Published online 7 May 2012 *Nucleic Acids Research, 2012, Vol. 40, No. 14* ***6765–6773***
doi:10.1093/nar/gks381

Programmable sequence-specific click-labeling of RNA using archaeal box C/D RNP methyltransferases

Miglė Tomkuvienė[1], Béatrice Clouet-d'Orval[2], Ignas Černiauskas[1], Elmar Weinhold[3] and Saulius Klimašauskas[1,*]

[1]Department of Biological DNA Modification, Institute of Biotechnology, Vilnius University, Vilnius LT-02241, Lithuania, [2]Laboratory of Microbiology and Molecular Genetics, CNRS-UMR 5100, Université Paul Sabatier, Toulouse F-31062, France and [3]Institute of Organic Chemistry, RWTH Aachen University, Aachen D-52056, Germany

Received February 7, 2012; Revised April 12, 2012; Accepted April 13, 2012

ABSTRACT

Biophysical and mechanistic investigation of RNA function requires site-specific incorporation of spectroscopic and chemical probes, which is difficult to achieve using current technologies. We have *in vitro* reconstituted a functional box C/D small ribonucleoprotein RNA 2′-*O*-methyltransferase (C/D RNP) from the thermophilic archaeon *Pyrococcus abyssi* and demonstrated its ability to transfer a prop-2-ynyl group from a synthetic cofactor analog to a series of preselected target sites in model tRNA and pre-mRNA molecules. Target selection of the RNP was programmed by changing a dodecanucleotide guide sequence in a 64-nt C/D guide RNA leading to efficient derivatization of three out of four new targets in each RNA substrate. We also show that the transferred terminal alkyne can be further appended with a fluorophore using a bioorthogonal azide-alkyne 1,3-cycloaddition (click) reaction. The described approach for the first time permits synthetically tunable sequence-specific labeling of RNA with single-nucleotide precision.

INTRODUCTION

RNA plays a multitude of roles in cellular metabolism such as the translation of the genetic code. Although many types of RNA species and their complexes, even such as ribosomes, have been structurally characterized in great detail, their mechanism of action cannot be well understood without the spatial and temporal description of the underlying molecular motions and interactions. Biophysical and mechanistic investigation of RNA function requires site-specific incorporation of spectroscopic and chemical probes, which is difficult and/or restricted to limited types of sites to be modified using current technologies (1–3).

In nature, numerous RNA methyltransferases (MTases) catalyze the transfer of methyl groups from *S*-adenosyl-L-methionine (AdoMet) to specific target nucleotides in RNA. A particular class of RNA MTases, box C/D ribonucleoprotein complexes (C/D RNPs), directs AdoMet-dependent site-specific 2′-*O*-methylation to numerous biological targets in ribosomal, spliceosomal and transfer RNAs [(4–6) reviewed in (7)]. A minimal structural unit of an archaeal C/D RNP comprises of one box C/D guide RNA (C/D RNA or guide RNA) and two copies of the L7Ae, Nop5p and aFib proteins (Figure 1A). The archaeal C/D RNAs are typically 50–70 nt long and contain conserved sequences C (RUG AUGA) and D (CUGA) at their 5′ and 3′-ends, respectively, and internal conserved C′ and D′ boxes. These motifs pair to form C/D and C′/D′ elements that fold into K-turn and K-loop structures, respectively (8,9). Both of them equally serve as a platform for the RNP's core protein binding. The K-structures are first recognized and stabilized by L7Ae, which is then followed by Nop5p and aFib—the methyltransferase, eventually forming a higher RNP dimer (10,11). Precise base pairing of the guide sequences, usually 12 nt long in functional archaeal C/D RNAs (12), upstream of the D or D′ boxes of the C/D RNA with its cognate RNA substrate directs the methyltransferase reaction to the 2′-hydroxyl of a target nucleotide determined by nucleotide at the fifth position upstream from the D or D′ box (13) (Figure 1A). Since the specificity of the reaction is defined by the guide sequence of the C/D RNA, it is possible to target *de novo* any position within an RNA molecule as it was done even *in vivo* in mammalian and yeast cells (14,15).

Although the methyl group is widely used by the cells to modify RNA for stabilization and folding, it is hard to detect and/or to further derivatize, which is needed for

*To whom correspondence should be addressed. Tel: +370 52602114; Fax: +370 52602118; Email: klimasau@ibt.lt

