The EMBO Journal Vol.16 No.13 pp.3757–3766, 1997

The crystal structure of the signal recognition particle *Alu* RNA binding heterodimer, SRP9/14

Darcy E.A.Birse, Ulrike Kapp, Katharina Strub[1], Stephen Cusack[2] and Anders Åberg[3]

European Molecular Biology Laboratory, Grenoble Outstation c/o ILL 156 X, 38042 Grenoble Cedéx 9, France and [1]Département de Biologie Cellulaire, Université de Genève, Sciences III, CH-1211, Geneva 4, Switzerland

[3]Present address: Department of Molecular Biophysics, Centre for Chemistry and Chemical Engineering, Lund University, PO Box 124, S-221 00 Lund, Sweden

[2]Corresponding author

The mammalian signal recognition particle (SRP) is an 11S cytoplasmic ribonucleoprotein that plays an essential role in protein sorting. SRP recognizes the signal sequence of the nascent polypeptide chain emerging from the ribosome, and targets the ribosome–nascent chain–SRP complex to the rough endoplasmic reticulum. The SRP consists of six polypeptides (SRP9, SRP14, SRP19, SRP54, SRP68 and SRP72) and a single 300 nucleotide RNA molecule. SRP9 and SRP14 proteins form a heterodimer that binds to the *Alu* domain of SRP RNA which is responsible for translation arrest. We report the first crystal structure of a mammalian SRP protein, that of the mouse SRP9/14 heterodimer, determined at 2.5 Å resolution. SRP9 and SRP14 are found to be structurally homologous, containing the same α-β-β-β-α fold. This we designate the *Alu* binding module (*Alu* bm), an additional member of the family of small α/β RNA binding domains. The heterodimer has pseudo 2-fold symmetry and is saddle like, comprising a strongly curved six-stranded amphipathic β-sheet with the four helices packed on the convex side and the exposed concave surface being lined with positively charged residues.

Keywords: *Alu* domain/crystal structure/RNA binding/signal recognition particle (SRP)/translation regulation

Introduction

The mammalian Signal Recognition Particle (SRP) plays an essential role in targeting of secretory and membrane proteins to the rough endoplasmic reticulum (RER) membrane (for reviews, see Walter and Johnson, 1994; Lütcke, 1995). Targeting occurs co-translationally and translocation across the RER membrane begins before polypeptide synthesis is complete. The SRP acts in three distinct ways: (i) it binds the signal sequence of the nascent polypeptide to be translocated, which is exposed on the surface of the translating ribosome; (ii) it temporarily retards the nascent polypeptide from further elongation; and (iii) it mediates docking of the SRP–ribosome–nascent polypeptide chain complex to the RER membrane via the heterodimeric SRP-receptor (SR). With the engagement of this machinery, the SRP is detached from the complex and recycled, and co-translational translocation proceeds. GTP hydrolysis plays an important role in the SRP cycle, one SRP protein and both subunits of the SR contain G-domains. The GTPase cycle of SRP, modulated by the ribosome, provides the regulatory link between translation and translocation machineries (Bacher *et al.*, 1996).

In vitro studies show that mammalian SRP9 and SRP14 proteins form a tight heterodimer SRP9/14 in the absence of SRP RNA and bind specifically to a region of the SRP RNA which includes both the 3′ and 5′ ends (Figure 1) (Strub *et al.*, 1991). The binding of SRP9/14 to the SRP *Alu* RNA is stoichiometric (Walter and Blobel, 1983; Bovia *et al.*, 1994), of high affinity (<0.1 nM; Janiak *et al.*, 1992) and independent of other SRP proteins (Strub and Walter, 1990). The part of SRP comprising SRP9/14 complexed with RNA forms a distinct structural domain known as the *Alu* domain due to the homology of the RNA sequences with the *Alu* family of repetitive DNA sequences and the small cytoplasmic *Alu* RNAs (sc*Alu*s) (Weiner, 1980; Chang *et al.*, 1996). The *Alu* domain of SRP mediates the specific pause(s) in the synthesis of nascent ER-targeted proteins whose signal sequence has been bound by SRP54 (Siegel and Walter, 1988). Both the mechanism and the functional rationale for the elongation arrest activity are unknown.

For structural and functional studies a fusion protein, denoted SRPΦ14-9, has been constructed using SRP9 and SRP14 from *Mus musculus*, which can functionally replace the SRP9/14 heterodimeric subunit in the SRP (Bovia *et al.*, 1994). SRPΦ14-9 binds SRP *Alu* RNA and functions, indistinguishably from wild-type, in elongation arrest and release of elongation arrest (Bovia *et al.*, 1994). Here, we describe the crystal structure of the SRPΦ14-9 fusion protein at 2.5 Å resolution and discuss features of the structure which may be relevant to the RNA binding and heterodimerization properties of SRP9/14.

Results and discussion

Solving the SRP9/14 structure by MIR methods

SRPΦ14-9 was purified and crystallized as described elsewhere (Birse *et al.*, 1996). The structure of the SRP9/14 heterodimer was determined by multiple isomorphous replacement (MIR) methods using mercury, platinum and selenium as heavy-atom derivatives (Table I and Materials and methods). All data were collected on flash-frozen crystals at 100 K. Solvent-flattened electron density maps calculated to 2.8 Å resolution using phases from the three heavy-atom derivatives were suitable to trace the SRP9/14 polypeptide chain. The structure has been refined to 2.5 Å resolution giving a model with a final R_{cryst} =

