SCIENCE CHINA PRESS

SCIENCE CHINA Life Sciences
life.scichina.com link.springer.com
https://doi.org/10.1007/s11427-022-2412-3

COVER ARTICLE

Complicated target recognition by archaeal box C/D guide RNAs

Jiayin Wang[1,2†], Songlin Wu[1,2†] & Keqiong Ye[1,2*]

[1]Key Laboratory of RNA Science and Engineering, CAS Center for Excellence in Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China;
[2]University of Chinese Academy of Sciences, Beijing 100049, China

†Contributed equally to this work
*Corresponding author (email: yekeqiong@ibp.ac.cn)

Received 26 October 2022; Accepted 14 December 2022; Published online 30 November 2023

Box C/D RNAs guide the site-specific formation of 2′-*O*-methylated nucleotides (Nm) of RNAs in eukaryotes and archaea. Although C/D RNAs have been profiled in several archaea, their targets have not been experimentally determined. Here, we mapped Nm in rRNAs, tRNAs, and abundant small RNAs (sRNAs) and profiled C/D RNAs in the crenarchaeon *Sulfolobus islandicus*. The targets of C/D RNAs were assigned by analysis of base-pairing interactions, *in vitro* modification assays, and gene deletion experiments, revealing a complicated landscape of C/D RNA-target interactions. C/D RNAs widely use dual antisense elements to target adjacent sites in rRNAs, enhancing modification at weakly bound sites. Two consecutive sites can be guided with the same antisense element upstream of box D or D′, a phenomenon known as double-specificity that is exclusive to internal box D′ in eukaryotic C/D RNAs. Several C/D RNAs guide modification at a single non-canonical site. This study reveals the global landscape of RNA-guided 2′-*O*-methylation in an archaeon and unexpected targeting rules employed by C/D RNA.

C/D RNA | 2′-*O*-methylation | archaea | guide RNA | RNA modification

INTRODUCTION

Most RNAs undergo post-transcriptional modifications, which can impact their structure, function, and stability (Boccaletto et al., 2018). In particular, tRNAs and rRNAs, which are core components of the protein translation machinery, harbor a wide range of modifications. RNA modifications are synthesized by protein enzymes or two distinct classes of RNA-guided enzymes. These RNA-guided enzymes depend on either box H/ACA RNA or box C/D RNA for substrate recognition and synthesize pseudouridines or 2′-*O*-methylated nucleotides (Nm) in rRNAs, snRNAs, and tRNAs (Karijolich and Yu, 2010; Kiss, 2001; Watkins and Bohnsack, 2011; Yu and Meier, 2014). H/ACA and C/D guide RNAs are present in eukaryotes and archaea but are absent in bacteria.

Box C/D RNAs contain terminal box C (RUGAUGA, R=purine) and D (CUGA) motifs, along with internal box C′ and D′ motifs. These motifs are connected by two spacers that can bind to complementary sequences in substrate RNAs. According to the "D+5" rule (Cavaille et al., 1996; Kiss-Laszlo et al., 1996; Tycowski et al., 1996), the nucleotide paired with the position 5 upstream of box D or D′ is selected for modification. To form an active RNA-protein complex (RNP), each C/D RNA associates with Nop5, fibrillarin and L7Ae in archaea or Nop56, Nop58, fibrillarin and Snu13 in eukaryotes (Omer et al., 2002; Yang et al., 2020).

The 2′-*O*-methylation modifications in rRNAs and their corresponding guide RNAs have been extensively studied in several model eukaryotes. In the yeast *Saccharomyces cerevisiae*, rRNAs contain a total of 55 Nm (Taoka et al., 2016; Yang et al., 2016a). Except for a single site (Gm2922 in 25S rRNA) being modified by the stand-alone enzyme Sbp1 (Lapeyre and Purushothaman, 2004), the remaining Nm are synthesized by 42 C/D RNA-guided enzymes. These RNAs commonly select their target sites following the D+5 rule, but four yeast C/D RNAs can simultaneously target two neighboring or closely located sites using a single antisense sequence upstream of box D′ (Kiss-Laszlo et al., 1996; Lowe and Eddy, 1999; van Nues and Watkins, 2016). A recent study also identified five C/D RNAs in *Arabidopsis* with similar double-specificity (Wu et al., 2021). Moreover, besides binding directly to their target sites, a significant fraction of yeast and *Arabidopsis* C/D RNAs form extra base pairs with sequences close to the target sites. This auxiliary base pairing facilitates substrate binding and modification processes (Cao et al., 2022; van Nues et al., 2011; Wu et al., 2021).

Box C/D RNA-guided 2′-*O*-methylation enzymes are conserved in archaea, where they modify rRNAs and tRNAs (Omer et al., 2000; Omer et al., 2002). The activity and structure of archaeal C/D RNPs have been extensively studied using *in vitro* reconstituted complexes (Aittaleb et al., 2003; Lin et al., 2011; Omer et al., 2002; Tran et al., 2003; Yang et al., 2016b; Yang et al., 2020). However, the *in vivo* target sites of C/D RNAs and the rules governing their targeting have not been systemically analyzed in any archaeon. C/D RNAs have been profiled in several archaeal species (Bernick et al., 2012; Daume et al., 2017; Dennis et al., 2015; Omer et al., 2000; Tang et al., 2005; Zago et al., 2005). Their targets in rRNAs and tRNAs were commonly predicted according to sequence complementarities with spacers and the D+5 rule. It is unknown to what extent those predicted sites are faithful.

In this study, we profiled Nm and C/D RNAs in *Sulfolobus islandicus*, which is an aerobic acidothermophilic model archaeon

belonging to the *Crenarchaeota* kingdom (Guo et al., 2011; Peng et al., 2017). By determining the guide RNAs for the majority of detected Nm and validating non-canonical targeting rules, we obtained a global view of C/D RNA-guided 2′-*O*-methylation in an archaeon and revealed unexpected rules governing target recognition.

RESULTS

Small RNAs in *S. islandicus*

We profiled small RNAs (sRNAs) in *S. islandicus* by small RNA sequencing (sRNA-seq) (Table S1 in Supporting Information). RNAs with a length of less than 600 nt were purified from gel, converted to libraries and subjected to 150-bp paired-end sequencing. A total of 60 C/D RNAs (sR1–sR60) were identified based on their conserved sequence motifs (Table S2 in Supporting Information). Similar to other archaea, *S. islandicus* possesses a C/D RNA within the intron of precursor tRNA-Trp, which guides the methylation of C34 and C39 in the same tRNA molecule (Bortolin et al., 2003; Clouet d'Orval et al., 2001). The total number of identified C/D RNAs reaches 61 when including the tRNA-Trp intron.

In addition, we identified RNase P RNA, 7S RNA, 5S RNA, 46 tRNAs, 208 clustered regularly interspaced short palindromic repeats (CRISPR) RNAs (crRNA), and 612 other sRNAs (sR101–sR712) (Figure S1A and Table S1 in Supporting Information). The other sRNAs include 2 putative H/ACA RNAs (sR115 and sR435), 39 derived from insertion elements (IS), 69 hypothetical coding sequences (hCDS), 368 antisense RNAs, and 134 sRNAs of unknown type. The antisense RNAs include 273 sequences complementary to mRNAs, 17 to tRNAs, 18 to C/D RNAs, 6 to crRNAs, 1 to H/ACA RNA and 53 to IS RNAs (Figure S1B in Supporting Information). Antisense RNAs of IS are highly abundant, representing 86.5% of the total antisense RNAs, which is consistent with the previous observation in *S. solfataricus* (Tang et al., 2005).

RiboMeth-seq analysis of *S. islandicus* RNAs

We mapped Nm in *S. islandicus* RNAs with the high-throughput RiboMeth-seq method (Tables S3–S5 in Supporting Information) (Birkedal et al., 2014; Marchand et al., 2016). The method utilizes an RNA alkaline hydrolysis reaction, where the 3′-phosphate undergoes nucleophilic attack by the 2′-OH group and the 2′-*O*-methylation blocks this reaction. Total RNA was hydrolyzed at pH 9.2 and 95°C into RNA fragments which were converted into libraries and sequenced. After the sequencing reads were mapped to individual RNA sequences, the 5′ and 3′ ends were counted at each position and used to calculate MethScores. MethScore is defined as 1 minus the ratio of the end count at a specific position to the background end count at neighboring sites (Birkedal et al., 2014). MethScores take a maximum of 1 for fully methylated sites and can vary considerably for unmodified sites.

To address potential issues with artificially high scores caused by significant variations in end counts along RNA sequences, especially in terminal regions and tRNAs, we modified the calculation of MethScore. Specifically, we processed the 5′ and 3′ end counts separately and balanced the background counts on the left and right sides of each position (See MATERIALS AND METHODS). The revised formula yielded scores similar to the original formula for Nm in rRNAs, which have stable background patterns. However, it effectively suppressed the majority of artificial high scores from terminal regions (Figure 1A).

Nm in rRNAs

To assess the baseline MethScores for unmodified sites, we also analyzed *in vitro* transcribed rRNAs. The MethScores displayed a bimodal distribution for endogenous modified rRNAs, while all MethScores for the *in vitro* transcribed rRNAs were found to be below 0.75 (Figure 1B). A modification site was assigned if its MethScore was >0.7 and significantly different from that of the unmodified site (Figure 1C). As many sites with low MethScores still exhibited significantly different MethScores from the unmodified rRNAs, low score modified sites cannot be distinguished.

We identified a total of 68 Nm sites, which include 24 in 16S rRNA, 43 in 23S rRNA, and 1 in 5S rRNA (Table 1; Table S3 in Supporting Information). As a comparison, previous studies using chromatography and mass spectrometry in the phylogenetically related species *S. solfactaricus* reported similar numbers of Nm sites: 24 in 16S rRNA, 43 in 23S rRNA, and 1 in 5S rRNA (Bruenger et al., 1993; Noon et al., 1998). However, the precise locations of these Nm sites were not determined in the previous studies. Out of the 68 Nm sites identified in *S. islandicus* rRNAs, ten sites are conserved in *S. cerevisiae*, *Arabidopsis*, and Human (Figure 1D; Table S6 in Supporting Information). *S. islandicus* shares additional 1 to 4 Nm sites with each of these eukaryotic species. Overall, *S. islandicus* is distantly related to eukaryotes in rRNA 2′-*O*-methylation.

Nm in other sRNAs

We calculated MethScores for other sRNAs with sufficient read coverage and focused on the most confident modification sites (Table S5 in Supporting Information). We identified 2′-*O*-methylation at G225 of RNase P RNA and C56 of sR471 and no modification in 7S RNA (Figures S2A, S2B and S3A in Supporting Information). sR471 is a highly abundant small RNA that is antisense to sR38 C/D RNA (Table S1 in Supporting Information) and folds into a hairpin structure (Figure S3B in Supporting Information). Despite the complementarity between sR38 and sR471, *in vitro* modification assay showed that sR38 cannot guide methylation of C56, which is located at position 8 upstream box D′ (Figure S3C and D in Supporting Information). sR19 is predicted to guide the methylation of C56 with a weak interaction (Table S5 in Supporting Information).

