The K-loop, a general feature of the *Pyrococcus* C/D guide RNAs, is an RNA structural motif related to the K-turn

Sophie Nolivos, Agamemnon J. Carpousis and Béatrice Clouet-d'Orval*

Laboratoire de Microbiologie et Génétique Moléculaires, Centre National de la Recherche Scientifique, UMR 5100 Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse cedex 9, France

ABSTRACT

The C/D guide RNAs predicted from the genomic sequences of three species of *Pyrococcus* delineate a family of small non-coding archaeal RNAs involved in the methylation of rRNA and tRNA. The C/D guides assemble into ribonucleoprotein (RNP) that contains the methyltransferase. The protein L7Ae, a key structural component of the RNP, binds to a Kink-turn (K-turn) formed by the C/D motif. The K-turn is a structure that consists of two RNA stems separated by a short asymmetric loop with a characteristic sharp bend (kink) between the two stems. The majority of the *pyrococcal* C/D guides contain a short 3 nt-spacer between the C′/D′ motifs. We show here that conserved terminal stem–loops formed by the C′/D′ motif of the *Pyrococcus* C/D RNAs are also L7Ae-binding sites. These stem–loops are related to the K-turn by sequence and structure, but they consist of a single stem closed by a terminal loop. We have named this structure the K-loop. We show that conserved non-canonical base pairs in the stem of the K-loop are necessary for L7Ae binding. For the C/D guides with a 3 nt-spacer we show that the sequence and length is also important. The K-loop could improve the stability of the C/D guide RNAs in *Pyrococcal* species, which are extreme hyperthermophiles.

INTRODUCTION

The C/D and H/ACA guide RNAs direct the site-specific 2′-*O*-methylation and pseudouridylation of ribosomal RNA in the Archaea and Eukarya, respectively (1–5). Additional substrates include snRNA in the vertebrates and tRNA in the Archaea (1,2,4,6,7). Base pairing of an antisense guide sequence to the RNA substrate targets the modifying enzyme to the site of methylation or pseudouridylation. For each guide RNA, the modifying enzyme is a component of the corresponding ribonucleoprotein (RNP). The C/D guide RNAs are defined by two conserved sequence elements: box C (RUGAUGA) and box D (CUGA) (8–10). The C/D guide RNAs often have a second set of motifs named C′ and D′ with sequences related to the C and D motifs, respectively. The C and D boxes pair to form an RNA structure that is the site of RNP assembly (5,11–15). Eukaryal C/D RNPs contain four core proteins: 15.5 kDa (or Snu13p in yeast), Nop56p, Nop58p and fibrillarin (or Nop1p), which is the methyltransferase (16–19). The distribution of these proteins between the C/D and C′/D′ motifs is apparently asymmetric (12,20). Analysis of the archaeal C/D RNPs has revealed a simpler composition and a symmetric distribution of proteins between the C/D and C′/D′ motifs (14,15,21). Archaeal C/D RNPs contain only three proteins: aFib and L7Ae, are homologs of the eukaryal fibrillarin and 15.5 kDa, respectively, whereas Nop5p is structurally related to eukaryal Nop56p and Nop58p (1). L7Ae, a member of the L30 family of ribosomal proteins, is a multifunctional protein that is a component of the archaeal ribosome, the C/D RNPs and the H/ACA RNPs (9,11).

The C/D motifs that have been characterized to date form an RNA structure known as a Kink-turn (K-turn) (9,22). The K-turn, found in all three domains of life, was initially characterized in the 5′ stem of U4 snRNA and in 23S ribosomal RNA (23–25). Defined by two helical stems, the canonical stem (C-stem) and the non-canonical stem (NC-stem) are separated by a 3 nt internal loop, which imposes a sharp bend (kink) between the two stems. The C-stem contains a variable number of Watson–Crick (WC) base pairs with no apparent conservation of sequence. For the C/D guide RNAs, the internal loop is asymmetric with a highly conserved flipped-out U and the NC-stem usually contains two tandem sheared GA base pairs, adjacent to the flipped-out U, followed by a third non-WC base pair. The K-turn is stabilized by a network

of contacts between bases and sugars in NC-stem and internal loop, and interactions between the C-stem and the NC-stem (22,24,26). The binding of L7Ae to the K-turn formed by the C/D motif is believed to be the initial step in the assembly of the C/D RNP since the other proteins do not bind to the guide RNA in their absence (14,27). Mutation of the non-WC sheared GA base pairs disrupts the interaction with the eukaryal Snu13p *in vitro* and affects guide RNA activity *in vivo* (13). Recent crystallographic studies have elucidated the structure of the K-turn in a C/D guide RNA bound to L7Ae from *Archeoglobulus fulgidus* and have confirmed previously proposed secondary structure of the box C/D RNA (28).

Although L7Ae and Snu13p bind to K-turns formed by the C/D motif, recent work suggests that their specificity differs. The crystal structure of L7Ae bound to the terminal stem–loop of an archaeal H/ACA guide RNA demonstrated an interaction with a structure related to the K-turn containing the NC-stem and the flipped-out U (29). However, this structure lacks the C-stem, which is replaced by a short terminal loop. The ability of L7Ae to recognize this structural motif accounts for its binding to the archaeal H/ACA guide RNAs, which do not contain discernable C/D or C′/D′ motifs (29–31). In contrast, Snu13p is unable to bind to K-turns lacking the C-stem (32). The structure of the terminal stem–loop of the archaeal H/ACA guide RNA has been referred to as a 'minimal' K-turn (29). However, we believe this is a misnomer since the lack of the C-stem precludes the formation of a sharp bend between two RNA stems, which is the definition of the K-turn. For this reason, we suggest here that the terminal stem–loop recognized by L7Ae be named a K-loop to clearly distinguish it from the K-turn. It has been suggested that the C′/D′ motifs of the archaeal C/D guide RNAs can form a K-loop (29). In addition L7Ae binding to an RNA stem–loop derived from an archaeal C′/D′ motif has been demonstrated *in vitro* recently (32). Here we show that K-loops formed by the C′/D′ motif are a general feature of the C/D guide RNAs predicted for three *Pyrococcus* species whose genomes have been sequenced. In the majority of these RNAs, the C′/D′ motif forms a K-loop with a conserved RNK (G/A-N-G/U) sequence in the terminal loop. We demonstrate the importance for L7Ae binding of the non-WC tandem sheared GA base pairs as well as the sequence and length of the 3 nt-spacer. Our results support the

hypothesis that the symmetric binding of the archaeal proteins to the C/D guide RNAs is due to the relaxed specificity of L7Ae, which is able to recognize K-loops formed by the C′/D′ motif. The K-loop motif could improve the stability of the C/D guide RNA in *Pyrococcus* species, which are extreme hyperthermophiles.

