VILNIUS UNIVERSITY

MIGLĖ TOMKUVIENĖ

METHYLTRANSFERASES AS TOOLS FOR SEQUENCE-SPECIFIC LABELING OF RNA AND DNA

Summary of doctoral dissertation

Physical sciences, biochemistry (04 P)

Vilnius, 2013

This study has been carried out at Vilnius University Institute of Biotechnology during 2006-2013.

Scientific supervisor

prof. habil. dr. **Saulius Klimašauskas** (Vilnius University; Physical sciences, Biochemistry – 04 P)

Evaluation board of dissertation of Biochemistry trend

Chairman

prof. dr. **Virginijus Šikšnys** (Vilnius University, Physical sciences, Biochemistry – 04 P)

Members

prof. dr. **Vilmantė Borutaitė** (Lithuanian Universitty of Health Sciences, Physical sciences, Biochemistry – 04 P)

habil. dr. **Sonata Jarmalaitė** (Vilnius University, Biomedical sciences, Biology – 01 B)

dr. **Giedrė Tamulaitienė** (Vilnius University, Physical sciences, Biochemistry – 04 P)

dr. **Arūnas Lagunavičius** (Thermo Fisher Scientific Baltics, Physical sciences, Biochemistry – 04 P)

Opponents

dr. **Jaunius Urbonavičius** (Vilnius University, Physical sciences, Biochemistry – 04 P)

dr. **Kęstutis Sužiedėlis** (Vilnius University Institute of Oncology, Physical sciences, Biochemistry – 04 P)

The thesis defence will take place at Vilnius University Institute of Biotechnology on the 5th of December, 2013, at 11 a.m.

Address: V. A. Graičiūno 8, Vilnius LT-02241, Lithuania

The summary of doctoral dissertation was distributed on 5th of November, 2013.

The thesis is available at the Library of Vilnius University and at the Library of Vilnius University Institute of Biotechnology.

VILNIAUS UNIVERSITETAS

MIGLĖ TOMKUVIENĖ

RNR IR DNR SPECIFINIS ŽYMĖJIMAS PANAUDOJANT METILTRANSFERAZES

Daktaro disertacijos santrauka

Fiziniai mokslai, biochemija (04P)

Vilnius, 2013

Disertacija rengta 2006-2013 m. Vilniaus universiteto Biotechnologijos institute.

Mokslinis vadovas

prof. habil. dr. **Saulius Klimašauskas** (Vilniaus universitetas; fiziniai mokslai, biochemija – 04 P)

Disertacija ginama Vilniaus universiteto Biochemijos mokslo krypties taryboje:

Pirmininkas

prof. dr. **Virginijus Šikšnys** (Vilniaus universitetas, fiziniai mokslai, biochemija – 04 P)

Nariai

prof. dr. **Vilmantė Borutaitė** (Lietuvos sveikatos mokslų universitetas, fiziniai mokslai, biochemija – 04 P)

habil. dr. **Sonata Jarmalaitė** (Vilniaus universitetas, biomedicinos mokslai, biologija – 01 B)

dr. **Giedrė Tamulaitienė** (Vilniaus universitetas, fiziniai mokslai, biochemija – 04 P)

dr. **Arūnas Lagunavičius** (UAB „Thermo Fisher Scientific Baltics“, fiziniai mokslai, biochemija – 04 P)

Oponentai

dr. **Jaunius Urbonavičius** (Vilniaus universitetas, fiziniai mokslai, biochemija – 04 P)

dr. **Kęstutis Sužiedėlis** (Vilniaus universiteto Onkologijos institutas, fiziniai mokslai, biochemija – 04 P)

Disertacija bus ginama viešame Biochemijos mokslo krypties tarybos posėdyje 2013 metų gruodžio mėn. 5 d. 11 val. Vilniaus universiteto Biotechnologijos instituto konferencijų salėje.

Adresas: V. A. Graičiūno 8, Vilnius LT-02241, Lietuva.

Disertacijos santrauka išsiuntinėta 2013 metų lapkričio mėn. 5 d.

Disertaciją galima peržiūrėti Vilniaus universiteto ir Vilniaus universiteto Biotechnologijos instituto bibliotekose.

Contents

INTRODUCTION

DNA and RNA are composed of four major nucleotides: adenine (A), guanine (G), cytosine (C) and thymine (T, DNA) or uracil (U, RNA). Sequence of nucleotides in polynucleotide chains encodes genetic information. Moreover, nucleotides in polynucleotide chains can be further covalently modified with various additional chemical groups, among which the methyl group is the most common. These modifications further expand the encoded information and refine the function of nucleic acids.

Investigation of nucleic acid function sometimes requires sequence-specific incorporation of various reporter and affinity probes. This is difficult to achieve using current technologies which compromise on efficiency, sequence- and site-specificity and the selection of functional groups to be incorporated. Specific labeling of biomolecules can be achieved using *methyltransferase-directed transfer of activated groups*, mTAG. This approach is based on activity of AdoMet-dependent methyltransferases (MTases) with synthetic AdoMet analogues equipped with transferable groups larger than the methyl group (Dalhoff et al., 2006; Lukinavičius et al., 2007; Klimašauskas and Weinhold, 2007). The transferrable chains usually carry reactive groups that can be further chemically appended with required reporters. As the methyltransfer reactions are widely used in cells, MTases of various specificities can be found. Their substrates range from small metabolites to proteins, DNA and RNA (Cheng and Blumenthal, 1999). Regarding the functional groups that can be used for specific coupling of reporters, azide and alkyne groups appear particularly attractive. Azide-alkyne 1,3-cycloaddition (AAC) is effective in physiological conditions and is bioorthogonal, i.e. these groups are not found in any biological compounds and do not react with biological functional groups (Kolb et al., 2001).

The mTAG approach has already been applied for DNA labeling (Lukinavičius et al., 2007). The DNA MTase HhaI was genetically engineered to accept synthetic cofactors with transferable linear chains as long as twelve atom units. The engineered M.HhaI is more active with synthetic cofactors than with AdoMet,

which allows the mTAG reaction to proceed even in the presence of the natural cofactor (Lukinavičius et al., 2012).

Nevertheless, the selection of cofactors that could be used with the engineered M.HhaI was limited. New stable cofactors with azide and alkyne groups were synthesized in our department. The investigation of their activity in mTAG reactions with the engineered M.HhaI and the possibility of using this MTase-cofactor system for DNA labeling in complex biological mixtures is described in this dissertation.

Similarly, mTAG labeling of RNA labeling can be achieved using RNA MTases. In archaea and eukaryotes, a particular class of RNA MTases, C/D ribonucleoprotein complexes (C/D RNPs), is found. These complexes consist of a guide RNA and several proteins, one of which has the methyltransferase activity. Precise base pairing of guide RNA sequences, usually 11 nt long, 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 conserved box D. Since the specificity of the reaction is defined by the changeable guide sequence of the C/D RNA, it is in principle possible to target methylation *de novo* to any predetermined position within RNA (Caveille et al., 1996). Many other types of RNA MTases recognize their substrates by protein-RNA interactions which cannot be readily manipulated to recognize predetermined targets (reviewed in Motorin and Helm, 2011). In this work, we sought to combine the unique specificity of C/D RNP machinery with synthetic AdoMet analogs to provide a novel molecular tool for programmable sequence-specific RNA labeling. For this purpose we have chosen a described, active *in vitro*, recombinant C/D RNP from a hyperthermophilic archaeon *Pyrococcus abyssi* (Bortolin et al., 2003; Nolivos et al., 2005).

The specific aims of this study were:

1) To investigate the activity of an engineered DNA methyltransferase HhaI (variant Q82A/Y254S/N204A) with synthetic cofactors that carry oct-2,7-diynyl and 6-azide-hex-2-ynyl transferable groups, and suitability of this reaction for DNA labeling in the presence of natural cofactor AdoMet *in vitro* and in cell lysates.

2) To examine the activity of the *Pyrococcus abyssi* C/D ribonucleoprotein complex (RNP) methyltransferase with synthetic AdoMet analogues and its suitability for site-specific labeling of RNA.
3) To examine if the C/D RNP-dependent mTAG reaction can be directed to new predetermined sites in model tRNA and iRNA substrates.
4) To evaluate the suitability of C/D RNP mTAG activity for site-specific labeling of programmed sites in model RNA substrates.

Scientific novelty

It was the first time that the activity of DNA methyltransferase HhaI (variant Q82A/Y254S/N204A) with synthetic cofactors which carry oct-2,7-diynyl and 6-azide-hex-2-ynyl transferable groups was shown. Second, it was investigated for the first time if RNA methyltransferase C/D RNP can be active with synthetic cofactors which carry groups longer than methyl. It was shown that the transfer of prop-2-ynyl group can be directed to new selected sites in substrate RNA.

Practical value

The described findings for the first time permit: 1) synthetically tunable sequence-specific labeling of RNA with single-nucleotide precision and 2) bioorthogonal sequence-specific DNA labeling possible to be achieved in complex biological mixtures. These new possibilities for sequence-specific labeling of nucleic acids can be adopted in biochemistry, biomedical, nanotechnology, etc. research.

Findings presented for defense:

1) The HhaI DNA methyltransferase variant (Q82A/Y254S/N204A) transfers the oct-2,7-diynyl and hex-2-ynyl-6-azide groups from synthetic cofactor analogues to target DNA *in vitro*, even in the presence of natural cofactor AdoMet. Using this methyltransferase, the cofactor carrying the hex-2-ynyl-6-azide group and azide-alkyne cycloaddition DNA can be specifically labeled in cell lysate.

2) *Pyrococcus abyssi* C/D RNP transfers prop-2-ynyl group from synthetic cofactor SeAdoYn on a specific target nucleotide in an RNA substrate. C/D RNP-dependent transfer of various other groups 2-5 carbon atoms long was less effective and longer groups were not transferred.

3) C/D RNP-dependent transfer of prop-2-ynyl group can be directed to predetermined sites in model tRNA and iRNA substrates with single-nucleotide precision by changing the sequence of guide RNA.
4) RNA molecules modified with prop-2-ynyl group can be specifically fluorescently labeled using azide-alkyne cycloaddition reaction.

MATERIALS AND METHODS

Reagents

AdoMet, DMSO, CuBr, TBTA and MegaStokes608 cyclooctine were purchased from *Sigma-Aldrich.* AdoMet was further purified by Dr. Rūta Gerasimaitė to remove traces of AdoHcy. Solvents for HPLC-MS were from *Roth.* Alexa647 azide was obtained from *Invitrogen*. Eterneon480/635 azide – from *Baseclick*. [metil-^{3}H]-AdoMet, 10 Ci/mmol, 55 µM - *PerkinElmer*. [γ-^{33}P]-ATP, >2500 Ci/mmol, 3.3 µM - *Hartmann Analytic.*

Nucleic acids

Genomic *Pyrococcus abyssi* DNA and recombinant plasmids pET15b-L7Ae, pET15b-Nop5p-aFib (Bortolin et al., 2003), pUC18-sR47 (Nolivos et al., 2005) were obtained from dr. Beatrice Clouet-d'Orval. Plasmid carrying a rabbit β-globin RNA gene (Kanopka et al., 1996) was a kind gift of dr. Arvydas Kanopka. Plasmid pΔGH$_6$E119H was derived from pΔGH$_6$ (Gerasimaitė et al., 2011) by replacing the codon of the active site amino acid Glu119 with codon of His by means of site-directed mutagenesis.

Reference DNA and RNA ladders were purchased from *Fermentas* (*Thermo Scientific).*

Synthetic DNA and RNA oligonucleotides were purchased from *Metabion*, 5'..3'.

For *P.abyssi* sR47 guide RNA reprogramming

Target nucleotide	Oligonucleotide name	Oligonucleotide sequence (D-guide section underlined) 5'..3', DNA
tRNA-Leu(CAA) sites		
U17a	U18a	GATCCAGTAAGGTGAATCAGAGCCTGGTCAAATCATCACTCC
	U18b	TGAGGAGTGATGATTTGACCAGGCTCTGATTCACCTTACTG
A31	A34a	GATCCAGTAAGGTGAATCAGGGGGACTCAAGATCATCACTCC
	A34b	TGAGGAGTGATGATCTTGAGTCCCCCTGATTCACCTTACTG
Ae3	A53a	GATCCAGTAAGGTGAATCAGCCGTAGGGGTTCTCATCACTCC
	A53b	TGAGGAGTGATGAGAACCCCTACGGCTGATTCACCTTACTG
C56	C68a	GATCCAGTAAGGTGAATCAGGGTTCAAATCCCTCATCACTCC
	C68b	TGAGGAGTGATGAGGGATTTGAACCCTGATTCACCTTACTG
Rabbit β-globin truncated pre-mRNA sites		
A102	RabGl_102a	GATCCAGTAAGGTGAATCAGAGTGAGGAGAAGTCATCACTCC
	RabGl_102b	TGAGGAGTGATGACTTCTCCTCACTCTGATTCACCTTACTG
U168	RabGl_168a	GATCCAGTAAGGTGAATCAGGCCCTGGGCAGGTCATCACTCC
	RabGl_168b	TGAGGAGTGATGACCTGCCCAGGGCCTGATTCACCTTACTG
C234	RabGl_234a	GATCCAGTAAGGTGAATCAGGAAACAGAGTAGTCATCACTCC
	RabGl_234b	TGAGGAGTGATGACTACTCTGTTTCCTGATTCACCTTACTG
G308	RabGl_308a	GATCCAGTAAGGTGAATCAGGGTTGTCTACCCTCATCACTCC
	RabGl_308b	TGAGGAGTGATGAGGGTAGACAACCCTGATTCACCTTACTG

DNA splint for substrate RNA oligonucleotide ligation

SplD_24 GGAGGGGATCTTGAGTGCCGCCGC.

