Structure and assembly of the Alu domain of the mammalian signal recognition particle

Weichenrieder, O; Wild, K; Strub, Katharina; Cusack, S

How to cite

WEICHENRIEDER, O et al. Structure and assembly of the Alu domain of the mammalian signal recognition particle. In: Nature, 2000, vol. 408, n° 6809, p. 167–173. doi: 10.1038/35041507

This publication URL: https://archive-ouverte.unige.ch/unige:17516
Publication DOI: 10.1038/35041507

Structure and assembly of the *Alu* domain of the mammalian signal recognition particle

Oliver Weichenrieder*, Klemens Wild*, Katharina Strub† & Stephen Cusack*

** European Molecular Laboratory Biology, Grenoble Outstation, BP 156X, F-38042, Grenoble Cedex 9, France*
† Département de Biologie Cellulaire, Université de Genève, Sciences III, CH-1211, Geneva 4, Switzerland

The *Alu* domain of the mammalian signal recognition particle (SRP) comprises the heterodimer of proteins SRP9 and SRP14 bound to the 5′ and 3′ terminal sequences of SRP RNA. It retards the ribosomal elongation of signal-peptide-containing proteins before their engagement with the translocation machinery in the endoplasmic reticulum. Here we report two crystal structures of the heterodimer SRP9/14 bound either to the 5′ domain or to a construct containing both 5′ and 3′ domains. We present a model of the complete *Alu* domain that is consistent with extensive biochemical data. SRP9/14 binds strongly to the conserved core of the 5′ domain, which forms a U-turn connecting two helical stacks. Reversible docking of the more weakly bound 3′ domain might be functionally important in the mechanism of translational regulation. The *Alu* domain structure is probably conserved in other cytoplasmic ribonucleoprotein particles and retroposition intermediates containing SRP9/14-bound RNAs transcribed from *Alu* repeats or related elements in genomic DNA.

The signal recognition particle (SRP) is an essential cytoplasmic ribonucleoprotein particle[1] found in all kingdoms of life that is involved in the targeting of signal-peptide-containing proteins to membranes[2,3]. In higher eukaryotes, SRP consists of two domains, the *Alu* domain and the S domain, which are linked by the SRP RNA (Fig. 1a). The *Alu* domain comprises the protein heterodimer of SRP9 and SRP14 (denoted SRP9/14) and the 5′ and 3′ terminal sequences of SRP RNA[4]. This domain is necessary[5–7] for retarding the ribosomal elongation ('elongation arrest') of nascent chains once their amino-terminal signal sequence is bound by the S domain of SRP. Elongation arrest is thought to enlarge the time window during which SRP can target nascent-chain/ribosome complexes to the membrane of the endoplasmic reticulum, thus ensuring the co-translational mode of protein translocation[2,8]. Apart from its role in translational control, the *Alu* RNA tertiary structure and probably SRP9/14 binding are crucial in the transcription[9], maturation[10], nucleolus localization and transport[11] of SRP RNA.

The *Alu* RNA tertiary structure is also relevant to understanding the spread of repetitive *Alu* elements, which now account for some 10% of primate genomic DNA (16.8% of human chromosome 22; ref. 12). *Alu* elements probably originated from the SRP RNA gene and propagate through retroposition[13], an ongoing process that has had a significant impact on the evolution of the human genome[14]. Transcriptionally active dimeric *Alu* elements give rise to stable small cytoplasmic *Alu* ribonucleoprotein particles (sc*Alu*RNPs) containing bound SRP9/14 (refs 15, 16). A cellular function of such sc*Alu* RNPs remains to be identified. A related example is the neuron-specific and regulated expression in primates of the monomeric *Alu*-like BC200 RNA, which binds SRP9/14 *in vitro*[17] and *in vivo*[18].

We previously determined the structure of a murine SRP9/14 fusion protein, which showed that the heterodimer has remarkable pseudo-two-fold symmetry as a consequence of the similar (β)αββββα fold of both proteins. The four α-helices are packed on the convex side of the continuous and highly concave central six-stranded, antiparallel β-sheet[19]. We also defined a minimal *Alu* RNA folding domain[20], SA86 (Fig. 1b), which specifically binds SRP9/14 with affinity comparable to that of complete SRP RNA.

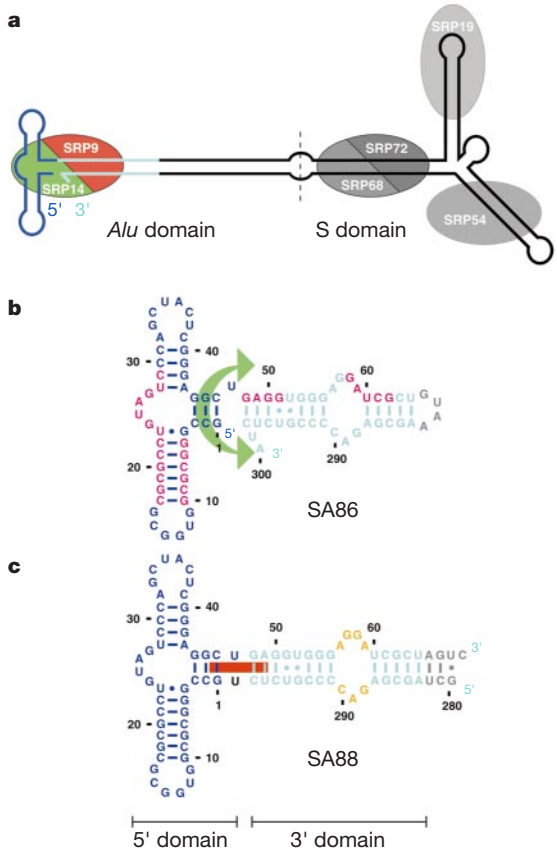

Figure 1 Mammalian SRP and *Alu* domain RNA constructs. **a**, Mammalian SRP with SRP RNA bound to six proteins. The *Alu* domain is coloured red (SRP9), green (SRP14), blue (*Alu* RNA 5′ domain) and cyan (*Alu* RNA 3′ domain). **b**, SA86, the minimal *Alu* RNA folding domain. S-domain RNA is replaced by a GUAA tetraloop (grey). The green arrow indicates flexibility between the RNA 5′ and 3′ domains. Hydroxyl-radical footprints of SRP9/14 on SRP RNA are highlighted in magenta[21]. **c**, SA88, a circular permutation of SA86. The original 5′ and 3′ ends are connected by a mono-uridine linker (black), constraining the RNA 5′ and 3′ domains to stack in an extended conformation (red bar). The GUAA tetraloop is replaced by a terminal stem (grey). The asymmetric internal loop is highlighted (yellow).

