Structure of the complete bacterial SRP Alu domain

Georg Kempf, Klemens Wild[*] and Irmgard Sinning[*]

Heidelberg University Biochemistry Center (BZH), INF 328, D-69120 Heidelberg, Germany

ABSTRACT

The Alu domain of the signal recognition particle (SRP) arrests protein biosynthesis by competition with elongation factor binding on the ribosome. The mammalian Alu domain is a protein–RNA complex, while prokaryotic Alu domains are protein-free with significant extensions of the RNA. Here we report the crystal structure of the complete Alu domain of *Bacillus subtilis* SRP RNA at 2.5 Å resolution. The bacterial Alu RNA reveals a compact fold, which is stabilized by prokaryote-specific extensions and interactions. In this 'closed' conformation, the 5′ and 3′ regions are clamped together by the additional helix 1, the connecting 3-way junction and a novel minor groove interaction, which we term the 'minor-saddle motif' (MSM). The 5′ region includes an extended loop–loop pseudoknot made of five consecutive Watson–Crick base pairs. Homology modeling with the human Alu domain in context of the ribosome shows that an additional lobe in the pseudoknot approaches the large subunit, while the absence of protein results in the detachment from the small subunit. Our findings provide the structural basis for purely RNA-driven elongation arrest in prokaryotes, and give insights into the structural adaption of SRP RNA during evolution.

INTRODUCTION

The signal recognition particle (SRP) plays an essential role in co-translational targeting of newly synthesized membrane proteins [1,2]. SRP is a ribonucleoprotein complex conserved in all three kingdoms of life with a high diversity regarding composition and complexity [3]. Eukaryotic SRP contains six proteins assembled on a 7SL RNA [4] and can be divided into two functional domains. While the S domain recognizes SRP targets through their N-terminal signal sequences as soon as they emerge from the ribosomal tunnel exit, the Alu domain imposes an elongation arrest by blocking the elongation factor entry site [5–8]. By retarding translation, SRP prevents membrane proteins from being prematurely released from the ribosome before the ribosome-nascent chain complex (RNC) has correctly en-

gaged with the translocation channel at the endoplasmic reticulum membrane [1,9]. The Alu domain of higher eukaryotes is composed of the 5′ and 3′ regions of SRP RNA and the two Alu RNA-specific proteins SRP9/14 (Figure 1A). The proteins stabilize the complex tertiary structure of the Alu RNA and contribute to ribosome binding [10–12]. In a cryo-EM structure of mammalian SRP bound to the RNC, the SRP9/14 proteins were shown to interact with the small ribosomal subunit, while the Alu RNA establishes a contact with the large ribosomal subunit [12,13]. The structure of the Alu RNA is instructive also for understanding of the retrotransposable, repetitive *Alu* elements, which comprise more than 10% of the primate genome and are derived from the 7SL RNA [14–18]. Despite their abundance, the precise roles of *Alu* elements are still poorly understood, and their function in gene regulation or as templates for the production of new exons is just emerging [18,19].

In most prokaryotes, the SRP RNA (long SRP RNA) also contains an Alu domain, and the Alu RNA has significant extensions compared to higher eukaryotes [3,20]. However, homologs of the SRP9/14 proteins have not been found in the genomes of archaea or bacteria so far [21]. In case of *Bacillus subtilis*, the DNA-binding protein HU1, which belongs to the family of histone-like proteins, has been suggested to be part of the bacterial Alu domain [22,23]. Due to the pleiotropic roles in nucleic acid binding and genome maintenance described for HU1 and homologous proteins, specific Alu RNA association or function in SRP could not be established [24–27]. SRP RNA in Gram-negative bacteria (4.5S RNA) does not contain an Alu domain (except in Thermotogae) whose presence in prokaryotes does not precisely correlate with criteria such as endospore formation or the presence of a cell wall. In protists like the ancient Trypanosoma parasites, SRP contains two RNA molecules, a 7SL RNA with a truncated 5′ region of the Alu domain and a tRNA-like molecule [28], while Alu domain binding proteins are absent [29]. In order to provide the structural basis for the prokaryotic Alu RNA and the general blueprint of the complete Alu RNA fold, we determined the crystal structure of the complete *B. subtilis* Alu domain.

[*]To whom correspondence should be addressed. Tel: +49 6221 544781; Fax: +49 6221 544790; Email: irmi.sinning@bzh.uni-heidelberg.de
Correspondence may also be addressed to Dr Klemens Wild. Tel: +49 6221 544785; Fax: +49 6221 544790; Email: klemens.wild@bzh.uni-heidelberg.de
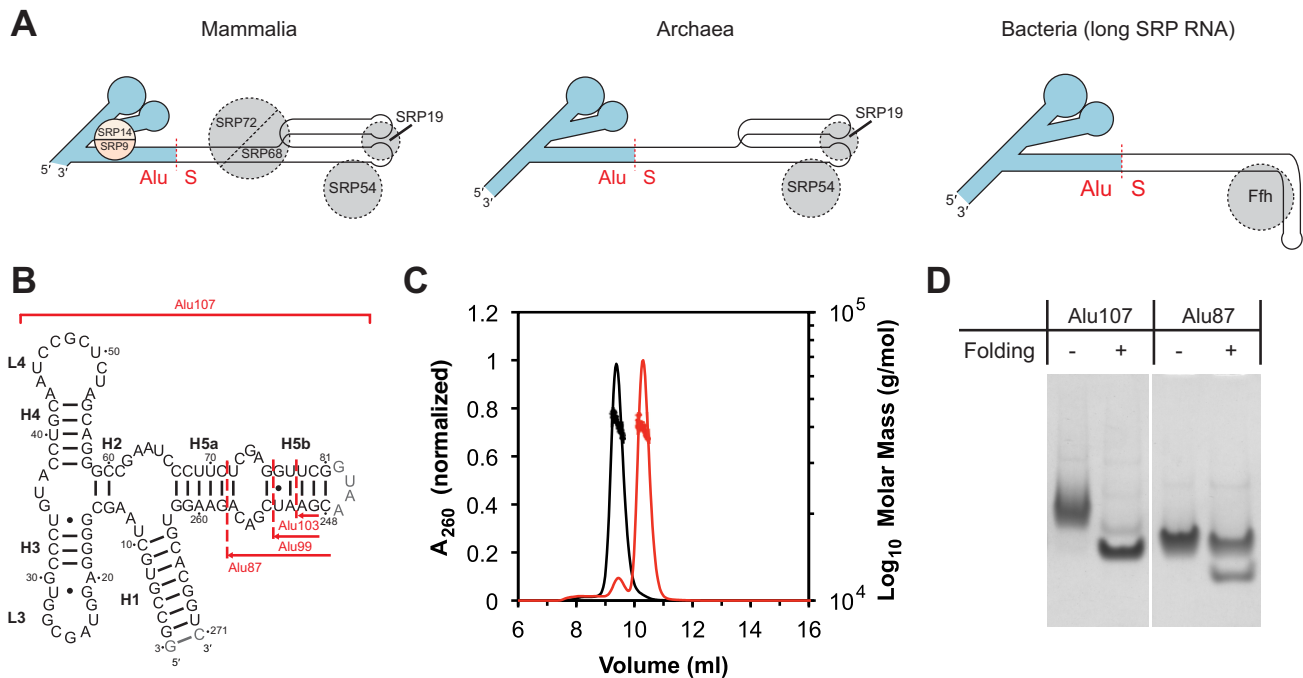
Figure 1. The bacterial SRP Alu domain. (**A**) Schematic representation of the mammalian, archaeal and bacterial SRP. 'Long SRP RNA' refers to the presence of a 6S RNA in many gram-positive bacteria, in contrast to the short 4.5S RNA in most gram-negative bacteria. Mammalian and archaeal SRP share a 7S RNA. The Alu RNA is highlighted in blue. (**B**) Constructs of the *Bacillus subtilis* Alu domain RNA. The S domain is replaced by a tetraloop at the end of helix 5 (gray). (**C**) Size-exclusion chromatography coupled with MALS of unfolded (black) and folded (red) Alu107 RNA. The lines correspond to the UV signal and the dots to the molar mass distribution at the respective peak. (**D**) Electrophoretic mobility shift assay demonstrating the homogeneous folding of Alu107 RNA and comparison with deletion construct Alu87 (shortened helix 5).

