Small RNA Modifications: Integral to Function and Disease

Xudong Zhang,1,3 Aaron E. Cozen,2,3 Ying Liu,1 Qi Chen,1,* and Todd M. Lowe2,*

Small RNAs have the potential to store a secondary layer of labile biological information in the form of modified nucleotides. Emerging evidence has shown that small RNAs including microRNAs (miRNAs), PIWI-interacting RNAs (piRNAs) and tRNA-derived small RNAs (tsRNAs) harbor a diversity of RNA modifications. These findings highlight the importance of RNA modifications in the modulation of basic properties such as RNA stability and other complex physiological processes involved in stress responses, metabolism, immunity, and epigenetic inheritance of environmentally acquired traits, among others. High-resolution, high-throughput methods for detecting, mapping and screening these small RNA modifications now provide opportunities to uncover their diagnostic potential as sensitive disease markers.

A Glimpse into RNA Modifications

RNAs are post-transcriptionally altered with a variety of chemical modifications in all domains of life. More than 100 types of RNA modification have been identified to date, in many cases showing context-dependent functions that are only beginning to become clear [1]. Post-transcriptional modifications are known to be especially prevalent in stable, structured RNAs such as tRNA and rRNA where they play important roles in biogenesis, correct structural folding, and function [2]. Emerging evidence indicates that a diverse set of modifications are also found in transient carriers of information, including mRNAs and an expanding catalog of small and large non-coding RNAs. These are exemplified by the recent discovery of reversible 6-methyladenosine (m6A) modifications in mRNAs [3,4] as well as key enzymes for their dynamic regulation: writer (see Glossary) [5,6], eraser [7,8], and reader [9]. Other studies have documented pseudouridine [10–12], 5-methylcytidine (m5C) [13,14], and, most recently, 1-methyladenosine (m1A) [15,16] in mRNAs. RNA modifications are also prevalent in small RNAs, where they could similarly carry information important for various cellular functions (Figure 1). Here we briefly discuss the roles of representative RNA modifications and RNA editing events in the biogenesis and function of small RNAs. We also highlight the potential clinical relevance of such modified small RNAs and their possible utilization as diagnostic markers of disease. Lastly, we summarize recent technical advances in high-throughput approaches for profiling RNA modifications in small RNAs, with a particular focus on heavily modified tRNAs and tsRNAs. New sequencing-based methods are highly complementary to established mass spectrometry-based analyses, with the advantages of positional resolution and wide availability.

2’-O-Methylation in Small RNAs: Stability and Beyond

miRNAs are key players in RNA silencing pathways inducing mRNA degradation, translational repression, and transcriptional silencing [17]. 2’-O-Methylation, mediated by HUA...
enhancer 1 (HEN1), was first recognized in plants [18], where it protects miRNAs from small RNA decay pathways that are triggered by 3′-uridylation [19,20], a well-characterized RNA decay pathway [21] (Figure 2A). Although mature miRNAs in animals lack 2′-O-methylation, other mammalian small RNAs such as siRNAs and P-element-induced wimpy testis (Piwi)-interacting RNAs (piRNAs) can be methylated by HEN1 orthologs, which similarly protect these small RNAs from 3′-uridylation [19]. This conserved mechanism provides a rationale in translational research for promoting the stability of small RNA therapeutic agents, with the potential for more efficient RNAi therapy [22]. 2′-O-methylation in plant miRNAs also provides a potential mechanism for prolonging their functional half-life in the human body when ingested, supporting the intriguing – and still controversial – idea that small RNAs derived from plant-based diets can modulate human physiology [23–28].

In addition to increasing small RNA stability, specific 2′-O-methylation of guanosine in bacterial tRNA at position 18 (Gm18) is necessary and sufficient to suppress immune activation mediated by Toll-like receptor 7 (TLR-7) during human host–pathogen responses [29,30]. Most bacteria lack the gene responsible for G18 tRNA 2′-O-methylation, TrmH; thus, it may be a positively selected trait in symbiotic and pathogenic species. This example demonstrates an important role for RNA modification in modulating recognition of self- versus non-self-derived RNAs [29,30] (Figure 2B). Overall, the functions of 2′-O-methylation suggest myriad potential translational applications related to the stability and immunogenicity of small RNA-based therapeutics.

