Identification of a minimal Alu RNA folding domain that specifically binds SRP9/14

WEICHENRIEDER, Oliver Ulrich, et al.

Abstract
We have identified functionally and analyzed a minimal Alu RNA folding domain that is recognized by SRPphi14-9. Recombinant SRPphi14-9 is a fusion protein containing on a single polypeptide chain the sequences of both the SRP14 and SRP9 proteins that are part of the Alu domain of the signal recognition particle (SRP). SRPphi14-9 has been shown to bind to the 7SL RNA of SRP and it confers elongation arrest activity to reconstituted SRP in vitro. Alu RNA variants with homogeneous 3’ ends were produced in vitro using ribozyme technology and tested for specific SRPphi14-9 binding in a quantitative equilibrium competition assay. This enabled identification of an Alu RNA of 86 nt (SA86) that competes efficiently with 7SL RNA for SRPphi14-9 binding, whereas smaller RNAs did not. The secondary structure of SA86 includes two stem-loops that are connected by a highly conserved bulge and, in addition, a part of the central adaptor stem that contains the sequence at the very 3’ end of 7SL RNA. Circularly permuted variants of SA86 competed only if the 5’ and 3’ ends were joined with an extended linker of four nucleotides. SA86 can [...]
Identification of a minimal Alu RNA folding domain that specifically binds SRP9/14

OLIVER WEICHENRIEDER,1,2 ULRIKE KAPP,1 STEPHEN CUSACK,1 and KATHARINA STRUB2
1European Molecular Biology Laboratory (EMBL), Grenoble Outstation, c/o I.L.L., B.P.156, F-38042 Grenoble Cedex 9, France
2Département de Biologie Cellulaire, Université de Genève, Sciences III, CH-1211 Genève 4, Switzerland

ABSTRACT

We have identified functionally and analyzed a minimal Alu RNA folding domain that is recognized by SRP9/14-9. Recombinant SRP9/14-9 is a fusion protein containing on a single polypeptide chain the sequences of both the SRP14 and SRP9 proteins that are part of the Alu domain of the signal recognition particle (SRP). SRP9/14-9 has been shown to bind to the 7SL RNA of SRP and it confers elongation arrest activity to reconstituted SRP in vitro. Alu RNA variants with homogeneous 3' ends were produced in vitro using ribozyme technology and tested for specific SRP9/14-9 binding in a quantitative equilibrium competition assay. This enabled identification of an Alu RNA of 86 nt (SA86) that competes efficiently with 7SL RNA for SRP9/14-9 binding, whereas smaller RNAs did not. The secondary structure of SA86 includes two stem-loops that are connected by a highly conserved bulge and, in addition, a part of the central adaptor stem that contains the sequence at the very 3' end of 7SL RNA. Circularly permuted variants of SA86 competed only if the 5' and 3' ends were joined with an extended linker of four nucleotides. SA86 can thus be defined as an autonomous RNA folding unit that does not require its 5' and 3' ends for folding or for specific recognition by SRP9/14-9. These results suggest that Alu RNA identity is determined by a characteristic tertiary structure, which might consist of two flexibly linked domains.

Keywords: circular permutation; equilibrium competition; fusion protein; induced fit; ribozyme technology; RNA-protein; tertiary structure

INTRODUCTION

Alu RNA sequences are found as part of the 7SL RNA (Weiner, 1980; Ullu et al., 1982) of the signal recognition particle (SRP) (Walter & Blobel, 1982) and also as independent Pol III transcripts that are derived from short interspersed elements (SINEs) in the genomic DNA of primates or rodents (Adeniyi-Jones & Zasloff, 1985; Matera et al., 1990; Marais et al., 1993). In primates, these elements are called Alu repeats and code for units of two tandemly arranged Alu sequences. In rodents, the elements are called B1 repeats and are monomeric. SINEs are believed to propagate by genomic reintegration of their reverse-transcribed RNAs, a process named retroposition and reviewed in Weiner et al. (1986). Evolutionary studies suggest an ancient 7SL RNA to be the original parent RNA of both Alu and B1 repeats (Ullu & Tschudi, 1984).

7SL RNA is the central component of SRP, a cytoplasmic ribonucleoprotein (RNP) that plays an essential role in sorting proteins to the membrane of the endoplasmic reticulum (ER). SRP recognizes and binds to the signal sequence of the nascent chain as it emerges from the ribosome. This interaction triggers a pause in the synthesis of the polypeptide chain and the ribosome-nascent chain-SRP complex is then targeted to the membrane of the ER via the interaction of SRP with its receptor. Protein synthesis is resumed and the nascent chain engages in the actual translocation process. Canine SRP was discovered first and is the most studied. In addition to its 300-nt long 7SL RNA, it consists of two heterodimeric protein subunits (SRP68/72 and SRP9/14) and two monomeric polypeptides (SRP54 and SRP19). The 7SL RNA links two enzymatically separable domains of SRP. The Alu domain consists of the heterodimer SRP9/14 bound to the Alu sequences at the 5' and extreme 3' end of the 7SL RNA. The S domain contains the remaining proteins bound to the central RNA sequence, which is unique to 7SL RNA. Functionally,
the Alu domain is essential to mediate the delay in the translation of secretory proteins (Siegel & Walter, 1986, 1988; Thomas et al., 1997), whereas the S domain is sufficient for signal sequence recognition and targeting to the ER membrane (see Walter & Johnson, 1994; Lütcke, 1995; Bovia & Strub, 1996 for review).

Alu repeats are present at nearly one million copies in primate DNA and thus represent about 5% of the genome by mass. Dimerization must have happened early and seems to have been favorable for the proliferation process (Quentin, 1992a, 1992b). In human DNA, Alu repeats can be classified into families of different evolutionary age (see Shen et al., 1991; Schmid & Marraia, 1992 for review; Batzer et al., 1996 for nomenclature). Two forms of Alu RNA sequences can be found in cells. The first form are full-length, dimeric Alu RNAs (flAlus). They are heterogeneous, contain A-rich tracts and variable 3' ends (Matera et al., 1990). RNA 3' processing of the rather labile flAlus gives rise to the second form, the stable, monomorphic, and more homogeneous small cytoplasmic Alu RNAs (scAlus). These correspond to the left monomer of flAlus and do not act as retroposons. 3' End processing would thus enable the cell to reduce the level of flAlus and consequently the probability for successful retroposition (Marraia et al., 1993; Sarrova et al., 1997).