MATERIALS AND METHODS

Cloning

DNA coding for Alu domain sequences (GenBank: D11417.1) fused to an upstream T7 promoter were obtained by annealing overlapping oligonucleotides followed by filling the overhangs using *Taq* polymerase. DNA coding for a hammerhead (HH) ribozyme as used previously (30) at the 3′ end of the constructs was obtained by annealing complementary oligonucleotides. Sequences coding for Alu87 (nucleotides 3 to 71 and 258 to 271) and Alu107 (nucleotides 3 to 81 and 248 to 271) were cloned into pUC19 via EcoRI/XbaI and are followed by the HH ribozyme sequence that was cloned via XbaI/HindIII yielding pUC19-Alu87-HH and pUC19-Alu107-HH. pUC19-Alu103-HH (nucleotides 3 to 79 and 250 to 271) and pUC19-Alu99-HH (nucleotides 3 to 77 and 252 to 271) were obtained by site-directed mutagenesis using pUC19-sc107-HH as a template. In all constructs, the SRP S domain was replaced by a closing GUAA tetraloop.

RNA synthesis and purification

SRP Alu RNA was produced by *in vitro* transcription as described earlier (30) using T7 polymerase and HindIII-linearized plasmid DNA. The Alu RNA was purified by denaturing polyacrylamide gel electrophoresis as described (30) and desalted in water using a PD-10 desalting column (GE Healthcare). For crystallization, the RNA was concentrated to 10–30 mg/ml.

Crystallization

The RNA was diluted in folding buffer containing 20-mM Tris pH 8.0, 10-mM $MgCl_2$, 10-mM KCl, 200-mM NaCl and 8% (v/v) glycerol to a concentration of 3–10 mg/ml, heated to 65°C and then slow-cooled to 20°C. Prior to setting up crystallization experiments, the RNA was centrifuged.

Crystallization was performed by sitting drop vapor diffusion, by robotized mixing 150 nl of RNA with 150 nl of mother liquor and incubating at 18°C. Crystals of Alu103 (space group $C222_1$) grew within 1–3 days from 50-mM sodium cacodylate pH 6.5, 10-mM sodium acetate pH 4.6, 20-mM $MgCl_2$, 120-mM sodium thiocyanate and 1.3 M $LiSO_4$ or 50-mM sodium cacodylate pH 6, 20-mM $MgCl_2$, 350-mM sodium formate, 1.3 M $LiSO_4$. Crystals of Alu107 RNA (space group $P2_12_12_1$) grew from 50-mM sodium cacodylate pH 6.5, 200-mM ammonium acetate, 10-mM magnesium acetate, 30% (w/v) PEG8000 and 2.5% (v/v) jeffamine-M600 within one week.

Derivative crystals used for phasing were prepared by soaking the crystals directly in the crystallization drop (300 nl) to which either 1.5 μl of mother liquor and 200 nl of 100-mM cobalt (III) hexamine chloride solution (10 mM final concentration) or 1.5 μl of mother liquor and 200 nl of 5 M cesium chloride solution (500 mM final concentration) were added followed by incubation for 1–2 h. Another cobalt (III) hexamine derivative, which was not used for phasing, was prepared by transferring a $C222_1$ crystal into 50-mM sodium cacodylate pH 6, 10-mM magnesium acetate, 1.3-M

Table 1. Data collection, phasing and refinement statistics

	Alu103 (Cesium derivative)	Alu107 (Co[NH$_3$]$_6$ derivative)		Alu107 (Co[NH$_3$]$_6$ derivative)
Data collection				
Space group	C222$_1$	P2$_1$2$_1$2$_1$		C222$_1$
Cell dimensions				
a, b, c (Å)	180.1, 194.1, 83.4	62.5, 68.3, 83.0		176.3, 194.3, 83.0
α, β, γ (°)	90, 90, 90	90, 90, 90		90, 90, 90
		Peak	*Remote*	
Wavelength (Å)	2.0664	1.60489	1.03321	0.8726
Resolution (Å)	57.4–3.4 (3.6–3.4)	49.9–2.86 (3.0–2.86)	46.1–2.49 (2.59–2.49)	49.05–3.10 (3.31–3.10)
R_{pim} (%) [a]	2.3 (50.6)	6.8 (33.5)	5.6 (50.9)	6.4 (47.0)
I / σI	18.7 (1.4)	12.9 (2.6)	12.0 (2.0)	13.8 (2.0)
Completeness (%)	99.8 (98.8)	100.0 (100.0)	99.9 (100.0)	99.8 (99.8)
Redundancy	36.0 (34.5)	15.6 (15.9)	10.5 (10.8)	6.4 (6.5)
Phasing				
FOM before DM	0.26			
FOM after DM	0.75			
Refinement				
Resolution (Å)			2.49	3.1
No. reflections			12953	26252
R_{work} / R_{free} (%)			18.0/20.9	19.1/20.8
Luzzati error (Å) [b]			0.46	0.52
No. atoms				
RNA			2302	4434
Ligand/ion			120	162
Water			34	4
B-factors (Å^2)				
RNA			41.65	89.37
Ligand/ion			69.64	108.32
Water			35.11	80.9
R.m.s deviations				
Bond lengths (Å)			0.002	0.002
Bond angles (°)			0.587	0.524

Each structure was determined from one crystal. Values in parentheses are for highest-resolution shell. FOM: figure of merit; DM: density modification.
[a]The multiplicity-weighted R_{pim} is calculated for all I+ & I−.
[b]Error in coordinates by Luzzati plot (calculated by sfcheck (43) of the CCP4 suite).

lithium acetate pH 6, 100-mM cobalt (III) hexamine acetate followed by incubation for 2 h at room temperature. Cobalt (III) hexamine acetate was prepared essentially as described (31). Crystals were cryo-protected by dipping into mother liquor supplemented with 15% to 25% (v/v) glycerol and flash-cooled in liquid nitrogen.

Data collection and processing

Datasets of the cobalt (III) hexamine derivative crystal with space group P2$_1$2$_1$2$_1$ were collected at beamline ID23-1 of the European Synchrotron Radiation Facility (ESRF) at the peak (1.60489 Å), inflection point (1.60549 Å) and high remote wavelength (1.03321 Å). The dataset of the cesium derivative crystal was collected at beamline ID29 of the ESRF at a wavelength of 2.0664 Å, which is close to the peak wavelength of the cesium absorption edge. Data of the cobalt (III) hexamine derivative in space group C222$_1$ was collected at beamline ID23-2 of the ESRF at a wavelength of 0.8726 Å. All data were integrated using XDS (32) and scaled using AIMLESS (33,34). Initial phases could be obtained for the cesium derivative crystal, which contained two molecules in the asymmetric unit, using single-wavelength anomalous dispersion (SAD) phasing

in Phenix.autosol (35,36). An initial model was then used in MR-SAD to calculate phases for the cobalt (III) hexamine derivative crystal belonging to space group P2$_1$2$_1$2$_1$ and containing one molecule in the asymmetric unit by using Phenix.phaser (37) and Phenix.autosol. The structure of the cobalt (III) hexamine derivative crystal with space group C222$_1$ was solved by MR using Molrep (38) and the refined model of the P2$_1$2$_1$2$_1$ crystal structure as search model.