**m5C: A Stress Sensor and Transgenerational Mark in tRNA and tsRNAs**

m5Cs are frequently detected in various non-coding RNAs such as tRNA, rRNA, and, more recently, miRNAs. While the full range of potential functions for m5C in RNA is not completely understood, bisulfite RNA sequencing, which pinpoints m5C residues [31], has helped reveal the importance of these residues in tRNAs [32–36]. One group has shown that m5C modifications of tRNAs by the methyltransferases DNA methyltransferase 2 (Dnmt2) and NOP2/Sun RNA methyltransferase family member 2 (Nsun2) are required for normal growth and development in mice [33] and flies [36], supporting a close relationship between m5C modification and tRNA stability. Moreover, m5C alters the tRNA-binding activity of the endonuclease angiogenin, as loss of m5C RNA methylation increases angiogenin-mediated cleavage of tRNAs in response to a variety of cellular stresses, generating tsRNAs [32–34]. For instance, accumulation of 5′-tsRNAs reduces protein translation rates [37] and tRNAs lacking Dnmt2-dependent methylation lead to codon mistranslation, which also alters protein expression [35]. These tRNA changes disrupt protein synthesis at a systemic level, negatively affecting body growth as well as neurological and stem cell functions [32–35]. Furthermore, Dnmt2-mediated virus RNA methylation has also been implicated in efficient antiviral defense in *Drosophila* sp. [38]. Together, these studies highlight an intricate interplay between m5C modification of tRNAs, tsRNA biogenesis, cellular stress responses, and protein synthesis – fundamental processes relevant to diverse disease conditions (Figure 2C).

Recent discoveries have also revealed important and surprising roles for m5C in RNA-mediated transgenerational epigenetic inheritance [39,40]. Studies have demonstrated that the m5C methyltransferase Dnmt2 is required for RNA-mediated epigenetic alterations in fur color and an overgrowth phenotype in mice [39]. Both phenotypes were shown to be induced in mice with wild-type Dnmt2 whereas neither phenotype was transmitted in mutants deficient in Dnmt2 [39]. More recent work has also implicated m5C in intergenerational epigenetic transmission of metabolic alterations arising from paternal dietary exposures. Male mice fed high-fat diets (HFDs) were reported to transmit impaired glucose tolerance and insulin resistance to offspring and some of these effects could also be produced by injecting total RNA or tsRNAs from HFD sperm into normal murine zygotes. Notably, tsRNAs constitute the dominant small RNA species in mature sperm and were shown to exhibit elevated m5C and m2G levels in HFD sperm.

---

**Glossary**

**Bisulfite RNA sequencing:** method to detect m5C methylation patterns in RNAs by bisulfite treatment of RNAs.

**cDNA library preparation:** synthesis of single-strand complementary DNAs from RNA samples by reverse transcription.

**Eraser:** a demethylase or associated enzymatic complex that removes RNA modifications.

**P-element-induced wimpy testis (Piwi)-interacting RNAs (piRNAs):** the largest class of small non-coding RNAs expressed in animal cells; form RNA–protein complexes with Piwi proteins.

**Reverse transcription:** synthesis of single-strand complementary DNA using single-strand RNA as the template, mediated by RNA-dependent DNA polymerase (reverse transcriptase).

**RNA silencing pathways:** mechanisms to suppress gene expression by RNA-induced degradation of mRNA or inhibition of protein synthesis.

**3′-Uridylation:** the enzymatic addition of uridines to the 3′ end of small RNAs (miRNA, pRNA, etc.); can trigger the small RNA degradation pathway.

**Writer:** a methyltransferase or associated enzymatic complex that adds chemical modifications to RNAs.
compared with those from males fed normal diets [40]. These separate lines of evidence involving tRNAs/tsRNAs implicate m5C in transgenerational epigenetic inheritance. Whereas synthetic tsRNAs lacking RNA modifications degrade rapidly when incubated in zygote lysate or serum, endogenous tsRNAs are more stable, suggesting possible roles for modifications in promoting the stability [40,41] and transgenerational effects of tsRNAs after fertilization [40]. These modifications may also contribute to the secondary structures of tsRNAs that lead to altered binding specificities to their targets (DNA, RNAs, or proteins) [40]. Together, these observations suggest that small RNA modifications are important in modulating the efficiency of RNA-mediated transmission from fathers to offspring.

