Published in final edited form as:

Angew Chem Int Ed Engl. 2017 April 24; 56(18): 5017–5020. doi:10.1002/anie.201700537.

Selective enzymatic demethylation of N^2,N^2 -dimethylguanosine in RNA and its application in high-throughput tRNA sequencing

Qing Dai^{+,1}, Guanqun Zheng^{+,2}, Michael H. Schwartz^{+,2}, Wesley C. Clark², and Tao Pan^{+,2}

¹Department of Chemistry, The University of Chicago, Chicago, IL 60637, USA

²Department of Biochemistry and Molecular Biology, The University of Chicago, Chicago, IL 60637, USA

Abstract

The abundant Watson-Crick face methylations in biological RNAs such as N^1 -methyladenosine (m^1A), N^1 -methylguanosine (m^1G), N^3 -methylcytosine (m^3C), and N^2,N^2 -dimethylguanosine (m_2^2G) cause significant obstacles for high throughput RNA sequencing by impairing cDNA synthesis. One strategy to overcome this obstacle is to remove the methyl group on these modified bases prior to cDNA synthesis using enzymes. The wild-type *E. coli* AlkB and its D135S mutant can remove most of m^1A , m^1G , m^3C modifications in transfer RNA (tRNA), but they work poorly on m_2^2G . Here we report the design and evaluation of a series of AlkB mutants against m_2^2G -containing model RNA substrates that we synthesize using an improved chemical synthesis method. We show that the AlkB D135S/L118V mutant efficiently and selectively converts m_2^2G modification to N^2 -methylguanosine (m^2G). We also show that this new enzyme lowers the amount of m_2^2G in tRNA and improves the efficiency of tRNA sequencing.

Keywords

demethylase; tRNA seq; AlkB mutant; N^2,N^2 -dimethylguanosine; N^2 -methylguanosine

tRNAs decode mRNA codons and are essential for cells. In humans, tRNA abundance is tissue dependent^[1] and tRNA expression and mutations are known to be associated with neurological pathologies and cancer development and proliferation.^[2] Efficient and quantitative tRNA sequencing methods are crucial for biological studies of tRNA. However, standard sequencing methods are very ineffective for tRNA-seq, primarily due to the high levels of four Watson-Crick face methylations, m^1A , m^1G , m^3C and m_2^2G that impair cDNA synthesis, although attempts have been made.^[3] Recently, we reported an efficient and quantitative high-throughput tRNA sequencing method (DM-tRNA-seq) by using a highly processive thermostable group II intron reverse transcriptase, and importantly, two demethylases that remove most base methylations and thus significantly improved the tRNA sequencing efficiency: for example, the amount of longer and full-length cDNA was substantially increased after the treatment of tRNA with demethylases.^[4] Despite this success, our sequencing data showed that the m_2^2G modification at position 26 of tRNAs

* taopan@uchicago.edu.

⁺These authors contributed equally to this work.

was not effectively demethylated and caused a significant stop in cDNA synthesis. To alleviate this problem, it would be highly desirable to identify a third demethylase with enhanced activity toward m_2^2G modification in RNA and apply it for tRNA sequencing.

Previously, Zhu *et al* showed that mutating active-site residues of an RNA/DNA demethylase switched its demethylation activity from m^1A to m^6A modification^[5]. We have also shown that an AlkB D135S mutant demethylates m^1G much more efficiently than the wild type enzyme.^[4] Therefore, we reasoned that further mutagenesis of residues around the active site of AlkB might produce a demethylase with selectively higher activity for our target m_2^2G modification.

