Published in final edited form as:

J Am Chem Soc. 2017 December 06; 139(48): 17213–17216. doi:10.1021/jacs.7b06837.

N⁶-Allyladenosine: A New Small Molecule for RNA Labeling Identified by Mutation Assay

Xiao Shu^{†,‡}, Qing Dai^{†,‡}, Tong Wu^{†,‡}, Ian R. Bothwell[§], Yanan Yue[†], Zezhou Zhang[†], Jie Cao[†], Qili Fei[‡], Minkui Luo[§], Chuan He^{*,‡}, and Jianzhao Liu^{*,†,||}

[†]MOE Key Laboratory of Macromolecular Synthesis and Functionalization, Department of Polymer Science and Engineering, Zhejiang University, Zheda Road 38, Hangzhou 310027, China

[‡]Department of Chemistry, Department of Biochemistry and Molecular Biology, Institute for Biophysical Dynamics, Howard Hughes Medical Institute, The University of Chicago, Chicago, Illinois 60637, United States

[§]Molecular Pharmacology and Chemistry Program, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, United States

^{||}Life Sciences Institute, Zhejiang University, Yuhangtang Road 866, Hangzhou 310058, China

Abstract

RNA labeling is crucial for the study of RNA structure and metabolism. Herein we report N⁶-allyladenosine (a⁶A) as a new small molecule for RNA labeling through both metabolic and enzyme-assisted manners. a⁶A behaves like A and can be metabolically incorporated into newly synthesized RNAs inside mammalian cells. We also show that human RNA N⁶-methyladenosine (m⁶A) methyltransferases METTL3/METTL14 can work with a synthetic cofactor, namely allyl-SAM (*S*-adenosyl methionine with methyl replaced by allyl) in order to site-specifically install an allyl group to the N⁶-position of A within specific sequence to generate a⁶A-labeled RNAs. The iodination of N⁶-allyl group of a⁶A under mild buffer conditions spontaneously induces the formation of N¹,N⁶-cyclized adenosine and creates mutations at its opposite site during complementary DNA synthesis of reverse transcription. The existing m⁶A in RNA is inert to methyltransferase-assisted allyl labeling, which offers a chance to differentiate m⁶A from A at individual RNA sites. Our work demonstrates a new method for RNA labeling, which could find applications in developing sequencing methods for nascent RNAs and RNA modifications.

*Corresponding Authors: chuanhe@uchicago.edu; liujz@zju.edu.cn.

#Author Contributions

X.S., Q.D., and T.W. contributed equally to this work.

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.7b06837.

Experimental section (PDF)

ORCID

Minkui Luo: 0000-0001-7409-7034

Jianzhao Liu: 0000-0001-9465-6075

Notes

The authors declare no competing financial interest.

Graphical abstract

It is known RNA, as the courier of genetic information, is a structurally and spatiotemporally dynamic biomacromolecule. RNA labeling offers useful tools for the study of RNA structure^{1,2} and metabolism,^{3–7} as well as to identify short-lived RNAs.⁸ Recent interest in the study of naturally occurring RNA chemical modifications has increased thanks to their previously uncovered biological functions.^{9–11} Among typical examples are studies of *N*⁶-methyladenosine (m⁶A),^{12–14} pseudouridine,^{15–17} and *N*¹-methyladenosine.^{18,19} We were inspired to discover that artificially designed base analogs with a small modifiable chemical group might be employed by cellular machines as the starting material for RNA synthesis and labeling. As known, 4-thiouridine (4SU), a close analog of uridine, represents an excellent example for RNA labeling,^{3,4} but compared to uridine, its base pairing property cannot undergo further manipulation and is thus not ideal for absolute differentiation of labeled RNA from unlabeled RNAs by using reverse transcription (RT)-induced mutation assays.