MATERIALS AND METHODS

Unless otherwise noted, all techniques for cloning and manipulating nucleic acids were performed according to standard protocols. The oligonucleotides used in this study are described in Table 1.

DNA templates

The sR47 wild-type and mutant DNA templates were obtained by PCR amplification using oligonucleotides 1 and 2 for each construct. The sR6 wild-type and mutant templates were obtained by hybridization of T7 promoter top strand oligonucleotide with the corresponding bottom strand oligonucleotide using a 2:1 ratio (33).

RNA synthesis

RNAs were synthesized by *in vitro* transcription of the cognate PCR template or top strand-bottom strand hybrid. Large-scale synthesis of sR47 for RNA footprinting was performed in a 100 µl-volume reaction containing 200 nM of template in 0.04 M of DTT, 40 mM Tris, pH 8.0, 14 mM $MgCl_2$, 2 mM spermidine, 40 U of RNasin (Promega), 5 mM of each rNTP (Amersham Pharmacia) and 160 U of T7 RNA polymerase (Promega). The reaction was incubated overnight at 37°C. After ethanol precipitation in presence of 14 mM EDTA, RNA was purified on an 8% denaturing gel. Dephosphorylated RNA was 5′ end-labeled with T4 polynucleotide kinase and [γ-^{32}P]ATP as described (15). Labeled RNAs were then re-purified on a 6% denaturing polyacrylamide gel. Uniformly [^{32}P]-labeled RNAs were synthesized as described (6).

RNA–protein interaction by electrophoretic mobility shift assay (EMSA)

Cloning, expression and purification of recombinant *Pyrococcus abyssi* L7Ae has been described (15). Uniformly

Table 1. Oligonucleotides used in this study

sR47-1	5′-AAGGTGAATCAGCACTCAAGATCCTCATCACTCCTCAGGACC-3′
sR47-2	5′-CCGGAATTCTAATACGACTCACTATAGATGAAGATGATGAGCTCGGCAGGTCCTGAGGAGTGATGAGG-3′
sR47 mut C/D-1	5′-AAGGTGAAGAAGCACTCAAGATCCTCATCACTCCTCAGGACC-3′
sR47 mut C/D-2	5′-CCGGAATTCTAATACGACTCACTATAGATGAAGATTCTGAGCTCGGCAGGTCCTGAGGAGTGATGAGG-3′
sR47 mut C′/D′-1	5′-AAGGTGAATCAGCACTCAAGATCCTCAGAACTCCGAAGGACC-3′
sR47 mut C′/D′-2	5′-CCGGAATTCTAATACGACTCACTATAGATGAAGATGATGAGCTCGGCAGGTCCTTCGGAGTTCTGAGG-3′
sR47 mut C/D&C′/D′-1	5′-AAGGTGAAGAAGCACTCAAGATCCTCAGAACTCCGAAGGACC-3′
sR47 mut C/D&C′/D′-2	5′-CCGGAATTCTAATACGACTCACTATAGATGAAGATTCTGAGCTCGGCAGGTCCTTCGGAGTTCTGAGG-3′
sR47 loopCCC-1	5′-AAGGTGAATCAGCACTCAAGATCCTCATCACGGGTCAGGACC-3′
sR47 loopCCC-2	5′-CCGGAATTCTAATACGACTCACTATAGATGAAGATGATGAGCTCGGCAGGTCCTGACCCGTGATGAGG-3′
T7 promoter top	5′-TAATACGACTCACTATA-3′
sR6 bottom	5′-GGGTCATCACTGCTCAGTCCCTATAGTGAGTCGTATTA-3′
sR6 bottom mut C′/D′	5′-GGGTCAGAACTGCGAAGTCCCTATAGTGAGTCGTATTA-3′
sR6 bottom L1	5′-GGGTCATCACTGATCAGTCCCTATAGTGAGTCGTATTA-3′
sR6 bottom LC	5′-GGGTCATCACGGGTCAGTCCCTATAGTGAGTCGTATTA-3′
sR6 bottom L2	5′-GGGTCATCAATGCTCAGTCCCTATAGTGAGTCGTATTA-3′

[^{32}P]-labeled transcripts were denatured and renatured as described (15). For binding, 0.25 fmol of radiolabeled RNA, mixed with 5 µg of *Escherichia coli* tRNA, was incubated with increasing amounts of L7Ae in a 10 µl reaction in buffer A (20 mM HEPES–KOH, pH 7.9, 150 mM KCl, 1.5 mM MgCl$_2$, 0.2 mM EDTA and 10% glycerol). RNP complexes were assembled by incubation at room temperature for 10 min and then at 70°C for 5 min. RNA and RNP complexes were separated on a native 8% (19:1) polyacrylamide gel containing 0.5× TBE and 5% glycerol. Electrophoresis was performed at room temperature at 250 V in 0.5× TBE running buffer containing 5% glycerol. The gels were dried and quantified using a Fuji-Bas 1000 phosphorimager. Dissociation constants were calculated from the binding isotherms of at least three independent determinations. The standard deviation ranged from 15 to 30%.

Footprint with RNase T1

End-labeled sR47 (0.1 pmol), 1 µg of *E.coli* tRNA (Sigma) and 1 µM L7Ae were incubated in a 10 µl reaction at 70°C for 10 min in buffer A (see above) and then quenched on ice. RNaseT1 (0.01 U) was added, the reaction was incubated at 50°C for 5 min and then stopped by ethanol precipitation as described (6). RNase T1 probing of free sR47 was performed in parallel under the same conditions. The RNase T1 ladder was generated by incubation of the 5′ end-labeled transcript (100 000 c.p.m.), 1 µg of tRNA and 0.1 U of RNase T1 in 20 mM sodium citrate, pH 5, 1 mM EDTA, 7 M urea for 10 min at 55°C. The hydroxyl ladder was generated by incubation of the 5′ end-labeled transcript (100 000 c.p.m.) and 1 µg of tRNA in 50 mM sodium carbonate, pH 10.3, 1 mM EDTA for 10 min at 70°C. Electrophoresis was on a 10% denaturing polyacrylamide gel.