Unlike m^1A , m^1G and m^3C which only have one methyl group in their structures, m_2^2G has two methyl groups at the exocyclic amine (Figure 1A). For sequencing purposes, removal of one methyl group to N^2 -methylguanosine (m^2G) would be sufficient for highly efficient cDNA synthesis by the thermophilic reverse transcriptase.^[4] m_2^2G modification is present in about half of human tRNAs at position 26 at the junction between the D-arm and the anticodon arm,^[6] as indicated in the tertiary structures of tRNA (Figure 1B). The $m_2^2G_{26}$ residue is buried in the structure, making it particularly difficult to be demethylated. The co-crystal structure of *E. coli* AlkB with a nucleic acid substrate^[7] was used to model the m_2^2G base into the active site (Figure 1C). AlkB D135S mutant we used previously for DM-tRNA-seq works efficiently toward m^1G , in which the shorter side chain of S135 not only creates more room to accommodate guanosine but also forms the crucial hydrogen bond to stabilize guanosine in the active site. Since such interaction would remain favored to accommodate m_2^2G in the pocket, the D135S mutant is used as our starting point of m_2^2G demethylase design. We reasoned that shortening the side chain of amino acids within the van de Waal's radii of the methyl groups on m_2^2G would allow better accommodation for m_2^2G , while leaving the rest of the active site intact would potentially enhance the activity towards m_2^2G . Therefore, three mutants, Arg210-to-Lys, Leu118-to-Val, and Met61-to-Ser were prepared and tested towards m_2^2G containing RNA substrates.

To test the activity of these potential m_2^2G demethylases, we synthesized three 9mer RNA oligos as demethylase substrates. The oligo sequence mimics the human tRNA^{Phe} sequence around the position of 26 (5'-GAGCXUUAG, X= m_2^2G , m^2G or G) (Scheme 1). We first improved the synthesis of the m_2^2G and m^2G phosphoramidites as follows. In the previous reports of m_2^2G and m^2G phosphoramidite synthesis,^[8] it was generally believed that the protection of O^6 of guanosine with *p*-nitrophenylethyl (NPE) group would not only lower the polarity and increase the solubility of guanosine derivatives to facilitate their purification, but also enhance the nucleophilicity of the exocyclic amine. Subsequent conversion of 2-amine to 2-methylamine and 2-dimethylamine were accomplished through a multiple-step reaction to convert 2-NH₂ to 2-F, followed by substitution of 2-F with methylamine and dimethylamine. The use of fluoride reagent to introduce 2-F precluded the application of Beigelman's optimal silyl protecting strategy for phosphoramidite synthesis^[9] and therefore several extra steps were needed for protecting group exchanges.

Several methods have been reported to prepare m^2G and m_2^2G derivatives^[10] but the most straightforward one used a "one-step" reductive amination reaction to methylate the

exocyclic amine directly.^[10b] To improve the synthesis, we attempted to combine the reductive amination reaction with Beigelman's protecting strategy to synthesize both phosphoramidites **7a** and **7b** efficiently at the same time. The main concern is whether the silyl protecting groups can survive the reaction conditions of protecting *O*⁶ and the acidic conditions used for reductive amination. We started our synthesis from guanosine **1** by selective protecting 3',5'-hydroxyls with di-*t*-butylsilyl group and 2'-hydroxyl with TBDMS in a "one-pot" reaction to give intermediate **2** as a pale solid precipitated from ethyl ester in 75% yield.^[9] Protection of *O*⁶ using *p*-nitrophenylethyl group by the treatment of **2** with *p*-nitrophenylethanol in the presence of triphenylphosphine (Ph₃P) and diisopropyl azodicarboxylate (DIAD) gave **3** in 92% yield. We found that at elevated temperature (100 °C), the reaction took place very quickly as indicated by the instant disappearance of cloudy **2** in 1,4-dioxane with the addition of DIAD.^[11] Under such conditions, both silyl protecting groups were stable. The reductive amination reaction between the exocyclic amine in **3** with paraformaldehyde and sodium cyanoborohydride (NaBH₃CN) in acetic acid at 40 °C resulted in sequential methylation of exocyclic amine to give a mixture of *N*²-methyl and *N*²,*N*²-dimethyl guanosine derivatives **4a** and **4b**. Depending on the amount of paraformaldehyde and NaBH₃CN used, either **4a** or **4b** could be obtained as major products. All the protecting groups in intermediate **3** remained stable under the acidic reaction conditions. Treatment of **4a/4b** with hydrogen fluoride pyridine removed the 3',5'-di-*t*-butylsilyl group quantitatively while keeping 2'-TBDMS intact to give intermediates **5a/5b**. 5'-trytylation gave the mixture of **6a/6b**. Although two types of methylated derivatives could be readily separated by column chromatography at **4a/4b** or **5a/5b** stage, it is more efficient to separate **6a** and **6b** by column chromatography. Subsequent 3'-phosphitylation **6a** and **6b** under standard conditions provided m₂²G phosphoramidite **7a** and m²G phosphoramidite **7b** in 88% and 85% yield, respectively.

