Title

Nucleoside analogs in the study of the epitranscriptome

Authors

Palumbo, Cody M
Beal, Peter A

Publication Date

2019-03-01

DOI

10.1016/j.ymeth.2018.10.014

Peer reviewed

HHS Public Access

Author manuscript

Methods. Author manuscript; available in PMC 2020 March 01.

Published in final edited form as:

Methods. 2019 March 01; 156: 46–52. doi:10.1016/j.ymeth.2018.10.014.

Nucleoside Analogs in the Study of the Epitranscriptome

Cody M. Palumbo and **Peter A. Beal**[1,*]

[1]Department of Chemistry, University of California, One Shields Ave, Davis, CA 95616, USA

Abstract

Over 150 unique RNA modifications are now known including several nonstandard nucleotides present in the body of messenger RNAs. These modifications can alter a transcript's function and are collectively referred to as the epitrancriptome. Chemically modified nucleoside analogs are poised to play an important role in the study of these epitranscriptomic marks. Introduced chemical features on nucleic acid strands provide unique structures or reactivity that can be used for downstream detection or quantification. Three methods are commonly used in the field to synthesize oligonucleotides containing chemically modified nucleoside analogs. Nucleoside analogs can be introduced via phosphoramidite-based chemical synthesis, via polymerases with modified nucleotide triphosphates or by metabolic labeling. In this review, these methods for incorporation of nucleoside analogs will be discussed with specific recently published examples pertaining to the study of the epitranscriptome.

Keywords

ADAR; Mettl3; epitranscriptome; nucleoside analog; m^6A; inosine

1. Introduction

Modifications to the four common nucleosides present in RNA have been known for over five decades [1–4]. However, modern analytical techniques, such as mass spectrometry and next generation sequencing (NGS), have caused a renaissance in their study. Naturally occurring RNA modifications are introduced into the polynucleotide chain by RNA modifying enzymes and are ubiquitous. To date, over 150 unique and structurally distinct RNA modifications have been discovered and characterized [5]. The vast majority of these modifications are found in tRNA and rRNA, but several nonstandard nucleotides including inosine (I) [6, 7], N^6-methyladenosine (m^6A) [8], N^6-2'-O-dimethyladenosine (m^6Am) [9] and pseudouridine (Ψ) [10], have been located in mRNA as well (Figure 1). These modifications have been referred to as the epitrancriptome to highlight their modulation of transcript function in analogy to the ways epigenetic modifications can alter genome function [11]. Epitranscriptomic modifications can cause changes in the fate or function of

*Corresponding author, tel: (530) 752-4132. pabeal@ucdavis.edu.

RNA by altering base-pairing, pi-stacking, or metal-chelation [12] (among others). In this review, we will highlight strategic uses of chemically synthesized nucleoside analogs and analog-containing nucleic acids in the study of the epitranscriptome. Nucleoside analogs have been developed that facilitate the detection of specific modifications in the transcriptome, the quantification of different types of modifications and for the study of the structure of modified RNA or RNA modifying enzyme/substrate complexes.

Nucleoside analog-containing nucleic acids can provide insight on mechanism or binding specificity which would otherwise be hard to elucidate by traditional biochemical techniques. Likewise, these analogs can be used to directly show the importance of functional group contacts at a nucleic acid-protein interface. Most commonly, nucleoside analogs are incorporated into oligonucleotides by synthesizing the corresponding phosphoramidite monomer and using it in automated chemical synthesis. Alternatively, triphosphates bearing the desired chemically modified analog can be incorporated via polymerases [13–15]. In addition to these methods, metabolic labeling can be used to prepare nucleoside analog-containing nucleic acids [16, 17]. In this method, the nucleoside analog or its precursor is taken up by metabolically active cells and incorporated into newly synthesized strands.

2. Metabolic Labeling Method

Metabolic labeling has been broadly applied to label different types of biomolecules (carbohydrates, proteins, lipids..etc) [18], but this review will focus on metabolically labeling RNA. Unnatural metabolites can enter cellular metabolic pathways where they function much like their native counterparts. Downstream detection is made possible because of unique chemical features of the unnatural metabolite. A well-established method for identifying RNA modifications is mass spectrometry. For example, supplying cellular cultures with isotopically labeled glucose results in isotopically labeled mRNA, which can be detected by mass spectrometry [19]. Furthermore, several nucleoside analogs have been developed for the purpose of labeling nascent RNA [16, 17].

A major requirement of this method is that the metabolite bearing the chemical modification must be cell permeable and non-toxic to the cells. Metabolic labeling provides a unique opportunity to probe cellular processes *in vivo* that would otherwise be difficult to study.

2.1 SAM analog generated by metabolic labeling to identify m^{6}A sites

S-Adenosylmethionine (SAM) is the primary source of methyl groups in nearly all methylation reactions within the cell [20]. The most common internal modification of eukaryotic mRNA is N^6-methyladenosine (m^{6}A) and is of particular interest because of its role in a vast array of biological processes but mostly, due to its impact on mRNA stability and abundance of mRNA transcripts [21]. Adenosine methylation is thought to be a dynamic process that is controlled by three classes of proteins, “writers”, “readers” and “erasers”. The modification is introduced by a “writer” complex comprised of a SAM-dependent heterodimer of methyltransferases, Mettl3 and Mettl14 (Mettl3–14) and prefers the consensus sequence “DR(m^{6}A)CH” (D=A or G or U; R=G or A; H= A,C or U) [22]. While early studies showed both Mettl3 and Mettl14 as catalytically active, more recently it has

been shown that Mettl3 is solely responsible for the methyltransferase activity [23]. The Wilms tumor-1 associated protein (WTAP), an RNA binding protein, recruits the Mettl3–14 complex and aides in its localization to nuclear speckles [24–26]. A different SAM-dependent methyltransferase, Mettl16, has recently been shown to have a distinct substrate specificity compared to the Mettl3–14 complex [27–29]. The most studied m^6A binding proteins ("readers") are the YTH protein family which have been shown to play roles in nearly every mRNA function including splicing, stability, translation and silencing [30–32]. "Erasers" are responsible for the removal of the m^6A methylation and include two proteins in the AlkB family. Fat mass and obesity-associated protein (FTO) and ALKBH5 are known to demethylate m^6A both *in vitro* and *in vivo* [33, 34]. Most reported methods for mapping m^6A modifications require immunoprecipitation and RNA-sequencing because m^6A is indistinguishable from A using readily available polymerases. While these methods allow whole transcriptome coverage, a lack of nucleotide resolution, scalability and antibody bias are major limitations [8, 35–38].

