Journal Pre-proof

A chemical labelling of N^6-formyl adenosine (f^6A) RNA

Li-jun Xie , Cui-Lian Lin , Li Liu , Liang Cheng

PII: S1001-8417(21)00740-3
DOI: https://doi.org/10.1016/j.cclet.2021.09.028
Reference: CCLET 6740

To appear in: *Chinese Chemical Letters*

Received date: 1 July 2021
Revised date: 3 September 2021
Accepted date: 7 September 2021

Please cite this article as: Li-jun Xie , Cui-Lian Lin , Li Liu , Liang Cheng , A chemical labelling of N^6-formyl adenosine (f^6A) RNA, *Chinese Chemical Letters* (2021), doi: https://doi.org/10.1016/j.cclet.2021.09.028

Communication

A chemical labelling of N^6-formyl adenosine (f^6A) RNA

Li-jun Xie[a], Cui-Lian Lin[a,b], Li Liu[a, b], Liang Cheng[a,b*]

[a]*Beijing National Laboratory for Molecular Sciences (BNLMS), CAS Key Laboratory of Molecular Recognition and Function, CAS Research/Education Center for Excellence in Molecular Sciences, Institute of Chemistry, Chinese Academy of Sciences, Beijing 100190, China*
[b] *University of Chinese Academy of Sciences, Beijing 100049, China*

ARTICLE INFO

Article history:
Received
Received in revised form
Accepted
Available online

Keywords:
RNA epigenetics
Chemical modulation
Labelling
Nucleic acids
Alkylation

ABSTRACT

N^6-methyl adenosine (m^6A) is an eminent epigenetic mark in mRNAs that affects a broad range of biological functions in diverse species. However, the chemically inert methyl group prevents a direct labelling of this modification for subsequent detection and sequencing. Therefore, most current approaches for the labelling of m^6A still have limitations of relying on the utilization of corresponding methyltransferases, which resulted in the lacking of efficiency. Here we present an approach which selectively alkylated the N^6-formyl adenosine (f^6A), the key intermediate during chemical oxidation of m^6A, with an alkyne functionality that can be further labelled with click reactions. This covalent labelling approach will be able to facilitate in the affinity purification, detection and genome-wide profiling studies.

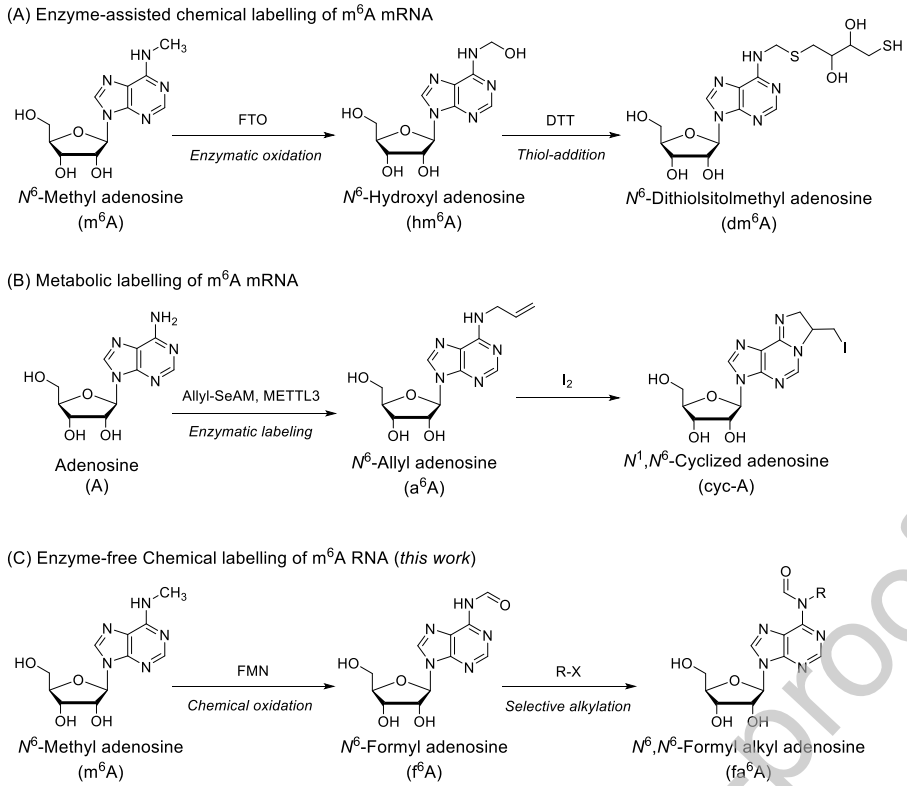
N^6-Methyl adenosine (m^6A) is a widely studied epigenetic mark that was discovered in the early 1970s in messenger RNAs (mRNAs) from eukaryotes [1-4]. The methylation process is catalyzed by a multicomponent methyltransferase complex, including METTL3, METTL14, WTAP and other "writers" [5, 6]. FTO and ALKBH5 are so-far two identified m^6A demethylases ("erasers") that can remove m^6A methylated groups from RNA, which makes the epigenetic modification a dynamic reversible process [7-9]. On the other hand, regulatory proteins ("readers") like YTHDF and YTHDC subtypes can bind to the m^6A modification site in RNA and initiate different downstream effects [10-12]. However, the comprehensive biological functions of m^6A modification are not fully understood at present [13, 14], largely because of the difficulties of identifying m^6A sites. The N^6-methyl substituent does not affect its reverse transcription during the PCR process. Therefore, traditional m^6A RNA fragments are usually captured and enriched by immunoprecipitation and then identified by second-generation sequencing [15-19]. However, these approaches are limited to the sources of antibodies or recognizing enzymes (reading proteins or restriction enzymes) and the specific sites in the transcriptome. Thus, there is a great need for a simple, sensitive, antibody-free method for m^6A detection.