For 7SL RNA, the secondary structure is well established and a three-dimensional model has been proposed, including a pseudoknot in the very 5' domain (Zwieb et al., 1996). The flAlus and scAlus seem to have grossly conserved the same base pairing pattern as the Alu domain of 7SL RNA and the conservation of the secondary structure of divergent Alu primary sequences has been claimed to be correlated with both retroposition of genomic Alu repeats and generation of scAlus (Chang & Marraia, 1993; Labuda & Zietkiewicz, 1994). The secondary structure of the very 5' domain of 7SL RNA also has been shown recently to mediate efficient transcription by RNA polymerase III and some still unidentified proteins specifically bind to an RNA fragment comprising the first 55 nt (Emde et al., 1997).

The proteins SRP9 and SRP14, named according to the apparent molecular weight of the canine homologues, form a stable heterodimer in the absence of 7SL RNA and only bind RNA as a dimer (Strub & Walter, 1990). The apparent dissociation constant \(K_d\) for canine SRP9/14 and synthetic canine SRP RNA—in this study called SRP301—is estimated to be in the subnanomolar range (Janiak et al., 1992). Four regions (I-IV) have been identified on SRP301 that are protected from hydroxyl radical cleavage in the presence of SRP9/14 (Strub et al., 1991) and mutations in SRP9/14 that affect RNA binding exclusively have been identified (Bui et al., 1997). Human and murine SRP9/14 also have been shown to bind strongly to synthetic Alu RNAs and scAlus in vitro (Chang & Marraia, 1993; Bovia et al., 1995, 1997). In vivo evidence for an interaction of SRP9/14 with Alu RNAs comes from an 8.5S ribonucleoprotein particle (RNP) that exists in HeLa cells and contains SRP9/14 and scAlus RNA (Bovia et al., 1995) as well as from immunoprecipitation experiments that identify SRP9/14 to be associated with scAlus and flAlus in HeLa cells (Chang et al., 1996). For the functional and structural analysis of the SRP9/14 heterodimer, a fusion protein (SRPPhi4-9) has been constructed. This single polypeptide is expressed in Escherichia coli and can be purified to homogeneity. Assembled into SRP, it is active in in vitro assays on translation arrest and translocation (Bovia et al., 1994), and the crystal structure of SRPPhi4-9 has been solved recently (Birse et al., 1997).

Besides the evolutionary aspect of the propagation of Alu repeats, Alu RNA sequences might have adopted various physiological roles. In the context of SRP, the elongation arrest function of the Alu domain is well established. But the existence of stable scAlus RNPs and the disproportional rise of flAlus over scAlus under cellular stress (Chang et al., 1996) indicate a larger spectrum of functions and Alu sequences that are part of the 3' untranslated regions of cytoplasmic messenger RNAs might serve unknown tasks as well. In the present study, we wanted to confine the minimal RNA target that is recognized specifically by SRP9/14. We wished to determine to what extent the protein–RNA recognition is base-sequence specific as opposed to dependency on a characteristic RNA tertiary structure. To identify a minimal Alu RNA folding domain, we used the recognition of Alu RNAs by SRPPhi4-9 as a criterion for correct folding and analyzed in vitro transcribed and highly purified Alu RNA variants that were derived from SRP301 in an equilibrium competition assay based on nitrocellulose filter binding. In order to compare the secondary structures of interesting variants, we digested them with RNAase V1. We identify in a first step an SRP Alu RNA of 86 nt length (SA86) and show that its interaction with SRPPhi4-9 is as specific as that of SRP301, which is functional in reconstituted SRP. In a second step, we assay deletion variants, asking whether SA86 can be reduced significantly further. In a third and final step, we use circular permutations of SA86 to test whether the 5' and 3' ends are essential for SRPPhi4-9 binding and whether there exists an autonomous Alu RNA folding unit.

RESULTS

Experimental strategy

Competition experiments are essential to prove the specificity of an interaction. We therefore decided to do quantitative equilibrium competition experiments based on nitrocellulose filter binding (Lin & Riggs, 1972; Carey et al., 1983) to test the specificity of SRPPhi4-9 for its RNA target and to compare the affinities of different Alu RNAs for SRPPhi4-9. For the quan-
titative analysis, we consider two binding reactions to proceed simultaneously and assume a 1:1 stoichiometry between SRP\(\Phi\)4-9 and its RNA target. The two mass-action equations are combined to express the fraction saturation (\(\nu\)) of the protein with specific RNA as a function of the ratio (\(\rho\)) of competitor RNA to specific RNA, and a correction is made for the activity of the protein. The fraction saturation (\(\nu\)) of the protein with specific RNA can be determined experimentally for each ratio (\(\rho\)) of competitor RNA to specific RNA; specific RNA is radioactively labeled with \(^{32}\)P and is retained on the filter only if complexed to protein and addition of unlabeled competitor RNA reduces the fraction saturation (\(\nu\)). With the help of a nonlinear least-squares fitting algorithm (KaleidaGraph), a theoretical curve can be drawn to fit the data, using the ratio of the two dissociation constants as the fit parameter. This strategy is very useful because of the automatic correction for retention efficiency and specific \(^{32}\)P activity resulting from the introduction of (\(\nu\)). For simplicity, the activities of specific and competitor RNA, i.e., the ratio of RNA molecules able to bind SRP\(\Phi\)4-9, are assumed to be equal. As a consequence, the experimentally determined value for \(\kappa\) reflects also possible differences in RNA activity. The activity of the protein has to be measured experimentally. Furthermore, it is essential to work at full saturation of the protein and under equilibrium conditions. To this end, \(\text{app}K_d\) and the half-life of the complex with specific RNA have to be determined.

The purity of both protein and RNA is another critical point. SRP\(\Phi\)4-9 is expressed in E. coli and purified to homogeneity. It is taken from a stock that served to produce crystals for the determination of the X-ray structure of this protein (Birse et al., 1996). The RNA variants are synthesized in vitro by run-off transcription with T7 RNA polymerase from linearized DNA plasmid vectors (Milligan & Uhlenbeck, 1989). Initial RNA transcripts contain sequences at their 3' ends that can fold into “hammerhead” ribozymes that autocleave themselves at precise positions. Alu RNAs with a precise 3' end containing a 2' 3' cyclic phosphate are obtained by this procedure (Price et al., 1995). In order to achieve conformational homogeneity, RNAs are annealed before each experiment in a denaturation renaturation protocol. With this experimental setup, it is possible to determine essential identity elements of an Alu RNA by synthesizing the respective unlabeled competitor RNA and testing it together with a labeled specific reference RNA.