To overcome some of these limitations, new chemically modified nucleic acids were created for the purpose of chemically labelling m^6A to improve direct detection [39, 40]. Shu et. al synthesized a modified SAM cofactor that replaced the methyl group with an allyl group [39] (Figure 2A). Isolated polyadenylated RNAs from HeLa cells could be incubated in the presence of Mettl3–14 and allyl-SAM to transfer the allyl group to specific adenosines. The allyl group can be treated with iodine (I_2) to promote a cyclization reaction that blocks the Watson-Crick face of the nucleobase (Figure 2A). This causes reverse transcriptases (RTs) to misread the modification site during extension (unlike m^6A, which is read as A by RTs). The resulting cDNA is then sequenced and Mettl3–14 sites can be identified as A to T/C/G mutations, which can further be quantified. Allyl labeling was used to confirm the location of a known m^6A site in total RNA, but at a lower level than when using the previously reported SCARLET method [8]. The difference in the two methods was suggested to be low allyl transfer or incomplete cyclization of the allyl group when treated with iodine.

Building on the allyl modification idea, Hartstock et. al developed a method for directly detecting m^6A sites by replacing the SAM methyl group with a propargyl group [40] (Figure 2B). The propargyl group is of particular interest because of its ability to react in the highly efficient and bio-orthogonal copper-catalyzed azide/alkyne cycloaddition (CuAAC) reaction. The authors first demonstrated that Mettl3–14 could efficiently transfer a propargyl group from a synthetic selenium-containing SAM analog. The selenium-based SAM analog has two main advantages over sulfur-containing analogs: greater stability and better activation [41]. The technique was further expanded by feeding the precursory propargyl-L-selenohomocysteine amino acid to HeLa cells. The metabolic precursor is incorporated into a SAM analog and the propargyl is donated by methyltransferases to cellular RNA. The alkyne RNA can react with an azide-biotin and be pulled down with streptavidin beads. The resulting RNA can be sequenced through next-generation sequencing (NGS) and methylation sites can be determined by termination during reverse transcription (RT) (Figure 2B). The authors note three main advantages of using propargyl over the allyl: 1) CuAAC is bio-orthogonal and highly specific towards azides, 2) allows for non-natural modifications to be introduced in vivo and 3) the propargyl was more efficiently introduced, likely due to the

selenium-based SAM analog. While this method was used to validate m^6A rRNA sites, future studies are likely to apply this method to study the m^6A in mRNA.

3. Non-standard nucleotide triphosphate incorporation

Amplification of RNA and DNA *in vitro* by polymerases has transformed the landscape of biotechnology and enabled the development of many widely used techniques such as routine sequencing, library generation and new ways of detecting infectious diseases [42–45]. Nucleotide triphosphates (NTPs) are the building blocks in the synthesis of oligonucleotides using polymerases. Fortunately, NTPs are chemically very manipulatable and can be functionalized at a variety of locations. This allows creation of a diverse library of NTPs that are unique and can be used for downstream applications like labeling RNAs for pull-down assays [17]. Recognition of non-standard NTPs by DNA polymerase has been heavily researched with the goal being determining which NTPs are suitable for PCR and chemical biology applications [46–49]. The general consensus is non-standard NTPs are less efficiently incorporated than the canonical bases, however, some modifications are more tolerated than others. It has now been established that atom substitutions and major groove modifications are more efficiently incorporated compared to minor groove modifications. For adding linkers, like an alkyne functional group, the C7 and C5 are best suited for purines and pyrimidines, respectively [50]. Evolution of polymerases that have higher fidelity with non-standard NTPs is a promising strategy for increasing yields of long oligonucleotides bearing multiple non-standard bases[51]. *In vitro* incorporation of non-standard NTP incorporation into DNA and RNA provides a low-cost but highly informative biochemical technique that can be widely adopted by any lab.

3.1 Selenium-containing triphosphate to probe m^6A sites

The study of the epitranscriptome frequently involves variants of next generation sequencing (NGS) approaches. For the detection of RNA modifications, samples for NGS are often enriched prior to sequencing using modification-specific antibodies or chemically modified to alter the structure of the modification. Strategies that involve immunoprecipitation for enrichment can suffer from antibody bias and do not provide nucleotide resolution in modification site determination[52, 53]. For modifications that are not distinguishable from the native nucleoside by RT, chemical conversion has been used to generate a species that causes reverse transcriptase to stop or to introduce mutations in the cDNA, allowing the precise location of the modification to be determined from the sequence reads [54]. However, the pre-treatment chemistry can involve harsh reagents and/or incomplete conversion. Chemical conversion or immunoprecipitation have been used to identify the presence of RNA modifications such as 1-methyl adenosine [55], m^6A[53], I [56], and Ψ [10, 57], however their exact levels are highly debated [37]. The m^6A modification is typically reverse transcribed like A, so detection requires a chemical modification to be installed, like the p^6A, or immunoprecipitation for enrichment.

Recently, the Zhou group synthesized a C4 selenium-containing thymidine 2'-deoxytriphosphate (4SeT) to allow for the direct detection of m^6A in RNA by RT [58] (Figure 3). Their data shows that 4SeT, when it is subjected to RT, is able to base pair with

normal adenosines, but the strand will be truncated when the 4SeT is base paired across m^6A. The atomic radius of selenium is thought to disrupt proper base-paring between 4SeT and m^6A. As a result of the cDNA truncation during RT, there is a detectable decrease in coverage at the position of m^6A. In order to prepare RNA for NGS, an FTO-assisted sequencing technique was developed. As mentioned in the previous section, FTO is capable of demethylating m^6A. The m^6A-containing RNA can be treated with or without FTO. FTO treated samples act as controls so a reduction in coverage can be detected (Figure 3). A change in coverage would then be indicative of the presence of an m^6A modification.