DNA primers for the reverse transcription analysis
RT_tLeu CGGGGGCGGGGATTTGAACC, RT_tLeu2 ACCCCTACGGGAGGGGA,
RT_tLeu3 TCTTGAGTCCCCCGCCTT, RT_tLeu4 CGGGGATTTGAACCCCGGA,
RT_tLeu5 CGGGGATTTGAACCCCGGAAC.
For eM.HhaI mTAG activity assay double stranded DNA pATTGCGC was annealed from ATTATTATTATTAGCGCATTATTA and TAATAATGCGCTAATAATAATAAT (HhaI recognition sequence is underlined).

DNA primers for PCR production of *in vitro* transcription templates

Oligonucleotide name	Oligonucleotide sequence (T7 promoter is underlined), 5'..3'
P.abyssi tRNA-Leu(CAA)	
P25'	CCGGAATTCTAATACGACTCACTATAGCGGGGGTTGCCGAGCCTGGT
P23'	CGCGGATCCTGGTGCGGGGGCGGGGATTT
Rabbit β-globin truncated pre-mRNA	
T7 β-globin	ATTAATACGACTCACTATAGAATACAAGCTTGGGCTG
β-globin (ex2)	GAGGACAGGTCCCCAAAG

RNA substrate oligonucleotide and parts for substrate construction
S22 CCACAGACUCAAGAUCCCCUCC, D11 GCGGCGGCACU, D13 CAAGAUCCCCUCC.

Enzymes

Restriction endonucleases, RNase A/T1 mix, T4 polynucleotide kinase, T4 DNA ligase, DNaseI, nuclease BAL31, High Fidelity PCR Enzyme Mix, phosphatases CIAP, SAP and FastAP, reverse transcription RevertAid premium mix, protein size reference ladder were purchased from *Fermentas (Thermo Scientific).* Thrombin - from *GE Healthcare.* Bovine serum albumin – *Pierce*. Nuclease P1 and lysozyme – *Sigma-Aldrich.* eM.HhaI was prepared by Giedrė Urbanavičiūtė as described (Daujotytė et al., 2003). Endonuclease McrBC was a kind gift from dr. Giedrius Sasnauskas (Sukackaitė et al., 2012).

Synthetic cofactors

SeAdoYn (Willnow et al., 2012), AdoEnYn (Peters et al., 2010), SeAdoEt (unpub.), AdoButen (unpub.) – obtained from prof. Elmar Weinhold, were prepared by dr. Michael Martin, dr. Sophie Willnow and others. AdoEt, AdoPropen, AdoButin (all three – Dalhoff et al., 2006), Ado-6-amine, Ado-6-alkyne, Ado-11-amine (all three – Lukinavičius et al., 2013) were prepared by dr. Gražvydas Lukinavičius. Ado-6-$SCOCH_3$ (unpub.) and Ado-6-azide (Lukinavičius et al., 2013) were prepared by dr. Viktoras Masevičius. Ado-6-azide preparation was of pure S-stereoisomer, others were mixtures of S- and R- stereoisomers. In the reactions with MTases only S-isomer concentration is calculated. All the synthetic cofactors were of >60% purity.

Reaction buffers

C/D RNP buffer for methylation and modification reactions *in vitro* (1×): 50 mM HEPES·NaOH pH=7,9, 150 mM KCl, 10 mM $MgCl_2$.

M.HhaI buffer for methylation and modification reactions *in vitro* (1×): 10 mM Tris·HCl pH=7,4, 50 mM NaCl, 0,5 mM EDTA, 0,2 mg/ml BSA.

C/D RNP protein expression and purification

The *P. abyssi* L7Ae, Nop5p and aFib proteins were expressed from recombinant pET15b vectors (Bortolin et al., 2003) in Escherichia coli BL21 (DE3) RIL (Stratagene) cells induced with 1mM isopropyl-b-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 (Hardin and Batey, 2006). 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, Thermo Scientific) using the following templates: a linearized pUC18-based plasmid carrying a recombinant sR47 gene (Nolivos et al., 2005) (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 (Kanopka et al., 1996) (for β-globin pre-mRNA). After ethanol precipitation, RNAs were purified on a 6% PAA denaturing gel.

Electrophoretic analysis of RNP assembly

C/D guide RNA sR47 was 5'-radiolabeled using T4 polynucleotide kinase and [γ-^{33}P]-ATP and passed through G-25 sephadex columns (GE Healthcare). Radiolabeled RNA was incubated with increasing concentrations of proteins L7Ae and Nop5p-aFib in reaction buffer (20 mM HEPES·NaOH pH=7.9, 150 mM NaCl, 1.5 mM $MgCl_2$, 10% glycerol, 0.2 mg/ml BSA and 0.5 mg/ml *E.coli* tRNA) at 70°C for 15 min. Complexes were resolved on native 8% polyacrylamide gel buffered with 0.5× TBE, and 5% glycerol. Gels were vacuum-dried and visualized using a phosphorimager.

In vitro methylation assay

C/D RNP was pre-assembled by combining 1 µM guide RNA, 3 µM L7Ae and 2.5 µM Nop5p and aFib in C/D RNP buffer in the presence of 50 µM [^{3}H-methyl]-AdoMet (1.6 Ci/mmol, obtained by diluting [^{3}H-methyl]-AdoMet (10 Ci/mmol, Perkin Elmer) with cold AdoMet) and pre-incubation at 68°C for 7 min. Substrate RNA was added to 1 µM and the reaction was incubated at 68°C. Aliquots were taken at indicated time points and cooled on ice. Samples were then applied onto Cellulose DE-81 (Whatman) filters, filters washed with 50 mM Na-PO_4 pH 7.4

and dried. ^{3}H-methyl incorporation into RNA was determined by scintillation counting in *EcoPlus* scintillation fluid (Roth). Assays were done at least in triplicate.

Assaying RNP activity toward synthetic cofactor analogs

The approach used substrate oligonucleotide with internally ^{33}P-labeled target nucleotide. RNA oligonucleotide D13 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 was annealed on a splint DNA oligonucleotide 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 mM sR47, 2 mM L7Ae and 2 mM 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 mM), or a synthetic cofactor analog (200–600 mM), and the internally ^{33}P-labeled substrate oligonucleotide (0.01 mM) supplemented with substrate oligonucleotide S22 (0.5 mM) 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 (Grosjean et al., 2007). 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 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

RNA modification reactions for HPLC analysis typically contained 1 mM preassembled RNP, 1 mM tRNA-Leu and 50 mM AdoMet or 400 mM 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 (ZymoResearch) 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 in P1 buffer (10 mM Tris-HCl, 10 mM magnesium chloride, and 1 mM zinc acetate, pH 7.5).

DNA modification reactions were performed with eM.HhaI (variant Q82A/Y254S/N304A, 0.5 μM), oligonucleotide duplex (pATTGCGC, 33 μM), and cofactors Ado-6-alkyne or Ado-6-azide (300 μM) in reaction buffer at 37 °C for 4 h and then heated at 80 °C for 20 min. The modified duplexes were desalted by ethanol precipitation. Nuclease P1 buffer (see above) containing Nuclease P1 (0.05 au) and shrimp alkaline phosphatase (0.1 au) was added, and the samples were incubated at 37 °C for 4 h.

After digestion to nucleosides all the samples were heated 10 min at 80°C to inactivate the enzymes and were centrifuged 30 min 10600×g. Hydrolysate consistent with 40 pmol of undigested RNA or 130 pmol DNA was loaded onto an integrated HPLC/ESI-MS Agilent 1200 series system equipped with a multichannel absorption detector and a Discovery C18 column (75×2.1 mm, Supelco) and eluted with a linear gradient of solvents A (20mM 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 in a positive ionization range 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 areas of unmodified nucleosides followed by measurement of the peak areas of the corresponding propynylated and methylated nucleosides.

Two-step labeling of RNA

Modification reaction. RNP (1 mM) was assembled by combining 1 mM guide RNA, 3 mM L7Ae and 2.5 mM Nop5p and aFib in methylation buffer (20mM Hepes-NaOH, pH 7.9, 150mM NaCl and 10mM MgCl2) and pre-incubation at 68°C for 7 min. Afterwards the synthetic cofactor analog SeAdoYn (400 mM) and 1 mM substrate RNA (and 10 mM 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 ml; 20–220 pmol of substrate RNA) was diluted with 0.8 volume (32 ml) of DMSO/t-BuOH 3:1 (Baseclick). 0.03M CuBr (Sigma-Aldrich) was freshly prepared each time in 0.1M tris(benzyltriazolylmethyl)amine (TBTA) (Sigma-Aldrich) (in DMSO/t-BuOH), and 12 ml of Cu-TBTA solution was added to the mixture followed by 8 ml of 2.5mM (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 473nm laser. Gels were then stained with ethidium bromide for bulk RNA imaging.

Reverse transcription analysis

Reverse transcription analysis was done following a described procedure (Motorin et al., 2007). 1.4 mM ^{33}P-5'-end-labeled primers were hybridized to 0.4 mM unmodified or C/D RNP guide Eterneon-labeled tRNA-Leu. Extension step was performed by addition of a mixture containing 400 mM or 4 mM 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 mM dNTP, except having 0.7 mM 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.

mTAG modification and labeling of DNA in the presence of competing AdoMet *in vitro*

Modification. eM.HhaI (0.125 μM) and pΔGH6E119H plasmid DNA (0.25 μM total GCGC sites) were combined with cofactor Ado-6-alkyne or Ado-6-azide and AdoMet (total cofactor concentration 50 μM) at specified ratios. Reactions were incubated at 37 °C for 1 h. Completeness of DNA modification was monitored by incubating reaction aliquots with restriction endonuclease R.Hin6I (2−5 au per 1 μg of DNA) in accordance to manufacturer recommendations (Fermentas).

DNA labeling via Cu(I)-assisted click chemistry (cofactor Ado-6-alkyne). An aliquot of a mTAG modification reaction (5 μL) was diluted with an equal volume of DMSO and 1.5 μL of freshly prepared Cu-TBTA solution was added followed by 1 μL of 2.5 mM Alexa647 azide in DMSO. The reaction mixture was incubated 45 min at 37°C, and DNA purified by ethanol precipitation. Cu-TBTA solutions were prepared by combining 0.03 M CuBr (Sigma-Aldrich) in DMSO with 100 mM tris-(benzyltriazolylmethyl)amine (TBTA) (Sigma-Aldrich) in DMSO/tert-butanol. For analysis of site-specific labeling, modified DNA was digested with R.HindII and R.PscI restriction enzymes for 5 h at 37 °C to generate three fragments: 1900 bp with 9 HhaI targets, 1500 bp with 5 HhaI targets, and 970 bp with no HhaI targets. Fragmented DNA was resolved on a 1% agarose gel and stained with EtBr. Gels were scanned on a Fuji FLA-5100 gel imagining system using 635 (for Alexa647) and 473 nm (for EtBr) lasers.

DNA labeling via Cu-free click chemistry. Cofactor Ado-6-azide mTAG modification reactions were directly treated with Alkyne MegaStokes608 for Cu-free click-chemistry (Sigma-Aldrich). MegaStokes608 dye (Sigma-Aldrich) was added to 100 μM, and the mixtures were incubated at 37°C for 1 h. Afterwards, the DNA was column-purified (Zymo Research IIIC columns). Finally, it was fragmented and agarose-gel resolved as described above (DNA labeling via Cu-assisted click chemistry in vitro). For MegaStokes608, fluorescence gels were scanned using a 473 nm laser and then stained with EtBr (0.1 mg mL^{-1}) and scanned again to visualize bulk DNA.

DNA modification and labeling in cell lysate

E.coli ER2267 cells were transformed with the pΔGH6E119H plasmid. After overnight growth in LB medium, the cells were spun down and resuspended in TE buffer (100 mM Tris·HCl, pH=8.0, and 100 mM EDTA). Proteinase inhibitor PMSF to 2.5 mM, RNase A/T1 mix, and lysozyme to 0.25 mg mL^{-1} were added, and the suspension was incubated 1 h at 37°C. Cell debris and unlysed cells were spun down, supernatant was combined with eM.HhaI (final concentration 2 μM) and cofactor Ado-6-azide or AdoMet (final concentration 50 μM) in TE buffer and incubated for 3 h at 37°C. PMSF was replenished every hour. Finally, plasmids were purified using a GeneJET Plasmid Miniprep kit (Fermentas) and digested with R.Hin6I or McrBC. For fluorescent labeling of the azide derivatized DNA, an equal volume of DMSO and MegaStokes608 alkyne dye to 200 μM were added to the lysate after modification and incubated for 3 h at 37°C. Labeled DNA was column-purified (Zymo Research IIIC columns) and digested with R.HincII and R.PscI as described above (DNA labeling via click chemistry *in vitro*).