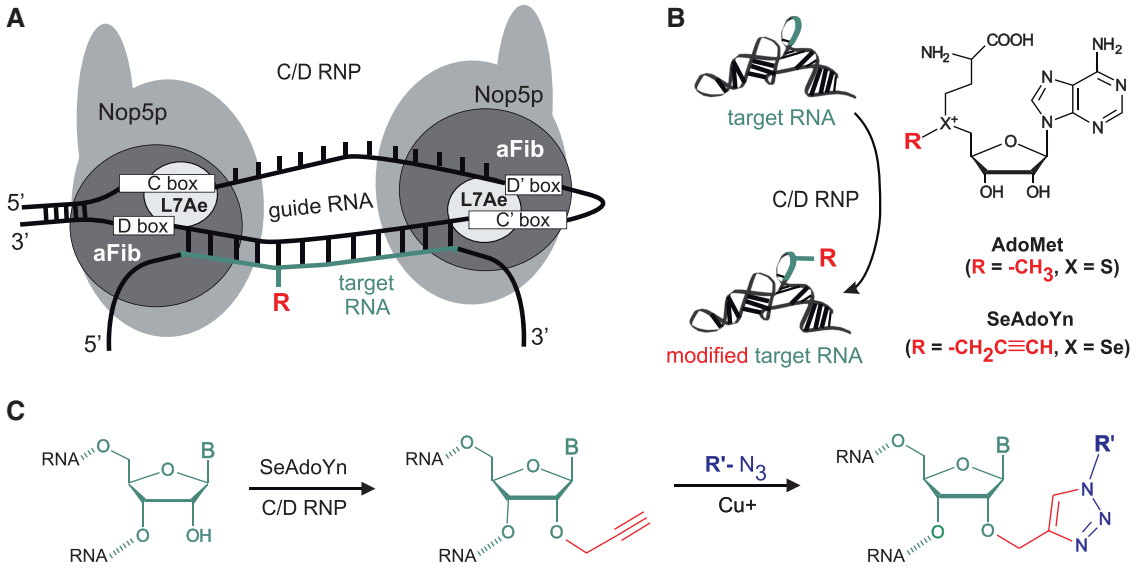


Figure 1. Archaeal C/D RNP-directed sequence-specific modification and labeling of target RNA. (**A**) Schematic structure of a C/D RNP complex with substrate RNA. Core proteins L7Ae, Nop5p and aFib are bound at the C/D and C′/D′ sites of a guide RNA. One of the variable guide sequences is shown base-paired to a target sequence (green) of a substrate RNA. Modification occurs at a nucleotide complementary to the fifth position upstream from the D box. (**B**) C/D RNP-directed transfer of an activated side chain (red) from a cofactor *S*-adenosyl-L-methionine (AdoMet, X = S and R = methyl) or its analog SeAdoYn (X = Se and R = prop-2-ynyl) onto an RNA substrate. (**C**) Two-step 'click' labeling of target RNA via a C/D RNP-directed alkynylation, followed by Cu(I)-assisted 1,3-cycloaddition of a fluorogenic azide derivative (blue).

site-specific RNA labeling in mechanistic and functional studies of RNA. To circumvent poor reporter properties of the methyl group, a series of synthetic AdoMet analogs with sulfonium-bound extended chains have been developed for covalent transfer of the side chains to various biomolecules by AdoMet-dependent methyltransferases (16). A recently devised mTAG (methyltransferase-directed Transfer of Extended Groups) approach permits site-specific functionalization and labeling of DNA (17,18), RNA (19), proteins (20–22) and small molecules (23). In this work, we sought to combine the specificity of C/D RNP machinery with synthetic AdoMet analogs to provide a novel molecular tool for programmable sequence-specific RNA labeling (Figure 1B). This system has the inherent advantage that the target nucleotide can be simply specified by supplying a suitable guide RNA. Therefore, we have *in vitro* reconstituted a functional box C/D RNP from a thermophilic archaeon, *Pyrococcus abyssi*, and demonstrated its ability to site-specifically transfer a selenium-bound alkynyl group from a synthetic cofactor to both the wild-type and newly programmed target sites in RNA molecules. We also show that the transferred group can be further appended with a fluorophore using a bio-orthogonal azide-alkyne 1,3-cycloaddition (click chemistry) reaction (Figure 1C).

MATERIALS AND METHODS

Methyltransferase cofactors

The selenium cofactor SeAdoYn (diastereomeric mixture at the selenonium center) was prepared in two steps starting from 5′-chloro-5′-deoxyadenosine (24). AdoButyn (25) and Ado-11-amine (18) cofactors were prepared as previously described. Unlabeled *S*-adenosyl-L-methionine (AdoMet) was purchased from Sigma.

Expression and purification of proteins

The *P. abyssi* L7Ae, Nop5p and aFib proteins were expressed from recombinant pET15b vectors (26) in *Escherichia coli* BL21 (DE3) RIL (Stratagene) cells induced with 1 mM isopropyl-β-D-thiogalactoside (IPTG). After sonication, bulk of *E. coli* proteins was removed by thermodenaturation at 65°C for 15 min and spinning down. Nucleic acids from L7Ae preparations were removed by 0.15% polyethylenimine (27). His_6-tagged L7Ae was purified individually, and His_6-tagged Nop5p and aFib were co-purified, by Ni^{++} chelate chromatography on a HiTrap IMAC HP column (GE Healthcare) according to manufacturer's recommendations. The His_6-tag was removed by thrombin (Amersham) cleavage. L7Ae preparation was additionally passed over a HiTrap SP-sepharose column (GE Healthcare) according to manufacturer's recommendations. Reconstituted C/D RNP methyltransferase preparations retained ~1 mol% of bound endogenous AdoMet leading to the appearance of detectable methylation products in enzymatic reactions containing synthetic cofactor analogs.

RNA synthesis and purification

All guide and substrate RNAs were produced by *in vitro* transcription using a TranscriptAid T7 High Yield Transcription kit (Fermentas) using the following templates: a linearized pUC18-based plasmid carrying a recombinant sR47 gene (9) (for guide RNAs), or a PCR fragment generated from a corresponding region of

P. abyssi genomic DNA (for tRNA-Leu) or a recombinant rabbit β-globin gene (28) (for β-globin pre-mRNA); PCR primers are listed in Supplementary Table S2. After ethanol precipitation, RNAs were purified on a 6% PAA denaturing gel. Synthetic RNA oligonucleotide S22 (5′-CCACAGACUCAAGAUCCCCUCC) was obtained from Metabion.

Assaying RNP activity toward synthetic cofactor analogs

The approach used substrate oligonucleotide with internally ^{33}P-labeled target nucleotide. Starting unmodified RNA oligonucleotides and DNA splint were purchased from Metabion. RNA oligonucleotide D13 (5′-CAAGA UCCCCUCC) was ^{33}P-labeled at the 5′-end using [γ-^{33}P]-ATP (Hartman Analytic) and T4 polynucleotide kinase (Fermentas). Following phosphorylation the reaction mixture was adjusted to meet the conditions of ligation, and the labeled D13 together with RNA oligonucleotide D11 (5′-GCGGCGGCACU) was annealed on a splint DNA oligonucleotide (5′-GGAGGGGATCTTG AGTGCCGCCGC) and were ligated by T4 DNA ligase (Fermentas), treated with DNaseI (Fermentas), and the resulting 24 nt RNA strand was purified on a 6% PAA gel. For the modification reaction, the RNP was assembled of 1 μM sR47, 2 μM L7Ae and 2 μM Nop5p and aFib in methylation buffer (20 mM Hepes-NaOH, pH 7.9, 150 mM NaCl and 10 mM $MgCl_2$) by incubation at 65°C for 10 min. Following sRNP assembly, AdoMet (50 μM), or a synthetic cofactor analog (200–600 μM), and the internally ^{33}P-labeled substrate oligonucleotide (~0.01 μM) supplemented with substrate oligonucleotide S22 (0.5 μM) were added. The reaction mix was incubated at 65°C 20–60 min. RNA was digested to nucleoside-5′-monophosphates by addition of nuclease Bal31 (Fermentas) and analyzed by one-dimensional TLC (29). TLC plates were autoradiographed to phosphorimager screens and analyzed with a FLA-5100 scanner and MultiGauge software (Fujifilm).