SA86 can be separated into a 5′ domain, comprising the first 47 nucleotides of SRP RNA, and a 39-nucleotide 3′ domain, comprising a helical stem with an asymmetric internal loop (Fig. 1). The *Alu* RNA 5′ domain, which contains the most highly conserved region of *Alu* RNA[21], is the primary SRP9/14 binding target and the 3′ stem contributes significantly to protein affinity only if covalently linked in a flexible way (unpublished data). This result prompted us to crystallize the ternary complex of human SRP9/14 with a 50-nucleotide 5′ domain RNA, denoted SA50 (Fig. 2a). We subsequently solved the structure of SRP9/14 in complex with a larger RNA, from which we have deduced a model for the native *Alu* RNP that is consistent with extensive biochemical data.

Structure of the 5′ domain complex

Microcrystals of the SA50 *Alu* RNP diffracted X-rays to 3.0 Å resolution at the microfocus beamline[22] of the European Synchrotron Radiation Facility (see Supplementary Information). The structure was solved by molecular replacement using the murine SRP9/14 fusion protein[19] as a model. The asymmetric unit contains a specific SRP9/14–SA50 complex with an additional SRP9/14 heterodimer nonspecifically bound to the RNA. Despite the modest resolution (3.2 Å), there is no ambiguity in the RNA fold (Fig. 2b).

The *Alu* RNA core contains a three-way τ-junction. The compactly folded RNA consists of two helical stacks connected by a central U-turn (Fig. 2b, d)—a tertiary structural motif that we call the 'τ-junction' owing to its shape. This fold is common to two other three-way junction RNAs, the hammerhead ribozyme[23,24] and the 23S ribosomal RNA-binding site of ribosomal protein L11 (L11 RNA)[25,26]. In all three structures (Fig. 2d), the two helical stacks (H1 and H2) are formed by three RNA strands, S1, S2 and S3. Strand S3 belongs entirely to helix H1 and forms base pairs with both strands S1 and S2, thereby defining helices H1.1 and H1.2. Strand S1 switches from helix H1.1 to H2, bending strongly at the junction, and strand S2 switches from helix H2 to H1.2, with a sharp loop in between containing the characteristic U-turn.

In SA50, strand S1 (nucleotides −2 to 10) is base paired throughout, bending and switching helices between C3 and G4 (Fig. 2). Strand S3 (nucleotides 40–48) is also completely base paired, and adopts helical geometry. The central U-turn is formed by U25 stacking on G24 and hydrogen bonding to the U28 phosphate. This reverses the chain direction of strand S2 (nucleotides 17–31), constraining the relative orientation of helices H2 and H1.2. The unpaired G24 makes a cross-strand purine stack on G4 of the opposing strand S1. G4·U23 is a universally conserved wobble base pair, which terminates helix H2. The conserved purine bases A26 and G27 loop out of the U-turn and are free for potential interactions.

The *Alu* RNA stem-loops make tertiary interactions. The relative orientation of subdomain 1 (helices H1.1 and H1.2 with loop L1.2) to subdomain 2 (helix H2 with loop L2) is reinforced by tertiary interactions between loops L1.2 and L2 (Fig. 2c). Loop L2 is a rigid hexa-loop that is internally stabilized by a sheared G11·G16 pair and a U-turn at U12, a structural motif similar to the 1095-loop of L11 RNA[25,26]. Bases G13, G14 and C15 face outward and form antiparallel stacked Watson–Crick pairs with C37, C34 and G33 of loop L1.2. Loop L1.2 has an unusual conformation, which may be induced only on interaction with loop L2 as the absence of internal stabilizing interactions suggests that it may be somewhat plastic. It is sharply bent with respect to helix H1.2, with the unpaired base A32 of helix H1.2 being almost perpendicular to G33 in loop L1.2. The

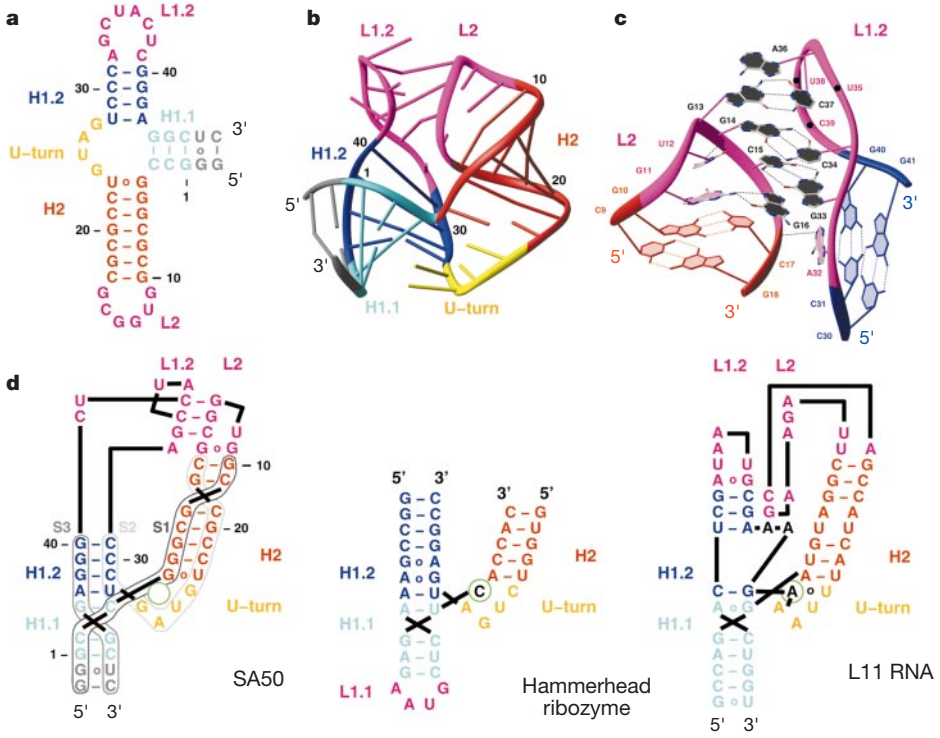

Figure 2 Structure of the *Alu* RNA 5′ domain (SA50). **a**, Secondary structure with the two artificial base pairs in grey. H, helix; L, loop; S, strand. **b**, Three-dimensional ribbon representation. **c**, Close-up of the interaction of the two peripheral loops. The view from the back emphasizes the central sandwich of stacked bases (black). U35, U38 and C39 are omitted for clarity. Dotted lines indicate putative hydrogen bonds. **d**, The τ-junction, a three-way junction comprising two helical stacks connected by a U-turn. Two-dimensional topology diagrams derived from the crystal structures of SA50, the hammerhead ribozyme[23,24] and L11 RNA (a fragment of 23S rRNA)[25,26] show the common architecture of the cores and the different interactions of the peripheral loops. An extra nucleotide in strand S1 of the ribozyme and in strand S2 of L11 RNA is absent in SA50 (green circle).

three interloop base pairs are sandwiched between two lids, G16 at the bottom and A36 at the top. U35, U38 and C39 in loop L1.2 are solvent exposed.

