

Box C/D guide RNAs recognize a maximum of 10 nt of substrates

Zuxiao Yang (杨祖晓)^{a,b,c,d}, Jinzhong Lin (林金钟)^b, and Keqiong Ye (叶克穷)^{c,d,e,1}

^aCollege of Biological Sciences, China Agricultural University, Beijing 100193, China; ^bNational Institute of Biological Sciences, Beijing 102206, China; ^cKey Laboratory of RNA Biology, CAS Center for Excellence in Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China; ^dBeijing Key Laboratory of Noncoding RNA, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China; and ^eUniversity of Chinese Academy of Sciences, Beijing 100049, China

Edited by Denis L. J. Lafontaine, Université Libre de Bruxelles, Brussels, Belgium, and accepted by Editorial Board Member Dinshaw J. Patel August 2, 2016 (received for review March 24, 2016)

Box C/D RNAs guide site-specific 2'-O-methylation of RNAs in archaea and eukaryotes. The spacer regions between boxes C to D' and boxes C' to D contain the guide sequence that can form a stretch of base pairs with substrate RNAs. The lengths of spacer regions and guide-substrate duplexes are variable among C/D RNAs. In a previously determined structure of C/D ribonucleoprotein (RNP), a 12-nt-long spacer forms 10 bp with the substrate. How spacers and guide-substrate duplexes of other lengths are accommodated remains unknown. Here we analyze how the lengths of spacers and guide-substrate duplexes affect the modification activity and determine three structures of C/D RNPs assembled with different spacers and substrates. We show that the guide can only form a duplex of a maximum of 10 bp with the substrate during modification. Slightly shorter duplexes are tolerated, but longer duplexes must be unwound to fit into a capped protein channel for modification. Spacers with <12 nucleotides are defective, mainly because they cannot load the substrate in the active conformation. For spacers with >12 nucleotides, the excessive unpaired sequences near the box C/C' side are looped out. Our results provide insight into the substrate recognition mechanism of C/D RNA and refute the RNA-swapped model for dimeric C/D RNP.

crystal structure | box C/D snoRNA | RNA modification | RNA-protein complex

Box C/D RNAs are a large family of noncoding RNAs conserved in archaea and eukaryotes (1–3). The majority of these RNAs function as guides in site-specific 2'-O-methylation of ribosomal RNAs (rRNAs) and small nuclear RNAs (4–7). A few special members, including U3 and U14, small nucleolar RNAs are required for pre-rRNA processing and ribosome assembly. One of the most abundant modifications in rRNAs, 2'-O-methylated nucleotides cluster in functionally important regions and are believed to fine-tune ribosome structure and function (8–11).

Box C/D RNAs have a bipartite structure and contain the C (RUGAUGA; R is a purine) and D (CUGA) box motifs at two ends and the related C' and D' box motifs in the internal regions. Each pair of boxes C and D combine into a characteristic kink turn (K-turn) or K-loop structure (12–15). The spacer regions located between boxes D and C' (D spacer) and between boxes D' and C (D' spacer) harbor the guide sequence. The guide sequence can potentially form 10–21 bp with complementary sequences in substrate RNAs and select the fifth nucleotide apart from box D/D' for modification (5).

To form a functional ribonucleoprotein (RNP) enzyme, each C/D RNA associates with four core proteins [Snu13, fibrillarin (Nop1 in yeast), Nop56, and Nop58] in eukaryotes (16–23) and with three proteins (L7Ae, fibrillarin, and Nop5) in archaea (24). The homologous eukaryotic Nop56 and Nop58 proteins are replaced by a single protein, Nop5, in archaea. Fibrillarin is the catalyst that transfers a methyl group from S-adenosylmethionine to the target ribose (21, 25). The RNA-guided methylation activity has been reconstituted for archaeal C/D RNPs (24), but not yet for eukaryotic complexes (26).

Structural studies have revealed the organization of archaeal C/D RNPs. Nop5 dimerizes through its central coiled-coil domain and binds fibrillarin with its N-terminal domain (NTD) (27). The K-turn/K-loop structure of C/D RNA is sandwiched between L7Ae and the C-terminal domain (CTD) of Nop5 (28, 29). The linker between the NTD and the coiled-coil domain of Nop5 is flexible, allowing fibrillarin to access the bound substrate in a dynamic manner (27, 29–31).

We have previously determined a crystal structure of substrate-bound C/D RNP in the active conformation (30). The structure illustrates the organization of a monomeric C/D RNP (mono-RNP), in which a single bipartite C/D RNA is anchored by its two K-turn motifs onto the Nop5 dimer platform. In the structure, the spacer region of the crystallized C/D RNA, previously known as CD45, is 12 nt long and forms a 10-bp duplex with the substrate. The guide-substrate duplex is placed on the dimerized coiled-coil domains of Nop5 and capped at two ends by the CTDs of Nop5. The spacer sequence located between boxes D and C' and boxes D' and C is of variable length in C/D RNAs. In archaeal C/D RNAs, the spacer is most frequently 12 nt long (32); however, in eukaryotes, spacer length varies widely and can potentially form 10–21 bp with the substrate (5, 32, 33). This raises an important question of how C/D RNPs can accommodate differently sized spacers and guide-substrate duplexes while still maintaining the specificity of modification.

In vitro reconstitution of archaeal C/D RNPs also yielded a dimeric RNP (di-RNP) (30, 34). The structural organization and biological relevance of di-RNP is currently a matter of debate.

Significance

Box C/D RNAs are a large family of noncoding RNAs that guide 2'-O-methylation of RNAs. These RNAs associate with three or four proteins into C/D ribonucleoproteins (RNPs). The guide region of C/D RNAs is variable in length, particularly in eukaryotes, and by prediction, it can form 10–21 bp with substrates. Crystallographic and biochemical analyses revealed that the guide recognizes only a maximum of 10 nt in a substrate. Longer guide-substrate duplexes need to be unwound to fit into a size-limiting protein channel for modification. Our study reveals an aspect of the substrate recognition mechanism of C/D RNA. This mechanism is incompatible with the RNA-swapped model for dimeric C/D RNP.

