP72). Also downloadable are sample SRP RNA secondary structure diagrams, a three-dimensional model of the human SRP RNA, search motifs and software.**

## DESCRIPTION

The Signal Recognition Particle Database (SRPDB) is located at the University of Texas Health Science Center at Tyler. It provides 96 aligned and annotated sequences of SRP RNAs and 39 SRP protein sequences. Alignments were created using the rules described in ref. (1), and the sequences are grouped and ordered according to the phylogeny derived by the Ribosomal Database Project (2).

### SRP RNAs

In order to update the SRP RNA alignment, we searched release 96 of GenBank (3) using the software PatScan (4) and a series of motifs that each describe the most conserved primary and secondary structure features, with broad allowance for mismatches. Only four new SRP RNAs were identified, one from *Oryza sativa* (Chittenden *et al*., unpublished), two identical ones from *Caulobacter crescentus* (an alpha-purple bacterium), and one from *Clostridium perfringens* (a low-GC Gram-positive bacterium). The structure of the latter, the gene of which follows immediately after the 3′CCA sequence of a serine tRNA, is 70% similar to the SRP RNA from *Bacillus*, and has been implicated in sporulation (5).

The SRP RNA alignment is available as concatenated GenBank (6) and EMBL (7) entries with gaps inserted in the sequences, as a human readable text format, and as printable PostScript where helices are numbered and highlighted. The corresponding sequences are also available as separate GenBank and EMBL entries. Representative SRP RNA secondary structure models (in PostScript format) are from *Bacillus subtilis* (Bacteria), *Halobacterium halobium* (Archaea) and *Canis* sp. (Eukaryota). A tentative three-dimensional model [(8) in PDB format] is offered for human SRP RNA.

### SRP proteins

New SRP protein sequences include human SRP9 (a total of four SRP9 sequences), a canine SRP14 sequence and SRP19 from *Methanococcus jannaschii*. The latter was identified in the complete genome of *M.jannaschii* (9) as a polypeptide of 87 residues using the motif P–(SVCA)–Y–X(8)–E–G–R–X(10)–P (10). The existence of an archaeal SRP19 homolog (Fig. 1) was predicted earlier from structural studies of the SRP19–SRP RNA complex (11,12) and is now confirmed. *Methanococcus jannaschii* also contains a homolog for SRP54 (9), but equivalents of SRP9, SRP14, SRP68 and SRP72 were not identified in our search; these proteins therefore may be absent in the archaeal SRP. Other new SRP54 sequences were from *Mycoplasma genitalium*, barley (*Hordeum vulgare*, three closely related sequences), *Caenorhabditis elegans* (a partial sequences that aligns near the C-terminus of SRP54), and human SRP54. The SRP72 sequence from *Saccharomyces cerevisiae* (13) was aligned. If possible, protein entries were linked to their primary source at GenBank (6) or EBI (7). The SRP protein alignments are available in textual (ASCII) or PostScript format.

### Pattern matching

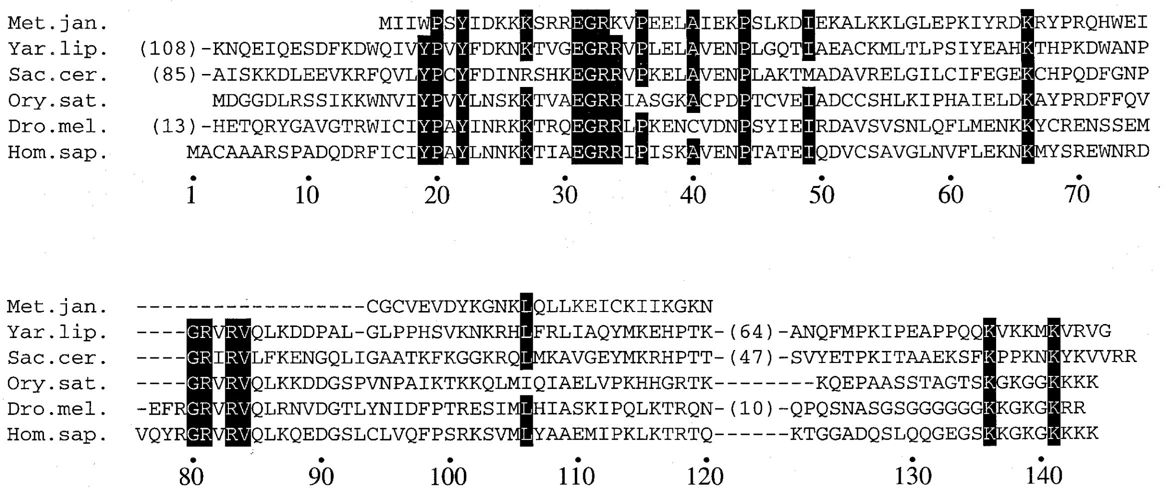
PatScan, a flexible pattern matching program (4), is available together with the search motifs used to identify new sequences. This enables anyone to repeat the search, and to change our pattern definitions.