Assignment of C/D RNA

To assign guide RNAs to the detected Nm sites, the sequence flanking each methylation site was paired with all spacers of C/D RNAs. The most suitable spacer-substrate duplexes were selected by assessing their stability, the position of methylation site relative to box D/D′, and the presence of dual-targeting. The selected duplexes commonly consisted of 8–12 Waston-Crick (WC) pairs, 0–2 GU wobble pairs, and no mismatches. For several methylation sites that did not locate at the canonical D+5 position, their guide RNAs were further validated by *in vitro* modification assays and genetic deletion experiments. Interest-

https://doi.org/10.1007/s11427-022-2412-3

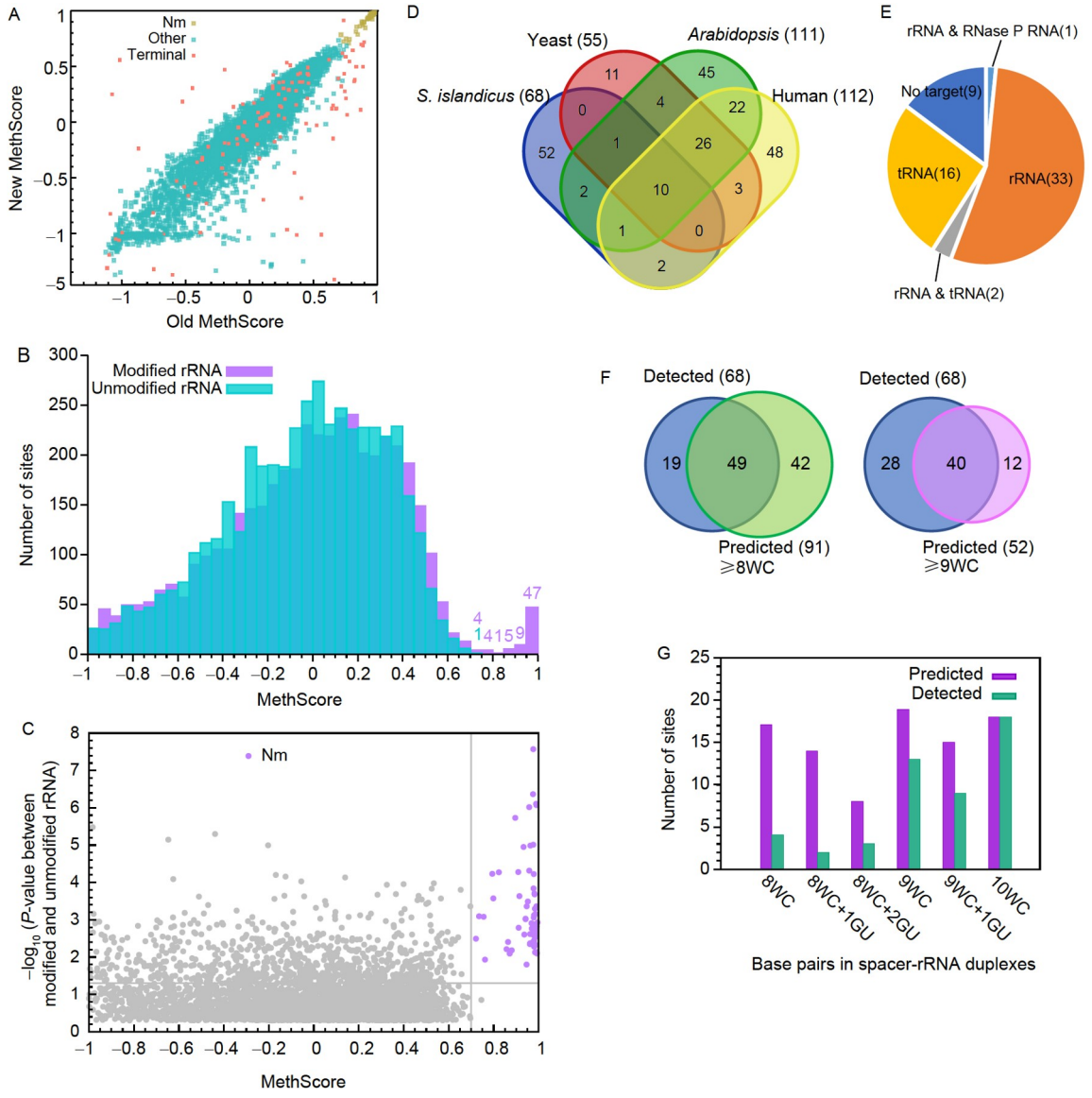

Figure 1. Nm in *S. islandicus* rRNAs and assignment of C/D RNA. A, MethScores of rRNAs calculated using the new formula versus the old formula. The scales of x- and y-axes are different for the range of −1 to 1 and −5 to −1. Identified Nm sites are shown in yellow, 20 nt terminal sites of each rRNA in red, and other sites in cyan. B, Histogram showing the distribution of MethScores for endogenous modified and *in vitro* transcribed unmodified *S. islandicus* rRNAs. The means of n=3 replicates were counted. Sites with MethScore<−1 are not shown. The numbers of sites are indicated for high-value bins. C, Plot of $-\log_{10}$ of P-value of differences between MethScores of endogenous and *in vitro* transcribed 16S and 23S rRNAs against MethScores from endogenous rRNAs. The horizontal line represents a P-value of 0.05, while the vertical line represents a MethScore threshold of 0.7. The P-values are derived from one-tail t-test with n=3. The MethScore of *in vitro* transcribed 5S rRNA was measured once and is not displayed. D, Venn diagram showing the conservation of Nm sites in four species. E, Classification of 61 C/D RNAs based on their substrate types. F, Venn plots showing the overlap between the detected Nm sites of rRNAs and the predicted sites that form at least 8 or 9 WC base pairs with C/D RNAs. G, Comparison of predicted and detected Nm sites in rRNAs categorized by the number of base pairs between substrates and positions 2–11 of spacers.

ingly, we found when two sites in close proximity along the primary sequence were targeted by dual antisense elements of a C/D RNA, one of the sites could be bound with relatively weak interactions (dual-targeting, see below). We also validated a few unconventional or weak interactions between C/D RNAs and their target sites, including a bulge for sR47, a 7-bp interaction for sR51, and a wobble pair at the methylation site for sR27 (Figure S4A and B in Supporting Information).

Overall, a total of 36 C/D RNAs were assigned to guide the modification of 65 Nm sites in rRNAs (Figure 1E; Table 1; Table S2 in Supporting Information). This includes the assignment of C2746 in 23S rRNA to sR19 by genetic evidence. However, three sites, namely C2191 and C2733 in 23S rRNA, and C32 in 5S rRNA, remain unassigned to specific guide RNAs. Additionally, 18 C/D RNAs were predicted to guide the methylation of 55 sites in tRNAs (Table S4 in Supporting Information, see below). Notably, sR25 and sR45 were found to target both rRNAs and tRNAs, while sR51 guides modification of both rRNA and RNase P RNA.

There are still 9 C/D RNAs (sR1, sR5, sR15, sR16, sR29, sR49, sR50, sR54, and sR59) that have no assigned target. sR1, sR16 and sR29 are one order of magnitude less abundant compared to other C/D RNAs, suggesting that their amounts may be insufficient for targeting abundant rRNAs and tRNAs. Moreover, sR16 and sR29 are located within the SIRE_0632 and SIRE_1167 genes, unlike other C/D RNAs that are typically

Table 1. Nm in *S. islandicus* rRNAs and assigned C/D RNAs

Site	Guide _Spacer[a)]	Target postion[b)]	Site	Guide _Spacer[a)]	Target Position[b)]
16S			23S		
U33	sR32_1	5	U1999	sR55_1	5
U52	sR32_2	5	C2040	sR55_2	5
G473	sR43_2	5	G2050	sR27_1	5
C481	sR11_2	5	A2056	sR48_2	5
A494	sR43_1	5	C2063	sR35_1	5
A635	sR37_1	5	A2081	sR27_2	5
G672	sR37_2	5	C2089	sR48_1	7
G857	sR53_1	5	G2093	sR41_2	5
A877	sR53_2	5	U2099	sR35_2	5
G905	sR4_1	5	U2110	sR39_2	5
G906	sR4_1	6	G2126	sR41_1	5
G926	sR6_1	5	C2191	NA	
G1007	sR52_2	5	G2388	sR34_1	5
G1018	sR12_1	5	A2414	sR25_2	5
U1032	sR52_1	5	A2424	sR11_1	5
G1194	sR6_2	5	C2451	sR11_2	5
A1199	sR20_2	5	C2564	sR31_1	6
A1253	sR46_2	5	C2581	sR31_2	5
C1254	sR46_2	6	G2586	sR8_2	5
C1272	sR46_1	5	U2600	sR34_2	5
U1344	sR10_1	5	A2609	sR9_2	6
C1366	sR20_1	5	A2618	sR8_1	5
G1372	sR33_2	5	U2619	sR7_2	5
C1490	sR36_2	5	U2631	sR9_1	5
23S			U2644	sR7_1	5
A681	sR18_1	5	G2649	sR47_1	5
A684	sR25_2	5	U2692	sR47_2	5
G886	sR45_1	5	G2708	sR51_2	5
A894	sR21_2	5	A2717	sR19_1	4
G934	sR21_1	5	C2730	sR51_1	5
G1112	sR3_1	5	U2733	NA	
A1132	sR3_2	5	G2739	sR24_2	5
C1842	sR60_1	5	C2746	sR19_2	5
A1942	sR17_2	5	5S		
G1993	sR17_1	5	C32	NA	

a) 1 for D′ spacers; 2 for D spacers; NA: not assigned. b) Position of target site relative to box D/D′.

found in intergenic regions and frequently overlap with a gene (Table S2 in Supporting Information). Among the 61 C/D RNAs identified in *S. islandicus*, sR16 and sR38 are not conserved in *S. solfataricus*, and 25 C/D RNAs are not conserved in *S. acidocaldarius* (Table S2 in Supporting Information). Furthermore, except for sR5, 8 of the 9 orphan C/D RNAs are not conserved in *S. acidocaldarius*, suggesting that they may have evolved more recently.

Performance of target prediction of C/D RNA

The identification of Nm sites in *S. islandicus* rRNAs allowed us to evaluate the accuracy of target prediction of C/D RNAs. We predicted 91 target sites on rRNAs based on the D+5 rule and the requirements that the core interaction between substrates and positions 2–11 of spacers should contain a minimum of 8 WC pairs, 0–2 GU pairs at non-targeted sites, and no mismatch (Table S3 in Supporting Information). The prediction yielded a true positive rate of 53.8% (49/91) and a false negative rate of 30.8% (19/68) (Figure 1F). Upon analyzing the reasons for the 19 missing sites, we found that 6 were located at non-canonical sites, 10 were targeted by weak interactions that did not meet the pairing requirement of prediction, and 3 had no assigned C/D RNA. The true positive rate is positively correlated to the number of WC pairs in the spacer-substrate duplexes, while the presence of GU pairs seemed to have a neutral effect (Figure 1G). It was also observed that several sites with good match to C/D RNAs were not modified likely due to limited substrate accessibility. By imposing a more stringent requirement of at least 9 WC pairs, 52 sites are predicted. This resulted in an increased overall true positive rate of 76.9% (40/52), but also a higher false negative rate of 41.2% (28/68) (Figure 1F). Therefore, due to moderate accuracy and coverage, the common practice of target prediction from C/D RNA sequences is inadequate for accurately identifying Nm.