Phosphoramidite **7a** or **7b** was then incorporated into RNA oligos with similar efficiency as the commercial unmodified guanosine phosphoramidite to prepare the oligos **ON-m₂²G** and **ON-m²G**. As a control, we also synthesized an unmodified oligo **ON-G** with the same sequence. After deprotection and HPLC purification, the structures of three RNA oligos were confirmed by Maldi-TOF MS, which showed 14 Dalton difference in mass between **ON-m₂²G** and **ON-m²G** and between **ON-m²G** and **ON-G**, as expected (Figure 2B).

We tested the demethylation activities of the AlkB mutant proteins using these RNA oligos as substrates. Oligo **ON-m₂²G** (1 nmol) was incubated with each enzyme in the reaction buffer containing 300 mM KCl, 2 mM MgCl₂, 50 μM of (NH₄)₂Fe(SO₄)₂·6H₂O, 300 μM α-ketoglutarate (*α*-KG), 2 mM L-ascorbic acid, 50 μg/ml BSA and 50 mM MES buffer (pH 5.0) for 2 h at room temperature, and then quenched by the addition of 5 mM EDTA. RNA oligos in the reaction were recovered by ethanol precipitation and dissolved in water. Part of samples were digested to free nucleosides by treatment of nuclease P1 and phosphatase and then analyzed by HPLC (Figure 2A), while the rest of samples were directly analyzed by MALDI-TOF MS (Figure 2B). As shown in Figure 2A, the D135S/L118V mutant gave the best results among the five AlkB enzymes. The m₂²G peak disappeared completely with the appearance of m²G accordingly; suggesting that one methyl group from m₂²G was fully removed. Other enzymes such as the wild-type AlkB, D135S or D135S/R210K mutant

showed limited activities while D135S/M61S mutant showed no activity at all toward m^2_2G . The Maldi-TOF MS of the sample treated with D135S/L118V mutant also showed complete disappearance of the peak at m/z 2922 and appearance of a new peak at 2908 which is identical to that of oligo **ON- m^2G** , indicating that one of two methyl groups in the m^2_2G oligo **ON- m^2_2G** was removed completely. No peak identical to that of **ON-G** was observed ($m/z = 2894$), again indicating that only one methyl was removed. Indeed, when we treated **ON- m^2G** with AlkB D135S/L118V mutant, both HPLC and MALDI-TOF showed no reaction, further confirming that AlkB mutants can only remove one methyl group from m^2_2G . The shorter side chain of Val118 compared with original Leu118, which sits directly opposite to methyl groups on m^2_2G , results in more space for m^2_2G to fit in the active site to be demethylated.

We also used a 9mer oligo containing an m^1G modification^[4] to test whether the AlkB D135S/L118V mutant can also demethylate m^1G efficiently. Unlike AlkB D135S mutant which demethylated m^1G quantitatively, we observed that AlkB D135S/L118V mutant had only limited activity toward m^1G , suggesting that AlkB D135S and D135S/L118V mutants can have distinct substrate selectivity (SI, Figure 1S). To obtain the optimal result in tRNA sequencing, it is therefore best to use a three-demethylase mixture: wild-type AlkB (to remove m^1A and m^3C), D135S (to remove m^1G) and D135S/L118V (to remove m^2_2G).