SRP9/14 specifically recognizes the core of the τ-junction. In the RNP, SRP9/14 is bound on the opposite side of the RNA with respect to the interacting loops (Fig. 3a). The human SRP9/14 heterodimer closely resembles the murine fusion protein, and RNA complexation does not change the backbone significantly. The N- and C-terminal residues of SRP9 (residues 1–5 and 76–86) and the C-terminal residues of SRP14 (amino acids 96–107) are disordered, and the internal loop of SRP14 (residues 36–53) is untraceable. The centres of mass of SRP9/14 and SA50 are well separated either side of the large interface (1,500 Å^2 buried surface area), and the positively charged concave β-sheet surface is the primary RNA-binding site. The protein interacts specifically with the core of the τ-junction, contacting both helical subdomains and the U-turn exclusively on RNA strand S2. Despite its pseudosymmetry, SRP9/14 is bound asymmetrically: the bulk of SA50 leans towards the SRP14 side of the heterodimer, with no contacts to the β2 and β3 strands of SRP9. The protein contacts RNA principally by basic residues to the RNA backbone rather than through specific hydrogen bonds to bases, and there are no base-stacking interactions with aromatic residues.

The backbone of RNA strand S2 follows a curved canyon formed mostly by basic amino acids, providing a complementary shape and charge distribution (Fig. 3b). Putative hydrogen bonds, often multiple, are made to every phosphate from C19 to C29, except G20 (Fig. 3c). Between nucleotides C21 and U25, the backbone of RNA strand S2 parallels the SRP9–SRP14 interface. Residues from the β1 strands of both SRP9 and SRP14 contact phosphates C22 to A26, highlighting the need for protein heterodimerization as a prerequisite for RNA binding[27].

Around the U-turn, the RNA backbone runs perpendicular to the β-sheet, such that nucleotides G24 to C29 contact residues from five different β-strands, including the short N-terminal β-strand of SRP14 (Fig. 3c). Phosphates G24 to G27 are embedded in a dense hydrogen-bond network with basic side chains typically bridging adjacent phosphates. Of note is the highly conserved SRP14 Arg 59, which bridges phosphates U25 and G27, thus recognizing the particular backbone conformation of the U-turn. G24 and U25, key residues within the τ-junction, are the most extensively contacted and the only base-specifically recognized nucleotides. The centrally positioned SRP14 Lys 66 bridges O6 of G24 and O4 of U25 and, in addition, contacts the U28 phosphate. This reinforces the fold of SA50 by doubling the crucial U-turn hydrogen bond to the U28 phosphate. Topologically, SRP14 Lys 66 occupies the position of the extra nucleotide present in the τ-junctions of the hammerhead ribozyme and of L11 RNA (Fig. 2d). The position of RNA helix H1.2 is also maintained by the SRP14 N-terminus (SRP14 Val 2) which contacts the C29 phosphate. The role of SRP14 Lys 64 (which forms a remarkable triskelion with Lys 65 and Lys 66) in pinning a distant region of helix H2 (C19 phosphate) to the central core of the complex is of particular note (Fig. 3c).