RESULTS AND DISCUSSION

Methyltransferases (MTases) naturally catalyze the transfer of methyl group from the cofactor S-adenosyl-L-methionine (AdoMet, Figure1B) to a variety of biomolecules. To expand the practical utility of this highly specific enzymatic reaction, synthetic AdoMet analogues have been developed, which permit transfer of extended side chains with reactive functional groups. The approach has been named methyltransferase-directed Transfer of Activated Groups, mTAG (Lukinavičius et al., 2007). Reporter and affinity labels can then be coupled to the transferred functional groups resulting in site-specifically labeled biomolecules.

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

In this work, we sought to combine the specificity of archaeal C/D ribonucleoprotein complex (RNP) machinery with synthetic AdoMet analogs to provide a novel molecular tool for programmable sequence-specific RNA labeling (Figure 1). 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 (RUGAUGA) 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 (Klein et al., 2001; Nolivos et al., 2005). 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 (Bleichert et al., 2009). Precise base pairing of the guide sequences, usually 12 nt long (11 nt complementarity) in functional archaeal C/D RNAs (Tran et al., 2005; Lin et al., 2011), upstream of the D or D' boxes of the C/D RNA with its cognate RNA substrate directs the methyltransferase reaction to the 2'-O-ribose of a target nucleotide determined by nucleotide at the fifth position upstream from the D or D' box (Appel and Maxwell, 2007). C/D RNP 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 tested its ability to site-specifically transfer alkynyl groups from a range of synthetic cofactors. The resulting alkyltransferase activity was directed to both the wild-type and newly programmed target sites in RNA molecules. We also show that the transferred group with a terminal triple C-C bond can be further appended with a fluorophore using a bioorthogonal azide-alkyne 1,3-cycloaddition (click chemistry) (Rostovsev et al., 2002; Tornoe et al., 2002) reaction (Figure 1C).

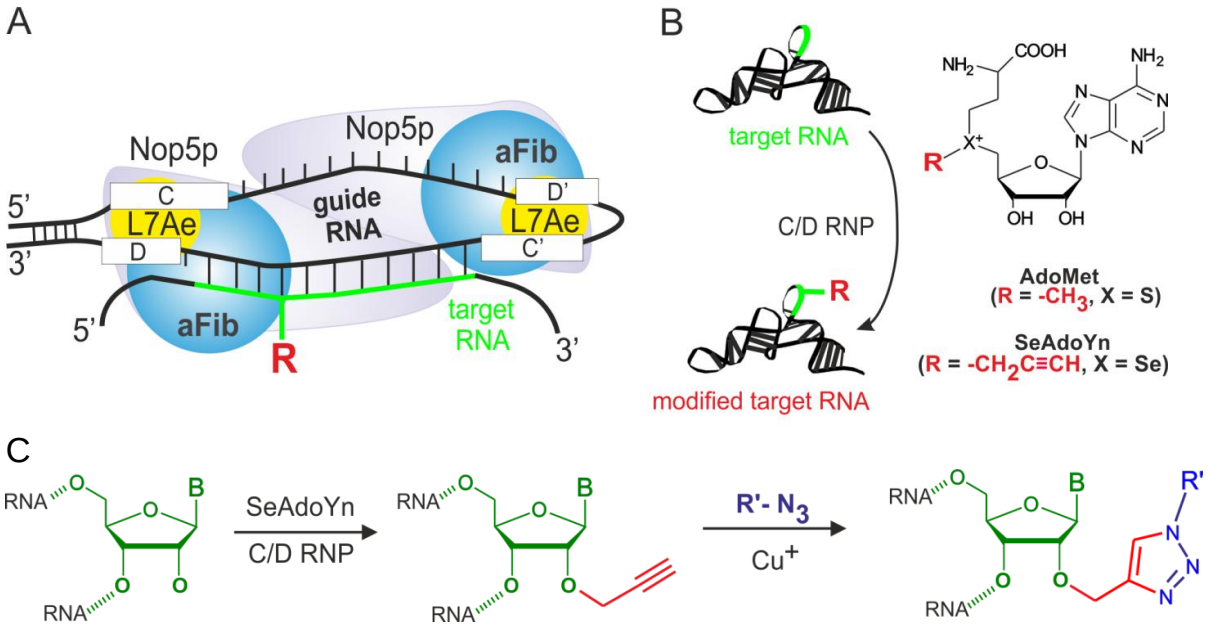


Figure 1. Archaeal C/D RNP-directed sequence-specific modification and labeling of target RNA. **(A)** Schematic structure of a C/D RNP complex with substrate RNA. Core proteins L7Ae, Nop5p and aFib are bound at the C/D and C‘/D‘ sites of a guide RNA. One of the variable guide sequences is shown base-paired 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).

1.1. 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 (Bortolin et al., 2003; Nolivos et al., 2005). 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) (Nolivos et al., 2005). 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 (see below, Figure 4). 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, Table 1). 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 (Grosjean et al., 2007) permitted selective determination of the modification status of the radiolabeled target nucleotide (Figure 2). Modification reaction conditions, including buffer composition, pH, temperature and others, were varied to achieve optimal RNP activity with synthetic cofactors (Figure 3). Finally, our analyses showed that, none of the synthetic cofactors with side chains larger than 5 carbon units rendered detectable modification of the RNA substrate (Table 1). It turned out that a synthetic cofactor, SeAdoYn (Figure 1B), carrying a selenium bound three carbon chain (Willnow et al., 2012) was well accepted by the archaeal methyltransferase aFib in the RNP guide reaction (Figure 2, Lane 2, and Figure 3A,C,D, Table 1). The transfer of the propynyl group on the cytosine was dependent on the C/D RNP enzyme (Figure 2, Lane 3).

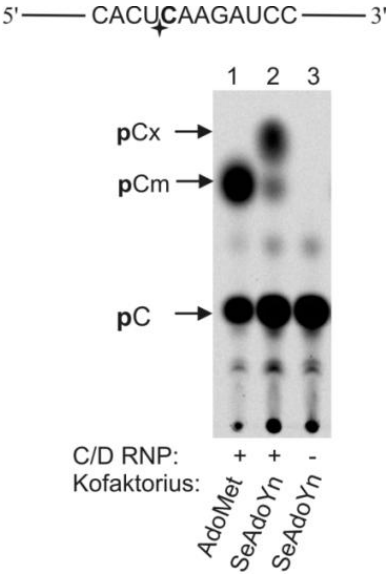
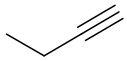
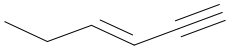
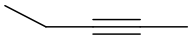
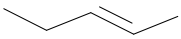
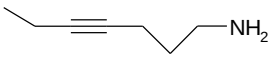
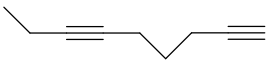
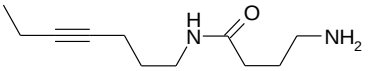
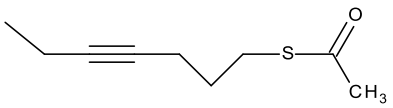


Figure 2. Formation of modified nucleotides in a 24-mer RNA oligonucleotide incubated with the C/D sR47 RNP in the presence of cofactors AdoMet and SeAdoYn. The target sequence of the substrate oligonucleotide (top) contains a ^{33}P-labeled target nucleotide (shown in bold and an asterisk). Substrate 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).

1.2 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 (Figure 4B) except for the sR47-A31guide, which showed a considerably lower incorporation of methyl groups into

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.

S/Se center	Transferable group (R)	Reactive group	% [a]	Cofactor name	Reference
S	-CH_3	-	50-97	AdoMet	-
Se		+	5-10	SeAdoYn	Willnow et al., 2012
S		+	2-7	AdoEnYn	Peters et al., 2010
S		-	5-8	AdoPropen	
S		-	2-5	AdoButin	Dalhoff et al., 2006
S		-	1-3	AdoEt	
Se			1-3	SeAdoEt	Weinhold group, unpub.
S		-	3	AdoButen	
S		+	0	Ado-6-amine	
S		+	0	Ado-6-alkyne	Lukinavičius et al., 2013
S		+	0	Ado-11-amine	
S		+	0	Ado-6-$SCOCH_3$	Klimašauskas group, unpub.

[a] modification efficiency of the target nucleotide (amount of the modified nucleotide in the total amount of the target nucleotide in a sample) was determined by TLC analysis of ^{33}P-labeled 5‘-mononucleotides as shown in Figures 2 and 3.

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 5). To determine 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 5), whereas a control sample in the absence of the functional MTase (omission of the Nop5 and aFib

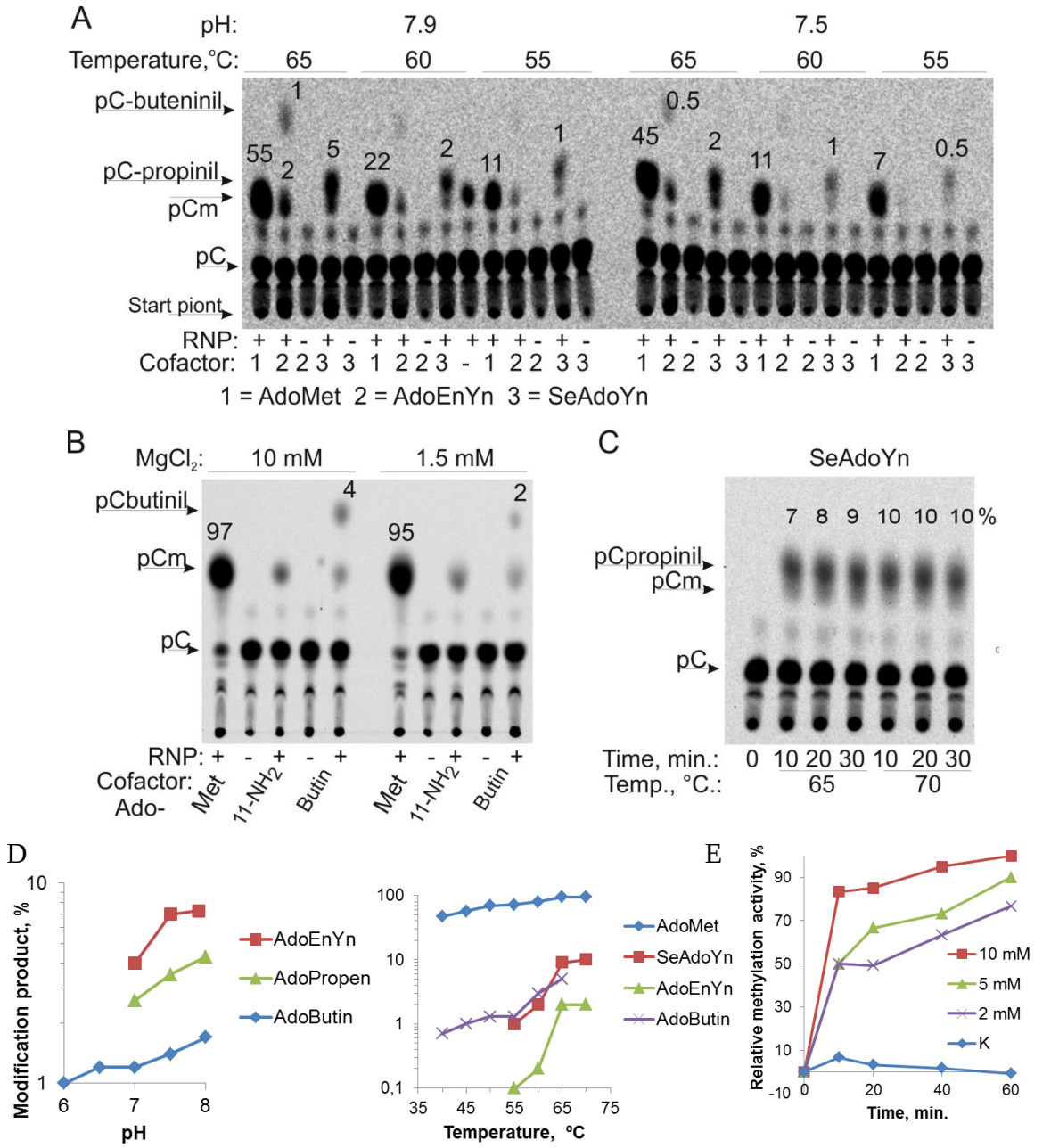


Figure 3. C/D RNP activity with AdoMet and synthetic cofactors depending on varied reaction conditions. Substrate oligonucleotide 24-mer RNA oligonucleotide (1 μM) was incubated with 1 μM pre-assembled RNP and 50 μM AdoMet or 400-600 μM synthetic cofactors AdoEnYn, SeAdoYn, Ado-11-NH_2 and AdoButin at pH=7.9, 65°C, 10 mM $MgCl_2$, 30-40 min., if not shown otherwise. Samples were subjected to nuclease Bal31 digestion and TLC analysis of ^{33}P-labeled mononucleotides. Arrows point to target nucleotide products. Numbers near spots indicate percentage of their relative signal strength in the total amount of spots corresponding to the target nucleotide. **(A)** Reactions with AdoMet, AdoEnYn or SeAdoYn at varied pH and temperature. **(B)** Reactions with AdoMet, Ado-11-NH_2 or AdoButin in the presence of $MgCl_2$ concentrations 10 mM and 1,5 mM. **(C)** Reactions with SeAdoYn at varied incubation time and temperature. **(D)** Modification activity with indicated cofactors depending on pH and temperature as shown using this assay. **(E)** Methylation activity on tRNA-Leu depending on varied $MgCl_2$ concentrations (indicated on the plot) as shown using tritium incorporation assay (as in fig. 4B) K – polyU as a substrate in the presence of 10 mM $MgCl_2$.

