

A homologue of the 19 kDa signal recognition particle protein locus in *Drosophila melanogaster*

Chaoqiang Lai *, Charles H. Langley

Center for Population Biology, University of California, Davis, CA 95616, USA

Received 24 January 1997; accepted 1 August 1997; Received by C.M. Kane

Abstract

A homologue of 19 kDa signal recognition particle locus (SRP19) was cloned and molecularly characterized in *Drosophila melanogaster*. It is located in the 65F region of the left arm on the third chromosome, approx. 500 bp 5' of the *quemao* locus. The *SRP19* transcript was determined from cDNA clones, Northern blot analysis, and the 5' rapid amplification of cDNA end method. *SRP19* was expressed in all the developmental stages of *Drosophila*. The predicted amino acid sequence (163 aa) shows that *SRP19* of *Drosophila* shares 44%, 29%, 17% and 19% identity with the homologues from human, rice and two yeast species (*Saccharomyces* and *Yarrowia*), respectively. The most conserved amino acid residues across these species are located at those sites required for in vitro association with the 7S RNA component of the SRP. © 1997 Elsevier Science B.V.

Keywords: Protein secretion; Translocation; RNA binding; Gene expression

1. Introduction

Cellular protein secretion is a ubiquitous and essential process in all organisms. Translocation of the nascent polypeptides synthesized by ribosomes across the endoplasmic reticulum (ER) membrane is the first step in the eukaryotic secretory pathway. This process requires the signal recognition particle (SRP), the signal peptide of the nascent protein on the ribosome and the SRP-receptor in the ER-membrane (reviewed in Walter and Lingappa, 1986; Rapoport, 1990). During translocation, the SRP binds to the signal sequence of the nascent polypeptide chain as it emerges from the ribosome. This slows protein synthesis, which is thought to facilitate the docking of the ribosome to the receptor on the ER. This ensures that the protein is not released into the cytosol (reviewed in Walter and Lingappa, 1986; Nunnari and Walter, 1992; Wolin, 1994). The targeting of the SRP-bound ribosome to the ER membrane involves three steps: (1) the SRP binds to the SRP receptor, or docking protein, which is an integral membrane protein exposed on the cytosolic surface of the rough ER membrane; (2) the ribosome binds to the

ribosome receptor in the ER membrane; (3) the synthesized peptide is then transferred to the translocation machinery in the ER membrane and the SRP is released from the SRP receptor. Protein synthesis continues and the polypeptide is transferred directly into the lumen of the ER.

Eukaryotic SRP is a cytosolic ribonucleoprotein complex composed of a 7S RNA and six polypeptide subunits: SRP9, SRP14, SRP19, SRP68, SRP72 and SRP54 (Walter and Blobel, 1980, 1982). The RNA molecule consists of eight helices which form a functional unit (backbone) that is contacted directly by all SRP-proteins, with the exception of SRP54 (Larsen and Zwieb, 1991; Zwieb, 1989). Protein SRP19 is suggested to be associated with the initiation of SRP assembly, as it is the only SRP protein that binds to free SRP RNA independently of other SRP components (Lingelbach et al., 1988). Systematic site-directed mutagenesis with human SRP19 has identified an essential amino acid cluster in five regions of SRP19 that are required for in vitro binding of 7S RNA (Chittenden et al., 1994). RNA-bound SRP19 is also a component required for assembly of SRP54 (Siegel and Walter, 1988b; Zwieb, 1991). The gene coding for SRP19 has been isolated from human (Lingelbach et al., 1988), yeast (Stirling and Hewitt, 1992) and rice (AC X65783). The null mutant of *SRP19* in yeast shows slow growth and is a

* Corresponding author. Present address: Affymetrix, Inc., 3380 Central Expressway, Santa Clara, CA 95051, USA. Tel.: +1 408 7315033; Fax: +1 408 4810422; e-mail: chaoqiang_lai@affymetrix.com

low temperature conditional-lethal (Stirling and Hewitt, 1992).

We identified and cloned a homologue of *SRP19* locus in *Drosophila melanogaster*, while investigating an adjacent new bristle locus, *quemao* (*qm*, Lai et al., 1997). Its chromosomal position and genomic structure were determined. Northern blot analysis suggests the *SRP19* homologue is highly expressed in all stages of *Drosophila* development. Comparison with other *SRP19* proteins from human, rice and two yeast species, reveals that *Drosophila SRP19* shares highly conserved amino acids that are required for 7S RNA binding. This information is of interest to those working in the field.