2′-*O*-methylation in tRNAs

The analysis of tRNA using RiboMeth-seq is challenging due to its short length, extensive modifications, and stable secondary structure (Marchand et al., 2017). To adapt RiboMeth-seq for tRNAs and other short RNAs, we purified small RNAs with <100 nt, combined multiple datasets to increase the read coverage and applied the revised formula to calculate MethScores. Consequently, MethScores were determined for 3,604 sites in 46 tRNAs, excluding 68 sites with low end coverage (Figure 2A; Table S4 in Supporting Information). In contrast to rRNAs, the distribution of MethScores for tRNAs does not exhibit clearly defined high and low score regions. This lack of distinction may be attributed to noisy data arising from highly variable end counts of tRNAs (Figures S2C, S2D, and S5 in Supporting Information) and incomplete modification catalyzed by stand-alone protein enzymes.

In archaea, 2′-*O*-methylation of tRNAs is catalyzed by both C/D RNA-guided enzymes and stand-alone protein enzymes (Hori et al., 2018). The sequences of intron-containing pre-tRNAs and mature tRNAs were used to search guide RNAs for 269 sites with MethScores>0.7. 55 sites at 14 positions were assigned to 18 C/D RNAs, 45 sites at positions 32 and 56 were found to be probably catalyzed by the protein enzymes TrmJ and Trm56 (Renalier et al., 2005; Somme et al., 2014), and 169 sites remain unassigned (Figure 2B). A significant fraction of unassigned sites are likely false positives due to the noisy data of tRNA. sR57, sR38 and sR25 were assigned with less certainty to three weakly modified sites tRNA-Ser(GCT)_U13, tRNA-Lys(CTT)_G22, and tRNA-Ser (CGA)_A49, respectively.

The tRNA model contains 76 consensus positions and a few variable positions, which have few high-scored sites (Figure 2D; Table S4 in Supporting Information). MethScores are shown as a heatmap for 76 consensus positions of 46 tRNAs (Figure 2E). The methylated sites with assigned C/D RNAs are located at positions 13, 17, 18, 20, 22, 34, 39, 44, 47, 49, 50, 51, 68, and 70. While most positions are modified in 1–6 tRNAs, positions 34 and 50 exhibit widespread methylation. Multiple tRNA isoacceptors are

https://doi.org/10.1007/s11427-022-2412-3

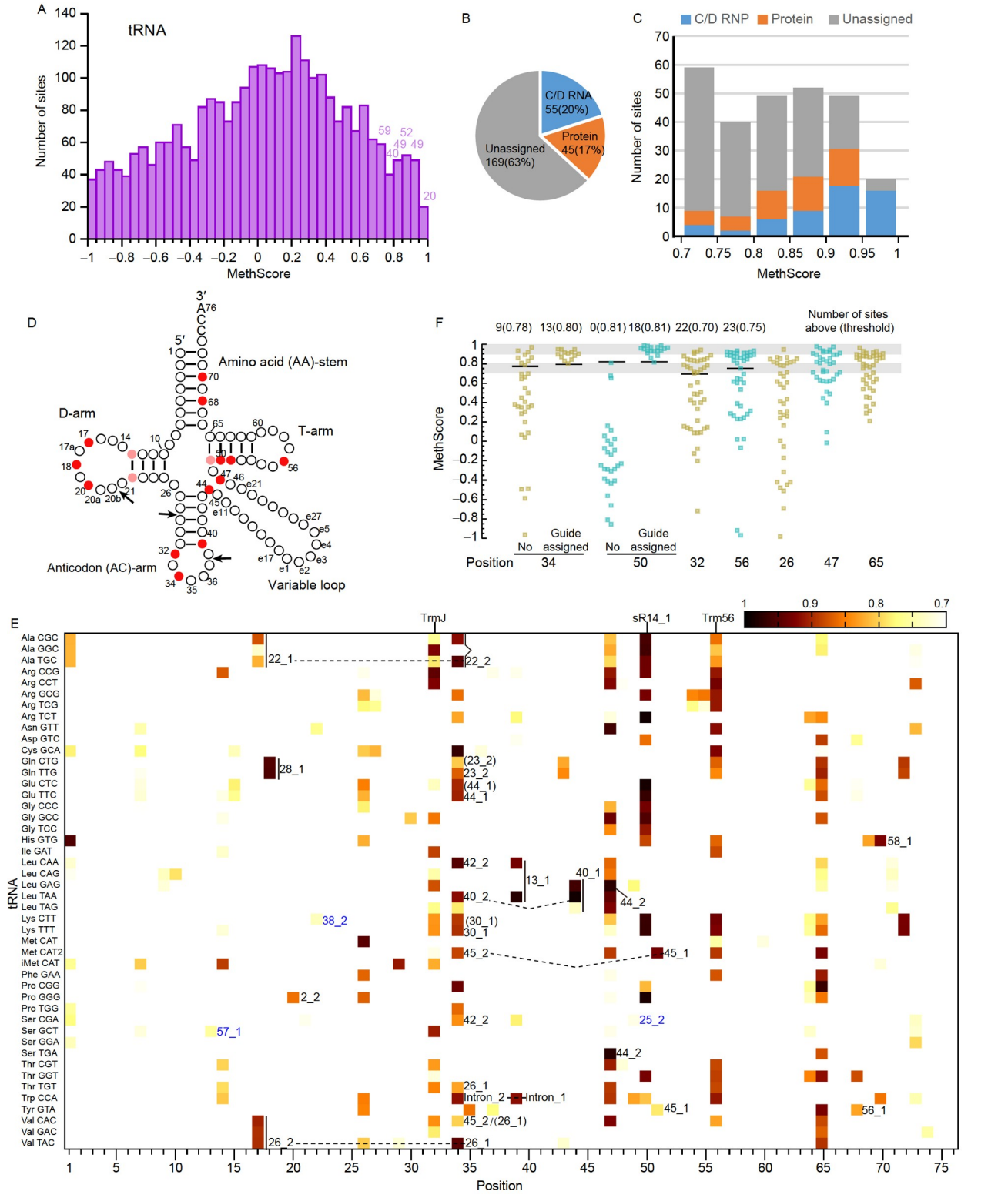

Figure 2. Nm in tRNAs and assignment of modification enzymes. A, Histogram showing the distribution of MethScores for 46 tRNAs. Sites with a score<−1 are not shown. B, Pie plot showing the assignment of modification enzymes for 269 tRNA sites with MethScores>0.7. C, Stacked bar plot showing the assignment of modification enzymes for tRNA sites categorized into different ranges of MethScores. D, Cloverleaf representation of tRNA. Methylation positions with assigned enzymes are colored red for MethScore>0.8 and pink for MethScore=0.7–0.8. Key residues are numbered. E, Heatmap showing the MethScores at model positions 1–76 for tRNAs. The assigned C/D RNAs (indicated by numbers) and spacers (1 for D′ spacer and 2 for D spacer) are labeled on the right side of the methylation sites. The putative guides that form an AC mismatch with substrates at the modification site are shown in parentheses. C/D RNAs assigned to a single weakly modified site are shown in blue. Two sites targeted by dual spacers of a C/D RNA are connected by dashed lines. F, Bee swarm plot showing the MethScores at selected positions. Sites at positions 34–50 are divided by guide RNA assignment. Thresholds for identifying modifications at individual positions and the numbers of identified Nm sites are indicated at the top. Thresholds were not determined for positions 26, 47 and 65 due to lack of enzyme assignments.

sometimes targeted by a single C/D RNA due to their conserved sometimes. Due to variations in the distribution of MethScores at different positions, individual thresholds for defining modification were determined (Figure 2F).

Position 34, the wobble position of anticodon, harbors a diverse array of modifications that enhance the interaction with the third position of codon. We found that 22 tRNAs are likely 2′-*O*-methylated at position 34 (MethScore>0.78), involving 10 C, 10 U and 2 G (Figure 2E and F). Due to sequence variability in the anticodon loop, 9 C/D RNAs have been found to guide the modification of 13 tRNAs at position 34. The remaining 9 unassigned sites may either be recognized by C/D RNAs by weak interactions (Table S4 in Supporting Information) or modified by protein enzymes as observed in *E. coli* and *S. cerevisiae* (Benitez-Paez et al., 2010; Pintard et al., 2002).

Position 50 is found to be modified in 18 tRNAs (MethScore>0.81) (Figure 2E and F). All of these modified nucleotides are cytosines, and they are all targeted by sR14, which happens to be the most abundant C/D RNA in *S. islandicus* (Table S1 in Supporting Information). The D′ spacer of sR14 pairs with nucleotides 47–56 of the tRNAs, encompassing the T-arm, which is highly conserved except at positions 47 and 49–51 (Figure S6 in Supporting Information). For the modification to occur, the pairing at position 47 of the tRNAs is dispensable, but it is crucial that the sequences at positions 49–51 are CCG in order to form base pairs with sR14. Any mutation of the CCG sequence would block the modification. The sR14-guided modification of tRNA demonstrates the strict requirement for pairing at positions 4–6 in the interaction between C/D RNAs and their target sites.

In the three tRNA-Val isoacceptors with the CUG sequence at positions 49–51 (Figure S6 in Supporting Information), the middle target site U50 cannot be modified despite forming a wobble pair with sR14 (Figure 2E). Interestingly, we have also observed a case where a wobble pair is allowed at the target position in sR27-guided modification of 23S G2050 (Figure S4 in Supporting Information). The distinction between these cases could be influenced by factors such as the location of the nucleotide involved in the wobble pair (U in the tRNA substrate versus G in the rRNA substate) and the sequence context of spacer-substrate duplexes.

Position 32 has a MethScore>0.7 in 22 tRNAs, while position 56 has a MethScore>0.75 in 23 tRNAs (Figure 2E and F). Interestingly, these positions tend to be partially modified, with MethScores ranging from 0.7 to 0.95 (Figure 2C and F). Previous studies have identified the protein enzymes TrmJ and Trm56 as responsible for 2′-*O*-methylation at positions 32 and 56, respectively, in archaeal tRNAs (Renalier et al., 2005; Somme et al., 2014). In *S. islandicus*, the homologs of these enzymes, SiRe_1729 and SiRe_1263, are likely involved in the modification at these positions. However, the specific determinants governing enzyme specificity remains unclear based on the available modification profiles (Figure S6 in Supporting Information).

Positions 26, 47 and 65 show high MethScores in numerous tRNAs (Figure 2E and F); however, the presence of modification at these positions remains uncertain for most cases. Only position 47 of two tRNAs is predicted to be targeted by sR44, while no other guide RNA has been assigned to these positions. In archaeal tRNAs, guanosine at position 26 is frequently modified to m2,2G and further to m2, 2Gm by an unknown enzyme (Hori et al., 2018; Wolff et al., 2020; Yu et al., 2019). In 43 *S. islandicus* tRNAs, position 26 is occupied by G and could potentially undergo 2′-*O*-methylation. However, due to significant fluctuating 3′ end counts from position 25 to 26 (Figure S5B in Supporting Information), the MethScore for position 26 cannot be confidently determined. As for position 65, 2′-*O*-methylation has not been detected in previous studies (Hori et al., 2018; Wolff et al., 2020; Yu et al., 2019). This nucleotide is located at a GC-rich region and its hydrolysis may be affected by strong secondary structures in this region.