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 (Table 2). 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. 10% at the wt (C34) target, 35% at U17a, 24% at A31, 17% 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.

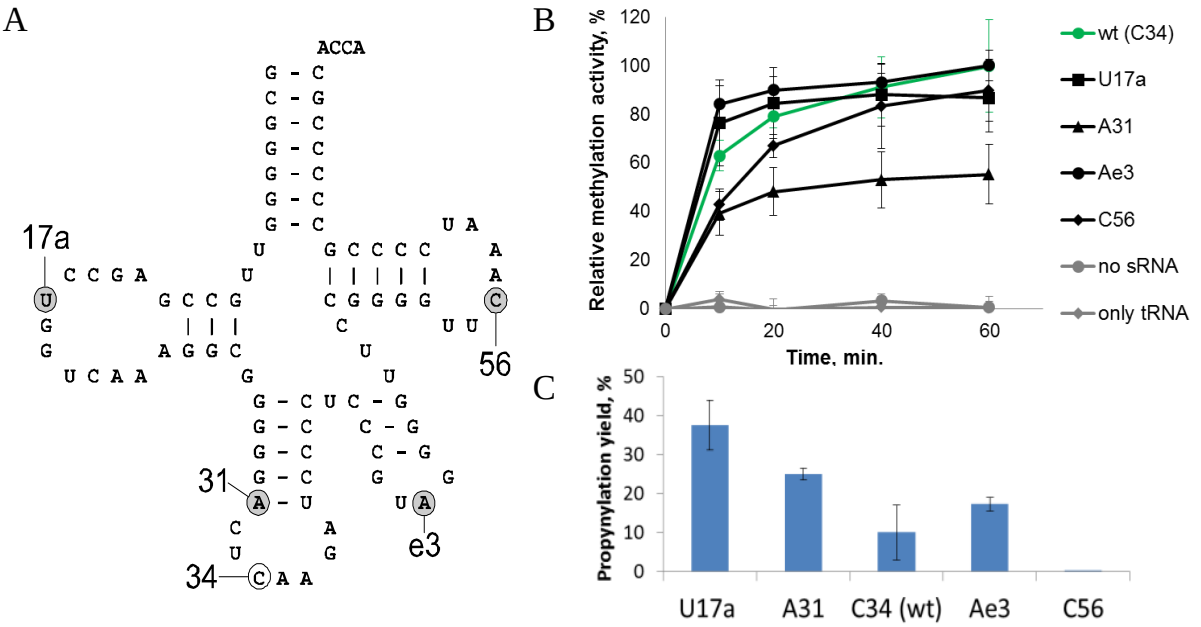


Figure 4. Targeting C/D RNP reactions to predetermined sites in a model tRNA substrate. **(A)** Schematic representation of *P. abyssi* tRNA-Leu(CAA) (Schneider et al., 2006). Position of target nucleotide for the wt sR47 RNA (C34) is shown as an empty circle, and positions of newly programmed target nucleotides (U17a, A31, Ae3 and C56) are shown as filled circles. **(B)** Time course of C/D RNP-dependent methylation of tRNA-Leu targets in the presence of 1 µM of pre-assembled C/D RNPs, 50 µM of [^{3}H-*methyl*]AdoMet and 1 µM of target RNA, at 68°C. **(C)** C/D RNP-dependent propynylation of tRNA at predetermined sites. 1 µM C/D RNP was incubated with 1 µM model tRNA substrate and SeAdoYn (400 µM) at 68°C for 30 min.

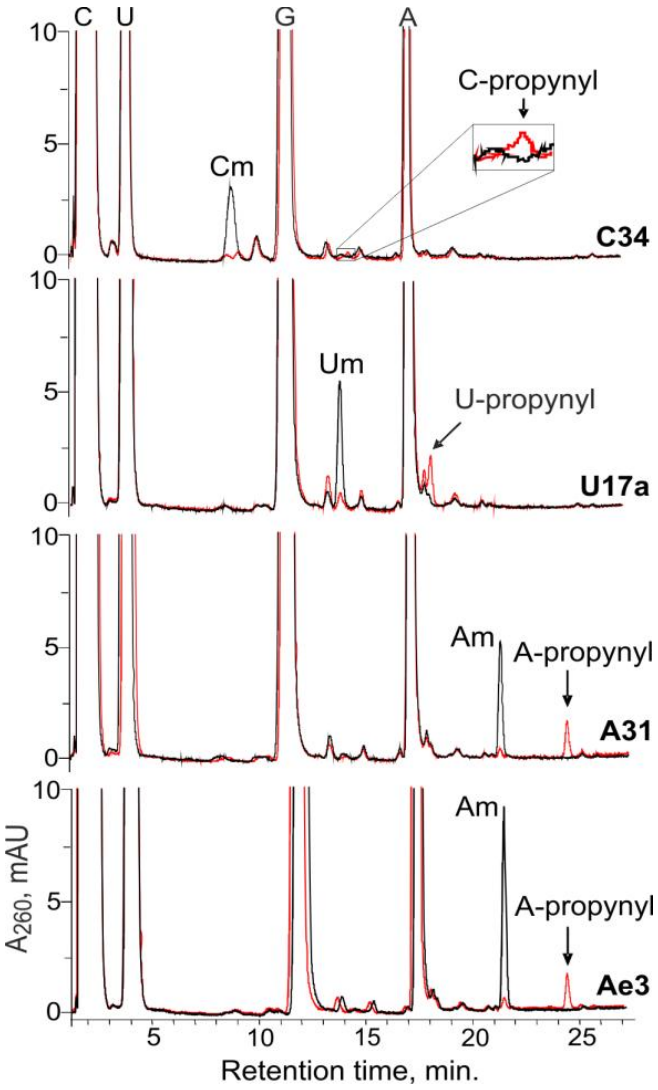


Figure 5. HPLC-MS analysis of products obtained after C/D RNP-directed modification of a model tRNA substrate with four variants of sR47 guide RNA (targeting C34, U17a, A31, Ae3 positions as indicated) in the presence of the AdoMet or SeAdoYn cofactor. Pre-assembled C/D RNP (1 μM), tRNA (1 μM) and AdoMet (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.

Table 2. Mass spectrometry analysis of HPLC fractions corresponding to modified nucleosides, produced in tRNA substrate after C/D RNP-directed modification with wt and reprogrammed variants of guide-RNA and in the presence of AdoMet or SeAdoYn cofactors.

Target modification		Theoretical mass of protonated nucleoside	Experimental m/z
C34(wt)	Methyl	258.1084	258.1081
	Propynyl	282.1084	282.1080
U17a	Methyl	259.0925	259.0919
	Propynyl	283.0925	283.0924
A31	Methyl	282.1202	282.1196
	Propynyl	306.1202	306.1192
Ae3	Methyl	282.1202	282.1198
	Propynyl	306.1202	306.1194

1.3 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 6). 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. 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 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.

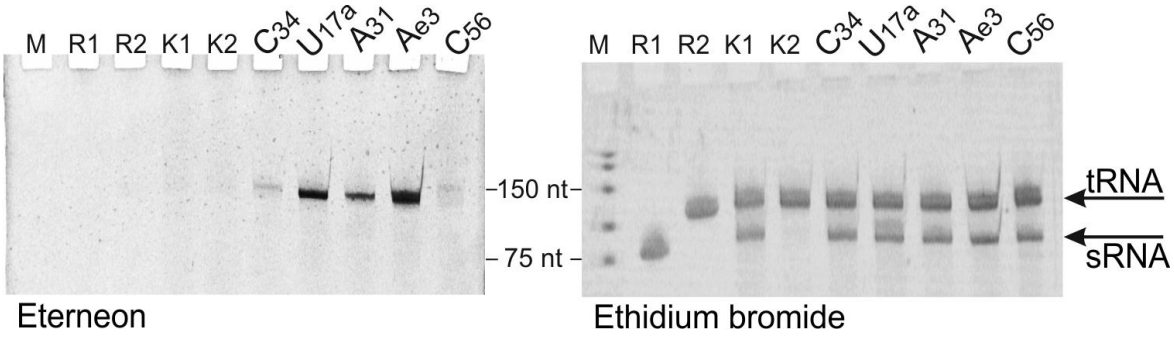


Figure 6. Fluorescent sequence-specific labeling of tRNA via guide RNA-directed enzymatic propynylation and copper-assisted coupling of an Eterneon(480/635) azide. The reactions (1 μM pre-assembled RNP with a corresponding guide RNA, 1 μM substrate RNA and 400 μM SeAdoYn cofactor incubated at 68°C for 30 min, followed by click chemistry reaction with Eterneon480/635 azide for 1 h at 37°C) were resolved on a denaturing polyacrylamide gel, scanned with a 473nm laser for Eterneon480/635 fluorescence (left panel), and then stained with ethidium bromide to reveal bulk RNA (right 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 and C/D RNA (sRNA) are shown by arrows.

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 (Figure 7), 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.

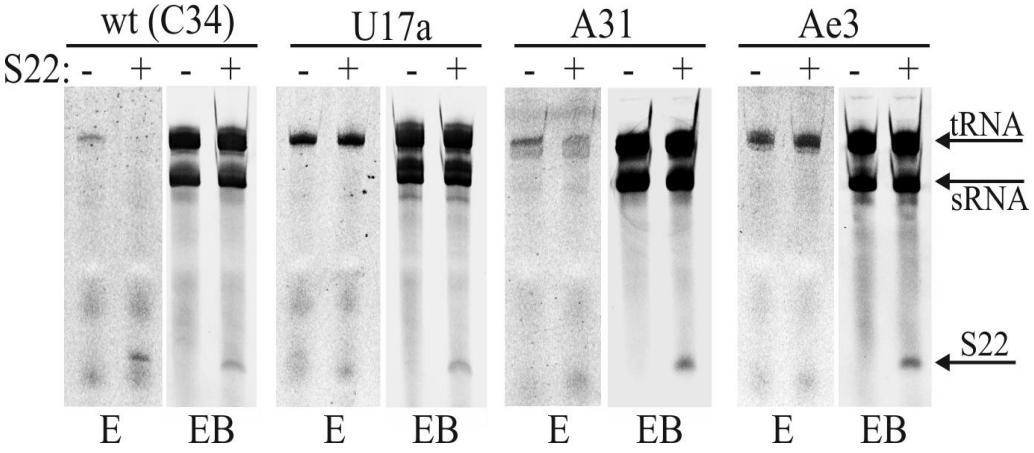


Figure 7. Specificity of the reprogrammed guide modification reaction in the presence of a competing oligonucleotide. Modification reactions were performed as in Figure 6 except that a tenfold molar excess of a 22-mer RNA oligonucleotide (S22) competitor, having the recognition sequence for the wt guide RNA sR47, was added comparing to the tRNA-Leu substrate. The wt C/D RNA sR47 and the three new variants were used as guide RNAs to target different sites as indicated on the top lanes. The PAA gel was scanned for the Eterneon fluorescence (left panel for each RNP variant) and then stained with ethidium bromide (right panel).

The positions of the RNP-directed labels were further confirmed by reverse transcription (RT) primer extension analysis. The best fluorescently labeled tRNAs (at U17a, A31 and Ae3 sites) were subjected to RT reactions along with sequencing reactions on an unmodified tRNA-Leu (Figure 8). 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.

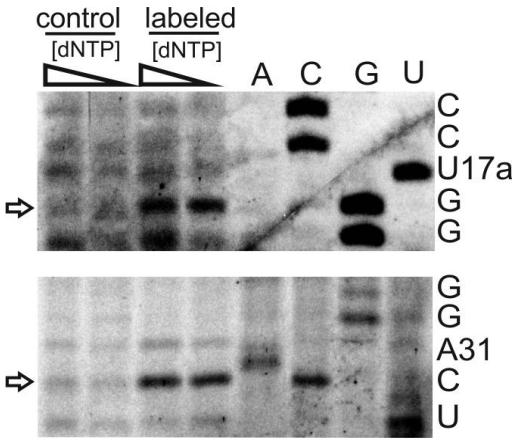


Figure 8. Reverse transcription (RT) mapping of RNP modification sites. tRNA-Leu was modified in the presence of a C/D RNP and SeAdoYn cofactor followed by click coupling of an Eterneon fluorophore, as in Figure 6. RT primer extension analysis shows polymerase halting (arrows) one position ahead of the target (U17a or A31) nucleotides in samples labeled using corresponding guide C/D RNAs.