Herein we present an adenosine (A) analog *N*⁶-allyladosine (a⁶A) for RNA-labeling applications (Scheme 1). On the one hand, a⁶A behaves like A and can be metabolically incorporated into newly synthesized RNAs inside mammalian cells. On the other hand, we used the idea of mammalian mRNA m⁶A methyltransferase-assisted chemical labeling²⁰ to convert specific A within RNA into a⁶A. After the a⁶A-containing RNA is subjected to iodination treatment, RT into complementary DNA (cDNA), and polymerase chain reaction (PCR), a⁶A is read as other bases T/C/G whereas the unmodified A is still read as A. We designed model reactions of the a⁶A nucleoside and the a⁶A-containing RNA oligo in order to prove the labeling-derived mutation assay. This work reveals a new RNA-labeling method and demonstrates its potential for future high-throughput sequencing applications in the RNA field.

First, the metabolic RNA-labeling behavior of a⁶A was investigated. After incubating a⁶A nucleoside with HeLa cells for a variable length of time, we extracted the total RNAs, digested them into single nucleosides, and utilized ultrahigh performance liquid chromatography coupled with triple-quadrupole tandem mass spectrometer (UHPLC-QQQ-MS/MS) in order to quantify a⁶A level (see Supporting Information (SI)).^{20,21} The a⁶A nucleoside was synthesized by a reaction of 6-chloropurine riboside with allylamine as a standard in order to generate the calibration curve (see SI).²² Results show that the a⁶A/A ratio of total RNA increases from 0.003% to 0.03% in a tested period of time from 5 min to 16 h (Figure 1A), indicating the incorporation of a⁶A into RNA through cellular metabolic pathways. Meanwhile, we colabeled RNAs with a⁶A and 4SU. 4SU-labeled RNAs can be biotinylated and enriched (see SI).^{23,24} We found that a⁶A/A values after enrichment were much higher than those of nonenriched total RNAs at the same labeling time points. a⁶A/A

increases to 0.1% for 2 h and 0.27% for 16 h. These data suggests that cellular RNAs could indeed be metabolically labeled by a⁶A.

Next, we explored m⁶A methyltransferase-assisted a⁶A labeling in potential applications. m⁶A is the most prevalent internal modification in mRNAs of higher eukaryotes at 1–3 residues per message, and carries fundamental biological functions.^{12,13} The mRNA m⁶A modification is sequence specific with GGACU predominant inside cells.^{12,13} The modified m⁶A behaves like A and reversely transcribes to T during RT. We speculated if we could synthesize an analog of methyltransferase cofactor *S*-adenosyl methionine (SAM), which carries the allyl group rather than the methyl, and combines its use with the methyltransferase in order to specifically modify A within the methylation consensus motif to form a⁶A. Prompted by this idea, we synthesized the allyl-SAM cofactor that replaces the methyl in normal SAM with the allyl group.²⁵ After purifying m⁶A methyltransferases METTL3/METTL14 heterodimer (MT) complex and synthesizing random-structured RNA oligos with/without m⁶A modification and/or consensus motif GGACU (probes **1–3**, Figure 1B),²⁰ we went on to test their allyl-transfer activity. Results show that RNA probe **1** with GGACU motif was active, whereas probe **2** with GGm⁶ACU and probe **3** with GGAUU were inactive for the allyl transfer (see SI). Next, we quantitatively characterized the modification yield for probe **1** with different equivalents of MT (Figure 1C). Increased load of MT from 1 to 3 equiv enhanced the allyl modification yield from 0.4% to around 8%. The MT complex exhibits a much lower k_{cat} and a higher K_{m} values from Michaelis–Menten kinetics studies in the presence of allyl-SAM than with normal SAM cofactor (see SI), which explains the low reaction yield for allyl-SAM. The allyl substitution may not fit the MT cofactor-binding pocket of the native methyltransferase. Future protein engineering will be performed in order to address this limit of the current system.