In addition, high MethScores (>0.9) are found at position 1 of tRNA-His, position 29 of initiator tRNA and position 72 of tRNA-Gln and tRNA-Lys. In *S. acidocaldarius*, Nm are also likely to present at position 29 of initiator tRNA and position 72 of tRNA-Gln(TTG) (Wolff et al., 2020).

Prevalent occurrences of dual-targeting

To guide the modification of 65 sites in rRNAs, a total of 36 C/D RNAs form 63 spacer-substrate duplexes (Tables S2 and S3 in Supporting Information). Two cases, namely sR4_2 and sR46_1, involve one duplex directing the modification of two adjacent sites (double-specificity, see below). Remarkably, 25 C/D RNAs employ their dual spacers to simultaneously target two sites that are closely located in the primary sequences, called as dual-targeting, accounting for 50 interactions and 79% of assigned RNA-guided methylations in rRNAs (Figure 3A). Dual-targeting is also observed in 5 cases for tRNAs (Figure 2E). The remaining 13 interactions are mediated by single spacers of 12 C/D RNAs (single-targeting) (Figure 3A). It is noteworthy that in some cases, a single spacer can target two distant sites. For example, the D spacer of sR25 targets both A684 and A2412 in 23S rRNA. Additionally, sR11 meditates dual-targeting of A2424 and A2451 of 23S and single-targeting of 16S_C481.

In dual-targeting, the two sites are mostly separated by a distance of 17–51 nt, with an average of 29.2 nt. However, for sR6, sR20, and sR34, the separation can be larger, ranging from 167 to 268 nt (Figure 3B). The stability of the spacer-substrate duplexes in single-targeting has a Gibbs free energy (ΔG) ranging from -15 to -24 kcal mol^{-1}. The stronger duplexes in dual-targeting have similar stabilities as the duplexes in single-targeting, but the weaker duplexes in dual-targeting are significantly less stable (ΔG=-7 to -18 kcal mol^{-1}) (Figure 3C). This suggests that dual-targeting generates a cooperative effect on substrate binding, allowing one of the targets to be bound with weak interactions.

To study the biochemical advantages of dual-targeting, we measured *in vitro* modification activities of assembled *S. solfartaricus* C/D RNPs (Lin et al., 2011; Yang et al., 2020; Ye et al., 2009). sR43 was predicted to guide the methylation of G473 and A494 of 16S rRNA using its D′ and D spacers, respectively (Figure 3D). When using short substrates, the modification efficiency was high for A494, but poor for G473, whose flanking sequences form a 7-bp duplex with sR43 (ΔG= -12.2 kcal mol^{-1}) (Figure 3E, lanes 1 and 2). However, G473 was efficiently and specifically modified when a long substrate containing both target sites and a pre-methylated A494 was used (Figure 3E, lanes 3 and 4). Interestingly, the modification of the short G473 substrate was also enhanced when the short A494 substrate was present *in trans*, although not to the same extent as with the long substrate (Figure 3E, lanes 3 and 5).

sR8 was predicted to target G2586 and A2618 of 23S rRNA

https://doi.org/10.1007/s11427-022-2412-3

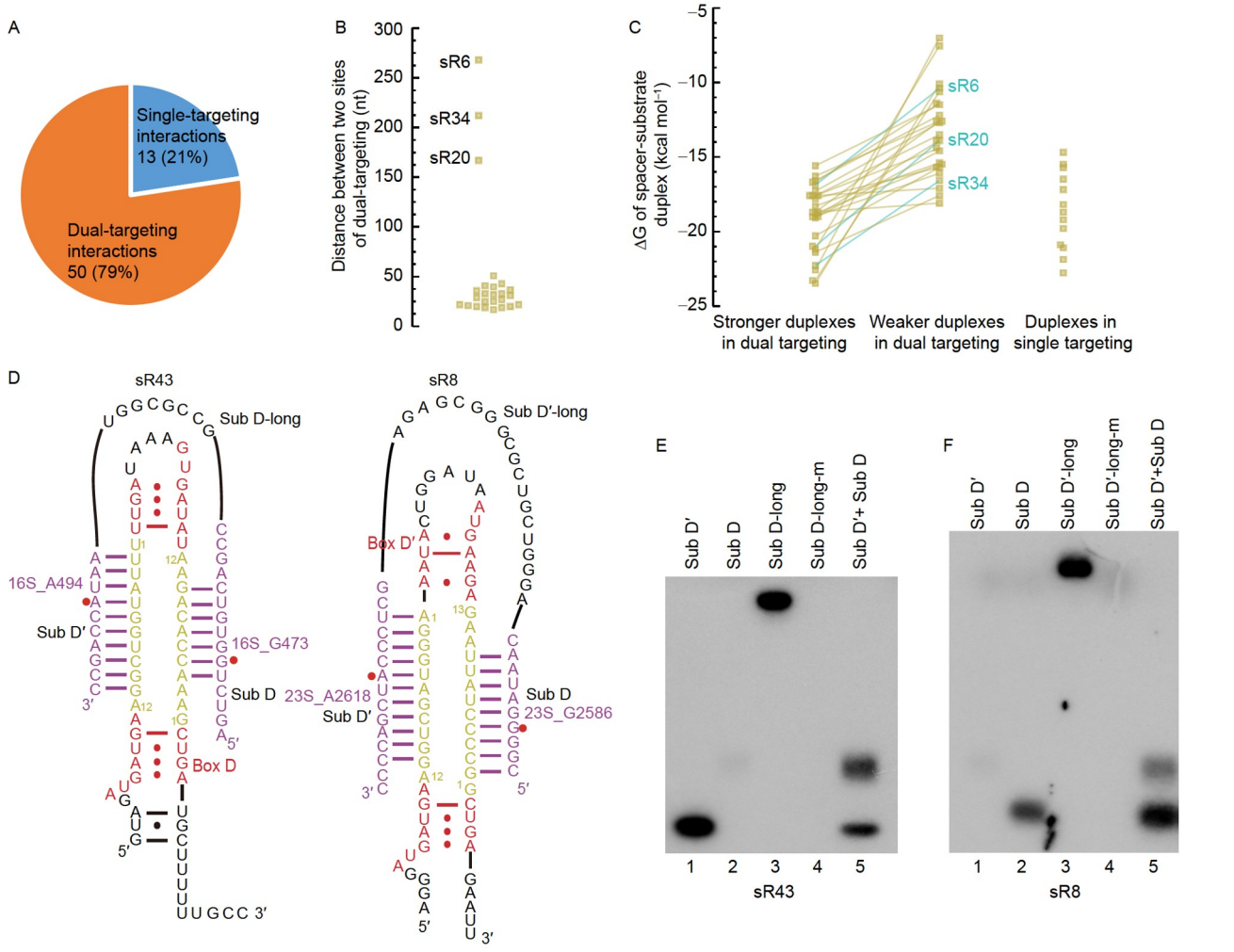

Figure 3. C/D RNAs targeting dual adjacent sites of rRNA. A, Pie plot showing the classification of 63 interactions between C/D RNA spacers and rRNAs. B, Bee swarm plot showing the distances between the two sites targeted by two spacers of a C/D RNA. C, Bee swarm plot showing the Gibbs free energy of spacer-substrate duplexes. Stronger and weaker duplexes in a dual-targeting event are grouped separately and connected by a line. Duplexes from single-targeting are in a different group. D, Secondary structures of sR43 and sR8 and their interactions with substrates. Box C/D and C′/D′ motifs are colored red, spacers are yellow, and substrates are purple. Detected Nm sites are marked by red circles. Nucleotides in spacers are numbered according to their distance to the coupled box D/D′. E, Activity of sR43-RNP. Sub D-long includes residues 468–500 of 16S rRNA and a pre-methylated A494. Sub D-long-m additionally contains a pre-methylated G473. F, Activity of sR8-RNP. Sub D′-long includes residues 2,583–2,626 of 23S rRNA and a pre-methylated G2586. Sub D′-long-m additionally contains a pre-methylated A2618.

using its dual spacers (Figure 3D). When using short substrates containing individual targets, G2586 was efficiently modified by sR8-loaded enzymes, whereas A2618 showed poor modification (Figure 3F, lanes 1 and 2). sR8 contains a degenerated box D′ motif (AAUA), which may potentially interfere with the guiding activity of the D′ spacer. However, when A2618 was placed in a long substrate that can bind both spacers of sR8, it was efficiently and specifically modified (Figure 3F, lanes 3 and 4). Moreover, the presence of the D substrate *in trans* also promoted the modification of the D′ substrate (Figure 3F, lane 5).

These data demonstrate that simultaneous interactions with dual spacers can enhance the modification of substrates that would otherwise be poorly modified. The enhancement of dual-targeting is likely due to two factors. First, the binding of one substrate to C/D RNPs, even *in trans*, may induce structural changes in C/D RNPs, facilitating the binding and modification of the other substrate. Second, the binding of the other substrate can be kinetically promoted when the two substates are covalently connected.

Double-specificity: two consecutive sites targeted by one antisense element

In 16S rRNA, two consecutive sites, G905 and G906, as well as A1253 and C1254, are both modified (Table 1; Figure S2E and F in Supporting Information). sR4 and sR46 were predicted to guide the modification of G905 and A1253, respectively, based on the D+5 rule. However, their downstream sites did not have any assigned guide. In yeast and *Arabidopsis thaliana*, several C/D snoRNAs guide methylation of two consecutive sites by the same antisense element upstream of box D′ (Kiss-Laszlo et al., 1996; Lowe and Eddy, 1999; van Nues and Watkins, 2016; Wu et al., 2021). We hypothesized that sR4 and sR46 might possess similar double-specificity and guide the methylation of two sites at positions 5 and 6 (Figure 4A). To test this hypothesis, we analyzed the substrate specificity of reconstituted sR4 and sR46 RNPs. We found that a substrate covering G905 and G906 was efficiently modified by the sR4 RNP (Figure 4B, lane 1). The modification was partially reduced when G905 or G906 was