1.4 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) (Figure 9A and B). 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 (Figure 9C). The two-step fluorescent labeling was also successful for all four targeted sites (Figure 9D), 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 10). The presented results convincingly show that the C/D RNP-directed labeling is targeted to a programmed RNA substrate.

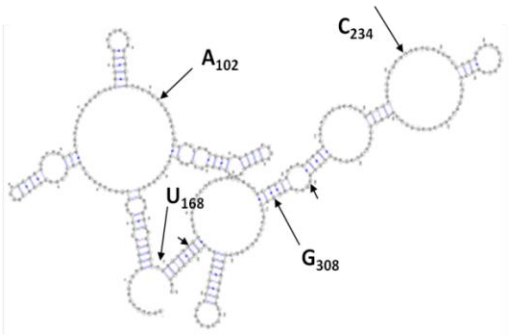
2 Bioorthogonal sequence-specific labeling of DNA

mTAG approach is being extensively applied for DNA labeling. DNA cytosine-5 MTase HhaI has been engineered at the active site and has become an alkyltransferase effectively transferring extended groups from the synthetic AdoMet analogues even in the presence of the natural cofactor AdoMet. The mutant M.HhaI has retained its specificity: it recognizes G<u>C</u>GC sequence and modifies the first cytosine (underlined) (Lukinavičius et al., 2012). This M.HhaI with a changed transferase activity has been named eM.HhaI (*<u>e</u>ngineered*). In parallel to MTase

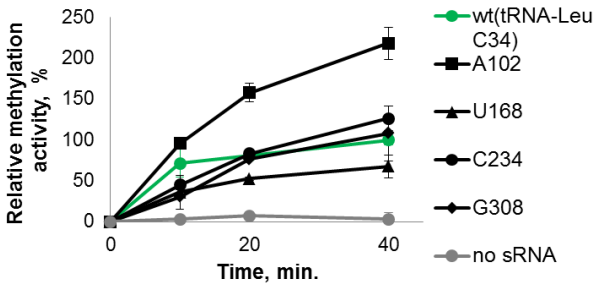
A

```
  1 GAAUACAAGC UUGGGCUGCA CUGCUGCUUA CACUUGCUUU UGACACAACU GUGUUUACUU
 61 GCAAUCCCCC AAAACAGACA GAAUGGUGCA UCUGUCCAGU GAGGAGAAGU CUGCGGUCAC
121 UGCCCUGUGG GGCAAGGUGA AUGUGGAAGA AGUUGGUGGU GAGGCCCUGG GCAGguuggu
181 auccuuuuua cagcacaacu uaaugagaca gauagaaacu ggucuuguag aaacagagua
241 gucgccugcu uuucugccag gugcugacuu cucuccccug ggcuguuuuc auuuucucag
301 GCUGGUUGUC UACCCAUGGA CCCAGAGGUU CUUCGAGUCC UUUGGGGACC UGUCCUC
```

B



C



D

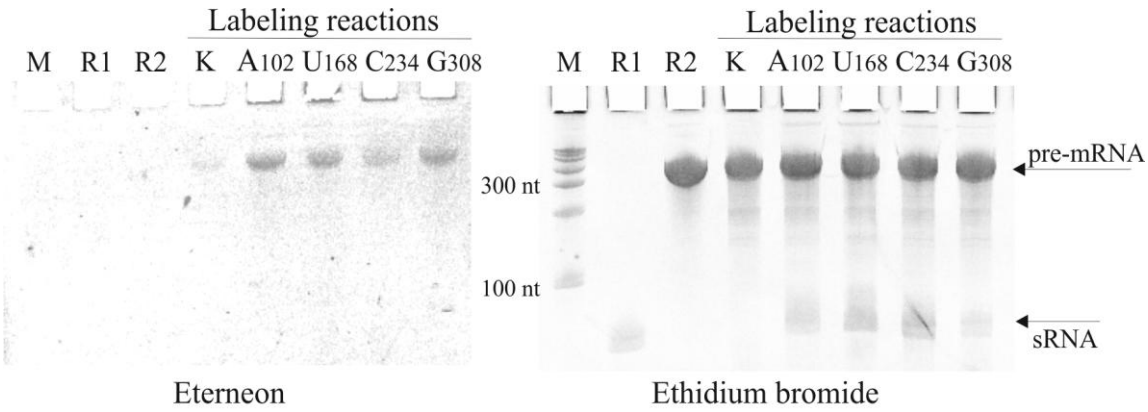


Figure 9. Programmable sequence-specific labeling of model pre-mRNA. **(A)** The sequence of 357-nt fragment of rabbit β-globin pre-mRNA, encompassing the first exon and intron and a part of the second exon. Exon sequence is in uppercase, intron – in lowercase. We had chosen four sites in this sequence, the according target sequences are marked in grey and the target nucleotides are in black. **(B)** Secondary structure of the pre-mRNA, obtained by *in silico* modeling by *Gensilico RNA* metaserver. Long arrows point to target nucleotides, the short ones point to boundaries of the intron. **(C)** Time course of C/D RNP-dependent methylation of pre-mRNA targets. 1 μM C/D RNP, 100 μM [^{3}H-methyl]-AdoMet and 1 μM substrate RNA were incubated at 68°C. Control reaction contained no guide RNA. **(D)** Fluorescent sequence-specific labeling of rabbit β-globin truncated pre-mRNA via guide-RNA directed enzymatic propynylation and copper-assisted coupling of an Eterneon480/635 azide. RNA molecules from the reaction (1 μM pre-assembled RNP with a corresponding guide RNA, 1 μM substrate RNA and 400 μM SeAdoYn cofactor incubated at 68°C for 30 min, followed by click chemistry reaction with Eterneon480/635 azide for 1 h at 37°C) were separated on a denaturing PAA gel. The gel was scanned for Eterneon fluorescence (left panel) and stained with ethidium bromide (right panel). M is for the RiboRuler RNA Ladder Low Range from Fermentas, R1 for box C/D guide RNA, R2 for unmodified substrate, and K for control in the absence of C/D guide RNA. Above the gel, lanes are labeled according to the nucleotides targeted in the reaction.

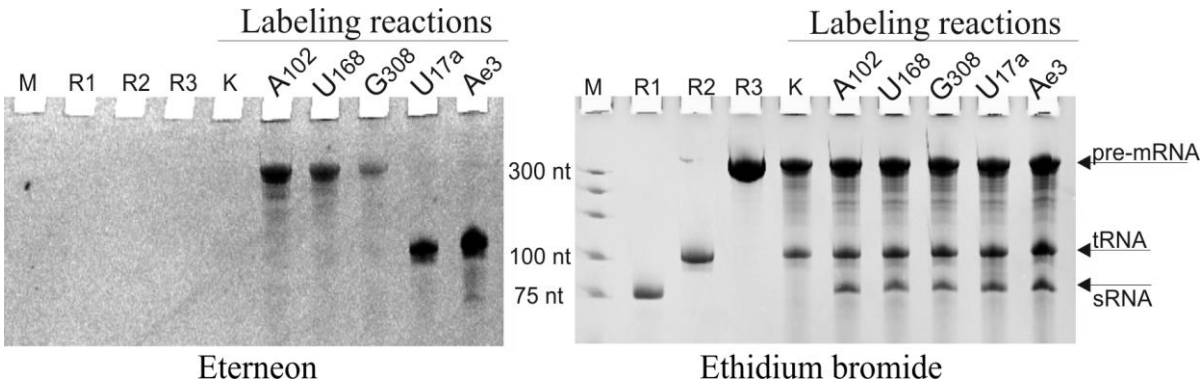


Figure 10. Substrate selectivity of C/D RNP-directed labeling of RNA. The assay was performed as in Figures 6 and 9D except that both the tRNA-Leu and rabbit β-globin pre-mRNA substrates were included in the reaction. RNA guides targeting A102, U168 and G308 in the pre-mRNA, and U17a and Ae3 in tRNA-Leu 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.

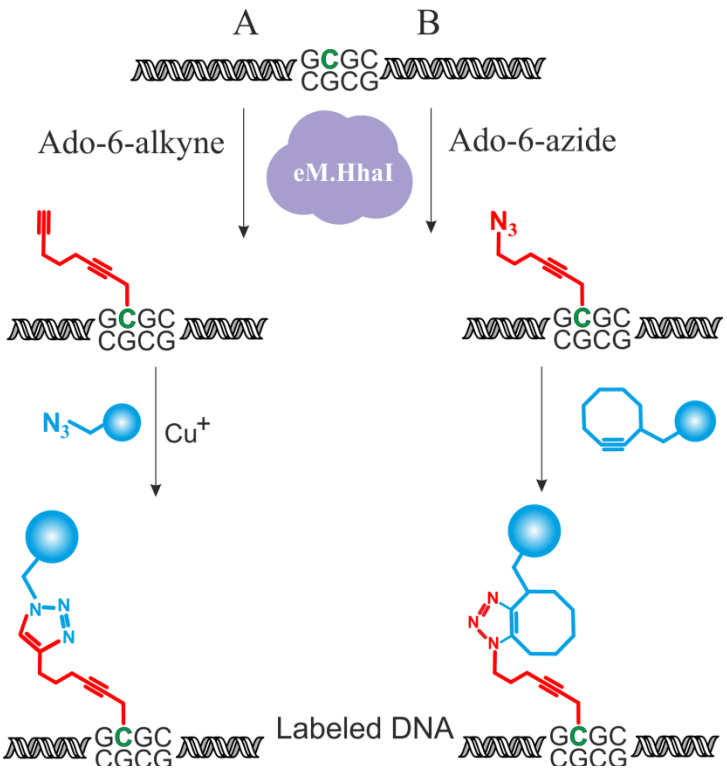


Figure 11. DNA mTAG labeling using MTase eM.HhaI and synthetic cofactors Ado-6-alkyne **(A)** and Ado-6-azide **(B)**, and azide-alkyne cycloaddition in the second step.

engineering synthetic cofactors with extended transferable chains carrying different functional groups were developed. Among these were two with functional groups suitable for azide-alkyne cycloaddition (click chemistry) reaction. Ado-6-alkyne carries a terminal C-C triple bond and Ado-6-azide – an azide group. In both cases

these groups are attached to transferable chains consistent of six carbon units and an activating triple bond at the β position to the sulfonium center. (Figure 11). It has been shown that synthetic cofactors of this design are well suitable for mTAG reactions: they are stable in reaction buffer at 37°C and active in eM.HhaI catalysis (Lukinavičius et al., 2013). On the basis of eM.HhaI activity with synthetic cofactors in the presence of AdoMet and the bioorthogonality of azide-alkyne cycloaddition we show that this system of molecular tools is suitable for DNA labeling in complex biological mixtures including cell lysate.

2.1 eM.HhaI activity with synthetic cofactors

Enzymatic activity of eM.HhaI with Ado-6-alkyne and Ado-6-azide was first assayed with a model substrate 24 bp long double stranded DNA oligonucleotide with one HhaI recognition site. After the modification reaction DNA was digested to nucleosides and analyzed by an HPLC-MS method which confirmed enzymatic transfer of whole side chains from both the cofactors onto a cytosine base (Figure 12).

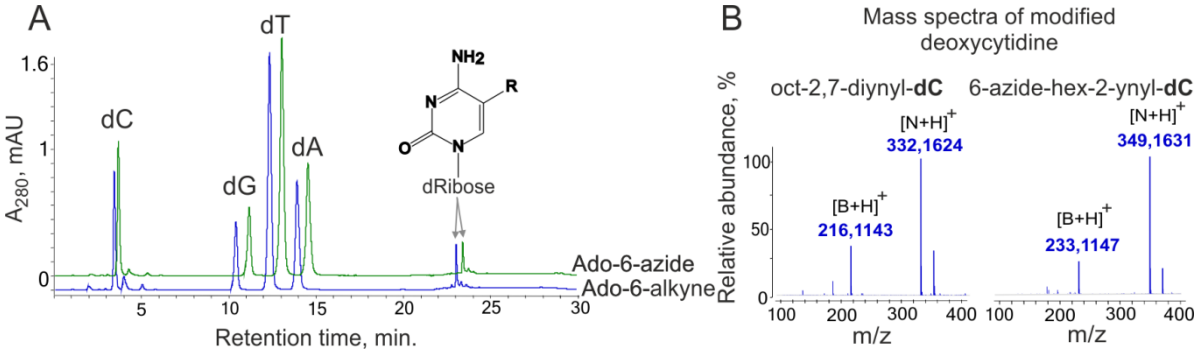


Figure 12. HPLC-MS analysis of mTAG transalkylation products formed in a duplex oligodeoxynucleotide with cofactors Ado-6-alkyne and Ado-6-azide, and eM.HhaI. **(A)** Nucleoside HPLC UV traces of enzymatically fragmented duplex oligodeoxynucleotide obtained after modification with eM.HhaI in the presence of AdoMet analogs. **(B)** ESI-MS analysis of HPLC fractions corresponding to modified nucleosides. N denotes 2'-deoxynucleoside; B – nucleobase.