Labeling of nucleic acids is required for many studies aiming to elucidate their functions and dynamics *in vitro* and in live cells. To date, two different strategies have been developed to label, profile and analysis genome-wide m^6A methylation patterns in live cells (Scheme 1). Jia *et al.* utilized m^6A demethylease FTO that are responsible to m^6A demethylation at the RR$\underline{m^6A}$CH sequence and converted the inert methyl substituent to chemically reactive hydroxymethyl group (hm^6A), which was sensitive to nucleophilic substitution with thiol compounds like dithiothreitol (DTT) to afford the N^6-dithiolsitolmethyl adenosine (dm^6A) [20]. Thus, bioorthgonally functional groups like alkyne/azide containing thiols can be exploited to collect those fragments and then subjected to library construction and deep-sequencing (Scheme 1A). Similarly, Liu *et al.* adapted the enzymatic methylation process by replacing the natural methyl donor SAM to Se-allyl-L-selenohomocysteine (Ally-SeAM) [21]. Under the promotion of the methyltransferase METTL3, the original m^6A sites will be replaced with N^6-allyl adenosines (a^6A). With the chemically functional alkenyl substituent in hand, they initiated the iodine-catalyzed intramolecular cyclization to generate the N^1,N^6-cyclized adenosine (cyc-A) followed by sequencing (Scheme 1B). In spite of above two methods, further technology development of a robust, efficient, unbiased approach for whole-genome methylation profiling of m^6A is still highly desirable. The development of such an approach without using antibodies or modifying enzymes will aid the general community in consistent profiling of RNA epigenetic modifications, and in developing disease-specific diagnoses as well as establishing biomarkers.

Here we propose a new approach inspired by the discovery that m^6A can be chemically oxidized by the flavin mononucleotide (FMN) promoted oxidation [22]. The inert C-H bonds at the N^6-methyl sites can be selective activated to generate hm^6A, just like FTO enzyme. On the other hand, hm^6A can be further converted to N^6-formyl adenosine (f^6A). We and others have invented chemical-labelling approaches to selectively label C5-formyl cytidines (f^5C/5fC) with functional groups, such as amines, for robust affinity enrichment, determination and sequencing [23-27]. We envisioned that such a chemical labelling strategy could be combined with FMN-mediated conversion of m^6A to f^6A for a selective labelling of m^6A for genome-wide detection and profiling (Scheme 1C). In our new approach, we took advantage of the electron-withdrawing propriety of the formyl group by employing the nucleophilic substitution. Utilizing Huisgen cycloaddition (click) chemistry [28], a possible