Guide reprogramming

pUC18 vector, carrying a sR47 sequence and accordingly positioned T7 RNA pol. promoter, was digested with BamHI and Eco81I restriction endonucleases (Fermentas) to remove a 38-bp fragment encompassing D-guide sequence and leave sticky ends for further insertions. The vector was dephosphorylated by FastAp (Fermentas) and purified on a 1% agarose gel using GeneJET Gel Extraction Kit (Fermentas). Synthetic oligonucleotides (Supplementary Table S1) were designed to form duplexes with sticky ends to restore a full guide RNA gene in the pUC18 vector incorporating a desired D-guide sequence. Oligonucleotide pairs were annealed in water, phosphorylated by T4 Polynucleotide Kinase (Fermentas) and ligated into the vector at 17°C overnight. The ligate was transformed into *E. coli* ER2267 and the new guide RNA gene sequences confirmed by sequencing. The guide RNAs were further transcribed as described above.

HPLC-MS analysis

Modification reactions for HPLC analysis typically contained 1 μM preassembled RNP, 1 μM tRNA-Leu and 50 μM AdoMet or 400 μM SeAdoYn, if any. After incubation at 68°C for 30 min, RNA (containing equimolar amounts of modified substrate and guide RNA) was purified on ZR RNA MicroPrep™ columns (Zymo Research) and digested to nucleosides by incubation with nuclease P1 (0.5 units, Sigma) for 2 h at 55°C and then with FastAP phosphatase (0.5 units, Fermentas) and nuclease P1 (0.25 units) overnight at 37°C. RNA hydrolyzate (40 pmol) was loaded onto an integrated HPLC/ESI-MS Agilent 1200 series system equipped with a Discovery C18 column (75 × 2.1 mm, Supelco) and eluted with a linear gradient of solvents A (20 mM ammonium formate, pH 3.5) and B (80% aqueous methanol) at a flow of 0.3 ml/min at 30°C as follows: 0–20 min, 0–20% B; 20–22 min, 20–100% B. High-resolution mass spectra of modification products were acquired on a Q-TOF 6250 mass spectrometer (Agilent) equipped with a Dual-ESI source. For quantitative analysis of the alkylation and methylation activities, UV-chromatograms were sample size normalized based on the peak area of the Ado nucleoside followed by measurement of the peak areas of the corresponding propynylated and methylated nucleosides.

Two-step labeling of RNA

Modification reaction

RNP (1 μM) was assembled by combining 1 μM guide RNA, 3 μM L7Ae and 2.5 μM Nop5p and aFib in methylation buffer (20 mM Hepes-NaOH, pH 7.9, 150 mM NaCl and 10 mM $MgCl_2$) and pre-incubation at 68°C for 7 min. Afterwards the synthetic cofactor analog SeAdoYn (400 μM) and 1 μM substrate RNA (and 10 μM S22 competitor, if any) were added, and the reaction was incubated at 68°C for 30 min. The modification was stopped by adding an equal volume of RNA loading dye (Fermentas). The resulting mixture was stored at −20°C.

Labeling reaction

Reaction conditions were adapted from (30) and manufacturer's recommendations (www.baseclick.eu). The reaction modification mixture (40 μl; 20–220 pmol of substrate RNA) was diluted with 0.8 volume (32 μl) of DMSO/t-BuOH 3:1 (Baseclick). 0.03 M CuBr (Sigma-Aldrich) was freshly prepared each time in 0.1 M tris(benzyltriazolylmethyl)amine (TBTA) (Sigma-Aldrich) (in DMSO/t-BuOH), and 12 μl of Cu-TBTA solution was added to the mixture followed by 8 μl of 2.5 mM (20 nmol; in DMSO/t-BuOH 3:1) Eterneon-480/635 azide (Baseclick). Reaction mixture was incubated for 1 h at 37°C with intermittent vortexing. Modified RNA was purified using ZR RNA MicroPrep columns (Zymo Research) and analyzed on a 15% PAA denaturing gel. Eterneon fluorescence was monitored on a FLA-5100 (Fujifilm) fluorescent image analyzer using a 473 nm laser. Gels were then stained with ethidium bromide for bulk RNA imaging.

Reverse transcription analysis

Reverse transcription analysis was done following a described procedure (31). The following primers (Metabion) were used: RT_tLeu2 5′-ACCCCTACGGGAGGGGA for A31 and RT_tLeu3 5′-TCTTGAGTCCCCCGCCTT for U17a. 1.4 μM ^{33}P-5′-end-labeled primer was hybridized to 0.4 μM unmodified or C/D RNP guide Eterneon-labeled tRNA-Leu. Extension step was performed by addition of a mixture containing 400 μM or 4 μM dNTP, 1/14 of the total reaction volume of RevertAid Premium Enzyme Mix (Fermentas) and incubation at 55°C for 30 min, followed by 5 min at 85°C. Sequencing reactions were done in parallel and were identical to the label-mapping ones with 4 μM dNTP, except having 0.7 μM unmodified tRNA-Leu in the hybridization step and 0.2 of the total reaction volume of Termination Mix G/A/T/C (CycleReader DNA Sequencing Kit, Fermentas). After ethanol precipitation, DNA was resolved on a denaturing 15% PAA gel and visualized by radioautography.