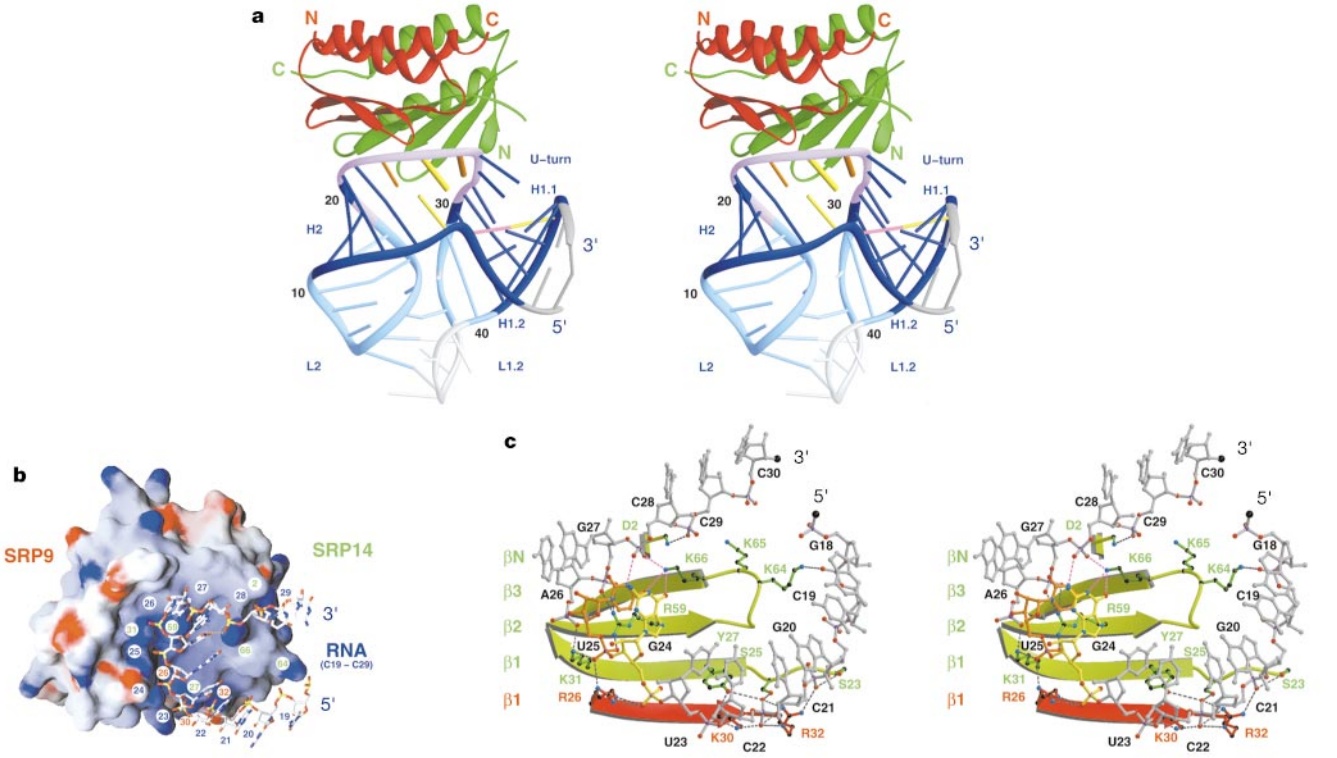

Figure 3 Protein recognition of the conserved fold of the *Alu* RNA τ-junction. **a**, Stereo view of the ternary complex of SRP9 (red), SRP14 (green) and the *Alu* RNA 5′ domain (SA50, blue with the peripheral loops pointing down and the two artificial base pairs in grey). Bases highly conserved in all *Alu*-like RNAs are coloured orange (G), magenta (C) or yellow (U); conserved base pairs and U-turn are dark blue; nucleotides occasionally absent even in higher eukaryotes (for example, plants, nematodes, insects) are silver; the RNA backbone contacted by protein is violet. Thick cylinders indicate RNA bases specifically recognized by protein. **b**, The GRASP[48] surface potential of SRP9/14 with RNA strand S2 in ball-and-stick representation shows the complementarity of shape and charge. The RNA backbone (phosphates numbered in blue) follows a canyon formed by basic residues from both SRP9 and SRP14 (red and green numbers, respectively). The U-turn hydrogen bond is dotted orange. See **c** for residue identity. **c**, Stereo diagram (view similar to **b**) of the RNA–protein interaction with part of the protein β-sheet in a ribbon representation and putative hydrogen bonds dotted. Important protein side chains and RNA strand S2 are in ball-and-stick representation. G24 (yellow) and U25 (orange) are base-specifically recognized. Interactions contributing to the positioning of phosphate U28 are highlighted in magenta.

Structure of SRP9/14 complex with 88-nucleotide RNA

To extend our structural understanding to the complete *Alu* domain, we crystallized purified *Alu* RNPs made from human SRP9/14 and SA86, the minimal *Alu* RNA, or one of two circularly permutated RNA variants. In SA91, which mimics the 5′–3′ domain flexibility of wild-type RNA, the natural 5′ and 3′ ends are connected by four uridines (not shown), whereas in SA88 (Fig. 1c) a rigidifying linker of one uridine is used. SA88 RNP crystals grown in the presence of europium nitrate diffracted beyond 4 Å resolution, but anisotropically. A good, unbiased, 4 Å resolution experimental map (Fig. 4) was obtained by the method of multiple anomalous dispersion (MAD) using the europium L(III) edge, and a complete model with an *R* factor of 38.9% was constructed, lacking only the last 2 of the 88 nucleotides (see Supplementay Information).

The structure of the SA88 RNP (Fig. 5a) shows the protein and RNA 5′ domain to be essentially the same as in the SA50 RNP, apart from a slight variation in RNA loop L1.2 probably caused by crystal contacts. As expected, the *Alu* RNA 3′ domain of SA88 (denoted helix H3) stacks onto helix H1.1 forming an extended helical structure. The proximal (helix H3.1) and distal (helix H3.3) parts of helix H3 have rather regular, A-helical parameters. The central part (helix H3.2), corresponding to the asymmetric internal loop in the secondary structure, has a unique geometry that causes considerable widening of the major groove and bends the helical axis between helices H3.1 and H3.3.

A model for the native *Alu* domain

The SA88 RNP forms a crystallographic dimer in which SRP9/14 clamps within its concave β-sheet surface not only the conserved region of the RNA 5′ domain, but also the 3′ domain of a two-fold related neighbouring molecule (Fig. 5b). The RNP dimer results entirely from the dual recognition by each SRP9/14 of both 5′ and 3′ domains from different RNA molecules, there being no contact between the two SRP9/14 molecules. By a number of analytic techniques, we have found that in concentrated solution the SA88 RNP is dimeric, whereas the SA86 and SA91 RNPs are monomeric, as are the three free RNAs (unpublished results). We therefore propose that intermolecular interactions leading to dimerization allow the rigidly constrained SA88 RNP to mimic intramolecular interactions in the SA86 and SA91 RNPs (and native SRP), where the 3′ domain is linked flexibly to the 5′ domain.

We generated a model for the monomeric *Alu* RNP by adjusting solely the position of U47 to cross-connect its ribose O3′ to the phosphate of G48 of the neighbouring molecule in the crystallographic dimer, 9 Å away. This places the two domains of a single RNA molecule side by side in the concave β-sheet surface of a single SRP9/14 heterodimer, the 3′ domain occupying the SRP9 surface left vacant in the SA50 RNP structure (Fig. 5d). Thus, we propose that SRP9/14 binding induces the RNA to fold back on itself like a jack-knife, such that the stem leading to the SRP S domain emerges on the SRP14 C-terminal side of SRP9/14 (Fig. 5b, c). Most of the 3′ domain contacts are made to the minor groove of distorted helix H3.2 (the asymmetric internal 'loop'). The β2–β3 loop of SRP9 is inserted into it, and the minor groove side of helix H2 of the RNA 5′ domain packs against it. The implied back-folding of the RNA 3′ domain by 180° is probably the final, but reversible, step in the assembly pathway of the *Alu* domain (unpublished data).