Author contributions: Z.Y., J.L., and K.Y. designed research; Z.Y. and J.L. performed research; Z.Y., J.L., and K.Y. analyzed data; and K.Y. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. D.L.J.L. is a Guest Editor invited by the Editorial Board.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 5GIN, 5GIO, and 5GIP).

¹To whom correspondence should be addressed. Email: yekeqiong@ibp.ac.cn.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1604872113/-DCSupplemental.

C/D RNA has proposed to be swapped between two different protein complexes in the di-RNP (34–38), rather than associating with the same Nop5 dimer as shown in the mono-RNP structure (30). It also has been argued that the RNA-swapped di-RNP model is advantageous for accommodating long guide-substrate duplexes compared with the mono-RNP model.

In this study, we investigated the mechanism by which C/D RNAs recognize differently sized spacers and guide-substrate duplexes. We varied the length of spacers and guide-substrate duplexes and examined how the modification activity of C/D RNP was affected. We also determined three additional structures of C/D RNPs assembled with different spacers and substrates. Contrary to what was previously thought, we found that C/D RNAs recognize only a maximum of 10 nt of a substrate. The protein channel that accommodates the guide-substrate duplex is rather rigid and restricts the size of the duplex.

Results

C/D RNP Activities Are Affected by the Length of Spacers and Guide-Substrate Duplexes. We first asked how the modification activity of C/D RNP is affected by the length of spacers and guide-substrate duplexes. We varied the spacer length by deleting or inserting nucleotides at the box C/C' side based on a two-piece C/D RNA backbone (Fig. 1A and B). A series of C/D RNAs was constructed with spacer lengths (SL) ranging from 10–15 nt. For clarity, the constructed C/D RNAs are named based on their spacer length (SL10–SL15), and the CD45 RNA is renamed SL12. These C/D RNAs were assembled with purified recombinant proteins Nop5, fibrillarin, and L7Ae from *Sulfolobus solfataricus* (24, 30). The modification activities of C/D RNPs were measured on cognate substrates at multiple-turnover conditions (30-fold excess of substrate over enzyme) by monitoring incorporation of ³H-methyl.