RESULTS

The C′/D′ motifs of the *Pyrococcus* C/D guide RNAs are separated by a short spacer that precludes the formation of a K-turn

Analysis of genomic sequences relying on the presence of hallmark motifs and ribosomal RNA antisense elements has been a powerful means for detecting C/D guide RNAs in the Eukarya and the Archaea (34–38). Small RNAs, predicted from the genomic sequences of three hyperthermophilic *Pyrococcus* species, constitute a family of prototypical C/D guides [http://rna.wustl.edu/snoRNAdb/ (35,38)]. They exhibit remarkable structural homogeneity, extended consensus motifs and the quasi-systematic presence of antisense guides upstream of the D and D′ elements, which often correspond to sequences close together in the structure of the rRNA (Figure 1A). For this reason, it has been suggested that in addition to RNA methylation, these guide RNAs also act as chaperones in the folding of the rRNA (38). An alignment of the sequences of the pyrococcal C/D guide RNAs shows that 40 out of 59 (*P.abyssi*), 29 out of 55 (*Pyrococcus furiosus*) and 32 out of 54 (*Pyrococcus horikoshii*) C/D RNAs have a short 3 nt-spacer between the D′ and C′ elements (Figure 1B). Is this length constraint an artifact of the computational searches? In the search by Gaspin *et al*. (38), the overall length from the C box to the D box was limited to 90 nt. Since the average length in predicted C/D guides is ∼60 nt (Figure 1B), it should have been possible to identify guides with larger D′/C′ spacers if they exist. Furthermore, D′/C′ spacers up to 12 nt are detected. A new search without the 90 nt limit has been performed (C. Gaspin, personal communication). In this search, half C/D RNAs containing a C box followed by a D′ box or a C′ box followed by a D box were identified in an initial step. The half elements were then joined by permitting a distance between the half RNAs of up to 100 nt.
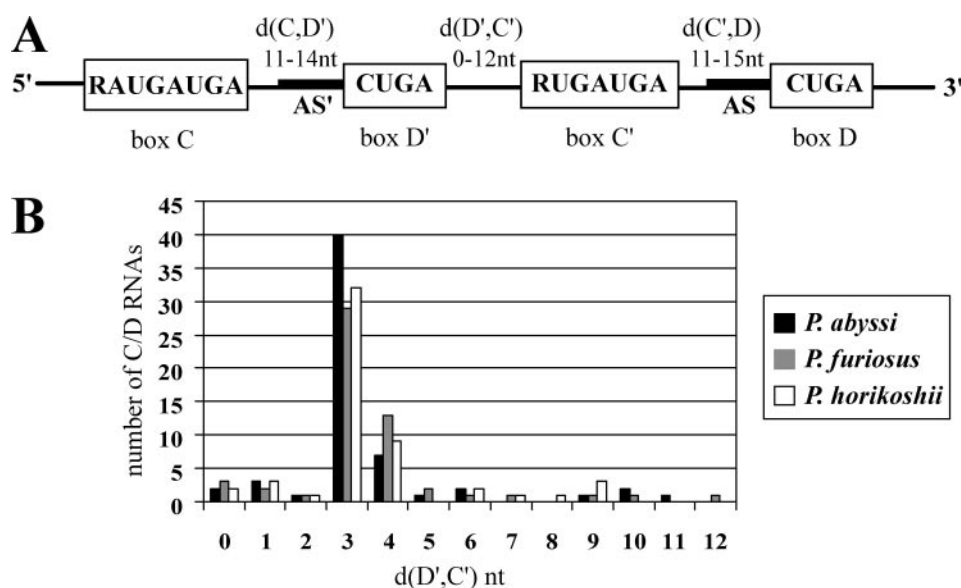

Figure 1. (**A**) Consensus features of *P.abyssi* C/D guide sRNA (38). The consensus motifs are boxed. The notations d(C,D′), d(D′,C′) and d(C′,D) indicate the distance in nucleotides between the conserved motifs. The antisense guides, AS and AS′, upstream of D and D′, respectively, are represented by the thick line. (**B**) Graph representing the number of *Pyrococcus* C/D sRNAs versus the nucleotide distance between D′ and C′.
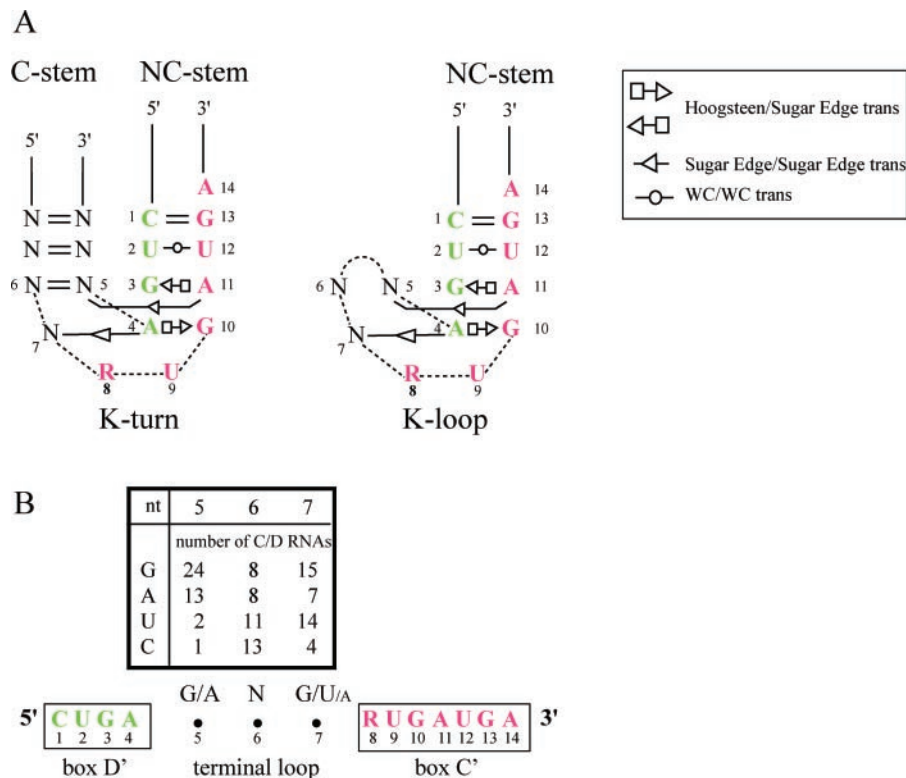
Figure 2. (**A**) K-turn modeled from the consensus C/D motif (28) with the C and D motifs shown in red and green, respectively. K-loop modeled from a consensus C′/D′ motif with a 3 nt-spacer. The nucleotide positions are numbered from 5′ to 3′ starting with the first one in the D or D′ motif. Positions 7, 8 and 9 correspond to the asymmetric internal loop of the K-turn with the flipped-out U (position 9). The symbols for the non-canonical base pairs (26,44) are shown in the legend at upper right. The non-canonical base pairs involving positions 5 and 7 of the K-loop are hypothetical. (**B**) Distribution of nucleotides in the terminal loops of 40 C/D guide RNAs from *P.abyssi* with a 3 nt-spacer.

