Nucleic Acids Research, 1994, Vol. 22, No. 11 1933–1947

Molecular evolution of SRP cycle components: functional implications

Steven Althoff, David Selinger+ and Jo Ann Wise*
University of Illinois, Department of Biochemistry, 600 S. Matthews Avenue, Urbana, IL 61801, USA

Received March 4, 1994; Accepted April 13, 1994

ABSTRACT

Signal recognition particle (SRP) is a cytoplasmic ribonucleoprotein that targets a subset of nascent presecretory proteins to the endoplasmic reticulum membrane. We have considered the SRP cycle from the perspective of molecular evolution, using recently determined sequences of genes or cDNAs encoding homologs of SRP (7SL) RNA, the Srp54 protein (Srp54p), and the α subunit of the SRP receptor (SRα) from a broad spectrum of organisms, together with the remaining five polypeptides of mammalian SRP. Our analysis provides insight into the significance of structural variation in SRP RNA and identifies novel conserved motifs in protein components of this pathway. The lack of congruence between an established phylogenetic tree and size variation in 7SL homologs implies the occurrence of several independent events that eliminated more than half the sequence content of this RNA during bacterial evolution. The apparently non-essential structures are domain I, a tRNA-like element that is constant in archaea, varies in size among eucaryotes, and is generally missing in bacteria, and domain III, a tightly base-paired hairpin that is present in all eucaryotic and archeal SRP RNAs but is invariably absent in bacteria. Based on both structural and functional considerations, we propose that the conserved core of SRP consists minimally of the 54 kDa signal sequence-binding protein complexed with the loosely base-paired domain IV helix of SRP RNA, and is also likely to contain a homolog of the Srp68 protein. Comparative sequence analysis of the methionine-rich M domains from a diverse array of Srp54p homologs reveals an extended region of amino acid identity that resembles a recently identified RNA recognition motif. Multiple sequence alignment of the G domains of Srp54p and SRα homologs indicates that these two polypeptides exhibit significant similarity even outside the four GTPase consensus motifs, including a block of nine contiguous amino acids in a location analogous to the binding site of the guanine nucleotide dissociation stimulator (GDS) for *E.coli* EF-Tu. The conservation of this sequence, in combination with the results of earlier genetic and biochemical studies of the SRP cycle, leads us to hypothesize that a component of the Srp68/72p heterodimer serves as the GDS for both Srp54p and SRα. Using an iterative alignment procedure, we demonstrate similarity between Srp68p and sequence motifs conserved among GDS proteins for small *Ras*-related GTPases. The conservation of SRP cycle components in organisms from all three major branches of the phylogenetic tree suggests that this pathway for protein export is of ancient evolutionary origin.

INTRODUCTION

A model for the role of a cytoplasmic ribonucleoprotein known as signal recognition particle (SRP) as an adaptor between the soluble protein synthetic machinery and the translocation apparatus of the endoplasmic reticulum (ER) membrane emerged from studies with a heterologous *in vitro* assay system (for reviews, see Refs. 1–6). The SRP cycle is an ordered series of events that begins with recognition of proteins destined for export through specific binding of the particle to ER-specific signal sequences as they emerge from the ribosome (7), triggering a transient pause in elongation (8, 9). At the ER, SRP interacts with an integral membrane protein known as the SRP receptor (10) or docking protein (11) and the signal sequence is released, probably to a component of the translocation machinery (12). The ribosome-nascent chain complex then forms a tight junction with membrane proteins (13) and the presecretory protein is translocated co-translationally into the ER lumen (14). Finally, SRP is released through a GTP-dependent process (15, 16) whose details remain to be elucidated (see below).

Canine signal recognition particle, the first example to be discovered and by far the most extensively characterized, is composed of one RNA molecule (originally designated 7SL but now generally referred to as SRP RNA) and six polypeptides organized into two heterodimeric (Srp9/14p and Srp68/72p) and two monomeric (Srp19p and Srp54p) proteins (17, 18, 19). Through chemical modification and photocrosslinking

*To whom correspondence should be addressed

+Present address: Institute of Molecular Biology, University of Oregon, Eugene, OR 97403, USA

experiments combined with *in vitro* reconstitution assays, general functions have been assigned to each of the SRP proteins (reviewed in 20). First, particles lacking the Srp9/14 heterodimeric protein do not display elongation arrest activity but still promote translocation (19). Second, alkylation of the Srp68/72 heterodimer produces a particle competent to arrest translation, yet unable to mediate translocation into the ER (21). Third, the absence of both activities in SRP reconstituted with modified Srp54p (21), in combination with the observation that this protein can be cross-linked to a nascent polypeptide (22, 23), implies that signal sequence recognition is a prerequisite for translation arrest and translocation promotion. Srp19p has not been ascribed a function other than SRP RNA binding, although *in vitro*, and probably *in vivo* as well, it facilitates assembly of Srp54p with the RNA (18, 24, 25, 26). While it seems likely that SRP RNA plays a role in the SRP cycle beyond orienting the protein subunits, no specific function has been proven experimentally.

The recent emergence of sequence data for homologs of SRP RNA, the Srp54 protein and the α subunit of the SRP receptor from a broad spectrum of organisms, as well as for the remaining five polypeptides of mammalian SRP, prompted us to examine in detail the molecular evolution of SRP cycle components. Our analysis provides new insights into the functional significance of structural differences between bacterial, eucaryotic and archaeal homologs of SRP RNA, and identifies sequence motifs in the protein components of the SRP cycle that lead us to propose novel roles and interactions. Finally, we review genetic and biochemical data suggesting that this pathway for protein export is of ancient evolutionary origin.

RESULTS AND DISCUSSION

Evolution of SRP RNA

SRP RNA can be divided into four structural domains. To gain insight into the function of SRP RNA, we undertook a comparison of homologs from all three major branches of the evolutionary tree, the eucaryotes, the archaea (formerly designated archaebacteria), and the bacteria (formerly designated eubacteria) (nomenclature according to 27). The original phylogenetic comparison which led to the proposal that SRP RNA folds into a conserved structure included only two sequences (*Drosophila melanogaster* and *Homo sapiens*) (28). Subsequent evolutionary comparisons of SRP-like RNAs from a broader spectrum of organisms indicated that this structure can be divided into four domains (29). As illustrated in the schematic representation of the human homolog shown at the top of Figure 1, domain I is a tRNA-like structure (30) formed by sequences near the 5′ end; domain II is a long helix in which sequences immediately downstream from domain I are paired with the 3′ end of the RNA; and domains III and IV are hairpin structures, the first highly base paired and the second interrupted by several internal loops. We have adopted the nomenclature proposed by Poritz *et al.* (28), which reflects the functional domains of SRP as defined by the activities of the proteins to which the RNA is bound, rather than an alternative system that assigns each helix a separate number (31). Because a comprehensive review of the sequence and structure of SRP-like RNAs was published relatively recently (31), the following discussion focuses on novel findings that emerged from our analysis and on functional aspects not covered in the earlier survey.

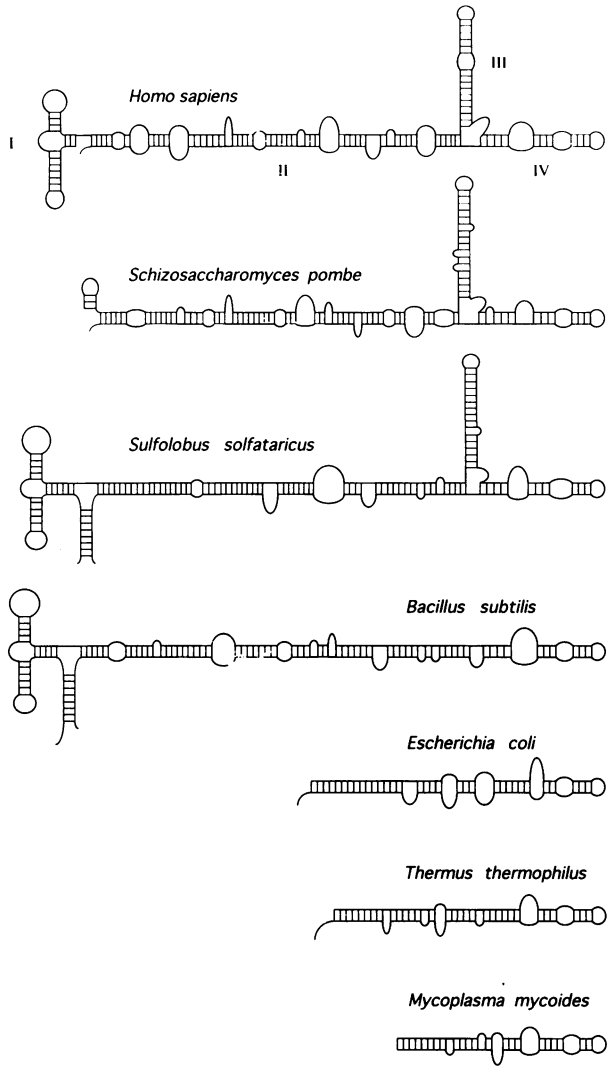

Figure 1. Evolutionary comparison of SRP RNA secondary structures. This figure shows the structures of SRP RNA homologs from two representative eucaryotes, *Homo sapiens* and *Schizosaccharomyces pombe*, the archaeum *Sulfolobus sulfataricus*, and the bacteria *Bacillus subtilis*, *Escherichia coli*, *Thermus thermophilus*, and *Mycoplasma mycoides*. The *H.sapiens*, *S.pombe*, *B.subtilis* and *E.coli* structures were adapted from Poritz *et al.* (29), and the remainder were derived on the basis of information in the original paper describing each RNA. References: *S.solfataricus* (47); *T.thermophilus* (36), and *M.mycoides* (34).

SRP RNA has undergone a dramatic size reduction several times during bacterial evolution. Each of the eight bacterial RNAs homologous to mammalian 7SL is missing one or more structural domains; five examples, including the earliest recognized member of the family, 4.5S RNA from *Escherichia coli*, are depicted schematically in Figure 1. *E.coli* 4.5S is 114 nucleotides in length and folds into a single hairpin whose apical portion displays significant sequence identity with domain IV of human SRP RNA (29, 32, 33; see below); however, this RNA lacks structures resembling domains I and III, and more than half of domain II is also missing. Shown at the bottom of Figure 1 is the even smaller (77 nucleotides) *Mycoplasma mycoides* homolog (34); a 79 nucleotide RNA from *M.pneumoniae* (not shown) also folds into a structure similar to *E.coli* 4.5S (35). At the other size extreme (271 nucleotides) is the SRP-like RNA from *Bacillus*

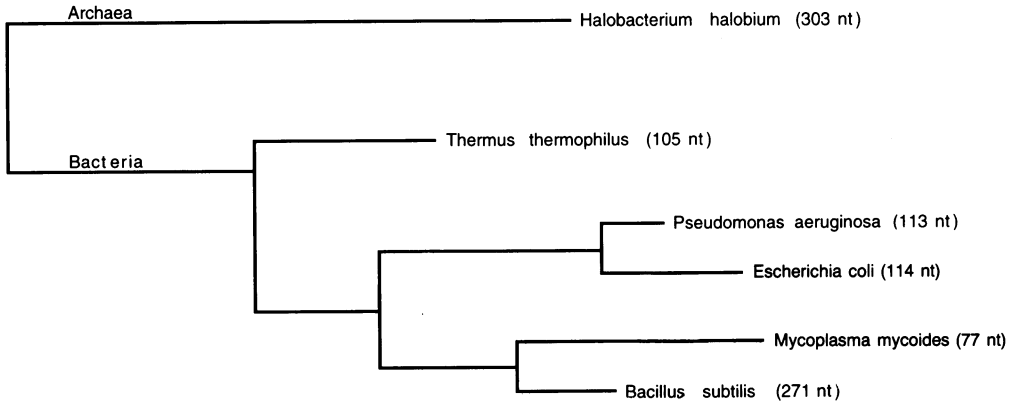

Figure 2. Evolutionary relationships and size variation among bacterial SRP RNAs. Shown is a phylogenetic tree based on comparative sequence analysis of 16S ribosomal RNA (obtained from the Ribosomal RNA Database Project; 39) that includes bacteria from which SRP RNAs have been sequenced, as well as a representative archaeum. In this unrooted tree, derived by distance matrix methods, the horizontal component of separation represents the evolutionary distance between organisms. The number in parentheses following each species name denotes the length of the SRP-like RNA from that organism. References: *H.halobium* (42); *T.thermophilus* (36); *E.coli* (32); *M.mycoides* (34), and *B.subtilis* (33).

subtilis (middle, Figure 1), which contains structures corresponding to domains I, II and IV of the mammalian RNA. Also depicted in Figure 1 is a 105 nucleotide SRP-like RNA from the extreme thermophile *Thermus thermophilus* (36) which, like the 113 nucleotide 4.5S RNA from *Pseudomonas aeruginosa* (37; not shown), can be folded into an unbranched secondary structure similar to *E.coli* 4.5S RNA. In addition to these well-characterized homologs, DNA fragments that could specify transcripts of ca. 100 nucleotides with the potential to fold into a single hairpin have been isolated from the bacteria *Legionella pneumophila* and *Micrococcus lysodeikticus* (38) by complementation of 4.5S depletion in *E.coli*.