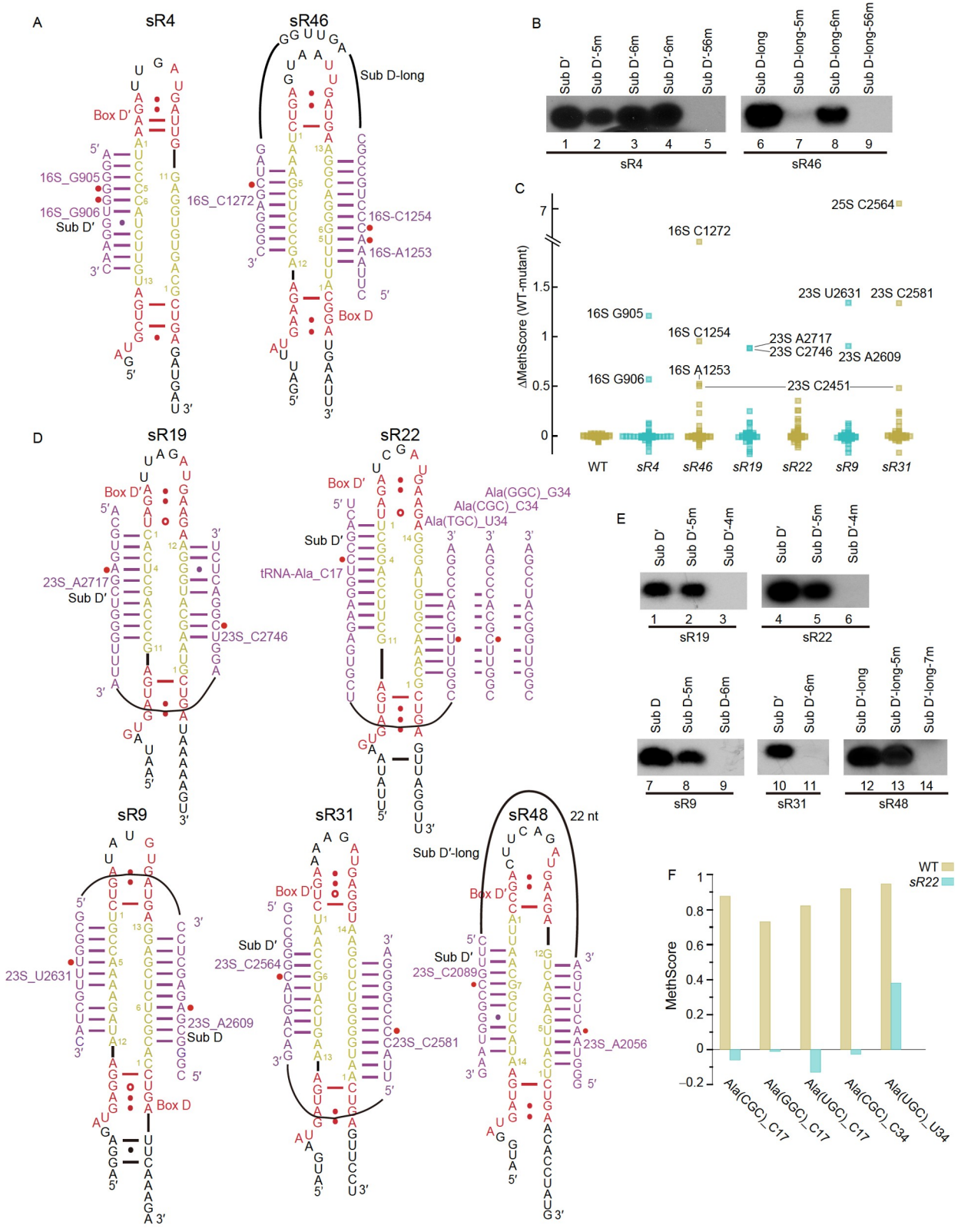

Figure 4. Special targeting rules of *S. islandicus* C/D RNAs. A, Secondary structures of sR4 and sR46 with double-specificity. Box C/D and C′/D′ motifs are colored red, spacers yellow, and substrates purple. Detected Nm sites are marked by red circles. Nucleotides in spacers are numbered according to their distance to the coupled box D/D′. B, Modification activity of sR4- and sR46-loaded RNPs. Substates are pre-methylated at position 5 (-5m) or 6 (-6m), or both (-56m). C, Bee swarm plot showing changes in MethScore for detected Nm sites in rRNAs between wild-type and C/D RNA deletion strains. The WT group was calculated between two independent wild-type samples. MethScore of 23S-C2451 was non-specifically reduced in both sR31 and sR46 deletion strains. D, Secondary structure models of five C/D RNAs targeting a non-canonical site. The assayed substrates are shown. E, Modification activity assay of sR19-, sR22-, sR9-, sR31-, and sR48-loaded RNPs. The substrates are pre-methylated at position 4 (-4m), 5 (-5m), 6 (-6m), or 7 (-7m) to assess site specificity. F, MethScores at positions 17 and 34 of tRNA-Ala in WT and sR22 deletion strains.

https://doi.org/10.1007/s11427-022-2412-3

individually pre-methylated and completely lost when both sites were pre-methylated (Figure 4B, lanes 2–5), validating that sR4 targets both G905 and G906.

As the D′ guide of sR46 additionally targets C1272 of 16S rRNA, the activity of sR46-loaded RNP was assayed on a long substrate that spans the targets of the D′ and D guides and contains a pre-methylated C1272. The substrate was efficiently modified by sR46-loaded RNP (Figure 4B, lane 6). When the substrate was pre-methylated at A1253 and A1254 individually, it was modified at a low and moderate level, respectively. However, when both sites were pre-methylated, no modification was observed (Figure 4B, lanes 7–9). These results indicate that both A1253 and A1254 were methylated by sR46-RNP with A1253 at position 5 being modified *in vitro* more efficiently than A1254 at position 6. Therefore, sR46 targets three sites with its dual spacers, resembling U24 in yeast and U24 and SnoR29 in *Arabidopsis* (Wu et al., 2021). Unlike eukaryotic C/D RNAs with double-specificity, sR46 employs its D spacer, rather than the D′ spacer, to guide the modification of two consecutive sites.

To further validate the targets of sR4 and sR46, we individually deleted their genes using the CRISPR gene editing method and profiled Nm in the mutant strains (Figure S7 in Supporting Information). The RiboMeth-seq data show that deletion of sR4 dramatically reduces the MethScores of both G905 and G906, while deletion of sR46 similarly decreased the MethScores of both A1253 and C1254 (Figure 4C). These results provide further evidence for the double-specificity of sR4 and sR46.

Targeting a single non-canonical position

We identified five C/D RNAs that target a non-canonical site at position 4 (sR19 and sR22), 6 (sR9 and sR31) or 7 (sR48) upstream of box D or D′ instead of the canonical position 5 (Figure 4D). These RNAs select only a single non-canonical site for modification using a spacer, distinguishing them from sR4 and sR46, which exhibit double-specificity by targeting both positions 5 and 6.

sR19 was predicted to guide the methylation of A2717 of 23S rRNA, which pairs with position 4 upstream of box D′. A cognate substrate was modified efficiently by sR19 RNP, and this modification was eliminated when A2717 at position 4 was pre-methylated, while pre-methylation at position 5 had no effect (Figure 4E, lanes 1–3). This validates the specific targeting of A2717 by sR19. Moreover, deletion of sR19 in cells abolished the methylation of A2717 (Figure 4C), consistent with the *in vitro* results. Surprisingly, the deletion also caused the loss of methylation at a proximate site, C2746, which was not originally assigned as a target of sR19 based on our criteria. The dual-targeting capability of sR19 appears to enable it to target C2746 with a weak interaction that involves an AG mismatch (Figure 4D).

sR22 was predicted to guide the methylation of C17 in three tRNA-Ala isoacceptors, which is located at position 4 upstream of box D′ (Figures 2E and 4D). This non-canonical target site was confirmed both by *in vitro* modification assays (Figure 4E, lanes 4–6) and by the genetic deletion of sR22 (Figure 4F). The D spacer of sR22 was predicted to guide the modification of U34 in tRNA-Ala(TGC), while C34 in tRNA-Ala (CGC) was not expected to be targeted due to a mismatch at the modification site (Figures 2E and 4D). Consistent with the prediction, deletion of sR22 abolished the modification at U34 in tRNA-Ala(TGC). Intriguingly, C34 of tRNA-Ala(CGC) was also not modified in the absence of sR22, indicating that it is another authentic target of the D spacer of sR22. This unusual case suggests that an AC mismatch between the guide and the substrate at the modification site is tolerable. It is worth noting that an AG mismatch is disallowed at the target site, as G34 in tRNA-Ala(GGC) was not modified at all (Figures 2E and 4D).

sR9 and sR31 were predicted to guide the methylation of A2609 and C2564 of 23S rRNA, respectively, which pair with position 6 of the spacers upstream of box D or D′ (Figure 4D). The *in vitro* modification assays confirmed the specific methylation of A2609 by sR9-RNP and C2564 by sR31-RNP (Figure 4E, lanes 7–11). Additionally, the genetic deletion of sR9 and sR31 specifically abolished the modification at the predicted non-canonical target sites, as well as the canonical target sites of their other spacers (Figure 4C).

sR48 was predicted to guide the modification of C2089 of 23S rRNA at position 7 using the D′ spacer, and the nearby site A2056 at position 5 with the D spacer (Figure 4D). Using a long substrate containing both target sites and a pre-methylated A2056, we confirmed that sR48 guides the specific modification of C2089 (Figure 4E, lanes 12–14). Unfortunately, the deletion of sR48 gene was not successful.

DISCUSSION

We have systemically mapped Nm in rRNAs, tRNAs and other sRNAs and identified 61 C/D RNAs in *S. islandicus*. The comprehensive lists of Nm and C/D RNAs, along with the extensive assignment of guide RNAs, provide a global view about C/D RNA-guided 2′-*O*-methylation in this archaeon. The majority of Nm sites in rRNAs (65/68) can be assigned with guide RNAs, indicating that 2′-*O*-methylation in rRNAs is predominantly RNA-guided. The three unassigned sites could be synthesized by C/D RNPs through weak interactions or by stand-alone protein enzymes. In yeast, a conserved Spb1 protein is responsible for the 2′-*O*-methylation of G2922 in 25S rRNA (Lapeyre and Purushothaman, 2004). The equivalent nucleotide G2693 in *S. islandicus* is unmodified, suggesting a lack of Spb1-mediated methylation in this archaeon.

Despite the noisy nature of the MethScore data for tRNAs, we were able to identify ~100 Nm sites in tRNA that could be assigned with guide RNAs or known protein enzymes. Positions 34 and 50 emerged as hot spots for 2′-*O*-methylation. Due to the conserved sequences around position 50, it is targeted by a single abundant C/D RNA, sR14, in all modified tRNAs. Modification of position 34 engages at least 9 C/D RNAs, representing a significant portion of the C/D RNA pool. In *S. acidocaldarius*, 2′-*O*-methylation at position 34 appears to be also frequent and detected in 6 of 22 analyzed tRNAs (Wolff et al., 2020). However, in other archaea, 2′-*O*-methylation at position 34 is only occasionally observed in tRNA-Met(CAU), tRNA-Gln(CUG) and tRNA-Trp (Gupta, 1984; Wolff et al., 2020; Yu et al., 2019). We also present evidence that archaeal C/D RNAs can guide methylation on RNA molecules beyond rRNAs and tRNAs. RNase P RNA contains a Gm225 modification, which is likely guided by sR51.

Mismatches were rarely found at the modification site in guide-substrate duplexes, as they would disrupt the conformation of the nucleotide to be modified. Artificially introduced mismatches at

the modification site generally block the modification, although they can be tolerated in a highly stable duplex (Cavaille and Bachellerie, 1998; Cavaille et al., 1996). We have found examples of naturally occurring non-WC pairs at modification sites, including a UG wobble pair (sR27) and an AC mismatch (sR22). Moreover, if an AC mismatch (A in guide and C in substrate) is allowed, modification at position 34 of tRNA-Gln (CTG), tRNA-Glu(CTC), tRNA-Lys(CTT), and tRNA-Val(CAC) can be assigned to a C/D RNA that originally targets their respective isoacceptors (Figure 2E; Figure S8 in Supporting Information). These findings indicate that the base pairing interaction between C/D RNAs and substrates exhibits greater flexibility than previously thought.