tRNAs are increasingly recognized as functional signaling and regulatory molecules in a diversity of contexts [32–34,37,40–47]. Mature tRNAs are among the most abundant and most heavily modified RNA species. Aberrations in tRNA modification play roles in cellular glucose metabolism [48] and have been implicated in various pathologies, including cancer, type 2
Figure 2. Representative Biological Functions of RNA Modifications and RNA Editing Events in Small RNAs. (A) 2′-O-Methylation at the 3′ end of small RNAs can protect these from uridylation, preventing degradation by RNA decay pathways. (B) Discrimination of self and non-self RNAs. Bacterial tRNAs lacking 2′-O-methylation at position 18 (Gm) are recognized by Toll-like receptor 7 (TLR-7) on the endosomal surface of relevant immune cells, activating downstream innate immunity. (C) m5C fine-tunes angiogenin (ANG)-mediated tRNA cleavage under conditions of cellular stress (nutritional starvation etc.). tRNAs lacking m5C are preferentially cleaved by ANG, leading to accumulation of tRNA-derived small RNAs (tsRNAs) and associated pathology (as in the case of certain tumors). (D) Adenosine-to-inosine (A-to-I) editing influences miRNA biogenesis. pri-miRNAs harboring inosines [generated by adenosine deaminase acting on RNA 1 (ADAR1)] are resistant to cleavage by double-strand RNA (dsRNA)-specific endoribonuclease (Drosha), the core nuclease in miRNA processing, thus decreasing mature miRNA production. (E) Inosine discriminates endogenous versus viral RNAs. Melanoma differentiation-associated protein 5 (MDA5) is a RIG-I-like receptor dsRNA helicase enzyme that binds viral dsRNAs lacking inosine bases to trigger antiviral responses.
diabetes, neurological disorders, and mitochondria-linked disorders [49]. A frequently changing population of tsRNAs is also ubiquitous in cells and contains some portion of the RNA modifications made in the parent tRNA molecules. The functions of these tsRNA modifications and the extent to which they are altered after tsRNA biogenesis are largely unexplored but are likely to alter in vivo tsRNA dynamics including stability, secondary structure, and interactions with proteins [50]. Thus, the extensive potential repertoire of tsRNA modifications could be important in a wide diversity of physiological processes and pathological conditions across species. Similarly, rRNAs may also give rise to abundant and highly modified small RNAs (rsRNAs) [51], but these harbor as-yet-unknown functions in normal physiology and disease.

**RNA Editing Events in Small RNAs: Biology and Disease Associations**

RNA editing is a process that alters the primary nucleotide sequence of RNA by specific enzymes. The chief distinction between RNA editing and RNA modification is that editing is usually irreversible and generally understood to produce altered base-pairing behavior. RNA editing events are widely observed in mRNA, tRNA, rRNA, and miRNA in all kingdoms of life. The most frequent editing events are deaminations that convert adenosine to inosine (A to I) and cytidine to uridine (C to U) [52]. Adenosine deaminase acting on RNA (ADAR) is an enzyme that mediates A-to-I editing by hydrolytic deamination of adenosine in double-strand RNAs (dsRNAs); this can occur within both the coding sequence and the non-coding regions of an mRNA transcript [53].

Deficiency of A-to-I editing activity in certain mRNA transcripts has been implicated in human diseases including amyotrophic lateral sclerosis (ALS) [54] and cancer [55]. miRNAs are also subject to A-to-I editing [53]. Primary miRNA (pri-miRNA) transcripts are typically ~60–70 nt long and can form dsRNA structures that can be edited by ADARs [53]. Deletion of ADAR1 or ADAR2 in mice has been shown to prevent degradation of the miRNA precursor pri-miR-142 by a component of the RNA-induced silencing complex (RISC) resulting in increased levels of the mature miRNA miR-142, thus illustrating a molecular mechanism by which miRNA biogenesis can be controlled [56] (Figure 2D). A-to-I editing also occurs, although less frequently, in mature miRNAs [57]. Edited mature miRNAs (e.g., miR-376) can alter their targeting specificity [58], a change that has been found to promote invasiveness of glioblastoma cells in a mouse model [59]. Moreover, alterations in ADAR editing activity levels have been linked to stem cell maintenance, hereditary autoimmune diseases, and disease progression in cancer, suggesting the potential of using A-to-I editing as a putative biomarker of disease severity for certain conditions [60].