Refinement

The structure of the cobalt (III) hexamine derivative crystal with space group P2$_1$2$_1$2$_1$ was refined against anomalous data (remote dataset) using Phenix.refine (39) and iterative model building in COOT (40). In addition, experimental phases of the peak dataset were included in early refinement steps. Before the final refinement steps, the model was automatically rebuilt with ERRASER (41). In late refinement steps, the X-ray/stereochemistry and the X-ray/atomic displacement parameter weights were optimized and the refinement strategy additionally included refinement of translation-libration-screw groups (calculated with phenix.refine). Cobalt and magnesium ions were essentially placed according to their locations determined by
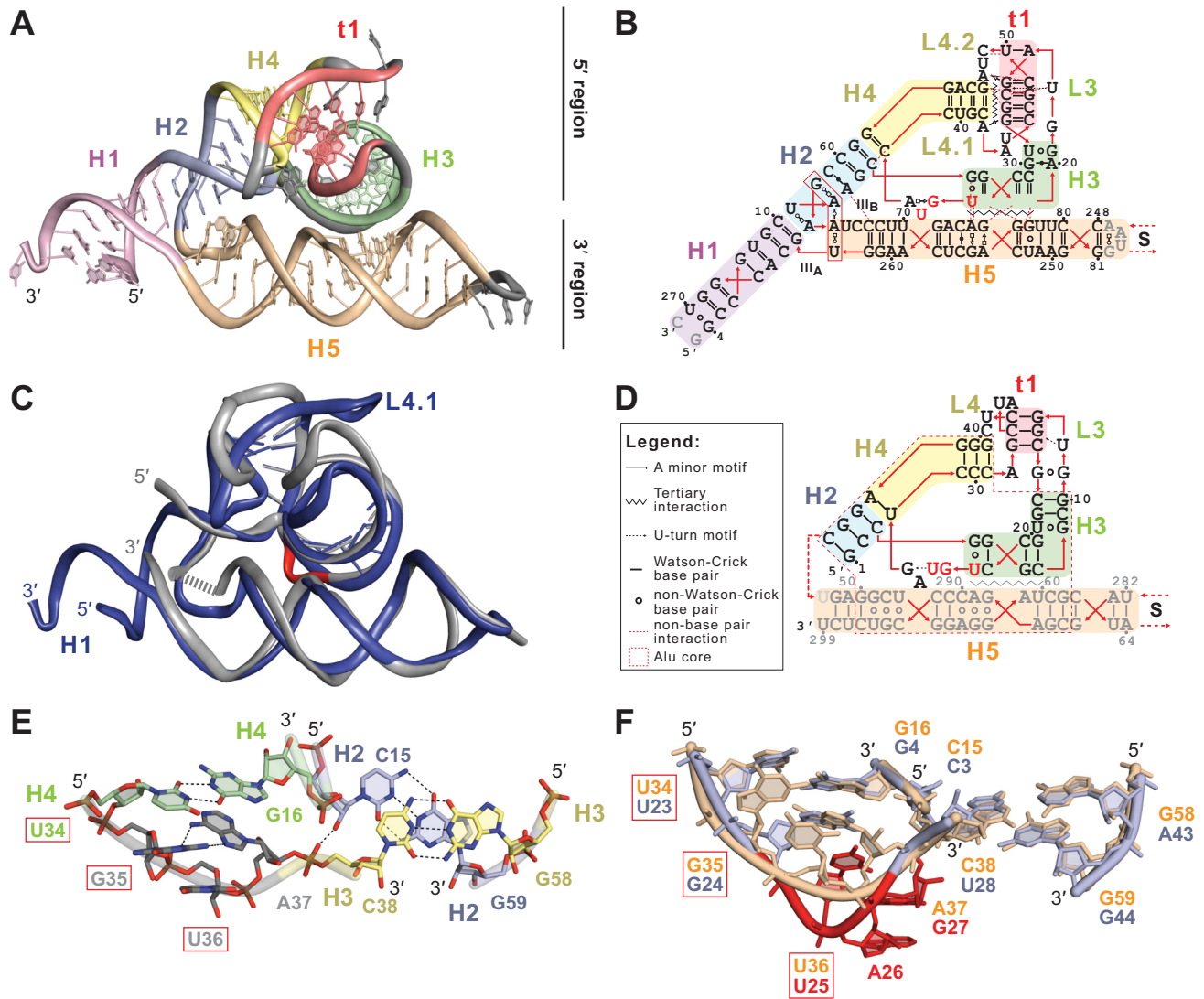
Figure 2. Structure of the Alu domain. (**A**) The structure of *Bacillus subtilis* Alu domain. Helices and the tertiary interaction (t1) in the loop-loop pseudoknot are indicated. (**B**) Two-dimensional representation of the Alu domain with helix and junction numbering and tertiary interactions. The four-base platform is boxed and the UGU sequence is highlighted in red. Artificial bases at the ends are shown in gray. Base-pairs are classified according to the nomenclature by Leontis and Westhof (46). (**C**) Superposition of *B. subtilis* and human Alu domains (combined human model based on PDB entries 1E8O and 1E8S; (10)) based on highly conserved nucleotides within helices 3 and 4 (shown as sticks). The additional helix 1 and L4.1 lobe in bacteria are labeled, and the U-turn including the UGU sequence in human SRP is shown in red. Human helix 5 belongs to a crystallographic neighbor (10). (**D**) Two-dimensional representation of the human Alu domain in the closed conformation (as in (B)). A dashed red line boxes the conserved Alu domain core. Labels also used in (B) are given in the inset. (**E**) The 3-way junction III_B between helices 2, 3 and 4 and the UGU sequence (nucleotides 34 to 36, boxed in red) shown for the *B. subtilis* Alu domain. An intra-strand *trans* Hoogsteen/sugar-edge base pair (G35/A37) bends the connecting loop between helices 3 and 4. Color-coding as in Figure 2B. (**F**) Superposition of the *B. subtilis* UGU sequence and 3-way junction III_B (colored in brown) with the corresponding part of the human Alu domain (blue, PDB code 1E8O) including the U-turn motif (U25-G27, red) instead of an intra-strand base pair.

log-likelihood gradient map completion in Phenix.phaser. The structure of the cobalt (III) hexamine derivative crystal with space group C222₁ was refined against anomalous data as described above. Ions were placed manually. The model quality was analyzed using MolProbity (42) and SFCHECK (43). Interaction interfaces were calculated with PISA (44). Base pair types were annotated with RNAview (45,46). Hydrogen bonds were annotated manually and with the help of DSSR of the 3DNA package (47,48). Helix parameters were obtained using the Curves+ web server (49). Structural figures were prepared using PyMol (50).

Sequence alignments

Sequence alignments were performed using LocARNA (51) integrated in the Freiburg RNA tools web server (52). For alignments, secondary structure predictions obtained from the SRPDB (21) were used in combination with manually defined constraints. The secondary structure prediction for *B. subtilis* SRP RNA was updated with the information obtained from the crystal structure.
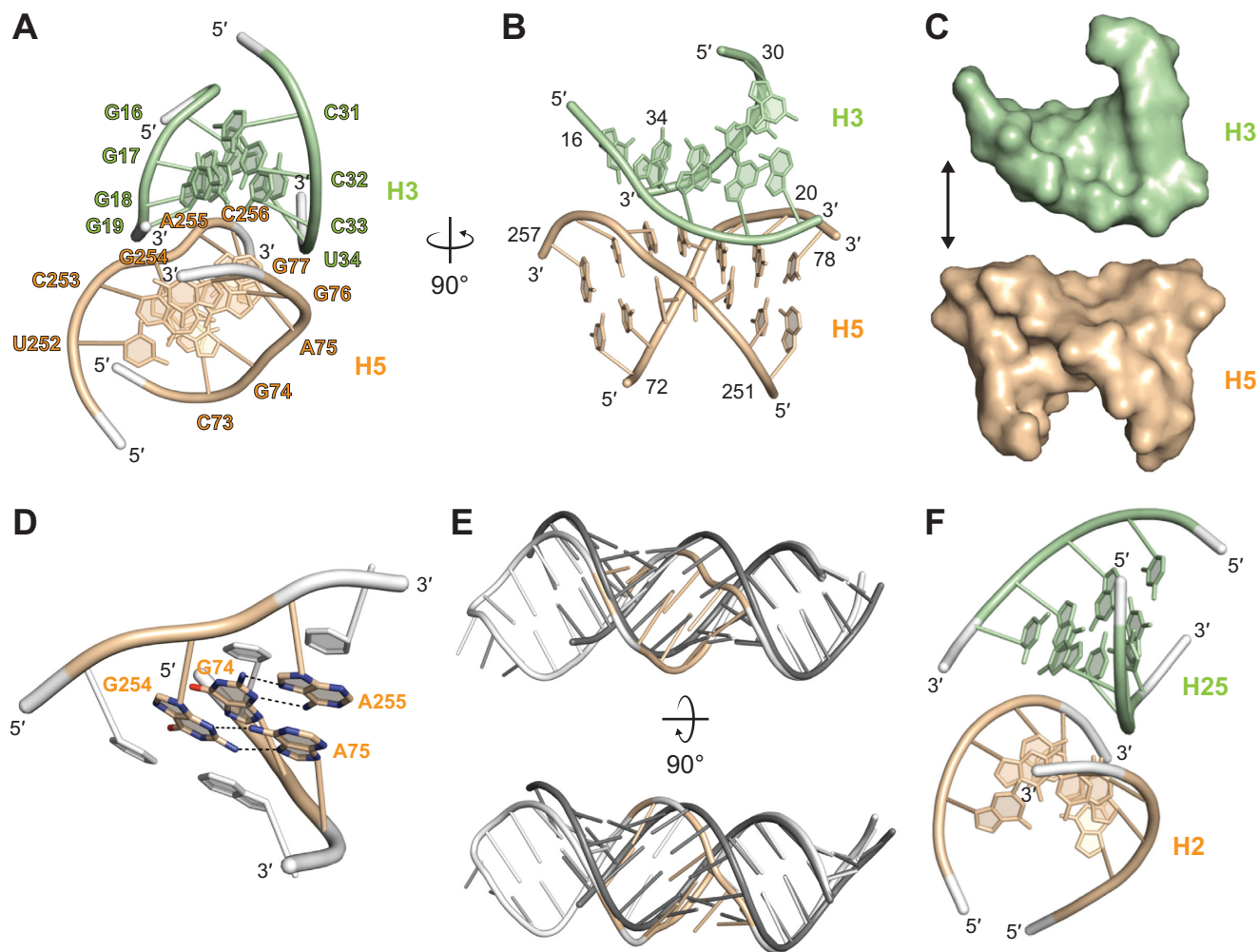
Figure 3. The minor-saddle motif. (**A**) The minor-saddle motif (MSM) between helices 3 and 5 including five non-Watson–Crick base pairs and purine cross-strand stacking. The twisted helix 5 forms a saddle for the perpendicular aligned minor groove of helix 3. (**B**) The MSM viewed perpendicular to the helix axis. (**C**) Surface representation of Figure 3B highlighting the flat saddle-like appearance of helix 5 and the aligned positioning of the two helices. The two helices are separated vertically for clarity. (**D**) The cross-strand purine stacking involving the two central sheared G-A base pairs. (**E**) Adaptation of the minor groove of helix 5 within the MSM. The groove is flattened and narrowed compared to an ideal A-RNA helix (gray). (**F**) A standard minor-groove interaction without saddle present in the large ribosomal subunit (PDB entry 2J01, (56)).