2. Results and discussion

2.1. Isolation of *SRP19* transcript and RNA blot analysis

The genomic DNA region of 24 kb that contains the *qm* locus was identified from λ Dash SAM genomic DNA library using a probe, the 5.5 kb S/S fragment (see Fig. 1) flanking the site of a *P* element insertion associated with the *qm*⁸⁹ mutation (Lai et al., 1997). The restriction map of the 5.5 S/S region is given in Fig. 1. In situ hybridization experiments map this DNA segment to the 65F polytene chromosomal band of the left arm of the third chromosome (Lai et al., 1997). When the 3.4 kb O/O fragment (see Fig. 1) was used as a probe to screen a cDNA library of the early pupa, four cDNA clones were identified and sequenced. Three of them are identical and are approx. 0.5 kb, while the fourth is identical but truncated at the 5' end. A poly (A)⁺ tail was found at the 3' end of each cDNA clone. The BLAST search of the protein database of the National Center for Biotechnology Information with the predicted protein sequence derived from these cDNA

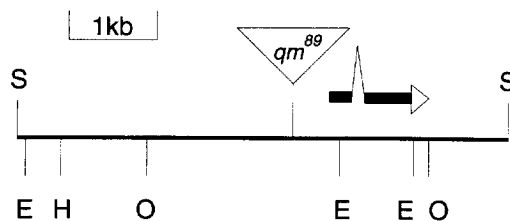


Fig. 1. Map of the *SRP19* gene region. E = *EcoRI*, O = *XhoI*, S = *SacI*. The *SRP19* transcript is indicated as the solid bar broken by the intron. The direction of transcription is indicated by the open arrow head. The position of the *P* element insertion in *qm*⁸⁹ is shown by a triangle. This restriction map of 5.5 S/S fragment is a part of 24 kb genomic DNA including the *SRP19* and *qm* locus that is derived from phage clones isolated from *qm*⁸⁹ and SAM genomic DNA libraries (Lai et al., 1997). **Method:** The cDNA library of the early pupa (5.5–7.5 days) from Dr Thomas Kornberg (University of California at San Francisco) was screened using the 3.4 kb *XhoI* fragment as a probe. cDNA clones and genomic DNA clones were sequenced by dideoxy sequencing using a Sequenase kit (USB).

sequences reveals a significant match (44% identity) to the 19 kDa signal recognition particle protein (*SRP19*) of human. Northern blot analysis using one of the cDNA clones revealed a 0.7 kb transcript (see Fig. 2). In order to identify the complete transcription unit of *SRP19*, the rapid amplification of cDNA end (RACE) method was used to isolate the 5' end sequence of *SRP19* mRNA from the poly(A)⁺ RNA of the early pupa. Five clones that had the largest inserts were isolated and sequenced (see Fig. 3). Among them, one has an insert which is approx. 59 bp longer compared with that of other four clones. It was later determined to carry two anchor primer sequences at the 5' end while others contain one. All clones are similar in length and identical for their sequences (approx. 162 bp excluding the primer), which includes 74 bp of sequence identified in the cDNA clones. This suggests that the remaining 88 bp is the extended sequence at the 5' end of the identified *SRP19* cDNA. Together, the RACE extension and cDNA sequences yield an mRNA length of 658 bp,