While this method was capable of probing m^6A sites in a small RNA construct, secondary structure of longer mRNA may interfere with the demethylation reaction and m^6A sites might not be detected. Also, the seven-step synthesis of 4SeT may hinder this technique from being widely adopted by labs who lack equipment and expertise to carry out the synthesis. Despite these downsides, this method could provide single nucleotide resolution probing of m^6A sites and not require immunoprecipitation, potentially avoiding antibody bias.

4. Solid phase oligonucleotide synthesis to incorporate nucleoside analogs

Starting in the 1950's, Merrifield began developing solid phase synthesis for polypeptides [59]. By anchoring the first amino acid to an insoluble solid support, the protected amino acid monomers could be added one at a time for an iterative, chemical approach to peptide synthesis. Thirty years later, an efficient solid phase, automated method was developed to generate oligonucleotides by utilizing phosphoramidite monomers as building blocks [60]. Now the phosphoramidite-based solid phase synthesis strategy is one of the most commonly used for the generation of RNA strands bearing nucleoside analogs. Because of the coupling efficiencies of a typical RNA solid phase synthesizer, strand length is restricted to less than ~ 120 nucleotides and is usually much shorter, in the ~ 50 nucleotide range. While oligonucleotide length can be a shortcoming, the phosphoramidite method is more compatible with heavily functionalized nucleoside analogs compared to the other methods discussed above since recognition by metabolic enzymes or polymerases is not required. This technique does rely on having a synthetic chemistry skill-set. However, many chemically modified phosphoramidite monomers are currently commercially available.

4.1 An adenosine deamination transition state analog to trap ADAR-RNA complexes

Recently, our lab applied the phosphoramidite method to generate nucleoside analog-containing RNAs useful for the study of adenosine deaminases that act on RNA (ADARs), the enzymes responsible for the incorporation of inosine in mRNA [61–63]. Since conversion of adenosine to inosine changes the base pairing properties of the nucleobase, this reaction is a form of RNA editing [64]. Due to the change in Watson-Crick pairing, this conversion can have profound downstream consequences such as codon changes, alternative splicing and alteration of protein-RNA interactions [65–67]. In humans, two active RNA editing deaminases exist, (ADAR1 and ADAR2) and one that is catalytically inactive (ADAR3) [68]. It is known that ADARs play essential roles in proper cellular functions and

aberrant RNA editing is linked to disease [69]. Indeed, mutations in the *ADAR1* gene are known to be a cause of Acardi Goutieres Syndrome [70]. A variety of methods exist for transcriptome-wide sequencing of A-to-I editing events [6, 7, 56]. However, these methods are insufficient for determining ADAR's biochemical preferences for binding and catalytic activity.

To address these knowledge gaps, high resolution structures of ADAR bound to RNA were needed. These were recently provided in the form of crystal structures of the human ADAR2 deaminase domain bound to RNA [71]. In order to solve such structures, two advances were required. The first was a method to trap ADAR bound to RNA in a complex relevant to the editing reaction. For this purpose, the purine analog 8-azanebularine (8-azaN) was used [72]. The covalent hydrate of 8-azaN is an excellent mimic of the high energy tetrahedral intermediate expected for the adenosine deamination reaction (Figure 4A, 4B). The other breakthrough was the identification of RNA sequences that were efficiently edited by an ADAR catalytic domain in vitro and formed a tight and specific binding complex with the protein when modified with 8-azaN [73]. The combination of 8-azaN and highly efficient ADAR2 substrate sequences enabled structural characterization of ADAR2-RNA complexes (Figure 4C).

Generation of 8-azaN-modifed RNA requires synthesis of an 8-azaN phosphoramidite monomer [74]. The synthesis of this compound starts with a Vorbruggen glycosylation reaction between the modified base 8-azaadenine and a tetraacetate protected ribofuranose. Subsequent steps remove the C6 exocyclic amine and deprotection of the sugar affording 8-azanebularine free nucleoside. The next two synthetic steps are to install the protecting groups required for automated synthesis, 5'-O-DMTr and 2'-O-TBDMS. Lastly, the nucleoside analog is subjected to the phosphoramidite reaction. This reaction must be anhydrous and anaerobic because oxidation of the phosphoramidite nucleoside will result in a poor coupling efficiency in the solid phase synthesizer. The purified phosphoramidite monomer bearing the nucleoside modification is analyzed by ^{31}P NMR to assess the oxidation level. The phosphoramidite monomer is then incorporated into a specific location within the RNA via the solid phase synthesis method. The RNA bearing the nucleoside analog is purified away from failure sequences by denaturing urea poly-acrylamide gel electrophoresis and confirmed by mass spectroscopy. The desired RNA is hybridized with its complementary strand to form the duplex recognized by ADAR. The ADAR2 deaminase domain is overexpressed in *Saccharomyces cerevisiae* and purified. *S. cerevisiae* is a preferred expression system for producing large amounts of ADARs, primarily because *E. coli* does not produce inositol hexakisphosphate (IP_6), a co-factor essential for proper folding and activity of ADAR [63]. To obtain protein of suitable quality for crystallography, it must be purified over Ni-NTA, heparin and finally size-exclusion chromatography. Lastly, crystal trays can be set up with crystallography quality ADAR2 deaminase and the RNA duplex containing 8-azanebularine.

The ADAR2-RNA structural studies enabled by this method helped explain editing site selectivity and revealed previously unknown contacts made between protein and RNA [71]. In each of the four crystal structures obtained of ADAR2-bound to dsRNA, the 8-azaN hydrate was observed in the active site of the protein. While some of the important residues

were previously identified by sequence alignment of ADARs and subsequent alanine scan [75], the new structures provided atomic-resolution detail for their roles and revealed previously uncharacterized protein loops important for RNA recognition.