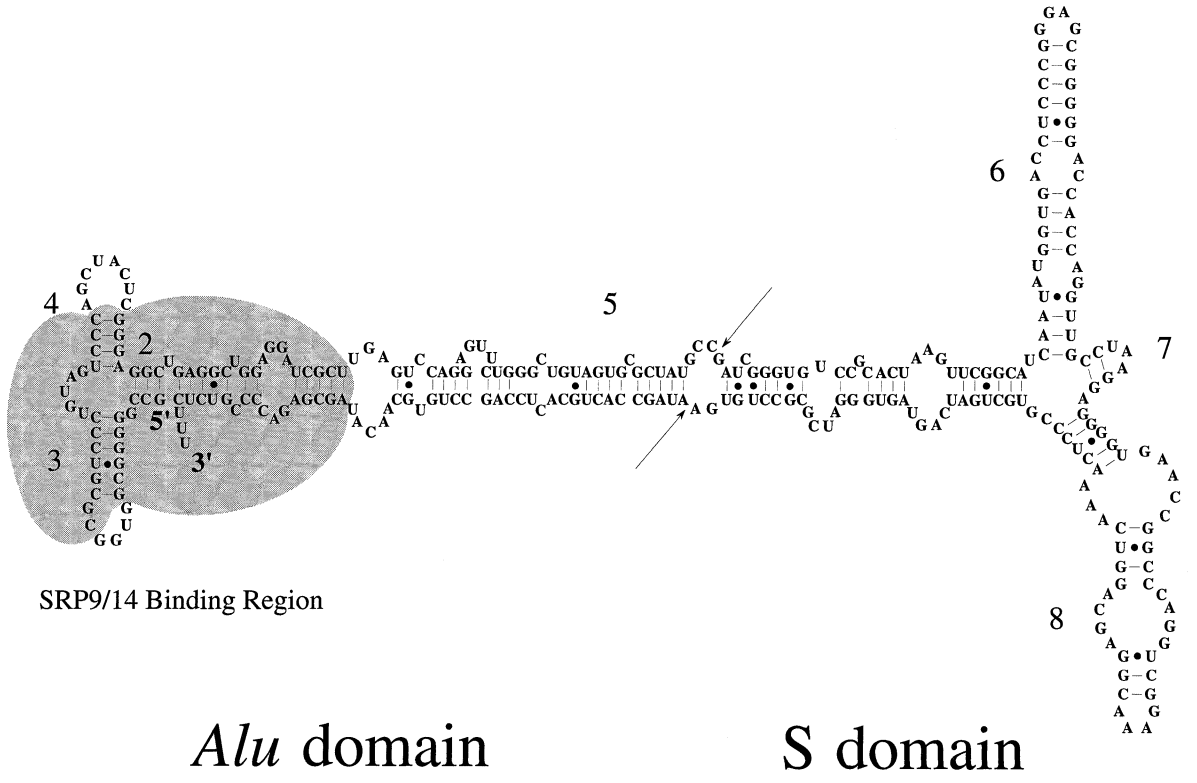

Fig. 1. Model of the secondary structure of SRP RNA from *Homo sapiens*. The shaded area illustrates the proposed SRP9/14 binding region as determined by chemical footprinting (Strub *et al.*, 1991). Base pairs supported by comparative sequence analysis are indicated with straight lines and G–U pairs are indicated with filled circles. Domains 2–8 of the SRP RNA are marked according to the nomenclature of Larsen and Zwieb (1991). The arrows indicate the experimentally determined micrococcal nuclease cleavage sites (Gundelfinger *et al.*, 1983) which separates the *Alu* domain from the S domain of the signal recognition particle. The diagram is adapted from those of Larsen and Zwieb (1991, 1996), Bovia and Strub (1996) and Zwieb *et al.* (1996).

24.8% and R_{free} = 29.9% with excellent stereochemistry (see Materials and methods).

Overall structure

The main chains of SRP9 and SRP14 proteins both fold into three β-strands and two α-helices with a α_1-β_1-β_2-β_3-α_2 connectivity. As a heterodimer, the two proteins form a six-stranded anti-parallel β-sheet stacked against the four α-helices with pseudo 2-fold symmetry (Figure 2A). The structure is compact with a central hydrophobic core sandwiched between the β-sheet and the four α-helices. Both the β-sheet and the four α-helices are amphipathic, having predominantly non-polar side chains interacting in the hydrophobic core and polar residues in solvent-exposed regions. The four α-helices are inclined by ~27° relative to the β-strands. The anti-parallel strands (β_1, β_2 and β_3) are connected via short hairpin loops except for the larger loop (22 residues) between β_1 and β_2 in SRP14, which is partially disordered in the crystal structure. The 10 C-terminal residues of SRP9 extend outward from the heterodimer, having well-defined electron density, forming a long arm that makes contact with the β-sheet of a symmetry-related molecule. The four N-terminal residues of SRP14 form a short β-strand ($14\beta_N$) at one end of the β-sheet, parallel to $14\beta_3$. At the other extremity of the β-sheet, clear electron density is observed for an eighth strand anti-parallel to $9\beta_3$, but with no visible connections to any other part of the molecule. The shape of the electron density for the side chains and stereochemical considerations suggest that this most likely corresponds to five residues of the linker peptide in the SRPΦ14-9 fusion construct (QGGEQK). The four N-terminal residues of SRP9 are disordered. They may have been displaced by the linker peptide and thereby prevented from forming a parallel strand analogous to $14\beta_N$. There is no electron density for the lysine-rich C-terminus of SRP14, nor the artificial N-terminal Φ extension of the fusion construct.

The heterodimer interface is formed by the anti-parallel strands $9\beta_1$ and $14\beta_1$ and the two anti-parallel helices $9\alpha_2$ and $14\alpha_2$. Interdigitation of hydrophobic residues from these strands and helices and six main chain hydrogen bonds between $9\beta_1$ and $14\beta_1$ (residues 27–31 of SRP9 and 26–30 of SRP14) are the major contributions to the dimer interface stability. In addition, there is a hydrogen bond between the side chains of SRP9-His66 and SRP14-Tyr83 buried within the hydrophobic core. Outside the hydrophobic core, additional SRP9–SRP14 interactions are made between the peptide 91–95 near the C-terminus of SRP14 which wraps along the edge of SRP9 in the vicinity of the loops between $9\beta_1$ and $9\beta_2$ and $9\beta_3$ and $9\alpha_2$. SRP14-Lys95 makes two main chain hydrogen bonds to SRP9-Asp54 and SRP9-Ala56 and a salt bridge with SRP9-Asp54, and the main chain amide of SRP14-Gly93 is hydrogen bonded to the hydroxyl group of SRP9-Tyr31. Also, SRP14-Leu94 is inserted into the hydrophobic core. The α-helices provide additional dimer stability through the polar side chain interactions of SRP9-Glu63 with

Table I. Summary of data collection statistics, phase calculation and refinement[a]

	Native	Mercury (I)	Mercury (II)	SeMet	Platinum
Cell parameters (Å)	$a = b = 69.0$	$a = b = 69.2$	$a = b = 69.0$	$a = b = 69.0$	$a = b = 68.8$
	$c = 90.4$	$c = 89.8$	$c = 90.3$	$c = 90.3$	$c = 89.4$
Collected	ID2-ESRF	Siemens R.A.	ID2-ESRF	ID2-ESRF	BM14-ESRF
Exposure time	10 s/°osc	3600 s/°osc	15 s/°osc	25 s/°osc	60 s/°osc
Detector	Mar Research IP	Mar Research IP	Mar Research IP	Mar Research IP	I.I/CCD
Resolution (Å)	2.53	3.08	2.90	2.90	2.46
Reflections	36 956	18 618	24 031	22 438	42 427
Unique reflections	7634	4358	5042	5063	8188
Average $I/\sigma I$	5.8 (2.4)	9.2 (3.0)	8.8 (11.9)	5.8 (2.8)	7.6 (2.4)
$I/\sigma I > 3$	91.6 (74.8)	84.2 (65.8)	93.3 (79.4)	91.3 (78.0)	83.7 (57.5)
Completeness (%)	98.7 (97.2)	99.6 (99.9)	78.9 (81.8)	97.5 (99.9)	99.0 (99.9)
Multiplicity	5.0 (4.2)	4.2 (4.3)	5.9 (5.0)	4.4 (4.7)	5.13 (4.8)
R_{merge} (%)[b]	6.9 (21.8)	7.6 (20.4)	8.8 (27.5)	8.1 (28.9)	7.3 (28.4)
Number of sites	–	3	3	6	1
Phasing power[c]					
centric	–	0.53	0.50	0.71	0.75
accentric	–	0.68	0.71	0.95	0.96
R_{Cullis} [d]					
centric	–	0.85	0.91	0.85	0.81
accentric	–	0.92	0.92	0.86	0.86
Figure of merit[e]					
centrics	0.6867				
all data	0.4941				
Refinement (∞–2.5 Å)					
R_{cryst} (%)[f]	24.8				
R_{free} (%)[g]	29.9				

[a]Values within parentheses indicate data in highest resolution bin (2.64–2.53 Å resolution).
[b]$R_{merge} = \Sigma_{hkl}\ \Sigma_i|I_i–I|/\Sigma_{hkl}\ \Sigma_i\ I_i$
[c]Phasing power = $<F_H>/<E>$, where $<F_H>$ is the mean calculated heavy-atom structure factor amplitude, and $<E>$ is the mean estimated lack of closure.
[d]$R_{Cullis} = <E>/<\text{iso}>$, where $<E>$ is the mean estimated lack of closure and $<\text{iso}>$ is the isomorphous difference.
[e]Figure of merit = $m = |F(\text{hkl})_{best}|/|F(\text{hkl})|$ for a reflection (hkl), where $F(\text{hkl})_{best} = \Sigma P(\alpha_l)F_{hkl}(\alpha_i)/\Sigma P(\Sigma_i)$, where (α_i) is all phase angles and $P(\alpha_i)$ is the probability for a reflection F(hkl) to take a phase angle α.
[f]$R_{cryst} = \Sigma_{hkl}|F_{obs}–F_{calc}|/\Sigma_{hkl}|F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes, respectively.
[g]R_{free} is calculated using 8% of the data chosen randomly and omitted from refinement.