An increased steric bulk of the AdoMet analogues makes them poor substrates for wild type MTases (Lukinavičius et al., 2012). This is a useful feature for numerous *in vivo* and *ex vivo* studies in which selective transalkylation activity with a designated tailor-engineered MTase is sought in the context of a vast variety of endogenous AdoMet-dependent MTases. However, this would be especially useful if an mTAG labeling reaction could be performed in the presence of the natural AdoMet cofactor. Thus we next tested the competition of natural cofactor and Ado-6-alkyne or Ado-6-azide in the eM.HhaI catalyzed reaction. For this we used a multicopy recombinant plasmid pΔGH$_6$E119H which has 14 HhaI targets unevenly distributed in the sequence (Figure 13A). Modification reactions contained AdoMet and synthetic cofactor Ado-6-alkyne or Ado-6-azide mixtures in different ratios but

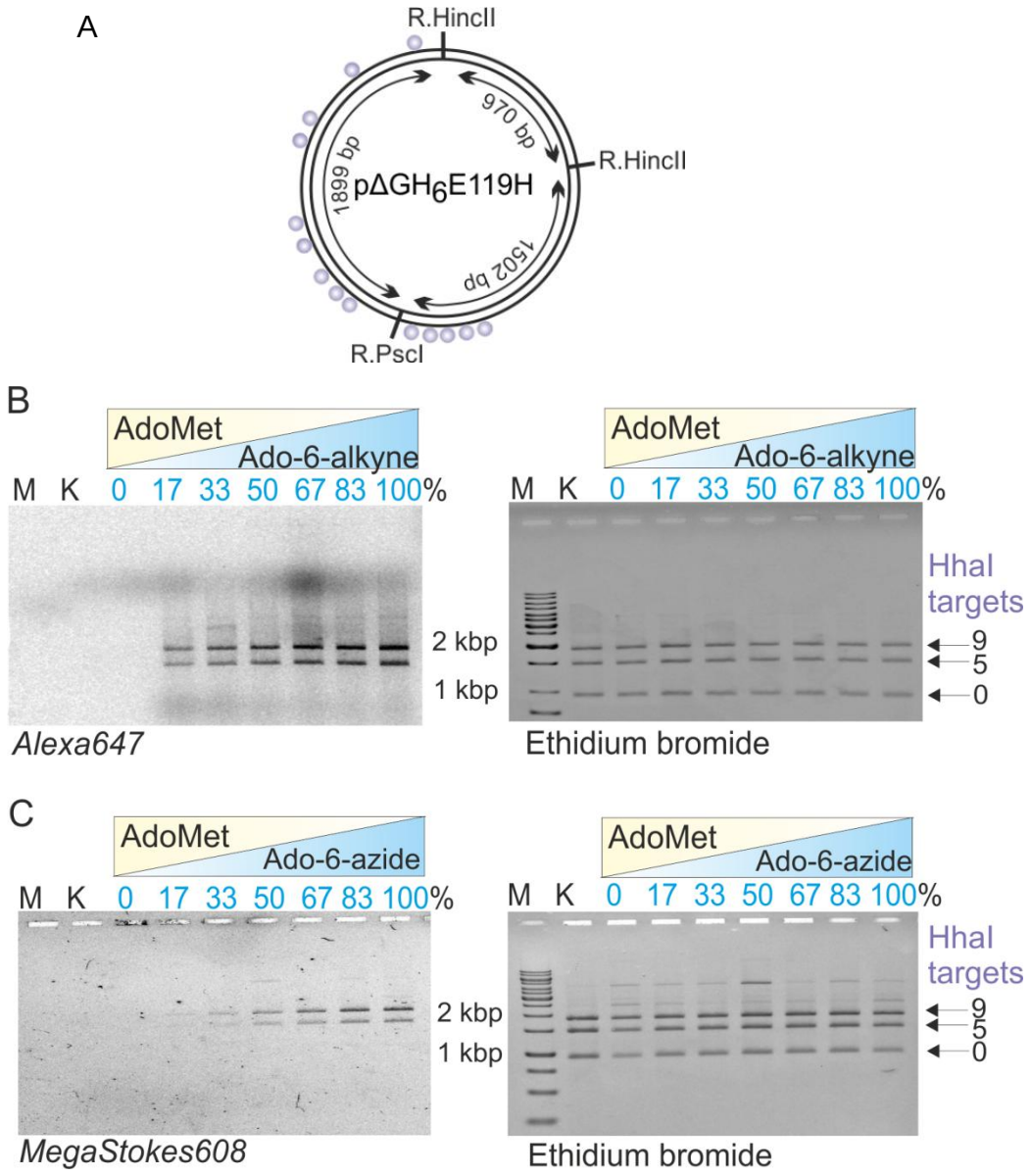


Figure 13. Sequence-specific click-labeling of DNA using eM.HhaI and synthetic cofactors Ado-6-alkyne or Ado-6-azide in mixtures with AdoMet *in vitro*. Plasmid pΔGH6E119H DNA (0.25 μM targets) was modified with eM.HhaI (0.125 μM) in the presence of AdoMet and synthetic cofactor Ado-6-alkyne or Ado-6-azide combined in different ratios as indicated (total cofactor concentration 50 μM). Control sample (lanes C) contained no cofactor. Modified DNA was labeled with a suitable dye and then treated with R.HincII and R.PscI endonucleases to give three DNA fragments containing 9, 5 or 0 HhaI target sites. The resulting fragments were separated on an agarose gel and scanned for labeling dye and EtBr fluorescence. M – 1 kb DNA Ladder. **(A)** A scheme representing distribution of HhaI targets (spheres) and restriction endonuclease digestion sites in pΔGH$_6$E119H plasmid. **(B)** mTAG labeling with cofactor Ado-6-alkyne and Alexa647 azide. Modified DNA was analyzed by agarose gel electrophoresis then stained with ethidium bromide and scanned for Alexa647 (635 nm laser) (left panel) and EtBr (473 nm laser) (right panel) fluorescence; **(C)** mTAG labeling with cofactor Ado-6-azide and Alkyne MegaStokes608 dye. The gel was first scanned for MegaStokes fluorescence (473 nm laser) (left panel), then stained with ethidium bromide and scanned again to visualize bulk DNA (right panel).

constant total cofactor concentration. After modification the reaction mixtures were supplemented with click chemistry reagents to couple fluorophores. Alexa647 azide was used in case of Ado-6-alkyne and the click reaction needed Cu^{+} catalysis. In case of Ado-6-azide reaction MegaStokes608 cyclooctine was chosen which is active enough because of ring tension and needs no extra catalysis for azide-alkyne cycloaddition (Nagy et al., 2010). Purified labeled DNA was treated with R.HincII and R.PscI endonucleases to give three DNA fragments containing 9, 5 or 0 HhaI target sites and resolved on an agarose gel (Figure 13B,C). Enzymatic transalkylation is well apparent even in samples with synthetic cofactor part as low as 17% and further increasing accordingly to the increasing part of a synthetic cofactor in a mixture with AdoMet. The relative fluorescence intensity of labeled DNA fragments was in accordance with the number of HhaI targets they had. The fragment with no HhaI targets had no detectable fluorescence. It argues for high specificity of our system.

2.2 Sequence-specific fluorescent DNA labeling in cell lysate

Since bioorthogonal conjugation reactions offer added benefits, we therefore examined if similar two-step labeling of DNA is possible in bacterial cell lysate. Pilot experiments showed that Ado-6-azide is the best choice for this purpose because the further click chemistry does not need Cu(I) catalysis.

First, *E.coli* ER2267 cells carrying the pΔGH6E119H plasmid were gently lysed with lysozyme and the lysate was complemented with eM.HhaI and cofactor AdoMet or Ado-6-azide. After modification the plasmids were purified from the lysate and aliquots digested with the R.Hin6I or McrBC nuclease. R.Hin6I cleaves DNA at unmodified HhaI target sites, but does not cleave at methylated sites; McrBC fragments DNA in the vicinity of methylated HhaI targets; both nucleases are inactive on GCGC sites modified with extended alkyl groups (Lukinavičius et al., 2012). This analysis showed that the mTAG reaction with the eM.HhaI is strongly in favor of transferring the extended groups, resulting in most of the HhaI sites carrying azide modification (Figure 14).

As a final proof of principle, we performed mTAG transalkylation and subsequent copper-free click labeling of DNA both in a bacterial cell lysate. For this, *E. coli* cells were gently lysed and treated with eM.HhaI and cofactor Ado-6-azide as in the previous experiment. In the second step, incubation was continued in the presence of MegaStokes608 cyclooctine. Purified plasmid DNA was then digested with R.HincII and R.PscI as described above and analyzed by agarose gel electrophoresis to reveal three clearly resolved DNA fragments. The observed fluorescence intensity distribution was fully consistent with the number of the GCGC target sites in the plasmid-derived fragments demonstrating a sequence-specific

labeling of the cellular DNA (Figure 15). The experimental procedure is simple and specific even in cell lysate attesting its suitability for routine laboratory use.

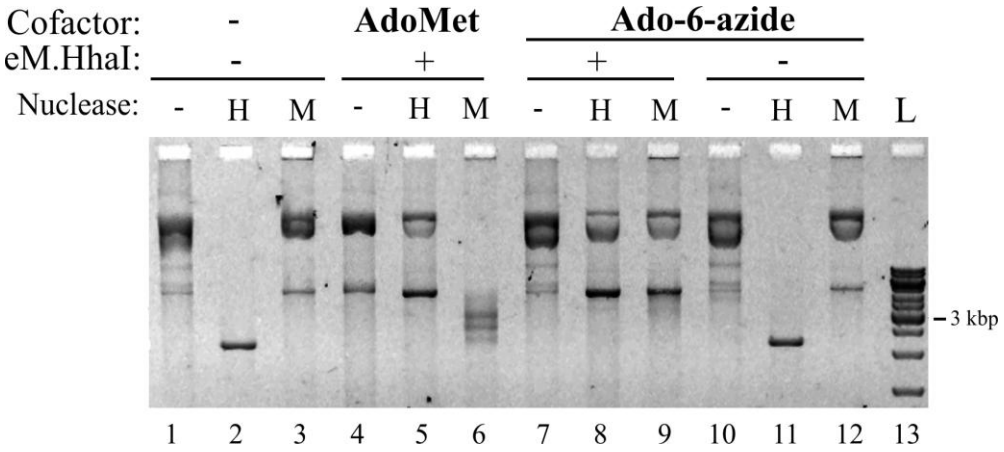


Figure 14. Methylation and modification in cell lysate. *E.coli* ER2267 cells carrying the pΔGH6E119H plasmid were lysed with lysozyme and the lysate was complemented with eM.HhaI (2 μM) and cofactor AdoMet (control) or Ado-6-azide (50 μM). After modification the plasmids were purified from the lysate and aliquots digested with the R.Hin6I (lanes H), or McrBC (lanes M) nuclease. R.Hin6I cleaves DNA at unmodified HhaI target sites (lanes 2 and 11), but does not cleave at modified sites (lanes 5 and 8); McrBC fragments DNA in the vicinity of methylated HhaI targets (lane 6); both nucleases are inactive on GCGC sites modified with extended alkyl groups. Resistance of DNA modified with eM.HhaI and cofactor Ado-6-azide to the action of both nucleases (lanes 8 and 9) indicates its complete alkylation of the GCGC sites. Lane 13 – GeneRuler DNA Ladder mix.

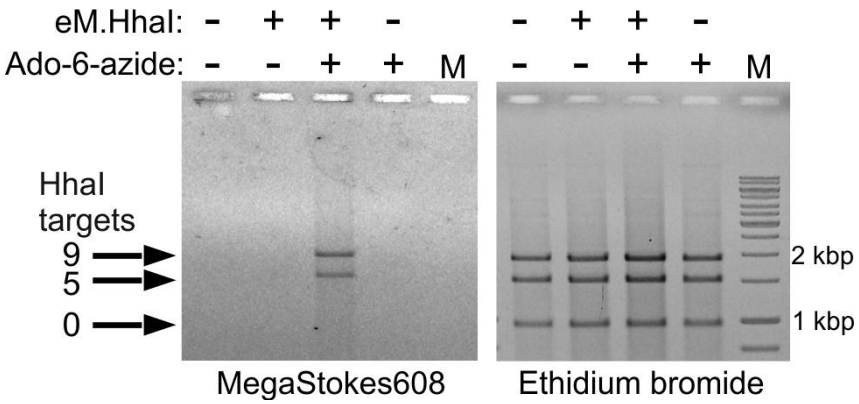


Figure 15. Sequence-specific two-step mTAG labeling of plasmid DNA using copper-free click reaction in bacterial cell lysate. *E. coli* ER2267 cells carrying the pΔGH6E119H plasmid were lysed and the lysate was treated with eM.HhaI (2 μM) and cofactor Ado-6-azide (50 μM) for 3 h at 37 °C. Alkyne MegaStokes608 dye cyclooctine was added (200 μM), and incubation continued for another 3 h, followed by plasmid DNA isolation by column purification. The plasmid was fragmented with R.HincII and R.PscI endonucleases and resolved on an agarose gel. The gel was first scanned for MegaStokes fluorescence (left panel), then stained with ethidium bromide and scanned again to visualize bulk DNA (right panel). Control samples lacked M.HhaI or cofactor as indicated. M, DNA size marker 1 kb DNA Ladder.