We then tested whether AlkB D135S/L118V mutant improves m^2_2G demethylation in tRNA. Like our previous DM-tRNA-seq method,^[4] we used the combination of three demethylases, wild-type, D135S and D135S/L118V to treat tRNA samples from HEK293T cells and constructed sequencing libraries to evaluate the effect of AlkB D135S/L118V mutant on high-throughput tRNA sequencing. As controls, we also sequenced the same samples without demethylase treatment or treated with just two demethylases, wild-type and D135S. To evaluate the effect of the D135S/L118V demethylase on m^2_2G , we applied a quantitative measure, modification index (MI) which is the sum of the fraction of sequencing reads derived from RT mutations (from G to A/C/T in this case) and stop at the m^2_2G 26 in each tRNA^[6b] (Figure 3A). Removal of the m^2_2G methyl group would result in increased cDNA synthesis through this nucleotide and incorporation of the correct nucleotide. Therefore, reduced level of m^2_2G 26 would decrease the MI values at this position. Indeed, treatment using the three-demethylase mixture (+3DM) reduced MI values for the majority of all known tRNAs with known m^2_2G 26 modifications at greater extent compared to using the two-demethylase mixture (+2DM). Because m^2_2G 26 is located in a very structured region of tRNA (Figure 1B), the absolute demethylation fraction is still substantially lower than the m^1A , m^1G and m^3C modifications shown in the previous study^[4]. Among the cytosolic tRNAs, tRNA^{Thr}, tRNA^{Trp} and tRNA^{Tyr} show a greater extent of MI reduction than other tRNAs (Figure 3A); this result may be attributed to the variable stabilities of cytosolic tRNAs that carry many modifications of distinct types at distinct sites. Consistent with the idea that the low absolute m^2_2G demethylation fraction is related to tRNA structure, the mitochondrial tRNA^{Ile} has the highest level of MI value reduction (Figure 3B), because mitochondrial tRNAs are markedly less structured than the cytosolic tRNAs^[6b].

In summary, we have identified an AlkB mutant (D135S/L118V) that can efficiently remove one methyl group from the m^2_2G modification to convert it to m^2G . Unlike m^2_2G , m^2G does

not block reverse transcription reaction under our sequencing conditions, resulting in improved tRNA sequencing efficiency. Different AlkB mutants showed demethylation selectivity on different base methylations, suggesting a role of the active-site residues in substrate recognition and the possibility to find other AlkB mutants to selectively demethylate other base modifications such as dimethyl adenosine. Besides the application of these demethylases in tRNA sequencing, these selective demethylases may also serve as useful tools for developing single-base resolution sequencing methods for base methylations in other RNA species such as mRNA,^[12] as well as screening for selective inhibitors that are capable of discriminating between these closely related demethylases.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by National Institute of Health grants in USA (K01 HG006699 to QD, and RM1HG008935 to TP). We thank Dr. Matthew Eckwahl-Sanna for helping edit the figures.

References

1. Dittmar KA, Goodenbour JM, Pan T. PLoS Genet. 2006; 2:e221. [PubMed: 17194224]
2. a Kirchner S, Ignatova Z. Nat Rev Genet. 2015; 16:98–112. [PubMed: 25534324] b Abbott JA, Francklyn CS, Robey-Bond SM. Front Genet. 2014; 5:158. [PubMed: 24917879]
3. Pang YL, Abo R, Levine SS, Dedon PC. Nucleic Acids Res. 2014; 42:e170. [PubMed: 25348403]
4. Zheng G, Qin Y, Clark WC, Dai Q, Yi C, He C, Lambowitz AM, Pan T. Nat Methods. 2015; 12:835–837. [PubMed: 26214130]
5. Zhu C, Yi C. Angew Chem Int Ed Engl. 2014; 53:3659–3662. [PubMed: 24596302]
6. a Edqvist J, Straby KB, Grosjean H. Biochimie. 1995; 77:54–61. [PubMed: 7599276] b Clark WC, Evans ME, Dominissini D, Zheng G, Pan T. RNA. 2016; 22:1771–1784. [PubMed: 27613580]
7. Yang CG, Yi C, Duguid EM, Sullivan CT, Jian X, Rice PA, He C. Nature. 2008; 452:961–965. [PubMed: 18432238]
8. a Porcher S, Pitsch S. Helv Chim Acta. 2005; 88:2683–2704. b Rife JPC, Charles S, Moore Peter B, Strobel Scott A. Nucleosides & Nucleotides. 1998; 17:2281–2288. c Claudia Höbartner CK, Flecker Elke, Elke Ottenschläger WP, Grubmayr Karl, Micura aR. Monatshefte für Chemie. 2003; 134:851–873.
9. Serebryany V, Beigelman L. Tetrahedron Lett. 2002; 43:1983–1985.
10. a Karolina Piecyk RED, Jankowska-Anyszka Marzena. Bioorganic & Medicinal Chemistry Letters. 2012; 22:4391–4395. [PubMed: 22658555] b Sekine M, Aoyagi M, Ushioda M, Ohkubo A, Seio K. J Org Chem. 2005; 70:8400–8408. [PubMed: 16209584] c Abeydeera ND, Chow CS. Bioorg Med Chem. 2009; 17:5887–5893. [PubMed: 19628400] d Sekine M, Takahiko S. Journal of Organic Chemistry. 1991; 56:1224–1227.
11. Kamaike K, Kinoshita K, Niwa K, Hirose K, Suzuki K, Ishido Y. Nucleosides Nucleotides Nucleic Acids. 2001; 20:59–75. [PubMed: 11303563]
12. a Dominissini D, Nachtergaele S, Moshitch-Moshkovitz S, Peer E, Kol N, Ben-Haim MS, Dai Q, Di Segni A, Salmon-Divon M, Clark WC, Zheng G, Pan T, Solomon O, Eyal E, Hershkovitz V, Han D, Dore LC, Amariglio N, Rechavi G, He C. Nature. 2016; 530:441–446. [PubMed: 26863196] b Li X, Xiong X, Wang K, Wang L, Shu X, Ma S, Yi C. Nat Chem Biol. 2016; 12:311–316. [PubMed: 26863410]