SRP14-Arg88, and SRP14-Gln80 which hydrogen bonds with SRP9-Ser67 and SRP9-Arg71. Finally, the side chains of SRP9-Arg32 and SRP14-Ser25 also form a hydrogen bond.

Structural homology from dissimilar sequences

SRP9 and SRP14 are structurally homologous but differences are observed in the loop between strands β_1 and β_2 (residues 33–54 in SRP14) and the offset angles between the two α-helices (~15° in SRP9 and ~7° in SRP14). The root mean square (r.m.s.) distance of the α-carbon positions between α_1, β_1, β_2 and β_3 (38 residues) of the two polypeptides is 1.0 Å. Using the structure, we have constructed a multiple sequence alignment for SRP9 and SRP14 proteins from various eukaryotes (Figure 3). Although the number of identical residues between SRP9 and SRP14 is very low, the pattern of hydrophobic residues which gives rise to the hydrophobic core is conserved and to a lesser extent, the positions of solvent-exposed arginines and lysines. A section of the longer loop between strands β_1 and β_2 loop of SRP14 is apparently important for specific SRP *Alu* RNA binding (N.Bui *et al.*, 1997) consistent with the fact that it is largely disordered in the crystal structure in the absence of RNA. Another significant difference between SRP9 and SRP14 is the interaction with SRP9 of the C-terminal region following helix α_2 of SRP14, as described above. These interactions could account for the much more detrimental effect on heterodimer stability of C-terminal truncations of SRP14 beyond –10 residues compared with SRP9. The arginine- and lysine-rich extreme C-terminus of SRP14 (Figure 2A) is also a distinguishing feature, but is disordered in the crystal structure and apparently dispensable for RNA binding (Bovia *et al.*, 1994; N.Bui *et al.*, 1997).

The structural homology between SRP9 and SRP14 suggests that the SRP *Alu* RNA binding protein may have originally been a homodimer which evolved into the presently observed heterodimer by gene duplication. This could perhaps be explained by the need to accommodate asymmetry in the co-evolving SRP *Alu* RNA. This argument could also account for the situation in *Saccharomyces cerevisiae* where the SRP14 homologue is larger than its higher eukaryotic counterparts (Hann and Walter, 1991) yet the 5′ end of the SRP RNA (scR1; Felici *et al.*, 1989), as in the case of *Schizosaccharomyces pombe* (Strub *et al.*, 1991), is much simpler. A search through the entire yeast genome has revealed no SRP9 homologue, nor has one been characterized experimentally (Hann and Walter, 1991; Brown *et al.*, 1994). From these observations one can hypothesize that either the yeast SRP14 functions as a monomer, possibly with its longer loop replacing SRP9, or that it forms a homodimer.

Putative RNA binding surface

Electrostatic interactions between basic residues and the phosphates of nucleic acids are generally thought to

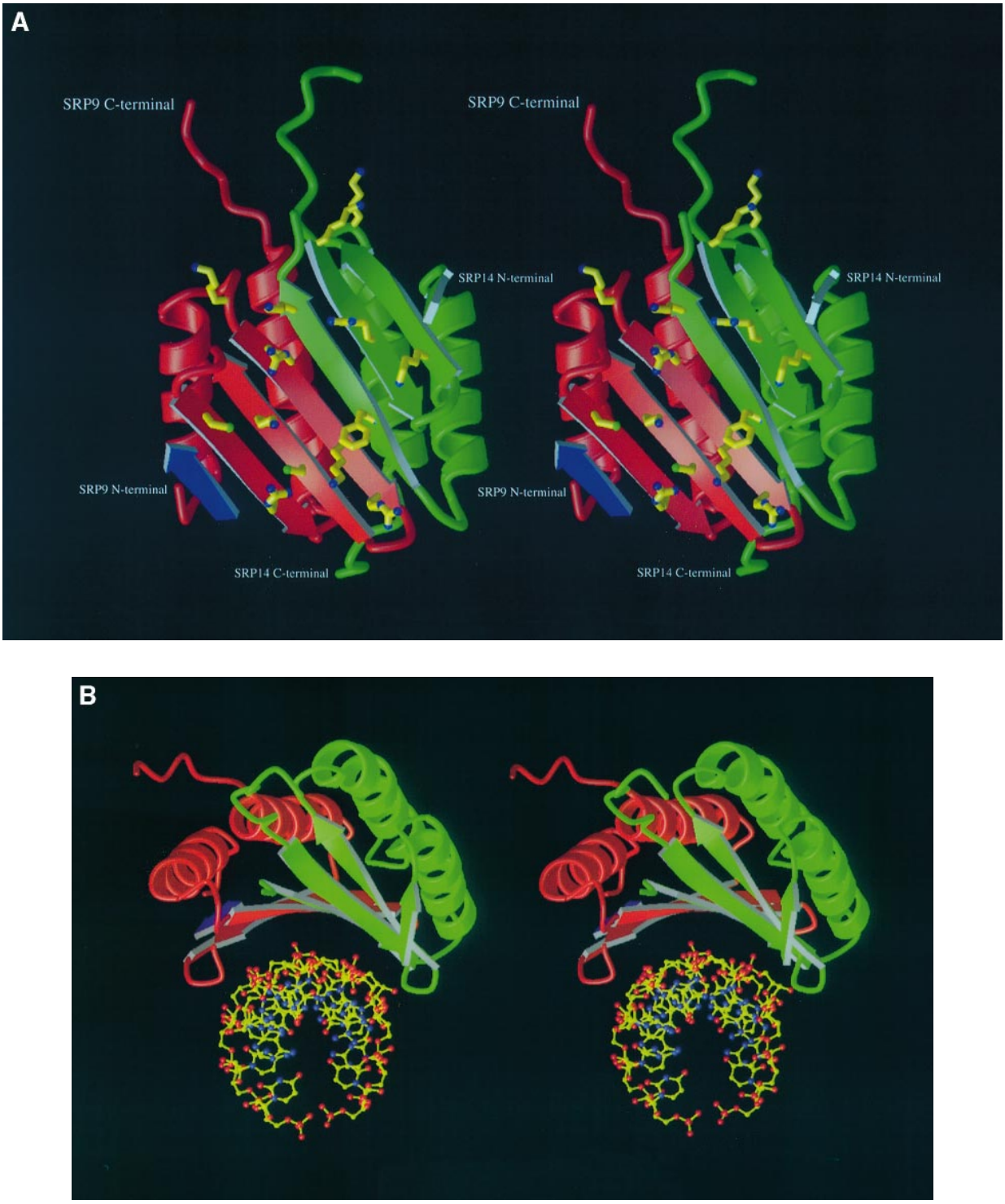

Fig. 2. (**A**) Stereo diagram of the SRP9/14 heterodimer, viewed looking onto the β-sheet surface, showing secondary structure elements. Basic residues which project out of the β-sheet surface are depicted with their side chains (SRP9-Arg26, SRP9-Lys30, SRP9-Arg32, SRP9-Lys41, SRP9-Arg52, SRP14-Lys31, SRP14-Lys55, SRP14-Arg59, SRP14-Lys66 and loop residues SRP9-Lys24 and SRP14-Lys74). Also shown is residue SRP14-Phe27 protruding out from the β-sheet surface and two cysteines (SRP9-Cys39 and SRP9-Cys48) implicated in NEM studies of SRP9. The blue strand represents a region of the fusion linker which may displace a putative SRP9 N-terminal parallel β-strand. Diagrams were made using the program MOLSCRIPT (Kraulis, 1991) and RASTER3D (Merritt and Murphy, 1994). (**B**) A side view of the SRP9/14 heterodimer superimposed with a helical RNA molecule. The curvature of the β-sheet conforms to that of the RNA (modelled with a six based-paired double-stranded helical region of tRNA).