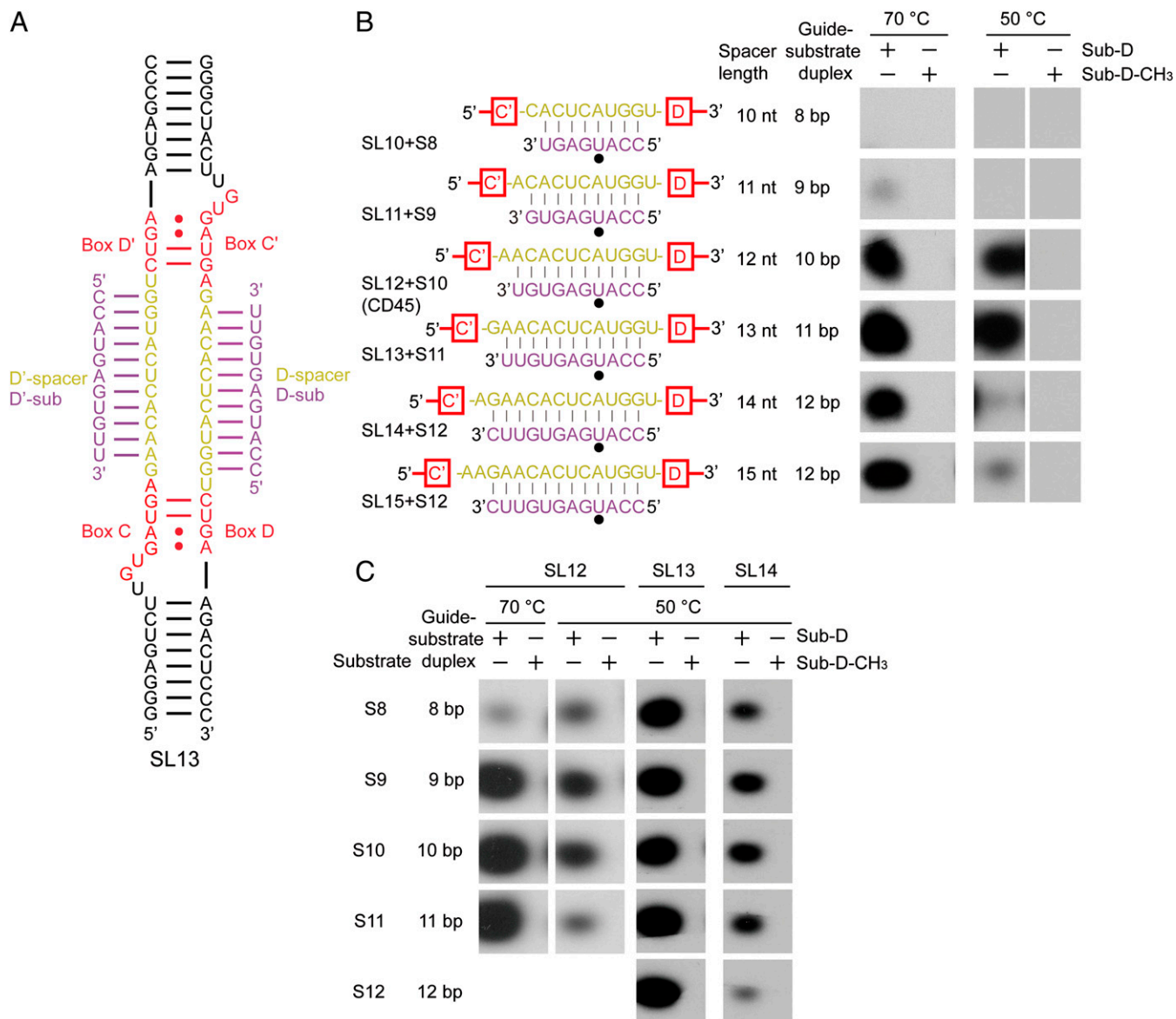


Fig. 1. C/D RNP activities depend on the length of spacers and guide-substrate duplexes. (A) Secondary structure of the two-piece C/D RNA SL13. Other C/D RNAs differ only in the spacer regions. Boxes C, D, C', and D' are shown in red, the spacer regions are in yellow, and the bound substrates are in purple. (B and C) Activity of C/D RNPs. C/D RNPs were reconstituted with guide RNAs SL10–SL15 (1 μM) and incubated with substrates S8–S12 or their premethylated versions (30 μM), [methyl-³H] SAM, and cold SAM (30 μM) at 70 °C or 50 °C for 20 min. RNAs were separated by denaturing PAGE and visualized by ³H autoradiography. Base pairing interactions between the guide and the substrate are displayed. The target site is marked by a black circle.

The SL12 RNA was highly active in guiding modification of a 10-nt substrate S10 (Fig. 1*B*), as shown previously (30). No modification was observed for the substrate premethylated at the target site, indicating that the modification was site-specific. The activity was significantly reduced when the spacer and the substrate were both shortened by 1 nt (SL11 + S9), and was completely abolished when they were further shortened by 1 nt (SL10 + S8). The loss of activity likely occurred because the shorter substrates were unable to form a thermodynamically stable duplex with the shorter spacers, or because the associated substrates cannot adopt an active conformation to be modified. To distinguish the two possibilities, we measured the modification of substrates S8 and S9 by C/D RNPs with a spacer length of 12–14 nt (Fig. 1*C*). Significant activities were detected, indicating that a guide-substrate duplex as short as 8 bp is tolerated. Nevertheless, the S8 substrate was less modified than longer substrates, particularly at high temperature of 70 °C (Fig. 1*C*), as expected for reduced binding of short substrates. In addition, the guiding activity of SL10 and SL11 RNAs was not restored at a lower temperature of 50 °C (Fig. 1*B*). We can conclude that short spacers (<12 nt long) are defective primarily because the substrate cannot be loaded into a modification-competent conformation.

Modification activity was not noticeably affected at 70 °C and 50 °C when the spacer was lengthened to 13 nt and the guide-substrate duplex was lengthened to 11 bp (Fig. 1*B*) (SL13 + S11). Further increases of the spacer length to 14 nt and 15 nt and of the substrate length to 12 nt slightly reduced the modification at 70 °C, but dramatically decreased the modification at 50 °C (Fig. 1*B*) (SL14 + S12, SL15 + S12). The data show that extensive guide-substrate pairing (12 bp in this case) inhibits the

modification in a temperature-dependent manner. An archaeal C/D RNA with spacers shorter or longer than 12 nt was previously shown to be defective in guiding modification (32). We confirmed that short spacers are not functional, but did not find that long spacers (up to 15 nt) are defective.

Structures of SL12 RNP Bound with Different Substrates. To provide structural insight into the guiding mechanism of C/D RNA, we determined two crystal structures of SL12 RNP bound to a 9-nt substrate at 3.3-Å resolution and a 13-nt substrate at 3.6-Å resolution (Fig. 2 and Table S1). The two structures were crystallized in the same space group as the previously determined structure of SL12 RNP bound to 10-nt substrates. The three SL12 RNP structures bound with different substrates are nearly identical, with an rmsd of 0.195–0.324 Å (Fig. 2*A* and *B*). In the previous structure, the 10-nt substrate pairs with the spacer at positions 2–11 (counted from box D/D') (Fig. 2*C*).

The 9-nt substrate lacks one nucleotide at the 3' end, resulting in the absence of 1 bp at position 11 (Fig. 2*D* and Fig. S1*A*). Nevertheless, the missing base pair causes no structural change in the remainder of the complex, including the now-unpaired guide nucleotide at position 11. The structure illustrates how a 9-bp guide-substrate duplex is accommodated in C/D RNP and accounts for the efficient modification of 9-nt substrates by SL12 RNP (Fig. 1*C*).

The 13-nt substrate contains three additional nucleotides at the 5' end (Fig. 2*E*). The extension is largely disordered in the structure, with only a phosphate group visible (Fig. S1*B*). The extension was designed to pair with the spacer at position 1; however, this