RESULTS

The Pyrococcus abyssi box C/D RNP uses SeAdoYn as a cofactor

A previously well characterized *in vitro* C/D guide RNP system from the thermococcal archaeon *P. abyssi* was chosen as a model for the RNA-guided RNA modification experiments (9,26). It has been previously shown that the sequence upstream of the D box of the C/D RNA sR47 guides the methylation of cytosine at position 34 in the anticodon loop of the *P. abyssi* tRNA-Leu(CAA) (tRNA-Leu or simply tRNA further here) (9). The three pyrococcal C/D RNP proteins, L7Ae, Nop5p and aFib, were expressed in *E. coli* and His-Tag-purified. The activity of the *in vitro* reconstituted sR47 RNP was confirmed by assaying tritium incorporation into RNA in the presence of the tritiated AdoMet cofactor with *in vitro* transcribed *P. abyssi* tRNA-Leu (Supplementary Figure S1). Subsequently, we verified that methyl groups have been transferred on the target cytosine of RNA substrate oligonucleotide by thin layer chromatography (TLC) analysis of the digested RNA products (see below, Figure 2, Lane 1). To achieve transfer of extended reactive groups onto the RNA substrates, we examined several synthetic cofactor analogs for their ability to replace AdoMet in the RNP-directed reaction (Figure 1B). For this, we constructed a 24-mer RNA substrate containing a sequence complementary to the D guide region of sR47 and a 5′-^{33}P-labeled nucleotide at the target position. After enzymatic digestion of the modified RNA, TLC analysis (29) permitted selective determination of the modification status of the radiolabeled target nucleotide. Our analyses showed that none of the synthetic cofactors with side chains larger than 5 carbon units rendered a detectable modification of the RNA substrate (see examples in Table 1 and Supplementary Figure S2). It turned out that a newly designed synthetic cofactor, SeAdoYn (Figure 1B), carrying a selenium-bound three carbon chain (24) was well accepted by the archaeal methyltransferase aFib in the RNP guide reaction (Figure 2, Lane 2, and Supplementary Figure S2B). The transfer of the propynyl group on the cytosine was dependent on the C/D RNP enzyme (Figure 2, Lane 3).

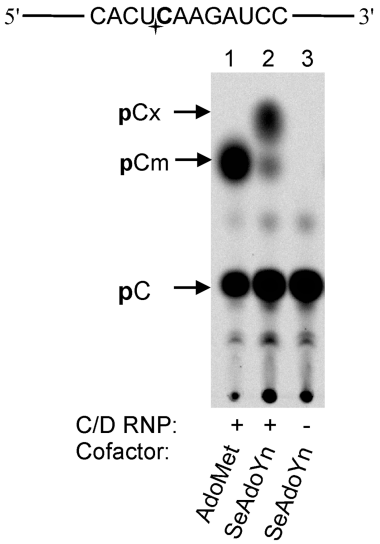
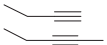
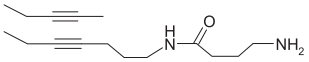


Figure 2. Formation of modified nucleotides in a 24-mer RNA oligonucleotide incubated with the C/D sR47 RNP in the presence of cofactors AdoMet and SeAdoYn. The target sequence of the substrate oligonucleotide (top) contains a ^{33}P-labeled target nucleotide (shown in bold and an asterisk). Substrate oligonucleotide (1 μM) was incubated with 1 μM pre-assembled RNP and 50 μM AdoMet or 400 μM SeAdoYn for 40 min at 65°C. Samples were subjected to nuclease Bal31 digestion and TLC analysis of ^{33}P-labeled mononucleotides. Arrows point to target nucleotide products formed in the presence of cofactors AdoMet (lane 1, constitutes 45% of total nucleotide counts) or SeAdoYn (lane 2, constitutes 5% of total nucleotide counts). A control with SeAdoYn was carried in the absence of the Nop5p-aFib RNP core proteins (Lane 3).

Reprogramming RNP-directed alkylations to defined sites in tRNA

A series of guide RNAs were designed to target selected positions in tRNA-Leu: U17a in the D loop, A31 in the anticodon stem, Ae3 in the variable loop and C56 in the Ψ loop (Figure 4A) by changing the 12-nt guide sequence in the sR47 guide RNA. We analyzed the ability of the C/D RNPs containing the new guide RNA variants to transfer methyl groups in the presence of the tritiated AdoMet cofactor onto the tRNA-Leu substrate. The extent of target methylation with the new guide RNAs was comparable to that of the wt sR47 (Supplementary Figure S1A) except for the A31 guide, which showed a considerably lower incorporation of methyl groups into the substrate RNA. The latter observation is consistent with its target nucleotide location in the anticodon stem, as opposed to hairpin loop locations of the other target nucleotides. The specificity of the reaction was then confirmed by nucleoside composition analysis of methylated tRNA-Leu using HPLC-MS, which showed incorporation of a methyl group at the target nucleosides (Figure 3). To determine

Table 1. *In vitro* modification of an internal ^{33}P-labeled target nucleotide (corresponding to C34 in tRNA-Leu) in a 24-mer fragment of tRNA-Leu by the wild-type *P. abyssi* C/D RNP with AdoMet or its synthetic analogs

Cofactor	Onium center (X^+)	Transferable group (R)	Reactive functionality	Modification of the target nucleotide (%)
AdoMet	S	$-CH_3$	None	45–97
SeAdoYn	Se		Terminal alkyne	5–10
AdoButyn	S		None	2–5
Ado-11-amine	S		Primary amine	0

Modification efficiency was determined by TLC analysis of ^{33}P-labeled 5′-mononucleotides as shown in Figure 2 and Supplementary Figure S2.

if the C/D RNP methyltransferase can be used as a tool to incorporate specific labels on RNA molecule at defined specific positions, the C/D RNP variants were queried in the presence of the SeAdoYn cofactor. The HPLC-MS analysis of nucleoside composition of modified tRNA revealed new compounds with similar UV spectra but longer retention times than the methylated nucleosides (Figure 3), whereas a control sample in the absence of the functional MTase (omission of the Nop5 and aFib proteins) revealed no detectable modified nucleosides. The mass spectra of the modification products were in good agreement with the theoretical masses of corresponding propynylated target nucleosides (Supplementary Table S3). In all cases, only one type of propynylated nucleoside was detectable (note that in all the cases adjacent 3' and 5' nucleotides were different from a target nucleotide itself) consistent with a highly faithful C/D RNP-directed reaction. The extent of propynylation was determined by weighing areas of corresponding peaks in normalized UV chromatograms to those of methylation products. The extent of propynylation reached a maximum of ~30% of the methylation level, i.e. 5% at the wt (C34) target, 33% at U17a, 24% at A31, 16% at Ae3, and no product was observed at C56. The observed extent of alkylations may be somewhat underestimated due to less efficient enzymatic hydrolysis of the 2′-*O*-propyn-modified RNAs as compared with methylated RNA by nuclease P1 which is used for production of monomeric RNA constituents prior to HPLC analysis. However, a slight retardation of the transalkylation reaction could well be expected due to steric hindrance of the bulkier transferable group. Altogether, these results show that the C/D RNP-dependent 2′-*O*-propynylation can be targeted to newly defined RNA sites by reprogramming the guide RNA.