be the important forces involved in the formation and stabilization of protein–RNA complexes. We have therefore examined the nature of the SRP9/14 heterodimer surface using electrostatic potential calculations (Nicholls *et al.*, 1993). As shown in Figure 4, these calculations strikingly reveal that the β-sheet possesses a highly positive charged concave surface due to the abundance of exposed basic residues (SRP9-Arg26, SRP9-Lys30, SRP9-Arg32,

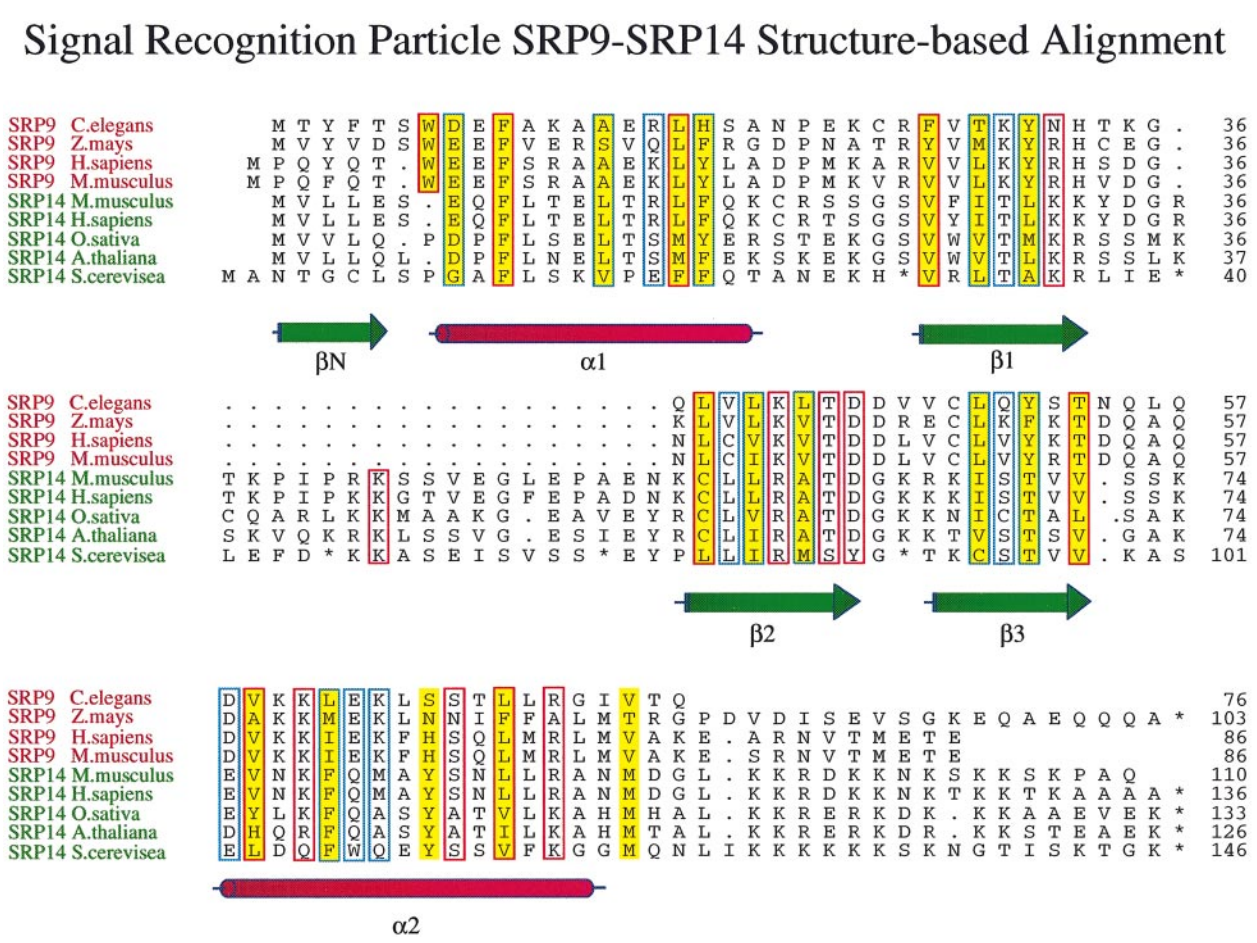

Fig. 3. Structure-based sequence alignment comparing signal recognition particle SRP9 and SRP14 proteins. Alignment follows SRP9 and SRP14 sequences from *Mus musculus* for which the secondary structures are represented by β-strands (green arrows) and α-helices (red cylinders). Secondary structure assignments are based on DSSP (Kabsch and Sander, 1983). Highly conserved residues are shown inside red boxes, residues with conservative substitutions having the same physical–chemical properties are shown in blue boxes and residues contributing to the hydrophobic core are shown on yellow background. Asterisks indicate omitted sequence.

SRP9-Lys41, SRP9-Arg52, SRP14-Lys31, SRP14-Lys55, SRP14-Arg59, SRP14-Lys66 and loop residues SRP9-Lys24 and SRP14-Lys74). In the crystal structure, a phosphate ion is observed bound to SRP9-Arg26, and SRP14-Lys31 and SRP14-Arg59. The proposed binding site of SRP9/14 on SRP *Alu* RNA (Figure 1), as determined by chemical footprinting (Strub *et al.*, 1991), covers a substantial region of RNA and at least 86 nucleotides of RNA are required to maintain highly specific binding (O.Weichenrieder, S.Cusack and K.Strub, unpublished results). Some of this RNA is likely to be double-stranded (Zwieb *et al.*, 1996) and Figure 2B shows that the heterodimer has indeed the right curvature to form a saddle on a double-stranded domain of *Alu* RNA. Also protruding from the concave β-sheet surface is a solvent-exposed aromatic residue, SRP14-Phe27, surrounded by positively charged residues. This is a candidate to interact with an RNA base as is often observed in protein–RNA complexes (Mattaj and Nagai, 1995).

The hypothesis that important RNA binding interactions occur at the concave β-sheet surface is consistent with experiments using the specific sulphydryl alkylating reagent *N*-ethylmaleimide (NEM; Siegel and Walter, 1988). These experiments showed that SRP9 contains two sulphydryl groups accessible to alkylation in the free protein, but which are protected from modification in the RNA complex. The conclusion was that these cysteines lie along an RNA binding region in the SRP9/14 heterodimer. The structure supports this observation by showing that the two highly conserved residues, SRP9-Cys39 and SRP9-Cys48, are indeed on the solvent-exposed β-sheet surface, whereas the two cysteines of SRP14 are buried. In the structure, the SRP9-Cys39 is covalently bound to a β-mercaptoethanol molecule arising from the purification procedure (Birse *et al.*, 1996).