ADARs have preferences for specific nucleotides neighboring editing sites [76], however the reasons these preferences were unknown. ADAR2-RNA crystal structures revealed an interaction between the 2-amino group of the 3' nearest-neighbor G and the carbonyl backbone of S486. Because this was not an amino acid side chain interaction to the RNA, a point mutation in ADAR2 would not be expected to have a substantial effect. Instead, to further validate the observations in the crystal structure, RNA substrates containing the guanosine analogs inosine (I), N^2-methylguanosine (N^2MeG) and 2-aminopurine (2AP) adjacent to the editing site were prepared by solid phase chemical synthesis, allowing for an investigation of the importance of the guanosine 2-amino group. Deamination reactions between these model substrates and ADAR2 were then performed. Together, the observed results highlighted the importance of this contact and its likely role in determining ADAR2's 3' nearest neighbor preference. Thus, nucleoside analog-containing RNA proved to be critical both in the formation of a stable complex for crystallization (e.g. 8-azaN RNA) and in testing the importance of specific contacts observed in the ADAR2-RNA crystal structures.

5. Concluding Remarks

In this review, we have highlighted recent examples of the use of nucleoside analog containing nucleic acids in the study of the epitranscriptome. Nucleoside analogs are introduced into nucleic acid strands by three commonly used methods; metabolic labeling, via polymerases and modified NTPs or by phosphoramidite-based solid phase synthesis. These methods allow for the introduction of unique chemical features into the strand imparted by the structure of the nucleoside analog. These features can then enable novel approaches to the study of the epitranscriptome such as new detection methods for m^6A or mechanism-based trapping of an inosine-generating ADAR enzyme.

Acknowledgements

The authors acknowledge funding from the US National Institutes of Health (NIH) grant R01GM061115 (P.A.B). The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of NIH.

References

1. Davis FF and Allen FW, Ribonucleic acids from yeast which contain a fifth nucleotide. J. Biol. Chem, 1957 227(2): p. 907–915. [PubMed: 13463012]
2. Holley RW, et al., Nucleotide sequences in the yeast alanine transfer ribonucleic acid. J. Biol. Chem, 1965 240: p. 2122–2128. [PubMed: 14299636]
3. Desrosiers R, Friderici K, and Rottman F, Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells. Proc. Natl. Acad. Sci. USA, 1974 71(10): p. 3971–5. [PubMed: 4372599]
4. Amos H and Korn M, 5-Methyl cytosine in the RNA of Escherichia coli. Biochimica et Biophysica Acta, 1958 29(2): p. 444–445. [PubMed: 13572373]

5. Boccaletto P, et al., MODOMICS: a database of RNA modification pathways. 2017 update. Nucleic Acids Res, 2018 46(D1): p. D303–D307. [PubMed: 29106616]
6. Cattenoz PB, et al., Transcriptome-wide identification of A > I RNA editing sites by inosine specific cleavage. RNA, 2013 19(2): p. 257–270. [PubMed: 23264566]
7. Kiran A and Baranov PV, DARNED: a DAtabase of RNa EDiting in humans. Bioinform, 2010 26(14): p. 1772–1776.
8. Liu N, et al., Probing N(6)-methyladenosine RNA modification status at single nucleotide resolution in mRNA and long noncoding RNA. RNA, 2013 19(12): p. 1848–1856. [PubMed: 24141618]
9. Mauer J, et al., Reversible methylation of m6Am in the 5′ cap controls mRNA stability. Nature, 2016 541: p. 371. [PubMed: 28002401]
10. Carlile TM, et al., Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. Nature, 2014 515: p. 143. [PubMed: 25192136]
11. Meyer KD, et al., Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. Cell, 2012 149(7): p. 1635–46. [PubMed: 22608085]
12. Chawla M, et al., An atlas of RNA base pairs involving modified nucleobases with optimal geometries and accurate energies. Nucleic Acids Res, 2015 43(14): p. 6714–6729. [PubMed: 26117545]
13. Eremeeva E, et al., Base-Modified Nucleic Acids as a Powerful Tool for Synthetic Biology and Biotechnology. Chem. – A European J, 2017 23(40): p. 9560–9576.
14. Bergen K, et al., Structures of KlenTaq DNA Polymerase Caught While Incorporating C5Modified Pyrimidine and C7-Modified 7-Deazapurine Nucleoside Triphosphates. J. Am. Chem. Soc, 2012 134(29): p. 11840–11843. [PubMed: 22475415]
15. Carroll SS, et al., Inhibition of Hepatitis C Virus RNA Replication by 2′-Modified Nucleoside Analogs. J. Biol. Chem, 2003 278(14): p. 11979–11984. [PubMed: 12554735]
16. Curanovic D, et al., Global profiling of stimulus-induced polyadenylation in cells using a poly(A) trap. Nat. Chem. Biol, 2013 9(11): p. 671–673. [PubMed: 23995769]
17. Zheng Y and Beal PA, Synthesis and evaluation of an alkyne-modified ATP analog for enzymatic incorporation into RNA. Bioorg. Med. Chem. Lett, 2016 26(7): p. 1799–1802. [PubMed: 26927424]
18. Siegrist MS, et al., Illumination of growth, division and secretion by metabolic labeling of the bacterial cell surface. FEMS Microbiol. Rev, 2015 39(2): p. 184–202. [PubMed: 25725012]
19. Kellner S, et al., Absolute and relative quantification of RNA modifications via biosynthetic isotopomers. Nucleic Acids Res, 2014 42(18): p. e142–e142. [PubMed: 25129236]
20. Shima H, et al., S-Adenosylmethionine Synthesis Is Regulated by Selective N6-Adenosine Methylation and mRNA Degradation Involving METTL16 and YTHDC1. Cell Rep., 2017 21(12): p. 3354–3363. [PubMed: 29262316]
21. Lin S, et al., METTL3 promotes translation in human cancer cells. Mol. Cell, 2016 62(3): p. 335–345. [PubMed: 27117702]
22. Liu J, et al., A METTL3–METTL14 complex mediates mammalian nuclear RNA N6adenosine methylation. Nat. Chem. Biol, 2013 10: p. 93. [PubMed: 24316715]
23. Śledź P and Jinek M, Structural insights into the molecular mechanism of the m(6)A writer complex. eLife, 2016 5: p. e18434. [PubMed: 27627798]
24. Ping X-L, et al., Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. Cell Res, 2014 24(2): p. 177–189. [PubMed: 24407421]
25. Jia G, et al., N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. Nat. Chem. Biol, 2011 7(12): p. 885–7. [PubMed: 22002720]
26. Little NA, Hastie ND, and Davies RC, Identification of WTAP, a novel Wilms' tumour 1-associating protein. Hum. Mol. Genet, 2000 9(15): p. 2231–9. [PubMed: 11001926]
27. Pendleton KE, et al., The U6 snRNA m(6)A methyltransferase METTL16 regulates SAM synthetase intron retention. Cell, 2017 169(5): p. 824–835.e14. [PubMed: 28525753]
28. Warda AS, et al., Human METTL16 is a N6-methyladenosine (m6A) methyltransferase that targets pre-mRNAs and various non-coding RNAs. EMBO Rep., 2017 18(11): p. 2004–2014. [PubMed: 29051200]