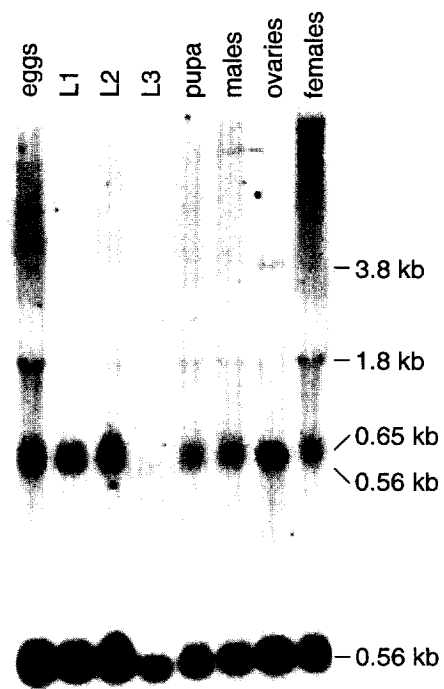


Fig. 2. Northern blot analysis (Sambrook et al., 1989) of *SRP19* transcription. Lanes 1–8 are loaded respectively with approx. 15 μ g of total RNA from the embryo, the first instar larva, the second instar larva, the third instar larva, the pupa, the adult male, the ovary and the adult female without ovaries. They were probed with the isolated 0.6 kb *SRP19* cDNA clone. The sizes of standard RNA markers are indicated to the right of the blot. There is a single 0.7 kb transcript detected in lanes 1–8, with a weak signal in lane 4, the third instar larva. This is probably due to degradation of RNA (see the reduced control signal in the lower panel), as in a separate Northern blot (not shown), the third instar larva shows a high expression of *SRP19*. An additional transcript of 3.8 kb was detected in the ovaries only. Lower panel: for normalization of loading, the same blot was reprobed with *Drosophila* ribosomal protein 49 gene (*RpL32*, O'Connell and Rosbash, 1984).

-333 Vattgaagttcttccgcgccctgcatttgagatgatcgcaagtgcgatgt
 -283 ttctatcgatagggttagggggtttttcttagaggtatcatcgatataga
 -253 caagagatgggaaatgtggcttaatagatgttggctatttccgccaactt
 -203 ctattaaaccagataagccaaaataaataatcctaataaattttattt
 -153 catgtactctttaaatttaagtcgcttataaaataacctcttaacgtagtag
 -103 ccagaaaagccatataaattatggcttctctcgctaaatatacgattact
 -27 cgggtccagctgtcatcgtgccatccctATTCAATTGGGTTGCTTGGGTG
 24 GAAAAATAGCTTCTAGCTCAATTTTGTAAATTACACAGGAATGCAATA
 75 AACCAACCAAGAAATTCACCTGAAAGAAATCGCCGACAGGTTCCCATGAAAAA
 M A T G S H E K

125 AACTGGAGCCCCAGCATGAAACACAACGATATGGAGCGGTAGGAACCCAGg
 K L E P Q H E T Q R Y G A V G T R

175 cactggccaacaaaactcaaagtctgatcttgcgatgtaactctttattgtt

225 gttaacacagATGGATTTCATATATCCCGCTACATCAATAGGAAAAAG
 W I C I Y P A Y I N R K K

275 ACACGCCAGGAGGGTCCGACGGCTGCCAAGGAGAACTGTGTAGACAATCC
 T R Q E G R R L P K E N C V D N P

325 CAGCTATATTGAGATTTCGCGATGCTGTGTCCGTTTCCAACCTGTCAGTTCC
 S Y I E I R D A V S V S N L Q F

375 TCATGGAGAACAAGAAGTACTGTCCGAGACAGCAGCGAGATGGAGTTC
 L M E N K K Y C R E N S S E M E F

425 CGTGGTCTGTGCGCGTCCAGCTCGCTAATGTCGATGGCACTTTGTACAA
 R G R V R V Q L R N V D G T L Y N

475 TATTGACTTTCACCGCGGATCCATGCTGCATATCGCCAGCAAGA
 I D F P T R E S I M L H I A S K

525 TACCGCAGCTGAAGACGCGACAGAACAATCCGGCGACTCGTACCACCAG
 I P Q L K T R Q N K S G D S Y H Q

575 CAGTCGCAGCCACAGTCAAGTGCATCCGGATCTGGGGTGGTGGTGGCGG
 Q S Q P Q S N A S G S G G G G G G

625 CAAGAAGGGAAGGGCAAGCGCCGCTAATGACTTTCAGTGGCAACTCGTTGT
 K K G K G K R R *

675 TTGGTTTCTTTGTAAATAAACCGCTGTAACAAA

Fig. 3. DNA and amino acid sequences of the *SRP19* locus of *D. melanogaster*. The GEN BANK accession number for *SRP19* is U35682. The inferred start codon (ATG) is underlined, and a putative TATA box (TATAAAT) is underlined and in lower-case letters. The open reading frame is indicated by the roman upper-case letters, whereas the intron and the 5' untranslated sequences are in lower-case letters. The 5' and 3' untranslated regions are in italics and upper case. The predicted amino acid sequence is given in the upper-case letter under the DNA sequence positioned at the centre of each codon. * is the stop codon. The *P* element (inverted small triangle) was inserted after position -333. **Method:** Total RNA of pupae was extracted by the phenol method (Jowett, 1986). mRNA was isolated using an mRNA Separator Kit from Clontech. RACE (rapid amplification of 5'-cDNA end) method was described by Frohman et al. (1988) and Belyavsky et al. (1989). cDNA was reverse-transcribed from mRNA of the early pupae using an antisense *P*₁ primer (5'-CGAGT-CGCCGATTTGTTCTGTCTCGCTCT-TCAGC-3'), then ligated to an anchor sequence (5'-CACGAATTCACCTA-TCGATTCTGGA-ACCTTCAGAGG-3'). The cDNA sequence was then amplified by PCR using the anchor primer (5'-CTGGTTCGGCCACCT-CTGAAGGTTCCA-GAATCGATAG-3') and a second antisense primer *P*₂ (5'-CTGGATCCATATC-GTTGTGTTTCATGC-3'), which is upstream of *P*₁. The PCR product was cloned into the *Sma*I site of the Bluescript vector for sequencing. The above procedures were performed using 5'-AmpliFINDER RACE kit from Clontech and SureClone Ligation Kit from Pharmacia Biotech.