A number of deletion and point mutations have been made on both SRP14 and SRP9 with the aim of identifying determinants for RNA binding and heterodimerization (N.Bui *et al.*, 1997). These results do not positively identify the basic surface described above as being involved in specific *Alu* RNA binding, nor do they exclude it since neither the basic residues cited above, nor SRP14-Phe27, have yet been investigated in a conclusive fashion. Furthermore, it may be that multiple mutations are required to get a significant deterioration of RNA binding ability. On the other hand, it is shown that deletions of the first part (residues 33–43) of the long SRP14 loop as well as point mutations of SRP9-Glu15 on $9\alpha_1$ and SRP9-Asp21-Pro22, all of which are on one edge of the molecule, are detrimental to specific *Alu* RNA binding without affecting

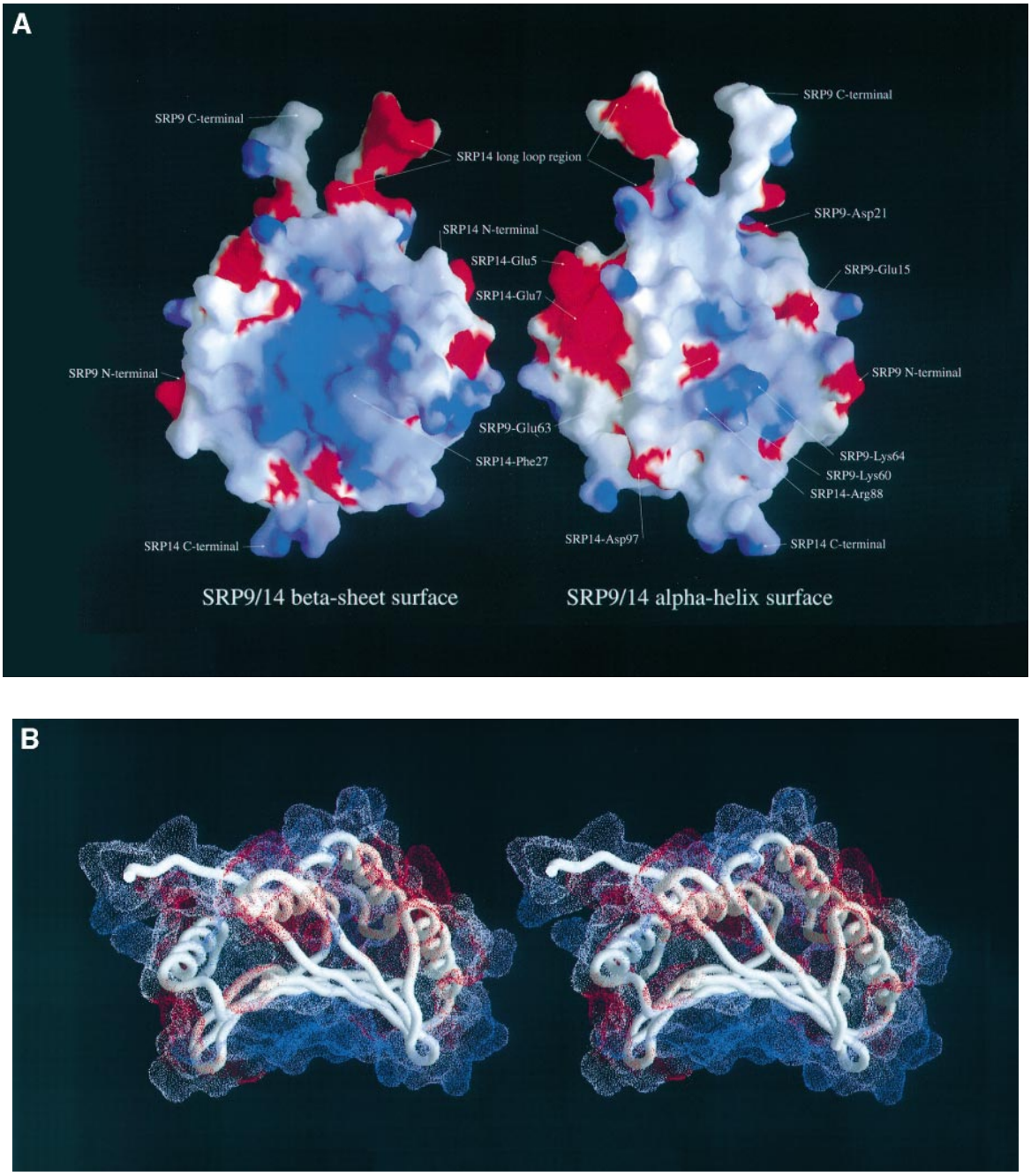

Fig. 4. (**A**) Electrostatic potential representation of the SRP9/14 heterodimer calculated using the program GRASP (Nichollis *et al.*, 1993). The solvent-exposed surface of the β-sheet is shown illustrating the positive (blue), negative (red) and neutral (white) electrostatic potentials. Depicted are the β-sheet and α-helical surfaces of the SRP9/14 model. (**B**) A stereo diagram showing the electrostatic potential of the solvent-exposed surface of the SRP9/14 model.

heterodimerization (N.Bui *et al.*, submitted). These results, together with the ability of *Alu* RNA to rescue certain dimerization-deficient mutants, point to the fact that the RNA binds across both subunits of the heterodimer. Consideration of the size of the RNA relative to that of the SRP9/14 heterodimer, suggests that several regions, including those identified by mutation, and the β-sheet surface most probably interact with the SRP *Alu* RNA, but it should be borne in mind that SRP9/14 may functionally interact, not only with *Alu* RNA, but also with other RNAs, for example rRNA.

The growing family of RNA binding modules

The SRP9/14 heterodimer is the latest member of a growing family of small α/β RNA binding proteins examples of which are: the ribonucleoprotein (RNP) domain (Nagai *et al.*, 1990; Oubridge *et al.*, 1994); the double-stranded RNA binding domain (dsRBD) (Farrandon *et al.*, 1994; Bycroft *et al.*, 1995; Kharrat *et al.*, 1995); the K homology (KH) domain (Musco *et al.*, 1996); the coat protein of bacteriophage MS2 (Valegård *et al.*, 1990); the translational initiation factor IF3 (Biou *et al.*, 1995); the S1 RNA binding domain (Bycroft *et al.*,

1997); and many ribosomal proteins (Nagai, 1996). The RNP and KH domains, as well as several ribosomal proteins (Liljas and Garber, 1995), belong to the so-called split α-β-α motif differing from the dsRBD, MS2 and SRP9/14 where the sheet is a β-meander. In aminoacyl-tRNA synthetases, a number of different tRNA anti-codon binding modules have also been characterized (Cusack, 1995; Moras and Poterszman, 1996). Interestingly, RNA and DNA binding modules appear to be in general structurally distinct and therefore to have evolved independently.

The α-β-β-α topology of SRP9 and SRP14 proteins is similar to the α-β-β-β-α motif of the dsRBDs but differs in α-helical connectivity and stacking of β-strands with α-helices (Kharrat *et al.*, 1995). A single dsRBD does not bind double-stranded RNA (dsRNA) in a sequence-specific manner (Farrandon *et al.*, 1994; Bycroft *et al.*, 1995), and it has been proposed that multiple modules may be necessary for specific RNA recognition and binding (St Johnson *et al.*, 1992; Farrandon *et al.*, 1994) as for instance found for dsRBDs in stau protein (Kim-Ha *et al.*, 1995) or KH domains in Bic-C protein (Mahone *et al.*, 1995). Furthermore, it has been suggested that dsRBDs interact with dsRNA solely via the loop regions (Farrandon *et al.*, 1994; Bycroft *et al.*, 1995; Kharrat *et al.*, 1995), since there are few positively charged residues on the β-sheet surface accessible to interact with dsRNA. The occurrence of a highly positively charged β-sheet surface in SRP9/14 thus suggests that RNA binding properties of SRP9/14 differ from dsRBDs, although the results of N.Bui *et al.* (1997) show that one edge of the molecule is also involved in specific interactions with *Alu* RNA.

The MS2 protein makes contact with an RNA hairpin via residues in the 10-stranded anti-parallel β-sheet formed by a dimer of the MS2 coat protein (Valegård *et al.*, 1994). This dimer bears some resemblance to the SRP9/14 heterodimer but includes two α-helices, one from each subunit, interdigitating in an anti-parallel fashion. Indeed, it has been shown that individual MS2 subunits do not fully fold, the final conformation depending on their mutual association (Peabody and Lim, 1996). In the case of SRP9/14 it is not known to what extent the individual proteins fold before heterodimerization, although the determination of the crystal structure of the murine SRP9 protein alone is in progress and may give some indications (Doublié *et al.*, 1996). The MS2 coat protein binds the RNA hairpin primarily by hydrogen bonding with residues of the β-sheet surface facilitated by a tyrosine which stacks with a cytosine base in the RNA. In the case of SRP9/14 it seems likely, considering the large number of exposed basic residues, that contacts to the phosphate backbone may play a more important role in RNA interactions.