29. Ruszkowska A, et al., Structural insights into the RNA methyltransferase domain of METTL16. Sci. Rep, 2018 8(1): p. 5311. [PubMed: 29593291]
30. Meyer KD and Jaffrey SR, Rethinking m(6)A Readers, Writers, and Erasers. Ann. Rev. Cell Dev. Biol, 2017 33: p. 319–342. [PubMed: 28759256]
31. Wang X, et al., m(6)A-dependent regulation of messenger RNA stability. Nature, 2014 505(7481): p. 117–120. [PubMed: 24284625]
32. Lence T, Soller M, and Roignant J-Y, A fly view on the roles and mechanisms of the m6A mRNA modification and its players. RNA Biol., 2017 14(9): p. 1232–1240. [PubMed: 28353398]
33. Zheng G, et al., ALKBH5 Is a Mammalian RNA Demethylase that Impacts RNA Metabolism and Mouse Fertility. Mol. Cell, 2013 49(1): p. 18–29. [PubMed: 23177736]
34. Jia G, et al., N(6)-Methyladenosine in Nuclear RNA is a Major Substrate of the ObesityAssociated FTO. Nat. Chem. Biol, 2011 7(12): p. 885–887. [PubMed: 22002720]
35. Wang T, et al., Design and bioinformatics analysis of genome-wide CLIP experiments. Nucleic Acids Res, 2015 43(11): p. 5263–5274. [PubMed: 25958398]
36. Wheeler Emily C, Van Nostrand Eric L, and Yeo Gene W, Advances and challenges in the detection of transcriptome-wide protein–RNA interactions. Wiley Interdiscip. Rev.: RNA, 2017 9(1): p. e1436.
37. Legrand C, et al., Statistically robust methylation calling for whole-transcriptome bisulfite sequencing reveals distinct methylation patterns for mouse RNAs. Genome Res., 2017 27(9): p. 1589–1596. [PubMed: 28684555]
38. Mauer J and Jaffrey SR, FTO, m6Am, and the hypothesis of reversible epitranscriptomic mRNA modifications. FEBS Lett., 2018 592(12): p. 2012–2022. [PubMed: 29754392]
39. Shu X, et al., N6-Allyladenosine: A New Small Molecule for RNA Labeling Identified by Mutation Assay. J. Am. Chem. Soc, 2017 139(48): p. 17213–17216. [PubMed: 29116772]
40. Hartstock K, et al., Enzymatic or In Vivo Installation of Propargyl Groups in Combination with Click Chemistry for the Enrichment and Detection of Methyltransferase Target Sites in RNA. Angew. Chem. Int. Ed, 2018 57(21): p. 6342–6346.
41. Willnow S, et al., A selenium-based click AdoMet analogue for versatile substrate labeling with wild-type protein methyltransferases. Chembiochem, 2012 13(8): p. 1167–73. [PubMed: 22549896]
42. Lanciotti RS, et al., Rapid detection and typing of dengue viruses from clinical samples by using reverse transcriptase-polymerase chain reaction. J Clin Microbiol, 1992 30(3): p. 545–551. [PubMed: 1372617]
43. Sanger F, Nicklen S, and Coulson AR, DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA, 1977 74(12): p. 5463. [PubMed: 271968]
44. Shendure J and Ji H, Next-generation DNA sequencing. Nat. Biotech, 2008 26: p. 1135.
45. Byron SA, et al., Translating RNA sequencing into clinical diagnostics: opportunities and challenges. Nat. Rev. Genet, 2016 17: p. 257. [PubMed: 26996076]
46. Morales JC and Kool ET, Minor Groove Interactions between Polymerase and DNA: More Essential to Replication than Watson–Crick Hydrogen Bonds? J. Am. Chem. Soc, 1999 121(10): p. 2323–2324. [PubMed: 20852718]
47. Laos R, Thomson JM, and Benner SA, DNA polymerases engineered by directed evolution to incorporate non-standard nucleotides. Front Microbiol., 2014 5: p. 565. [PubMed: 25400626]
48. Hirao I and Kimoto M, Unnatural base pair systems toward the expansion of the genetic alphabet in the central dogma. Proc. Jpn. Acad. Ser. B. Phys. Biol. Sci, 2012 88(7): p. 345–367.
49. Chen T and Romesberg FE, Directed Polymerase Evolution. FEBS Lett., 2014 588(2): p. 19–229.
50. Seo YJ, et al., Site-Specific Labeling of DNA and RNA Using an Efficiently Replicated and Transcribed Class of Unnatural Base Pairs. J. Am. Chem. Soc, 2011 133(49): p. 1987819888.
51. Chim N, et al., Structural basis for TNA synthesis by an engineered TNA polymerase. Nat. Commun, 2017 8(1): p. 1810. [PubMed: 29180809]
52. Frye M, et al., RNA modifications: what have we learned and where are we headed? Nature Rev. Genet, 2016 17: p. 365. [PubMed: 27140282]