This analysis did not reveal a new class of guide RNAs with longer spacing between the D′ and C′ elements. It is thus very unlikely that the short spacing is an artifact of the length constraint in the original computational searches.

In the case of *P.abyssi*, six other box C/D guide RNAs have a 4 nt linker and the spacing of the remaining 13 RNAs ranges from 0 to 11 nt. Linkers shorter than 8 nt effectively preclude the formation of a K-turn since a stable C-stem is expected to require at least a 2 bp stem and a 4 nt loop (Figure 2A). For comparison, we show a hypothetical K-loop in which the C-stem of the K-turn is replaced by a 3 nt-loop (Figure 2B). Sequence comparison of the 40 guide RNAs from *P.abyssi* with a 3 nt-spacer revealed a consensus RNK (G/A-N-G/U) sequence (Figure 2B) that is also conserved in the box C/D RNAs with a 3 nt-spacer in *P.horikoshii* and *P.furiosus* (data not shown). This conservation suggests either that the K-loop is stabilized by specific contacts between the terminal loop and the NC-stem or that the conserved nucleotides serve another function such as providing contacts for L7Ae binding.

L7Ae binds independently to two sites on sR47

sR47 is a C/D RNA from *P.abyssi* (Figure 3A) with antisense guides that target tRNA for 2′-*O*-ribose methylation (6). This RNA is typical of the *Pyrococcus* C/D RNAs in that the sequences of its motifs are well conserved, it has extended antisense guides directly 5′ of the D′ and D motifs, and there is a short 3 nt-linker between the D′ and C′ motifs. To test

binding, sR47 was incubated with increasing concentrations of L7Ae and complex formation was analyzed by EMSA. sR47 clearly shifts into two complexes, RNP1 and RNP2, showing that sR47 binds at least two molecules of L7Ae (Figure 3B). In order to determine the role of the C/D and C′/D′ motifs in L7Ae binding, the tandem sheared AG pairs in each motif were mutated (referred as *mut C/D* and *mut C′/D′*). Disruption of either motif abolished formation of RNP2 whereas disruption of both motifs (referred as *mut C/D* and *C′/D′*) abolished RNA binding (Figure 3B). These data strongly suggest that the C/D and C′/D′ motifs form two independent binding sites for L7Ae. Furthermore, they suggest that the C′/D′ motif, which cannot form a K-turn, nevertheless binds L7Ae and that the tandem sheared AG base pairs are important for this binding. The observation that RNP1 and RNP2 are formed sequentially with increasing L7Ae concentration shows that binding to the C/D and C′/D′ sites is not cooperative. This differs from the previously reported cooperative binding of L7Ae to sR8 of *Methanococcus jannashii* (14) and to the pre-tRNATrp of *Haloferax volcanii* (15). Based on those results, it has been suggested that the initial binding of L7Ae helps to stabilize the RNA structure of the second site. Nevertheless, our results clearly show that L7Ae binds independently to each site in sR47 under the conditions employed here. This conclusion is further supported by the dissociation constants calculated from the EMSA data (Figure 3B). The overall affinity of L7Ae for wild-type, *mut C/D* and *mut C′/D′* ranged from 50 to 100 nM. These small differences in affinity
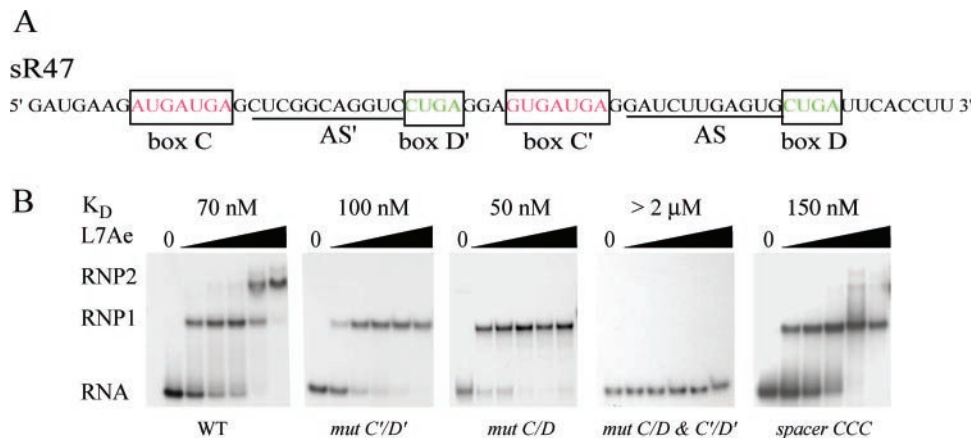
Figure 3. (**A**) The sequence of sR47, a C/D guide RNA from *P.abyssi* (6). Note that in the *Pyrococcus* sRNA database (http://rna.wustl.edu/snoRNAdb/), this RNA was annotated as sR31. The AS and AS′ guide sequences are complementary to putative methylation sites in tRNAGly(UCC) and tRNALeu(CAA), respectively. (**B**) Binding of L7Ae to sR47 as measured by EMSAs. Uniformly ^{32}P-labeled RNA was incubated in the presence of L7Ae from *P.abyssi* and then analyzed by native gel electrophoresis. The L7Ae concentrations were 62.5, 125, 250, 500 and 1000 nM. RNP1 and RNP2 are RNA–protein complexes. Dissociation constants (Materials and Methods) are indicated at the top of each panel. Wild-type (WT) sR47 and the various mutants are indicated at the bottom of each panel. In *mut C′/D′*, CUGA was changed to CUUC and GUGAUGA was changed to GUUCUGA, thus, disrupting the two AG sheared base pairs. Comparable changes were made in *mut C/D*. In *mut C/D* and *C′/D′* all four motifs were mutated. The GGA spacer between the D′ and C′ motifs was changed to CCC in the *spacer CCC* mutation.

are not significant. Thus, the initial binding to all three RNAs is equivalent.