**Synthesis and purification of SA86 using “ribozyme technology” and characterization of the SRP\(\Phi\)4-9–SA86 complex**

Based on the secondary structure model of SRP301, SA86 is designed to contain all four (I-IV) previously identified regions of SRP301 (Fig. 1) that are protected from hydroxyl radical cleavage by SRP9/14 (Strub et al., 1991) in their natural base pairing context. To this end, the stem linking the Alu domain and the S domain of SRP301, here called adaptor stem, is short cut. Nucleotides 64 and 283 are linked with the sequence GUAA, a structurally well-defined tetraloop of the GNRA type (Heus & Pardi, 1991; Pley et al., 1994; Cate et al., 1996) that should stabilize the base pairing in the adjacent stem (Uhlenbeck, 1990). A 3' ribozyme cleaves co-transcriptionally and to almost completeness as indicated by the time course experiment in Figure 2. As much as one milligram of SA86 with a homogeneous 2' 3' cyclic phosphate end can be purified from 1 mL of transcription reaction. With respect to flAlu and scAlu, the size of SA86 corresponds to the most highly conserved domain of flAlu left monomer RNA (Sarrowa et al., 1997).

To determine the thermodynamic and kinetic stability of possible complexes of SRP\(\Phi\)4-9 and SA86, we did quantitative nitrocellulose filter-binding experiments (Yarus & Berg, 1967; Carey et al., 1983; Carey & Uhlenbeck, 1983). For measuring an \(\text{app}K_d\) for the complex, the binding curve was first determined as a function of total protein concentration. Based on this curve, the activity of the protein was determined in a separate experiment and the value for the \(\text{app}K_d\) corrected correspondingly. For obtaining the binding curve, a very low concentration (~0.1 nM) of cotranscriptionally \(^{32}\)P-labeled SA86 was titrated with increasing concentrations of SRP\(\Phi\)4-9 in a volume of 50 \(\mu\)L, using a 50-fold excess in mass of poly(rG) over SA86 as non-specific competitor. Complex formation was allowed to proceed for 10 min at room temperature, more time than necessary for equilibration (see below). Twenty-microliter volumes were filtered in duplicate without dilution and washed once with 150 \(\mu\)L of reconstitut-

**FIGURE 1.** Putative secondary structures of Alu RNA variants. A: Synthetic canine SRP RNA, here called SRP301. B: The minimal Alu RNA SA86. Deletion variants SA79 (C), SA50 (D), and SA54 (E). Circularly permuted variants SA88 (F) and SA85 (G). Stars mark uridines introduced to link the original 5' and 3' ends. SRP301 (A), SA86 (B), and SA88 (F) are recognized specifically by SRP\(\Phi\)4-9, whereas SA79 (C), SA50 (D), SA54 (E), and SA85 (G) are not. Boxed regions show absence of cleavage or reduced cleavage of SA85 in comparison to SA88 in the otherwise identical RNase V1 digestion patterns. Every tenth nucleotide is marked with a dot and the numbering of residues corresponds to the numbering in SRP301. The secondary structures are drawn corresponding to Strub et al. (1991). I-IV: Regions in synthetic canine SRP RNA that are protected from hydroxyl radical cleavage by canine SRP9/14.
FIGURE 1.
tion buffer. This reduced background retention and additional washes did not have any further effects, indicating that complexes, once bound, were kinetically stable. The plateau of the curve at 65% of total input can be explained by a retention efficiency of the protein-RNA complex of less than 100% combined with a protein binding activity of the RNA of less than 1 (Fig. 3B). Because the level of the plateau has no effect on the determination of the $K_d$ value, we did not try to resolve these two effects experimentally. An initial value of 35 nM could be obtained for the $K_d$. For independent experiments, this value was within a factor of 2-3, which is common with this method (Carey et al., 1983). This initial value for $K_d$ still had to be corrected for protein activity. To this end, a constant, high concentration of SA86 (500 nM unlabeled RNA and $\approx 0.1$ nM labeled RNA) was titrated with increasing concentrations of SRP14-9 until RNA became limiting. Below RNA saturation, complex formation depends linearly on the protein concentration. From the slope of this line, the protein activity can be calculated. It is important to realize that the value for the protein activity does not depend on the filter retention efficiency and that it is maximal for an RNA activity of 100%. A protein activity of only 20% was measured (Fig. 3A). Consequently, the $K_d$ was determined to be 7 nM. The low activity of the protein came as a surprise and will be discussed later. With regard to the competition experiments, the low protein activity as such has no consequence.

For estimating roughly the kinetic stability of the SA86-SRP14-9 complex, 2,500 nM cold SA86 was added to an equilibrated mixture of 500 nM SRP14-9, 500 nM unlabeled SA86, and $\approx 0.1$ nM $^{32}$P-labeled SA86 in a total volume of 250 µL. The quenching of label in the retained complexes was followed over time by filtering aliquots of 20 µL, which stopped the competition between labeled and unlabeled RNA. The experiment was done at 6°C because at room temperature the reaction was too quick to be followed. Ten minutes were more than enough to reach the new equilibrium, which is an important prerequisite for the analysis of the following competition experiments that were done under very similar conditions. From the resulting 0.03 s$^{-1}$ for the time constant ($k_{eq}$) (Fig. 3C), a half-life of the complex at room temperature of less than 20 s can be estimated for the buffer conditions used in the assay. The thermodynamic stability of the complex, reflected by the $K_d$ of 7 nM, is thus combined with a very dynamic equilibrium reflected by the short half-life of less than 20 s.