Figure 1. AlkB mutant design for m²₂G demethylation

(A) Chemical structure of m²₂G. (B) Location of m²₂G modification in tRNA tertiary structure (PDB 4tna). The m²₂G residue is in orange, and the two methyl groups are in blue. (C) View of AlkB active site stereochemistry model with m²₂G coordination based on Protein data bank (PDB) ID 3BIE. m²₂G is labeled in cyan.

Figure 2.

Evaluations of a series of AlkB mutants for efficient m²G 5'-GAGCXUUAG, X= m²G (ON- m²₂G), m²G (ON-m²G) or G (ON-G). (A) HPLC traces of the digestion products after ON-m²₂G was treated with demethylases showed that AlkB D135S/L118V mutant quantitatively converts m²₂G to m²G. (B) Comparison of MALDI-TOF MS of unmodified RNA oligo ON-G, m²G-containing RNA oligo ON-m²G, m²₂G-containing RNA oligo ON-m²₂G, and ON-m²₂G which was treated with AlkB D135S/L118V mutant confirmed that the demethylation product of ON-m²₂G is ON-m²G instead of ON-G.

Figure 3.

Application of AlkB D135S/L118V to high throughput tRNA-seq. (A) tRNA-seq heatmap result showing the modification index (MI) change at position 26 in tRNAs with known m²G sites. 2DM = wild-type AlkB+D135S; 3DM = wild-type AlkB + D135S + D135S/L118V. The higher the MI value, the higher the fraction of modifications. Therefore, decreased modification fraction corresponds to lower MI value or more intense green color for the corresponding tRNA isoacceptor. (B) Full MI plot for the full-length mitochondrial tRNA^{Ile}. MI at the m²G nucleotide is 0.972, 0.343, 0.178 for untreated (black), +2DM (red), and +3DM (blue), respectively.

Scheme 1.

Improved syntheses of m²G and m²₂G phosphoramidites and their incorporation into RNA to prepare demethylase substrates. i). (a) (*t*-Bu)₂Si(OTf)₂, DMF; (b) imidazole, TBDMS-Cl, 90 °C, 75%. ii). *p*-nitrophenylethanol, Ph₃P, DIAD, 1,4-dioxane 100 °C, 92%. iii). (HCHO)_x, NaBH₃CN, AcOH, 40 °C. iv) HF·Py/THF. v). DMTr-Cl/Py. vi). (*i*-Pr)₂NP(Cl)OCH₂CH₂CN, CH₂Cl₂, (*i*-Pr)₂NEt, 1-methyl-imidazole. vii). RNA synthesis and deprotection.