To further examine the interaction of L7Ae with the C/D and C′/D′ motifs of sR47, footprinting with RNase T1 was performed (Figure 4A). Six positions at G7, G10, G13, G15, G29 and G36 are strongly protected from T1 digestion at 1 μM L7Ae (gray circles in Figure 4B). The first four sites are within or flanking the C motif whereas the other two sites are within the C′ and D′ motifs. Increased cleavage at G19 upon L7Ae binding suggests an alteration in the RNA structure of the spacer between the two motifs. These results provide direct physical evidence for contacts between L7Ae and the C/D and C′/D′ motifs of the sR47 RNA. Taken together with the work described in the previous paragraph, these results show that the sR47 guide RNA has two independent L7Ae-binding sites formed by the C/D motif and the C′/D′ motif.

The sequence and length of the C′/D′ spacer is important for L7Ae binding

To further explore the features of C′/D′ motif that are important for L7Ae binding, the GGA spacer of sR47 (Figure 3A) was mutated to CCC thus disrupting the RNK consensus sequence. This mutation affects formation of the RNP2 complex (Figure 3B) suggesting that the sequence of the spacer is important for L7Ae binding to the C′/D′ motif. The dissociation constant for L7Ae binding to the CCC mutant was 150 nM suggesting a slight reduction in affinity for L7Ae, which could be owing to a small destabilization of the overall structure of the RNA. Although it seems unlikely, we cannot exclude more complicated possibilities such as an effect of the C′/D′ spacer on L7Ae binding to the C/D motif. For this reason, the following experiments were performed with a model substrate derived from the C′/D′ motif of the sR6 guide RNA from *P.abyssi*. A 21 nt stem–loop (Figure 5A) containing the C′/D′ motif with a 3 nt GCA spacer has been shown recently to bind L7Ae (32). This fragment of sR6 RNA contains a C′/D′ motif closed at the 5′ and 3′

ends by three CG pairs, which stabilizes the secondary structure. As expected, L7Ae binds to this RNA (Figure 5B). Under our assay conditions, we measured a dissociation constant of 120 nM, which is not significantly different from the affinity of L7Ae for the sites in sR47. Mutation of the tandem sheared AG base pairs, *mut C′/D′*, abolishes L7Ae binding (Figure 5B). Furthermore, consistent with the results with the full-length sR47 RNA, changing the 3 nt spacer to CCC abolishes L7Ae binding (*LC*, Figure 5B). An RNA with a spacer in which the conserved G is changed to U, binds L7Ae with a dissociation constant of 350 nM (*L1*, Figure 5B). Moreover an RNA with a CCA spacer loses its ability to bind L7Ae (data not shown). Mutation in which the length of the spacer was reduced to 2 nt (GG) eliminates L7Ae binding (data not shown). The L2 mutation, in which the conserved G at position 8 was converted to an U, also eliminates L7Ae binding (*L2*, Figure 5B). These results confirm our conclusion based on the sR47 RNA that L7Ae can bind independently to the C′/D′ motif and demonstrate the importance of the sequence of the C′/D′ spacer for L7Ae binding.

DISCUSSION

We have shown by sequence comparisons that a C′/D′ motif with a short spacer is a general feature of the *Pyrococcus* C/D guide RNAs. A majority (40/59) of the C/D guide RNAs from *P.abyssi* contain a 3 nt spacer. The sequence of the 3 nt-spacer is not random. It displays an RNK consensus sequence. The preference for a purine in the first position is particularly strong (37/40). These conclusions hold for the C/D guide RNAs of two other *Pyrococcus* species, *P.furiosus* and *P.horikoshii*. In addition, the majority of C/D guide RNAs from other Archaea found in the database (except for *Aeropyrum pernix*) also show a preference for a 3 nt spacer (76% of *Pyrococcus aerophilum*, 54% of *Sulfolobus solfataricus*, 38% of *Sulfolobus acidocaldarius* and 66% of *A.fulgidus* and *Methanococcus jannaschii*). This feature is
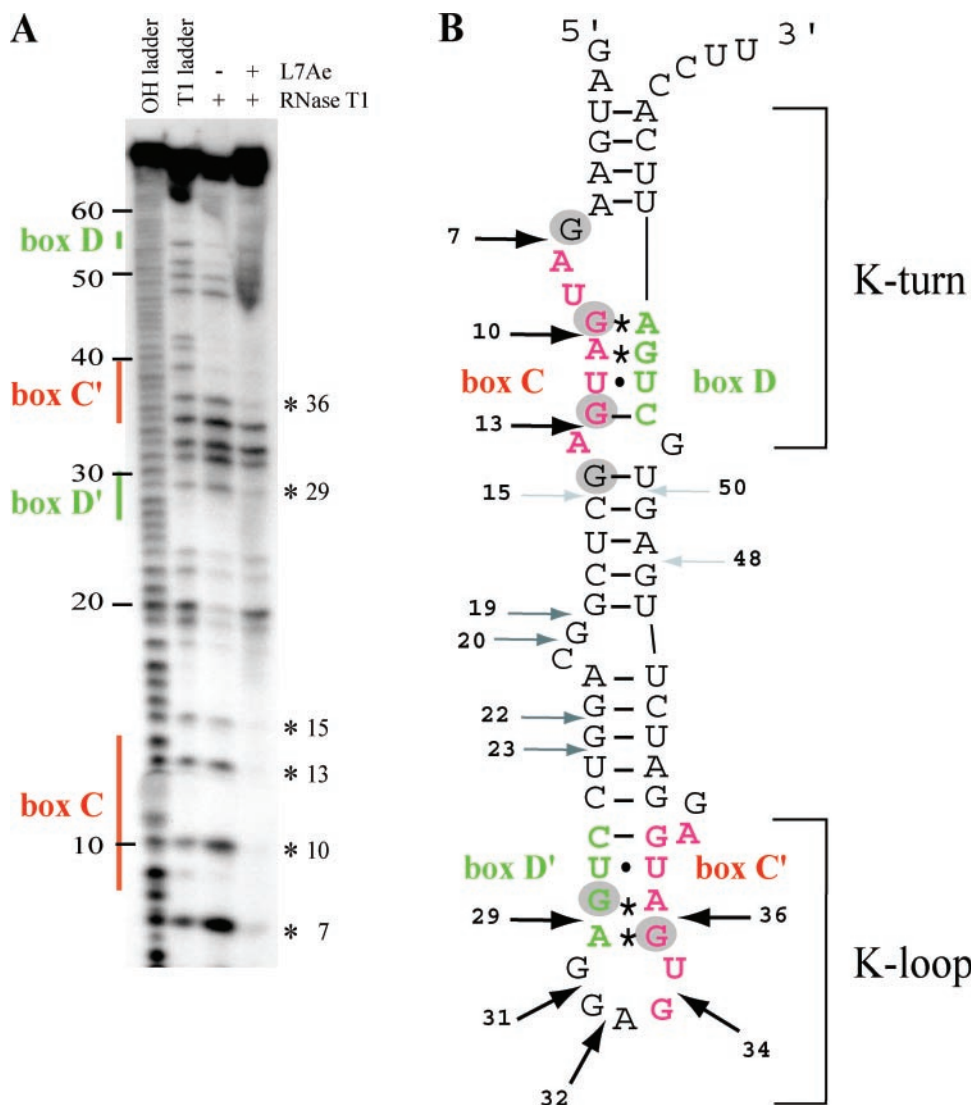
Figure 4. (**A**) Footprint of L7Ae on sR47 using RNase T1. T1 and alkaline hydrolysis (OH) ladders as well as control without L7Ae were analyzed in parallel. The location of the major sequence features in sR47 is indicated to the left. The asterisks to the right indicate the position of G residues protected from RNase T1 cleavage by L7Ae. (**B**) Schematic representation of the RNase T1 footprinting results. The AG sheared base pairs are indicated by asterisks; the other non-canonical base pairs are indicated by dots. Strong, medium and weak RNase T1 cleavages in absence of L7Ae are denoted by black, dark gray and light gray arrows, respectively. G residues that are protected from RNase T1 cleavage by L7Ae are indicated by gray circles. The position of the K-turn and the K-loop are as indicated.