Our study has uncovered several important features of substrate recognition by C/D RNAs that were previously unappreciated. One notable finding is the extensive utilization of dual spacers by C/D RNAs to guide the modification of two sites that are close in the primary sequence of rRNA. In fact, this dual-targeting mechanism accounts for 79% of RNA-guided methylation in rRNAs. While a few examples of archaeal C/D RNAs targeting two close sites with dual spacers have been noted previously, the prevalence of this dual-targeting mechanism was not fully appreciated (Dennis et al., 2015; Ziesche et al., 2004). In contrast, in *Arabidopsis*, only a small fraction of C/D RNAs (6 out of 108) guide methylation of two close sites using dual spacers (Wu et al., 2021). Many eukaryotic C/D RNAs recognize substrates with an extra pairing interaction that does not typically guide methylation (van Nues et al., 2011; Wu et al., 2021). In *S. islandicus*, dual-targeting is the exclusive form of extra pairing. This distinctive strategy of dual-targeting not only enhances the efficiency of modification but also serves as an economical way to introduce many modifications using a limited set of guide RNAs.

We have shown biochemically that interactions with a nearby target site can increase the modification of a target site that is otherwise poorly modified. Interestingly, this enhancement occurs when both substrates are provided either *in trans* or *in cis*. These findings suggest that the binding of one target site not only facilitates the binding of closely spaced sites but also induces conformational changes in the enzyme, thereby promoting efficient binding and modification. This cooperative mechanism may operate effectively for two target sites that are more distantly (~200 nt) located in the primary sequence, as observed for sR6, sR20 and sR34.

Classic C/D RNAs typically target a single site that pairs to position 5 upstream of box D/D′. In yeast and *Arabidopsis*, there are snoRNAs that exhibit double-specificity, targeting two sites with the same antisense element linked to box D′. These snoRNAs, such as U18, snR13, and U24 in yeast and U24, SnoR58Y, SnoR29, and SnoR10 in *Arabidopsis*, typically target two sites located at positions 5 and 6 upstream of box D′. We have found that *S. islandicus* also possesses C/D RNAs, namely sR4 and sR46, which exhibit double-specificity similar to the eukaryotic snoRNAs. Nevertheless, there are differences in the specific spacers involved. In *S. islandicus*, the double-specificity can be attributed to either the D′ or D spacers, whereas in eukaryotes, it is exclusively found in the D′ spacers. Additionally, yeast snR48 and its *Arabidopsis* homolog SnoR1 display another type of double-specificity, where two target sites are separated by one nucleotide. However, we have not found a similar example of this double-specificity in *S. islandicus*. These findings suggest that the phenomenon of double-specificity in C/D RNAs is not exclusive to eukaryotes but has its origin in archaea.

We have found for the first time that C/D RNAs can target a single site other than position 5. We identified five C/D RNAs that target sites located at positions 4, 6 or 7 upstream of box D or D′. This non-canonical targeting of a single site has not been previously reported for eukaryotic C/D RNAs, where non-canonical targeting is always associated with the double-specificity of D′ spacers.

Structural analysis of archaeal C/D RNPs has revealed the molecular mechanism behind the D+5 rule (Lin et al., 2011; Yang et al., 2016b). In the structures of substrate-bound C/D RNPs, the substrate forms a 10-bp duplex with positions 2–11 of the spacers. Positions 1 and 12 of the spacers are unpaired and contact Nop5. The first base-pair at position 2 stacks on the CTD of Nop5, creating a molecular ruler that places the ribose at position 5 precisely at the active site of fibrillarin. The newly discovered C/D RNAs that exclusively target position 6 (sR9 and sR31) or target both positions 5 and 6 (sR4 and sR46) have a 13-nt spacer, which is one nucleotide longer than the more common 12-nt spacer found in archaeal C/D RNAs (Yang et al., 2016b). By allowing one nucleotide of the spacer between box D/D′ and the spacer-substrate duplexes to loop out, the target site can be shifted to position 6. The first two to three bases of these spacers remain unpaired, facilitating the looping-out process, except for sR4. The reason why sR4 and sR46 guide methylation at both positions 5 and 6, while sR9 and sR31 exclusively target position 6, remains puzzling. In the case of sR48, the only identified C/D RNA targeting position 7, it has a 14-nt spacer with the first three bases unpaired. Looping out two bases would position the nucleotide at position 7 into the active site for methylation. However, the targeting mechanism of sR19 and sR22, which follow the D+4 rule and have an 11-nt spacer, cannot be explained by the looping-out hypothesis. Further structural analysis of non-canonical targeting C/D RNPs is expected to shed light on their mechanisms of action.

MATERIALS AND METHODS

Growth of *S. islandicus*

The uracil-auxotrophic E233S1 (Δ*pyrEF* Δ*lacS*) strain of *S. islandicus* REY15A, a kind gift from Dr. Qunxin She, was used in this study (Deng et al., 2009; Peng et al., 2017). The basic medium, prepared per liter, consisted of 3 g $(NH_4)_2SO_4$, 0.5 g K_2SO_4, 0.1 g KCl, 0.7 g glycine, 0.8 mg $MnCl_2$•$4H_2O$, 2.1 mg $Na_2B_4O_7$•$10H_2O$, 0.11 mg $ZnSO_4$•$7H_2O$, 0.025 mg $CuSO_4$•$5H_2O$, 0.015 mg Na_2MoO_4•$2H_2O$, 0.015 mg $VOSO_4$•$5H_2O$, 0.005 mg $CoSO_4$•$7H_2O$, 0.005 mg $NiSO_4$•$6H_2O$, 2 mg $FeSO_4$•$7H_2O$ (dissolved in 0.5 mol L^{-1} HCl), 1 mmol L^{-1} $MgCl_2$, and 0.3 mmol L^{-1} $Ca(NO_3)_2$ (Zillig et al., 1993). The pH was adjusted to 3.0 by adding concentrated sulfuric acid. To create the PSCVU medium, the basic medium was supplemented with 1 mmol L^{-1} KH_2PO_4, 0.2% sucrose, 0.2% vitamin-free casamino acids (Difco Vitamin Assay, BD, USA), 1% of a mixed vitamin solution, and 20 μg mL^{-1} uracil. The 100× vitamin solution per liter contains 10 mg niacin, 4 mg biotin, 10 mg pantothenate, 10 mg lipoic acid, 4 mg folic acid, 10 mg p-aminobenzoic acid, 10 mg vitamin B1, 10 mg vitamin B2, 10 mg vitamin B6, and 10 mg vitamin B12. The casamino acids solution was treated with 0.1% activated carbon (Sigma-Aldrich, USA) overnight and filtered

https://doi.org/10.1007/s11427-022-2412-3

through a 0.22 μm membrane. Liquid cultures were grown at 78°C in high-temperature shakers (HZ-9612K by Taicang Hualida Laboratory Equipment Co., Ltd, Shanghai, China or ZQWY-200S by Shanghai Zhichu Instrument Co., Ltd., Shanghai, China) with a rotation speed of 150 r min^{-1}.

Construction of sRNA deletion strains

Genome editing utilized the pSeRp plasmid which can be constructed to contain a mini-CRISPR array and a donor sequence for homology recombination (Li et al., 2016). Oligonucleotides used for generating sRNA mutants were purchased from Sangon Biotech (Table S7 in Supporting Information). The pSeRp plasmid was linearized using BspMI (Invitrogen, USA) and then ligated with a chemically synthesized spacer DNA using T4 DNA ligase. The resulting plasmid was transformed into *E. coli* DH5α strain and amplified. The spacer-containing plasmid was digested with SalI and NotI (NEB, USA) followed by gel purification. A donor DNA fragment was generated by overlap PCR from *S. islandicus* genomic DNA using KOD DNA polymerase and specific primers (Table S7 in Supporting Information). The donor DNA was then ligated to the linearized spacer-containing plasmid using the ClonExpress II One Step Cloning Kit (Vazyme, Nanjing, China). The resulting plasmids, known as pGE plasmids, were amplified in DH5α and confirmed by Sanger sequencing. Next, the pGE plasmids (~800 ng) were transformed into *S. islandicus* E233S1 strain by electroporation. The transformed cells were grown in liquid PSCV (uracil-less PSCVU) medium at 75°C for 14–16 d and then in PSCV plates for 5 d. Positive colonies were sequenced to confirm the desired mutation.

sRNA-seq

Total RNA (12 μg) was resolved using 8% urea polyacrylamide gels. Gel regions containing 20–600 nt RNAs were excised, chopped into small pieces, and soaked in 400 μL of RNA extraction buffer (300 mmol L^{-1} NaAc, pH 5.2, 1 mmol L^{-1} EDTA and 0.25% SDS) at 22°C for 12 h. The RNA was then precipitated with isopropanol.

To prepare sequencing libraries, RNA was dephosphorylated using 5 units of FastAP phosphatase (Invitrogen), followed by phenol/chloroform extraction and isopropanol precipitation. RNA was then phosphorylated at the 5′ end using 20 units of T4 polynucleotide kinase (NEB) and recovered as described above. The sequencing libraries were prepared with the NEBNext Small RNA Library Prep Set (NEB) following the manufacturer's instruction, with the use of custom adaptors and primers. The 3′ DNA adaptor (5′-PO_4-TGGAATTCTCGGGTGCCAAGG-NH_2-3′) was adenylated with Mth RNA Ligase (NEB). The 5′ adaptor sequence is 5′-ACACGArCrGrCrUrCrUrUrCrCrGrArUrCrUrNrNrNrUr-3′, where deoxynucleotides are represented by single letters, ribonucleotides are indicated by "r", "N" denotes a random sequence, and the underlined sequences represent barcodes. The reverse transcription primer and the indexed P5 and P7 primers have been previously described (Wu et al., 2021).

Processing of sRNA-seq data

The libraries of sRNA-seq were sequenced using the Illumina HiSeq X10 platform in the 150-bp paired-end mode by Annoroad Gene Technology. The reads were demultiplexed based on the indexes in the P5 and P7 primers. Adaptor sequences were removed with Flexbar (v3.1), and reads with a minimum of 14-nt actual RNA sequences (excluding barcode sequences) were retained (Roehr et al., 2017). A custom script was used to remove PCR duplicates that share the same barcode and RNA sequences, and the barcodes were trimmed and stored in the read names.

The sRNA-seq reads were mapped to the genome of *S. islandicus* (CP002425.1) (Guo et al., 2011) by HISAT2 using the no-spliced-alignment option (Kim et al., 2015). The identified sRNAs should have a combined read count of >30 from two biological repeats, well-defined termini, and a length ranging from 20 to 500 nt and generally should not be part of mRNAs. The sRNAs and their 5′ and 3′ ends were visualized and examined in IGV (Robinson et al., 2011). The genome was annotated with 48 tRNA genes from GtRNAdb (Chan and Lowe, 2016), but tRNA-Val(TAC-2) and tRNA-Leu(CAA-2) were pseudogenes with low model scores and showed no expression in our sRNA-seq data.