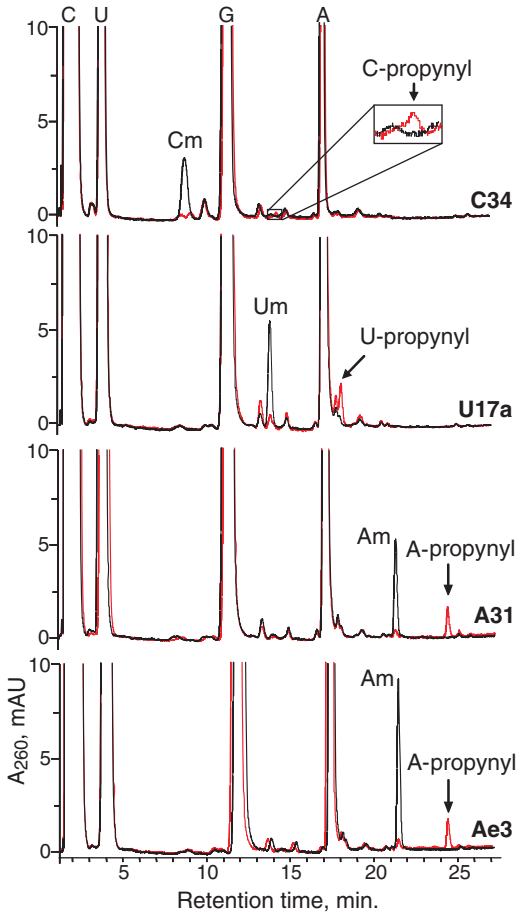


Figure 3. HPLC-MS analysis of products obtained after C/D RNP-directed modification of a model tRNA substrate with four variants of sR47 guide RNA (targeting C34, U17a, A31, Ae3 positions as indicated) in the presence of the AdoMet or SeAdoYn cofactor. Pre-assembled C/D RNP (1 μM), tRNA (1 μM) and AdoMet (50 μM) or SeAdoYn (400 μM) were incubated at 68°C for 30 min. Purified RNA products were enzymatically digested to nucleosides and analyzed by reversed-phase HPLC. Sample size-normalized A_{260} traces of HPLC chromatograms are shown for the methylation (black) and propynylation (red) reactions.

Fluorescent labeling of tRNA at programmed sites

The next step was to show that the derivatized nucleotides in the tRNA could be appended with a fluorophore *via* click chemistry. The click reaction was performed without purification (in the presence of all components from the propynylation reaction), and the RNA was afterward recovered in a single column-purification step. The tRNAs modified by the C/D RNP variants showed a clear fluorescence signal (Figure 4B). The fluorescent signal is only observed on the RNA species corresponding to the tRNA (substrate) and not to the guide RNA (part of the RNP enzyme). No fluorescence signal was detectable in control reactions without cofactor or guide RNA (Figure 4B). Thus, the fluorescent label was only attached to tRNAs that have been modified through a C/D RNP-dependent reaction. Notably, the intensity of fluorescent labeling at three newly programmed targets (U17a, A31 and Ae3) was stronger than at the wild-type site (C34), and the brightest fluorescence observed at Ae3 exceeded the C34

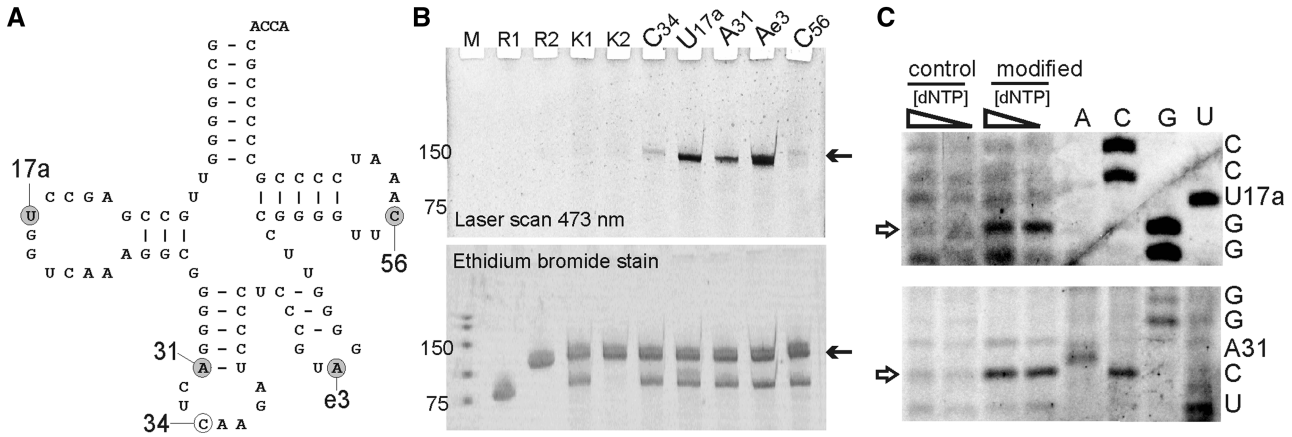


Figure 4. C/D RNP-dependent labeling of predetermined sites in a model tRNA substrate: (**A**) Schematic representation of *P. abyssi* tRNA-Leu(CAA) (41). Position of target nucleotide for the wt sR47 RNA (C34) is shown as an empty circle, and positions of newly programmed target nucleotides (U17a, A31, Ae3 and C56) are shown as filled circles. (**B**) Fluorescent sequence-specific labeling of tRNA via guide RNA-directed enzymatic propynylation and copper-assisted coupling of an Eterneon(480/635) azide. The reactions were resolved on a denaturing polyacrylamide gel, scanned with a 473 nm laser for Eterneon fluorescence (upper panel), and then stained with ethidium bromide to reveal bulk RNA (lower panel). M, DNA marker; R1, guide sR47 RNA; R2, unmodified tRNA; K1, control reaction without cofactor; K2, control reaction in the absence of C/D RNA. Bands corresponding to tRNA are shown by arrows. (**C**) Reverse transcription (RT) mapping of RNP modification sites. tRNA-Leu was modified in the presence of a C/D RNP and SeAdoYn cofactor followed by click coupling of an Eterneon fluorophore, as in **B**. RT primer extension analysis shows polymerase halting (arrows) one position ahead of the target (U17a or A31) nucleotides in samples labeled using corresponding guide RNAs.

signal by a factor of 7. The extent of alkylation of a target nucleotide observed by HPLC-MS was roughly proportional to the efficiency of fluorescent labeling, indicating consistent performance at the click chemistry step.