With the two above labeling strategies confirmed, we next studied the chemical treatment of a⁶A-modified RNA and explored its resultant effect on the RT process. We thought that the double bond of the allyl group at the N⁶-position of A (**4**) could be readily iodized under mild conditions and that the iodinated alkyl chain (**5**) could spontaneously attack the N¹-position of A in order to generate N¹,N⁶-cyclized form (**6**, **7**) (Scheme 2). Because the hydrogen bonding sites of A are occupied, mutations at opposite site can be readily induced during RT and also detected in cDNA sequencing.²⁶

In order to prove the proposed chemical treatment process as shown in Scheme 2, we characterized the consecutive iodination and cyclization steps of model nucleoside a⁶A by standard NMR and mass spectroscopy (see SI). The spectroscopic data indicated that the formation of five-membered N¹,N⁶-cyclized product was favored. To mimic mRNA, we synthesized a⁶A-containing RNA oligo²⁷ probe **8** with a GG(a⁶A)CU motif and converted it into cyclized form **9** using similar protocol to that of the a⁶A model nucleoside (see SI). The MALDI-TOF mass spectrometry result (Figure 2A) showed that probe **8** was converted to probe **9** with a molar mass increase of ~127, equivalent to an iodine atom mass, which clearly supports our proposed RNA post-treatment reaction mechanism.

Our next step was to investigate the RT reaction of these modified probes. By following NEB small RNA library protocol, we ligated probes **8** and **9** with a 3'-adapter, annealed the

5'-primer, and then ligated to the 5'-adaptor. The HIV enzyme was selected as the reverse transcriptase²⁶ because it could tolerate more structural variations and allowed for more read-through compared to normal enzymes that cause RT stop. After RT, the cDNAs were amplified and subjected to high-throughput sequencing. About 5.0 and 3.8 million reads were collected in order to calculate cDNA mutation rates of a⁶A and N¹,N⁶-cyclized A in probes **8** and **9**, respectively. Bioinformatic analysis revealed that a small portion (~5%) of a⁶A sites were mutated to T/C/G, whereas N¹,N⁶-cyclized A showed up to ~70% mutations to T/C/G (Figure 2B and see SI). The clear mutation contrast indicates that the combination of a⁶A labeling and mild postchemical treatment could represent a strategy for identifications of RNA methylations.

We further tested a biological sample in order to validate the MT-assisted a⁶A-labeling method. Following a typical protocol (see SI), the polyadenylated RNAs from HeLa cells were purified, labeled in the presence of METTL3/METTL14 and allyl-SAM cofactor, iodinated, cyclized, and reverse transcribed to cDNA using the HIV RT enzyme. A known GGACU site from mRNA ACTB was chosen as an example, whose m⁶A methylation fraction within HeLa cells has been determined to be ~20% (~80% A and ~20% m⁶A) using the SCARLET (site-specific cleavage and radioactive-labeling followed by ligation-assisted extraction and thin-layer chromatography) method.²⁸ The corresponding cDNAs, derived from labeled and unlabeled RNAs, were sequenced to compare their mutation patterns using standard TA cloning technique (see SI). The result displayed that a mutation frequency of 17% was observed for the selected ACTB GGACU site (see SI), whereas the nearby UGACU site did not reveal mutation, indicative of high sequence specificity. The low allyl transfer activity of the methyltransferase of the current system and the fact that not all N¹,N⁶-cyclized A generate mutations contributed to the low mutation frequency. This method did show its promise for differentiating m⁶A and A on cellular RNAs at potential base-resolution, however. Further efforts will focus on METTL3/METTL14 protein-engineering based on their crystal structure²⁹ in order to better accommodate allyl-SAM cofactor and increase RNA-labeling efficiency.

In summary, we found a⁶A as a new small molecule probe for RNA labeling through both metabolic and MT-assisted pathways. With monosubstitution of allyl group at N⁶-position of A, a⁶A in labeled RNA behaves like A, and can be chemically treated by iodine under mild buffer conditions in order to form N¹,N⁶-cyclized A, which induces mutations to other bases during RT to cDNA. The model reactions of single a⁶A nucleoside and a⁶A-containing RNA oligo fully support our proposed labeling-coupled, post-treatment-induced mutation mechanism.