Two distinct evolutionary scenarios might account for the small sizes of bacterial SRP RNA homologs. The first and, we believe, most likely explanation is that they represent truncated forms of an ancestral molecule similar in size and structure to the RNA found in modern eucaryotes and archaea. The second possibility, that an RNA similar to the bacterial RNAs is the predecessor of full-sized homologs, is rendered implausible by the presence of all but domain III in the SRP-like RNA from *B.subtilis*. In addition, this RNA exhibits a domain I tertiary pairing common to homologs from the other two branches of the phylogenetic tree and contains a terminal helix also found in archael 7S RNA (Figure 1 and see below). Since it is unlikely that these structures were acquired independently, the bacterial 7SL homologs must be derived from a molecule containing at least three of the four domains of mammalian SRP RNA. While these observations would predict that the *B.subtilis* RNA is most closely related to the common ancestor of the bacterial SRP-like RNAs and the *Mycoplasma* RNAs are most extensively diverged, comparison of the size distribution of bacterial SRP RNAs to a phylogenetic tree based on 16S ribosomal RNA (39) does not support this hypothesis (see Figure 2). In fact, the deepest branching bacterium from which SRP RNA has been sequenced, *T.thermophilus*, contains an RNA slightly smaller than *E.coli* 4.5S. This branch is followed by a division between the group represented by *P.aeruginosa* and *E.coli* and the one containing *B.subtilis* and *M.mycoides*. Because both lineages that diverged from the branch leading to *B.subtilis* culminate in organisms with truncated homologs, the rRNA phylogeny indicates that SRP RNA was reduced in size via at least three independent events during bacterial evolution. The extremely similar sizes of SRP-like RNAs in bacteria that are closely related based on the 16S rRNA phylogeny (the two purple bacteria, *E.coli* and *P.aeruginosa*, and the two *Mycoplasma* species) provides further support for this scenario. Although the simplest explanation for the currently available data is that most bacteria have adapted to life with a 'stripped-down' version of SRP RNA, an interesting but as yet unsupported hypothesis is that the missing domains are encoded by separate loci (40). The co-purification of a tRNA-sized RNA with trypanosomal SRP lends some credence to this proposal (41). Regardless of the precise mechanism by which entire domains of the ancestral SRP-like RNA were eliminated, the fact that similar events apparently occurred independently in multiple bacterial lineages implies that truncation confers a significant selective advantage in this phylogenetic group.

7S RNAs from archaea closely resemble mammalian SRP RNA. In contrast to *E.coli* 4.5S, an RNA discovered nearly a decade ago in the archaeum *Halobacterium halobium* (42) was immediately identifed as a homolog of human 7SL based on its similarity in both size and secondary structure. More recently, homologous RNAs have been sequenced from eight other archaeal species, including *Methanobacterium thermoautotrophicum* (43), *Methanococcus voltae* (44), *Pyrodictium occultum* (45), *Methanothermus fervidus* (46), *Archaeglobus fulgidus*, *Methanosarcina acetivorans*, *Sulfolobus solfataricus*, and *Thermococcus celer* (47). A gel electrophoretic survey of additional archaea revealed that all species examined contain a prominent 7S non-ribosomal RNA (48). In contrast to the diversity in size found among bacterial SRP-like RNAs, the archaeal 7S RNAs are quite homogeneous, with all known examples containing approximately 300 nucleotides. These RNAs, represented in Figure 1 by *S.solfataricus* 7S, can be folded into a series of helices that are virtually identical to the secondary structure of human SRP RNA. The only notable difference is that, in addition to the four domains common to eucaryotic homologs, archaeal 7S RNAs contain an extra terminal helix formed by pairing of their 5′ and 3′ ends. This structure, which is shared with the *B.subtilis* RNA, may represent a processing

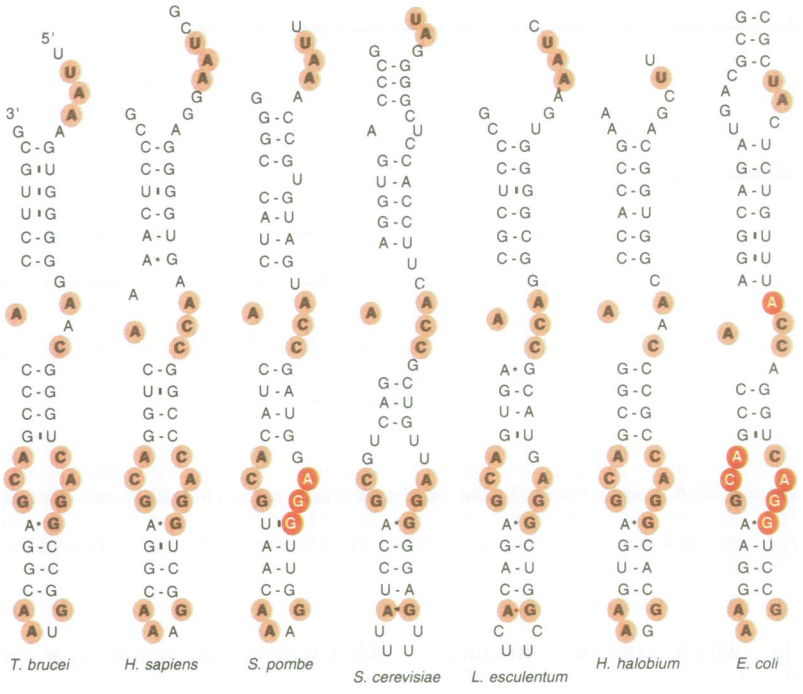

Figure 3. Evolutionary comparison of SRP RNA domain IV structures. For each RNA, the 5′ nucleotide (right) of the segment shown is the one just following the terminal domain III pair and the 3′ nucleotide (left) is the one just preceding the teminal pair in the domain II helix. The examples shown were selected to demonstrate several invariant and group-specific features of the domain IV structure, as well as to represent the phylogentic diversity of organisms from which SRP RNA has been sequenced. The *Trypanosoma brucei* sequence and structure were reproduced from Michaeli *et al.* (58); this is the most deeply rooted eukaryote (157) from which SRP RNA has been sequenced. The *H. sapiens* SRP RNA sequence was determined by Ullu *et al.* (49) and the *Lycopersicon esculentum* (tomato) sequence by Haas *et al.* (46); their domain IV structures were modified from those originally published (28, 46) to more closely resemble other eucaryotic SRP RNAs. The sequence and structure of *S.pombe* SRP RNA were determined by Brennwald *et al.* (52), Poritz *et al.* (53) and Ribes *et al.* (54). The sequence of *S.cerevisiae* scR1 RNA was determined by Felici *et al.* (55) and the domain IV structure derived in the present work as described in the text. The *H.halobium* (archaeal) 7S RNA sequence was determined by Moritz and Goebel (42) and the *E.coli* (bacterial) 4.5S sequence by Hsu *et al.* (32). Highly conserved nucleotides are shown in boldface type within a light red circle. Nucleotides in the *S.pombe* and *E.coli* RNAs that are known to be important for binding to the Srp54 protein (26, 69) are shown in reverse type within a bright red circle. Accession numbers for SRP RNA sequences can be found in Ref. (31).

site (45). Within the archaea, there is remarkable conservation of even minute details of the secondary structure, e.g. the positions and sizes of interior loops within domain II (47). This structural homogeneity is somewhat surprising in light of the phylogenetic diversity of the species from which homologs have been isolated, which includes representatives of each of the three major branches of the archaeal tree (47).

Structural domains of SRP RNA that are absent in bacteria contribute to function in eucaryotes. In the decade since the discovery that the RNA component of canine SRP is nearly identical to previously sequenced cytoplasmic RNAs from rat and HeLa cells (17, 49, 50), homologs have been identified from a variety of eucaryotic species including an insect, *Drosophila melanogaster* (28, 51); three yeasts, *Schizosaccharomyces pombe*, *Yarrowia lipolytica* and *Saccharomyces cerevisiae* (52, 53, 54, 55, 56), two protists, *Tetrahymena thermophila* and *Trypanosoma brucei* (57, 58) and the plants *Zea mays* (corn), *Triticum aestivum* (wheat), *Lycopersicon esculentum* (tomato), and *Cineraria hybrida* (46, 59, 60; unpublished data cited in 31). Although these RNAs share only limited sequence identity with mammalian 7SL RNA, with the exception of scR1 (small cytoplasmic RNA 1) from *S.cerevisiae* (see below), they are similar in size (250−300 nucleotides) and can be folded into the conserved secondary structure deduced by phylogenetic comparison (see Figure 1). In contrast to the truncated versions of SRP RNA found in bacteria, all four structural domains are invariably present even in homologs from rapidly dividing eucaryotes. Moreover, the results of mutagenesis experiments carried out principally on SRP RNA from the fission yeast *S.pombe* indicate that domains which are absent in bacteria nonetheless play critical roles in at least one eucaryote. Consistent with the pattern of primary sequence conservation, the majority of point mutations deleterious to function cluster in domain IV, although lethal or conditional mutations have been identified throughout SRP RNA (26, 61−63). The effects of mutating the putative binding site of a fission yeast Srp9/14 homolog indicate that both the identity of a conserved residue and the structure within which it is embedded are important (62). In both *Y.lipolytica* and *S.pombe* SRP RNA, mutating the domain III tetranucleotide loop (consensus GNAR) produces conditional lethality (62, 64, 65). Moreover, a precise deletion of domain III, which is absent in all bacterial SRP-like RNAs, is lethal in fission yeast (66). Although only a few point mutations have been characterized in domain II, which was partially excluded from the random mutagenesis (62), a 131 nucleotide insertion near its junction with domain III produces temperature-sensitivity (62). In combination, these observations indicate that all four structural domains of SRP RNA are necessary for wild-type growth in eucaryotes.

Domain IV of SRP RNA is universally conserved. As illustrated in Figure 3, the loosely base paired domain IV hairpin contains

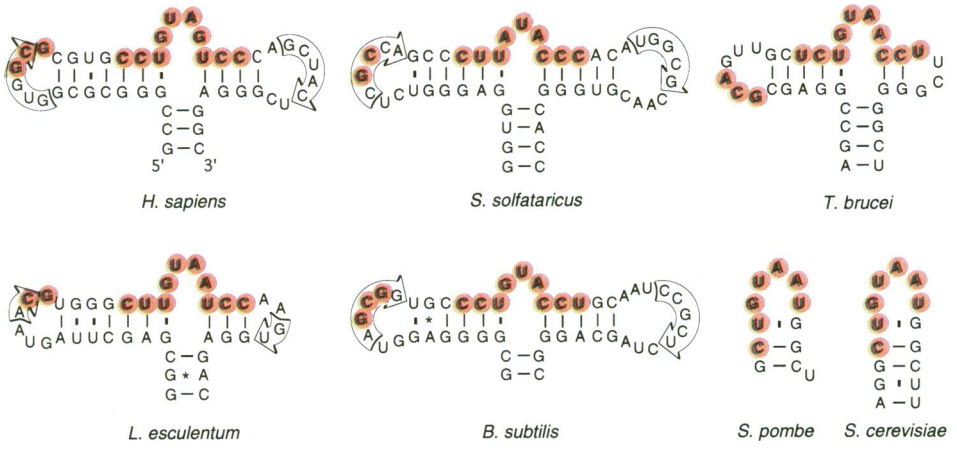

Figure 4. Evolutionary comparison of SRP RNA domain I structures. For each RNA, the segment shown extends from the 5′ nucleotide of the first potential Watson–Crick base pair to its 3′ partner. The tomato (*L.esculentum*), *B.subtilis*, and *T.brucei* structures were modified from those originally published to maximize placement of conserved nucleotides in similar secondary structure contexts. Highly conserved nucleotides are shown in boldface type and surrounded by a light red circle. Nucleotides enclosed within arrows in each of the two hairpin loops have the potential to form tertiary interactions through Watson–Crick base pairing. References: *H.sapiens* (49); *S.solfataricus* (47); *L.esculentum* (46); *T.brucei* (58); *S.pombe* (52–54); *S.cerevisiae* (55); *Y.lipolytica* (53). Accession numbers for SRP RNA sequences can be found in Ref. (31).

a significant number of nucleotides that are identical among SRP-like RNAs from all three major phylogenetic groups. Recent data indicate that the conservation of this region can be attributed at least part to the presence of the binding site for the Srp54 protein, which plays a central role in the function of signal recognition particle (see Introduction). While the ability of the human protein to bind *E.coli* 4.5S RNA (67, 68), which extends little beyond domain IV, was suggestive, more compelling evidence was provided by the observation that mutating conserved internal loop nucleotides in this region of both *E.coli* and *S.pombe* SRP RNAs disrupts assembly with the homologous Srp54 protein (69, 26). The identification of nucleotides required for SRP54p binding to bacterial 4.5S RNA relied on an *in vitro* DEAE–Sepharose retention assay, while the fission yeast studies used an *in vivo* assay in which native immunoprecipitation with α-Srp54p was followed by primer extension RNA sequencing. Despite the different methods employed, for each analogous mutation tested, the effect was similar in both organisms (26, 69). Nucleotides critical for Srp54p binding to *S.pombe* and *E.coli* SRP RNA are highlighted in Figure 3.