To assess the fidelity of the reprogrammed RNPs, we performed a competition assay by addition of a 10-fold molar excess of a 22-mer RNA probe in the reaction. The RNA competitor includes a dodecanucleotide stretch identical to the C34 target sequence (complementary to the upstream of the box D sequence of sR47). Under these conditions (Supplementary Figure S3), the tRNA substrate was no longer modified using the wt guide, but the label was transferred onto the competitor 22-mer RNA probe. In contrast, the guide RNA variants designed to target positions U17a and Ae3 were not affected by the addition of the RNA competitor. In the case of the A31 guide RNA, the efficiency of the tRNA-Leu labeling was reduced, but labeling of the competitor was not detectable. The latter observation could be explained by a partial overlap (9 bp) of the A31 guide RNA with the competing oligonucleotide. Altogether one can conclude that perfect complementarity of the guide sequence to its target site in RNA is required for the propynylation reactions. The positions of the RNP-directed labels were further confirmed by reverse transcription (RT) primer extension analysis. The best fluorescently labeled tRNAs (U17a, A31 or Ae3) were subjected to RT reactions along with sequencing reactions on an unmodified tRNA-Leu (Figure 4C). In the case of U17a and A31 labeled tRNAs, the primer elongation stops were clearly seen one position ahead of the target nucleotide, thus confirming the presence of modified nucleotides at the targeted positions. Modification at the third position could not be confirmed by these experiments due to inability of the RT to produce a sequencing ladder around Ae3 in control tRNA.

Programmed fluorescent labeling of a model pre-mRNA

We further attempted to label an even larger RNA substrate that was unrelated to any known targets of archaeal C/D RNP methyltransferases. For this purpose, we chose a 357-nt fragment of rabbit β-globin pre-mRNA, encompassing the first exon and intron and a part of the second exon. Four sites were chosen as targets: two in the first exon (A102 and U168), one in the intron (C234) and one on the second exon near the intron-exon boundary (G308). Appropriate guide RNAs were generated as described above. and the C/D RNP methylation activity with pre-mRNA was shown to be comparable to that of the wt guide-substrate pair, sR47 and tRNA-Leu, in all cases (Supplementary Figure S1B). The two-step fluorescent labeling was also successful for all four targeted sites (Supplementary Figure S4), although occurred with different intensities, as observed with the tRNA-Leu substrate. Finally, the same modification and labeling reactions were performed in the presence of both substrates, tRNA-Leu and pre-mRNA, using five guide RNAs targeting specific positions in pre-mRNA or tRNA (Figure 5). The presented results convincingly show that the C/D RNP-directed labeling is targeted to a programmed RNA substrate.

DISCUSSION

Here we demonstrate the first chemo-enzymatic system for synthetically programmable sequence-specific covalent functionalization and labeling of RNA molecules *in vitro*. In the first step, a prop-2-ynyl group is transferred

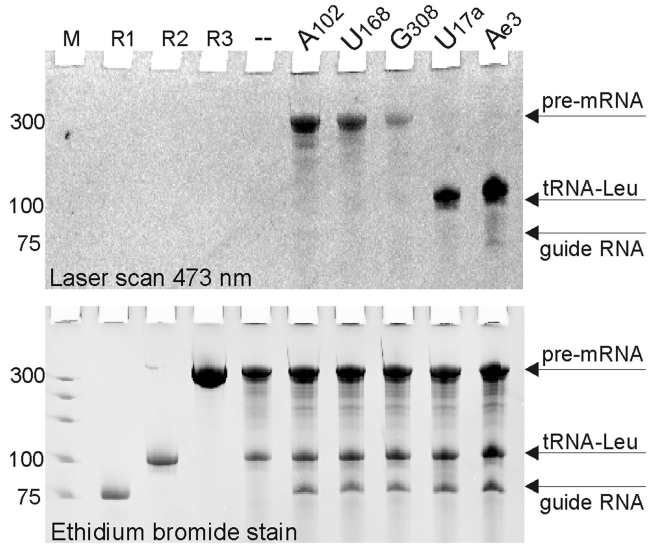


Figure 5. Substrate selectivity of C/D RNP-directed labeling of RNA. The assay was performed as in Figure 3B except that both the tRNA-Leu and rabbit β-globin pre-mRNA substrates were included in the reaction. RNA guides targeting A102, U168 and G308 in the pre-mRNA, and U17a and Ae3 in tRNA-Leu were used as indicated. M, DNA marker; R1, box C/D guide RNA; R2, tRNA-Leu; R3, rabbit β-globin pre-mRNA; '–', control reaction with no guide RNA.

from a synthetic analog of the AdoMet cofactor to a precisely defined position of a target RNA molecule; in the second step, the terminal triple bond can be further appended with a desired reporter group using a copper-catalyzed azide-alkyne 1,3-cycloaddition (32–34). One of the most important characteristic of this reaction is its unique bio-orthogonality, as neither azides nor terminal alkynes are generally present in natural compounds.