Similar to 2′-O-methylation, the inosines in RNAs produced by A-to-I editing can be used by the innate immune system to discriminate self versus non-self RNAs. When viral dsRNAs lacking inosine enter the cell, RNA sensors such as melanoma differentiation-associated protein 5 (MDA5), a RIG-I-like receptor dsRNA helicase enzyme, can bind dsRNA and activate antiviral responses [61,62] (Figure 2E). By contrast, endogenous dsRNAs edited by ADAR1 contain A-to-I edits that prevent MDA5 binding and antiviral immune responses [61,62] (Fig. 2E).

C-to-U RNA editing is mediated by cytidine deaminases, which was first observed acting on vertebrate apolipoprotein B (apoB) mRNA transcripts [52]. The C-to-U editing sites are usually enriched in the 3′ UTR of mRNAs and influence the targeting specificity of miRNAs [63]. C-to-U editing is also required for the maturation process of mitochondrial tRNAAsp in marsupials [64]. Future high-throughput sequencing studies to systematically map C-to-U RNA editing sites may lead to the discovery of novel biological functions and their roles in disease.

**Profiling Small RNA Modifications As Biomarkers of Disease**

Several different approaches have been used for high-throughput analysis of RNA modifications; each has different strengths in terms of its ability to monitor specific RNA modifications,
characterize global modification profiles in biological samples, or reveal the identity of modified RNAs with sequence-level specificity. Antibody-based detection of modified small RNAs has shown promise as a sensitive diagnostic for identifying early markers of disease. For example, m1A antibodies can detect tRNA conformational changes and circulating tsRNAs that arise from acute cellular stress, potentially providing an earlier indicator of tissue injury than other markers of apoptosis and DNA damage (Figure 3A) [65]. High-throughput liquid chromatography–tandem mass spectrometry (LC-MS/MS) provides a complementary approach that can be used for quantification of multiple RNA modifications within a single RNA sample [40,51,66–68]. LC-MS/MS has uncovered a diversity of previously overlooked RNA modifications among small RNAs (10–30 nt, 30–60 nt) in mouse liver [51], revealing dynamic changes for multiple RNA modifications in mouse models of diabetes (Gm, m6Cm, Cm, Am, Um) [51] (Figure 1B). LC-MS/MS profiling has also identified altered modification profiles, with increased levels of m5C and m2G, in the 30–40-nt fraction of sperm RNAs (predominantly tsRNAs) in mice fed HFDs (described above), suggesting that these chemical modifications may transmit information about paternal diet to offspring [40] (Figure 1C). Similar approaches have been used to detect altered modified nucleosides in the urine of cancer patients, where modification profiles could potentially be utilized as biomarkers for clinical diagnosis in the design of personalized therapy [69]. Thus, the combined use of antibody- and/or LC-MS/MS-based profiling of RNA modifications holds great

![Figure 3](image-url)
promise for noninvasive tests (e.g., of blood, sperm, or urine) to help guide the clinical diagnosis and treatment of a variety of disorders.

**Emerging Methods for Profiling and Mapping RNA Modifications in Small RNAs**

While LC-MS/MS is a powerful approach for quantifying the compositional changes of RNA modifications, this method requires digestion of RNAs into short segments or single nucleosides before detection, therefore making it difficult to unambiguously assign modifications to specific transcripts (Figure 4A) [51,66,67]. Thus, complementary approaches to map RNA modifications to specific RNAs are required to gain a more comprehensive view of modified small RNAs.