## ACCESS

All data and software are freely accessible and downloadable: connect your World Wide Web browser to the URL http://pegasus.uthct.edu/SRPDB/SRPDB.html. SRPDB can also be accessed directly by anonymous ftp to 'diana.uthct.edu'. (Currently 192.88.11.4; login with the name 'anonymous', without the quotes, and give your full electronic mail address as the password.) The complete SRPDB can be downloaded from the same site as a single tar-file. Hardcopies of the alignments are available through written contact or through e-mail to the authors at 'zwieb@jason.uthct.edu'; however electronic transfer is

---

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



```
Met.jan.                      MIIWPSYIDKKKSRREGRKVPEELAIEKPSLKDIEKALKKLGLEPKIYRDKRYPRQHWEI
Yar.lip.  (108)-KNQEIQESDFKDWQIVYPVYFDKNKTVGEGRRVPLELAVENPLGQTIAEACKMLTLPSIYEAHKTHPKDWANP
Sac.cer.   (85)-AISKKDLEEVKRFQVLYPCYFDINRSHKEGRRVPKELAVENPLAKTMADAVRELGILCIFEGEKCHPQDFGNP
Ory.sat.        MDGGDLRSSIKKWNVIYPVYLNSKKTVAEGRRIASGKACPDPTCVEIADCCSHLKIPHAIELDKAYPRDFFQV
Dro.mel.   (13)-HETQRYGAVGTRWICIYPAYINRKKTRQEGRRLPKENCVDNPSYIEIRDAVSVSNLQFLMENKKYCRENSSEM
Hom.sap.      MACAAARSPADQDRFICIYPAYLNNKKTIAEGRRIPISKAVENPTATEIQDVCSAVGLNVFLEKNKMYSREWNRD
                 1         10        20        30        40        50        60        70

Met.jan.  ------------------CGCVEVDYKGNKLQLLKEICKIIKGKN
Yar.lip.  ----GRVRVQLKDDPAL-GLPPHSVKNKRHLFRLIAQYMKEHPTK-(64)-ANQFMPKIPEAPPQQKVKKMKVRVG
Sac.cer.  ----GRIRVLFKENGQLIGAATKFKGGKRQLMKAVGEYMKRHPTT-(47)-SVYETPKITAAEKSFKPPKNKYKVVRR
Ory.sat.  ----GRVRVQLKKDDGSPVNPAIKTKKQLMIQIAELVPKHHGRTK--------KQEPAASSTAGTSKGKGGKKKK
Dro.mel.  -EFRGRVRVQLRNVDGTLYNIDFPTRESIMLHIASKIPQLKTRQN-(10)-QPQSNASGSGGGGGGKKGKGKRR
Hom.sap.  VQYRGRVRVQLKQEDGSLCLVQFPSRKSVMLYAAEMIPKLKTRTQ------KTGGADQSLQQGEGSKKGKGKKKK
              80        90        100       110       120            130       140
```