The RNA binding properties of SRP9 and SRP14 are reminiscent of the DNA binding properties of a group of transcriptional activators which require the formation of homo- or heterodimers before binding with high affinity to specific DNA sequences (Strub and Walter, 1990; Nelson, 1995). Indeed, the saddle-like form of SRP9/14 is reminiscent of known structures such as the DNA TATA binding protein (TBP) (Nikolov *et al.*, 1992) and the bifunctional PCD/DCoH protein (Endrizzi *et al.*, 1995; Ficner *et al.*, 1995). Furthermore, it has been shown that certain dimerization-deficient mutants of SRP14 can be rescued by *Alu* RNA binding (N.Bui *et al.*, submitted).

In conclusion, the crystal structure of the SRP *Alu* RNA binding SRP9/14 heterodimer reveals a novel RNA binding motif, designated *Alu* binding motif (*Alu* bm), a new member in the growing family of small α/β RNA binding proteins. It also provides the first step towards a structure-based understanding of how the SRP *Alu* domain functions in elongation arrest after signal peptide binding to the S domain of SRP. In order to fulfill its function in elongation arrest, the SRP *Alu* domain presumably has to interfere with the ribosome or other factors involved in elongation. It is currently unknown whether these interactions are made by the protein and/or RNA moieties of the SRP *Alu* domain. The next steps towards further understanding this stage of translational regulation are the determination of the atomic structure of the complete SRP *Alu* domain; the SRP9/14 heterodimer complexed with *Alu* RNA, and to identify with which components of the translation machinery the SRP *Alu* domain interacts.

Materials and methods

Expression and purification

SRPΦ14-9 and selenomethionine-incorporated SRPΦ14-9 protein were overexpressed and purified as described elsewhere (Birse *et al.*, 1996; Doublié *et al.*, 1996).

Crystallization

The SRPΦ14-9 protein was crystallized (Birse *et al.*, 1996) by the hanging drop method in 2.0 M NaH_2/K_2HPO_4, pH 7.7, 2% MPD, 1.0 mM NaN_3 at 4°C with a final protein concentration of 5–8 mg/ml. Crystals formed over 2–3 weeks and were typically $150\times150\times300$ μm^3 in space group $P4_322$ with cell dimensions $a = b = 69.02$ Å, $c = 90.44$ Å. There is one SRPΦ14-9 polypeptide per asymmetric unit. The crystals diffract to beyond 2.5 Å resolution flash-frozen at 100 K using 30% sucrose as cryoprotectant. Derivatives were prepared by soaking crystals in pre-equilibrated cryoprotectant-containing drops with 2.0 mM thimerosal ($C_9H_9HgO_2SNa$) or 1.0 mM $K_2Pt(CN)_4$ for ~24 h.

Data collection and processing

A native data set was collected to 2.5 Å resolution at ID2 (High Brilliance beamline-ESRF). Two mercury (thimerosal) derivative data sets were collected; mercury I, to 3.1 Å resolution on a Siemens rotating anode generator (EMBL-Grenoble) and mercury II, to 2.9 Å resolution at ID2-ESRF. The selenomethionine data set to 2.9 Å resolution was collected on ID2-ESRF. The native, mercury and selenomethionine data sets used 30 cm Mar Research image plate detectors (MAR Research, Hamburg, Germany). A platinum derivative data set was collected to 2.5 Å resolution on BM14 (MAD beamline-ESRF) using an image intensifier/CCD detector. Data were processed using DENZO (Otwinowksi, 1993) and MOSFLM (Leslie, 1992) and scaled using the CCP4 suite of programs (Collaborative Computing Project No. 4, 1994). According to the Matthews coefficient (Matthews, 1968), applying protein density estimations, solvent content in the tetragonal bipyramid crystals is ~34%. A single platinum site was located in difference Patterson maps. Three mercury sites were found by cross-Fourier methods using single isomorphous replacement (SIR) protein phases from the platinum derivative. MIR phases from platinum and mercury were used to locate six ordered selenomethionine sites (out of 10 methionines in the SRPΦ14-9 molecule). Heavy-atom positions were refined and MIR phases to 2.8 Å resolution were calculated using MLPHARE (Collaborative Computing Project No. 4, 1994) to produce the initial MIR electron density map. The map was further improved by solvent flattening, histogram matching and Sayres' equation using the program DM (Collaborative Computing Project No. 4, 1994).

Model building and refinement

Using DM solvent-flattened maps and density skeletons created using the program MAPMAN (Kleywegt and Jones, 1994, 1996), the SRP9