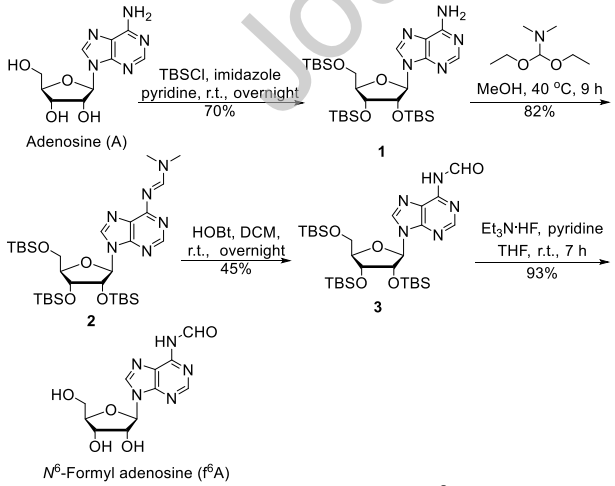
* Corresponding author.
E-mail address: chengl@iccas.ac.cn

tag (or any chemical tag) may be installed, thus facilitating an efficient and unbiased labelling of the original m^{6}A-containing RNA fragments for detection and genome-wide profiling. This new approach may provide a wider coverage of m^{6}A-containing genomic regions compared with other affinity-enrichment methods.



Scheme 1. Previously reported labelling of m^{6}A and our strategy.

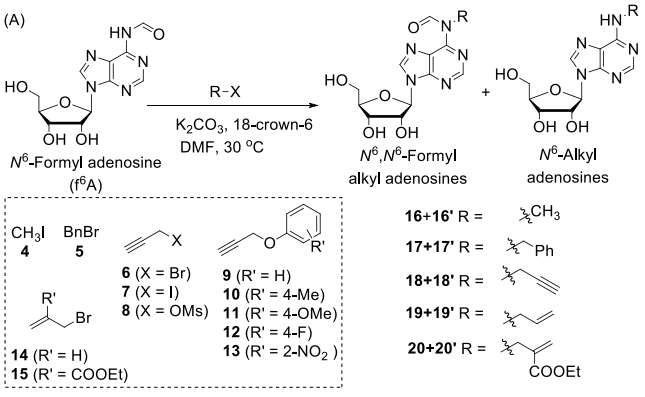
To investigate the reactivity of f^{6}A, we firstly prepared the adenosine derivative in a gram scale (Scheme 2) [8]. The hydroxyls at adenosine were protected with *tert*-butyl(dimethyl)silyl (TBS) group, and then the free amine in **1** was reacted with *N*,*N*-dimethyformamide diethy acetal to generate the dimethylformimidamide **2**. The latter was then subjected to hydrolysis with hydroxybenzotriazole to release the *N*-formyl group. We modified this step by using dichloromethane instead of methanol, which significantly improved the yield to 45%. Upon treatment with triethylamine trihydrofluoride, the unprotected f^{6}A was obtained in 23% yield (4 steps).



Scheme 2 Synthesis of f^{6}A.

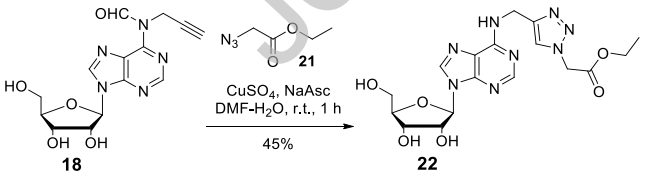
With the sufficient amount of f^{6}A in hand, we then investigated the chemical labelling of this central intermediate (Scheme 3A). We firstly intended to functionalize the carbonyl group with Wittig olefination, nucleophilic addition and other related reactions. However, due to the amide resonance with its enolate isomer, the carbonyl was stabilized in the *N*-formyl functionality and thus can not be labelled under extreme mild conditions (see Supporting information for details). We thus tuned to the *N*-formyl itself by adapting the *N*-alkylation of peptides/proteins. The *N*-H in f^{6}A was influenced by the heterocyclic adenine and the attached carbonyl group and should be acidic enough to be easily deprotonated. Thus by using mild potassium carbonate as the base with the crown ether 18-crown-6, we were able to screen a bunch of electrophiles, in most cases, alkyl halides **4-15**. With the simplest iodomethane **4**, the N^6-formyl-N^6-methyl adenosine **16** was obtained in a 30% yield with 3 h (Scheme 3B, Table S1 in Supporting information). However, no de-formylation product **16'** was observed, indicating that the alkylation indeed could selectively label f^{6}A while inhibiting its hydrolysis. A much more reactive benzyl bromide **5** would lead to a higher yield of the corresponding di-substituted adenosine in 40 min. Considering the bioorthgonal property of alkyne functionality, we next tested several alkynyl halides **6** and **7** and sulfonates **8-13**. While the propargyl chloride **6** afforded the N^6-formyl-N^6-propargyl adenosine **18** in a 33% yield, its bromide analogue **7** resulted in a much faster alkylation (within 15 min) along with a few de-formylation product **18'** (15%). We thus assume that the leaving group might play a crucial role in the efficiency of the propargylation and may influence the stability of the product **18** during the alkylation. Indeed, simple propargyl mesylate **8** and benzenesulfonate **9** generated uneven distribution of **18** and **18'**. Interestingly, the *p*-toluenesulfonate **10**, *p*-methoxyl benzenesulfonate **11** and *p*-fluorobenzenesulfonate **12**, all afforded the N^6-propargyl