Although nearly twice the size of mammalian SRP RNA, *S.cerevisiae* scR1 was suggested to be homologous (55). However, computer-assisted folding of scR1 (55) generated many potential hairpins, none obviously corresponding to domains II, III, and IV, and comparative sequence analysis (70) also failed to yield a structure with obvious similarity to other SRP RNA homologs. Evidence that this RNA is, in fact, part of an SRP-like particle emerged with the recent demonstration that it is immunoprecipitated by antisera directed against the *S.cerevisiae* Srp54 protein (70). In light of this and the coincidence of the Srp54p binding site in a bacterium and a eucaryote (see above), we re-examined the sequence of scR1 in search of the appropriate pattern of conserved nucleotides. This strategy succeeded in identifying the domain IV structure shown at the center of Figure 3, in which nucleotides critical for Srp54p binding are located in similar contexts as in other SRP RNAs. Notably, the domain IV tetranucleotide loop, which generally conforms to the consensus GNRA (reviewed in 31) and which had been the focus of previous attempts to fold scR1 RNA, is not conserved. In contrast to most SRP RNA homologs, including those from archaea and bacteria (see Figure 3), this hairpin in *S.cerevisiae* scR1 is capped by a loop of four pyrimidines closed by a non-canonical G*A base pair. The similarity of the terminal region of scR1 domain IV to the structure found in plant SRP RNAs, exemplified in Figure 3 by the tomato homolog, is unlikely to be due to a common evolutionary origin since, in a recently published phylogenetic tree determined by maximum likelihood methods (71), fungi group more closely with animals, and in the two other yeast SRP RNAs sequenced to date, domain IV is capped by a GNRA tetraloop. Consistent with the domain IV structures shown in Figure 3 for plant and budding yeast homologs, we have recently shown that a functional RNA results fom replacing the GAAA tetraloop in *S.pombe* SRP RNA with four pyrimidines (63). While a reasonable candidate for domain III, a tightly paired hairpin, emerged from the original computer-assisted folding of scR1 (55), assignment of a definite structure for the remainder of the RNA, as well as the precise terminus of the domain IV helix, must await sequencing of homologs from more closely related organisms.

Domain I of SRP RNA varies dramatically in size through evolution. The sequence of domain I of SRP RNA is conserved, albeit less so than that of domain IV. This region of human SRP RNA folds into two short hairpins joined by a stem and a single-stranded region (Figure 4, top left corner). Box I, the shorter of the two conserved motifs noted by Strub *et al.* (72) is located in the first hairpin loop, while Box II encompasses the single-stranded region joining the hairpins and surrounding nucleotides. The domain I structure of archaeal SRP RNAs, exemplified in the figure by *S.solfataricus* 7S, are strikingly similar to the corresponding region of mammalian SRP RNA. The absence of a third single-stranded purine in the joining region is atypical, but is shared by the *B.subtilis* domain I structure (Figure 4). The potential for an extended base pairing between the two domain

I hairpin loops is conserved not only in mammals and archaea, but in the *B.subtilis* RNA, the only bacterial homolog known to contain this structure, which provides compelling evidence for a common evolutionary origin.

Although the tomato and trypanosome domain I secondary structures are similar to the corresponding region of human SRP RNA except for the presence of a truncated second hairpin, they lack the potential for an extended tertiary interaction: only two Watson–Crick pairs can be proposed for the *L.esculentum* RNA, and no two contiguous bases are complementary in the *T.brucei* homolog. One possibility is that non-Watson–Crick interactions might occur; notably, the trypanosome RNA contains at this site a UUCG tetraloop that is likely to be highly structured (73). The domain I structures of the three fungal SRP RNAs sequenced to date are even more highly truncated, with Box I completely absent and the conserved nucleotides of Box II located in a hairpin loop and part of the adjacent stem rather than forming a single-stranded bulge between two hairpins. Although sequences just downstream from Box II in the *Y.lipolytica* RNA can be folded into a second hairpin (53), the stem contains G–U base pairs, in contrast to the second hairpin in metazoan and archaeal homologs, and a helix cannot form between sequences downstream from this proposed structure and the 5′ end of the RNA. Thus, a severe truncation of domain I in which most of the first and second hairpins were lost occurred subsequent to the divergence of the fungi from other eucaryotic lineages. Taken together, these observations indicate that domain I of SRP RNA, in contrast to domain IV, varies widely with respect to both its presence in a given RNA and the details of its structure.

Hydroxyl radical footprinting experiments demonstrate that, in the mammalian homolog, this region provides the binding site for the Srp9/14 heterodimeric protein (72), and mutating one of the two invariant residues is deleterious to the function of the *S.pombe* RNA *in vivo* (62). As noted above, mammalian SRP lacking Srp9/14p is competent to promote translocation but does not exhibit translation arrest activity (19, 74). As the latter function is arguably ancillary to the primary role of the particle, the complete absence of domain I from all but one bacterial SRP-like RNA (reviewed in 31) is not difficult to rationalize.

Homologs from eucaryotes and archaea can functionally replace bacterial SRP-like RNAs. An extensive series of genetic complementation experiments has been performed in *E.coli* to determine whether the RNAs described above are functionally as well as structurally related. Using a strain in which 4.5S RNA synthesis is under control of a regulated promoter, it was shown that expression of not only the similar-sized RNAs from *P.aeruginosa* and *T.thermophilus*, but also the much larger *B.subtilis* RNA and the very small homolog from *M.pneumoniae* support growth after repression of 4.5S synthesis (35, 40). Using a strain containng the native gene within a thermoinducible l prophage, the *P.aeruginosa* gene was re-identified, and complementing DNA fragments from *M.lysodeikticus* and *L.pneumophila* were isolated (38). More recently, this strategy was used to demonstrate that archaeal 7S RNAs, including homologs from *M.voltae*, *P.occultum* and *S.solfataricus*, as well as an internal fragment lacking domain I of the *P.occultum* RNA, can replace *E.coli* 4.5S (75). Thus, while all known bacterial SRP-like RNAs lack a structure homologous to the domain III hairpin, its presence is apparently compatible with function in *E.coli*.

The failure of the human 7SL gene inserted into the same vector to complement the 4.5S null allele was attributed to differences in three of the sixteen nucleotides absolutely conserved in all known bacterial and archaeal SRP-like RNAs (75); these include nucleotides within the domain IV tetraloop and the interior loops critical for Srp54p binding. However, the results of mutational, phylogenetic and structural analysis argue against these differences being responsible. First, mutating the G at the second position of the tetraloop in *E.coli* 4.5S RNA to an A, the nucleotide found in human 7SL, allowed fully wild-type growth (69; see Figure 3), and neither a transition nor a transversion at the corresponding nucleotide in *S.pombe* 7SL RNA had any effect on growth (61). Moreover, phylogenetic analysis of 16S-like rRNA sequences, in which tetraloops conforming to this consensus predominate, indicates that the second position can be occupied by all four nucleotides (76), and in the three-dimensional structure of a GCAA tetraloop determined by NMR spectroscopy, the cytidine is the sole base that does not participate in intra-loop hydrogen bonds (77). Second, reversal of the closing base pair in 4.5S RNA to the G–C found in the human RNA did not detectably affect growth of *E.coli* (69), although the orientation of this pair does appear to be important in *S.pombe* 7SL RNA (26). Since simple primary sequence divergence does not account for why archaeal SRP RNAs can functionally replace *E.coli* 4.5S RNA but a eucaryotic homolog cannot, a different explanation must be sought. Human SRP RNA does not accumulate to a high level in *E.coli* (75), possibly because its lack of the terminal stem-loop structure found in bacterial and archaeal homologs prevents efficient processing. Moreover, expression of the human RNA from a different promoter could extend the viability of cells in which 4.5S RNA synthesis is repressed by several generations, implying that it has at least some activity in *E.coli* (78). Further evidence for the functional equivalence of eucaryotic and bacterial SRP-like RNAs is the observation that a DNA fragment encoding human 7SL efficiently complements the depletion of an SRP-like RNA in *B.subtilis* (79). This strain grew nearly as well as one harboring *E.coli* 4.5S RNA, which allows virtually wild-type growth of *B.subtilis* despite being only half the size of the native RNA (79).

While these genetic complementation experiments strongly suggest that bacterial, archaeal and eucaryotic SRP-like RNAs are functionally equivalent, what has remained unclear is the precise nature of their role(s). Why is SRP a ribonucleoprotein and not, as originally believed (80), composed solely of protein? Among the RNA components of stable ribonucleproteins whose functions are known, many participate in RNA processing reactions in which they are known or believed to play catalytic roles (reviewed in 81, 82). Since it seems unlikely that SRP RNA plays a catalytic role, at least in modern organisms, other potential roles must be considered. Perhaps a reasonable paradigm is provided by the U1 small nuclear RNA, which recognizes the 5′ splice site of premessenger RNAs via Watson–Crick base pairing (reviewed in 83). By analogy, SRP RNA might contribute to specific and stable binding of the particle to another component of the SRP cycle, the most likely candidate being the ribosome, through base pairing interactions. Alternatively, the RNA may specifically interact with one or more ribosomal proteins or other non-SRP polypeptides. Within the context of the particle itself, the two extreme schools of thought are that the RNA subunit serves merely as a passive scaffold to ensure proper three-dimensional orientation of SRP proteins (17, 84), or that it

undergoes extensive structural rearrangements during the functional cycle of the particle (85, 86). These models are not necessarily mutually exclusive, since conformational changes in the proteins arising from their interactions with other components of the SRP cycle could produce structural perturbations in the RNA (87). Moreover, because SRP RNA extends throughout the length of the particle (88), it is also conceivable that it tramsmits the effects of protein conformational changes from one domain to another. Analysis of an extensive series of point mutations in *S.pombe* SRP RNA demonstrated that lethal and conditional lesions cluster in presumptive protein binding sites (26, 62, 63), supporting the idea that orienting the proteins is a critical role. Strikingly, mutations which lie at opposite ends of the RNA secondary structure exhibit synergistic lethality, implying that these domains interact, either directly or indirectly (66).

The core structure of SRP is conserved through evolution. Notably, a structure homologous to domain IV of mammalian 7SL is present in both the highly truncated SRP-like RNA from *M.mycoides* (34) and in the much larger *S.cerevisiae* scR1 RNA (this work). The ubiquity of the domain IV structure, together with the presence of conserved nucleotides (Figure 3), implies that it constitutes the essential RNA core of signal recognition particle. In RNAs from both size extremes, the nucleotides known to be important for binding of Srp54 protein to the *E.coli* and *S.pombe* homologs (see Figure 3) are conserved; indeed, both the *M.mycoides* and *S.cerevisiae* SRP RNAs bind have been shown to bind proteins homologous to Srp54p from their native organisms (70, 89). Consistent with the conservation of the Srp54p binding site in RNAs from a diverse phylogenetic spectrum, *SRP54* genes have been cloned from many different organisms (see below). Moreover, this protein plays critical roles during both the soluble and membrane-bound phases of the SRP cycle, being responsible for recognizing signal sequences at the ribosome and contacting the SRP receptor at the ER membrane (22, 23, 90). We therefore conclude that Srp54p is almost certainly an indispensable component of signal recognition particle.