**Specificity of SA86 in comparison to SRP301**

We next wanted to test the specificity of recognition of SA86 by SRP14-9 in comparison with SRP301. To this end, we did equilibrium competition experiments as described above. We used labeled SA86 as the specific reference RNA. Unlabeled SRP301 served as a positive control for competition and unlabeled SS160 served as a negative control. SS160 is an RNA of 160 nt, representing the central 5'-sequence of SRP 301. Competition experiments were generally done in a volume of 50 µL at room temperature, premixing 500 nM of unlabeled SA86, $\approx 0.1$ nM of labeled SA86, and up to 10 µM of unlabeled competitor RNA, and then adding 500 nM of total SRP14-9. After 10 min, 20-µL aliquots were filtered in duplicate and washed once with 150 µL of reconstitution buffer. Filters carrying up to 48 samples were analyzed on a phosphorimager and data were plotted according to Equation 4 using ($a = 0.2$) for the protein activity as measured previously (Fig. 3B). The initial competition experiment shows...
that SA86 is indeed recognized specifically by SRPΦ14-9. Unlabeled SA86 competes with labeled SA86, as already observed in the kinetic experiment. SS160 does not compete with SA86, its $\text{app}K_D$ being at least 1,000-fold higher. This shows that, once the labeled RNA has dissociated, it can rebind the protein during the period of the experiment. SRP301 competes efficiently with SA86 and its apparent dissociation constant is only marginally lower ($\kappa = 1.3$) (Fig. 4A). A very similar result was obtained using labeled SRP301 and unlabeled SA86 (data not shown). The experiment confirms that the binding site for SRP9/14 is located in the Alu domain of SRP RNA. SA86 is identified as a small Alu RNA that can replace SRP301 functionally in the binding reaction with SRPΦ14-9.

It remained to be verified whether any of the new elements that were introduced to create SA86, namely the 2' 3' cyclic phosphate at the 3' end resulting from the ribozyme cleavage and the artificial tetraloop to close the adaptor stem, had an important effect on binding to SRPΦ14-9. We therefore generated two new Alu RNAs, SA87 and SA86'. At the 3' end of SA87, there is an additional guanosine with a free 3' hydroxyl group, as in SRP301. In SA86', the stabilizing GUAA tetraloop is replaced by UGAA, the nucleotides at positions 65–66 and 281–282 in SRP301. In the competition assay, SA86, SA87, and SA86' behaved virtually identically (data not shown). Because we find no difference in the affinity of SA87 and SA86 for SRPΦ14-9, we conclude that the 3' end is not an essential Alu RNA identity element, a possible effect of the natural sequence at the 3' end of 7SL RNA (Zwieb & Larsen, 1997) not being considered here. From the behavior of SA86', we conclude that a stabilization of the terminal base pairs in the adaptor stem by a GNA-type tetraloop is not essential and that the structurally well-defined GUAA tetraloop of SA86 does not interfere with SRPΦ14-9 binding.

**Deletion analysis of SA86**

Having established SA86 as a small Alu RNA that is functionally active in SRPΦ14-9 binding, we asked the question whether we could confine this Alu RNA motif even further. Based again on the secondary structure model of SRP301, three other RNA variants, SA50, SA54, and SA79 (Fig. 1), were designed and tested. SA50 only contains the protection regions I and II. The short 5' stem is designed to be elongated by two additional base pairs that should provide additional stacking energy for the formation of this stem. SA54 contains only the small 5' stem and the adaptor stem with the protection regions III and IV. The small 5' stem is closed by a loop that occurs similarly in the SRP RNA from *Schizosaccharomyces pombe* (Zwieb & Larsen, 1997), but contains the conserved CUGUAGU motif that is com-
mon to the mammalian 7SL RNAs (Strub et al., 1991). The 3' end of SA54 is generated by ribozyme cleavage as in SA86. SA79 lacks the complementary strand of the adaptor stem corresponding to nt 283–301 at the 3' end of SRP301. The protection regions III and IV thus cannot be in their natural base pairing context. This construct includes the sequence and structure elements necessary for efficient transcription by RNA polymerase III and would bind to factors involved in this process (Emde et al., 1997).

None of the three new RNA variants competed efficiently with SA86 for SRPφ14-9 (Fig. 4B). The result with SA50 and SA54 shows that neither an RNA with a complete deletion of the adaptor stem fragment (SA50) nor this fragment alone (SA54) can replace SA86 effectively. A reduction of the Alu RNA motif to a simple stem-loop structure is therefore not easily feasible. In addition to regions I and II, a part of the adaptor stem seems to be required. The comparison of SA79 with SA86, both of which contain all four protection regions, indicates that the sequence at the 3' end of SRP301 is required as well and that regions III and IV are only recognized in the secondary structure context of the adaptor stem. A construct containing only regions I, II, and III in their natural base pairing context has not been tested explicitly here, but Sarrowa et al. (1997) have shown for scAlus that region IV contains important determinants for SRP9/14 binding. If region IV is required, region III is likely to be required as well, if only for maintaining the spacial arrangement of regions I and II with respect to region IV. Therefore, we reason that most likely all four regions (I–IV) of SRP301 that are protected by SRP9/14 from hydroxyl radical cleavage have to be contained in an Alu RNA for binding of SRPφ14-9 without loss of affinity. These regions seem necessary, but are not sufficient. The example with SA79 shows that an Alu RNA has to adopt a defined secondary and probably tertiary structure in order to bind SRP9/14 specifically. If specific base recognition exists, it can only occur within a very defined structural context or is otherwise not detectable. We conclude, therefore, that SA86 is a minimal Alu RNA functional for SRP9/14 binding.

Analysis of circularly permuted variants of SA86—Is SA86 an autonomous folding unit?

In order to learn more about the relative organization of the different secondary structure elements in this minimal Alu RNA and in order to test whether its free 3' and 5' ends are essential for SRPφ14-9 binding, we analyzed two circularly permuted variants of SA86.
that we named SA88 and SA85 according to their length. The original 3' and 5' ends of SA86 were joined with either a long linker (UUUU) in SA88 or a short linker (U) in SA85 and new ends were created instead of closing the adaptor stem with a tetraloop (Fig. 1). In both cases, a cis-acting hammerhead ribozyme creates a clean 3' end. Based on the secondary structure model of SRP301, SA85 is designed to be rather rigid, forcing the adaptor stem to stack on the short 5' stem. SA88 should allow more flexibility between the adaptor stem fragment and the rest of the molecule. Surprisingly, the affinities of SA88 and SA85 for SRPφ14-9 turned out to be very different. When compared directly with SA86, SA88 showed only a hardly detectable twofold increase in the appKd, whereas SA85 did not compete at all (Fig. 4C). This shows that it is not essential to have the RNA termini at their natural location as long as some flexibility is allowed in the linker region. The fact that SA86 contains all the necessary information for correct folding and that the RNA termini have no major influence on the tertiary structure defines the Alu RNA binding site for SRP9/14 as an autonomous folding unit, i.e., independent of RNA context.