The current a⁶A-labeling method will find potential applications in the RNA research field, such as nascent RNA sequencing and RNA decay dynamics. Meanwhile, the MT-assisted allyl-labeling method suggests a new approach to site-specifically label RNA with modifiable chemical group, which can be extended for other enzyme-assisted base modifications. We also postulate that the MT-assisted a⁶A labeling could be applied toward potentially differentiating m⁶A from A transcriptome-wide. On the basis of the principle that MT can install the allyl group on unmodified A within the consensus motif while the existing m⁶A in the motif is inert to allyl labeling, we may acquire information on m⁶A

stoichiometry at a specific mRNA modification site. We also postulate the introduction of a⁶A to generate mutations could be a concept quite useful in developing new sequencing technology to interrogate dynamics of transcriptomes in the future.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the National Key Research and Development Program of China (2017YFA0506800), National Natural Science Foundation of China (21642015), and National Institutes of Health of United States (HG008935 to C.H., R01GM096056 to M.L., and K01HG006699 to Q.D.). J.L. thanks the Thousand Young Talents Plan of China and Hundred Talents Program of Zhejiang University. C.H. is an investigator of the Howard Hughes Medical Institute. We thank S. Frank, MA for editing the paper.

References

1. Liu Y, Holmstrom E, Zhang J, Yu P, Wang J, Dyba MA, De C, Ying J, Lockett S, Nesbitt DJ, Ferré D'-Amaré AR, Sousa R, Stagno JR, Wang Y-X. *Nature*. 2015; 522:368. [PubMed: 25938715]
2. Lu K, Miyazaki Y, Summers MF. *J Biomol NMR*. 2010; 46:113. [PubMed: 19789981]
3. Rabani M, Levin JZ, Fan L, Adiconis X, Raychowdhury R, Garber M, Gnirke A, Nusbaum C, Hacohen N, Friedman N, Amit I, Regev A. *Nat Biotechnol*. 2011; 29:436. [PubMed: 21516085]
4. Cleary MD, Meiering CD, Jan E, Guymon R, Boothroyd JC. *Nat Biotechnol*. 2005; 23:232. [PubMed: 15685165]
5. Zheng Y, Beal PA. *Bioorg Med Chem Lett*. 2016; 26:1799. [PubMed: 26927424]
6. Curanovic D, Cohen M, Singh I, Slagle CE, Leslie CS, Jaffrey SR. *Nat Chem Biol*. 2013; 9:671. [PubMed: 23995769]
7. Grammel M, Hang H, Conrad NK. *ChemBioChem*. 2012; 13:1112. [PubMed: 22513998]
8. Schwalb B, Michel M, Zacher B, Frühauf K, Demel C, Tresch A, Gagneur J, Cramer P. *Science*. 2016; 352:1225. [PubMed: 27257258]
9. Roundtree IA, Evans ME, Pan T, He C. *Cell*. 2017; 169:1187. [PubMed: 28622506]
10. Harcourt EM, Kietrys AM, Kool ET. *Nature*. 2017; 541:339. [PubMed: 28102265]
11. Machnicka MA, Milanowska K, Osman Oglou O, Purta E, Kurkowska M, Olchowik A, Januszewski W, Kalinowski S, Dunin-Horkawicz S, Rother KM, Helm M, Bujnicki JM, Grosjean H. *Nucleic Acids Res*. 2013; 41:D262. [PubMed: 23118484]
12. Fu Y, Dominissini D, Rechavi G, He C. *Nat Rev Genet*. 2014; 15:293. [PubMed: 24662220]
13. Meyer KD, Jaffrey SR. *Nat Rev Mol Cell Biol*. 2014; 15:313. [PubMed: 24713629]
14. Gokhale NS, McIntyre ABR, McFadden MJ, Roder AE, Kennedy EM, Gandara JA, Hopcraft SE, Quicke KM, Vazquez C, Willer J, Ilkayeva OR, Law BA, Holley CL, Garcia-Blanco MA, Evans MJ, Suthar MS, Bradrick SS, Mason CE, Horner SM. *Cell Host Microbe*. 2016; 20:654. [PubMed: 27773535]
15. Schwartz S, Bernstein DA, Mumbach MR, Jovanovic M, Herbst RH, León-Ricardo BX, Engreitz JM, Guttman M, Satija R, Lander ES, Fink G, Regev A. *Cell*. 2014; 159:148. [PubMed: 25219674]
16. Carlile TM, Rojas-Duran MF, Zinshteyn B, Shin H, Bartoli KM, Gilbert WV. *Nature*. 2014; 515:143. [PubMed: 25192136]
17. Li X, Zhu P, Ma S, Song J, Bai J, Sun F, Yi C. *Nat Chem Biol*. 2015; 11:592. [PubMed: 26075521]
18. 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, Doré LC, Amariglio N, Rechavi G, He C. *Nature*. 2016; 530:441. [PubMed: 26863196]
19. Li X, Xiong X, Wang K, Wang L, Shu X, Ma S, Yi C. *Nat Chem Biol*. 2016; 12:311. [PubMed: 26863410]