Native gel electrophoresis

The RNA was either diluted in water or in folding buffer and, in the latter case, folded as described above. Two micrograms of RNA were loaded to an 8% magnesium acetate-polyacrylamide gel and the gel was run in 25-mM Tris, 7.5-mM MgOAc, pH 7.5 at 4°C first at 50 V for 15 min until the samples had completely entered the gel and then at 150 V for 1.5–2 h. After electrophoresis, bands were visualized by staining with methylene blue.

SEC-MALS experiments

Unfolded and folded RNA samples were obtained as described above and subjected to size-exclusion chromatography (SEC) using a Superdex™ 75 10/300 GL column equilibrated in a buffer containing 20-mM Tris/HCl pH 8.0, 200-mM NaCl, 10-mM KCl, 10-mM MgCl$_2$. SEC was coupled to a MALS (miniDAWN Tristar, Wyatt Technologies) and refractive index detector (RI-71, Shodex) allowing de-

termination of the absolute molar mass. For calculation of the molar mass, a dn/dc value of 0.170 was used.

RESULTS

Crystallization and structure determination of the *B. subtilis* Alu domain

A number of *B. subtilis* Alu domain variants were generated by replacing the S domain with a GUAA tetraloop and tested in folding trials. The variant comprising the complete Alu domain (Alu107) folded into a homogeneous species after annealing in buffered salt solution as judged from a significantly smaller hydrodynamic volume in size exclusion chromatography coupled with multi-angle light scattering (MALS) and a faster migration behavior in gel electrophoresis (Figure 1B–D). Alu107 as well as two variants with helix 5 shortened by two (Alu103) or four (Alu99) base pairs readily crystallized in several conditions. An initial native dataset for Alu103 could be collected to 3.5 Å resolu-
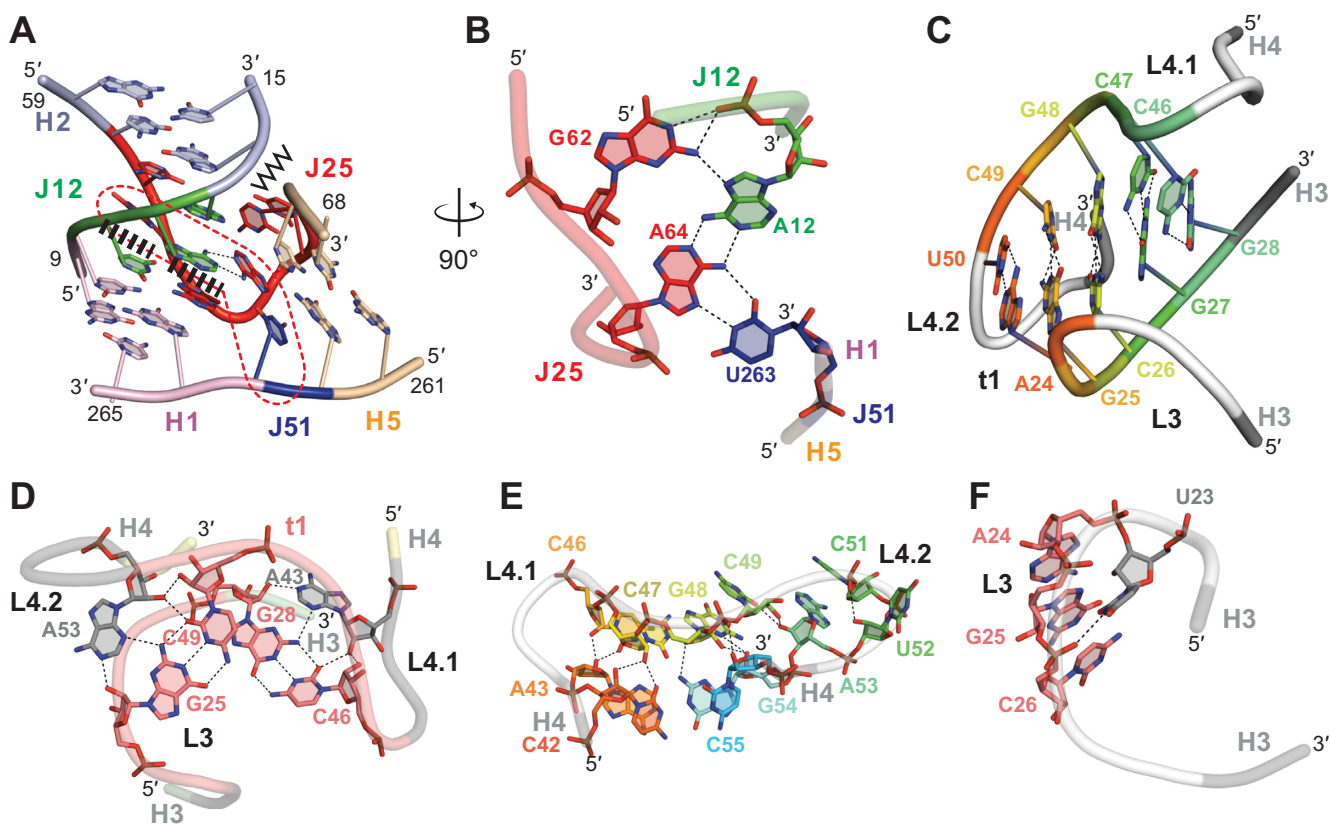
Figure 4. Prokaryote specific extensions. (**A**) The junction III$_B$ connecting helices 1, 2 and 5. Nucleotides within the junctions (J) are colored accordingly. The stabilizing elements are indicated: four-base platform (boxed in red), cross-strand stacking (black dashes), ribose-zipper (zig-zag). (**B**) The central four-base platform at the 3-way junction III$_A$ between helices 2 and 5. Coloring as in (A). (**C**) Close-up of helix t1 formed by the loop–loop pseudoknot interaction of loops L3 and L4. (**D**) The two A-minor motifs (type I) of A43 and A53 important for L4.1 and L4.2 lobe formation. (**E**) Two ribose-zipper motifs and additional non-base pair interactions shaping loop L4. (**F**) The U-turn motif in loop L3 generating a strong bending of the RNA backbone. Figure 4C and E are colored in a rainbow for clarity, other color codes are as in Figure 2B.