Are the remaining SRP proteins also ubiquitous? We have recently shown by non-denaturing immunoprecipitation that fission yeast Srp54p is tightly complexed not only with 7SL RNA, but with five polypeptides, each of which is similar in size to a component of mammalian SRP. Although these observations suggest that the subunit composition of SRP is conserved at least throughout the eucaryotic branch of the phylogenetic tree (66), the smaller sizes of bacterial SRP RNA homologs imply that they may not bind the full complement of SRP proteins. Specifically, domain I, which contains the Srp9/14p binding site (72), is present only in the *B.subtilis* RNA, and domain III, the major determinant of Srp19p binding (91), is invariably absent. The inability of human Srp19p to bind *E.coli* 4.5S (78) suggests that bacterial SRP may lack a counterpart of this protein. While the precise composition of bacterial SRP-like particles remains to be determined, the sedimentation properties of the *E.coli* complex suggest that it does not consist simply of Srp54p and 4.5S RNA. Indeed, the ability of canine Srp68/72p to bind 4.5S RNA under some *in vitro* conditions (A.E.Johnson, personal communication) suggests the existence of a bacterial homolog of this protein, consistent with our proposal that the Srp68 polypeptide plays a pivotal role in the SRP cycle (see below).

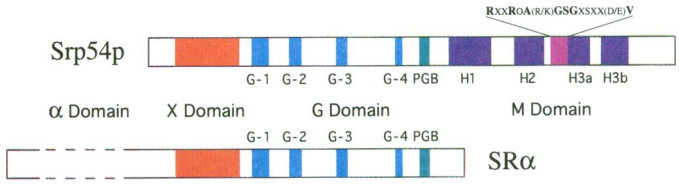

Figure 5. Domain structures of Srp54 and SRα proteins. (Top) The structure of the Srp54 protein is shown in a linear representation, with domains based on comparative sequence analysis (98, 99) indicated as follows: X domain (red), a conserved region of unknown function; G domain, the region containing the GTPase consensus motifs G-1 through G-4 (cyan) and an additional conserved sequence designated PGB (teal), which we propose to be the site of interaction with a common regulatory factor (see text); M domain, the methionine-rich region containing four proposed amphipathic helices H-1 through H-3b (purple) and a conserved element (fuschia) that we propose to be an RNA-binding motif (see text); X=any amino acid, J=hydrophillic, O=hydrophobic, amino acids shown in bold are >90% conserved. (Bottom) The structure of the SRα protein is shown in a linear representation, with domains indicated as follows: α domain (dashed lines), a region of variable size and unknown function with a conserved amino acid composition; X and G domains, defined as for Srp54p.

Evolution of SRP proteins

The sequences of the small SRP proteins provide few functional clues. The domains of SRP RNA that are absent in bacterial homologs correspond to the binding sites of the three smallest polypeptides in the mammalian ribonucleoprotein. These proteins are likely to directly contact SRP RNA, since they protect specific regions from nucleolytic digestion in the absence of other components (72, 74). However, the sequences of cloned cDNAs encoding the Srp9p, Srp14p, and Srp19p did not reveal homology to known RNA binding motifs, nor did they provide other hints about the functions of the proteins (92–94). Notably, the Srp9 and Srp19 proteins exhibit sequence similarity over a stretch of 36 amino acids, and a portion of this motif is conserved in Srp14p as well (94). Genes encoding proteins homologous to two of the small SRP subunits have now been identified from non-mammalian sources. The high degree of sequence conservation between a small *Caenorhabditis elegans* protein uncovered through the genome sequencing project (95) and mammalian Srp9p (47% amino acid identity; 66% similarity) precludes the delineation of functionally important motifs. A more distant relative of human Srp19p was identified through a screen for secretion-defective *S.cerevisiae* mutants (96, 97). While the functional significance of the two short patches of amino acid identity between this 31.5 kDa protein and its 19 kDa human counterpart (^{40}AVENP^{44} and ^{80}GRVR^{83}; human numbering) remains to be determined, we note that the second motif is found in the region conserved among Srp19p, Srp14p, and Srp9p; the suggestion that it may participate in RNA binding is reinforced by the presence of two conserved arginines.

Comparative sequence analysis reveals that Srp54p and SRα have modular structures. The sequences of mammalian cDNAs encoding the Srp54 protein provided compelling evidence that this component is highly conserved over vast evolutionary distances, since it shares extensive amino acid identity (31% identity; 60% similarity) with a previously uncategorized *E.coli* open reading frame (98, 99). The *ffh* locus encodes a protein of 48 kDa (p48) which is now known to be a functional homolog of Srp54p (100, 101; see below). The sequence conservation is

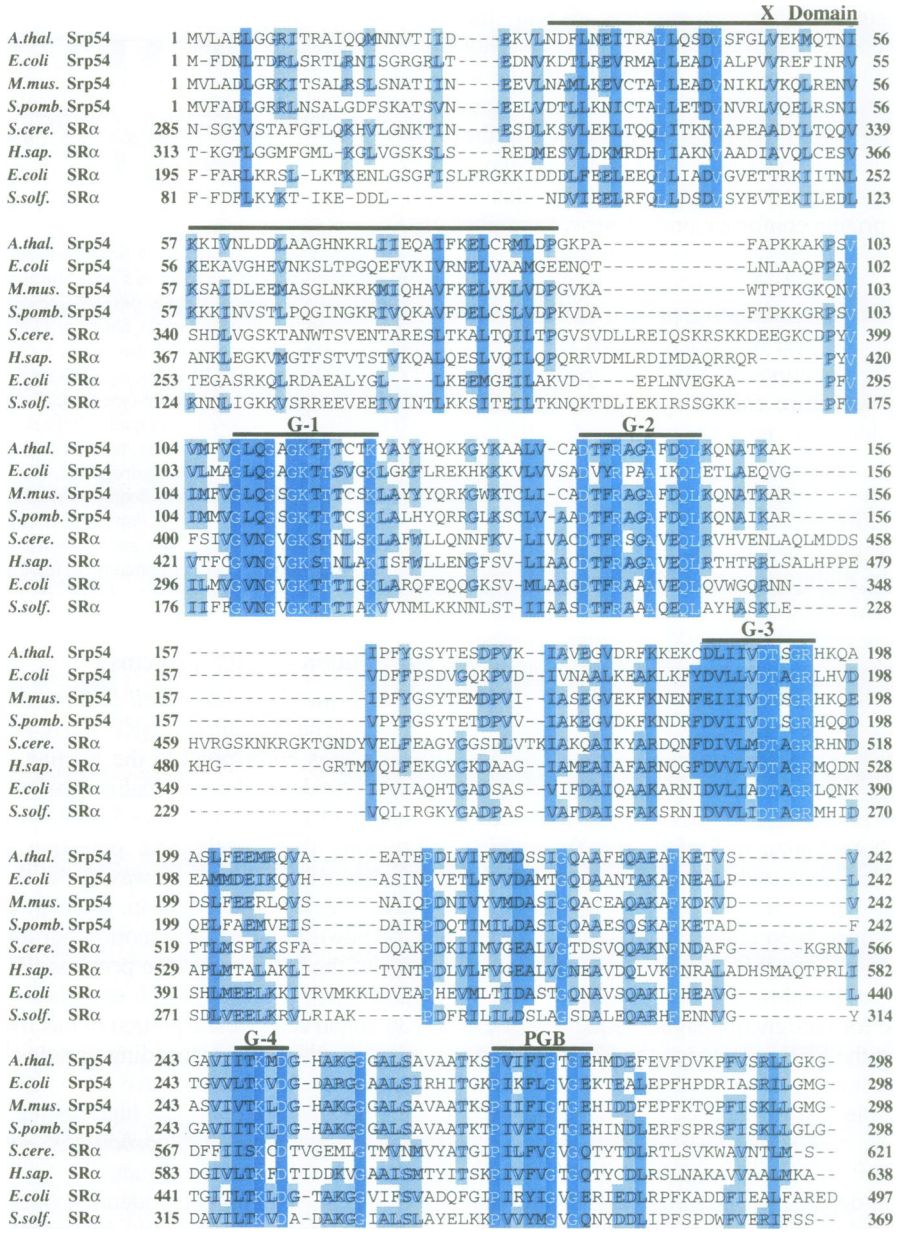

Figure 6. Sequence comparison of the X and G domains of a broad spectrum of Srp54p and SRα homologs. The alignment was generated using the Clustal V multiple sequence alignment program running on a Macintosh IIci personal computer. This software uses the algorithm of Higgins and Sharp (158). The sequences were retrieved from GenBank using either the NCBI RETRIEVE mail server (retrieve@ncbi.nih.gov) or, more recently, the GDB/accessor 1.0p program obtained from the EMBL mail server (netserv@embl-heidelberg.DE). The amino-terminal α domains of the SRα homologs and the carboxyl-terminal M domains of the Srp54p homologs are not depicted. The N-termini of the SRα and the C-termini of the Srp54p sequences were trimmed by eye initially, followed by alignment of the X and G domains using the default parameters. Slight adjustments were then made by hand to eliminate the few remaining unlikely gaps. Absolutely conserved amino acids are shown in the darkest blue shaded boxes in reverse type. Residues that are invariably similar are shown in black text shaded with medium blue (amino acids classified as similar are: K=R; S=T; V=I=L=M; D=E; Q=N; Y=F=W). Amino acid similarities present in ≥50% of the sequences are shaded in light blue to emphasize conservation between Srp54p and SRα. The four GTPase consensus motifs, as well as the PGB element that we propose to be the site of binding for a GDS shared by the two proteins (see text) are overlined. The sequences displayed were chosen to represent the phylogentic diversity of organisms from which homologs of these proteins have been sequenced, and include SRα from an archaeum. Accession numbers for the aligned sequences are provided in the Appendix.

sufficient to allow the isolation of *SRP54* genes from virtually any source using the polymerase chain reaction. This strategy was first used for molecular cloning of homologs from the fission yeast *S.pombe* and the budding yeast *S.cerevisiae* (102); the *S.cerevisiae* gene was also cloned serendipitously in an attempt to isolate a eucaryotic homolog of *E.coli* SecY (103). More recently, PCR has been used to clone *SRP54* genes from two additional bacteria, *M.mycoides* (89) and *B.subtilis* (104), as well as a cDNA encoding the cytoplasmic form of the protein from a plant, *Arabidopsis thaliana* (105). A cDNA encoding an *A.thaliana* chloroplast protein with striking similarity to bacterial Ffhp was identified during a screen for envelope polypeptides (106). Comparative sequence analysis revealed that the Srp54 protein can be divided into three distinct domains, illustrated schematically in Figure 5: an amino-terminal region that we have designated 'X', which is conserved between Srp54p and SRα

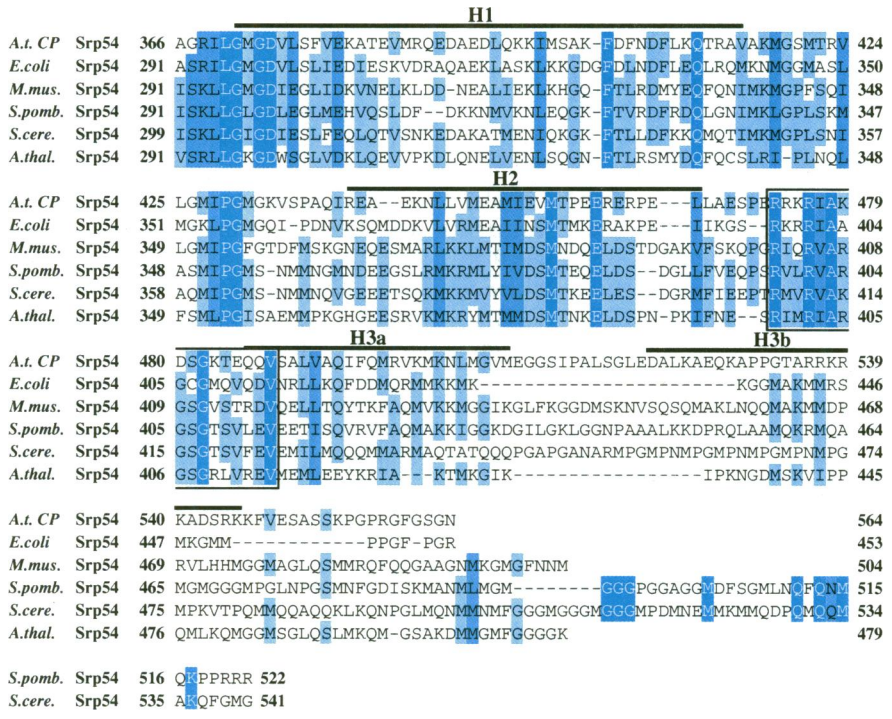

Figure 7. Sequence comparison of the M domains of divergent Srp54p homologs. The alignment was generated using the Clustal V multiple sequence alignment program as described in the legend to Figure 6. The extreme C-termini (beginning just upstream from helix H3b), which are the least similar, were aligned separately using Clustal V with both the pairwise and multiple alignment gap penalties increased, followed by final adjustments by hand. The shading is as in Figure 6. The proposed amphipathic helices for the fission yeast Srp54 protein are overlined, and the region of sequence similarity that we propose to be an RNA recognition motif (see text) is boxed. Accession numbers for the aligned sequences are provided in the Appendix.