Because the two Alu RNAs, SA88 and SA85, have quasi-identical primary sequences, the dramatic distinction by SRPφ14-9 must be based mainly on structural properties of the two RNAs. To answer the question of whether SA88 and SA85 have a different base pairing pattern or whether the differences are more subtle, for instance, in the relative orientation of secondary structure elements, we digested the 5' 32P-labeled RNAs SRP301, SA86, SA88, and SA85 with RNase V1, which cuts specifically after double-stranded or stacked nucleotides (Ehresmann et al., 1987). The separation of the digestion products on a denaturing polyacrylamide gel yields a very similar pattern for all four RNAs, including SA85, and corresponds to the results of Gundelfinger et al. (1984) with 7SL RNA. This indicates that the major secondary structure elements are identical in all four RNAs. In spite of the similarities in the digestion patterns of SA88 and SA85, there are differences in the intensities at some sites (Fig. 5). In SA85, there are no cuts at all at positions A43-G44, and, at positions C290-U299, the cutting frequencies are much less than in SA88. It might be that the affected nucleotides are less well stacked or less frequently base paired in SA85 due to conformational strain exerted by the short linkage of the original 5' and 3' ends. However, the nucleotides do not form any alternative detectable base pairing and the corresponding cuts at the opposite side of the stem (U47-G48, C52-U53) are unchanged. Probably the more rigid nature of SA86 simply limits the access of the RNase V1 to certain sites (Figs. 1, 5). It is thus conceivable that SA86 consists of two distinct structural domains, a 5', possibly pseudoknotted domain that is oriented flexibly with respect to the double-helical adaptor stem fragment.

FIGURE 5. Secondary structure probing of 5' 32P-labeled SA88 and SA85. V1: Digestion with RNase V1, cutting double-stranded or stacked nucleotides. C: Control incubation without RNase. L: Ladder, generated by alkaline hydrolysis. T1: Guanosine ladder, generated by digestion with RNase T1 under denaturing conditions. Detectable guanines are numbered according to their position in SRP301. Boxed regions show absence of cleavage or reduced cleavage of SA85 in comparison to SA88 in the otherwise identical RNase V1 digestion patterns. Unequal intensities of bands within one lane probably reflect regiospecific accessibility of the RNA. Stars mark uridines introduced to link the original 5' and 3' ends.
DISCUSSION

We used the specific Alu RNA binding activity of the SRPΦ14-9 fusion protein for a functional assay. We show that SA86, a fragment of 7SL RNA containing the 5' and 3' ends and the sites found to be protected by SRP9/14 from hydroxyl radical cleavage, is necessary and sufficient for efficient binding to SRPΦ14-9. A defined tertiary structure seems to be required for recognition by the protein. SA88, a circularly permuted form of SA86, demonstrates the dispensability of the 5' and 3' ends for correct RNA folding and for SRPΦ14-9 binding, but flexibility between the 5' domain and the adaptor stem fragment seems to be necessary.

We found equilibrium competition experiments (Lin & Riggs, 1972; Carey et al., 1983) particularly useful for this system because of the possibility of comparing the affinities of two RNAs quantitatively under exactly the same conditions. This leads to a clear distinction between specific and nonspecific RNAs. All data were treated on the basis of a 1:1 binding stoichiometry between SRPΦ14-9 and an Alu RNA (Bovia et al., 1994), and our data give us no reason to refer to more complex stoichiometries. To perform and analyze these experiments correctly, the ratio of active protein (a), the apparent dissociation constant $K_d$, and the half-life of the SA86–SRPΦ14-9 complex had to be determined.

The protein activity was found to be low. About 80% of the protein does not bind to RNA. This came as a surprise because crystals were grown from the same batch of protein (Birse et al., 1996), making it unlikely that a substantial fraction of molecules was misfolded randomly. RNA and protein concentrations were determined accurately. A possible explanation comes from the crystal structure of SRPΦ14-9, where a $\beta$-mercaptoethanol molecule is bound covalently to the solvent-exposed C39 in the SRP9 sequence. This cysteine is located on a concave $\beta$-sheet surface, that might bind to a double-helical part of the RNA (Birse et al., 1997). The $\beta$-mercaptoethanol could therefore block RNA binding by simple sterical interference and account for the low activity if bound to about 80% of the protein molecules. The fact that N-ethylmaleimide alkylation of the two conserved sulphydryl groups in canine SRP9 is only possible in the absence of RNA and also seems to interfere with subsequent RNA binding supports this interpretation (Siegel & Walter, 1988). With an $K_d$ of 7 nM, the interaction of SRPΦ14-9 with SA86 is quite strong. The $K_d$ is, however, at least one order of magnitude higher than that for canine SRP9/14 with SRP301, which has been assayed with a different technique under very similar conditions (Janiak et al., 1992). Because we find SRP301 and SA86 to behave almost alike, the difference in the $K_d$ is most probably due to the protein component, which, in this case, is a fusion protein instead of a heterodimer. The fact of using murine instead of canine proteins might play a role as well, because species differences have also been shown to affect the $K_d$ (Bovia et al., 1997). The SA86–SRPΦ14-9 complex is very dynamic, with a half-life of less than 20 s. Whereas this is not unusual for an RNA–protein complex (Carey & Uhlenbeck, 1983), the fact that a stable scAlu RNP can be purified from HeLa cells (Bovia et al., 1995) indicates that the half-life might be shorter under physiological conditions and with different RNA–protein partners.