The *Alu* domain model and biochemical data

Extensive biochemical data are explained well by the *Alu* domain model. Most notably, it can rationalize the complex footprint of SRP9/14 on *Alu* RNA in hydroxyl-radical cleavage experiments[21]. Whereas the SA50 RNP structure explains the protection seen on strand S2 (owing to direct binding of the protein to the RNA backbone), protection within the 5′ and 3′ domains of SRP *Alu* RNA can be explained only by the model of the compactly folded monomeric *Alu* RNP (Figs 1b, 5c and d). Protected regions I (G5–G10, defined in ref. 21), and III (G48–G51) appear to arise exclusively from protein-induced RNA tertiary contacts, whereas regions II (C17–C29) and IV (G58–G62) are caused by a combination of both induced RNA tertiary contacts and direct protein contacts. Furthermore, only in the compact particle are SRP9 Cys 39 and SRP9 Cys 48 buried in the RNA–protein interface (Fig. 5d), which explains their protection from alkylation in the presence of

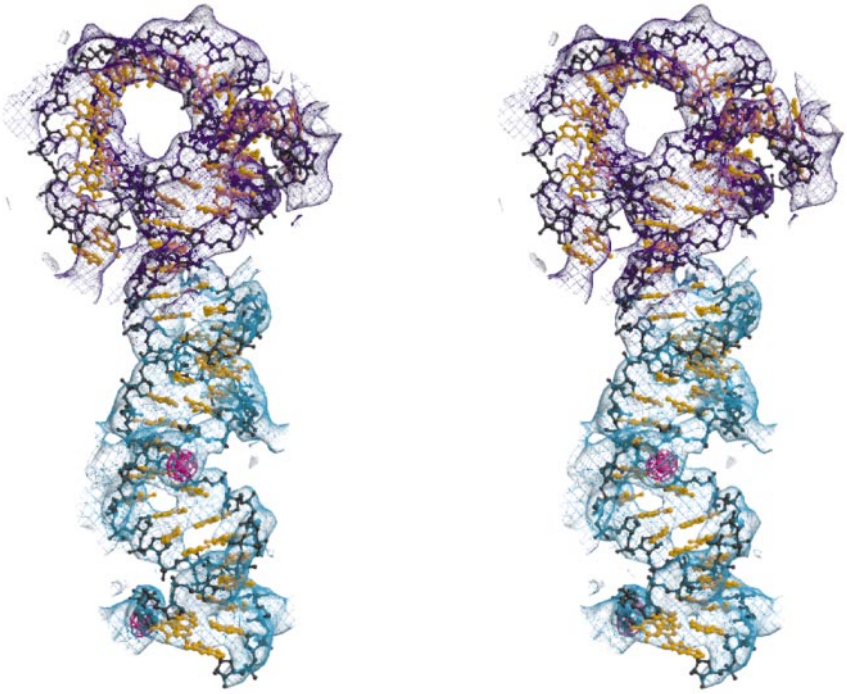

Figure 4 Stereo image of experimental electron density for SA88 RNA. The map around the model of SA88 *Alu* RNA (phosphate-ribose backbone in black, bases in orange) is contoured in blue (5′ domain) and cyan (3′ domain) at 1.0σ. The central (major) and peripheral (minor) europium sites are contoured in magenta at 8.0 and 3.0σ, respectively.

SRP RNA[28]. In addition, the model explains the need of protein hetero-dimerization for RNA binding[27], the phylogenetically supported RNA base-pairing scheme including tertiary interactions[29], the importance of G24 for protein binding[30] and the stacking of helix H1.1 with H1.2, as predicted from the RNase V1 cleavage pattern[20]. The interaction of the SRP14 Val 2 N-terminus with phosphate C29 is consistent with the observed absence of SRP14 Met 1 *in vivo*[31]. The model does not yet address the role of the disordered loop between the β1 and β2 strands of SRP14, which may order on interaction with the compact, monomeric *Alu* RNA in the region of the bend (Fig. 5a). Finally, the dimensions of the model ($52 \times 59 \times 70$ Å^3) correspond well to the small end domain of SRP observed in electron microscopy[32]. We therefore conclude that the compactly folded *Alu* RNP model is a close approximation to the true conformation of the SRP *Alu* domain.

Biological implications

We have shown that SRP9/14 binds primarily to the universally conserved core of the *Alu* RNA 5′ domain, which forms a U-turn in the context of a τ-junction. This RNA motif is highly conserved in SRP RNAs from higher eukaryotes to yeast and from archaea to some Gram-positive eubacteria[21]. Also conserved is the potential of the RNA 5′ domain to back-fold onto the RNA 3′ domain. The more peripheral parts of the RNA not involved in protein binding are much less conserved. Sequence variations or truncations in loop L1.2, for example, indicate that the RNA loop–loop tertiary interactions are likely to be idiosyncratic. With respect to SRP9/14, no eubacterial or archaeal sequence homologues have been found, but the histone-like protein HBsu from *Bacillus subtilis* might be a functional analogue in eubacteria[33]. In yeast, where elongation arrest is preserved[7] in the context of an apparently severely truncated *Alu* RNA, SRP9/14 is replaced by a homodimer of SRP14 (ref. 34). These variations on the basic architecture of the *Alu* domain, together with the structural information on the mammalian system, should help pinpoint features that are maintained owing to functional importance other than *Alu* RNP assembly. The conserved C-terminal tail of SRP14 might represent such an essential feature, as its removal specifically abolishes elongation arrest activity in both the mammalian[6] and yeast[7] systems.

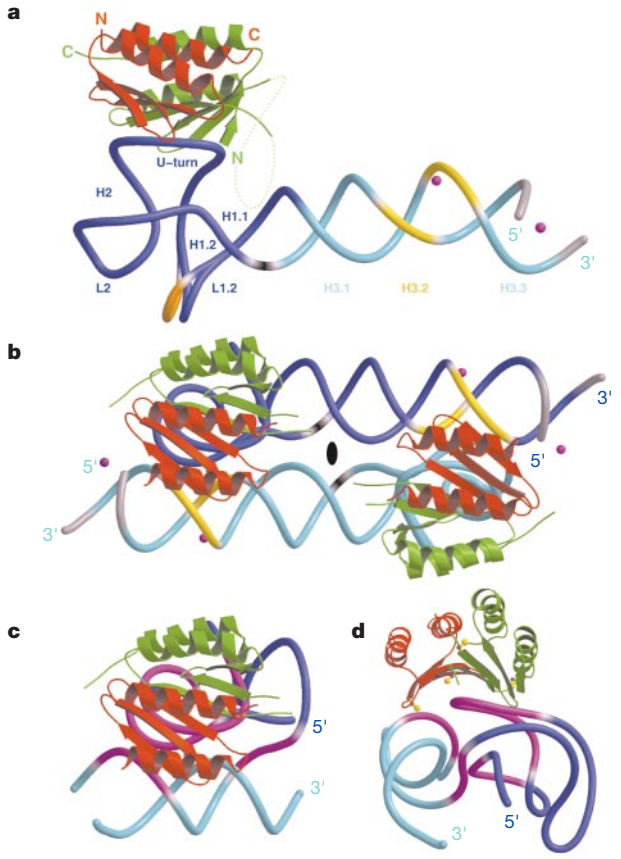

Figure 5 Structure of the *Alu* domain of the mammalian SRP (see Fig. 1 for colours). **a**, The SA88 *Alu* RNP with the alternative conformation of loop L1.2 in the SA50 *Alu* RNP in orange and the two europium sites in magenta. The disordered SRP14 β1–β2 loop is indicated by a dotted line; helix H3.2 is coloured yellow; H, helix; L, loop. **b**, The SA88 *Alu* RNP dimer viewed down its crystallographic two-fold axis. **c**, Model for an *Alu* RNP (SA86) in its fully folded, physiological conformation with hydroxyl-radical footprints of SRP9/14 on SRP RNA[21] in magenta. **d**, As **c**, viewed down the SRP9/14 β-sheet and illustrating the coverage of the previously exposed SRP9 surface with cysteines in yellow ball-and-stick representation.