**Figure 1.** Comparison of homologous protein SRP19 sequences. Sequences are from *Methanococcus jannaschi* (Met.jan.) (9), *Yarrowia lipolytica* (Yar.lip.; Dominguez, A., *et al*., unpublished), *Saccharomyces cerevisiae* (Sac.cer.) (14) (Barrell, B., *et al*., unpublished), *Oryza sativa* (Ory.sat.; Chittenden, K., *et al*., unpublished), *Drosophila melanogaster* (Dro.mel.; Lai, C.Q. and Langley, C.H., unpublished), and *Homo sapiens* (Hom.sap.) (15). Residue positions of the human SRP19 are indicated at the bottom of each panel; bracketed numbers represent the number of amino acids not shown; letters in reverse print are amino acid residues that are conserved in at least five of the six sequences. The full alignment is available from the SRPDB.

preferred. Please cite this article if your research is assisted by the SRPDB.

Submission of SRP-related data will be accepted in any form. We will align submitted sequences and return the alignment to the submitter in the requested format. The submitter may request that the data not be released until after a given date or upon notification. C.Z. can be contacted by e-mail at zwieb@jason.uthct.edu, and N.L. can be contacted by e-mail at niels@truth.mph.msu.edu.

## ACKNOWLEDGEMENTS

We thank the investigators who have submitted data prior to publication. This work was assisted by NIH grant GM-49034 to C.Z. and by the Center for Microbial Ecology at Michigan State University. We also thank the Free Software Foundation and Larry Wall for their excellent free software, and Kimberly Chittenden for critical reading of the manuscript.

## REFERENCES

1 Larsen, N. and Zwieb, C. (1991) *Nucleic Acids Res.*, **19**, 209–215.
2 Maidak, B.L., Olsen, G.J., Larsen, N., Overbeek, R., McCaughey, M.J. and Woese, C.R. (1996) *Nucleic Acids Res.*, **24**, 82–85.
3 Benson, D.A., Boguski, M.. Lipman, D.J. and Ostell, J. (1996) *Nucleic Acids Res.*, **24**, 1–5.
4 Overbeek, R. and Price, M. (1993) *Accessing Integrated Genomic Data Using GenoBase: A Tutorial* (Part 1). Report ANL/MCS-TM-173, Argonne National Laboratory.
5 Nakamura, K., Hashizume, E., Shibata, T., Nakamura, Y., Mala, S. and Yamane, K. (1995) *Microbiology*, **141**, 2965–2975.
6 Benson, D.A., Boguski, M., Lipman, D.J. and Ostell, J. (1996) *Nucleic Acids Res.*, **24**, 1–5.
7 Rodriguez-Tomé, Stoehr, P.J., Cameron, G.N. and Flores, T.P. (1996) *Nucleic Acids Res.*, **24**, 6–12.
8 Zwieb, C., Müller, F. and Larsen, N. (1996) *Folding Design*, **1**, 315–324.
9 Bult, C.J., White, O., Olsen, G.J., Zhou, L., Fleischmann, R.D., Sutton, G.G., Blake, J.A., FitzGerald, L.M., Clayton, R.A., Gocayne, J.D. *et al.* (1996) *Science*, **273**, 1058–1073.
10 Bairoch, A., Bucher, P. and Hofmann, K. (1996) *Nucleic Acids Res.*, **1**, 189–196.
11 Zwieb, C. (1994) *Eur. J. Biochem.*, **222**, 885–890.
12 Zwieb, C. (1992) *J. Biol. Chem.*, **267**, 15650–15656.
13 Brown, J.D., Hann, B.C., Medzihradszky, K.F., Niwa, M., Burlingame, A.L. and Walter, P. (1994) *EMBO J.*, **13**, 4390–4400.
14 Stirling, C.J. and Hewitt, E.W. (1992) *Nature*, **356**, 534–537.
15 Lingelbach, K., Zwieb, C., Webb, J. R., Marshallsay, C., Hoben, P. J., Walter, P. and Dobberstein, B. (1988) *Nucleic Acids Res.*, **16**, 9431–9442.