also characteristic of the C/D guides identified by co-purification with the C/D core proteins (35,39) or by the sequencing of cDNA libraries (3,40) further supporting the notion that the short spacer is a general feature of the archaeal box C/D guide RNAs.

As demonstrated here by EMSA and RNA footprinting using the full-length sR47 guide RNA of *P.abyssi*, the C′/D′ motif is a bona fide site for the binding of L7Ae. Muta-genesis of the tandem sheared AG base pairs disrupting L7Ae binding and the non-cooperative binding of L7Ae show that the C/D and C′/D′ motifs can fold independently into func-tional L7Ae-binding sites. This could be a general feature of the *Pyrococcus* C/D RNAs since recently published results have demonstrated that an RNA stem–loop derived from the C′/D′ motif of the sR6 guide RNA of *P.abyssi* binds L7Ae with high affinity (32). The C′ and D′ motifs of sR47

and sR6 are both separated by a 3 nt-spacer. We have presen-ted experimental evidence showing that the sequence of the 3 nt-spacer is important for L7Ae binding. Of the C/D guide RNAs from *P.abyssi*, six have a spacer shorter than 3 nt. Visual inspection of the sequence of these guides reveals that the C′ and D′ elements often deviate from the consensus suggesting that these C′/D′ motifs might not form a functional L7Ae-binding site. Eleven other C/D guide RNAs from *P.abyssi* have spacer lengths ranging from 4 to 11 nt. Most of these guides have well-conserved C′ and D′ motifs suggest-ing that they might form a functional L7Ae-binding site. Although it would be interesting to know if there is a con-sensus sequence for the longer spacers, meaningful compar-isons are not possible since there are too few examples.

The yeast homolog of L7Ae, Snu13p, requires an intact K-turn containing the C-stem for RNA binding whereas
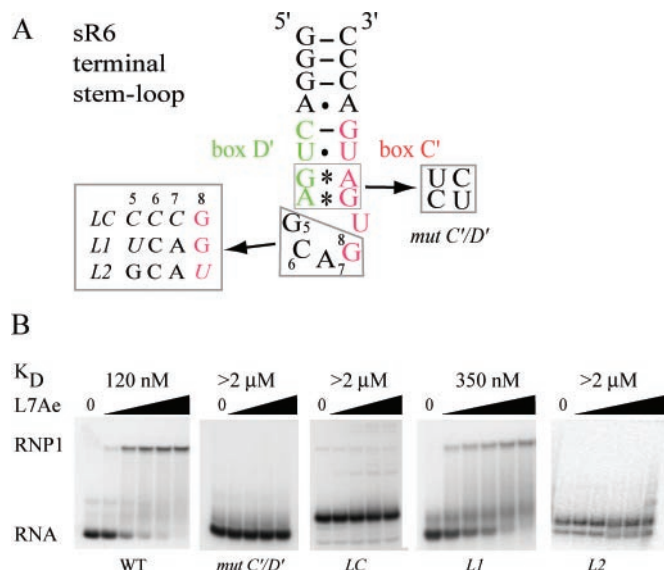
Figure 5. (**A**) An RNA stem–loop derived from the C′/D′ motif of the sR6 of *P.abyssi* (32). Mutations in the C′/D′ motif and in the spacer are indicated. (**B**) The binding of L7Ae measured by EMSA. Uniformly ³²P-labeled RNA was incubated in the presence of L7Ae from *P.abyssi* and then analyzed by native gel electrophoresis. The L7Ae concentrations were 62.5, 125, 250, 500 and 1000 nM for WT, *L1*, *L2* and *L4*; 62.5, 250, 500 and 1000 nM for *mut C′/D′* and *LC*. For *L1* and *L2* alternative conformers are observed which are not suitable for L7Ae binding.

L7Ae has a relaxed specificity permitting the recognition of the K-turn and the K-loop (32). Consistent with previous work in this field, we have assumed that the C/D motif of the *Pyrococcal* guide RNAs forms a K-turn. This is probable for the sR47 substrate employed here for *in vitro* experiments. The C-stem of the sR47 K-turn is formed by the pairing of bases at the 5′ and 3′ ends of the molecule (Figure 4B). However, with the notable exception of recently published work suggesting that most C/D guide RNAs from *P.furiosus* are circularized (41), few if any of the 5′ and 3′ ends of the *Pyrococcal* C/D guide RNAs have been mapped *in vivo*. The ends of the sR47 guide employed in this work were selected to keep the 5′ and 3′ terminal extensions as short as possible while maintaining the terminal C-stem. Whether the C/D motif actually forms a K-turn *in vivo* will depend on the position of 5′ and 3′ ends. In a circular guide RNA, the formation of a K-turn would be constrained by the spacing between C and D motifs. It is possible that the circularized C/D guide RNAs have L7Ae bound to K-loops at both motifs.