DISCUSSION

Programmable sequence-specific click-labeling of RNA

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 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 (Rostovsev et al., 2002; Tornoe et al., 2002). One of the most important characteristic of this reaction is its unique bioorthogonality, 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 C/D 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 [(Peters et al., 2010; Islam et al., 2011) and unpublished observations]. 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 (Lukinavičius et al., 2007; Neely et al., 2010). However, crystal structures of a C/D RNP from a related species (Lin et al., 2011) 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 (Deng et al., 2004); 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 (Feder et al., 2003). The catalytic Asp was shown to be essential for the methylation activity directly in an archaeal fibrillarin (Aittaleb et al., 2004). 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 (Hager et al., 2002); a temperature-sensitive phenotype was obtained upon replacing Ala-245 to a Val in *S. cerevisiae* (corresponds to Ala-152 in *P. abyssi* fibrillarin) (Tollervey et al., 1993). 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 (Appel and Maxwell, 2007) and yet remains poorly understood. Notably, the RNP-directed ribose modification does not employ nucleotide flipping (Klimašauskas et al., 1994; reviewed in Roberts and Cheng, 1998) 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 (Lin et al., 2011). 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 seven out of eight newly selected targets were successfully modified *in vitro* by the reconstituted box C/D RNP in two model RNA substrates, 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 (Motorin et al., 2010), DNA (Artyukhin and Woo, 2012) and proteins (Peters et al., 2010; Islam et al., 2011; Wang et al., 2011; Willnow et al., 2012; Bothwell et al., 2012) 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 *ex vivo* studies.

Selective DNA labeling in biological systems

We have shown that DNA mTAG labeling can be achieved using MTase eM.HhaI (Lukinavičius et al., 2012) and synthetic cofactors bearing functional groups for azide-alkyne cycloaddition reaction, namely Ado-6-alkyne with a terminal C-C

triple bond and Ado-6-azide with an azide group. Thus the arsenal of cofactors that can be used with eM.HhaI has been supplemented with two new cofactor analogues and a possibility to use azide-alkyne cycloaddition for coupling a label.

Wild type MTases are usually inactive with the synthetic cofactors bearing long transferable chains, as it has been shown with DNA (Lukinavičius et al., 2007, 2012) and protein MTases (Islam et al., 2011; Wang et al., 2011). Although some exceptions can occur, as DNA MTase TaqI (Dalhoff et al., 2006; Lukinavičius et al., 2007). Still, the engineered enzymes, including eM.HhaI, with a wider cofactor binding pocket use the synthetic cofactors more efficiently than the AdoMet. This makes the system partially bioorgthogonal: the enzyme is still active with AdoMet, but the natural cofactor is easily outcompeted by a synthetic one; and although some minor possibility that some wild type enzymes will also use the synthetic cofactor exists, most of them would not use it. We have shown that using this partially bioorthogonal cofactor-MTase system DNA can be specifically modified in cell lysate, regardless of the presence of other MTases and the natural cofactor. And finally when DNA is modified with functional groups suitable for Cu-free click chemistry, the second step of labeling can also be performed there in cell lysate.

During the past years, a number of mTAG cofactors with short side chains were synthesized and studied for derivatization and labeling of DNA, RNA, and protein targets (Artyukhin and Woo, 2012; Willnow et al., 2012; Bothwell 2012; Peters et al., 2010). These cofactors were typically intended for work with wild type MTases. This strategy permits interrogation of a wide spectrum of cellular enzymes and even entire methylomes in cells or cell extracts. However, our selection of a larger hex-2-ynyl moiety as the basic transfer unit permits a better accessibility and enhanced reactivity of the attached terminal group, which is beneficial for efficient conjugation of the derivatized molecule with other bulky compounds or large biomolecules. In the case of DNA modification, bulky moieties attached on longer linear side chains are generally better tolerated by DNA polymerases as compared to shorter linkers (Weisbrod and Marx, 2008) permitting efficient amplification and analysis of subsequently labeled DNA by PCR and sequencing techniques. Combination of different MTases and cofactor chemistries potentially permits a multicolor sequence-selective labeling of DNA.

Useful *in vitro* application of mTAG labeling using azide cofactor has recently been demonstrated for epigenomics studies by covalent tagging and enrichment of DNA at unmodified, epigenetically unlabeled sites of human genome (Kriukienė et al., 2013).

The variations of sequence-specific RNA and DNA labeling that we have presented here combining mTAG approach and bioorthogonal azide-alkyne cycloaddition (click chemistry) open new possibilities for use of these methods in life sciences, medical diagnostics, bionanotechnologies, etc.

CONCLUSIONS

1. The *Pyrococcus abyssi* C/D ribonucleoprotein complex (RNP) catalyzes sequence-specific transfer of extended side chains (up to 5 carbon units long) from synthetic analogues of the natural *S*-adenosyl-L-methionine (AdoMet) cofactor to a target nucleotide in substrate RNA *in vitro*.

2. C/D RNAs with synthetically reprogrammed guide sequences can be used to direct the C/D RNP-dependent transfer of a prop-2-ynyl group from the synthetic SeAdoYn cofactor to predetermined nucleotides in substrate RNAs.

3. The C/D RNP-dependent transfer of the prop-2-ynyl group followed by azide-alkyne cycloaddition can be used for sequence-specific "click" labeling of a variety of RNA substrates *in vitro*.

4. The HhaI DNA methyltransferase (variant Q82A/Y254S/N204A) catalyzes efficient sequence-specific transfer of hex-2-ynyl side chains containing terminal alkyne or azide groups from synthetic cofactor analogues (Ado-6-alkyne or Ado-6-azide, respectively) to DNA *in vitro*. Both the enzymatic transfer and subsequent "click" coupling of a fluorophore can be performed in bacterial cell lysates.

LIST OF PUBLICATIONS

The doctoral dissertation is based on the following publications:

1. **Tomkuvienė, M.**, Clouet-d'Orval, B., Černiauskas, I., Weinhold, E., Klimašauskas, S. (2012) Programmable sequence-specific click-labeling of RNA using archaeal box C/D RNP methyltransferases. Nucleic Acids Res. 40(14):6765-73.
2. Lukinavičius, G., **Tomkuvienė, M.**, Masevičius, V., Klimašauskas, S. (2013) Enhanced chemical stability of AdoMet analogues for improved methyltransferase-directed labeling of DNA. ACS Chem Biol. Epub. 04.10.

Other publications:

1. Sergeeva, O.V., Burakovsky, D.E., Sergiev, P.V., Zatsepin, T.S., **Tomkuvienė, M.**, Klimašauskas, S., Dontsova O.A. (2012) Usage of rRNA-methyltransferase for site-specific fluorescent labeling. Moscow University Chemistry Bulletin 67 (2), 88-93.

Related conference presentations:

1. **Tomkuvienė, M.**, Clouet-d'Orval, B., Černiauskas, I., Weinhold, E., Klimašauskas, S. Programmable sequence-specific click-labeling of RNA using archaeal box C/D RNP methyltransferases. „The 18th annual meeting of the RNA Society", Davos, Switzerland, 2013.06.11-16. Poster presentation.
2. **Tomkuvienė, M.**, Clouet-d'Orval, B., Černiauskas, I., Weinhold, E., Klimašauskas, S. Programmable sequence-specific click-labeling of RNA using archaeal box C/D RNP methyltransferases. „The 22nd IUBMB & 37th FEBS Congress", Seville, Spain, 2012.09.04-09. Poster presentation.
3. **Tomkuvienė, M.**, Clouet-d'Orval, B., Černiauskas, I., Weinhold, E., Klimašauskas, S. Programmable sequence-specific click-labeling of RNA using archaeal box C/D RNP methyltransferases. XIIth Meeting of the Lithuanian Biochemical Society. Molėtai, Lithuania, 2012.06.28-30. Poster presentation. „Thermo Scientific" best poster award.
4. **Tomkuvienė, M.**, Clouet-d'Orval, B., Černiauskas, I., Weinhold, E., Klimašauskas, S. Programmable sequence-specific click-labeling of RNA using archaeal box C/D RNP methyltransferases. IInd young scientist conference „Interdisciplinary research in physical and technological sciences" organized by Lithuanian Academy of Sciences. Vilnius, Lithuania. 2012.02.14. Oral presentation. Best presentation diploma.
5. **Gudeliauskaitė, M.**, Clouet-d'Orval, B., Lukinavičius, G. and Klimašauskas, S. Reconstitution of archaeal C/D Box sRNPs: towards synthetically programmable sequence-specific labelling of RNA. XIth Meeting of the Lithuanian Biochemical Society. Molėtai, Lithuania, 2010.06.15-17. Poster presentation.
6. **Gudeliauskaitė, M.**, Clouet-d'Orval, B., Lukinavičius, G. and Klimašauskas, S. Reconstitution of archaeal C/D Box sRNPs: towards synthetically programmable

sequence-specific labelling of RNA. „33rd FEBS Congress & 11th IUBMB Conference“, Athens, Greece, 2008.06.28-07.03. Poster presentation.

7. **Gudeliauskaitė, M.**, Clouet-d'Orval, B., Lukinavičius, G. ir Klimašauskas, S. Programuojama RNR molekulių modifikacija archėjų C/D sRNP pagalba. Xth Meeting of the Lithuanian Biochemical Society. Molėtai, Lithuania, 2008.06.20-22. Poster presentation.
8. **Gudeliauskaitė, M.**, Clouet-d'Orval, B. and Klimašauskas, S. Reconstitution of an archaeal C/D box RNP containing an sRNA K-loop structure. FEBS Workshop „DNA and RNA modification Enzymes: Comparative Structure, Mechanism, Function and Evolution“, Aussois, France, 2007.09.11-16. Poster presentation.

CURRICULUM VITAE

Name	Miglė Tomkuvienė
Date of birth	1982 01 11
Address	Department of Biological DNA modification, Vilnius University, Institute of Biotechnology, V.A. Graičiūno 8, LT-02241 Vilnius
Phone	+370 5 2602118
E-mail	migleg@ibt.lt migleg11@gmail.com

Education and Employment

2010-now Junior research fellow at Vilnius University Institute of Biotechnology, Department of Biological DNA modification

2006-2013 PhD student of Biochemistry (nucleic acids, protein synthesis - P 320) at Vilnius University, Vilnius, Lithuania.

2006 Master of Sciences. Biology, Genetics. Vilnius University, Vilnius, Lithuania. *Magna Cum Laude.*

2004 Bachelor of Sciences. Biology, Molecular Biology. Vilnius University, Vilnius, Lithuania. *Cum Laude.*

Outgoing research

2013, 2011, 2008, 2007, 2006. Studies of archaeal C/D RNP under supervision of dr. B. Clouet-d'Orval. 1-3 week-long visits to the CNRS Laboratory of Microbiology and Molecular Genetics, prof. A.J. Carpousis group, Paul Sabatier University, Toulouse, France.

FINANCIAL SUPPORT

This study was in part supported by grants from European Social Fund under the Global Grant measure (VP1-3.1-ŠMM-07-K-01-105), USA National Institute of Health (NIH, R21HG004535), Lithuania-France bilateral collaboration program „Žiliberas“ (Gilibert) (SUT-157) and FP7-RegPot-2009-1 (245721, MoBiLi).

ACKNOWLEDGEMENTS

I am grateful to prof. Saulius Klimašauskas for the offer of PhD studies under his supervision and interesting research topics. For excellent guidance and sharing his immense experience. For raising my competence, which lets me seek the PhD degree and gives a solid background for the future career.

I thank my first supervisor dr. Rūta Gerasimaitė for a thorough introduction and training on scientific work. I thank my students Ignas Černiauskas and Janina Ličytė for a joy of having a wonderful team. Dr. Zita Liutkevičiūtė for teaching me the TLC and HPLC-MS methods. I am thankful to all colleagues in the Institute of Biotechnology for a friendly and motivating atmosphere.

I thank dr. Gražvydas Lukinavičius, dr. Michael Martin, dr. Sophie Willnow, prof. Elmar Weinhold and dr. Viktoras Masevičius for synthetic cofactor analogues. Giedrė Urbanavičiūtė for preparation of eM.HhaI. Dr. Giedrius Sasnauskas for preparation of R.McrBC. Dr. Arvydas Kanopka for the plasmid carrying the rabbit β-globin gene. Dr. Gintautas Žvirblis for pET43.1a(+) plasmid.

Merci dr. Beatrice Clouet-d‘Orval and all colleagues in LMGM for both the work experience and a wonderful touch of life of the Southern France.

Finally my warmest appreciation goes to my family and friends for their continuous interest, understanding and support.