adenosine **18'** in inferior yields (24%-28%), while the *o*-nitro benzenesulfonate **13** delivered moderate amount of **18** (31%). We also verified if allylic motif can be introduced to f^6A since the terminal alkene is also a bioorthgonal handler in biocongugation chemistry. Thus f^6A was subjected to react with allylic bromide **14**, affording the corresponding labelled product **19** in good yield (72%) without any noticeable de-formylation. Another interesting example was using Morita-Baylis-Hillman adduct **15** as the electrophile. In that case, the acrylate motif was selectively introduced to the N^6-position, which might afford another type of chemical enrichment by conjugate addition. Thus different labelling groups could be hosted by varying the electrophiles.



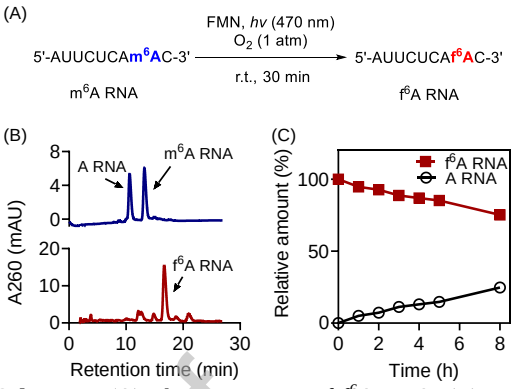
Scheme 3 Chemical labelling of f^6A.

As a demonstration of the utilization of this powerful reaction toward m^6A enrichment, we applied the 1,3-dipolar cycloaddition of the product **18** with ethyl azidoacetate **21** (Scheme 4). The click reaction smoothly afforded the triazole product **22** in good yield (45%). Interestingly, the formyl group was removed during the post-modification process. In a word, we have developed a wide range of labelling intermediates with f^6A *via* simple steps and rapidly creating new products for further applications.



Scheme 4 The click reaction of **18** with ethyl azidoacetate.

In order to explore the possibility of this approach at transcription RNAs, we prepared an m^6A containing oligo 5'-AUUCUCA$\underline{m^6A}$C-3' and subjected it to the standard chemical demethylation conditions (Scheme 5A). We have carefully optimized the amount of FMN and the oxidation time. Finally we were able to generate the f^6A oligo 5'-AUUCUCA$\underline{f^6A}$C-3' in 36% yield according to HPLC analysis (Scheme 5B). It was reported in literature that f^6A was unstable in aqueous solutions and rapidly hydrolyzed to adenosine. However, under our reaction conditions, f^6A oligo was stable in 8 h, with less than 30% of decomposition (Scheme 5C). With that in hand, we applied the aforementioned labelling to this oligo (see Supporting information for details). Unfortunately, the f^6A oligo remained untouched or decomposed completely under these conditions (pH 8.0 or pH 9.0), no matter what electrophile we have used. Thus, a more efficient method would be needed for RNA samples isolated from live cells. This optimization is currently underway in our laboratory and will be reported in due courses.



Scheme 5. (A) The preparation of f^6A RNA; (B) HPLC analysis of an ssRNA 5'-AUUCUCA$\underline{m^6A}$C-3' after treatment with FMN and irradiation with blue LED light of 470 nm under oxygen for 30 min (red line), compared to A RNA (5'-AUUCUCA$\underline{A}$C-3') and m^6A RNA (blue line); (C) Stability test of f^6A RNA (red line) at room temperature and the generation curve of A RNA (black line).

In summary, we have presented a practical approach for the selective labelling of f^6A, an essential intermediate during the oxidative demethylation of m^6A, with simple and easily available small organic molecules. The alkylation proceeded rapidly and selectively under mild conditions to covalently link a bunch of bioorthogonal components. With the successful establishment of click reaction pertaining to f^6A derivatives, it is expected to provide a useful strategy for chemically uniform and highly selective labelling of m^6A RNAs in an enzyme-free and covalent selective fashion.