which is approximately the same size as the transcript detected in Northern blot analysis (see below).

To determine the temporal pattern of *SRP19* expression, Northern blots of stage-specific total RNA were hybridized with *SRP19* cDNA (Fig. 2). It is clear that

the 0.7 kb *SRP19* transcript is abundant at all stages. This is consistent with expectations that *SRP19* is an essential 'house-keeping' gene. Interestingly, one distinct transcript of 3.8 kb was found in ovaries, but not in other developmental stages. This transcript has not yet been characterized.

2.2. Structure of the *SRP19* locus

Southern blot analysis (Sambrook et al., 1989) of the 3.4 kb O/O fragment indicates that the *SRP19* transcript is located in the 0.7 kb E/E and 0.95 kb E/P fragments (data not shown). The whole 0.7 kb E/E fragment and part of 0.95 kb E/P fragment were sequenced. The resultant sequence was aligned with that of the *SRP19* transcript. One small intron of 61 bp was identified between position 174 and 234. The direction of transcription and the intron of the *SRP19* locus are indicated in Figs. 1 and 3. The *P* element was inserted 333 bp 5' of the apparent start of *SRP19* transcription (see Fig. 3). The *P* element insertion has no detectable effect on the expression of *SRP19* (data not shown).

The identified *SRP19* transcript contains a single open reading frame of 492 bp, an untranslated region of 169 bp at the 5' end and an untranslated region of 76 bp at the 3' end (Fig. 3). The putative start codon (ATG) is characterized by a 5' flanking sequence (AAGA), which is similar to the *Drosophila* consensus sequence (C/A AA A/C) before the start codon (Cavener, 1987). A putative TATA box (TATAAAT) was located between residues -122 to -128 and is embedded in an AT-rich region. These structural characteristics and the amino acid sequence of the ORF support the conclusion that the identified gene is the true transcript of *SRP19* in *Drosophila*. Furthermore, the codon usage bias of this reading frame is consistent with the pattern of *D. melanogaster* (data not shown). The predicted *SRP19* protein is a 163 amino acid polypeptide and its estimated molecular weight is 18.5 kDa.

2.3. Comparison with other *SRP19* proteins

An alignment of *Drosophila* *SRP19* protein with the four known *SRP19* proteins from human (*Homo sapiens*), rice (*Oryza sativa*) and two yeasts (*Saccharomyces cerevisiae* and *Yarrowia lipolytica*) showed that *SRP19* in *D. melanogaster* shares respectively 44%, 29%, 17% and 19% identical amino acid sequences with *SRP19* of these species. *SRP19* of human, *Drosophila* and rice are more similar to each other (on average 37% identity). Surprisingly, the *SRP19* sequences of two yeast species are less similar (25%). Note also that human, *Drosophila* and rice have a shorter *SRP19* than yeast.

Among the six protein subunits of the SRP, *SRP19* and *SRP54* are monomers, while the other four proteins (*SRP68*, *SRP72*, *SRP9* and *SRP14*) form heterodimers:

SRP68/72 and SRP9/14 (Siegel and Walter, 1988a; Strub et al., 1991; Janiak et al., 1992). In the process of SRP assembly, heterodimers and SRP19 bind to the 7S RNA molecule directly, while the binding of SRP54 occurs only after SRP19 is already bound to the RNA. Therefore, SRP19 is particularly interesting as it plays a critical role in the assembly of SRP. The 7S RNA consists of eight helices. Helices 6 and 8 are the major binding sites of SRP19 (Zwieb, 1991, 1994). The study of the in vitro interaction between 7S RNA and human SRP19 (Zwieb, 1991, 1994; Chittenden et al., 1994) using site-directed mutagenesis has revealed that 60 of 144 (42%) human SRP19 amino acid residues were required for binding to the 7S RNA molecule. Most of these critical amino acid sites cluster in the five conserved regions shown in Fig. 4. All species are identical at 17 amino acid residues (14 in the domains shown in Fig. 4). Eleven of these 17 strictly conserved residues were shown to be required for binding of 7S RNA. Other residues (not shown in Fig. 4) located at the N- and C-termini and a small region adjacent to the C-terminal (Chittenden et al., 1994) are also essential for the proper RNA binding of SRP19. SRP19 contains a very hydrophilic C-terminal domain of seven lysine residues interrupted by two glycine residues, that is likely to be on the exposed surface (Zwieb, 1991). Furthermore, we have also shown that this lysine-rich C-terminus is also highly conserved across species (data not shown).

Comparison of five SRP19 amino acid sequences from human, *Drosophila*, rice and yeast indicates that these conservative regions are generally consistent with the predicted RNA binding sites (see Fig. 4). However, there are three conserved sites (two glycines at positions 43 and 90, and one glutamate at 42) across the five species that were not identified as necessary for RNA binding by site-directed mutagenesis. The simplest explanation for this result seems to be that the glycines at

```

RNA-bnd      ***** ** ** *****
Dro.mel.    (25)WICLYPAYINRKKTRQEGRR (5)CVDNPSYIEIRDAVS (8)
Hom.sap.    (14)oICLYPAYoNxxKKTxxEGRR (5)xVoNPoxxEIXDoxS (8)
Ory.sat.    (12)WxxIYPoYoNxxKKTxxEGRR (5)xxxxPoxxEIXDxxS (8)
Sac.cer.    (97)oxxoYPxYxxxxoxxEGRR (5)xVoNPxxxxxxDAVx (8)
Yar.lip.    (120)WxxoYPoYxxoxKTxEGRR (5)xVoNPxxxxIxox (8)

RNA-bnd      **** ** ***** *****
Dro.mel.    ENKK-YCRE (8)GRVVRQLRNVGD (5)DFPTRESIMLHIASK
Hom.sap.    ExxKxYoRE (8)GRVVRQLoxxDG (5)xFPoRxSoMLxoAxx
Ory.sat.    ExxKxYxRo (4)GRVVRQLoxxDG (5)xxxToxxoMoxIAxx
Sac.cer.    ExxKxxxxxo (4)GRoRVxxoxxoxx (6)xFxoooooxxoxxoxx
Yar.lip.    ExoKxxxxxo (4)GRVVRQLoxxxxo (6)xxxxoooooxxoxx

```

Fig. 4. The conserved domains in SRP19 of *Drosophila melanogaster* (AC U35682), *Homo sapiens* (AC X12791), *Oryza sativa* (AC U19030), *Saccharomyces cerevisiae* (AC X65783) and *Yarrowia lipolytica* (AC Z22570). The five conserved regions that are required for RNA binding are interrupted by a number of amino acids in parentheses. The shared amino acids among five species are indicated by repeating the letter code of the amino acid, o=chemically similar amino acid, x=dissimilar amino acid, —=deletion. The alignment were performed using CRUSTAL W (Higgins and Sharp, 1989). The sites in *Homo sapiens* required for in vitro binding to the 7S RNA (Zwieb, 1994) are indicated with "*".

positions 43 and 90 were not mutated, as glycine was used as a replacement for all experimentally mutated amino acid sites. At other less conserved sites at which the human residues were shown to be required for RNA binding, chemical properties of the various residues are often conserved (see Fig. 4).

Acknowledgement

We want to thank Christian Zwieb for the unpublished *SRP19* locus sequence of rice and helpful comments on this paper. We also thank Ulgo Mazina for the *Drosophila* Northern blot.