53. Dominissini D, et al., Transcriptome-wide mapping of N6-methyladenosine by m6A-seq based on immunocapturing and massively parallel sequencing. Nat. Prot, 2013 8: p. 176.
54. Suzuki T, et al., Transcriptome-wide identification of adenosine-to-inosine editing using the ICE-seq method. Nat. Prot, 2015 10: p. 715.
55. Dominissini D, et al., The dynamic N1-methyladenosine methylome in eukaryotic messenger RNA. Nature, 2016 530: p. 441. [PubMed: 26863196]
56. Kim DDY, et al., Widespread RNA Editing of Embedded Alu Elements in the Human Transcriptome. Genome Res., 2004 14(9): p. 1719–1725. [PubMed: 15342557]
57. Zaringhalam M and Papavasiliou FN, Pseudouridylation meets next-generation sequencing. Methods, 2016 107: p. 63–72. [PubMed: 26968262]
58. Hong T, et al., Precise Antibody-Independent m6A Identification via 4SedTTP-Involved and FTO-Assisted Strategy at Single-Nucleotide Resolution. J. Am. Chem. Soc, 2018 140(18): p. 5886–5889. [PubMed: 29489347]
59. Merrifield RB, Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. J. Am. Chem. Soc, 1963 85(14): p. 2149–2154.
60. Beaucage SL and Iyer RP, Advances in the Synthesis of Oligonucleotides by the Phosphoramidite Approach. Tetrahedron, 1992 48(12): p. 2223–2311.
61. George CX, John L, and Samuel CE, An RNA editor, adenosine deaminase acting on double-stranded RNA (ADAR1). J. Interferon Cytokine Res, 2014 34(6): p. 437–46. [PubMed: 24905200]
62. Bass BL, et al., A standardized nomenclature for adenosine deaminases that act on RNA. RNA, 1997 3(9): p. 947–949. [PubMed: 9292492]
63. Macbeth MR and Bass BL, Large-Scale Overexpression and Purification of ADARs from Saccharomyces cerevisiae for Biophysical and Biochemical Studies. Methods Enzymol., 2007 424: p. 319–331. [PubMed: 17662848]
64. Zinshteyn B and Nishikura K, Adenosine-to-inosine RNA editing. Wiley interdiscip. Rev. Sys. Biol. Med, 2009 1(2): p. 202–209.
65. Nishikura K, Functions and Regulation of RNA Editing by ADAR Deaminases. Annu. Rev. Biochem, 2010 79: p. 321–349. [PubMed: 20192758]
66. Peng PL, et al., ADAR2-dependent RNA editing of AMPA receptor subunit GluR2 determines vulnerability of neurons in forebrain ischemia. Neuron, 2006 49(5): p. 71933.
67. Hu T, et al., TLR8 activation and inhibition by guanosine analogs in RNA: Importance of functional groups and chain length. Bioorg. Med. Chem, 2018 26(1): p. 77–83. [PubMed: 29174509]
68. Galipon J, et al., Differential Binding of Three Major Human ADAR Isoforms to Coding and Long Non-Coding Transcripts. Genes, 2017 8(2): p. 68.
69. Slotkin W and Nishikura K, Adenosine-to-inosine RNA editing and human disease. Genome Med., 2013 5(11): p. 105. [PubMed: 24289319]
70. Rice GI, et al., Mutations in ADAR1 cause Aicardi-Goutières syndrome associated with a type I interferon signature. Nature Genet., 2012 44: p. 1243. [PubMed: 23001123]
71. Matthews MM, et al., Structures of human ADAR2 bound to dsRNA reveal base-flipping mechanism and basis for site selectivity. Nat. Struc. Mol. Biol, 2016 23: p. 426.
72. Maydanovych O and Beal PA, C6-Substituted Analogues of 8-Azanebularine: Probes of an RNA-Editing Enzyme Active Site. Org. Lett, 2006 8(17): p. 3753–3756. [PubMed: 16898809]
73. Eifler T, Pokharel S, and Beal PA, RNA-Seq analysis identifies a novel set of editing substrates for human ADAR2 present in Saccharomyces cerevisiae. Biochemistry, 2013 52(45): p. 7857–69. [PubMed: 24124932]
74. Haudenschild BL, et al., A Transition State Analogue for an RNA-Editing Reaction. J. Am. Chem. Soc, 2004 126(36): p. 11213–11219. [PubMed: 15355102]
75. Mizrahi RA, et al., Nucleoside analog studies indicate mechanistic differences between RNA-editing adenosine deaminases. Nucleic Acids Res., 2012 40(19): p. 9825–9835. [PubMed: 22885375]
76. Eggington JM, Greene T, and Bass BL, Predicting sites of ADAR editing in doublestranded RNA. Nature Comm., 2011 2: p. 319.

Highlights

- Three commonly used methods for generating nucleoside analog-containing nucleic acids are described
- Examples of the use of each method in the study of the epitranscriptome are examined
- Strengths and limitations to each method discussed