tion, however attempts to solve the structure by molecular replacement in combination with the human Alu RNA 5′ region (PDB code 1E8O) failed. Therefore, crystals were soaked with various heavy metal ions including cobalt (III) hexamine, europium, samarium, and cesium and several single-wavelength anomalous dispersion (SAD) and multi-wavelength anomalous dispersion (MAD) datasets were collected. An interpretable electron density map could be calculated from a highly redundant SAD dataset collected from a cesium derivative of Alu103 at 3.5 Å resolution. The initial model was used as a search model in MR-SAD to obtain the phases for a MAD dataset collected from a cobalt (III) hexamine-soaked Alu107 crystal that diffracted to the substantially higher resolution of 2.5 Å. As Alu103 crystallized in a different space group, its structure was also built at a final resolution of 3.1 Å. The two structures could be refined to an excellent quality with R_{work}/R_{free} (%) of 18.0/20.9 (Alu107) and 19.1/20.8 (Alu103) (Table 1 and Supplementary Figure S1). In our folding trials of Alu domain variants we observed a strong dependence on magnesium ions. Magnesium could be substituted by the trivalent cobalt hexamine, which exerts the same effects at ∼10-fold lower concentration (Supplementary Figure S2). The metal ions stabilize the tertiary fold mainly by bridging adjacent

helices or loops underlining their crucial role in folding the Alu RNA.

B. subtilis Alu domain adopts a compact fold

The *B. subtilis* Alu domain consists of five RNA helices with helices 1, 2 and 4 being continuously stacked (Figure 2A and B). Helices 5 and 3 are connected via two 3-way junctions with helices 1 and 2 (junction III$_A$) and helices 2 and 4 (III$_B$), respectively. Junction III$_A$ is specific to prokaryotes due to the absence of helix 1 in eukaryotic SRP RNA and can be assigned to the family A of 3-way junctions, so far only described for rRNA (53). As in case of the human Alu domain, junction III$_B$ belongs to the more widespread family C of 3-way junctions, but does not contain a U-turn motif (see below). The closing loops of helices 3 and 4 interact in an extended loop–loop pseudoknot forming an additional helix (t1) of five consecutive Watson–Crick base pairs. The structure comprises both the 5′ and 3′ regions that adopt a compact 'closed' conformation with the 5′ region folded back onto the 3′ region. This closed conformation is established by perpendicular packing of helix 3 on helix 5 via minor groove interactions. In case of the human Alu domain, the complete structure is not available, as this
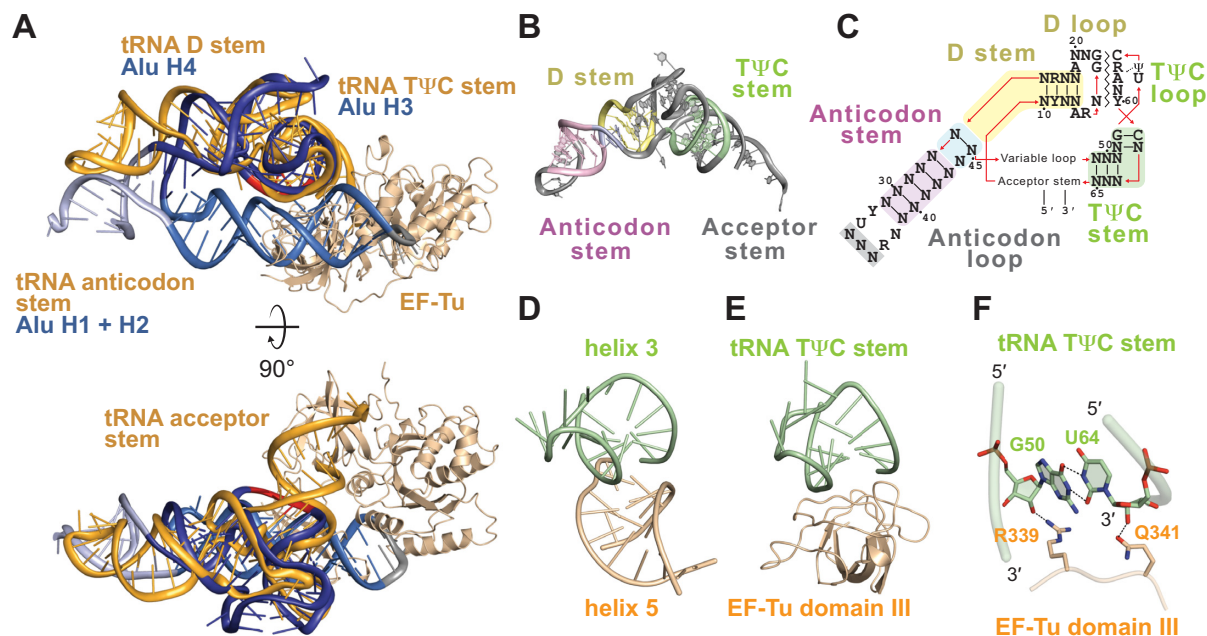
Figure 5. The Alu domain resembles tRNA. (**A**) Superposition of bacterial Alu RNA (split into 5′ and 3′ parts and helix 1) with the tRNA/EF-Tu complex (PDB entry 2WRN). The 5′ domain and helix 1 match the tRNA structure, whereas parts of the 3′ domain superpose with EF-Tu. (**B**) Structure of threonine tRNA (PDB entry 2WRN) with domains structurally similar to Alu RNA colored as in Figure 2B. The first two nucleotides of the variable loop and nucleotide 26 would correspond to Alu RNA helix 2 (light blue). The UGU sequence (red) shortcuts the tRNA acceptor stem. (**C**) Two-dimensional representation of tRNA highlighting the structural similarity to Alu RNA (color code as in Figure 2B and D). (**D**) The Alu domain-specific MSM between helices 3 and 5. (**E**) The similar interaction of the tRNA TΨC-stem with EF-Tu domain III. (**F**) Detail of the tRNA/EF-Tu interaction highlighting the corresponding read-out of a G-U wobble base pair.

conformation could not be crystallized. It was however predicted based on biochemical data and inferred from crystal packing (10,54,55) (Figure 2C and D). In general, the 5′ region consists of helices 2 to 4 and the 3′ region of helix 5. The additional helix 1 including the 5′ and 3′ termini is specific to bacterial and archaeal Alu domains (Supplementary Figure S3A). In our 3.1 Å structure solved in a different space group, helix 1 is tilted due to different crystal packing at the 5′, 3′ end revealing some plasticity of the fold (Supplementary Figure S1C). Taken together, our structure of *B. subtilis* Alu RNA is the first complete structure of an Alu domain and provides the basis to understand the function and evolution of Alu domains in general.

The Alu domain core is conserved

Comparison of bacterial and human Alu domain structures shows a common core formed by helices 2 to 5 (Figure 2C and D), and sequence alignment of representative Alu domains of all three kingdoms of life confirms that these regions form a common framework (Supplementary Figure S3A and S3B). The second 3-way junction III$_B$ (also denoted as τ-junction (10)) is conserved in all Alu domains and forms the hub of the core connecting helices 2, 3 and 4 (Figure 2D). The most highly conserved part of the Alu domain core is defined by the 'UGU' sequence (U34-G35-U36), which is part of the connecting loop between helices 3 and 4 (Figure 2B and D), and even has been retained in SRP RNA of certain fungi that lack the two helices. In the human Alu domain, the 'UGU' sequence constitutes the central binding site for the SRP9/14 proteins and has

been characterized by a U-turn motif (Figure 2C and Supplementary Figure S3C). However, in *B. subtilis* the 'UGU' sequence does not form a U-turn and bending of the loop is achieved by intra-strand base pairing of G35 and A37 forming a sheared non-Watson–Crick base pair (Figure 2E and Supplementary Figure S3B). The bridging U36 is the only bulged-out base compared to four unpaired bases in the human Alu domain (Figure 2F). Compared to eukaryotes and archaea, most bacteria lack one nucleotide that is inserted between the two U-turn forming nucleotides and therefore also the U-turn (Supplementary Figure S3A and S3B). In summary, the Alu domain folds into a compact structure with a conserved core. Although the 'UGU' sequence is present in all kingdoms of life, its structure is not strictly conserved. Archaeal SRP lacks the SRP9/14 homologs, however, the insertion is retained and therefore probably also the U-turn.