C/D RNAs were identified based on the conserved sequence motifs (box C/C′: RUGAUGA, box D/D′: CUGA) and a length range of 50–103 nt. The presence of the first UGA in box C/C′ and the GA in box D/D′ was required, but one box motif could be mutated. Antisense RNAs were identified by their complementarity to other assigned genes. All transcripts without annotations were aligned in nucleotide-to-protein mode to the NCBI Non-redundant protein database using BLAST. RNAs that were aligned with >80% sequence and had an *E*-value<1×10^{-5} or an identity>95% were considered hCDS. Read 1 of properly paired reads were counted by BEDTools (Quinlan and Hall, 2010). C/D RNAs were named as sR1 to sR60, and sRNAs of other types were named starting from sR101.

Homology analysis of C/D RNAs

As the list of C/D RNAs was incomplete in *S. solfataricus* (Omer et al., 2000; Tang et al., 2005; Zago et al., 2005), its genome sequence (NC_002754.1) was used for alignment. The previously reported C/D RNA list of *S. acidocaldarius* was used for alignment (Daume et al., 2017; Dennis et al., 2015). The *S. islandicus* C/D RNA sequences were aligned by BLASTN with a word size of 7 and an *E*-value cutoff of 1×10^{-4} for *S. solfataricus* or 1×10^{-3} for *S. acidocaldarius*. Two spacers were considered homologous if they shared at least 7 continuous identical sequences.

RiboMeth-seq

RiboMeth-seq was conducted and processed as previously described (Wu et al., 2021). Total RNA was extracted from 100 mL of cells grown to A_{600}=0.2–0.4 using the Trizol method. Unmodified rRNAs were *in vitro* transcribed using T7 RNA polymerase and purified by 8 mol L^{-1} urea polyacrylamide gel electrophoresis (PAGE). The 16S and 23S rRNA genes were PCR-amplified from *S. islandicus* genomic DNA and cloned downstream of a T7 promoter in plasmids pEASY-Blunt or pBCSK. The templates for *in vitro* transcription were amplified from these plasmids. The template for 5S rRNA was directly amplified from the genomic DNA using a forward primer containing a T7 promoter sequence. Individual libraries were prepared for *in vitro* transcribed 16S, 23S and 5S rRNAs. To specifically analyze small

RNAs, 150 μg of total RNA was separated in an 8% urea PAGE. Gel bands corresponding to 30–100 nt RNAs were excised, chopped into pieces, and incubated in extraction buffer (300 mmol L^{-1} NaAc, pH 5.2, 1 mmol L^{-1} EDTA, and 0.25% SDS) at 22°C for 13 h. RNA was then precipitated using isopropanol. The purified RNA was dissolved in 50 mmol L^{-1} sodium carbonate/bicarbonate buffer (pH 9.2) and fragmentated by heating at 95°C for 40 min.

New formula of MethScore

In the RiboMeth-seq analysis, paired-end reads were aligned to individual RNA sequences. The 5′ end count was determined based on the 5′ end of read 1 and shifted upstream by 1 nt along the direction of RNA sequence. The 3′ end count was determined based on the 5′ end of read 2. For tRNA reads, the reads aligned m times were counted with a weight of $1/m$.

The MethScore S was determined by the formula $S = 1 - \frac{C_n}{B}$, where C_n is the end count at position n, and B is the background level of end counts around that position (Birkedal et al., 2014). The original formula combined the 5′ and 3′ end counts and used a weighted average of flanking 12 residues (weight=1, 0.9, 0.8, 0.7, 0.6 and 0.5 for positions $n\pm i$, i=1 to 6) to determine the background. In principle, MethScore can be calculated from 5′ or 3′ end counts alone or their combination and B can be determined from points at 5′ or 3′ sides of position n alone or from both sides. End counts could vary by>100-fold between adjacent positions in tRNAs (Table S4 in Supporting Information), causing a problem for calculating MethScore for transition regions. To reduce false positives, we revised the formula of MethScore. Basically, the 5′ and 3′ end counts are processed separately. For each end count, backgrounds at 5′ and 3′ sides are separately calculated, normalized, and combined.

The 5′ or 3′ end counts, C, are processed separately. The weighted averages of the left and right backgrounds, B_l and B_r, are determined separately from a 3-nt region upstream or downstream of position n:

$$B_l = \frac{\sum_{i=1}^{3} w_i \times C_{n-i}}{\sum_{i=1}^{3} w_i},\ B_r = \frac{\sum_{i=1}^{3} w_i \times C_{n+i}}{\sum_{i=1}^{3} w_i},$$

where the weight w_i=1, 0.8 and 0.4 for i=1 to 3.

To balance the left and right background, the MethScore of one track of end count is calculated as:

$$S = 1 - \frac{C_n + \frac{B_r}{B_l} \times C_n}{B_r + \frac{B_r}{B_l} \times B_l} = 1 - \frac{C_n}{\frac{2B_r B_l}{B_r + B_l}} = 1 - \frac{C_n}{B_m},$$

where the left and right backgrounds are mixed into $B_m = \frac{2B_r B_l}{B_r + B_l}$.

The final MethScore is determined from the 5′ and 3′ end counts at position n and their mixed backgrounds:

$$S = 1 - \frac{C_n^{5'} + C_n^{3'}}{B_m^{5'} + B_m^{3'}}.$$

If one side of the background points is missing, such as in terminal sites, only the background from the available side is used. If one track of end count has a low coverage (⩾3 points have<20 end counts in a 6-nt background), it is not used. MethScore is considered undetermined (set as −10) if both end count tracts around a position have insufficient coverage.

All high score sites were also manually inspected to analyze the distribution of 5′ and 3′ end counts. Any sites with artificially high scores were marked with "F" and excluded for analysis. For rRNA sites (MethScore>0.7), those with MethScores that did not significantly (P>0.05, one-side t-test) different from the MethScores of unmodified 16S and 23S rRNA (based on three measurements), or had a difference of <0.2 compared to unmodified 5S rRNA (one measurement), were marked with "B" and considered unmodified. High MethScores that were based on only one track of end counts were marked with "L". Reliable MethScores were labeled with "T". For analysis of tRNAs and other sRNAs, three datasets of WT total RNAs and two datasets of sRNAs were pooled.

Conservation analysis of Nm in rRNAs

The rRNA sequences were aligned using the MUSCLE program in Jalview. The Nm in rRNAs of *Saccharomyces cerevisiae* (Birkedal et al., 2014), *Arabidopsis* (Wu et al., 2021), and human (Taoka et al., 2018) were compared.

Assignment of C/D RNA to the detected Nm

For each detected methylation site, a 14-nt sequence flanking the site (from position n–6 to n+7, where n is the modification site), was extracted. These target sequences were paired with the spacer sequences of all C/D RNAs using the RNAup program (Lorenz et al., 2011), with parameters set as "-b -d2 -T 37". In cases where the output spacer-substrate duplexes were trimmed by RNAup, they were extended to position 1 at the 3′ end of the spacers and by 1 nt at the 5′ end of the spacers. All duplexes that have at least 6 WC or GU pairs, at most 1 mismatch or bulge, a target at positions 4–7 and a ΔG<−5 kcal mol^{-1} were listed in the "Potential" column and sorted based on their free energy. These duplexes were scored based on the number of WC pairs (2 scores), GU pairs (0.5 scores), bulges (−1 scores), and mismatches (−2 scores). The most reasonable duplex was selected to assign a guide. Generally, spacer-substrate duplexes needed to contain at least 8 WC pairs, 0–2 GU pairs, no mismatch, and no bulge. However, in cases of dual-targeting, weak duplexes were allowed. A guide targeting a non-canonical site was only assigned if validated by biochemical or genetic assays.

In addition, for each identified Nm site in rRNAs, the flanking 100-nt sequences were aligned to complementary sequences of C/D RNAs by BLASTN. This allowed for the identification of primary interactions at the target sites as well as extra interactions at nearby sequences (Wu et al., 2021).

Predicting Nm from C/D RNA sequences

The spacer sequences ranging from positions 2–11 were extracted from C/D RNAs and paired to a 10-nt sequence window that moved along the entire length of the RNA of interest. The number and location of WC pairs, GU wobble pairs, and mismatches were counted for each resulting 10-bp duplexes. Among these spacer-substrate duplexes, only those that have a minimum of 8 WC pairs, 0–2 GU pairs, and no mismatch were selected. The nucleotides at position 5 were predicted as potential

https://doi.org/10.1007/s11427-022-2412-3

methylation sites.

In vitro modification assay

C/D RNAs were prepared by *in vitro* transcription. Substrate RNAs were synthesized by TaKaRa (Japan). The assembly of *S. solfartaricus* C/D RNPs and the methylation activity assay were performed as previously described (Lin et al., 2011; Yang et al., 2020; Ye et al., 2009). The C/D RNPs were assembled from C/D RNAs (1 μmol L^{-1}), Nop5/Fib complex (~2 μmol L^{-1}) and L7Ae (3 μmol L^{-1}). The assembled RNPs were mixed with substrates or pre-methylated substrates (30 μmol L^{-1}) in the presence of 30 μmol L^{-1} cold *S*-adenosyl methionine (SAM) and trace amounts of [methyl-^{3}H] SAM and incubated for 20 min at 70°C. After modification, the RNAs were resolved in denaturing PAGE and visualized through ^{3}H autoradiography. The gels were exposed to X-ray films for 5–6 d.

Data availability

The raw data of sRNA-seq and RiboMeth-seq have been deposited into the National Genomics Data Center (bigd.big.ac.cn) under GSA accession code CRA011490.

Compliance and ethics

The author(s) declare that they have no conflict of interest.

Acknowledgement

This work was supported by the Strategic Priority Research Program of Chinese Academy of Sciences (XDB0570000, XDB37010201), the Basic Research Program Based on Major Scientific Infrastructures of Chinese Academy of Sciences (JZHKYPT-2021-05), the National Natural Science Foundation of China (91940302, 91540201, 31430024, 31325007) and the National Key Research and Development Program of China (2017YFA0504600). We thank Qunxin She for providing *S. islandicus* strains, Li Huang for help in *S. islandicus* culturing, Hongjie Zhang for help in radioactivity experiments and Xiuling Gao for technical assistance.

Supporting information

The supporting information is available online at https://doi.org/10.1007/s11427-022-2412-3. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

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. Nat Struct Mol Biol 10, 256–263.

Benitez-Paez, A., Villarroya, M., Douthwaite, S., Gabaldon, T., and Armengod, M.E. (2010). YibK is the 2′-*O*-methyltransferase TrmL that modifies the wobble nucleotide in *Escherichia coli* $tRNA^{Leu}$ isoacceptors. RNA 16, 2131–2143.

Bernick, D.L., Dennis, P.P., Lui, L.M., and Lowe, T.M. (2012). Diversity of antisense and other non-coding RNAs in archaea revealed by comparative small RNA sequencing in four *Pyrobaculum* species. Front Microbio 3, 231.