SA86 includes all the necessary identity elements for specific recognition by SRPΦ14-9 in the correct tertiary structure context and corresponds exactly to the most conserved domain of flAlus (Sarowwa et al., 1997). The sequence at the 3' end of SA86 is necessary for efficient binding of SRPΦ14-9, but dispensable for efficient transcription by RNA polymerase III. This argues against a role of SRP9/14 in transcriptional activation and explains its absence in RNA–protein complexes containing only the very 5' domain of 7SL RNA (Emde et al., 1997). The result with the circularly permuted SA88 shows that SA86 is an autonomous folding unit and that the sequence termini are neither critical for the tertiary structure nor for SRPΦ14-9 binding. Systematic circular permutation analysis of tRNA$^{The}$ has shown that most isomers fold correctly and has suggested that an autonomous tRNA folding unit can occur internally within other RNA sequences (Pan et al., 1991; Pan & Uhlenbeck, 1993). Internal Alu RNA sequences that are part of the untranslated 3' regions of mRNA could consequently adopt the correct tertiary structure to be recognized specifically by SRP9/14, and there is experimental support for this hypothesis in at least one case (Hsu et al., 1995). Such an internal recognition could affect any aspect of mRNA metabolism, an effect on translational control of the involved mRNA in cis (Siegel & Walter, 1986) being a particularly attractive idea.

Because SA88 and SA85 seem to have very similar secondary structure elements, suggesting a very similar tertiary structure, it is probably the flexibility between the original 5' domain and the adaptor stem fragment that permits SA88 to bind to SRPΦ14-9. Protein binding might involve an induced fit of the RNA, as observed for the ribosomal protein S15 and its RNA target (Batey & Williamson, 1996). Binding sites on the 5' domain and the 3' domain of SA86 could be brought together or, alternatively, one domain of SA86 could be moved out of the way to get access to a binding site on the other domain. Both kinds of movements would be blocked in SA85, but be allowed by the flexibility in SA88. Recognition would therefore be defined rather by what the RNA can be folded into than by the structure of the naked RNA. If it were true, however, that only one domain contains positive identity elements, this domain should fold independently and should work in the competition assay. The negative results with SA50 and SA79, containing the 5' domain, and
SA54, containing the adaptor stem fragment, indicate that this is probably not the case. Both domains of SA86 are necessary in their integrity and both seem to carry binding determinants for SRP9/14.

The recognition of Alu RNA variants by SRPΦ14-9 is highly specific, the protein being able to discriminate against subtle differences in RNA tertiary structure. Highly specific recognition of nucleic acid tertiary structure at stem junctions is known from DNA junction resolving enzymes (Duckett et al., 1995; Pöhler et al., 1996; White & Lilley, 1996), and indirect sequence readout via secondary and tertiary structure elements may be very important in general for RNA binding proteins, because, in contrast to DNA the major groove in A-helix, double-stranded RNA is not wide enough for direct sequence readout. The conservation of the 7SL RNA-like secondary structure of divergent flAlu and scAlu primary sequences, which has been claimed to correlate with the efficiency of both retroposition of genomic Alu elements and generation of small cytoplasmic RNAs (Chang & Maraita, 1993; Labuda & Zietkiewicz, 1994), is therefore probably just a reflection of the conservation of a tertiary structure, which might be necessary for the interaction with SRP9/14 or other cellular components. It has been suggested that binding of SRP9/14 could modulate the rate of flAlu retroposition (Sarrowsa et al., 1997) and, in this context, the correct secondary structure of the hydroxyl radical protection region IV has been shown to be important for efficient recognition by SRP9/14. In light of this result, one can speculate that in rigid SA85, region IV, although folded correctly, is not recognized at all because of a misorientation with respect to regions I and II.

SRPΦ14-9 is much smaller and much more compact than SA86 and might be contacted by an Alu RNA in several regions. These regions would include the concave β-sheet surface seen in the crystal structure of SRPΦ14-9 (Birse et al., 1997) and the regions abolishing RNA binding identified by mutational analysis (Bui et al., 1997). The mutations comprise the first part of the long SRP14 loop (residues 33–43) and point mutations in SRP9 (E15, D21, and/or P22), and they cluster on one edge of the heterodimer in the crystal structure. In a very speculative model, SRP9/14 would be sitting, with its concave surface on a double-helical part of SA86 like a saddle, being asymmetrically contacted from one edge by another, flexibly linked part of SA86. Current work on a crystal structure of a complete Alu ribonucleoparticle will show if this model holds true.

MATERIALS AND METHODS

Plasmids for in vitro transcription of RNA

Synthetic DNA oligonucleotides with the following sequences were ordered gel-purified without modifications and on 0.2 μmol scale from Genosys: oEcoT7 (5'-gggatccttaacgtacactca-3'), oSA86B (5'-gtcctagagcgctgctcggctcacacc-3'), oSA86BHA (5'-agctgctgtcttccggtatgggactcatttcggt-3'), oSA86HB (5'-gcgctgtctgcgatagctggtccatctcctcgc-3'), oSA86A (5'-gggatccttaactacgactcatttcggt-3'), oSA79B (5'-gggatccttaactacgactcatttcggt-3'), oSA79A (5'-atcttactatagctggtcctgc-3'), oSA50A (5'-gcgctgctgtccgtggtcctgc-3'), oSA50B (5'-gtcctagagcgctgctcggctcacacc-3'), oSA45A (5'-gggatccttaactacgactcatttcggt-3'), oSA881B (5'-gggatccttaactacgactcatttcggt-3'), oSA881A (5'-atcttactatagctggtcctgc-3'), oSA882A (5'-gcgctgctgtccgtggtcctgc-3'), oSA882B (5'-gtcctagagcgctgctcggctcacacc-3').