Our structural results lead to three conclusions. First, the existence of a stable 5′ domain RNP shows the possibility for the *Alu* RNA 5′ domain to form a specific complex with nuclear SRP9/14 immediately after transcription by RNA polymerase III. This is consistent with mutational data[9], which show that the integrity of the tertiary structure and the SRP9/14 binding determinants of the first 50 nucleotides of SRP RNA are necessary for its efficient transcription. SRP9/14 and SRP RNA might therefore associate into a functional unit at the very beginning of transcription, and probably remain tightly associated throughout subsequent RNA processing, assembly and transport steps as well as in the functional cytoplasmic SRP.

Second, regarding the mechanism of elongation arrest, the flexible, rod-like appearance of SRP with its two functional domains at opposing ends has prompted speculations that SRP might reach from near the ribosome polypeptide exit site to the elongation factor/transfer-RNA-binding site[32,35]. Elongation arrest might then arise from steric interference or active competition of the *Alu* domain with elongation factor (EF)-1α, EF2 or tRNA. Although the *Alu* domain model shows no structural mimicry with tRNA or elongation factors, its size and shape do not preclude it from entering the ribosomal elongation-factor-binding site. Our hypothesis that the fairly weak interaction of the RNA 3′ domain with the stable RNP 5′ domain is dynamic suggests that a reversible switch in the folding of the *Alu* domain may also be of functional importance, for example, permitting alternating onset or release of elongation arrest. Further investigation of the mechanism of elongation arrest will require structural work on ribosome–SRP complexes and specific crosslinks to be found between the *Alu* domain and ribosomal components.

Third, RNA sequence analysis in the light of the crystal structure illustrates that non-SRP *Alu* RNAs, such as monomeric BC200 RNA and left and right arm monomers of actively retroposing human *Alu* element RNA, have indeed conserved both the three-dimensional architecture found in SRP *Alu* RNA and the determinants for binding SRP9/14. The *Alu* RNP 5′–3′ domain flexibility proposed for SRP might therefore also be important for the processing and function of sc*Alu* RNPs. Furthermore, the observation that in a crystalline environment even SA86 *Alu* RNPs can form the domain-switch non-covalent dimers observed for SA88 RNPs (our own unpublished data) may be of relevance to *Alu* retroposition. Whereas there is no rationale to support the occurrence of dimerization of SRP, a dimeric *Alu* RNP complex might be important in the origin or propagation of tandemly arranged *Alu* retroposons, as retropositional success was clearly correlated with the emergence of dimeric *Alu* elements during primate evolution[36].

Conclusions

We have described the structure of two RNPs derived from the mammalian SRP permitting us to propose a model for the *Alu* domain occurring in SRP or other scRNPs. The structures provide further insight into the principles determining the architecture of

compactly folded RNAs and the hierarchical assembly of RNPs. They also provide a solid structural basis to assist further functional studies aimed at understanding the role of the *Alu* domain of SRP and possibly other *Alu*-like particles, such as the BC200 RNP, in translational regulation. The finding that SRP9/14 binds specifically to the 5′ domain of SRP RNA gives a possible rationalization for the required integrity of this domain in the early steps of SRP RNA transcription. Finally, the structure will contribute to an understanding of the remarkable success of *Alu* retroposons in colonizing primate genomes. □

Methods

Preparation of RNA and protein samples

RNA molecules with 3′-terminal hammerhead ribozymes were synthesized by *in vitro* transcription with T7 RNA polymerase from *Hin*dIII linearized plasmid templates and purified as described[20]. Plasmids pSA86(2)H, pSA88(2)H and pSA50H, coding for SA86, SA88 and SA50, were obtained by introducing synthetic oligonucleotides into *Ava*I/*Xba*I digested plasmids pSA86H and pSA50 or *Xba*I/*Hin*dIII digested plasmids pSA88H and pSA85 (ref. 20). Human SRP9 and C-terminally truncated SRP14 were overexpressed in *Escherichia coli* from plasmids pEh9 and pEh14ΔR (ref. 17), separately purified by Heparin and MonoS ion exchange chromatography and finally combined and repurified on a MonoS column as an essentially RNase-free heterodimer.

Crystallization, structure determination and refinement

For the SA50 complex, crystals ($30\times\sim5\times<5\ \mu m^3$) were grown at 12 °C in 3 d in a 3-μl hanging drop (130 μM SA50, 180 μM SRP9/14, 2.5 mM HEPES pH 7.5, 25 mM $NaCH_3COO$ pH 5.0, 10 mM $MgCl_2$, 75 mM NaCl, 210 mM $(NH_4)_2SO_4$, 5 mM dithiothreitol, 11% PEG 2000) equilibrated over a 500-μl reservoir (50 mM $NaCH_3COO$ pH 5.0, 10 mM $MgCl_2$, 140 mM NaCl, 390 mM $(NH_4)_2SO_4$, 21% PEG 2000). They were flash frozen in liquid nitrogen from 40 mM $NaCH_3COO$ pH 5.0, 15 mM $MgCl_2$, 120 mM NaCl, 350 mM $(NH_4)_2SO_4$, 22% PEG 2000 supplemented with 20% glycerol. Crystals are of space group $C222_1$ with unit-cell dimensions $a = 57.4$ Å, $b = 186.6$ Å and $c = 189.8$ Å, and 71% solvent. Diffraction data were collected on ESRF beamline ID13 from a single crystal centred in a 10-μm beam. This dataset was integrated with MOSFLM 6.0 (ref. 37) and completed with low-resolution data (50 Å to 8.5 Å) from a second crystal measured at beamline ID14eh2. Molecular replacement was done with AMoRe[38] using data between 15 and 5.5 Å resolution from the second crystal and searching for two copies of a modified polyalanine model of murine SRP9/14. The top solution was rigid-body refined in CNS[39]. Two rounds of density modification by program ARP/wARP[40] yielded the electron density map into which the model was built manually with the program O[41]. Refinement (including atomic *B* factors) was continued with CNS, correcting for anisotropy of the data, applying a bulk solvent correction and restraining identified Watson–Crick and G·U wobble base pairs as well as the ribose sugar puckers (all C3′ endo except nucleotides 128, 132, 135–136 and 138–139, which are C2′ endo).