The minor-saddle motif

One major determinant of the closed conformation of the Alu RNA is the RNA–RNA tertiary contact (interface of 485 Å^2) between helices 3 and 5 connecting the 5′ and 3′ regions (Figure 3 and Supplementary Figure S4). This contact is conserved between bacterial, mammalian and probably also archaeal SRP. Four non-Watson–Crick base pairs in helix 5 (C73/C256, G74/A255, A75/G254 and G77/U252) remodel the minor groove geometry to a flat saddle-like appearance (Figure 3A–C and Supplementary Figure S4A) establishing a unique interaction surface with the perpendicular oriented minor groove of helix 3 (in the
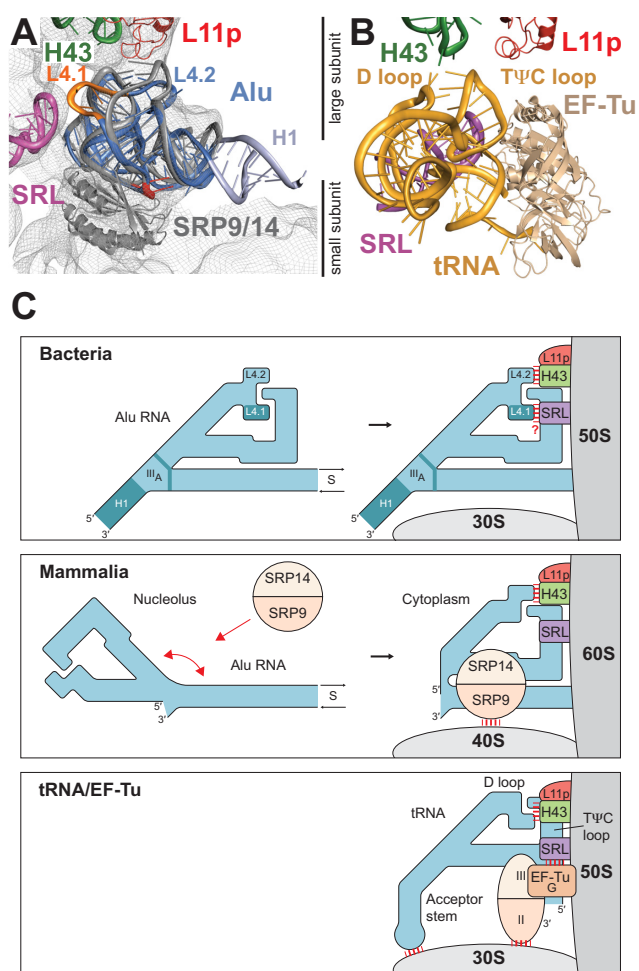
Figure 6. Ribosome binding of the Alu domain. (**A**) Superposition of *Bacillus subtilis* Alu RNA (blue) with the human Alu domain (10) (gray; combined model based on PDB entries 1E8O and 1E8S) fitted into the cryo-EM density of ribosome-bound mammalian SRP (13) (EMDB entry 1217). The bacterial L4.1 lobe approaching the α-sarcin-ricin loop (SRL) or H43 in the stalk-base is highlighted in orange and the U-turn in human Alu RNA binding to SRP9/14 is shown in red. Helix 1 points to the ribosomal inter-subunit space. For the ribosome only the stalk-base (H43 and L11p) and the SRL are shown for clarity. (**B**) Binding of the tRNA/EF-Tu initiation complex to the large ribosomal subunit (57) (PDB entries 2WRN, 2WRO). (**C**) Top: Scheme for the bacterial Alu domain. Prokaryote-specific Alu RNA features (helix 1, the junction III_A including the four-base platform, and lobe 4.1) are shown in darker blue. Middle: Scheme for the mammalian Alu domain. SRP9/14 induced folding into the closed conformation occurring in the nucleolus is indicated by red arrows. Bottom: Scheme for tRNA/EF-Tu (split in domains G, II and III).

following denoted minor-saddle motif (MSM)). While flattening of the minor groove reduces the groove depth up to 1.5 Å, respective cross-strand purine stacking between the G74/A255 and A75/G254 base pairs (Figure 3D) also leads to a decrease of the minor groove width of up to 4 Å (Figure 3E) concomitant with an increased base inclination, base-pair opening toward the minor groove, and helical twist changes. The negative twist induces a helical depression, the characteristic feature of the MSM, which allows the phospho-ribose backbone of helix 5 to align with helix 3 and to form an extensive hydrogen-bonding net-

work between bases, riboses and phosphates (Supplementary Figure S4B–G). Ribose-zippers on both sides and exclusive non-Watson–Crick base pairing with two diametric opposed G-U wobble pairs form the outer rim of the saddle. While Alu107 migrates faster and nearly as a single species in native gel electrophoresis after annealing, only a small portion of a variant with helix 5 shortened beyond the contact interface (Alu87) shows this migration behavior after folding under the same conditions. As most of Alu87 migrates at the same velocity before and after annealing, the MSM seems to be essential for proper folding of the Alu domain (Figure 1D). A 'similar' minor groove interaction can be found between helices H2 and H25 of 23S rRNA (Figure 3F and Supplementary Figure S4H and I) (56). However, in this case, the helix geometry is largely maintained and not twisted, thus only allowing for an interaction in which the minor grooves are offset with respect to each other. In summary, the occurrence of a series of conserved non-Watson–Crick base pairs and cross-strand purine stacks results in a shape complementarity of two opposing minor grooves and the formation of a saddle-like RNA–RNA interface, here denoted as the MSM. This conserved interface seems to be formed in all kingdoms of life when the 5′ and 3′ regions are locked in place.

B. subtilis **Alu RNA is stabilized by prokaryote-specific extensions**

For the human Alu domain, it has been shown that the 5′ region is flexible with respect to the 3′ region via a hinge between helices 2 and 5 and that SRP9/14 are required to induce and stabilize the closed conformation (10,11). In contrast, our structure shows that the bacterial Alu domain adopts the closed conformation in the absence of proteins. The folding into the closed conformation in the absence of proteins can thus be explained by the presence of stabilizing elements within the Alu RNA of the prokaryotic SRP. First, the coaxial stacking of helix 1 onto helix 2 reduces the flexibility between the 5′ and 3′ regions. Second, the stabilizing function of helix 1 is supported by tertiary interactions between helices 2 and 5 that lock the 3-way junction III_A in a fixed position including cross-strand stacking, a ribose/ribose zipper between A13 and C67, and a central four-base platform (Figure 4A). The platform involves two *trans* Hoogsteen/Watson–Crick base pairs A12/G62 and A64/U263 in helices 2 and 5, respectively, that are connected by a *trans* Watson–Crick base pair between the two adenines (Figure 4B). Helix 1 and the four-base platform are absent in eukaryotic SRP RNA (Figure 2D and Supplementary Figure S3A), thus allowing for a high degree of flexibility between the 5′ and 3′ regions in the absence of SRP9/14.

The most striking structural feature of the bacterial Alu domain is the extended loop–loop pseudoknot tertiary interaction (t1) between the closing loops of helices 3 and 4 (Figure 4C; see also Figure 2). While in the human Alu domain the interaction involves three Watson–Crick base pairs, in *B. subtilis* five continuous base pairs are formed. The closing loop of helix 4 contains three additional nucleotides and folds into two lobes (L4.1 and L4.2) compared to a single one corresponding to L4.2 in the hu-

man case. The lobes are tightened to the tip of helix 4 by combined A-minor (type I)/ribose-zipper motifs (Figure 4D and E), which are absent in the human Alu domain. The shape of the closing loop of helix 3 (L3) is largely conserved when compared to the human Alu domain, but its relative position is shifted to accommodate the altered shape of the L4 loop. As in the human Alu domain, a U-turn motif forms the tip of the L3 loop by generating a strong orthogonal bending of the RNA backbone (Figure 4F). The sequence comparison of these loops from *B. subtilis* with *Methanococcus jannaschii* (Supplementary Figure S3A) predicts an almost identical conformation for the archaeal Alu domain. Thus, the extended loop–loop pseudoknot seems to be preserved in prokaryotic Alu RNAs. Taken together, the bacterial Alu RNA folds into its native conformation independent of protein due to in-built stabilizing elements that are absent in eukaryotic Alu RNAs. The absence of SRP9/14 homologs in the genomes of bacteria and archaea correlates with the stable fold of the Alu RNA as shown here for *B. subtilis*.