Plasmids pSA86H, pSA86H', pSA79, pSA50, pSA54H, pSA88H, and pSA85H, the "H" indicating a hammerhead ribosome sequence at the 3' end of the initial transcript, were obtained using E. coli strain DH5α cells and standard cloning techniques. For pSA86H, a PCR reaction on plasmid p75s wt (Strub et al., 1991) with oligos oEcoT7 and oSA86B yielded the coding sequence for SA86. The product was cut with EcoR I and Xba I and ligated to the respective sites in plasmid p75swt. The resulting plasmid contains a unique Nhe I site that we destroyed by double-digesting the plasmid with Nhe I and Nco I, filling up the Nhe I 3' overhang and religating the blunt ends. Next, the ribozyme sequence was introduced as a pair of synthetic DNA oligonucleotides (oSA86HA and oSA86HB), ligated directly to the Xba I and Hind III sites of the plasmid, giving rise to pSA86H. pSA86' H was obtained by replacing the EcoR I-Xba I fragment of pSA86H with a PCR fragment cut with the same enzymes. The PCR fragment was generated using oligos oEcoT7 and oSA86B on pSA86H. Plasmid pSA79 was obtained by a blunt-end insertion of a PCR product (oligos oSA79A and oSA79B on p75swt) into the Hind II site of SP65 (Promega). For pSA50, the coding sequence was obtained by PCR with oligos oSA50A and oSA50B on plasmid p75swt. The PCR product, digested with EcoR I and Xba I, was inserted into the respectively digested plasmid p75swt. For pSA54H, the PCR was done with oligos oSA54A and oSA86B on plasmid pSA86H and the digested product was inserted into pSA86H, again using the EcoR I and Xba I restriction sites. Plasmid pSA88H was generated in three ligation and amplification steps from three pairs of synthetic DNA oligonucleotides. Oligos oSA883A and oSA883B were annealed and ligated to the Xba I/Hind III double-digested plasmid pSA86H. The resulting plasmid was digested with BssH II and Xba I and the second pair of oligos (oSA882A and oSA882B) was inserted. Finally, oligos oSA881A and oSA881B were ligated to the EcoR I and BssH II sites to give pSA88H. For pSA85H, oligos oSA851A and oSA851B were inserted in pSA88H, making use of the BssH II and Nhe I sites.

Large-scale in vitro transcription and purification of RNA

For run-off in vitro transcription with T7 RNA polymerase (Milligan & Uhlenbeck, 1989), plasmids were linearized with

1271
Hind III, except for SRP301 (p7Swt cut with Xba I), SS160 (p7S1-A [Bovia et al., 1995] cut with Xba I), SA87 (pSA86H cut with Xba I), SA79 (pSA79 cut with Xba I), and SA50 (pSA50 cut with Sma I). Reactions were done in 1 mL of 40 mM Tris-HCl, pH 8.0, 5 mM DTT, 1 mM spermidine, 0.01% Triton X-100, 28 mM MgCl2, 4 mM of each NTP, 100 ng/mg DNA template, and 20 μLM T7 RNA polymerase. Specially designed hammerhead ribozymes (Price et al., 1995) autocleave themselves co-translationally. Nucleotides (Sigma) were stored as 100 mM solutions at −80 ºC. T7 RNA polymerase was expressed and purified as described by Davanloo et al. (1984) from E. coli strain BL21/pPAR122. After 2 h at 37 ºC, the transcription products were mixed with one volume of 95% formamide and loaded directly and fractionated on denaturing (8 M urea) 6%-8% polyacrylamide gels. The RNA was visualized with UV shadowing, excised, and eluted from the gel at room temperature for 14 h in 0.3 M sodium acetate. After ethanol precipitation, the RNA was dissolved (10 mg/mL) and stored in water at 4 ºC (Jones et al., 1996).

**Co-transcriptional and 5′ end labeling of RNA**

Co-transcriptional labeling of SA86 (Milligan & Uhlenbeck, 1989) for nitrocellulose filter binding was done in 10 μL using 2 mM each ATP, CTP, UTP, 20 μM cold GAP, and 20 μCi [α-32P]GTP (Amersham, 3,000 Ci/mmol). The general transcription conditions and the purification were the same as for large reactions, except that the detection of the RNAs in the gel was done by autoradiography. The concentration of the RNA was calculated from the specific activity of the [α-32P]GTP and the number of guanines present in the transcript. A 50-fold excess in mass of poly(rG) (Boehringer) was added as a nonspecific competitor to the eluted RNA before ethanol precipitation.

5′ End labeling of RNAs with [γ-32P]ATP (Amersham, 3,000 Ci/mmol) (Silberling et al., 1977) was done for the secondary structure probing experiments. Five micrograms of purified unlabeled RNA was dephosphorylated in a volume of 40 μL for 30 min at 60 ºC with 1 U of alkaline phosphatase (Boehringer), phenol extracted, ethanol precipitated, and labeled in 10 μL with 50 μCi of [γ-32P]ATP for 30 min at 37 ºC, using 10 U of T4 polynucleotid kinase (Boehringer). Labeled RNAs were re-purified on a 6% denaturing (8 M urea) polyacrylamide gel. Radioactively labeled RNAs were stored at −100,000 cpm/μL in reconstitution buffer at −20 ºC.

**Nitrocellulose filter binding**

SRP14-9 was expressed and purified as in Birse et al. (1996). It was dialyzed against reconstitution buffer (20 mM HEPES-KOH, pH 7.5, 5 mM DTT, 300 mM potassium acetate, 10 mM magnesium acetate, 0.015% Nikkol) and the final concentration was determined spectrophotometrically, using a molar extinction coefficient of ε280 = 12,500 M−1 cm−1 (Gill & von Hippel, 1989). Labeled and unlabeled RNAs (20 μM) were annealed in reconstitution buffer by heating the samples 10 min to 65 ºC and letting them cool over 1 h to the assay temperature. Concentrations of unlabeled RNA were determined spectrophotometrically using a value of 40 μg/mL for 1 OD260. The filter binding experiments were done in reconstitution buffer as described in the text using a Milliblot S microfiltration device (Millipore), allowing treatment of 48 samples simultaneously on a single filter. The 0.45-μm pore filters (Schleicher & Schuell) had been soaked previously for 10 min in reconstitution buffer. In order to reduce pipetting errors, protein was never pipetted in volumes less than 5 μL, and premixes were done wherever possible. The total [32P] input per filtered aliquot was approximated to be 10,000 cpn.