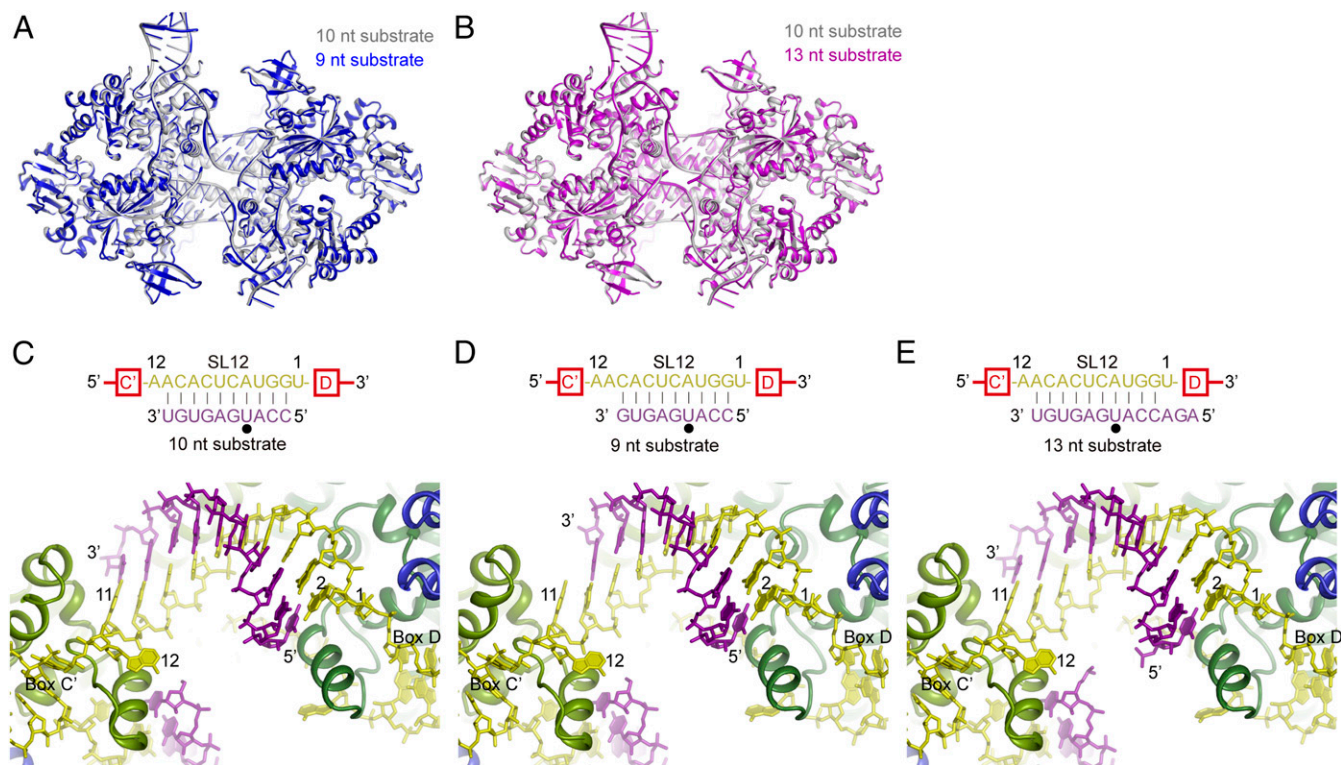


Fig. 2. Structures of substrate-bound SL12 C/D RNPs. (A) Alignment of SL12 RNP structures bound to 10-nt (silver) and 9-nt (blue) substrates. (B) Alignment of SL12 RNP structures bound to 10-nt (silver) and 13-nt (magenta) substrates. (C–E) Conformation of guide-substrate duplexes in SL12 RNP structures bound to 10-nt (C), 9-nt (D), and 13-nt (E) substrates. These structures share the same orientation. Proteins are shown as ribbons; RNAs, as sticks. Fibrillarin is omitted for clarity. The two Nop5 subunits are colored in dark green and light green, L7Ae is in blue, C/D RNA is in yellow, and the substrates are in purple. Base pairing interactions between the guide and the substrate are displayed at the top. The target site is marked by a black circle. Nucleotides in spacers and substrates are numbered by their distances to box D.

base pair was not formed in the structure, confirming that the first nucleotide in the spacer is not involved in binding substrate.

Structure of SL13 RNP. We further determined the crystal structure of SL13 RNP bound with the 11-nt substrate S11 at 3.1-Å resolution (Fig. 3A–C). The overall structure of SL13 RNP is highly similar to that of the SL12 RNP structure despite the different spacers and substrates used. The two structures can be aligned with an rmsd of 0.976 Å over 1,441 C α atom pairs (Fig. 3A). Of note, the SL12 and SL13 RNPs are crystallized in different space groups ($P4_12_1$ vs. $P1$), excluding the possibility that their structural similarity is imposed by the same crystal packing environment.

The spacer region of SL13 RNA is 13 nt long and is expected to form a 11-bp duplex with the S11 substrate (Fig. 3B); however, in the S11–SL13 RNP structure, the substrate forms only a 10-bp duplex with the guide at positions 2–11. The 3' end base of S11 is unpaired and sandwiched between its 5' base and the imidazole ring of Nop5 His327 (Fig. 3C). The nucleotide at position 12 of the spacer is also unpaired and adopts a similar conformation as its counterpart in the SL12 RNP structure. The extra spacer nucleotide at position 13 is looped out and stacks on the guanidine group of Nop5 Arg313. The structure shows that C/D RNP is unable to accept an 11-bp guide-substrate duplex. For even longer spacers, the structure also suggests that additional spacer sequences between position 12 and box C'/C can be looped out in a similar manner as for nucleotide 13.

The size of the RNA-binding channel is preserved as well. For example, the C α –C α distance between His327 and Phe325 from the opposing Nop5 subunits, which are two aromatic residues capping the two ends of the guide-substrate duplex, is 37.1 ± 0 Å for the SL13 RNP structure and 38.0 ± 0.2 Å for the SL12 RNP structure (Fig. 3B).

Discussion

In principle, the CTDs of the Nop5 dimer could adjust their positions somewhat to accommodate the guide-substrate duplex

of various sizes. Our structural and biochemical data reveal a different scenario, in which the substrate-binding channel formed by the Nop5 dimer is rather rigid and exactly fits a 10-bp duplex. Slightly shorter duplexes (8–9 bp) are tolerated as long as they are thermodynamically stable. Longer duplexes must be partially unwound at the C/C' side to fit into the size-limiting channel. Consequently, fibrillar can approach the guide-substrate duplex in the same manner, ensuring the efficiency and specificity of modification. The substrate recognition mechanism of C/D RNA explains that excessive guide-substrate pairing inhibits the modification in a temperature-dependent manner (Fig. 1B), because duplex unwinding slows at low temperatures.

Archaeal C/D RNAs have a compact structure, with the spacer region length constrained to 12 nt (32). Recently, 489 C/D RNAs were compiled from the RNA-Seq data from seven archaea, and 719 guide–rRNA interactions were predicted (39). Analysis of this large dataset showed that 57.8% of the spacers are 12 nt long (Fig. 4A), confirming a previous report (32). Only 14.2% of spacers are <12 nt long; however, these short spacers have significantly lower percentages (20.1–36%) of predicted rRNA targets compared with the spacers of ≥ 12 nt (46.7–70%). This is consistent with the finding that short spacers are defective in guiding modification (Fig. 1B). The predicted guide-substrate duplexes terminate mainly at positions 10–12 relative to box D or D' (Fig. 4B). A guide-substrate duplex that ends at or before position 11 can be directly loaded onto the substrate-binding channel and modified efficiently. Although longer spacers are functional (Fig. 1), they potentially can make extra base pairs with the substrate and are not optimal. Archaeal C/D RNAs appear to be under strong evolutionary pressure to maintain high efficiency in modification by restraining the spacer length to 12 nt.