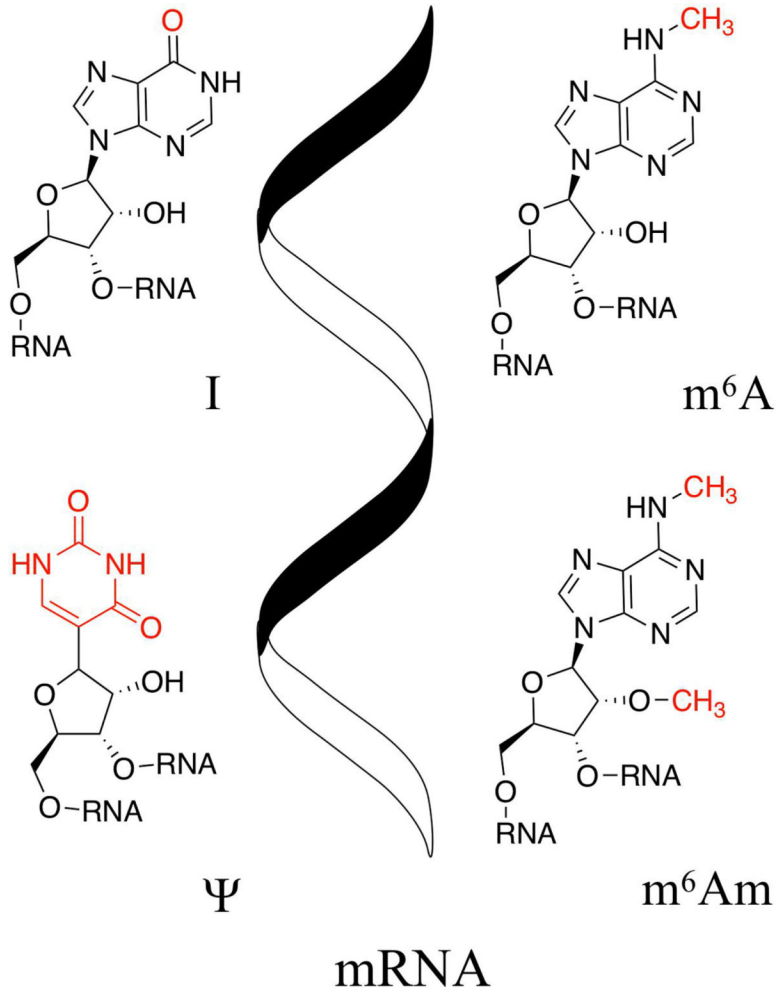

Figure 1.

The structures of naturally occurring mRNA modifications found in mammals. For clarity, the RNA modification is highlighted in red. Abbreviations of the modification are indicated below each molecule. Inosine (I), N^6-methyladenosine (m^6A), pseudouridine (Ψ), N^6-2'-O-dimethyladenosine (m^6Am).

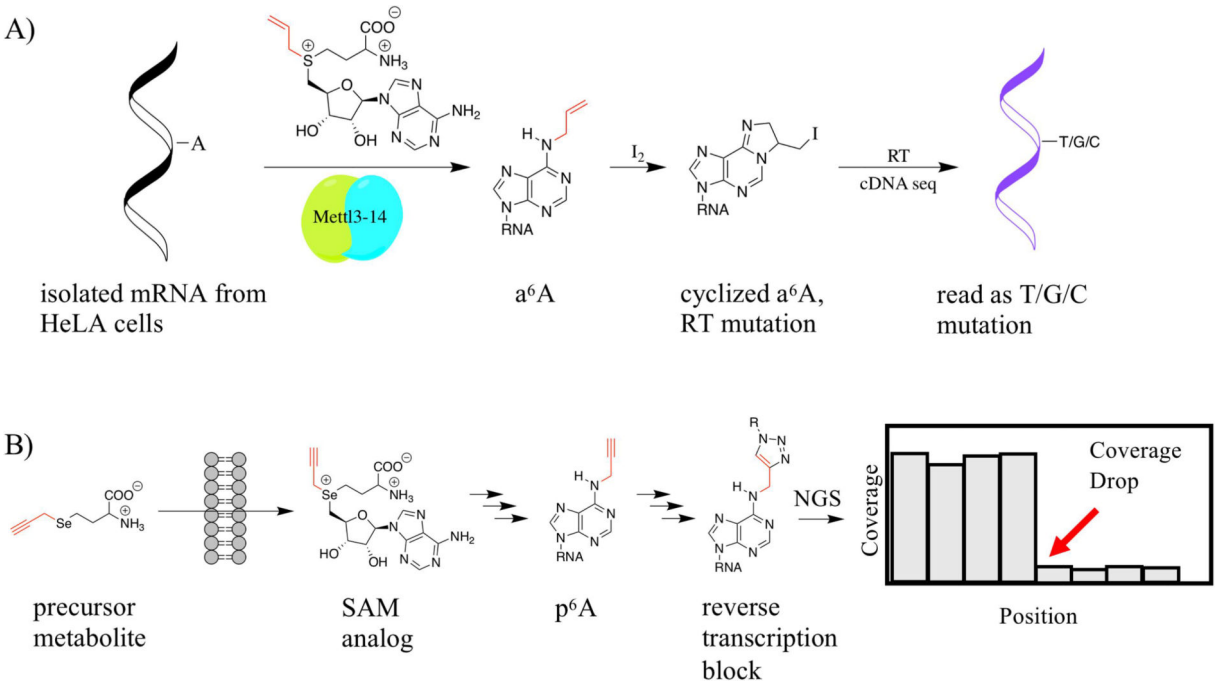

Figure 2.

The study of m^{6}A using SAM analogs to identify Mettl3–14 modification sites. **A**) Isolated mRNA from HeLa cells can be incubated in the presence of an allyl SAM cofactor and the Mettl3–14 complex to selectively label m^{6}A sites with a^{6}A [39]. Subsequently, mRNA can be reacted with iodine to produce a cyclized a^{6}A molecule that is no longer RT silent. After mRNA is converted into cDNA and sequenced, there will be an A to T/G/C mutation which is indictive of a m^{6}A site. **B**) Propargyl-L-selenohomocysteine can be taken up by HeLa cells and incorporated into a propargyl SAM analog [40]. The propargyl group is transferred by cellular methyltransferases to m^{6}A sites. The isolated mRNA containing p^{6}A can be subjected to CuAAC click reactions to install bulky groups that block RT. cDNA obtained from RT can be sequenced using NGS. As a result of installation of the bulky group, coverage drops will occur where there are m^{6}A sites.