(see below) but is unrelated to other proteins in the database, followed by a G domain which contains the four signature motifs of the GTPase superfamily, and a carboxyl-terminal M domain rich in methionine (98, 99).

The first evidence that the membrane-bound as well as the soluble components of the SRP cycle might be ubiquitous came from the surprising observation that the *E.coli ftsY* gene, which is located within an operon specifying components of the DNA replication machinery, is conserved over its entire length with the α subunit of the mammalian SRP receptor (98, 99, 107). Comparative sequence analysis indicates that this protein, too, can be divided into three distinct domains (Figure 5). Significantly, two of these are shared with Srp54p, although in SRα the X and G domains are located at the carboxyl rather than the amino terminus. A recently identified gene from the archaeum *S.solfataricus* was proposed to encode a homolog of SRα based on its sequence similarity to the carboxyl-terminal half of the mammalian protein (108). While the dramatically smaller size of the archaeal protein relative to mammalian SRα (369 vs. 638 amino acids) might at first seem incompatible with a common role, the product of the *E.coli ftsY* gene, which is known to be a functional homolog (100), is also significantly smaller (497 amino acids) than the mammalian protein. The size variation among the three proteins occurs in their unique amino terminal domain, designated α in Figure 5. Notably, although the sequence of this region is not well-conserved, the α domain of the *S.solfataricus* protein is similar to known SRα homologs in amino acid composition, which is strongly biased toward a mixture of basic and acidic residues. Further evidence that the proposed archaeal protein is a true homolog of SRα is provided by the presence of a region with sequence similarity to the X domain (see Figure 6), which precedes the G domain proper as defined by the GTPase consensus motifs. Our sequence comparison also revealed that PilA, a protein from *Neisseria gonorrhoeae* originally proposed to be a homolog of Srp54p (109), is very likely to be equivalent instead to SRα, since it contains all three domains found in the human, *E.coli* and *S.solfataricus* proteins, and exhibits a high degree of sequence conservation (37% identity; 55% similarity) with the product of the *E.coli ftsy* gene (data not shown). Finally, although FlhF, a protein required for motility in *B.subtilis*, contains a G domain more closely related to those of Srp54p and SRα than to other GTPases (110), it lacks sequences corresponding to the X domain, which they share, and has neither a C-terminal M domain in common with Srp54p homologs nor an N-terminal α domain in common with SRα homologs (data not shown).

The SRP receptor is a heterodimer in which the α subunit is a peripheral membrane protein anchored to the ER surface through the β subunit. To date, this organization has been directly demonstrated only for the mammalian homolog. However, the existence of a protein in *S.cerevisiae* that is functionally analogous to SRβ has been inferred from the membrane extraction properties of SRα (111), implying that the organization of the SRP receptor is similar from yeast to humans. The sequence of mammalian SRβ suggests that it is also a GTPase, but more closely related to the classical heterotrimeric G proteins than to Srp54p and SRα (unpublished data cited in 111). Recent evidence indicates that at least one component of the translocation apparatus that is presumed to receive proteins targeted to the membrane by SRP is conserved between bacteria and eucaryotes (reviewed in 4).

In combination with the identification of SRP RNA homologs from a diverse spectrum of bacteria and archaea, these observations suggest that an SRP-mediated pathway for protein export is present in all living cells.

A conserved sequence in Srp54p is similar to an RNA binding motif. Computer-assisted secondary structure analysis of the carboxyl-terminal region of Srp54p, designated the M domain based on its high content of methionine residues, suggested that it folds into four amphipathic α-helices, each with methionine and other hydrophobic residues clustered on one face (98). The flexible methionine side chains were proposed to form a binding pocket for signal peptides, whose sequences have only hydrophobicity in common (98); indeed, the M domain can be cross-linked to the signal sequence of a nascent presecretory protein (68, 112). The predicted boundaries between the proximal three helices virtually coincide between the human and *E.coli* proteins, despite scant (38%) amino acid identity in this region (98), and the validity of these assignments is further supported by the newly available sequences for plant cytoplasmic and chloroplast homologs of the protein (see Figure 7). The similarities among the M domains of Srp54 proteins break down at their extreme C-termini, even within the eucaryotes. Although bacterial homologs lack a clearly defined terminal methionine-rich amphipathic helix (H3b; see Figures 5 and 7), deletion of this structure is lethal in fission yeast (S.A. and J.A.W., unpublished data). While the M domain, in contrast to the G domain, lacks extensive sequence similarity to other known proteins, a database search identifies members of the Hsp70 family of stress proteins (S.A. and J.A.W., unpublished observations), which have in common the presence of multiple motifs of regularly spaced hydrophobic residues. These are likely to function in recognition of substrates, presumably signal sequences in the case of Srp54p (112, 113) and interior hydrophobic peptides of unfolded proteins in the case of Hsp70 proteins (114, 115).

Limited protease digestion of mammalian Srp54p demonstrates that the two domains postulated on the basis of sequence comparison are structurally distinct (68). *In vitro* studies with truncated proteins and proteolytic fragments indicate that, in addition to its proximity to the signal sequence, the M domain binds SRP RNA (67, 68). Nonetheless, this region lacks similarity to the ribonucleoprotein consensus RNA binding domain (116, 117) and, although it contains many basic residues that could contribute to electrostatic interactions with the phosphodiester backbone, none form a pattern corresponding to the arginine fork, another well-characterized RNA binding motif (118). Because the RNA binding region of Srp54p overlaps with the region necessary for signal sequence crosslinking, it was proposed that the hydrophobic faces of the amphipathic α-helices form a groove that binds signal peptides (100), while the hydrophilic faces contact SRP RNA (112, 119). This arrangement seems implausible, both because it requires two very different binding specificities to have become interspersed during evolution, and because RNA recognition by proteins is generally mediated by small, usually contiguous, sequence motifs (see, e.g., 118–123). We therefore searched for a local region of amino acid identity as a candidate RNA binding motif. A careful comparison of the M domains of all known Srp54p homologs, illustrated in Figure 7 for a representative set, uncovered a highly conserved region (consensus RXXROA(R/K)GSGXSXX(D/E)V; X=any amino acid, J=hydrophillic, O=hydrophobic) which, consistent with the earlier data, lies principally between two of the proposed methionine-rich α-helices (see Figure 7). The identity and spacing, although not the linear order, of these residues resembles a newly discovered consensus sequence (GJGXSKKOAK) found within the 23 amino acid RNA binding peptide of human immunodeficiency virus type 1 TRBP, which recognizes the TAR stem-loop structure; this consensus is also found in the *Drosophila* mRNA localization protein *staufen* and in human P1/dsI kinase (123), and a related sequence is found in *E.coli* tRNA CCA nucleotidyl transferase (S.A. and J.A.W., unpublished). The motif, which extends from positions 397–409 in the fission yeast protein, includes two invariant arginines and two conserved glycines. Remarkably, the arginines within this peptide are the only two absolutely conserved basic residues in the entire M domain. We have recently found that replacing either arginine with alanine produces growth defects at high temperature in *S.pombe* (S.A. and J.A.W., unpublished data), demonstrating that the motif is functionally important. We are currently determining whether these mutations disrupt binding of Srp54p to SRP RNA.

A G domain motif conserved between Srp54p and SRα may interact with a common GDS. Recent biochemical data demonstrating that both the Srp54 protein and the α subunit of the SRP receptor are active GTPases (90, 124), as had been predicted based on the presence of the four consensus GTP-binding motifs also found in small *Ras*-related GTPases, several translation factors, and the α subunits of heterotrimeric signaling proteins (98, 99). An earlier multiple sequence alignment revealed that Srp54p and SRα are more closely related to each other than to other members of the GTPase superfamily (125). We have extended these observations by aligning the G (and X) domains of the four most divergent examples of each protein for which sequence data are now available (Figure 6). Within the G-1 through G-4 motifs, the pattern of amino acid conservation in Srp54p and SRα places them in a small but expanding branch of the superfamily. In common with such diverse proteins as dynamin, the mammalian homolog of *Drosophila shibire*, which participates in receptor-mediated endocytosis (126, 127) and the yeast vacuolar protein sorting factor Vps1p (128), they contain, with the exception of a conservative substitution in *S.cerevisiae* SRα , a threonine at a position (248 in fission yeast Srp54p) within the G-4 motif where asparagine is found in hundreds of other GTPases (125). Srp54p and SRα differ from even their close relatives in the GTPase superfamily in that the first two conserved residues (DT) of the G2 motif are directly juxtaposed (Figure 6), while non-conserved amino acids intervene in other known GTPases (125).

In addition to possessing shared, unique features within their GTPase consensus motifs, Srp54p and SRα exhibit significant conservation in the surrounding amino acids. The amino acid similarities are most striking in the C-terminal portions of the sequences (Figure 6), with a region of extended conservation (9 contiguous identical or similar amino acids) lying downstream from G-4. In support of a significant role for this motif, it is highly conserved in proteins homologous to both Srp54p and SRα from all three branches of the phylogenetic tree; the poor match to this sequence in *B.subtilis* FlhFp underscores the notion that this protein does not play a role in the SRP cycle. We have recently found that mutating a threonine within this motif, which is conserved in all eukaryotic Srp54p homologs, is lethal in *S.pombe*, indicating that the element is functionally important.

The regions outside the four consensus motifs, which are unique to different GTPase subfamilies, contain the binding sites for specific modulatory factors such as guanine nucleotide dissociation stimulators and GTPase activating proteins. One well-characterized guanine nucleotide dissociation stimulator (GDS) is *E.coli* EF-Ts, which catalyzes release of GDP from EF-Tu; GTP then preferentially enters the active site due to its higher intracellular concentration, restoring the conformation that delivers charged tRNAs to the ribosome (reviewed in 125). The binding site for EF-Ts lies downstream from the G-4 motif in EF-Tu, in a position analogous to the extended region of similarity between Srp54p and SRα just described. Based on the location of this motif, in combination with studies of translocation *in vitro* (see below), we propose that it is recognized by a shared GDS for these proteins, and have therefore designated it the PGB (putative GDS binding) element. In support of the idea that the PGB motifs in Srp54p and SRα interact with a common third component, we note that its sequence is more similar between Srp54p-SRα pairs from the same or closely related organisms than among homologs of the same protein from different organisms (Figure 6).

The other major class of factors that modulate the activity of GTPase superfamily members, the GTPase activating proteins (GAPs), interact with amino acids surrounding the G-2 motif (reviewed in 125). Although these sequences are certainly conserved between Srp54p and SRα (see Figure 6), there is no biochemical evidence to support the notion that the two proteins share a GAP. Finally, the conservation of the X domain between Srp54p and SRα may indicate a common function; one tantalizing possibility is that this region is the site of contact between the two proteins, which is thought to have originated in homotypic interactions (98).