Although our data are based on archaeal C/D RNPs, recognition of a maximum of 10 nt in substrates is likely a general principle that holds for eukaryotic C/D RNPs as well. The size of

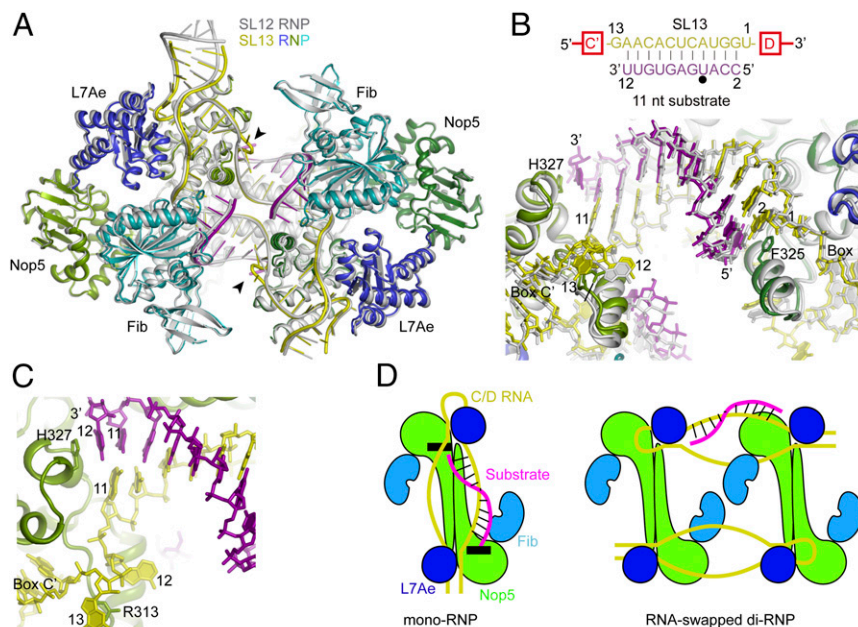


Fig. 3. Structure of SL13 RNP bound to 11-nt substrates. (A) Structural alignment of SL13 RNP bound with 11-nt substrates and SL12 RNP bound with 10-nt substrates. The SL12 RNP structure is colored in silver. The SL13 RNP structure is colored by molecule; the two Nop5 subunits are in dark green and light green, L7Ae is in blue, fibrillar (Fib) is in cyan, C/D RNA is in yellow, and substrate RNA is in purple. The looped-out guide nucleotide 13 is marked with an arrowhead. (B) Conformation of guide-substrate duplexes in the aligned SL12 RNP and SL13 RNP structures. Fibrillar is omitted for clarity. Proteins are shown as ribbons; RNAs, as sticks. Base pairing interactions between the guide and the substrate are displayed at the top. The target site is marked by a black circle. Nucleotides in spacers and substrates are numbered by their distances to box D. (C) Close-up view of the SL13 RNP structure at the box C' side. (D) Cartoons of C/D mono-RNP and RNA-swapped di-RNP. Black boxes indicate the structural elements that cap the guide-substrate duplex in the mono-RNP model.

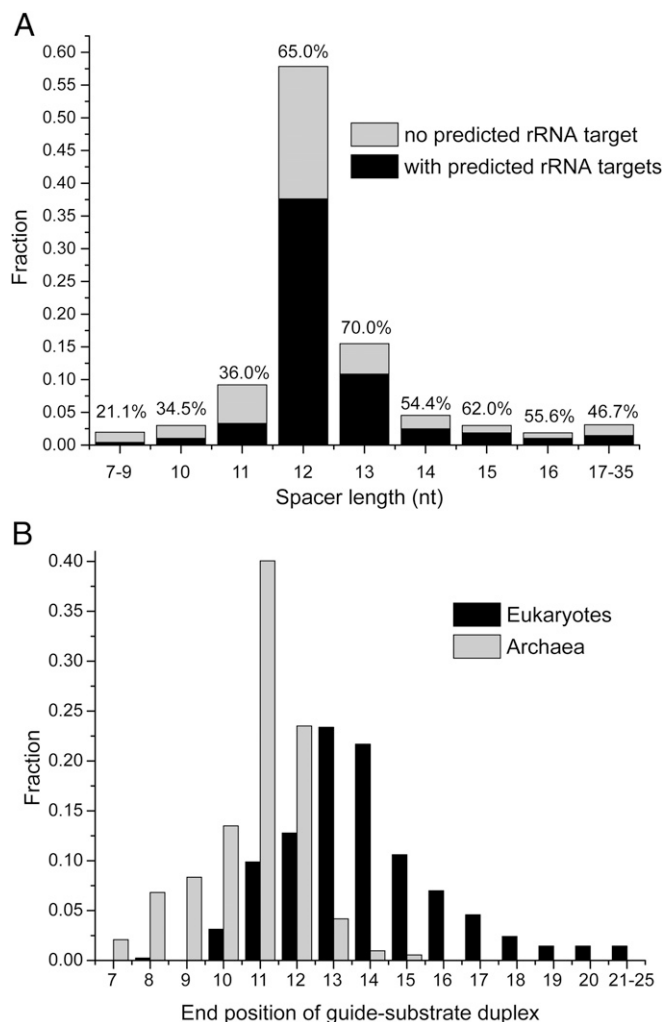


Fig. 4. Spacers and guide-substrate duplexes in archaeal and eukaryotic C/D RNAs. (A) Distribution of spacer length in archaeal C/D RNAs. A total of 968 spacer sequences from 489 archaeal C/D RNAs were counted (39). Ten spacers could be determined owing to missing box C' or/and D'. The percentage of spacers with predicted rRNA targets is indicated for each length group. (B) Distribution of end positions of guide-rRNA duplexes in archaeal and eukaryotic C/D RNAs. Distances to box D or D' are shown on the x-axis. The archaeal dataset include 719 predicted guide-rRNA duplexes from 489 archaeal C/D RNAs (39). The eukaryotic dataset includes 415 predicted guide-rRNA duplexes from *Homo sapiens*, *Mus musculus*, *Arabidopsis thaliana*, *Oryza sativa*, *Saccharomyces cerevisiae*, and *Schizosaccharomyces pombe* (33). The end of the duplex must be followed by a mismatch that must not be followed by a minimum of 2 Watson-Crick bp or G:U bp.