Among the cofactor analogs examined, SeAdoYn turned out to be the most efficient for the RNP-directed transfer of a reactive functional group onto RNA. Interestingly, the sulfur-containing prop-2-ynyl analog proved extremely labile in neutral and alkaline aqueous solutions due to the fast addition of a water molecule to the transferrable group [(20,22) and unpublished observations]. A larger atomic radius and a slightly lower electronegativity of Se as compared with S may both contribute to a lower reactivity of the side chain, and thus extend the lifetime of the cofactor in physiological buffers. Another advantage of SeAdoYn as compared with other stable cofactors carrying larger side chains is probably a compact size of the transferable propynyl group, which may better fit into the active site of the wild-type methyltransferase. Steric engineering of DNA methyltransferases proved a valuable approach for enhancing the catalytic transfer of linear side chains 12 atoms long (17,18). However, crystal structures of a C/D RNP from a related species (35) indicate that the transferable methyl group of the bound cofactor in the active site of the methyltransferase (fibrillarin) is surrounded by Lys-57, Asp-150, Lys-157, His-202, Ala-152 residues on one side and the bound RNA substrate on the other. All these residues are invariant in fibrillarin sequences from archaea to vertebrates (36); the catalytic K-D-K-H tetrad is structurally conserved in the fibrillarin family, and superimposes with the K-D-K-E tetrad found in RrmJ family (37). The catalytic Asp was shown to be essential for the methylation activity directly in an archaeal fibrillarin (38). All residues in the catalytic tetrad of the standalone rRNA methyltransferase RrmJ from *E. coli* were examined by mutagenesis and were shown to be essential or important for the methylation activity (39); a temperature-sensitive phenotype was obtained upon replacing Ala-245 to a Val in *S. cerevisiae* (corresponds to Ala-152 in *P. abyssi* fibrillarin) (40). Our attempt to sterically engineer the RNP methyltransferase by replacing Ala-152 with a smaller (Gly) residue did lead to reduced catalytic efficiency with both the natural and synthetic cofactors (data not shown). All this suggests that steric engineering of the active site may not be possible for this class of enzymes. Thus, SeAdoYn appears the cofactor of choice for C/D RNP.

Relative functionalization and labeling intensities are not readily predictable at different target sites. Although high labeling intensities of the tRNA substrate observed at U17a and Ae3 correlate well with high methylation rates observed in the presence of AdoMet, the other three positions show contrasting trends (methylation is higher at C34 and C56, whereas labeling is higher at A31). A strong dependence of the *in vitro* methylation efficiency on the structural context has been noted by others (13), and yet remains poorly understood. Notably, the RNP-directed ribose modification does not employ nucleotide flipping in its mechanism, which inevitably brings the RNA duplex formed by the substrate and guide RNA strands into close proximity with the bound cofactor and some of the catalytic residues (35). Sequence variations around the target nucleotide may thus substantially affect the chemical environment and conformation of these critical elements during catalysis. Altogether we found that three out of four newly selected targets were successfully modified *in vitro* by the reconstituted box C/D RNP in each model RNA substrate, indicating a high potential of this synthetically programmable system to label biologically relevant RNA targets for a variety of imaging and affinity-based experiments. Obviously, the plethora of natural modifications present in mature tRNAs may impose additional limitations as certain types of base modifications interfere with Watson-Crick base pairing required for the RNP activity.

Targeted labeling of RNA (19) and protein (20–22) using click chemistry have recently been reported using AdoMet-dependent methyltransferases. However, in all these systems, the target positions are determined by the substrate specificity of the methyltransferase, which cannot be changed by simple engineering. This may be a substantial shortcoming when labeling native biologically active macromolecules since the majority of methylation targets would be pre-methylated *in vivo*, and thus could no longer be targeted for labeling using corresponding natural enzymes. The flexibility in target site selection thus is of critical importance for a multitude of *in vivo* or *ex vivo* studies.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Tables 1–3 and Supplementary Figures 1–4.

ACKNOWLEDGEMENTS

The authors are grateful to Michael Martin for synthesis of the SeAdoYn cofactor, Gražvydas Lukinavičius for a gift of AdoButyn and Ado-11-amine cofactors, Zita Liutkevičiūtė for help with HPLC-MS analysis, Janina Ličytė for technical assistance and Dr Arvydas Kanopka for a gift of a rabbit β-globin gene.

FUNDING

European Social Fund under the Global Grant measure [VP1-3.1-ŠMM-07-K-01-105]; Lithuanian State Science and Studies Foundation and the French Ministry of foreign and European affairs under the Lithuania-France bilateral collaboration programme Gilibert [SUT-157]; Excellence Initiative of the German Federal and State Governments in part. Funding for open access charge: FP7-REGPOT-2009-1 program [245721, MoBiLi].

Conflict of interest statement. None declared.