For the SA88 complex, crystals ($200\times200\times150\ \mu m^3$) were grown at 12 °C in a week in 3-μl hanging drops (70 μM SA88 RNP, 28 mM HEPES pH 7.0, 10 mM $MgCl_2$, 75 mM NaCl, 0.40 mM $Eu(NO_3)_3$, 5 mM dithiothreitol, 210 mM $(NH_4)_2SO_4$, 12% PEG 400) equilibrated over 500 μl reservoirs (50 mM HEPES pH 7.0, 10 mM $MgCl_2$, 150 mM NaCl, 0.80 mM $Eu(NO_3)_3$, 390 mM $(NH_4)_2SO_4$, 23% PEG 400). They were flash frozen from mother liquor in liquid nitrogen. Crystals are of space group $P4_22_12$ with unit-cell dimensions of $a = b = 143.3$ Å and $c = 60.4$ Å, and 67% solvent. Electron density calculated from a molecular replacement solution with the SA50 RNP showed a specific europium-binding site. A three-wavelength MAD experiment at the europium L(III) edge was therefore undertaken on a single crystal at ESRF beamline BM14 and data reduced using the HKL package[42]. A dataset to 4 Å resolution was collected on the same crystal at beamline ID14eh2 and processed with MOSFLM 6.0 (ref. 37). SHARP[43] was used with the MAD data to identify a second europium-binding site, to refine the heavy metal positions and to calculate phases to 4 Å, followed by solvent flipping with SOLOMON[44]. The helical *Alu* RNA 3′ domain extending the structure of the SA50 RNP was built manually from RNA fragments of known structure. Refinement was done with CNS, applying a correction for data anisotropy ($B_{11} = -6.0$ Å^2, $B_{22} = -6.0$ Å^2, $B_{33} = 12.0$ Å^2), a bulk-solvent correction, rigid-body refinement and a final, highly restrained energy minimization step to improve the RNA geometry.

Figures were produced with RIBBONS[45] and BobScript[46]/Raster 3D[47].

Received 11 May; accepted 4 October 2000.