The Srp68 protein contains sequences similar to guanine nucleotide dissociation stimulators. A vast body of biochemical and genetic evidence supports the idea that proteins containing GTPase consensus motifs but otherwise different primary sequences employ a common mechanism to serve as binary molecular switches in a variety of biological contexts (reviewed in 129). In cycling between a GTP bound 'on' state and a GDP bound 'off' state, these proteins adopt dramatically different conformations (see, e.g., 130, 131) which have unequal affinities for upstream and downstream effectors in the pathway. Because GTPases exhibit very low intrinsic hydrolysis rates and very high affinities for guanine nucleotides, they typically employ two extrinsic factors in order to proceed through multiple catalytic cycles: a GTPase activating protein to stimulate GTP hydrolysis and a guanine nucleotide dissociation stimulator (also known as a guanine nucleotide release protein or GNRP) to promote displacement of GDP (reviewed in 132). In view of the requirement for GTP during ER import (16, 133), and the fact that one SRP subunit as well as both components of the SRP receptor contain GTPase consensus motifs, it seems likely that the SRP cycle is driven by sequential conformational changes in these proteins. However, specific models have not yet been proposed regarding which events in the SRP cycle are regulated by either subunit of the receptor, and GTP hydrolysis by Srp54p has been hypothesized to trigger both transfer of the signal sequence to the translocation machinery (134) and the subsequent release of SRP from its receptor (90). Whatever the precise functions of Srp54p and SRα , the extent of amino acid similarity between their G (and X) domains (Figure 6) suggests that they interact with one or more common partners during the SRP cycle. Of particular interest is the PGB motif, whose location suggests that it functions as a binding site for a guanine nucleotide dissociation stimulator shared by these two proteins. In seeking clues to the identity of such a factor, we took note of the early observation that alkylation of the 68/72 kDa heterodimeric SRP protein produces a particle deficient in promoting translocation (21). Several lines of evidence indicate that Srp54p is the SRP subunit that makes initial contact with the SRP receptor on the ER membrane (90), eliminating the most obvious role for Srp68/72p during translocation. We propose instead that disruption of translocation is due to a component of this heterodimer serving as the SRα GDS, which is consistent with the observation that guanine nucleotide exchange at SRα most likely occurs subsequent to SRP binding (16). The notion that the guanyl exchange factor for Srp54p is a component of signal recognition particle is compatible with our recent demonstration that a mutation designed to lock this protein in the GDP-bound state does not exhibit genetic dominance (134), in contrast to observations with GTPases that employ freely dissociable GDS factors (135–137).

Does the sequence of either subunit of the large heterodimeric SRP protein support a role as a guanine nucleotide exchange protein? Initially, the primary structure of neither cloned cDNA appeared illuminating: Srp72p was reported to lack similarity to any known protein (138), and the only landmark noted within Srp68p was a glycine-rich region near the amino terminus (139). Although a BLAST search (140) does not reveal obvious similarities to known GDS proteins, we were encouraged to proceed not only by the functional data, but by the fact that both subunits of the Srp68/72 heterodimer are similar in size to members of this class of proteins. Moreover, the sequences of GDS proteins, most of which are quite large, are not very closely related, and it is only recently that three short elements containing several conserved amino acids were uncovered (132, 141). Using an iterative alignment procedure (described in the legend to Figure 8), we were able to delineate regions of Srp68p similar to each of these motifs. The significance of the sequence similarity is underscored by the fact that all of the other known or suspected GDS's in the comparison serve as exchange factors for small GTPases from the *Ras* branch of the superfamily, to which Srp54p and SRα , the proposed substrates for Srp68p, are only distantly related. In addition to the amino acid conservation, it is noteworthy that the linear order of the Box I, II and III motifs is preserved (see Figure 8). Moreover, the spacing between Boxes I and II is similar between Srp68p and the other proteins, although the distance from Box II to Box III is greater. This sequence similarity, in combination with circumstantial evidence from genetic and biochemical studies of SRP cycle components and the existence of a conserved potential binding site, suggests that the Srp68 protein serves as the GDS for both Srp54p and the α subunit of the SRP receptor. We therefore expect that SRP from any organism will possess a homolog of this protein, and that functional studies will demonstrate that it promotes guanine nucleotide exchange at two steps in the SRP cycle.

Evolution of the SRP pathway

All SRP-like complexes are likely to be involved in secretion. While the presence of obvious sequence and structural homology between components of bacterial and human SRP-like particles indicates that they share a common evolutionary origin (142), a lingering controversy has surrounded the idea that they function

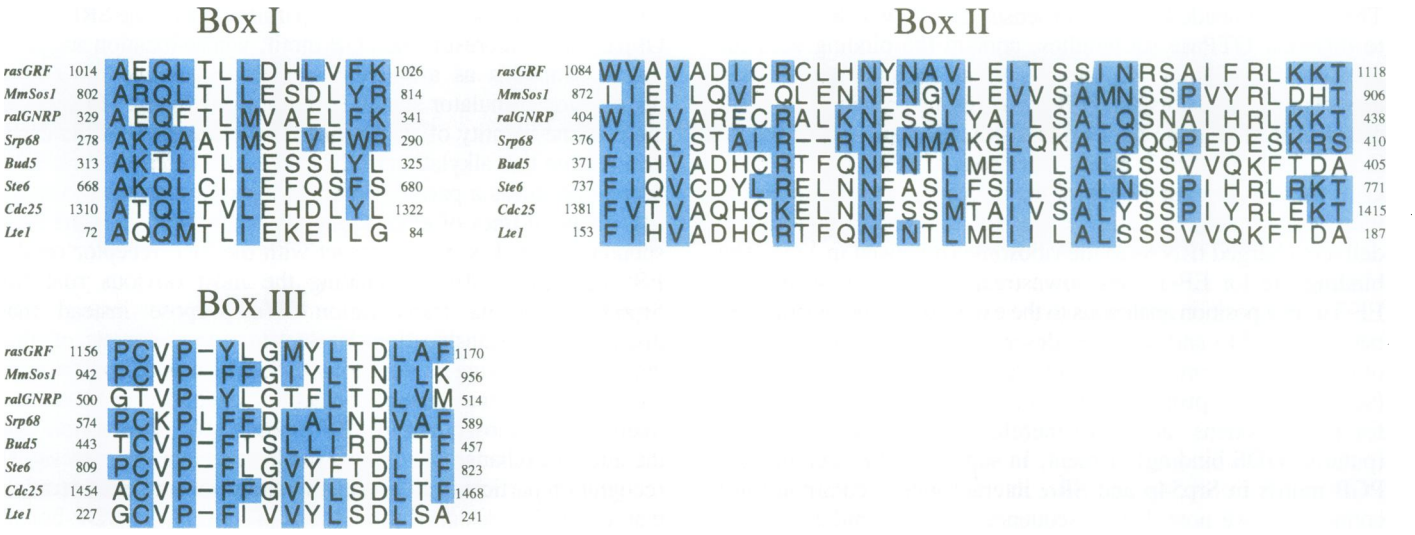

Figure 8. Sequence motifs conserved between the Srp68 protein and guanine nucleotide dissociation stimulators of small *Ras*-related GTPases. The sequence similarity between the regions of conservation shared by guanyl dissociation factors and the mammalian Srp68p was uncovered by first aligning the sequences of known or presumptive GDS proteins (132). Sequences outside the region designated Box II were then eliminated, and a consensus sequence for this motif derived. This element was then aligned with the Srp68p sequence using the Clustal V program set to default parameters. The consensus sequence aligned to the region shown. A new consensus derived from the known GDS Box II sequences and Srp68p Box II, when used to probe the non-redundant protein database using BLAST (140), returned all of the GDS sequences along with Srp68p as significant matches. Box I and Box III were then searched for and found by eye. Because Srp68p is likely to display the poorest conservation within this group (see text), amino acids which are similar or identical between this protein and any of the known GDS's are indicated by light blue shading. Amino acids considered similar are as in Figures 6 and 7, except that the hydrophobic group (V=I=L=M) is expanded to include A.

in a common pathway. The bacterial geneticists who first characterized 4.5S RNA homologs have generally interpreted their data in terms of a role in translation (e.g., 79, 143–147), while laboratories which also study eucaryotic SRP have postulated a more restricted functional engagement with ribosomes synthesizing presecretory proteins (e.g., 78, 142). We favor the latter possibility since, despite intensive investigation, no evidence has emerged to suggest that eucaryotic SRP is a general modulator of translation and, more significantly, the results of each experiment conducted with components of the bacterial RNP can be easily reconciled with a role in secretion. Upon depletion of 4.5S RNA, which is essential for growth in *E.coli*, protein synthetic activity decreases (146), and extracts prepared from these cells can translate poly(U) but not natural mRNAs (145). Although this was attributed to impaired initiation, purified ribosomes but not 4.5S could relieve the defect, implying that the small RNA does not directly promote translation. Moreover, DNA synthesis declined only minutes after the shut-down of translation, suggesting that both pathways were responding to a common stimulus. Consistent with the hypothesis that the signal is the presence of normally secreted proteins inside the cell, a later study showed that loss of 4.5S RNA induces the persistent synthesis of heat shock proteins (147); several Hsp's and their constitutively expressed counterparts are now known to serve as molecular chaperones that assist in protein folding (reviewed in 148) and are induced by the presence of unfolded or incorrectly folded proteins (149). To gain insight into factors that interact with 4.5S RNA, inducer-independent suppressors of a regulated gene were isolated; a mutation in the RNA polymerase region of the *E.coli* chromosome increased the concentration of 4.5S RNA, while another suppressor mapped to the gene encoding translation elongation factor G (EF-G), which promotes translocation of tRNA-linked nascent proteins from the A site to the P site on the ribosome (143). Although it was proposed that 4.5S acts in concert with EF-G either transiently or only for a restricted class of mRNAs, the sub-stoichiometric association of the RNA with polysomes is also compatible with a role in co-translational export of presecretory proteins. Additional suppressors of 4.5S depletion map to other components of the translation machinery, including two amino acyl-tRNA synthetases, alanyl tRNA and 23S ribosomal RNA (144). To explain these observations, 4.5S RNA was postulated to stabilize the ribosome following the action of EF-G by replacing 23S rRNA as a binding site for the elongation factor. However, the extended sequence identity between 4.5S and 23S, a major cornerstone of this model (144), is not conserved in homologs from other organisms. Notably, all of the suppressors isolated to date can only compensate for reduced levels of 4.5S RNA, not its complete absence. We therefore suggest an alternative model, *viz.*, that decreasing the rate of translation increases the interval during which SRP can bind to the ribosome before the growth of the nascent polypeptide renders the signal sequence inaccessible. Consistent with the existence of a mechanism in bacteria that couples secretion and translation, certain suppressors of *sec* mutants reduce the translation rate of exported proteins (reviewed in 150).

The strongest argument that the particle containing *E.coli* 4.5S RNA participates in secretory protein targeting derives from recent data that provide direct evidence for such a role. The search for an SRP-like function for this RNA was stimulated by the finding that it is complexed with the product of the *ffh* locus (Ffhp, also called p48), which is homologous to Srp54p (78, 142). Ffhp can be crosslinked to the β-lactamase signal sequence *in vitro* (151), and cells deficient in p48 exhibit a dramatic decrease in export of this protein (101). Depletion of 4.5S RNA *in vivo* also results in impaired β-lactamase secretion (78, 142). The normal

secretion of lamB and maltose binding protein in the Ffh-deficient strains may explain why SRP components were not discovered through genetic selections for export-defective *E.coli* mutants, since these proteins were the principal substrates employed; their export depends instead upon the action of the molecular chaperone SecB (reviewed in 152). In *B.subtilis*, as in *E.coli*, depletion of either SRP RNA or Srp54 protein results in a diminished rate of protein synthesis, as well as defective export of secretory proteins (79, 104).

Among the eucaryotes, the *in vivo* operation of SRP-mediated protein targeting has been most extensively characterized in *S.cerevisiae*. As predicted by the model derived from *in vitro* assays of mammalian components, the loss of SRP function in this organism results in cytoplasmic accumulation of precursors to ER-targeted polypeptides (70). Repression of Srp54p synthesis virtually eliminated ER transport of the vacuolar membrane protein dipeptidyl aminopeptidase B, while the secretory proteins invertase and prepro-α-factor were moderately affected and the import of a soluble vacuolar protein, carboxypeptidase Y, was not detectably impaired (70, 153). Evidence for the occurrence of translation arrest *in vivo*, also postulated on the basis of *in vitro* data, derives from experiments with two other yeasts. Strains of *Y.lipolytica* harboring temperature-sensitive mutations in the putative Srp19p binding site in domain III of SRP RNA exhibit reduced levels of mature alkaline extracellular protease (64, 65), but no accumulation of cytoplasmic precursor even with very short labeling times. Similarly, the most dramatic effect of a temperature-sensitive *S.pombe* mutant containing an insertion at the junction of domains II and III of SRP RNA, within the region protected by canine Srp68/72p, is a deficiency in the production of mature secreted acid phosphatase at the non-pemissive temperature (P.Brennwald and J.A.W., unpublished data), and diminished levels of glycosylated acid phosphatase are also observed in a cold-sensitive strain of fission yeast harboring a mutation in the gene encoding Srp54p (S.A. and J.A.W., unpublished data). These observations imply that mutations at different sites in SRP RNA and in Srp54p are compatible with assembly into particles that can recognize and bind to ribosomes translating presecretory proteins, but are either unable to bind the SRP receptor at the membrane or engage in non-productive interactions, with the net result that synthesis of the presecretory protein does not resume.