the substrate-binding channel is determined by the Nop5 structure in archaeal C/D RNP. Nop5 is highly conserved in sequence with its eukaryotic counterparts Nop56 and Nop58 (30), suggesting that a similar-sized channel exists in eukaryotic complexes.

The spacer region in eukaryotic C/D RNAs is much longer and of more variable size compared with archaeal C/D RNAs. Analysis of 415 predicted guide-rRNA duplexes from six distantly related eukaryotes shows that 86.7% of them are longer than the optimal duplex that ends at position 11 (33) (Fig. 4B). These duplexes most often end at positions 13 and 14. The long guide sequence may load substrates in two steps. At the first step, it binds the substrate with full range of complementarity; however, the resulting >10-bp duplex is placed out of the substrate-binding channel and cannot be modified. After unwinding of a few terminal base pairs to position 11 by breathing or other mechanisms, the

substrate is then loaded and subsequently modified and released. The long guide sequence may be advantageous for enhancing the specificity and rate of target recognition at the first binding step. In addition, the unwinding of long guide-substrate duplexes takes more time, allowing C/D RNP to bind the rRNA transcript for longer times. We propose that the extent of guide-substrate pairing provides a mechanism to control the association time of C/D RNP with rRNA. In eukaryotes, the precursor rRNA is cotranscriptionally modified by C/D RNPs and H/ACA RNPs and assembled into preribosomes (40–43). The association of guide RNAs could suppress misfolding of the target site sequence when other sequences required for assembling native structures have not been transcribed. In this way, eukaryotic C/D RNPs regulate the folding and assembly of rRNA by sequestering the target sequence. In contrast, the primary function of archaeal C/D RNP is apparently to efficiently modify rRNA.

We demonstrate that only the base pair interactions from positions 2–11 are effective for loading the substrate to the active site. Previous mutational studies have shown that the pairing interactions at the core region are mostly important for efficient modification of the substrate, whereas the pairing interactions outside the core region are dispensable (44, 45). Moreover, the pairing at the core region has been shown to impose strong constraints on the evolution of guide sequences of eukaryotic C/D RNAs. A previous analysis of conservation in eukaryotic guide-rRNA duplexes found that the duplex region from positions 3–11 is composed of most conserved Watson–Crick pairs (33). The other paired regions are less conserved than the core region (33), although they may aid in target recognition and rRNA folding.

Our reconstituted C/D RNP for activity measurement consists of a mixture of mono-RNP, di-RNP, and other higher-order species (30). Both mono-RNP and di-RNP are active in site-specific modification (30). The inhibitory effect of long guide-substrate duplexes was observed for the mixture, indicating that the di-RNP also recognizes only 10 nt of substrate as the mono-RNP.

The di-RNP likely contains two C/D RNAs and four copies of each of Nop5, L7Ae, and fibrillarin. In the proposed RNA-swapped model (34, 35), the spacer regions of C/D RNAs are placed between two Nop5 dimers and mediate dimerization of the RNP (Fig. 3D). As a result, the guide-substrate duplex cannot be loaded onto the substrate-binding channel present in the mono-RNP structure. The RNA-swapped model is also unlikely to limit the size of the guide-substrate duplex, because the ends of the guide-substrate duplex are now attached to two freely placed Nop5 molecules. Thus, the RNA-swapped di-RNP model cannot account for the recognition of 10 nt of substrate by C/D RNA.

The di-RNP structure has been analyzed by low-resolution approaches (34, 35), but the key information about the organization of C/D RNA cannot be determined directly from these studies. Without invoking the unconventional RNA-swapped model, the di-RNP could be accounted for simply by self-association of the mono-RNP through protein–protein interaction. In this way, each C/D RNA would still be bound to an Nop5 dimer as in the mono-RNP, and the di-RNP would adopt the same mechanism of action as the mono-RNP.

Methods

Activity Assay. Substrate RNAs and premethylated substrate RNAs were purchased from Takara. The new C/D RNAs were constructed from the CD45 RNA by QuikChange mutagenesis (30). The D and D' spacer sequences were different for the SL12/CD45 RNA and the same for other C/D RNAs. The modification reactions (10 μ L) contained 10 pmol C/D RNA, 30 pmol L7Ae, 20 pmol Nop5-Fib dimer, and 300 pmol D site substrate or premethylated substrate in buffer containing 20 mM Hepes-Na (pH 8.0), 150 mM NaCl, and 1.5 mM MgCl₂. The reactions were prepared in PCR tubes and incubated at 20 °C for 10 min. To start the modification, the tubes were added with 300 pmol unlabeled S-adenosylmethionine (SAM) and 0.25 μ Ci [methyl-³H] SAM (83 Ci mmol⁻¹; PerkinElmer) and incubated at 70 °C or 50 °C for 20 min.

The reactions were stopped by freezing in liquid nitrogen. Each sample was mixed with an equal volume of 8 M urea in Tris-borate-EDTA (TBE) buffer and heated at 95 °C for 2 min. The samples were then loaded onto a 10% (wt/vol) polyacrylamide gel containing 8 M urea and then separated in TBE buffer at a constant power of 13 W for 1 h. The gel was fixed in a solution containing 10% (vol/vol) acetic acid, 25% (vol/vol) isopropanol, and 65% (vol/vol) water for 30 min, then soaked in Amplify solution (GE Healthcare) with agitation for 15–30 min and dried by vacuum at 80 °C. For autoradiography, the gel was exposed to X-ray films for 2–3 d at –80 °C.