REFERENCES

1. Onizuka,K., Taniguchi,Y. and Sasaki,S. (2009) Site-specific covalent modification of RNA guided by functionality-transfer oligodeoxynucleotides. *Bioconjug. Chem.*, **20**, 799–803.
2. Chow,C.S., Mahto,S.K. and Lamichhane,T.N. (2008) Combined approaches to site-specific modification of RNA. *ACS Chem. Biol.*, **3**, 30–37.
3. Paredes,E. and Das,S.R. (2010) Click chemistry for rapid labeling and ligation of RNA. *ChemBioChem*, **12**, 125–131.
4. 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.
5. 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.
6. 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.
7. Reichow,S.L., Hamma,T., Ferre-D'Amare,A.R. and Varani,G. (2007) The structure and function of small nucleolar ribonucleoproteins. *Nucleic Acids Res.*, **35**, 1452–1464.
8. 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.
9. Nolivos,S., Carpousis,A.J. and Clouet-d'Orval,B. (2005) The K-loop, a general feature of the Pyrococcus C/D guide RNAs, is an RNA structural motif related to the K-turn. *Nucleic Acids Res.*, **33**, 6507–6514.
10. Bleichert,F., Gagnon,K.T., Brown,B.A. II, Maxwell,E.S., Leschziner,A.E., Unger,V.M. and Baserga,S.J. (2009) A dimeric structure for archaeal box C/D small ribonucleoproteins. *Science*, **325**, 1384–1387.
11. Xue,S., Wang,R., Yang,F., Terns,R.M., Terns,M.P., Zhang,X., Maxwell,E.S. and Li,H. (2010) Structural basis for substrate placement by an archaeal box C/D ribonucleoprotein particle. *Mol. Cell*, **39**, 939–949.
12. Tran,E., Zhang,X., Lackey,L. and Maxwell,E.S. (2005) Conserved spacing between the box C/D and C′/D′ RNPs of the archaeal box C/D sRNP complex is required for efficient 2′-*O*-methylation of target RNAs. *RNA*, **11**, 285–293.
13. Appel,C.D. and Maxwell,E.S. (2007) Structural features of the guide:target RNA duplex required for archaeal box C/D sRNA-guided nucleotide 2′-*O*-methylation. *RNA*, **13**, 899–911.
14. Cavaillé,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.
15. Liu,B., Ni,J. and Fournier,M.J. (2001) Probing RNA *in vivo* with methylation guide small nucleolar RNAs. *Methods*, **23**, 276–286.
16. Klimašauskas,S. and Weinhold,E. (2007) A new tool for biotechnology: AdoMet-dependent methyltransferases. *Trends Biotechnol.*, **25**, 99–104.
17. Lukinavičius,G., Lapienė,V., Staševskij,Z., Dalhoff,C., Weinhold,E. and Klimašauskas,S. (2007) Targeted labeling of DNA by methyltransferase-directed transfer of activated groups (mTAG). *J. Am. Chem. Soc.*, **129**, 2758–2759.
18. Neely,R.K., Dedecker,P., Hotta,J., Urbanavičiūtė,G., Klimašauskas,S. and Hofkens,J. (2010) DNA fluorocode: a single molecule, optical map of DNA with nanometre resolution. *Chem. Sci.*, **1**, 417–528.
19. Motorin,Y., Burhenne,J., Teimer,R., Koynov,K., Willnow,S., Weinhold,E. and Helm,M. (2011) Expanding the chemical scope of RNA:methyltransferases to site-specific alkynylation of RNA for click labeling. *Nucleic Acids Res.*, **39**, 1943–1952.
20. Peters,W., Willnow,S., Duisken,M., Kleine,H., Macherey,T., Duncan,K.E., Litchfield,D.W., Lüscher,B. and Weinhold,E. (2010) Enzymatic site-specific functionalization of protein methyltransferase substrates with alkynes for click labeling. *Angew. Chem. Int. Ed. Engl.*, **49**, 5170–5173.
21. Binda,O., Boyce,M., Rush,J.S., Palaniappan,K.K., Bertozzi,C.R. and Gozani,O. (2011) A chemical method for labeling lysine methyltransferase substrates. *Chembiochem*, **12**, 330–334.
22. Islam,K., Zheng,W., Yu,H., Deng,H. and Luo,M. (2011) Expanding cofactor repertoire of protein lysine methyltransferase for substrate labeling. *ACS Chem. Biol.*, **6**, 679–684.
23. Lee,B.W., Sun,H.G., Zang,T., Kim,B.J., Alfaro,J.F. and Zhou,Z.S. (2010) Enzyme-catalyzed transfer of a ketone group from an *S*-adenosylmethionine analogue: a tool for the functional analysis of methyltransferases. *J. Am. Chem. Soc.*, **132**, 3642–3643.
24. Willnow,S., Martin,M., Lüscher,B. and Weinhold,E. (2012) A selenium-based click AdoMet analog for versatile substrate labeling with wild-type protein methyltransferases. *ChemBioChem*, **13**, 1167–1173.
25. Dalhoff,C., Lukinavičius,G., Klimašauskas,S. and Weinhold,E. (2006) Direct transfer of extended groups from synthetic cofactors by DNA methyltransferases. *Nat. Chem. Biol.*, **2**, 31–32.
26. 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.
27. Hardin,J.W. and Batey,R.T. (2006) The bipartite architecture of the sRNA in an archaeal box C/D complex is a primary determinant of specificity. *Nucleic Acids Res.*, **34**, 5039–5051.
28. Kanopka,A., Mühlemann,O. and Akusjärvi,G. (1996) Inhibition by SR proteins of splicing of a regulated adenovirus pre-mRNA. *Nature*, **381**, 535–538.
29. Grosjean,H., Droogmans,L., Roovers,M. and Keith,G. (2007) Detection of enzymatic activity of transfer RNA modification enzymes using radiolabeled tRNA substrates. *Methods Enzymol.*, **425**, 55–101.
30. Gramlich,P.M., Warncke,S., Gierlich,J. and Carell,T. (2008) Click-click-click: single to triple modification of DNA. *Angew. Chem. Int. Ed. Engl.*, **47**, 3442–3444.
31. Motorin,Y., Muller,S., Behm-Ansmant,I. and Branlant,C. (2007) Identification of modified residues in RNAs by reverse transcription-based methods. *Methods Enzymol.*, **425**, 21–53.
32. Rostovtsev,V.V., Green,L.G., Fokin,V.V. and Sharpless,K.B. (2002) A stepwise huisgen cycloaddition process: copper(I)-catalyzed

regioselective ''ligation'' of azides and terminal alkynes. *Angew. Chem. Int. Ed. Engl.*, **41**, 2596–2599.
33. Tornoe,C.W., Christensen,C. and Meldal,M. (2002) Peptidotriazoles on solid phase: [1,2,3]-triazoles by regiospecific copper(I)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides. *J. Org. Chem.*, **67**, 3057–3064.
34. Gramlich,P.M., Wirges,C.T., Manetto,A. and Carell,T. (2008) Postsynthetic DNA modification through the copper-catalyzed azide-alkyne cycloaddition reaction. *Angew. Chem. Int. Ed. Engl.*, **47**, 8350–8358.
35. 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.
36. Deng,L., Starostina,N.G., Liu,Z.J., Rose,J.P., Terns,R.M., Terns,M.P. and Wang,B.C. (2004) Structure determination of fibrillarin from the hyperthermophilic archaeon *Pyrococcus furiosus*. *Biochem. Biophys. Res. Commun.*, **315**, 726–732.
37. Feder,M., Pas,J., Wyrwicz,L.S. and Bujnicki,J.M. (2003) Molecular phylogenetics of the RrmJ/fibrillarin superfamily of ribose 2′-*O*-methyltransferases. *Gene*, **302**, 129–138.
38. Aittaleb,M., Visone,T., Fenley,M.O. and Li,H. (2004) Structural and thermodynamic evidence for a stabilizing role of Nop5p in *S*-adenosyl-L-methionine binding to fibrillarin. *J. Biol. Chem.*, **279**, 41822–41829.
39. Hager,J., Staker,B.L., Bugl,H. and Jakob,U. (2002) Active site in RrmJ, a heat shock-induced methyltransferase. *J. Biol. Chem.*, **277**, 41978–41986.
40. Tollervey,D., Lehtonen,H., Jansen,R., Kern,H. and Hurt,E.C. (1993) Temperature-sensitive mutations demonstrate roles for yeast fibrillarin in pre-rRNA processing, pre-rRNA methylation, and ribosome assembly. *Cell*, **72**, 443–457.
41. Schneider,K.L., Pollard,K.S., Baertsch,R., Pohl,A. and Lowe,T.M. (2006) The UCSC Archaeal Genome Browser. *Nucleic Acids Res.*, **34**, D407–D410.