Indirect evidence was recently presented to implicate an SRP component from the third branch of the phylogenetic tree, 7SL RNA from the archaeum *H.halobium*, in secretion. This RNA, like mammalian SRP RNA, co-fractionates with membrane-bound polysomes, and this association correlates with expression of prebacterioopsin, the major membrane protein in this organism (154). In combination with the identification of a protein homologous to SRα in another archaeal species, *S.solfataricus* (108), this observation strongly suggests that these organisms employ a secretory protein targeting pathway analogous to the SRP cycle.

SRP function can be by-passed in vivo. Although it seems beyond dispute that the membrane-targeting function of SRP is conserved through evolution, the fact that secretion of only a subset of proteins is impaired in its absence implies that at least one additional mechanism for protein export must exist *in vivo*. Consistent with this idea, *S.cerevisiae* strains harboring disruptions of the genes encoding Srp54p, scR1 RNA and SRα, either singly or in combination, are viable (55, 70, 111). The presence of two SRP RNA genes in *Y.lipolytica* suggests a more important role for the pathway in this dimorphic yeast and, indeed, at least one intact copy is required for growth (155). Despite the ability of most proteins examined to be exported from *E.coli* after 4.5S or Ffhp depletion, disruption of the gene encoding either component is lethal (101, 146). While the homologs of these two SRP components are likewise essential for viability in *S.pombe* (52, 54, 134), we have recently obtained evidence that this organism, too, possesses a mechanism to compensate for SRP deficiency, since an adaptive response likely to involve heat-shock proteins can occur in mutants harboring conditional alleles of Srp54p (134). Even in mammals, despite the predominance of co-translational ER translocation, an SRP-independent pathway has been shown to operate for at least some presecretory proteins (e.g., 156). The SRP-mediated targeting mechanism has been postulated to be of more ancient evolutionary origin, with chaperone-assisted translocation arising later as a salvage pathway (3). However, it has also been noted that, if secretory protein trafficking had been studied first in bacteria and *S.cerevisiae* rather than in pancreatic cells, the opposite scenario might seem more plausible (152). We favor the former model, and speculate that, in the course of evolution, the requirement for SRP may be by-passed by altering a few essential secretory substrates to allow their export via alternative pathways.

ACKNOWLEDGEMENTS

The authors would like to thank Claudia Reich, Enrique Martinez-Forcé, and Daniel Celander for critical comments on the manuscript, and Faith Belanger and Art Johnson for sharing data prior to publication. This work was supported by NSF Grants #DCB 88-16325, awarded to J.A.W., and #BCS 92-13895, awarded to J.A.W. and K.Dane Wittrup.

REFERENCES

1. Walter, P. and Lingappa, V. (1986) *Annu. Rev. Cell Biol.*, **2**, 499–516.
2. Rapaport, T. A. (1991) *FASEB J.*, **5**, 2792–2798.
3. Rapaport, T. A. (1992) *Science*, **258**, 931–936.
4. Sanders, S. L. and Schekman, R. (1992) *J. Biol. Chem.*, **267**, 13791–13794.
5. Nunnari, J. and Walter, P. (1992) *Curr. Op. Cell. Biol.*, **4**, 573–580.
6. Gilmore, R. (1993) *Cell*, **75**, 589–592.
7. Walter, P. and Blobel, G. (1981) *J. Cell Biol.*, **91**, 551–556.
8. Walter, P. and Blobel, G. (1981) *J. Cell Biol.*, **91**, 557–561.
9. Wolin, S. L. and Walter, P. (1988) *J. Cell Biol.*, **109**, 2617–2622.
10. Gilmore, R., Walter, P. and Blobel, G. (1982) *J. Cell Biol.*, **95**, 463–469.
11. Meyer, D. I., Krause, E. and Dobberstein, B. (1982) *Nature*, **297**, 647–650.
12. Görlich, D., Hartmann, E., Prehn, S. and Rapoport, T. A. (1992) *Nature*, **357**, 47–52.
13. Crowley, K. S., Reinhart, G. D. and Johnson, A. E. (1993) *Cell*, **73**, 1101–1115.
14. Gilmore, R. and Blobel, G. (1983) *Cell*, **35**, 677–685.
15. Connolly, T. and Gilmore, R. (1986) *J. Cell Biol.*, **103**, 2253–2261.
16. Connolly, T. and Gilmore, R. (1989) *Cell*, **57**, 599–610.
17. Walter, P. and Blobel, G. (1982) *Nature*, **299**, 691–698.
18. Walter, P. and Blobel, G. (1983) *Cell*, **34**, 525–533.
19. Siegel, V. and Walter, P. (1985) *J. Cell Biol.*, **100**, 1913–1921.
20. Siegel, V. and Walter, P. (1988) *Trends Biochem. Sci.*, **13**, 376–378.
21. Siegel, V. and Walter, P. (1988) *Cell*, **52**, 39–49.
22. Krieg, U. C., Walter, P. and Johnson, A. E. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 8604–8608.
23. Kurzchalia, T. V., Wiedmann, M., Girshovich, A. S., Bochkareva, E. S., Bielka, H. and Rapoport, T. (1986) *Nature*, **320**, 634–636.
24. Siegel, V. and Walter, P. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 1801–1805.
25. Hann, B. C., Stirling, C. J. and Walter, P. (1992) *Nature*, **356**, 532–533.

26. Selinger, D., Brennwald, P., Liao, X. and Wise, J. A. (1993) *Mol. Cell. Biol.*, **13**, 1353–1362.
27. Woese, C. R., Kandler, O. and Wheelis, M L. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 4576–4579.
28. Gundelfinger, E. D., DiCarlo, M., Zopf, D. and Melli, M. (1984) *EMBO J.*, **10**, 2325–2332.
29. Poritz, M. A., Strub, K. and Walter, P. (1988) *Cell*, **55**, 4–6.
30. Zwieb, C. (1985) *Nucleic Acids Res.*, **13**, 6105–6124.
31. Larsen, N. and Zwieb, C. (1990) *Nucleic Acids Res.*, **19**, 209–215.
32. Hsu, L. M., Zagorski, J. and Fournier, M. J. (1984) *J. Mol. Biol.*, **178**, 509–531.
33. Struck, J. C. R., Vogel, D. W., Ulbrich, N. and Erdmann, V. A. (1988) *Nucleic Acids Res.*, **16**, 2719.
34. Samuelsson, T. and Guindy, Y. (1990) *Nucleic Acids Res.*, **18**, 4938.
35. Simoneau, P. and Hu, P.-C. (1992) *J. Bacteriol.*, **174**, 627–629.
36. Struck, J. C. R., Toschka, H. Y., Specht, T. and Erdmann, V. A. (1988) *Nucleic Acids Res.*, **16**, 7740.
37. Toschka, H. Y., Struck, J. C. R. and Erdmann, V. A. (1989) *Nucleic Acids Res.*, **17**, 31–36.
38. Brown, S., Thon, G. and Tolentino, E. (1989) *J. Bacteriol.*, **171**, 6517–6520.
39. Olsen, G. J., Overbeek, R., Larsen, N., Marsh, T. L., McCaughey, M. J., Maciukenas, M. A., Kuan, W.-M., Macke, T. J., Xing, Y. and Woese, C. R. (1992) *Nucleic Acids Res.*, **20**, 2199–2200.
40. Struck, J. C. R., Lempicki, R. A., Toschka, H. Y., Erdmann, V. A. and Fournier, M. J. (1990) *J. Bacteriol.*, **172**, 1284–1288.
41. Beja, O., Ullu, E. and Michaeli, S. (1993) *Mol. and Bioch. Parasitol.*, **57**, 223–230.
42. Moritz, A. and Goebel, W. (1985) *Nucleic Acids Research*, **13**, 6969–6979.
43. Østergård, L., Larsen, N., Leffers, H., Kjems, J. and Garrett, R. A. (1987) *System. Appl. Microbiol.*, **9**, 199–209.
44. Kaine, B. P. and Merkel, V. L. (1989) *J. Bacteriol.*, **171**, 4261–4265.
45. Kaine, B. P., Schurke, C. M. and Stetter, K. O. (1990) *System. Appl. Microbiol.*, **12**, 8–14.
46. Haas, E. S., Klanner, A., Ramm, K. and Sanger, H. L. (1988) *EMBO J.*, **7**, 4063–4074.
47. Kaine, B. P. (1990) *Mol. Gen. Genet.*, **221**, 315–321.
48. Luehrsen, K. R., Nicholson, D. E., Jr. and Fox, G. E. (1985) *Curr. Microbiol.*, **12**, 69–72.
49. Ullu, E., Murphy, S. and Melli, M. (1982) *Cell*, **29**, 195–202.
50. Li, W.-Y., Reddy, R., Henning, D., Epstein, P. and Busch, H. (1982) *J. Biol. Chem.*, **257**, 5136–5142.
51. Ullu, E. and Tschudi, C. (1984) *Nature*, **312**, 171–172.
52. Brennwald, P., Liao, X., Holm, K., Porter, G. and Wise, J.A. (1988) *Mol. Cell. Biol.*, **8**, 1580–1590.
53. Poritz, M. A., Siegel, V., Hansen, W. B. and Walter, P. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 4315–4319.
54. Ribes, V., Dehoux, P. and Tollervey, D. (1988) *EMBO J.*, **7**, 231–237.
55. Felici, F., Cesareni, G. and Hughes, J. M. X. (1989) *Mol. Cell. Biol.*, **9**, 3260–3268.
56. He, F., Beckerich, J.-M., Ribes, V., Tollervey, D. and Gaillardin, C. (1989) *Curr. Genet.*, **16**, 347–350.
57. Brennwald, P., Siegel, V., Walter, P. and Wise, J. A. (1991) *Nucleic Acids Res.*, **19**, 1942.
58. Michaeli, S., Podell, D., Agabian, N. and Ullu, E. (1992) *Mol. and Bioch. Parasitol.*, **51**, 55–64.
59. Campos, N., Palau, J. and Zwieb, C. (1989) *Nucleic Acids Res.*, **17**, 1573–1588.
60. Marshallsay, C., Prehn, S. and Zwieb, C. (1989) *Nucleic Acids Res.*, **17**, 1771.
61. Liao, X., Brennwald, P. and Wise, J. A. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 4137–4141.
62. Liao, X., Selinger, D., Althoff, S., Chiang, A., Hamilton, D., Ma, M. and Wise, J. A. (1992) *Nucleic Acids Res.*, **20**, 1607–1615.
63. Selinger, D., Liao, X. and Wise, J. A. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 5409–5413.
64. He, F., Beckerich, J.-M. and Gaillardin, C. (1992) *J. Biol. Chem.*, **267**, 1932–1937.
65. Yaver, D. S., Matoba, S. and Ogrydziak, D. M. (1992) *J. Cell Biol.*, **116**, 605–616.
66. Selinger, D., Brennwald, P., Althoff, S., Reich, C., Hann, B., Walter, P. and Wise, J. A. (1993) Submitted for publication.
67. Römisch, K., Webb, J., Lingelbäch, K., Gausepohl, H. and Dobberstein, B. (1990) *J. Cell Biol.*, **111**, 1793–1802.
68. Zopf, D., Bernstein, H. D., Johnson, A. E. and Walter, P. (1990) *EMBO J.*, **9**, 4511–4517.
69. Wood, H., Luirink, J. and Tollervey, D. (1992) *Nucleic Acids Res.*, **20**, 5919–5925.
70. Hann, B. C. and Walter, P. (1991) *Cell*, **67**, 131–144.
71. Wainwright, P. O., Hinkle, G., Sogin, M. and Stickel, S. K. (1993) *Science*, **260**, 340–342.
72. Strub, K., Barnett-Moss, J. and Walter, P. (1991) *Mol. Cell. Biol.*, **11**, 3949–3959.
73. Cheong, C., Varani, G. and Tinoco, I., Jr. (1990) *Nature*, **346**, 680–682.
74. Siegel, V. and Walter, P. (1986) *Nature*, **320**, 81–84.
75. Brown, S. (1991) *J. Bacteriol.*, **173**, 1835–1837.
76. Woese, C. R., Winker, S. and Guttell, R. R. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 8467–8471.
77. Heus, H. A. and Pardi, A. (1991) *Science*, **253**, 191–194.
78. Ribes, V., Römisch, K., Giner, A., Dobberstein, B. and Tollervey, D. (1990) *Cell*, **63**, 591–600.
79. Nakamura, K., Imai, Y., Nakamura, A. and Yamane, K. (1992) *J. Bacteriol.*, **174**, 2185–2192.
80. Walter, P. and Blobel, G. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 7112–7116.
81. Guthrie, C. (1991) *Science*, **253**, 157–163.
82. Pace, N. R. and Smith, D. (1990) *J. Biol. Chem.*, **265**, 3587–3590.
83. Guthrie, C. and Patterson, B. (1988) *Annu. Rev. Genet.*, **22**, 387–420.
84. Dreyfuss, G., Swanson, M. S. and Pinol-Roma, S. (1988) *Trends Biochem. Sci.*, **13**, 86–91.
85. Zwieb, C. and Ullu, E. (1986) *Nucleic Acids Res.*, **14**, 4639–4657.
86. Andreazzoli, M. and Gerbi, S. A. (1991) *EMBO J.*, **10**, 767–777.
87. Janiak, F., Walter, P. and Johnson, A. E. (1992) *Biochemistry*, **31**, 5830–5840.
88. Andrews, D. W., Walter, P. and Ottensmeyer, F. P. (1987) *EMBO J.*, **6**, 3471–3477.
89. Samuelsson, T. (1992) *Nucleic Acids Res.*, **20**, 5763–5770.
90. Miller, J. D., Wilhelm, H., Gierasch, L., Gilmore, R. and Walter, P. (1993) *Nature*, **366**, 351–354.
91. Zwieb, C. (1991) *Nucleic Acids Res.*, **19**, 2955–2960.
92. Lingelbäch, K., Zwieb, C., Webb, J. R., Marshallsay, C., Hoben, P. J., Walter, P. and Dobberstein, B. (1988) *Nucleic Acids Res.*, **16**, 9431–9442.
93. Strub, K. and Walter, P. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 9747–9751.
94. Strub, K. and Walter., P. (1990) *Mol. Cell. Biol.*, **10**, 777–784.
95. Hawkins, T. and Ainscough, R. (unpublished) Z22177.
96. Stirling, C. J., Rothblatt, J., Hosobuchi, M., Deshaies, R. and Schekman, R. (1992) *Mol. Biol. Cell*, **3**, 129–142.
97. Stirling, C. J. and Hewitt, E. W. (1992) *Nature*, **356**, 534–537.
98. Bernstein, H. D., Poritz, M. A., Strub, K., Hoben, P. J., Brenner, S. and Walter, P. (1989) *Nature*, **340**, 482–486.
99. Römisch, K., Webb, J., Herz, J., Prehn, S., Frank, R., Vingron, M. and Dobberstein, B. (1989) *Nature*, **340**, 478–482.
100. Bernstein, H. D., Zopf, D. and Walter, P. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 5229–5233.
101. Phillips, G. J. and Silhavy, T. J. (1992) *Nature*, **359**, 744–746.
102. Hann, B. C., Poritz, M. A. and Walter, P. (1989) *J. Cell Biol.*, **109**, 3223–3230.
103. Amaya, Y., Nakano, A., Ito, K. and Mori, M. (1990) *J. Biochem.*, **107**, 457–463.
104. Honda, K., Nakamura, K., Nishiguchi, M. and Yamane, K. (1993) *J. Bacteriol.*, **175**, 4885–4894.
105. Lindstrom, J. T., Chu, B. and Belanger, F. C. (1994) *Plant Mol. Biol.*, in press.
106. Franklin, A. E. and Hoffman, N. E. (1993) *J. Biol. Chem.*, **268**, 22175–22180.
107. Lauffer, L., Garcia, P. D., Harkins, R. N., Coussens, L., Ullrich, A. and Walter, P. (1986) *Nature*, **318**, 334–338.
108. Ramirez, C. and Matheson, A. T. (1991) *Mol. Microbiol.*, **5**, 1687–1693.
109. Taha, M. K., Dupuy, B., Saurin, W., So, M. and Marchal, C. (1991) *Mol. Microbiol.*, **5**, 137–148.
110. Carpenter, P. B., Hanlon, D. W. and Ordal, G. W. (1992) *Mol. Microbiol.*, **6**, 2705–2713.
111. Ogg, S. C., Poritz, M. A. and Walter, P. (1992) *Mol. Biol. Cell*, **3**, 895–911.
112. High, S. A., Dobberstein, B. (1991) *J. Cell Biol.*, **113**, 229–233.
113. Zopf, D., Bernstein, H. D. and Walter, P. (1993) *J. Cell Biol.*, **120**, 1113–1121.