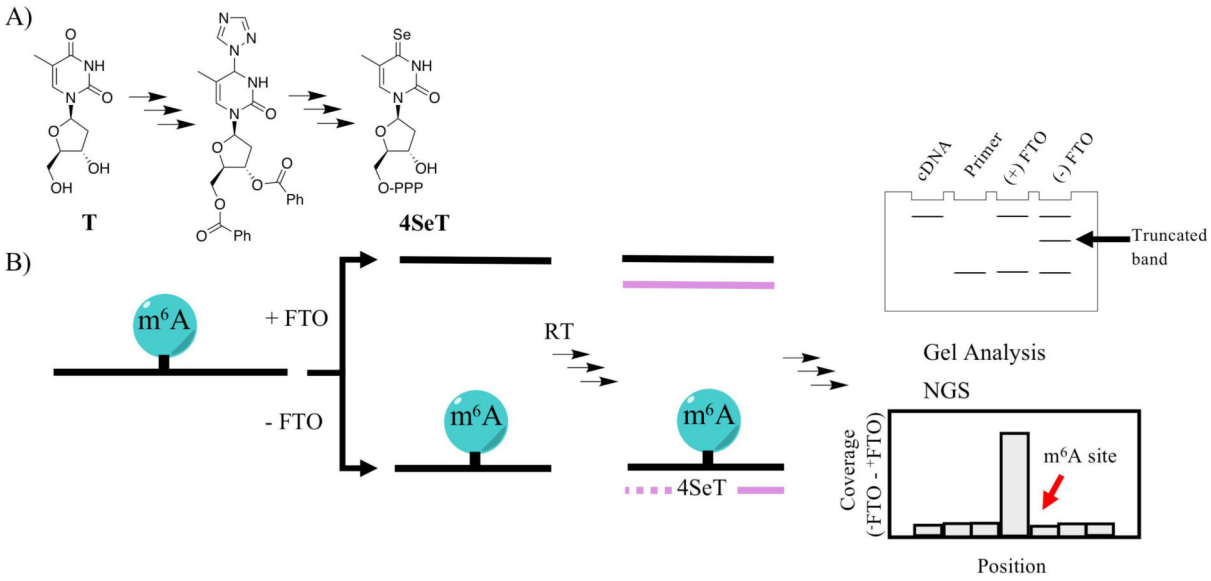

Figure 3.

The method of using a DNA triphosphate analog to distinguish m^6A sites during RT [58]. A C4 selenium containing thymidine DNA triphosphate can be synthesized in several steps from thymidine. During RT, 4SeT can successfully be incorporated across from adenosine but will fail to be incorporated across from m^6A, resulting in truncated cDNA. The resulting cDNA can be sequenced through NGS or specific targets can be evaluated by gel analysis. To aide in the determination of truncations, RNA can be treated with or without FTO prior to RT. Appearance of a truncated band by denaturing gel analysis is evidence of a m^6A site. In NGS data, this would be determined by a coverage drop due to truncation of cDNA.

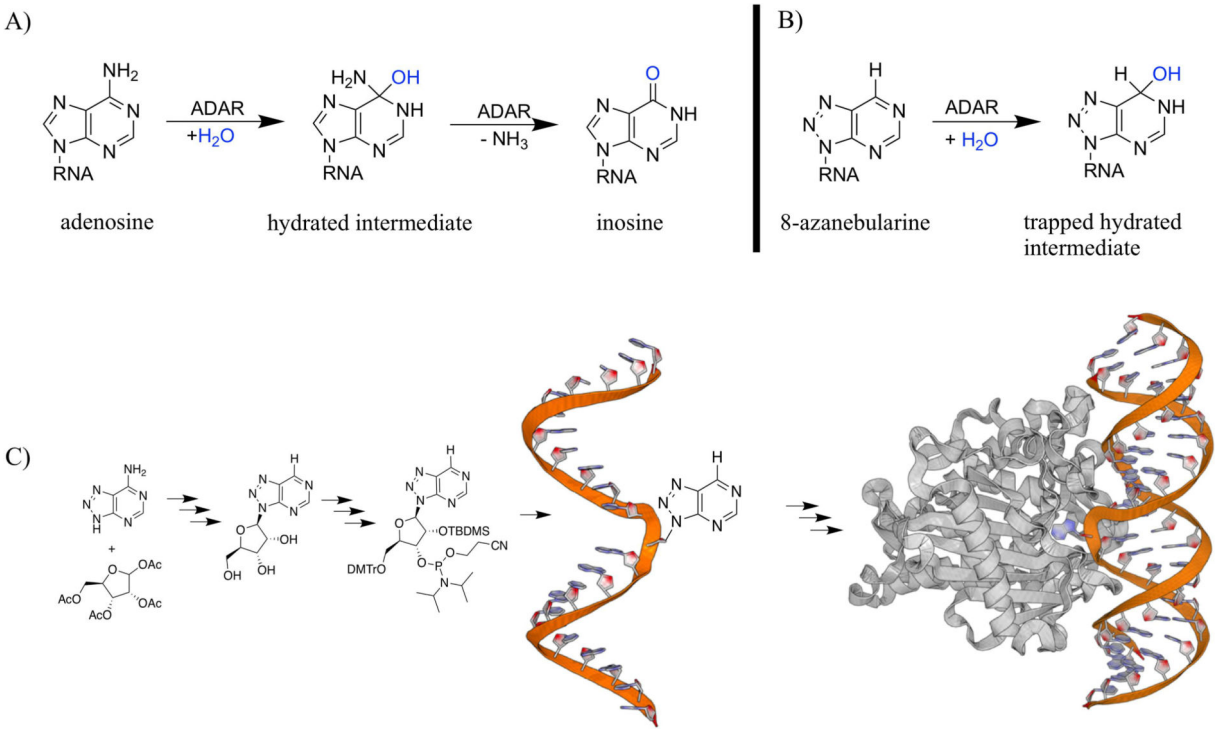

Figure 4.

Method for crystallization of ADAR2 bound to dsRNA containing a transition state analog [71]. **A**) A-to-I editing reaction pathway showing the hydrated intermediate that leads to the formation of inosine. **B**) 8-Aza-nebularine when subjected to the ADAR reactions becomes a trapped hydrated intermediate. **C**) 8-Azanebularine phosphoramidite can be synthesized from the starting materials of 8-azaadenosine and a tetraacetate protected ribofuranose. Subsequently, the exocyclic amine can be removed, and the sugar's protecting groups removed. The standard phosphoramidite protecting groups are installed prior to RNA solid phase synthesis. Finally, crystal structures can be solved of ADAR2 with dsRNA containing 8-azanebularine. This method afforded four high resolution crystal structures of ADAR2 bound to its substrate.