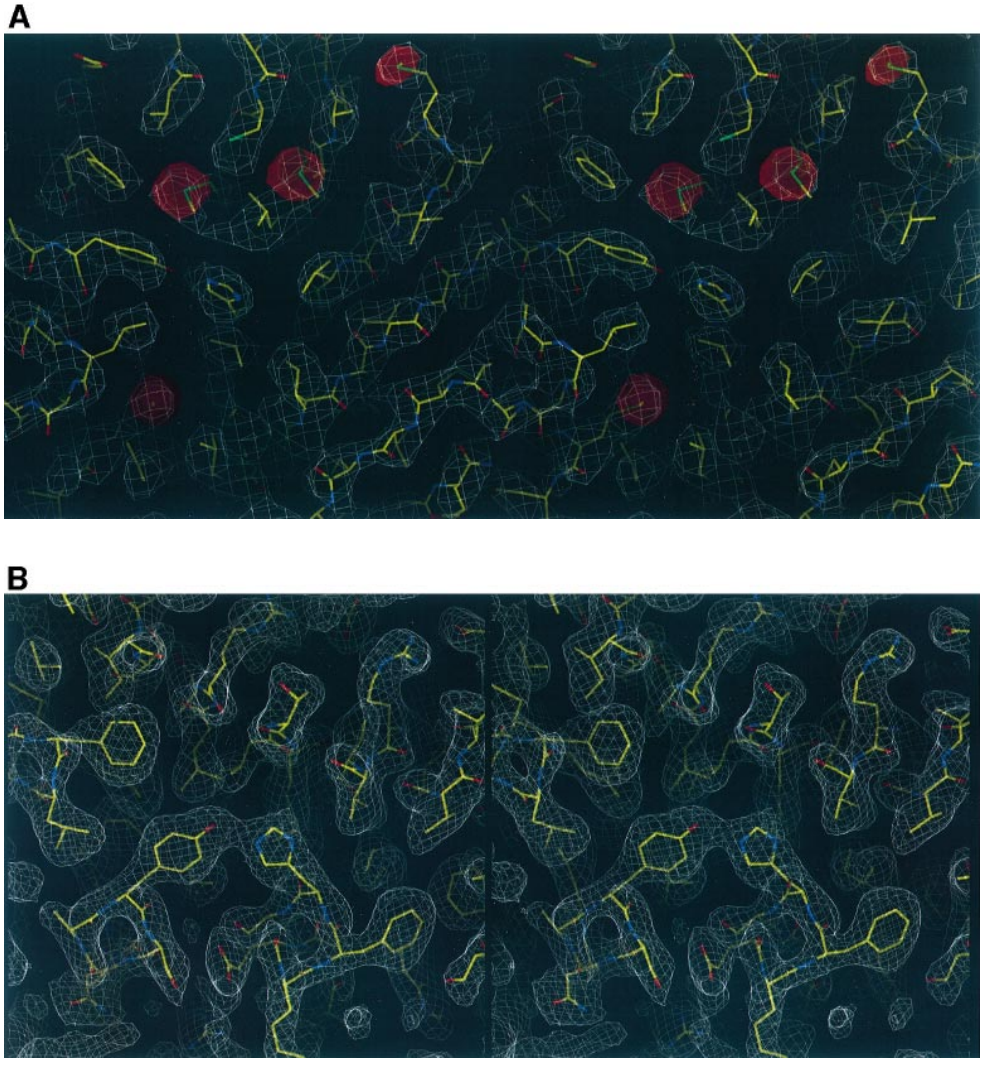

Fig. 5. (**A**) Stereo diagram of electron density representation of solvent-flattened experimental MIR maps contoured at 1.0 σ (in white) and difference Fourier maps of selenomethionine density (in red) with stick-model of SRP9/14 heterodimer. The difference Fourier map for selenomethionine density was calculated using MIR phases from mercury and platinum contoured at 6.0 σ. Methionine sites shown (SRP9-Met23, SRP9-Met73, SRP9-Met70 and SRP14-Met91) of the model superimpose with selenomethionine density. (**B**) Stereo diagram of $2F_{obs}–F_{calc}$ electron density maps with the refined model contoured at 1.2 σ showing the hydrogen bond between SRP9-His66 and SRP14-Tyr83 within the hydrophobic core of the SRP9/14 molecule.

and SRP14 polypeptide chains could be traced. Strand and helical density were clearly interpretable. A polyalanine chain was built into the density for SRP9 and SRP14, followed by the addition of 161 side chains. The characteristic density from the single tryptophan, SRP9-Trp7, provided a starting point for model building. The mercury sites provided sequence markers to locate three cysteines in the model (SRP9-Cys39, SRP9-Cys48 and SRP14-Cys56). The fourth cysteine (SRP9-Cys39) is covalently bound to a β-mercaptoethanol molecule (see 'Putative Alu RNA binding surface'). The single platinum site was found coordinated between SRP9-Arg34 and SRP9-Lys41. Difference Fourier maps located selenomethionine density to position six methionines, SRP9-Met23, SRP9-Met70, SRP9-Met73, SRP14-Met1, SRP14-Met81 and SRP14-Met91 (Figure 5A). The model was refined to 2.5 Å resolution using simulated annealing, positional refinement, *B*-factor refinement and manual rebuilding using the programs XPLOR (Brünger, 1992) and O (Jones *et al.*, 1991). The model includes 77 residues of SRP9 (4–81) and 84 residues of SRP14 (1–34) and (47–97). The SRP14 loop electron density connecting $14\beta_1$–$14\beta_2$ (34–47) is weak suggesting that the loop is flexible. Five C-terminal and three N-terminal residues of SRP9 and 13 C-terminal residues of SRP14 have poorly defined density. In addition, the model does not include the SRPΦ14-9 N-terminal Φ extension (18 residues) nor the entire linker region of which only eight of 17 residues are ordered. In total, disordered regions making up 10% of the SRP9/14 heterodimer (22% of the SRPΦ14-9 fusion protein) could not be modelled. As defined by PROCHECK (Laskowski *et al.*, 1993), there are no residues in disallowed main-chain torsion angle regions and only two residues in the generously allowed regions. The model includes one phosphate ion, a β-mercaptoethanol molecule and 39 water molecules. An electron density map, illustrating a hydrophobic core region of the final refined model, is shown in Figure 5B. The final R_{cryst} for all data is 24.8% and R_{free} is 29.9% for the SRPΦ14-9 model.

Accession numbers

Coordinates and structure factors for SRPΦ14-9 referred to in this paper will be deposited in the Brookhaven Protein Database within 1 year of publication. Sequence accession numbers from GenBank are as follows: 464807, 1362938, 1363236, 201063, 586035, 1362463. *Arabidopsis thaliana* (Y10116), *Oryza sativa* (Y10118) and *Zea mays* (Y10117) sequences (N.Bui, N.Wolff and K.Strub, unpublished results).

Acknowledgements

We would like to thank Andy Thompson (EMBL) and Sean McSweeney (EMBL) for help on the MAD Beamline (BM14), Bjarne Rasmussen (EMBL) at the High Brilliance Beamline (ID2), Michael Wulff (ESRF) at the White Beamline (ID9) at the European Synchrotron Radiation Facility (ESRF), Grenoble, France. We also wish to thank Sylvie Doublié

for helpful discussions. K.S. and S.C. are both members of SRPNET which is supported by the European Union Training and Mobility of Researchers programme and the Swiss Government. K.S. is a fellow of the START programme of the Swiss National Science Foundation. Correspondence and requests for material should be addressed to Stephen Cusack (cusack@embl-grenoble.fr).

References

Bacher,G., Lütcke,H., Jungnickel,B., Rapoport,T.A. and Dobberstein,B. (1996) Regulation by the ribosome of the GTPase of the signal-recognition particle during protein targeting. *Nature*, **381**, 248–251.

Biou,V., Shu,F. and Ramakrishnan,V. (1995) X-ray crystallography shows that translational initiation factor IF3 consists of two compact α/β domains linked by an α-helix. *EMBO J.*, **14**, 4056–4064.

Birse,D.E.A., Doublié,S., Kapp,U., Strub,K., Cusack,S. and Åberg,A. (1996) Crystallization and preliminary crystallographic analysis of the signal recognition particle SRPΦ14-9 fusion protein. *FEBS Lett.*, **384**, 215–218.

Bovia,F., Bui,N. and Strub,K. (1994) The heterodimeric subunit SRP9/14 of the signal recognition particle functions as permuted single polypeptide chain. *Nucleic Acids Res.*, **22**, 2028–2035.

Brown,J.D., Hann,B.C., Medzihradszky,K.F., Niwa,M., Burlingame,A.L. and Walter,P. (1994) Subunits of the *Saccharomyces cerevisiae* signal recognition particle required for its functional expression. *EMBO J.*, **13**, 4390–4400.

Brünger,A. (1992) *X-Plor Version 3.1*. A system for X-ray crystallography and NMR. Yale University, New Haven, CT.

Bui,N., Wolff,N., Cusack,S. and Strub,K. (1997) *RNA*, in press.

Bycroft,M., Grünert,S., Murzin,A.G., Proctor,M. and St Johnston,D. **(**1995) NMR structure of a dsRNA binding domain from *Drosophila* staufen protein reveals homology to the N-terminal domain of ribosomal protein S5. *EMBO J.*, **14**, 3563–3571.

Bycroft,M., Hubbard,T.J.P., Proctor,M., Freund,S.M.V. and Murzin,A.G. (1997) The solution structure of the S1 RNA binding domain: a member of an ancient nucleic acid-binding fold. *Cell*, **88**, 235–242.

Chang,D.-Y., Hsu,K. and Mararia,R.J. (1996) Monomeric sc*Alu* and nascent dimeric *Alu* RNAs induced by adenovirus are assembled into SRP9/14-containing RNPs in HeLa cells. *Nucleic Acids Res.*, **24**, 4165–4170.

Collaborative Computing Project No. 4. (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr.*, **D50**, 760–766.

Cusack,S. (1995) Eleven down and nine to go. *Nature Struct. Biol.*, **2**, 824–831.

Doublié,S., Kapp,U., Åberg,A., Brown,K., Strub,K. and Cusack,S. (1996) Crystallization and preliminary X-ray analysis of the 9 kDa protein of the mouse signal recognition particle and the selenomethionyl-SRP9. *FEBS Lett.*, **384**, 219–221.

Endrizzi,J.A., Cronk,J.D., Wang,W., Crabtree,G.R. and Aber,T. (1995) Crystal structure of DCoH, a bifunctional, protein-binding transcriptional coactivator. *Science*, **268**, 556–559.