114. Flynn, G. C., Chappell, T. G. and Rothman, J. E. (1989) *Science*, **245**, 385–390.
115. Blond-Elgundi, S. C., S. E., Dower, W. J., Lipshutz, R. J., Sprang, S. R., Sambrook, J. F. and Gething, M.-J. (1993) *Cell*, **75**, 717–728.
116. Dreyfuss, G., Philipson, L. and Mattaj, I. W. (1988) *J. Cell. Biol.*, **106**, 1419–1425.
117. Bandziulis, R. J., Swanson, M. S. and Dreyfuss, G. (1989) *Genes Dev.*, **3**, 431–437.
118. Calnan, B. J., Tidor, B., Biancalana, S., Hudson, D. and Frankel, A. D. (1991) *Science*, **252**, 1167–1171.
119. Lütcke, H., High, S., Romisch, K., Ashford, A. J. and Dobberstein, B. (1992) *EMBO J.*, **11**, 1543–1551.
120. Calnan, B. J., Biancalana, S., Hudson, D. and Frankel, A. D. (1991) *Genes Dev.*, **5**, 201–210.
121. Nagai, K., Oubridge, C., Jessen, H.-H., Li, J. and Evans, P. R. (1990) *Nature*, **348**, 515–520.
122. Weeks, K. M., Ampe, C., Schultz, S. C., Steitz T. A. and Crothers, D. M. (1990) *Science*, **249**, 1281–1285.
123. Gatignol, A., Buckler, C. and Jeang, K. (1993) *Mol. Cell. Biol.*, **13**, 2193–2202.
124. Connolly, T. and Gilmore, R. (1993) *J. Cell Biol.*, **123**, 799–807.
125. Bourne, H. R., Sanders, D. A. and McCormick, F. (1991) *Nature*, **349**, 117–127.
126. van der Blieck, A. M., Redelmeyer, T. E., Damke, H., Tisdale, E. J., Meyerowitz, E. M. and Schmid, S. L. (1993) *J. Cell Biol.*, **122**, 553–563.
127. Herskovits, J. S., Burgess, C., Obar, R. A. and Vallee, R. B. (1993) *J. Cell Biol.*, **122**, 565–578.
128. Rothman, J. H., Raymond, C. K., Gilbert, T., O'Hara, P. and Stevens, T. (1990) *Cell*, **61**, 1063–1074.
129. Bourne, H. R., Sanders, D. A. and McCormick, F. (1990) *Nature*, **348**, 125–132.
130. Milburn, M. V., Tong, L., deVos, A. M., Brunger, A., Yamaizumi, Z., Nishimura, S. and Kim, S.-H. (1990) *Science*, **247**, 939–945.
131. Berchtold, H., Reshetnikova, L., Reiser, C. O. A., Schirmer, N. K., Sprinzl, M. and Hilgenfeld, R. (1993) *Nature*, **365**, 126–132.
132. Boguski, M. S. and McCormick, F. (1993) *Nature*, **366**, 643–654.
133. Connolly, T., Rapiejko, P. J. and Gilmore, R. (1991) *Science*, **252**, 1171–1173.
134. Althoff, S., Stevens, S. and Wise, J. A. (1994) Submitted for publication.
135. Lai, C.-C., Boguski, M., Broik, D. and Powers, S. (1993) *Mol. Cell. Biol.*, **13**, 1345–1352.
136. Farnsworth, C. L. and Feig, L. A. (1991) *Mol. Cell. Biol.*, **11**, 4822–4829.
137. Ziman, M., O'Brien, J. M., Ouellette, L. A., Church, W. R. and Johnson, D. L. (1991) *Mol. Cell. Biol.*, **11**, 3537–3544.
138. Lütcke, H., Prehn, S., Ashford, A. J., Remus, M., Frank, R. and Dobberstein, B. (1993) *J. Cell Biol.*, **121**, 977–985.
139. Herz, J., Flint, N., Stanley, K., Frank, R. and Dobberstein, B. (1990) *FEBS Lett.*, **276**, 103–107.
140. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990) *J. Mol. Biol.*, **215**, 403–410.
141. Chant, J., Corrado, K., Pringle, J. R. and Herskowitz, I. (1991) *Cell*, **65**, 1213–1224.
142. Poritz, M. A., Bernstein, H. D., Strub, K., Zopf, D., Wilhelm, H. and Walter, P. (1990) *Science*, **250**, 1111–1117.
143. Brown, S. (1987) *Cell*, **49**, 825–833.
144. Brown, S. (1991) *J. Mol. Biol.*, **209**, 79–90.
145. Bourgaize, D. B. and Fournier, M. J. (1987) *Nature*, **325**, 281–284.
146. Brown, S. and Fournier, M. J. (1984) *J. Mol. Biol.*, **178**, 533–550.
147. Bourgaize, D. B., Phillips, T. A., VanBogelen, R. A., Jones, P. G., Neidhardt, F. C. and Fournier, M. J. (1990) *J. Bacteriol.*, **172**, 1151–1154.
148. Gething, M.-J. and Sambrook, J. (1992) *Nature*, **355**, 33–45.
149. Kohno, K., Normington, K., Sambrook, J., Gething, M.-J. and Mori, K. (1993) *Mol. Cell. Biol.*, **13**, 877–890.
150. Benson, S. A., Hall, M. N. and Silhavy, T. J. (1985) *Annu. Rev. Biochem.*, **54**, 101–134.
151. Luirink, J., High, S., Wood, H., Giner, A., Tollervey, D. and Dobberstein, B. (1992) *Nature*, **259**, 741–743.
152. Pugsley, A. P. (1993) *Microbiol. Rev.*, **5**, 50–108.
153. Amaya, Y. and Nakano, A. (1991) *FEBS Lett.*, **282**, 325–328.
154. Gropp, R., Gropp, F. and Betlach, M. C. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 1204–1208.
155. He, F., Yaver, D., Beckerich, J.-M., Ribes, V., Ogrydziak, D. and Gaillardin, C. (1990) *Curr. Genet.*, **17**, 289–292.
156. Schlenstedt, G., Gudmundson, G. H., Bonan, H. G. and Zimmerman, R. (1990) *J. Biol. Chem.*, **265**, 13960–13968.
157. Sogin, M. L., Gunderson, J. H., Elwood, H. J., Alonso, R. A. and Peattie, D. A. (1989) *Science*, **243**, 75–77.
158. Higgins, D. G. and Sharp, P. M. (1989) *CABIOS*, **5**, 151–153.

APPENDIX: ACCESSION NUMBERS FOR PROTEIN SEQUENCES

Srp54p homologs

Canine: sp | P13624; Mouse: sp | P14576; *S.cerevisiae*: sp | P20424; *S.pombe* : sp | P21565; *E.coli*: *ffh*: sp | P07019; *M.mycoides:* sp | Q01442; *B.subtilis*: gb | S64117; *A.thaliana* (cyto): gb | L19997; *A.thaliana* (chloro): gb | Z21970; *C.elegans* (partial): gb | Z14549.

SRα homologs

E.coli ftsY: sp | P10121; *S.solfataricus*: sp | P27414; *S.cerevisiae*: sp | P32916; Human: sp | P08240; Canine: sp | P06625; *N.gonorrhoeae pilA*: sp | P14929.

Other SRP proteins

Human Srp19p: sp | P09132; *S.cerevisiae SEC65*: sp | P29478; Mouse Srp14p: sp | P16254; Human Srp14p: gb | X73459; Canine Srp9p: sp | P21262; *C.elegans* Srp9p: gb | Z22177; Canine Srp68p: sp | Q00004; Canine Srp72p: gb | X67813.