References

- Belyavsky, A.T., Vinogradova, T., Rajewsky, K., 1989. PCR-based cDNA library construction: general cDNA libraries at the level of a few cells. *Nucleic Acids Res.* 17, 2919–2932.
- Cavener, D.R., 1987. Comparison of the consensus sequence flanking translational start sites in *Drosophila* and vertebrates. *Nucl. Acids Res.* 15, 1353–1361.
- Chittenden, K., Black, S.D., Zwieb, C., 1994. Systematic site-directed mutagenesis of protein SRP19—identification of the residues essential for binding to signal recognition particle RNA. *J. Biol. Chem.* 269, 20497–20502.
- Frohman, M.A., Dush, M.K., Martin, G.R., 1988. Rapid production of full-length cDNA from rare transcripts: amplification using single gene-specific oligonucleotide primers. *Proc. Natl. Acad. Sci. USA* 85, 8998–9002.
- Higgins, D.G., Sharp, P.M., 1989. Fast and sensitive multiple sequence alignments on a microcomputer. *Comput. Appl. Biosci.* 5, 151–153.
- Janiak, F., Walter, P., Johnson, A.E., 1992. Fluorescent-detected assembly of the signal recognition particle: the binding of the two SRP protein heterodimers to SRP RNA is noncooperative. *Biochemistry* 31, 5830–5840.
- Jowett, T., 1986. Preparation of nucleic acids. In: Roberts, D.B. (Ed.), *Drosophila, A Practical Approach*. RRL Press, Oxford, pp. 275–286.
- Lai, C.Q., McMahon, R., Young, C., Mackay, T.F., Langley, C.H., 1997. *quemo*, a *Drosophila* bristle locus, codes for geranylgeranyl pyrophosphate synthase. *Genetics*, submitted.
- Larsen, N., Zwieb, C., 1991. SRP-RNA sequence alignment and secondary structure. *Nucleic Acids Res.* 19, 209–215.
- Lingelbach, K., Zwieb, C., Webb, J.R., Marshallsay, C., Hoben, P.J., Walter, P., Dobberstein, B., 1988. Isolation and characterization of a cDNA clone encoding the 19 kDa protein of signal recognition particle (SRP): expression and binding to 7SL RNA. *Nucleic Acids Res.* 16, 9431–9442.
- Nunnari, J., Walter, P., 1992. Protein targeting to and translocation across the membrane of the endoplasmic reticulum. *Curr. Opin. Cell Biol.* 4, 573–580.
- O'Connell, P.O., Rosbash, M., 1984. Sequence, structure, and codon preference of the *Drosophila* ribosomal protein 49 gene. *Nucleic Acids Res.* 16, 5495–5513.
- Rapoport, T.A., 1990. Protein transport across the ER membrane. *Trends Biochem. Sci.* 15, 355–358.
- Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Siegel, V., Walter, P., 1988a. Each of the activities of signal recognition

- particle (SRP) is contained within a distinct domain: analysis of biochemical mutants of SRP. *Cell* 52, 39–49.
- Siegel, V., Walter, P., 1988b. Binding sites of the 19-kDa and 68/72-kDa signal recognition particle (SRP) proteins on SRP RNA as determined by protein–RNA "footprinting". *Proc. Natl. Acad. Sci. USA* 85, 1801–1805.
- Stirling, C.J., Hewitt, E.W., 1992. The *S. cerevisiae sec65* gene encodes a component of yeast signal recognition particle with homology to human SRP19. *Nature* 356, 534–537.
- Strub, K., Moss, J., Walter, P., Binding sites for the 9-kilodalton and 14 kilodalton heterodimeric protein subunit of the signal recognition particle (SRP) are contained exclusively in the Alu domain of SRP-RNA and contain a sequence motif that is conserved in evolution. 1991. *Mol. Cell Biol.* 11, 3949–3959.
- Walter, P., Blobel, G., 1980. Purification of a membrane-associated protein complex required for protein translocation across the endoplasmic reticulum. *Proc. Natl. Acad. Sci. USA* 77, 7112–7116.
- Walter, P., Blobel, G., 1982. Signal recognition particle contains a 7S RNA essential for protein translocation across the endoplasmic reticulum. *Nature* 299, 691–698.
- Walter, P., Lingappa, V.R., 1986. Mechanism of protein translocation across the endoplasmic reticulum. *Annu. Rev. Cell Biol.* 2, 499–516.
- Wolin, S.L., 1994. From the elephant to *E. coli*: SRP-dependent protein targeting. *Cell* 77, 787–790.
- Zwieb, C., 1991. Interaction of protein SRP19 with signal recognition particle RNA lacking individual RNA-helices. *Nucleic Acids Res.* 19, 2955–2960.
- Zwieb, C., 1994. Site-directed mutagenesis of signal-recognition particle RNA-identification of the nucleotides in helix 8 required for interaction with protein SRP19. *Eur. J. Biochem.* 222, 885–890.
- Zwieb, C., 1989. Structure and function of signal recognition particle RNA. *Prog. Nucleic Acids Res. Mol. Biol.* 37, 207–234.