1. Walter, P. & Blobel, G. Signal recognition particle contains a 7S RNA essential for protein translocation across the endoplasmic reticulum. *Nature* **299,** 691–698 (1982).
2. Walter, P. & Johnson, A. E. Signal sequence recognition and protein targeting to the endoplasmic reticulum membrane. *Annu. Rev. Cell Biol.* **10,** 87–119 (1994).
3. Lütcke, H. Signal recognition particle (SRP), a ubiquitous initiator of protein translocation. *Eur. J. Biochem.* **228,** 531–550 (1995).
4. Gundelfinger, E. D., Krause, E., Melli, M. & Dobberstein, B. The organization of the 7SL RNA in the signal recognition particle. *Nucleic Acids Res.* **11,** 7363–7374 (1983).
5. Siegel, V. & Walter, P. Removal of the *Alu* structural domain from signal recognition particle leaves its protein translocation activity intact. *Nature* **320,** 81–84 (1986).
6. Thomas, Y., Bui, N. & Strub, K. A truncation in the 14 kDa protein of the signal recognition particle leads to tertiary structure changes in the RNA and abolishes the elongation arrest activity of the particle. *Nucleic Acids Res.* **25,** 1920–1929 (1997).
7. Mason, N., Ciufo, L. F. & Brown, J. D. Elongation arrest is a physiologically important function of signal recognition particle. *EMBO J.* **19,** 4164–4174 (2000).
8. Rapoport, T. A., Heinrich, R., Walter, P. & Schulmeister, T. Mathematical modeling of the effects of the signal recognition particle on translation and translocation of proteins across the endoplasmic reticulum membrane. *J. Mol. Biol.* **195,** 621–636 (1987).
9. Emde, G., Frontzek, A. & Benecke, B. J. Secondary structure of the nascent 7S L RNA mediates efficient transcription by RNA polymerase III. *RNA* **3,** 538–549 (1997).
10. Chen, Y., Sinha, K., Perumal, K., Gu, J. & Reddy, R. Accurate 3′ end processing and adenylation of human signal recognition particle RNA and Alu RNA *in vitro*. *J. Biol. Chem.* **273,** 35023–35031 (1998).
11. Jacobson, M. R. & Pederson, T. Localization of signal recognition particle RNA in the nucleolus of mammalian cells. *Proc. Natl Acad. Sci. USA* **95,** 7981–7986 (1998).
12. Dunham, I. *et al.* The DNA sequence of human chromosome 22. *Nature* **402,** 489–495 (1999).
13. Weiner, A. M., Deininger, P. L. & Efstratiadis, A. Nonviral retroposons: genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information. *Annu. Rev. Biochem.* **55,** 631–661 (1986).
14. Kazazian, H. H. Jr Mobile elements and disease. *Curr. Opin. Genet. Dev.* **8,** 343–350 (1998).
15. Bovia, F., Fornallaz, M., Leffers, H. & Strub, K. The SRP9/14 subunit of the signal recognition particle (SRP) is present in more than 20-fold excess over SRP in primate cells and exists primarily free but also in complex with small cytoplasmic Alu RNAs. *Mol. Biol. Cell* **6,** 471–484 (1995).
16. Chang, D. Y., Hsu, K. & Maraia, R. J. Monomeric scAlu and nascent dimeric Alu RNAs induced by adenovirus are assembled into SRP9/14-containing RNPs in HeLa cells. *Nucleic Acids Res.* **24,** 4165–4170 (1996).
17. Bovia, F., Wolff, N., Ryser, S. & Strub, K. The SRP9/14 subunit of the human signal recognition particle binds to a variety of Alu-like RNAs and with higher affinity than its mouse homolog. *Nucleic Acids Res.* **25,** 318–326 (1997).
18. Kremerskothen, J. *et al.* Heterodimer SRP9/14 is an integral part of the neural BC200 RNP in primate brain. *Neurosci. Lett.* **245,** 123–126 (1998).
19. Birse, D. E., Kapp, U., Strub, K., Cusack, S. & Åberg, A. The crystal structure of the signal recognition particle *Alu* RNA binding heterodimer, SRP9/14. *EMBO J.* **16,** 3757–3766 (1997).
20. Weichenrieder, O., Kapp, U., Cusack, S. & Strub, K. Identification of a minimal *Alu* RNA folding domain that specifically binds SRP9/14. *RNA* **3,** 1262–1274 (1997).
21. Strub, K., Moss, J. & Walter, P. 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 (1991).
22. Cusack, S. *et al.* Small is beautiful: protein micro-crystallography. *Nature Struct. Biol.* **5,** 634–637 (1998).
23. Pley, H. W., Flaherty, K. M. & McKay, D. B. Three-dimensional structure of a hammerhead ribozyme. *Nature* **372,** 68–74 (1994).
24. Scott, W. G., Finch, J. T. & Klug, A. The crystal structure of an all-RNA hammerhead ribozyme: a proposed mechanism for RNA catalytic cleavage. *Cell* **81,** 991–1002 (1995).
25. Wimberly, B. T., Guymon, R., McCutcheon, J. P., White, S. W. & Ramakrishnan, V. A detailed view of a ribosomal active site: the structure of the L11–RNA complex. *Cell* **97,** 491–502 (1999).
26. Conn, G. L., Draper, D. E., Lattman, E. E. & Gittis, A. G. Crystal structure of a conserved ribosomal protein–RNA complex. *Science* **284,** 1171–1174 (1999).
27. Strub, K. & Walter, P. 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 (1990).
28. Siegel, V. & Walter, P. 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 (1988).
29. Zwieb, C., Müller, F. & Larsen, N. Comparative analysis of tertiary structure elements in signal recognition particle RNA. *Fold. Des.* **1,** 315–324 (1996).
30. Chang, D. Y., Newitt, J. A., Hsu, K., Bernstein, H. D. & Maraia, R. J. A highly conserved nucleotide in the Alu domain of SRP RNA mediates translation arrest through high affinity binding to SRP9/14. *Nucleic Acids Res.* **25,** 1117–1122 (1997).
31. Strub, K. & Walter, P. Isolation of a cDNA clone of the 14-kDa subunit of the signal recognition particle by cross-hybridization of differently primed polymerase chain reactions. *Proc. Natl Acad. Sci. USA* **86,** 9747–9751 (1989).
32. Andrews, D. W., Walter, P. & Ottensmeyer, F. P. Evidence for an extended 7SL RNA structure in the signal recognition particle. *EMBO J.* **6,** 3471–3477 (1987).
33. Nakamura, K., Yahagi, S., Yamazaki, T. & Yamane, K. *Bacillus subtilis* histone-like protein, HBsu, is an integral component of a SRP-like particle that can bind the *Alu* domain of small cytoplasmic RNA. *J. Biol. Chem.* **274,** 13569–13576 (1999).
34. Strub, K., Fornallaz, M. & Bui, N. The Alu domain homolog of the yeast signal recognition particle consists of an Srp14p homodimer and a yeast-specific RNA structure. *RNA* **5,** 1333–1347 (1999).
35. Siegel, V. & Walter, P. Binding sites of the 19-kDa and 68/72-kDa signal recognition particle (SRP) proteins on SRP RNA as determined in protein–RNA "footprinting". *Proc. Natl Acad. Sci. USA* **85,** 1801–1805 (1988).
36. Mighell, A. J., Markham, A. F. & Robinson, P. A. *Alu* sequences. *FEBS Lett.* **417,** 1–5 (1997).
37. Leslie, A. G. W. *Joint CCP4 & ESF-EACBM Newsletter on Protein Crystallography* No. 26 (Daresbury Laboratory, Warrington, UK, 1992).
38. Navaza, J. AMoRe: an automated package for molecular replacement. *Acta Crystallogr. A* **50,** 157–163 (1994).
39. Brünger, A. T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54,** 905–921 (1998).
40. Perrakis, A., Morris, R. & Lamzin, V. S. Automated protein model building combined with iterative structure refinement. *Nature Struct. Biol.* **6,** 458–463 (1999).
41. Jones, T. A., Zou, J. Y., Cowan, S. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47,** 110–119 (1991).
42. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276,** 307–326 (1997).

43. de la Fortelle, E. & Bricogne, G. Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods Enzymol.* **276,** 472 –494 (1997).
44. Abrahams, J. P. & Leslie, A. G. W. Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta Crystallogr. D* **52,** 30–42 (1996).
45. Carson, M. RIBBONS 2.0. *J. Appl. Crystallogr.* **24,** 958–961 (1991).
46. Esnouf, R. M. An extensively modified version of MolScript that includes greatly enhanced coloring capabilities. *J. Mol. Graph.* **15,** 133–138 (1997).
47. Merritt, E. A. & Bacon, D. J. Raster3D: photorealistic molecular graphics. *Methods Enzymol.* **277,** 505–524 (1997).
48. Nicholls, A., Sharp, K. A. & Honig, B. Protein folding and association: insight from the interfacial and thermodynamic properties of hydrocarbon. *Proteins* **11,** 281–296 (1991).

Supplementary information is available on *Nature*'s World-Wide Web site (http://www.nature.com) or as paper copy from the London editorial office of *Nature*.

Acknowledgements

O.W. received a pre-doctoral fellowship from the Boehringer-Ingelheim-Fonds. K.S. is supported by a START subsidy and a grant from the Swiss National Science Foundation and the Canton de Genève. K.S. and S.C. are both principal investigators belonging to SRPNET, a research network of the European Union Training and Mobility in Research program, which also supported K.W. We thank members of the EMBL/ESRF Joint Structural Biology Group (JSBG) and the staff of BM30 and ID13 for support on ESRF beamlines; in particular, A. Thompson for the europium MAD experiment and A. Perrakis for the ID13 measurements and help with wARP. C. Petosa commented critically on the manuscript.

Correspondence and requests for materials should be addressed to S.C. (e-mail: cusack@embl-grenoble.fr). Structures have been deposited in the PDB under accession numbers 1e8o (SA50 RNP) and 1e8s (SA88 RNP).