Crystallization. Purification of Nop5, fibrillarin, and L7Ae proteins; preparation of C/D RNAs; and assembly of C/D RNPs were done as described previously (29, 30). Crystallization was performed by the hanging-drop vapor diffusion method at 20 °C. The SL12 RNP was loaded with a 1.5 molar equivalent of the 9-nt substrates (5'-UCCAGUACU-3' and 5'-CCAUGAGUG-3') or the 13-nt substrates (5'-AGAUCAGUACUU-3' and 5'-AGACCAUGAGUGU-3'). The two SL12 RNPs were crystallized from 2.0 M (NH₄)₂SO₄, 2% (wt/vol) PEG-400, 10 mM MgCl₂, and 0.1 M Hepes-Na (pH 6.0), the same condition as the previous SL12 RNP bound to the 10-nt substrates.

The SL13 RNP (4 mg/mL in 5 mM Hepes-K, pH 8.0) bound to the 11-nt substrate (5'-CCAUGAGUGUU-3') was crystallized in 1.7 M DL-malic acid (pH 7.0) and 0.1 M

potassium sodium tartrate tetrahydrate. The crystal was cryoprotected in 30% (vol/vol) glycerol created in the well solution and flash-frozen in liquid nitrogen.

Structure Determination. Diffraction data were collected at the BL17U beamline of the Shanghai Synchrotron Radiation Facility and processed with the HKL-2000 program package (46). All structures were solved by molecular replacement and refined with PHENIX (47). Models were built in COOT (48). Structural figures were prepared with PyMol (49).

The two SL12 RNP structures bound to the 9-nt and 13-nt substrates contain 1.5 copies of C/D RNP in the asymmetric unit and have the same space group of *P4₁2₁2* as the previously determined SL12 RNP structure bound to the 10-nt substrates. The 9-nt substrate-bound structure was refined to 3.3 Å, and the 13 nt substrate-bound structure was refined to 3.6 Å. The SL13 RNP structure bound with the 11-nt substrate is in the *P1* space group with two copies of C/D RNPs in the asymmetric unit and was refined to 3.1 Å.

ACKNOWLEDGMENTS. We thank the staff at beamline BL17U of the Shanghai Synchrotron Research Facility for assistance in diffraction data collection. This work was supported by the National Natural Science Foundation of China (Grants 31325007, 31430024, and 91540201), the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant XDB08010203), and the Beijing Municipal Government.