We believe that the C′/D′ motifs of the *Pyrococcal* guide RNAs are probable to form a K-loop similar to the terminal stem–loop of the archaeal H/ACA guide RNAs (29). The RNK consensus could be involved in either the binding of L7Ae or the structure of the K-loop. Considering our RNA footprinting data, the former possibility seems unlikely since three Gs in the terminal loop of sR47 (nucleotide positions 31, 32 and 34) are not protected from RNase T1 by L7Ae (Figure 4). Thus, it seems probable that the sequence of the spacer is important for RNA stability suggesting that the terminal loop adopts a specific structure. It is possible that the guide RNA with longer spacers (Figure 1B) also form K-loops in which the extra nucleotides are somehow accommodated into the structure

of the loop. The stability of RNA is a particular challenge for hyperthermophiles such as the *Pyrococcus* species, which can be subjected to temperatures as high as 110°C in deep ocean thermal vents (42). Increasing the GC content of RNA, as observed for rRNA and tRNA in many thermophiles (43), is one adaptation that could stabilize RNA. The circularization of RNA could also improve stability owing to the constraint imposed by eliminating free ends. In a similar fashion, K-loops could improve stability by replacing the C-stem of the K-turn with a small intrinsically structured terminal loop. Thus, in general, an important role of the K-loop could be to increase the resistance of the *Pyrococcal* C/D guide RNAs to thermal denaturation.

REFERENCES

1. Omer,A.D., Ziesche,S., Decatur,W.A., Fournier,M.J. and Dennis,P.P. (2003) RNA-modifying machines in archaea. *Mol. Microbiol.*, **48**, 617–629.
2. Kiss,T. (2002) Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions. *Cell*, **109**, 145–148.
3. Tang,T.H., Bachellerie,J.P., Rozhdestvensky,T., Bortolin,M.L., Huber,H., Drungowski,M., Elge,T., Brosius,J. and Huttenhofer,A. (2002) Identification of 86 candidates for small non-messenger RNAs from the archaeon *Archaeoglobus fulgidus*. *Proc. Natl Acad. Sci. USA*, **99**, 7536–7541.
4. Decatur,W.A. and Fournier,M.J. (2003) RNA-guided nucleotide modification of ribosomal and other RNAs. *J. Biol. Chem.*, **278**, 695–698.
5. Fatica,A. and Tollervey,D. (2003) Insights into the structure and function of a guide RNP. *Nature Struct. Biol.*, **10**, 237–239.
6. Clouet d'Orval,B., Bortolin,M.L., Gaspin,C. and Bachellerie,J.P. (2001) Box C/D RNA guides for the ribose methylation of archaeal tRNAs. The tRNATrp intron guides the formation of two ribose-methylated nucleosides in the mature tRNATrp. *Nucleic Acids Res.*, **29**, 4518–4529.
7. Dennis,P.P., Omer,A. and Lowe,T. (2001) A guided tour: small RNA function in Archaea. *Mol. Microbiol.*, **40**, 509–519.
8. Caffarelli,E., Fatica,A., Prislei,S., De Gregorio,E., Fragapane,P. and Bozzoni,I. (1996) Processing of the intron-encoded U16 and U18 snoRNAs: the conserved C and D boxes control both the processing reaction and the stability of the mature snoRNA. *EMBO J.*, **15**, 1121–1131.
9. Watkins,N.J., Segault,V., Charpentier,B., Nottrott,S., Fabrizio,P., Bachi,A., Wilm,M., Rosbash,M., Branlant,C. and Luhrmann,R. (2000) A common core RNP structure shared between the small nucleoar box C/D RNPs and the spliceosomal U4 snRNP. *Cell*, **103**, 457–466.
10. Kiss-Laszlo,Z., Henry,Y., Bachellerie,J.P., Caizergues-Ferrer,M. and Kiss,T. (1996) Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs. *Cell*, **85**, 1077–1088.
11. Kuhn,J.F., Tran,E.J. and Maxwell,E.S. (2002) Archaeal ribosomal protein L7 is a functional homolog of the eukaryotic 15.5kD/Snu13p snoRNP core protein. *Nucleic Acids Res.*, **30**, 931–941.

12. Szewczak,L.B., DeGregorio,S.J., Strobel,S.A. and Steitz,J.A. (2002) Exclusive interaction of the 15.5 kD protein with the terminal box C/D motif of a methylation guide snoRNP. *Chem. Biol.*, **9**, 1095–1107.

13. Watkins,N.J., Dickmanns,A. and Luhrmann,R. (2002) Conserved stem II of the box C/D motif is essential for nucleolar localization and is required, along with the 15.5K protein, for the hierarchical assembly of the box C/D snoRNP. *Mol. Cell. Biol.*, **22**, 8342–8352.

14. Tran,E.J., Zhang,X. and Maxwell,E.S. (2003) Efficient RNA 2′-*O*-methylation requires juxtaposed and symmetrically assembled archaeal box C/D and C′/D′ RNPs. *EMBO J.*, **22**, 3930–3940.

15. Bortolin,M.L., Bachellerie,J.P. and Clouet-d'Orval,B. (2003) *In vitro* RNP assembly and methylation guide activity of an unusual box C/D RNA, cis-acting archaeal pre-tRNA(Trp). *Nucleic Acids Res.*, **31**, 6524–6535.

16. Galardi,S., Fatica,A., Bachi,A., Scaloni,A., Presutti,C. and Bozzoni,I. (2002) Purified box C/D snoRNPs are able to reproduce site-specific 2′-*O*-methylation of target RNA in vitro. *Mol. Cell. Biol.*, **22**, 6663–6668.

17. Lafontaine,D.L. and Tollervey,D. (1999) Nop58p is a common component of the box C+D snoRNPs that is required for snoRNA stability. *RNA*, **5**, 455–467.

18. Lyman,S.K., Gerace,L. and Baserga,S.J. (1999) Human Nop5/Nop58 is a component common to the box C/D small nucleolar ribonucleoproteins. *RNA*, **5**, 1597–1604.

19. Newman,D.R., Kuhn,J.F., Shanab,G.M. and Maxwell,E.S. (2000) Box C/D snoRNA-associated proteins: two pairs of evolutionarily ancient proteins and possible links to replication and transcription. *RNA*, **6**, 861–879.

20. Cahill,N.M., Friend,K., Speckmann,W., Li,Z.H., Terns,R.M., Terns,M.P. and Steitz,J.A. (2002) Site-specific cross-linking analyses reveal an asymmetric protein distribution for a box C/D snoRNP. *EMBO J.*, **21**, 3816–3828.