Declaration of Interest Statement

We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

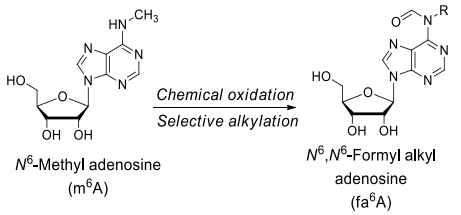
Acknowledgments

This work was supported by the National Key R&D Program of China (Nos. 2017YFA0208100 and 2020YFA0707901), National Natural Science Foundation of China (Nos. 22022704, 91853124, 21977097, and 21778057), Postdoctoral Innovative Talents Support Program (No. BX20200337, Dr. L.J. Xie) and Chinese Academy of Sciences.

References

[1] R. Desrosiers, K. Friderici, F. Rottman, Proc. Natl Acad. Sci. U. S. A. 71 (1974) 3971-3975.
[2] J.M. Adams, S. Cory, Nature 255 (1975) 28-33.
[3] W. Cha-Mer, G. Alan, M. Bernard, Cell 4 (1975) 379-386.
[4] Y. Furuichi, M. Morgan, A.J. Shatkin, et al., Proc. Natl Acad. Sci. U. S. A. 72 (1975) 1904-1908.
[5] H. Shi, J. Wei, C. He, Mol. Cell 74 (2019) 640-650.
[6] H. Huang, H. Weng, J. Chen, Trends Genet. 36 (2020) 44-52.
[7] G. Jia, Y. Fu, X. Zhao, et al., Nat. Chem. Biol. 7 (2011) 885-887.
[8] Y. Fu, G. Jia, X. Pang, et al., Nat. Commun. 4 (2013) 1798.
[9] G. Zheng, J.A. Dahl, Y. Niu, et al., Mol. Cell 49 (2013) 18-29.
[10] D.P. Patil, B.F. Pickering, S.R. Jaffrey, Trends Cell Biol. 28 (2018) 113-127.
[11] S. Berlivet, J. Scutenaire, J.M. Deragon, C. BBA-Gene Regul. Mech. 1862 (2019) 329-342.
[12] Y.L. Zhao, Y.H. Liu, R.F. Wu, et al., Mol. Biotech. 61 (2019) 355-364.
[13] W. Zhao, X. Qi, L. Liu, et al., Mol. Ther. Nucleic Acids 19 (2019) 405-412.
[14] I. Livneh, S. Moshitch-Moshkovitz, N. Amariglio, G. Rechavi, D. Dominissini, Nat. Rev. Neurosci. 21 (2020) 36-51.
[15] D. Dominissini, S. Moshitch-Moshkovitz, S. Schwartz, et al., Nature 485 (2012) 201-284.
[16] K.D. Meyer, Y. Saletore, P. Zumbo, et al., Cell 149 (2012) 1635-1646.
[17] M.A. Garcia-Campos, S. Edelheit, U. Toth, et al., Cell 178 (2019) 731-747.
[18] K.D. Meyer, Nature Methods 16 (2019) 1275-1280.
[19] Z. Zhang, L.Q. Chen, Y.L. Zhao, et al., Sci. Adv. 5 (2019) eaax0250.
[20] Y. Wang, Y. Xiao, S. Dong, Q. Yu, G. Jia, Nat. Chem. Biol. 16 (2020) 896–903.
[21] X. Shu, J. Cao, M. Cheng, et al., Nat. Chem. Biol. 16 (2020) 887–895.
[22] L.J. Xie, X.T. Yang, R.L. Wang, et al., Angew. Chem. Int. Ed. 58 (2019) 5028-5032.
[23] B. Samanta, J. Seikowski, C. Hobartner, Angew. Chem. Int. Ed. 55 (2016) 1912-1916.
[24] M.D. Lan, B.F. Yuan, Y.Q. Feng, Chin. Chem. Lett. 30 (2019) 1-6.
[25] C. Qi, J. Ding, B. Yuan, Y. Feng, Chin. Chem. Lett. 30 (2019) 1618-1626.
[26] R.L. Wang, X.Y. Jin, D.L. Kong, et al., Adv. Synth. Catal. 361 (2019) 5406-5411.
[27] Y. Wang, X. Zhang, G. Zou, et al., Acc. Chem. Res. 52 (2019) 1016-1024.
[28] M. Meldal, C.W. Tornøe, Chem. Rev. 108 (2008) 2952-3015.

Graphical Abstract



We present here an approach which selectively alkylated the N^6-formyl adenosine f^6A, the key intermediate during chemical oxidation of N^6-methyl adenosine m^6A, with an alkyne functionality that can be further labelled with click reactions.