- Kiss T (2001) Small nucleolar RNA-guided post-transcriptional modification of cellular RNAs. *EMBO J* 20(14):3617–3622.
- Watkins NJ, Bohnsack MT (2012) The box C/D and H/ACA snoRNPs: Key players in the modification, processing and the dynamic folding of ribosomal RNA. *Wiley Interdiscip Rev RNA* 3(3):397–414.
- Liang B, Li H (2011) Structures of ribonucleoprotein particle modification enzymes. *Q Rev Biophys* 44(1):95–122.
- Cavaillé J, Nicoloso M, Bachelier JP (1996) Targeted ribose methylation of RNA in vivo directed by tailored antisense RNA guides. *Nature* 383(6602):732–735.
- Kiss-László Z, Henry Y, Bachelier JP, Caizergues-Ferrer M, Kiss T (1996) Site-specific ribose methylation of preribosomal RNA: A novel function for small nucleolar RNAs. *Cell* 85(7):1077–1088.
- Tycowski KT, Smith CM, Shu MD, Steitz JA (1996) A small nucleolar RNA requirement for site-specific ribose methylation of rRNA in *Xenopus*. *Proc Natl Acad Sci USA* 93(25):14480–14485.
- Omer AD, et al. (2000) Homologs of small nucleolar RNAs in Archaea. *Science* 288(5465):517–522.
- Baudin-Baillieu A, et al. (2009) Nucleotide modifications in three functionally important regions of the *Saccharomyces cerevisiae* ribosome affect translation accuracy. *Nucleic Acids Res* 37(22):7665–7677.
- Liang XH, Liu Q, Fournier MJ (2009) Loss of rRNA modifications in the decoding center of the ribosome impairs translation and strongly delays pre-rRNA processing. *RNA* 15(9):1716–1728.
- Esguerra J, Warringer J, Blomberg A (2008) Functional importance of individual rRNA 2'-O-ribose methylations revealed by high-resolution phenotyping. *RNA* 14(4):649–656.
- Decatur WA, Fournier MJ (2002) rRNA modifications and ribosome function. *Trends Biochem Sci* 27(7):344–351.
- Watkins NJ, et al. (2000) A common core RNP structure shared between the small nucleolar box C/D RNPs and the spliceosomal U4 snRNP. *Cell* 103(3):457–466.
- Klein DJ, Schmeing TM, Moore PB, Steitz TA (2001) The kink-turn: A new RNA secondary structure motif. *EMBO J* 20(15):4214–4221.
- Vidovic I, Nottrott S, Hartmuth K, Lüthmann R, Ficner R (2000) Crystal structure of the spliceosomal 15.5-kD protein bound to a U4 snRNA fragment. *Mol Cell* 6(6):1331–1342.
- Nolivos S, Carpousis AJ, Clouet-d'Orval B (2005) The K-loop, a general feature of the *Pyrococcus* C/D guide RNAs, is an RNA structural motif related to the K-turn. *Nucleic Acids Res* 33(20):6507–6514.
- Newman DR, Kuhn JF, Shanab GM, Maxwell ES (2000) Box C/D snoRNA-associated proteins: Two pairs of evolutionarily ancient proteins and possible links to replication and transcription. *RNA* 6(6):861–879.
- Lafontaine DL, Tollervey D (2000) Synthesis and assembly of the box C+D small nucleolar RNPs. *Mol Cell Biol* 20(8):2650–2659.
- Lafontaine DL, Tollervey D (1999) Nop58p is a common component of the box C+D snoRNPs that is required for snoRNA stability. *RNA* 5(3):455–467.
- Wu P, Brockenbrough JS, Metcalfe AC, Chen S, Aris JP (1998) Nop5p is a small nucleolar ribonucleoprotein component required for pre-18 S rRNA processing in yeast. *J Biol Chem* 273(26):16453–16463.
- Gautier T, Bergès T, Tollervey D, Hurt E (1997) Nucleolar KKE/D repeat proteins Nop56p and Nop58p interact with Nop1p and are required for ribosome biogenesis. *Mol Cell Biol* 17(12):7088–7098.
- Tollervey D, Lehtonen H, Jansen R, Kern H, Hurt EC (1993) Temperature-sensitive mutations demonstrate roles for yeast fibrillarin in pre-rRNA processing, pre-rRNA methylation, and ribosome assembly. *Cell* 72(3):443–457.
- Schimmang T, Tollervey D, Kern H, Frank R, Hurt EC (1989) A yeast nucleolar protein related to mammalian fibrillarin is associated with small nucleolar RNA and is essential for viability. *EMBO J* 8(13):4015–4024.
- Galardi S, et al. (2002) Purified box C/D snoRNPs are able to reproduce site-specific 2'-O-methylation of target RNA in vitro. *Mol Cell Biol* 22(19):6663–6668.
- Omer AD, Ziesche S, Ehardt H, Dennis PP (2002) In vitro reconstitution and activity of a C/D box methylation guide ribonucleoprotein complex. *Proc Natl Acad Sci USA* 99(8):5289–5294.
- Wang H, Boisvert D, Kim KK, Kim R, Kim SH (2000) Crystal structure of a fibrillarin homologue from *Methanococcus jannaschii*, a hyperthermophile, at 1.6 Å resolution. *EMBO J* 19(3):317–323.
- Peng Y, Yu G, Tian S, Li H (2014) Co-expression and co-purification of archaeal and eukaryal box C/D RNPs. *PLoS One* 9(7):e103096.
- Aittaleb M, et al. (2003) Structure and function of archaeal box C/D sRNP core proteins. *Nat Struct Biol* 10(4):256–263.
- Xue S, et al. (2010) Structural basis for substrate placement by an archaeal box C/D ribonucleoprotein particle. *Mol Cell* 39(6):939–949.
- Ye K, et al. (2009) Structural organization of box C/D RNA-guided RNA methyltransferase. *Proc Natl Acad Sci USA* 106(33):13808–13813.
- Lin J, et al. (2011) Structural basis for site-specific ribose methylation by box C/D RNA protein complexes. *Nature* 469(7331):559–563.
- Oruganti S, et al. (2007) Alternative conformations of the archaeal Nop56/58-fibrillarin complex imply flexibility in box C/D RNPs. *J Mol Biol* 371(5):1141–1150.
- Tran E, Zhang X, Lackey L, Maxwell ES (2005) Conserved spacing between the box C/D and C/D' RNPs of the archaeal box C/D sRNP complex is required for efficient 2'-O-methylation of target RNAs. *RNA* 11(3):285–293.
- Chen CL, Perasso R, Qu LH, Amar L (2007) Exploration of pairing constraints identifies a 9 base-pair core within box C/D snoRNA-rRNA duplexes. *J Mol Biol* 369(3):771–783.
- Bleichert F, et al. (2009) A dimeric structure for archaeal box C/D small ribonucleoproteins. *Science* 325(5946):1384–1387.
- Lapinaite A, et al. (2013) The structure of the box C/D enzyme reveals regulation of RNA methylation. *Nature* 502(7472):519–523.
- Graziadei A, Masiewicz P, Lapinaite A, Carlomagno T (2016) Archaeal box C/D enzymes methylate two distinct substrate rRNA sequences with different efficiency. *RNA* 22(5):764–772.
- Bower-Phipps KR, Taylor DW, Wang HW, Baserga SJ (2012) The box C/D sRNP dimeric architecture is conserved across domain Archaea. *RNA* 18(8):1527–1540.
- Bleichert F, Baserga SJ (2010) Dissecting the role of conserved box C/D sRNA sequences in di-sRNP assembly and function. *Nucleic Acids Res* 38(22):8295–8305.
- Dennis PP, Tripp V, Lui L, Lowe T, Randau L (2015) C/D box sRNA-guided 2'-O-methylation patterns of archaeal rRNA molecules. *BMC Genomics* 16:632.
- Osheim YN, et al. (2004) Pre-18S ribosomal RNA is structurally compacted into the 5SU processome prior to being cleaved from nascent transcripts in *Saccharomyces cerevisiae*. *Mol Cell* 16(6):943–954.
- Zhang L, Wu C, Cai G, Chen S, Ye K (2016) Stepwise and dynamic assembly of the earliest precursors of small ribosomal subunits in yeast. *Genes Dev* 30(6):718–732.
- Chaker-Margot M, Hunziker M, Barandun J, Dill BD, Klinge S (2015) Stage-specific assembly events of the 6-MDa small-subunit processome initiate eukaryotic ribosome biogenesis. *Nat Struct Mol Biol* 22(11):920–923.
- Kos M, Tollervey D (2010) Yeast pre-rRNA processing and modification occur co-transcriptionally. *Mol Cell* 37(6):809–820.
- Appel CD, Maxwell ES (2007) Structural features of the guide:target RNA duplex required for archaeal box C/D sRNA-guided nucleotide 2'-O-methylation. *RNA* 13(6):899–911.
- Cavaillé J, Bachelier JP (1998) SnoRNA-guided ribose methylation of rRNA: Structural features of the guide RNA duplex influencing the extent of the reaction. *Nucleic Acids Res* 26(7):1576–1587.
- Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307–326.
- Adams PD, et al. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):213–221.
- Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126–2132.
- DeLano WL (2002) *The PyMOL User's Manual* (Delano Scientific, San Carlos, CA).