21. Aittaleb,M., Rashid,R., Chen,Q., Palmer,J.R., Daniels,C.J. and Li,H. (2003) Structure and function of archaeal box C/D sRNP core proteins. *Nature Struct. Biol.*, **10**, 256–263.

22. Klein,D.J., Schmeing,T.M., Moore,P.B. and Steitz,T.A. (2001) The kink-turn: a new RNA secondary structure motif. *EMBO J.*, **20**, 4214–4221.

23. Nottrott,S., Hartmuth,K., Fabrizio,P., Urlaub,H., Vidovic,I., Ficner,R. and Luhrmann,R. (1999) Functional interaction of a novel 15.5kD [U4/U6.U5] tri-snRNP protein with the 5′ stem–loop of U4 snRNA. *EMBO J.*, **18**, 6119–6133.

24. Vidovic,I., Nottrott,S., Hartmuth,K., Luhrmann,R. and Ficner,R. (2000) Crystal Structure of the Spliceosomal 15.5kD Protein Bound to a U4 snRNA Fragment. *Mol. Cell*, **6**, 1331–1342.

25. Ban,N., Nissen,P., Hansen,J., Moore,P.B. and Steitz,T.A. (2000) The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. *Science*, **289**, 905–920.

26. Lescoute,A., Leontis,N.B., Massire,C. and Westhof,E. (2005) Recurrent structural RNA motifs, isostericity matrices and sequence alignments. *Nucleic Acids Res.*, **33**, 2395–2409.

27. Omer,A.D., Ziesche,S., Ebhardt,H. and Dennis,P.P. (2002) *In vitro* reconstitution and activity of a C/D box methylation guide ribonucleoprotein complex. *Proc. Natl Acad. Sci. USA*, **99**, 5289–5294.

28. Moore,T., Zhang,Y., Fenley,M.O. and Li,H. (2004) Molecular basis of box C/D RNA–protein interactions; cocrystal structure of archaeal L7Ae and a box C/D RNA. *Structure (Camb.)*, **12**, 807–818.

29. Hamma,T. and Ferre-D'Amare,A.R. (2004) Structure of protein L7Ae bound to a K-turn derived from an archaeal box H/ACA sRNA at 1.8 A resolution. *Structure (Camb.)*, **12**, 893–903.

30. Rozhdestvensky,T.S., Tang,T.H., Tchirkova,I.V., Brosius,J., Bachellerie,J.P. and Huttenhofer,A. (2003) Binding of L7Ae protein to the K-turn of archaeal snoRNAs: a shared RNA binding motif for C/D and H/ACA box snoRNAs in Archaea. *Nucleic Acids Res.*, **31**, 869–877.

31. Charpentier,B., Muller,S. and Branlant,C. (2005) Reconstitution of archaeal H/ACA small ribonucleoprotein complexes active in pseudouridylation. *Nucleic Acids Res.*, **33**, 3133–3144.

32. Charron,C., Manival,X., Clery,A., Senty-Segault,V., Charpentier,B., Marmier-Gourrier,N., Branlant,C. and Aubry,A. (2004) The archaeal sRNA binding protein L7Ae has a 3D Structure very similar to that of its Eukaryal counterpart while having a broader RNA-binding specificity. *J. Mol. Biol.*, **342**, 757–773.

33. Milligan,J.F. and Uhlenbeck,O.C. (1989) Synthesis of small RNAs using T7 RNA polymerase. *Methods Enzymol.*, **180**, 51–62.

34. Lowe,T.M. and Eddy,S.R. (1999) A computational screen for methylation guide snoRNAs in yeast. *Science*, **283**, 1168–1171.

35. Omer,A.D., Lowe,T.M., Russell,A.G., Ebhardt,H., Eddy,S.R. and Dennis,P.P. (2000) Homologs of small nucleolar RNAs in Archaea. *Science*, **288**, 517–522.

36. Qu,L.H., Meng,Q., Zhou,H., Chen,Y.Q., Liang-Hu,Q., Qing,M., Hui,Z. and Yue-Qin,C. (2001) Identification of 10 novel snoRNA gene clusters from *Arabidopsis thaliana*. *Nucleic Acids Res.*, **29**, 1623–1630.

37. Qu,L.H., Henras,A., Lu,Y.J., Zhou,H., Zhou,W.X., Zhu,Y.Q., Zhao,J., Henry,Y., Caizergues-Ferrer,M. and Bachellerie,J.P. (1999) Seven novel methylation guide small nucleolar RNAs are processed from a common polycistronic transcript by Rat1p and RNase III in yeast. *Mol. Cell. Biol.*, **19**, 1144–1158.

38. Gaspin,C., Cavaille,J., Erauso,G. and Bachellerie,J.P. (2000) Archaeal homologs of eukaryotic methylation guide small nucleolar RNAs: lessons from the *Pyrococcus* genomes. *J. Mol. Biol.*, **297**, 895–906.

39. Zago,M.A., Dennis,P.P. and Omer,A.D. (2005) The expanding world of small RNAs in the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Mol. Microbiol.*, **55**, 1812–1828.

40. Tang,T.H., Polacek,N., Zywicki,M., Huber,H., Brugger,K., Garrett,R., Bachellerie,J.P. and Huttenhofer,A. (2005) Identification of novel non-coding RNAs as potential antisense regulators in the archaeon *Sulfolobus solfataricus*. *Mol. Microbiol.*, **55**, 469–481.

41. Starostina,N.G., Marshburn,S., Johnson,L.S., Eddy,S.R., Terns,R.M. and Terns,M.P. (2004) Circular box C/D RNAs in *Pyrococcus furiosus*. *Proc. Natl Acad. Sci. USA*, **101**, 14097–14101.

42. Cohen,G.N., Barbe,V., Flament,D., Galperin,M., Heilig,R., Lecompte,O., Poch,O., Prieur,D., Querellou,J., Ripp,R. *et al.* (2003) An integrated analysis of the genome of the hyperthermophilic archaeon *Pyrococcus abyssi*. *Mol. Microbiol.*, **47**, 1495–1512.

43. Galtier,N. and Lobry,J.R. (1997) Relationships between genomic G+C content, RNA secondary structures, and optimal growth temperature in prokaryotes. *J. Mol. Evol.*, **44**, 632–636.

44. Leontis,N.B. and Westhof,E. (2003) Analysis of RNA motifs. *Curr. Opin. Struct. Biol.*, **13**, 300–308.