Farrandon,D., Elphick,L., Nüsslein-Vohard,C. and St Johnston,D. (1994) Staufen protein associated with the 3′UTR of *bicoid* mRNA to form particles that move in a microtubule dependent manner. *Cell*, **79**, 1221–1231.

Felici,F., Cesarini,G. and Hughes,J.M.X. (1989) The most abundant small cytoplasmic RNA of *S. cerevisiae* has an important function required for normal cell growth. *Mol. Cell. Biol.*, **9**, 3260–3268.

Ficner,R., Sauer,U.H., Stier,G. and Suck,D. (1995) Three-dimensional structure of the bifunctional protein PCD/DCoH, a cytoplasmic enzyme interacting with transcription factor HNF1. *EMBO J.*, **14**, 2034–2042.

Gundelfinger,E.D., Krause,E., Melli,M. and Dobberstein,B. (1983) The organization of the 7SL RNA component of the signal recognition particle. *Nucleic Acids Res.*, **11**, 7363–7374.

Hann,B.C. and Walter,P. (1991) The signal recognition particle in *S. cerevisiae*. *Cell*, **67**, 131–144.

Janiak,F., Walter,P. and Johnson,A.E. (1992) Fluorescence-detected assembly of the signal recognition particle: binding of the two SRP protein heterodimers to SRP RNA is noncooperative. *Biochemistry*, **31**, 5830–5840.

Jones,T.A., Zou,J.Y., Cowan,S.W. and Kjeldgaard,M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr.*, **A47**, 110–119.

Kabsch,W. and Sander,C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, **22**, 2577–2637.

Kharrat,A., Macias,M.J., Gibson,T.J., Nilges,M. and Pastore,A. (1995) Structure of the dsRNA binding domain of *E.coli* RNase III. *EMBO J.*, **14**, 3572–3584.

Kim-Ha,J., Kerr,K. and Macdonald,P.M. (1995) Translational regulation of *oskar* mRNA by Bruno, an ovarian RNA-binding protein, is essential. *Cell*, **81**, 403–412.

Kleywegt,G.J. and Jones,T.A. (1994) Halloween... masks and bones. In Bailey,S., Hubbard,R. and Waller,D. (eds), *From First Map to Final Model*. SERC Daresbury Laboratory, Warrington, UK, pp. 59–66.

Kleywegt,G.J. and Jones,T.A. (1996) xdlMAPMAN and xdlDATA MAN—programs for reformatting, analysis and manipulation of biomacromolecular electron-density maps and reflection data sets. *Acta Crystallogr.*, **D52**, 826–828.

Kraulis,P.J., (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.*, **24**, 946–950.

Larsen,N. and Zwieb,C. (1991) SRP-RNA sequence alignment and secondary structure. *Nucleic Acids Res.*, **19**, 209–215.

Larsen,N. and Zwieb,C. (1996) The signal recognition particle database (SRPDB). *Nucleic Acids Res.*, **24**, 80–81.

Laskowski,R.A., MacAurthur,M.W., Moss,D.S. and Thornton,J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.*, **26**, 282–291.

Leslie,A.G.W. (1992) Recent changes to the MOSFLM package for processing films and image plate data. *CCP4 and ESF-EACMB Newsletter on Protein Crystallography*, Number 26.

Liljas,A and Garber,M. (1995) Ribosomal proteins and elongation factors. *Curr. Opin. Struct. Biol.*, **5**, 721–727.

Lütcke,H. (1995) Review: signal recognition particle (SRP), a ubiquitous initiator of protein translation. *Eur. J. Biochem.*, **228**, 531–550.

Mahone,M., Saffman,E.E. and Lasko,P.F. (1995) Localized *Bicaudal-C* RNA encodes a protein containing a KH domain, the RNA binding motif of FMR1. *EMBO J.*, **14**, 2043–2055.

Mattaj,I.W. and Nagai,K. (1995) Recruiting proteins to the RNA world. *Nature Struct. Biol.*, **2**, 518–522.

Matthews,B.M. (1968) Solvent content of protein crystals. *J. Mol. Biol.*, **33**, 491–497.

Merritt,E.A. and Murphy,M.E.P. (1994) Raster3D version 2.0. A program for photorealistic molecular graphics. *Acta Crystallogr.*, **D50**, 869–873.

Moras,D. and Poterszman,A. (1996) Protein–RNA interactions: getting into the major groove. *Curr. Biol.*, **6**, 530–533.

Musco,G., Stier,G., Joseph,C., Antonietta,M., Morelli,C., Nilges,M., Gibson,T.J. and Pastore,A. (1996) Three-dimensional structure and stability of the KH domain: molecular insights into the fragile X syndrome. *Cell*, **85**, 237–245.

Nagai,K. (1996) RNA–protein complexes. *Curr. Opin. Struct. Biol.*, **6**, 53–61.

Nagai,K., Oubridge,C., Jessen,T.H., Li,J. and Evans,P.R. (1990) Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A. *Nature*, **348**, 515–520.

Nelson,H.C.M. (1995) Structure and function of DNA-binding proteins. *Curr. Opin. Genet. Dev.*, **5**, 180–189.

Nicholls,A., Bharadwaj,R. and Honig,B. (1993) Grasp-graphical representation and analysis of surface properties. *Biophysics J.*, **64**, A166.

Nikolov,D.B., Hu,S.-H., Lin,J., Gasch,A., Hoffmann,A., Horikoshi,M., Chua,N.-H., Roeder,R.G. and Burley,S. (1992) Crystal structure of TFIID TATA-box binding protein. *Nature*, **360**, 40–46.

Otwinowski,Z. (1991) A film processing program for macromolecular crystallography. Yale University, New Haven, CT, USA.

Oubridge,C., Ito,N., Evans,P.R., Teo,C.H. and Nagai,K. (1994) Crystal structure at 1.92 Å resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature*, **372**, 432–438.

Peabody,D.S. and Lim,F. (1996) Complementation of RNA binding site mutations in MS2 coat protein heterodimers. *Nucleic Acids Res.*, **24**, 2352–2359.

Siegel,V. and Walter,P. (1988) 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.

St Johnston,D., Brown,N.H., Gall,J.G. and Janstach,M. (1992) A conserved double-stranded RNA-binding domain. *Proc. Natl Acad. Sci. USA*, **89**, 10979–10983.

Strub,K. and Walter,P. (1990) Assembly of the *Alu* domain of the signal recognition particle (SRP): dimerization of the two protein components is required for efficient binding to SRP RNA. *Mol. Cell. Biol.*, **10**, 777–784.

Strub,K., Moss,J. and Walter,P. (1991) Binding sites of the 9- 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. *Mol. Cell. Biol.*, **11**, 3949–3959.

Valegård,K., Liljas,L., Fridborg,K. and Unge,T. (1990) The three-dimensional structure of the bacterial virus MS2. *Nature*, **345**, 36–41.

Valegård,K., Murray,J.B., Stockley,P.G., Stonehouse,N.J. and Liljas,L. (1994) Crystal structure of an RNA bacteriophage coat protein-operator complex. *Nature*, **371**, 623–626.

Walter,P. and Blobel,G. (1983) Disassembly and reconstitution of signal recognition particle, *Cell*, **34**, 525–533.

Walter,P. and Johnson,A.E. (1994) Signal sequence recognition and protein targeting to the endoplasmic reticulum membrane. *Annu. Rev. Cell Biol.*, **10**, 87–119.

Weiner,A.M. (1980) An abundant cytoplasmic 7S RNA is complementary to the dominant interspersed middle repetitive DNA sequence family in the human genome. *Cell*, **22**, 209–218.

Zwieb,C., Müller,F. and Larsen,N. (1996) Comparative analysis of tertiary structure elements in signal recognition particle RNA. *Folding & Design*, **1**, 315–324.

Received on January 8, 1997; revised on February 21, 1997