The Alu RNA structure resembles tRNA

The Alu domain has been shown to impose elongation arrest by blocking the elongation factor entry site at the ribosome (12). Superposition of the complete *B. subtilis* Alu domain with the tRNA/EF-Tu complex (57) based on the loop–loop pseudoknot interaction shows that the 5′ domain including helix 1 resembles a tRNA structure (Figure 5A–C). Transfer RNA resemblance is of broad biological and evolutionary importance as highlighted for disease-related viral tRNA-like RNAs (58,59). Specifically, helix 4 corresponds to the D-stem, helix 3 to the TΨC-stem and helices 1 and 2 to the anticodon stem, although helix 1 is longer with respect to the canonical anticodon stem in mature tRNA (nine compared to five base pairs). In contrast, the acceptor stem is absent in the Alu RNA and shortcut by the 'UGU-loop' connecting helices 3 and 4. In addition to the structural similarity, the MSM coincides with the minor groove interaction of the tRNA TΨC-stem with EF-Tu domain III, and although the chemistry is different, hydrogen-bonding patterns are strikingly similar (Figure 5D–F).

The structural similarity mirrors the overlapping binding sites in the interface of the ribosomal subunits. The binding mode of the mammalian Alu domain has been determined by the cryo-EM structure of the complex between SRP and a stalled RNC (12). The superposition of the bacterial Alu domain on the conserved mammalian Alu domain core bound to the ribosome (Figure 6A) and the comparison with tRNA/EF-Tu ribosome-binding (Figure 6B) suggests two prokaryotic adaptations. First, the absence of the SRP9/14 heterodimer results in the detachment of prokaryotic SRP from the small ribosomal subunit. Second, the loss seems to be counterweighted by an additional RNA–RNA contact of the additional 4.1 lobe with either the so-called stalk base (H43, H44 and L11p) or the α-sarcin-ricin loop of the large subunit. In summary, the 5′ domain of the Alu RNA shares significant structural similarities with tRNAs, and in ribosome interaction, the absence of protein in prokaryotes seems compensated by Alu RNA extensions.

DISCUSSION

The biology, structure and evolution of the SRP Alu domain are still poorly understood. Our crystal structure of the complete *B. subtilis* Alu domain shows how conserved, prokaryote-specific extensions of the Alu RNA stabilize its structure, making protein components dispensable. The extensions comprise helix 1, the 3-way junction III$_A$ including the four-base platform, and the lobe 4.1, which stabilizes the extended loop–loop pseudoknot. There is a striking similarity of the common Alu domain core to the canonical tRNA structure. It is suggestive, that during evolution of the SRP RNA, a tRNA gene might have been internalized into primordial SRP, thus creating the 5′ region of the Alu RNA. Although several short interspersed elements (SINEs) could be related to tRNA sequences (60), no sequence homology between Alu RNA and tRNA has been found so far. However, as this event might have occurred early in evolution only important features of the tertiary structure might have remained conserved. Interestingly, in several protozoa such as trypanosomes, a tRNA-like RNA has been found to be associated with a significantly reduced Alu domain supporting a potential general relationship of the Alu domain with tRNA (28). The structural resemblance is clearly indicative for the functional correlation (Figure 6C), as shown for the mammalian Alu domain imposing 'elongation arrest' of translation (12), however, the biological implications still need to be addressed in detail for all prokaryotic systems.

The structural resemblance with tRNA goes even beyond the RNA itself and also concerns the interaction with the elongation factor. Within the tRNA/EF-Tu complex, the minor groove of the TΨC-stem is accommodated by domain III, an interaction perfectly mirrored by the MSM. The MSM reveals an unprecedented way of achieving shape and charge complementarity between the minor grooves of two perpendicularly aligned RNA helices. It creates a flattened surface within one helix (here helix 5) involving several non-Watson–Crick base pairs and a central block of cross-strand purine stacks (Supplementary Figure S5). This arrangement creates a twist within the phospho-ribose backbone, necessary for the establishment of ribose-zippers between the directly opposing strands. It remains to be seen, if this elegant solution is an Alu-specific invention or if it applies to other complicated RNA-folds.

ACCESSION NUMBERS

Coordinates and structure factors are deposited in the RCSB protein data bank (PDB) with the accession numbers: 4WFL and 4WFM.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGMENTS

REFERENCES

1. Nyathi,Y., Wilkinson,B.M. and Pool,M.R. (2013) Co-translational targeting and translocation of proteins to the endoplasmic reticulum. *Biochim. Biophys. Acta*, **1833**, 2392–2402.
2. Akopian,D., Shen,K., Zhang,X. and Shan,S.O. (2013) Signal recognition particle: an essential protein-targeting machine. *Annu. Rev. Biochem.*, **82**, 693–721.
3. Rosenblad,M.A., Larsen,N., Samuelsson,T. and Zwieb,C. (2009) Kinship in the SRP RNA family. *RNA Biol.*, **6**, 508–516.
4. Walter,P. and Blobel,G. (1982) Signal recognition particle contains a 7S RNA essential for protein translocation across the endoplasmic reticulum. *Nature*, **299**, 691–698.
5. Walter,P. and Blobel,G. (1981) Translocation of proteins across the endoplasmic reticulum III. Signal recognition protein (SRP) causes signal sequence-dependent and site-specific arrest of chain elongation that is released by microsomal membranes. *J. Cell Biol.*, **91**, 557–561.
6. Siegel,V. and Walter,P. (1986) Removal of the Alu structural domain from signal recognition particle leaves its protein translocation activity intact. *Nature*, **320**, 81–84.
7. Wolin,S.L. and Walter,P. (1989) Signal recognition particle mediates a transient elongation arrest of preprolactin in reticulocyte lysate. *J. Cell Biol.*, **109**, 2617–2622.
8. Mason,N., Ciufo,L.F. and Brown,J.D. (2000) Elongation arrest is a physiologically important function of signal recognition particle. *EMBO J.*, **19**, 4164–4174.
9. Walter,P., Gilmore,R. and Blobel,G. (1984) Protein translocation across the endoplasmic reticulum. *Cell*, **38**, 5–8.
10. Weichenrieder,O., Wild,K., Strub,K. and Cusack,S. (2000) Structure and assembly of the Alu domain of the mammalian signal recognition particle. *Nature*, **408**, 167–173.
11. Weichenrieder,O., Stehlin,C., Kapp,U., Birse,D.E., Timmins,P.A., Strub,K. and Cusack,S. (2001) Hierarchical assembly of the Alu domain of the mammalian signal recognition particle. *RNA*, **7**, 731–740.
12. Halic,M., Becker,T., Pool,M.R., Spahn,C.M., Grassucci,R.A., Frank,J. and Beckmann,R. (2004) Structure of the signal recognition particle interacting with the elongation-arrested ribosome. *Nature*, **427**, 808–814.
13. Halic,M., Gartmann,M., Schlenker,O., Mielke,T., Pool,M.R., Sinning,I. and Beckmann,R. (2006) Signal recognition particle receptor exposes the ribosomal translocon binding site. *Science*, **312**, 745–747.
14. Ullu,E. and Tschudi,C. (1984) Alu sequences are processed 7SL RNA genes. *Nature*, **312**, 171–172.
15. Weiner,A.M., Deininger,P.L. and Efstratiadis,A. (1986) Nonviral retroposons: genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information. *Annu. Rev. Biochem.*, **55**, 631–661.
16. Kazazian,H.H. Jr (1998) Mobile elements and disease. *Curr. Opin. Genet. Dev.*, **8**, 343–350.
17. Kriegs,J.O., Churakov,G., Jurka,J., Brosius,J. and Schmitz,J. (2007) Evolutionary history of 7SL RNA-derived SINEs in Supraprimates. *Trends Genet.*, **23**, 158–161.
18. Ule,J. (2013) Alu elements: at the crossroads between disease and evolution. *Biochem. Soc. Trans.*, **41**, 1532–1535.
19. Deininger,P. (2011) Alu elements: know the SINEs. *Genome Biol.*, **12**, 236–247.
20. Zwieb,C. and Bhuiyan,S. (2010) Archaea signal recognition particle shows the way. *Archaea*, **2010**, 1–11.
21. Andersen,E.S., Rosenblad,M.A., Larsen,N., Westergaard,J.C., Burks,J., Wower,I.K., Wower,J., Gorodkin,J., Samuelsson,T. and Zwieb,C. (2006) The tmRDB and SRPDB resources. *Nucleic Acids Res.*, **34**, D163–D168.
22. Nakamura,K., Yahagi,S., Yamazaki,T. and Yamane,K. (1999) 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.
23. Yamazaki,T., Yahagi,S., Nakamura,K. and Yamane,K. (1999) Depletion of Bacillus subtilis histone-like protein, HBsu, causes defective protein translocation and induces upregulation of small cytoplasmic RNA. *Biochem. Biophys. Res. Commun.*, **258**, 211–214.
24. Micka,B. and Marahiel,M.A. (1992) The DNA-binding protein HBsu is essential for normal growth and development in Bacillus subtilis. *Biochimie*, **74**, 641–650.
25. Alonso,J.C., Weise,F. and Rojo,F. (1995) The Bacillus subtilis histone-like protein Hbsu is required for DNA resolution and DNA inversion mediated by the beta recombinase of plasmid pSM19035. *J. Biol. Chem.*, **270**, 2938–2945.
26. Kohler,P. and Marahiel,M.A. (1997) Association of the histone-like protein HBsu with the nucleoid of Bacillus subtilis. *J. Bacteriol.*, **179**, 2060–2064.
27. Macvanin,M., Edgar,R., Cui,F., Trostel,A., Zhurkin,V. and Adhya,S. (2012) Noncoding RNAs binding to the nucleoid protein HU in Escherichia coli. *J. Bacteriol.*, **194**, 6046–6055.
28. Liu,L., Ben-Shlomo,H., Xu,Y.X., Stern,M.Z., Goncharov,I., Zhang,Y. and Michaeli,S. (2003) The trypanosomatid signal recognition particle consists of two RNA molecules, a 7SL RNA homologue and a novel tRNA-like molecule. *J. Biol. Chem.*, **278**, 18271–18280.
29. Lustig,Y., Goldshmidt,H., Uliel,S. and Michaeli,S. (2005) The Trypanosoma brucei signal recognition particle lacks the Alu-domain-binding proteins: purification and functional analysis of its binding proteins by RNAi. *J. Cell Sci.*, **118**, 4551–4562.
30. Wild,K., Bange,G., Bozkurt,G., Segnitz,B., Hendricks,A. and Sinning,I. (2010) Structural insights into the assembly of the human and archaeal signal recognition particles. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 295–303.
31. Keel,A.Y., Rambo,R.P., Batey,R.T. and Kieft,J.S. (2007) A general strategy to solve the phase problem in RNA crystallography. *Structure*, **15**, 761–772.
32. Kabsch,W. (2010) Xds. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 125–132.
33. Potterton,E., Briggs,P., Turkenburg,M. and Dodson,E. (2003) A graphical user interface to the CCP4 program suite. *Acta Crystallogr. D Biol. Crystallogr.*, **59**, 1131–1137.
34. Evans,P. (2006) Scaling and assessment of data quality. *Acta Crystallogr. D Biol. Crystallogr.*, **62**, 72–82.
35. Terwilliger,T.C., Adams,P.D., Read,R.J., McCoy,A.J., Moriarty,N.W., Grosse-Kunstleve,R.W., Afonine,P.V., Zwart,P.H. and Hung,L.W. (2009) Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. *Acta Crystallogr. D Biol. Crystallogr.*, **65**, 582–601.
36. Adams,P.D., Afonine,P.V., Bunkoczi,G., Chen,V.B., Davis,I.W., Echols,N., Headd,J.J., Hung,L.W., Kapral,G.J., Grosse-Kunstleve,R.W. et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 213–221.
37. McCoy,A.J., Grosse-Kunstleve,R.W., Adams,P.D., Winn,M.D., Storoni,L.C. and Read,R.J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.*, **40**, 658–674.
38. Vagin,A. and Teplyakov,A. (2010) Molecular replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 22–25.
39. Afonine,P.V., Grosse-Kunstleve,R.W., Echols,N., Headd,J.J., Moriarty,N.W., Mustyakimov,M., Terwilliger,T.C., Urzhumtsev,A., Zwart,P.H. and Adams,P.D. (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.*, **68**, 352–367.
40. Emsley,P., Lohkamp,B., Scott,W.G. and Cowtan,K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 486–501.
41. Chou,F.C., Sripakdeevong,P., Dibrov,S.M., Hermann,T. and Das,R. (2013) Correcting pervasive errors in RNA crystallography through enumerative structure prediction. *Nat. Methods*, **10**, 74–76.

42. Chen,V.B., Arendall,W.B. 3rd, Headd,J.J., Keedy,D.A., Immormino,R.M., Kapral,G.J., Murray,L.W., Richardson,J.S. and Richardson,D.C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 12–21.

43. Vaguine,A.A., Richelle,J. and Wodak,S.J. (1999) SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. *Acta Crystallogr. D Biol. Crystallogr.*, **55**, 191–205.

44. Krissinel,E. and Henrick,K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.*, **372**, 774–797.

45. Yang,H., Jossinet,F., Leontis,N., Chen,L., Westbrook,J., Berman,H. and Westhof,E. (2003) Tools for the automatic identification and classification of RNA base pairs. *Nucleic Acids Res.*, **31**, 3450–3460.

46. Leontis,N.B. and Westhof,E. (2001) Geometric nomenclature and classification of RNA base pairs. *RNA*, **7**, 499–512.

47. Lu,X.J. and Olson,W.K. (2003) 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res.*, **31**, 5108–5121.

48. Lu,X.J., Olson,W.K. and Bussemaker,H.J. (2010) The RNA backbone plays a crucial role in mediating the intrinsic stability of the GpU dinucleotide platform and the GpUpA/GpA miniduplex. *Nucleic Acids Res.*, **38**, 4868–4876.

49. Lavery,R., Moakher,M., Maddocks,J.H., Petkeviciute,D. and Zakrzewska,K. (2009) Conformational analysis of nucleic acids revisited: Curves+. *Nucleic Acids Res.*, **37**, 5917–5929.

50. Schrodinger,L. (2010) The PyMOL molecular graphics system, Version 1.5.0.4.

51. Will,S., Joshi,T., Hofacker,I.L., Stadler,P.F. and Backofen,R. (2012) LocARNA-P: accurate boundary prediction and improved detection of structural RNAs. *RNA*, **18**, 900–914.

52. Smith,C., Heyne,S., Richter,A.S., Will,S. and Backofen,R. (2010) Freiburg RNA Tools: a web server integrating INTARNA, EXPARNA and LOCARNA. *Nucleic Acids Res.*, **38**, W373–W377.

53. Lescoute,A. and Westhof,E. (2006) Topology of three-way junctions in folded RNAs. *RNA*, **12**, 83–93.

54. Strub,K., Moss,J. and Walter,P. (1991) Binding sites of the 9- and 14-kilodalton heterodimeric protein subunit of the signal recognition particle (SRP) are contained exclusively in the Alu domain of SRP RNA and contain a sequence motif that is conserved in evolution. *Mol. Cell. Biol.*, **11**, 3949–3959.

55. Siegel,V. and Walter,P. (1988) Each of the activities of signal recognition particle (SRP) is contained within a distinct domain: analysis of biochemical mutants of SRP. *Cell*, **52**, 39–49.

56. Selmer,M., Dunham,C.M., Murphy,F.V.t., Weixlbaumer,A., Petry,S., Kelley,A.C., Weir,J.R. and Ramakrishnan,V. (2006) Structure of the 70S ribosome complexed with mRNA and tRNA. *Science*, **313**, 1935–1942.

57. Schmeing,T.M., Voorhees,R.M., Kelley,A.C., Gao,Y.G., Murphy,F.V.t., Weir,J.R. and Ramakrishnan,V. (2009) The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA. *Science*, **326**, 688–694.

58. Dreher,T.W. (2010) Viral tRNAs and tRNA-like structures. *Wiley Interdiscip. Rev. RNA*, **1**, 402–414.

59. Colussi,T.M., Costantino,D.A., Hammond,J.A., Ruehle,G.M., Nix,J.C. and Kieft,J.S. (2014) The structural basis of transfer RNA mimicry and conformational plasticity by a viral RNA. *Nature*, **511**, 366–369.

60. Daniels,G.R. and Deininger,P.L. (1985) Repeat sequence families derived from mammalian tRNA genes. *Nature*, **317**, 819–822.