For determining the protein activity and the apparent dissociation constant, the filters, after drying, were cut and the amount of radioactivity was measured for each spot on the filter by liquid scintillation counting (Beckman), corrected for background retention, and divided by the total amount of radioactivity in the sample. The total input was determined accurately by spotting an aliquot on a nitrocellulose membrane and measuring it without filtering. For app.Kd, a curve based on a simple bimolecular reaction was fit to the data using Equation 1 for the nonlinear least-squares fitting algorithm of KaleidaGraph:

\[
\text{cpm}(P_a)/\text{cpm}_0 = (\text{cpm}_{\text{max}}/\text{cpm}_0) \times P_a/(P_a + \text{app.K}_d).
\]  

Equation 1

\[
\text{cpm}(t) = (\text{cpm}(0) - \text{cpm}(\infty)) \times e^{-k_{off}t} + \text{cpm}(\infty).
\]  

Equation 2

\[\kappa = \frac{K_{ds}}{K_{ds} - (R_{eq} - PR) \times PR}, \quad \text{cpm}(0) \quad \text{and} \quad \text{cpm}(\infty) \quad \text{are the filter-bound cpm before starting the competition and after re-equilibration respectively and} \quad K_{off} \quad \text{is the fit parameter.} \]

\[\kappa = \frac{K_{ds}}{K_{ds} - (R_{eq} - PR) \times PR}; \quad \text{cpm}(0) \quad \text{and} \quad \text{cpm}(\infty) \quad \text{are the filter-bound cpm before starting the competition and after re-equilibration respectively and} \quad K_{off} \quad \text{is the fit parameter. For the equilibrium competition assays, the dried filters were analyzed directly in a phosphor-imager (BioRad) using the associated software and taking care to be within the linear response of the detector. Measuring the input experimentally was not necessary here. The two mass-action equations for binding specific or competitor RNA to the protein are combined into Equation 3:}

\[\kappa = \frac{K_{ds}}{K_{ds} - (R_{eq} - PR) \times PR}; \quad \text{cpm}(0) \quad \text{and} \quad \text{cpm}(\infty) \quad \text{are the filter-bound cpm before starting the competition and after re-equilibration respectively and} \quad K_{off} \quad \text{is the fit parameter. For the equilibrium competition assays, the dried filters were analyzed directly in a phosphor-imager (BioRad) using the associated software and taking care to be within the linear response of the detector. Measuring the input experimentally was not necessary here. The two mass-action equations for binding specific or competitor RNA to the protein are combined into Equation 3:}

\[\kappa = \frac{K_{ds}}{K_{ds} - (R_{eq} - PR) \times PR}; \quad \text{cpm}(0) \quad \text{and} \quad \text{cpm}(\infty) \quad \text{are the filter-bound cpm before starting the competition and after re-equilibration respectively and} \quad K_{off} \quad \text{is the fit parameter. For the equilibrium competition assays, the dried filters were analyzed directly in a phosphor-imager (BioRad) using the associated software and taking care to be within the linear response of the detector. Measuring the input experimentally was not necessary here. The two mass-action equations for binding specific or competitor RNA to the protein are combined into Equation 3:}

**κ** is the apparent ratio of the dissociation constants, Kds and Keq for complexes of the protein with specific or competitor RNA, respectively. Rq and Kds are the total RNA concentrations for specific and competitor RNA and PR, and PR, are the equilibrium concentrations of complexes of the protein with specific or competitor RNA, respectively. The total concentrations of protein, Pq, and specific RNA, Rq, are pipetted to be equimolar and are chosen high enough to reach full saturation of the protein. A correction is made for the activity (a) of the protein. This is the fraction of protein capable of specific RNA binding. Pq is the total concentration of active protein; Pq = a × Rq. Now PR, can be expressed as:

\[\text{PR} = \frac{PR}{a \times R_{eq}} = \frac{pdu}{pdu_{\text{max}}};
\]  

Equation 4
\( \nu \) can be determined experimentally for each concentration of total competitor RNA by dividing the filter-bound radioactivity (expressed in phosphorimagery pixel density units, \( pda \)) by the filter-bound radioactivity in the absence of competitor (\( pda_{max} \)). Defining \( (\rho) \) as the ratio of total competitor RNA to total specific RNA, \( (\rho = \frac{R_c}{R_a}) \), Equation (3) can be regrouped into the basic competition equation:

\[
\kappa = \frac{1 - \nu}{\nu} \times \frac{1 - a \times \nu}{\rho - a \times (1 - \nu)}.
\]

Finally, Equation 5 can be solved for \( \nu \) as a function of \( \rho \), \( (\nu = f(\rho)) \), and then be used to fit \( \kappa \) to the data, again using a nonlinear least-squares fitting algorithm (KaleidaGraph):

\[
\nu = \frac{\kappa \times (a - \rho) - 1 - a + \sqrt{(\kappa \times (a - \rho) - 1 - a)^2 + 4 \times a \times (\kappa - 1)}}{2 \times a \times (\kappa - 1)}.
\]

### Probing of RNA secondary structure

5’ Labeled and annealed RNA (100,000 cpm) was digested with 0.05 U of RNase V1 (Pharmacia) in 20 \( \mu \)L of reconstitution buffer for 30 min at 37°C. A control without RNase V1 was run in parallel. After phenol extraction and ethanol precipitation, the fragments were analyzed on a 15% denaturing (8 M urea) polyacrylamide gel. To trace the sequence of the respective RNA, a guanosine ladder was created, digesting 100,000 cpm with 0.1 U of RNase T1 (Pharmacia) in 10 \( \mu \)L under denaturing conditions (4.5 M urea) for 15 min at 55°C. A complete ladder, serving the same purpose, was obtained by partial alkaline hydrolysis (100,000 cpm in 5 \( \mu \)L of 0.1 M NaHCO\(_3\), pH 9.5, for 3 min at 95°C). All samples contained 1 \( \mu \)g of unlabeled E. coli tRNA (Boehringer). For further details, see Ehresmann et al. (1987).

### ACKNOWLEDGMENTS

We thank Guillaume Bec (Grenoble) for purifying T7 RNA polymerase, Florence Baudin (Grenoble) for fruitful discussions, and Ludwig Lehle (Regensburg) for initial support. O.W. received a Boehringer-Ingelheim-Found predoctoral fellowship. S.C. and K.S. are both members of SRPNET, which is supported by the European Union Training and Mobility of Researchers program and the Swiss Government. K.S. is a fellow of the START program of the Swiss National Science Foundation.

Received June 27, 1997; returned for revision July 28, 1997; revised manuscript received August 14, 1997

### REFERENCES


Bovia F, Fornallaz M, Leffers H, Strub K. 1995. The SRP/14 subunit of the signal recognition particle (SRP) is present in more than 10-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.


Sarroja J, Chang DY, Maraia RJ. 1997. The decline in human Alu retroposition was accompanied by an asymmetric decrease in SRP9/14 binding to dimeric Alu RNA and increased expression of small cytoplasmic Alu RNA. *Mol Cell Biol* 17:1144–1151.