REZIUMĖ

DNR ir RNR polinukleotidinėse grandinėse išsidėstant keturiems pagrindiniams nukleotidams susidaro unikalios sekos. Tiriant nukleorūgštis, neretai svarbu prijungti įvairius reporterinius ar gimiringumo žymenis tik prie pasirinktų sekų, tam tikrose vietose – t.y. specifiškai. Vienas iš naujų molekulinių įrankių, skirtų specifiniam biomolekulių žymėjimui, pavadintas mTAG (angl. *methyltransferase-directed transfer of activated groups*), pagrįstas tuo, kad fermentai metiltransferazės (MTazės), natūraliai naudojančios kofaktorių AdoMet, gali būti aktyvios ir su sintetiniais jo analogais, turinčiais ilgesnes nei metil- pernešamas grandines (Dalhoff ir kt., 2006; Lukinavičius ir kt., 2007; Klimašauskas ir Weinhold, 2007). Jei šios grandinės turi galines funkcines grupes, prie jų vėliau cheminių reakcijų pagalba galima prijungti norimus žymenis. Tokiam prijungimui itin patogi yra azidų-alkinų 1,3-cikloprijungimo (AAC), dar vadinama „click“, reakcija, kuri efektyviai vyksta fiziologinėmis sąlygomis ir yra bioortogonali (šios funkcinės grupės nėra randamos jokiose žinomose gamtinėse molekulėse ir su jomis nereaguoja) (Kolb ir kt., 2001).

Pirmieji bandymai pritaikyti mTAG metodą DNR žymėjimui (Lukinavičius ir kt., 2007) parodė, kad siekiant efektyvios ilgų grandinių pernašos, reikalingas DNR MTazės katalitinio centro pertvarkymas baltymų inžinerijos būdu ir optimizuota sintetinių kofaktorių struktūra (Lukinavičius ir kt., 2012). Kolegoms susintetinus atitinkamus kofaktorius, čia aprašomų darbų metu buvo parodyta, kad DNR metiltransferazė HhaI (variantas Q82A/Y254S/N204A) efektyviai katalizuoja heks-2-inil- fragmentų, su 6-toje padėtyje esančiomis azido- arba alkinil- grupėmis, pernašą nuo sintetinių kofaktoriaus analogų ant DNR specifinėse sekose. Taip pat parodyta, kad naudojant šią MTazės ir kofaktorių sistemą bei AAC, visą specifinio DNR žymėjimo procesą galima atlikti sudėtinguose biologiniuose mišiniuose, įskaitant bakterijų ląstelių lizatus.

RNR žymėjimui naudotinos atitinkamos RNR MTazės. Archėjose ir eukariotuose randamos MTazės C/D ribonukleoproteininiai kompleksai (C/D RNP), sudaryti iš RNR ir keleto baltymų. Šie RNP kompleksai pasižymi unikaliu substrato atpažinimo mechanizmu: komplekso sudėtyje esanti RNR, vadinama kreipiančiąja, komplementarumo principu atpažįsta 11 nt ilgio seką substratinėje RNR ir nukreipia metiltransferazinį aktyvumą į konkretaus nukleotido ribozės 2’-O atomą. Taikinį atpažįstančią RNR dalį galima keisti, kartu atitinkamai pasikeičiant ir atpažįstamai RNR sekai (Appel ir Maxwell, 2007). Šiame darbe pirmąkart parodyta, kad archėjos *Pyrococcus abyssi* C/D RNP gali būti aktyvus su 2-5 atomų ilgio pernešamas grandines turinčiais sintetiniais kofaktoriais. Sintetiškai keičiant kreipiančiąją RNR, prop-2-inilgrupės pernašą galima nukreipti į norimas substratinių RNR vietas. Galiausiai, naudojant C/D RNP katalizuojamą prop-2-inilgrupės pernašą ir azidų-alkinų cikloprijungimo reakciją, buvo sėkmingai fluorescencine žyme pažymėti septyni iš aštuonių pasirinktų taikinių modelinėse tRNR ir iRNR molekulėse. Taigi, naujasis molekulinis įrankis gali būti sintetiškai užprogramuojamas RNR molekules pažymėti norimose sekose.

Šios naujos specifinio nukleorūgščių žymėjimo galimybės gali būti pritaikytos biochemijos, biomedicinos, nanotechnologijų ir kitose tyrimų srityse bei kuriant medicinos diagnostikos priemones.

REFERENCES

1. Aittaleb, M., Visone, T., Fenley, M. O., & Li, H. (2004). Structural and thermodynamic evidence for a stabilizing role of Nop5p in S-adenosyl-L-methionine binding to fibrillarin. *The Journal of biological chemistry*, *279*(40), 41822–9.
2. Appel, C. D., & 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 (New York, N.Y.)*, *13*(6), 899–911.
3. Artyukhin, A. B., & Woo, Y.-H. (2012). DNA extraction method with improved efficiency and specificity using DNA methyltransferase and "click" chemistry. *Analytical biochemistry*, *425*(2), 169–74.
4. Bleichert, F., Gagnon, K. T., Brown, B. A., Maxwell, E. S., Leschziner, A. E., Unger, V. M., & Baserga, S. J. (2009). A dimeric structure for archaeal box C/D small ribonucleoproteins. *Science (New York, N.Y.)*, *325*(5946), 1384–7.
5. Bortolin, M.-L., Bachellerie, J.-P., & 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 research*, *31*(22), 6524–35.
6. Bothwell, I. R., Islam, K., Chen, Y., Zheng, W., Blum, G., Deng, H., & Luo, M. (2012). Se-adenosyl-L-selenomethionine cofactor analogue as a reporter of protein methylation. *Journal of the American Chemical Society*, *134*(36), 14905–12.
7. Cavaillé, J., Nicoloso, M., & Bachellerie, J. P. (1996). Targeted ribose methylation of RNA in vivo directed by tailored antisense RNA guides. *Nature*, *383*(6602), 732–5.
8. Cheng, X. and Blumenthal, R. (1999). *S-adenosylmethionine-dependent Methyltransferases: Structures and Functions* (p. 400).
9. Dalhoff, C., Lukinavicius, G., Klimasăuskas, S., & Weinhold, E. (2006). Direct transfer of extended groups from synthetic cofactors by DNA methyltransferases. *Nature chemical biology*, *2*(1), 31–2.
10. Daujotyte, D., Vilkaitis, G., Manelyte, L., Skalicky, J., Szyperski, T., & Klimasauskas, S. (2003). Solubility engineering of the HhaI methyltransferase. *Protein engineering*, *16*(4), 295–301.
11. Deng, L., Starostina, N. G., Liu, Z.-J., Rose, J. P., Terns, R. M., Terns, M. P., & Wang, B.-C. (2004). Structure determination of fibrillarin from the hyperthermophilic archaeon Pyrococcus furiosus. *Biochemical and biophysical research communications*, *315*(3),
12. Feder, M., Pas, J., Wyrwicz, L. S., & Bujnicki, J. M. (2003). Molecular phylogenetics of the RrmJ/fibrillarin superfamily of ribose 2'-O-methyltransferases. *Gene*, *302*(1-2), 129–38.
13. Gerasimaitė, R., Merkienė, E., & Klimašauskas, S. (2011). Direct observation of cytosine flipping and covalent catalysis in a DNA methyltransferase. *Nucleic acids research*, *39*(9), 3771–80.
14. Grosjean, H., Droogmans, L., Roovers, M., & Keith, G. (2007). Detection of enzymatic activity of transfer RNA modification enzymes using radiolabeled tRNA substrates. *Methods in enzymology*, *425*, 55–101.
15. Hager, J., Staker, B. L., Bugl, H., & Jakob, U. (2002). Active site in RrmJ, a heat shock-induced methyltransferase. *The Journal of biological chemistry*, *277*(44), 41978–86.
16. Hardin, J. W., & 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 research*, *34*(18), 5039–51.

17. Islam, K., Zheng, W., Yu, H., Deng, H., & Luo, M. (2011). Expanding cofactor repertoire of protein lysine methyltransferase for substrate labeling. *ACS chemical biology*, *6*(7), 679–84.
18. Kanopka, A., Mühlemann, O., & Akusjärvi, G. (1996). Inhibition by SR proteins of splicing of a regulated adenovirus pre-mRNA. *Nature*, *381*(6582), 535–8.
19. Klein, D. J., Schmeing, T. M., Moore, P. B., & Steitz, T. A. (2001). The kink-turn: a new RNA secondary structure motif. *The EMBO journal*, *20*(15), 4214–21.
20. Klimašauskas, S, Kumar, S., Roberts, R. J., & Cheng, X. (1994). HhaI methyltransferase flips its target base out of the DNA helix. *Cell*, *76*(2), 357–69.
21. Klimašauskas, Saulius, & Weinhold, E. (2007). A new tool for biotechnology: AdoMet-dependent methyltransferases. *Trends in biotechnology*, *25*(3), 99–104.
22. Kolb, H. C., Finn, M. G., & Sharpless, K. B. (2001). Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angewandte Chemie (International ed. in English)*, *40*(11), 2004–2021.
23. Kriukienė, E., Labrie, V., Khare, T., Urbanavičiūtė, G., Lapinaitė, A., Koncevičius, K., … Klimašauskas, S. (2013). DNA unmethylome profiling by covalent capture of CpG sites. *Nature communications*, *4*, 2190.
24. Lin, J., Lai, S., Jia, R., Xu, A., Zhang, L., Lu, J., & Ye, K. (2011). Structural basis for site-specific ribose methylation by box C/D RNA protein complexes. *Nature*, *469*(7331), 559–63.
25. Lukinavičius, G., Lapienė, V., Stasevskij, Z., Dalhoff, C., Weinhold, E., & Klimašauskas, S. (2007). Targeted labeling of DNA by methyltransferase-directed transfer of activated groups (mTAG). *Journal of the American Chemical Society*, *129*(10), 2758–2759.
26. Lukinavičius, G., Lapinaite, A., Urbanaviciute, G., Gerasimaite, R., & Klimasauskas, S. (2012). Engineering the DNA cytosine-5 methyltransferase reaction for sequence-specific labeling of DNA. *Nucleic acids research*, *40*(22), 11594–602.
27. Lukinavičius, G., Tomkuvienė, M., Masevičius, V., & Klimašauskas, S. (2013). Enhanced Chemical Stability of AdoMet Analogues for Improved Methyltransferase-Directed Labeling of DNA. *ACS chemical biology*.
28. Motorin, Y., Burhenne, J., Teimer, R., Koynov, K., Willnow, S., Weinhold, E., & Helm, M. (2011). Expanding the chemical scope of RNA:methyltransferases to site-specific alkynylation of RNA for click labeling. *Nucleic acids research*, *39*(5), 1943–52.
29. Motorin, Y., & Helm, M. (2011). RNA nucleotide methylation. *Wiley interdisciplinary reviews. RNA*, *2*(5), 611–31.
30. Motorin, Y., Muller, S., Behm-Ansmant, I., & Branlant, C. (2007). Identification of modified residues in RNAs by reverse transcription-based methods. *Methods in enzymology*, *425*, 21–53.
31. Nagy, K., Orbán, E., Bősze, S. and Kele, P. (2010), Clickable Long-Wave "Mega-Stokes" Fluorophores for Orthogonal Chemoselective Labeling of Cells. Chem. Asian J., 5: 773–777
32. Neely, R. K., Dedecker, P., Hotta, J., Urbanaviciute, G., Klimasauskas, S., & Hofkens, J. (2010). DNA fluorocode: A single molecule, optical map of DNA with nanometre resolution. *Chemical Science*, *1*(4), 453–460.
33. Nolivos, S., Carpousis, A J., 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 Research*, *33*(20), 6507–6514.
34. Peters, W., Willnow, S., Duisken, M., Kleine, H., Macherey, T., Duncan, K. E., Weinhold, E. (2010). Enzymatic site-specific functionalization of protein

methyltransferase substrates with alkynes for click labeling. *Angewandte Chemie (International ed. in English)*, *49*(30), 5170–3.

35. Roberts, R. J., & Cheng, X. (1998). Base flipping. *Annual review of biochemistry*, *67*, 181–98.
36. 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.
37. Schneider, K. L., Pollard, K. S., Baertsch, R., Pohl, A., & Lowe, T. M. (2006). The UCSC Archaeal Genome Browser. *Nucleic acids research*, *34*(Database issue), D407–10.
38. Sukackaite, R., Grazulis, S., Tamulaitis, G., & Siksnys, V. (2012). The recognition domain of the methyl-specific endonuclease McrBC flips out 5-methylcytosine. *Nucleic acids research*, *40*(15), 7552–62.
39. Tollervey, D., Lehtonen, H., Jansen, R., Kern, H., & Hurt, E. C. (1993). Temperature-sensitive mutations demonstrate roles for yeast fibrillarin in pre-rRNA processing, pre-rRNA methylation, and ribosome assembly. *Cell*, *72*(3), 443–57.
40. 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.
41. Tran, E., Zhang, X., Lackey, L., & 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 (New York, N.Y.)*, *11*(3), 285–93.
42. Wang, R., Zheng, W., Yu, H., Deng, H., & Luo, M. (2011). Labeling substrates of protein arginine methyltransferase with engineered enzymes and matched S-adenosyl-L-methionine analogues. *Journal of the American Chemical Society*, *133*(20), 7648–51.
43. Weisbrod, S. H., & Marx, A. (2008). Novel strategies for the site-specific covalent labelling of nucleic acids. *Chemical communications (Cambridge, England)*, (44), 5675–85.
44. Willnow, S., Martin, M., Lüscher, B., & Weinhold, E. (2012). A selenium-based click AdoMet analogue for versatile substrate labeling with wild-type protein methyltransferases. *Chembiochem : a European journal of chemical biology*, *13*(8), 1167–73.