20. Liu J, Yue Y, Han D, Wang X, Fu Y, Zhang L, Jia G, Yu M, Lu Z, Deng X, Dai Q, Chen W, He C. *Nat Chem Biol.* 2014; 10:93. [PubMed: 24316715]
21. Jia G, Fu Y, Zhao X, Dai Q, Zheng G, Yang Y, Yi C, Lindahl T, Pan T, Yang Y-G, He C. *Nat Chem Biol.* 2011; 7:885. [PubMed: 22002720]
22. Ottria R, Casati S, Baldoli E, Maier JAM, Ciuffreda P. *Bioorg Med Chem.* 2010; 18:8396. [PubMed: 21035348]
23. Duffy EE, Rutenberg-Schoenberg M, Stark CD, Kitchen RR, Gerstein MB, Simon MD. *Mol Cell.* 2015; 59:858. [PubMed: 26340425]
24. Nguyen K, Fazio M, Kubota M, Nainar S, Feng C, Li X, Atwood SX, Bredy TW, Spitale RC. *J Am Chem Soc.* 2017; 139:2148. [PubMed: 28139910]
25. Wang R, Zheng W, Yu H, Deng H, Luo M. *J Am Chem Soc.* 2011; 133:7648. [PubMed: 21539310]
26. Calabretta A, Leumann CJ. *Biochemistry.* 2013; 52:1990. [PubMed: 23425279]
27. Allerson CR, Chen SL, Verdine GL. *J Am Chem Soc.* 1997; 119:7423.
28. Liu N, Parisien M, Dai Q, Zheng G, He C, Pan T. *RNA.* 2013; 19:1848. [PubMed: 24141618]
29. Wang X, Feng J, Xue Y, Guan Z, Zhang D, Liu Z, Gong Z, Wang Q, Huang J, Tang C, Zou T, Yin P. *Nature.* 2016; 534:575. [PubMed: 27281194]

Figure 1.

(A) Quantification of total RNA a^6A/A ratio versus different a^6A labeling time in HeLa cell using UHPLC-QQQ-MS/MS. The a^6A -labeled RNAs were enriched and their a^6A levels were measured for comparison. (B) Test of the allyl transfer ability of mRNA m^6A methyltransferases METTL3/METTL14 heterodimer (MT) in the presence of allyl-SAM cofactor on different RNA probes (1–3) with/without m^6A and/or m^6A consensus motif. (C) Quantification of the allyl-transfer efficiency on probe 1 with different equivalents of MT using UHPLC-QQQ-MS/MS. The modification yield was calculated by generated a^6A per probe 1.

Figure 2.

(A) MALDI-TOF spectra of a⁶A-modified RNA oligo (probe **8**) and its iodination-induced N¹,N⁶-cyclized form (probe **9**). (B) Percentages of mutation of a⁶A and its cyclized form into other bases after reverse transcription using HIV enzyme and high-throughput cDNA sequencing.

Scheme 1. Illustration of RNA Labeling with N^6 -Allyladosine and Its Mutation Assay

Scheme 2.

Iodination of N^6 -Allyl-adenosine Leads to Formation of N^1,N^6 -Cyclized Adenosine, Which Induces Mutation during Reverse Transcription