Birkedal, U., Christensen-Dalsgaard, M., Krogh, N., Sabarinathan, R., Gorodkin, J., and Nielsen, H. (2014). Profiling of ribose methylations in RNA by high-throughput sequencing. Angew Chem Int Ed 54, 451–455.

Boccaletto, P., Machnicka, M.A., Purta, E., Piątkowski, P., Bagiński, B., Wirecki, T.K., de Crecy-Lagard, V., Ross, R., Limbach, P.A., Kotter, A., et al. (2018). MODOMICS: a database of RNA modification pathways. 2017 update. Nucleic Acids Res 46, D303–D307.

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-tRNATrp. Nucleic Acids Res 31, 6524–6535.

Bruenger, E., Kowalak, J.A., Kuchino, Y., McCloskey, J.A., Mizushima, H., Stetter, K. O., and Crain, P.F. (1993). 5S rRNA modification in the hyperthermophilic archaea *Sulfolobus solfataricus* and *Pyrodictium occultum*. FASEB J 7, 196–200.

Cao, Y., Wang, J., Wu, S., Yin, X., Shu, J., Dai, X., Liu, Y., Sun, L., Zhu, D., Deng, X.W., et al. (2022). The small nucleolar RNA SnoR28 regulates plant growth and development by directing rRNA maturation. Plant Cell 34, 4173–4190.

Cavaille, J., and Bachellerie, J.P. (1998). SnoRNA-guided ribose methylation of rRNA: structural features of the guide RNA duplex influencing the extent of the reaction. Nucleic Acids Res 26, 1576–1587.

Cavaille, J., Nicoloso, M., and Bachellerie, J.P. (1996). Targeted ribose methylation of RNA *in vivo* directed by tailored antisense RNA guides. Nature 383, 732–735.

Chan, P.P., and Lowe, T.M. (2016). GtRNAdb 2.0: an expanded database of transfer RNA genes identified in complete and draft genomes. Nucleic Acids Res 44, D184–D189.

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.

Daume, M., Uhl, M., Backofen, R., and Randau, L. (2017). RIP-Seq suggests translational regulation by L7Ae in archaea. mBio 8, e00730-17.

Deng, L., Zhu, H., Chen, Z., Liang, Y.X., and She, Q. (2009). Unmarked gene deletion and host-vector system for the hyperthermophilic crenarchaeon *Sulfolobus islandicus*. Extremophiles 13, 735–746.

Dennis, P.P., Tripp, V., Lui, L., Lowe, T., and Randau, L. (2015). C/D box sRNA-guided 2′-*O*-methylation patterns of archaeal rRNA molecules. BMC Genomics 16, 632.

Guo, L., Brügger, K., Liu, C., Shah, S.A., Zheng, H., Zhu, Y., Wang, S., Lillestøl, R.K., Chen, L., Frank, J., et al. (2011). Genome analyses of Icelandic strains of *Sulfolobus islandicus*, model organisms for genetic and virus-host interaction studies. J Bacteriol 193, 1672–1680.

Gupta, R. (1984). Halobacterium volcanii tRNAs. Identification of 41 tRNAs covering all amino acids, and the sequences of 33 class I tRNAs. J Biol Chem 259, 9461–9471.

Hori, H., Kawamura, T., Awai, T., Ochi, A., Yamagami, R., Tomikawa, C., and Hirata, A. (2018). Transfer RNA modification enzymes from thermophiles and their modified nucleosides in tRNA. Microorganisms 6, 110.

Karijolich, J., and Yu, Y.T. (2010). Spliceosomal snRNA modifications and their function. RNA Biol 7, 192–204.

Kim, D., Langmead, B., and Salzberg, S.L. (2015). HISAT: a fast spliced aligner with low memory requirements. Nat Methods 12, 357–360.

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.

Kiss, T. (2001). Small nucleolar RNA-guided post-transcriptional modification of cellular RNAs. EMBO J 20, 3617–3622.

Lapeyre, B., and Purushothaman, S.K. (2004). Spb1p-directed formation of Gm2922 in the ribosome catalytic center occurs at a late processing stage. Mol Cell 16, 663–669.

Li, Y., Pan, S., Zhang, Y., Ren, M., Feng, M., Peng, N., Chen, L., Liang, Y.X., and She, Q. (2016). Harnessing Type I and Type III CRISPR-Cas systems for genome editing. Nucleic Acids Res 44, e34.

Lin, J., Lai, S., Jia, R., Xu, A., Zhang, L., Lu, J., and Ye, K. (2011). Structural basis for site-specific ribose methylation by box C/D RNA protein complexes. Nature 469, 559–563.

Lorenz, R., Bernhart, S.H., Höner zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P. F., and Hofacker, I.L. (2011). ViennaRNA package 2.0. Algorithms Mol Biol 6, 26.

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

Marchand, V., Blanloeil-Oillo, F., Helm, M., and Motorin, Y. (2016). Illumina-based RiboMethSeq approach for mapping of 2′-*O*-Me residues in RNA. Nucleic Acids Res 44, e135.

Marchand, V., Pichot, F., Thüring, K., Ayadi, L., Freund, I., Dalpke, A., Helm, M., and Motorin, Y. (2017). Next-generation sequencing-based RiboMethSeq protocol for analysis of tRNA 2′-*O*-methylation. Biomolecules 7, 13.

Noon, K.R., Bruenger, E., and McCloskey, J.A. (1998). Posttranscriptional modifications in 16S and 23S rRNAs of the archaeal hyperthermophile *Sulfolobus solfataricus*. J Bacteriol 180, 2883–2888.

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.

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.

Peng, N., Han, W., Li, Y., Liang, Y., and She, Q. (2017). Genetic technologies for extremely thermophilic microorganisms of *Sulfolobus*, the only genetically tractable genus of crenarchaea. Sci China Life Sci 60, 370–385.

Pintard, L., Lecointe, F., Bujnicki, J.M., Bonnerot, C., Grosjean, H., and Lapeyre, B. (2002). Trm7p catalyses the formation of two 2′-*O*-methylriboses in yeast tRNA anticodon loop. EMBO J 21, 1811–1820.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842.

Renalier, M.H., Joseph, N., Gaspin, C., Thebault, P., and Mougin, A. (2005). The Cm56 tRNA modification in archaea is catalyzed either by a specific 2′-*O*-methylase, or a C/D sRNP. RNA 11, 1051–1063.

Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat Biotechnol 29, 24–26.
Roehr, J.T., Dieterich, C., and Reinert, K. (2017). Flexbar 3.0-SIMD and multicore parallelization. Bioinformatics 33, 2941–2942.
Somme, J., Van Laer, B., Roovers, M., Steyaert, J., Versées, W., and Droogmans, L. (2014). Characterization of two homologous 2′-*O*-methyltransferases showing different specificities for their tRNA substrates. RNA 20, 1257–1271.
Tang, T.H., Polacek, N., Zywicki, M., Huber, H., Brugger, K., Garrett, R., Bachellerie, J. P., and Hüttenhofer, A. (2005). Identification of novel non-coding RNAs as potential antisense regulators in the archaeon *Sulfolobus solfataricus*. Mol Microbiol 55, 469–481.
Taoka, M., Nobe, Y., Yamaki, Y., Sato, K., Ishikawa, H., Izumikawa, K., Yamauchi, Y., Hirota, K., Nakayama, H., Takahashi, N., et al. (2018). Landscape of the complete RNA chemical modifications in the human 80S ribosome. Nucleic Acids Res 46, 9289–9298.
Taoka, M., Nobe, Y., Yamaki, Y., Yamauchi, Y., Ishikawa, H., Takahashi, N., Nakayama, H., and Isobe, T. (2016). The complete chemical structure of *Saccharomyces cerevisiae* rRNA: partial pseudouridylation of U2345 in 25S rRNA by snoRNA snR9. Nucleic Acids Res 44, 8951–8961.
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.
Tycowski, K.T., Smith, C.M., Shu, M.D., and Steitz, J.A. (1996). A small nucleolar RNA requirement for site-specific ribose methylation of rRNA in *Xenopus*. Proc Natl Acad Sci USA 93, 14480–14485.
van Nues, R.W., Granneman, S., Kudla, G., Sloan, K.E., Chicken, M., Tollervey, D., and Watkins, N.J. (2011). Box C/D snoRNP catalysed methylation is aided by additional pre-rRNA base-pairing. EMBO J 30, 2420–2430.
van Nues, R.W., and Watkins, N.J. (2016). Unusual C′/D′ motifs enable box C/D snoRNPs to modify multiple sites in the same rRNA target region. Nucleic Acids Res 45, 2016–2028.
Watkins, N.J., and Bohnsack, M.T. (2011). The box C/D and H/ACA snoRNPs: key players in the modification, processing and the dynamic folding of ribosomal RNA. WIREs RNA 3, 397–414.
Wolff, P., Villette, C., Zumsteg, J., Heintz, D., Antoine, L., Chane-Woon-Ming, B., Droogmans, L., Grosjean, H., and Westhof, E. (2020). Comparative patterns of modified nucleotides in individual tRNA species from a mesophilic and two thermophilic archaea. RNA 26, 1957–1975.
Wu, S., Wang, Y., Wang, J., Li, X., Li, J., and Ye, K. (2021). Profiling of RNA ribose methylation in *Arabidopsis thaliana*. Nucleic Acids Res 49, 4104–4119.
Yang, J., Sharma, S., Watzinger, P., Hartmann, J.D., Kotter, P., and Entian, K.D. (2016a). Mapping of complete set of ribose and base modifications of yeast rRNA by RP-HPLC and mung bean nuclease assay. PLoS ONE 11, e0168873.
Yang, Z., Lin, J., and Ye, K. (2016b). Box C/D guide RNAs recognize a maximum of 10 nt of substrates. Proc Natl Acad Sci USA 113, 10878–10883.
Yang, Z., Wang, J., Huang, L., Lilley, D.M.J., and Ye, K. (2020). Functional organization of box C/D RNA-guided RNA methyltransferase. Nucleic Acids Res 48, 5094–5105.
Ye, K., Jia, R., Lin, J., Ju, M., Peng, J., Xu, A., and Zhang, L. (2009). Structural organization of box C/D RNA-guided RNA methyltransferase. Proc Natl Acad Sci USA 106, 13808–13813.
Yu, N., Jora, M., Solivio, B., Thakur, P., Acevedo-Rocha, C.G., Randau, L., de Crecy-Lagard, V., Addepalli, B., and Limbach, P.A. (2019). tRNA modification profiles and codon-decoding strategies in *Methanocaldococcus jannaschii*. J Bacteriol 201, e00690-18.
Yu, Y.T., and Meier, U.T. (2014). RNA-guided isomerization of uridine to pseudouridine-pseudouridylation. RNA Biol 11, 1483–1494.
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.
Ziesche, S.M., Omer, A.D., and Dennis, P.P. (2004). RNA-guided nucleotide modification of ribosomal and non-ribosomal RNAs in Archaea. Mol Microbiol 54, 980–993.
Zillig, W., Kletzin, A., Schleper, C., Holz, I., Janekovic, D., Hain, J., Lanzendörfer, M., and Kristjansson, J.K. (1993). Screening for *Sulfolobales*, their plasmids and their viruses in icelandic solfataras. Syst Appl Microbiol 16, 609–628.