Sequencing-based protocols have been developed to identify RNAs modified with m⁵C, pseudouridine, m⁶A, and m⁷A, in some cases also providing precise positional information about the location of these modifications [70]. For example, RNA bisulfite sequencing can detect m⁵C in both large and small RNAs by converting unmethylated cytosine (C) to uracil (U) with bisulfite, which will be read as thymine (T) during sequencing, while methylated C (m⁵C) remains unconverted. By comparing the sequencing results of bisulfite-treated samples versus untreated samples, the position of m⁵C in RNAs can be reproducibly and quantitatively analyzed [31,71]. The mapping of both m⁶A and m⁷A modifications in mRNAs is based on immunoprecipitation (IP) of methylated RNAs with antibodies specific to each modification, followed by RNA sequencing of the IP-enriched RNA fragments. When these sequencing reads are mapped to the genome, enrichment peaks correspond to the locations of m⁶A and m⁷A in the mRNAs [3,4,16]. These ‘first-generation’ antibody-based approaches can identify modified mRNAs and can localize modified residues to ~100–200-nt regions within transcripts; however, the precise m⁶A positions are not known to single-nucleotide resolution and this method cannot be applied.

**Figure 4. Emerging Methods for Profiling RNA Modifications in Small RNAs.** (A) Profiling and quantification of RNA modifications using liquid chromatography–tandem mass spectrometry (LC-MS/MS). (B) Identification and mapping of methylated RNAs using Escherichia coli AlkB treatment. AlkB-facilitated RNA methylation sequencing (ARM-seq) measures changes in the abundance (either full read or no read) of tRNA-derived small RNAs (tsRNAs) after AlkB treatment, which removes RNA modifications blocking reverse transcriptase during cDNA library construction. Demethylase-thermostable group II intron reverse transcription tRNA sequencing (DM-tRNA-seq) can identify modifications via comparison of early termination (short) reads in non-treated samples with full reads from samples treated with AlkB; DM-tRNA-seq preferentially sequences full-length tsRNAs, which are in much higher abundance than tsRNAs. RT, reverse transcription. TGIRT, thermostable group II intron reverse transcriptase.
to mapping modifications on small RNAs. A recent improvement of this technique crosslinks the antibody to the bound RNA fragment with UV light, producing a specific pattern of read truncations and base changes adjacent to the m\(^{\text{6}}\)A site and thus allowing single-base resolution of m\(^{\text{6}}\)A sites [72]. This improved method can be used to map human and mouse small nucleolar RNAs (snoRNAs) of approximately 70–90 nt in length [72] and future improvements may lead to precise mapping of even smaller RNAs (<40 nt). Consequently, methodologies based on the combination of chemical treatments with RNA sequencing hold great promise for precisely mapping a variety of RNA modifications in small RNAs.

Notably, other modifications, including m\(^{7}\)A, m\(^{1}\)G, and m\(^{3}\)C, are common in tRNAs (but not miRNAs) and interfere with normal intra- and intermolecular base pairing. Reverse transcription, an essential step in cDNA library preparation for RNA-seq, is halted at a high rate at these so-called ‘hard-stop’ modifications. This significantly impairs efficient detection of small RNAs containing these modifications, suggesting that a large but unknown fraction of biologically important tsRNAs has been missed in all prior studies and large surveys (e.g., ENCODE, TCGA) employing standard small RNA-seq methods to profile small RNA populations.

Recently, two complementary new approaches have been developed to tackle the problem of not overlooking tsRNAs. By using pretreatment of RNAs with the Escherichia coli deamidating enzyme AlkB to demethylate m\(^{7}\)A, m\(^{1}\)G, and m\(^{3}\)C before sequencing, detection of methylated mature iRNAs [demethylase-thermostable group II intron reverse transcription tRNA sequencing (DM-tRNA-seq)] [73] and tsRNAs [AlkB-facilitated RNA methylation sequencing (ARM-seq)] [74] has been possible (Figure 4B). Comparative methylation analysis using ARM-seq has revealed a wealth of previously undetected methylated tsRNAs and, importantly, demonstrated that these are modified very similarly to the mature tRNAs from which they are derived [74]. Combining these sequencing-based approaches with LC-MS/MS is a promising means of mapping specific types of RNA modifications with both compositional and positional information (Figure 4A,B). A detailed map of the many types of RNA modifications across the major types of small RNAs presents one among many challenges ahead. Charting the dynamics of these modifications in different cell types and physiological processes is another. Together, these new views are critical cornerstones to addressing further questions about the functional roles of modifications of small RNA and their potential use in medicine as both diagnostic markers and therapeutic targets (see Outstanding Questions) (Box 1).

**Box 1. The Clinician’s Corner**

RNAs extracted from biological samples such as blood and urine have shown great diagnostic value. In addition to the specific profile of small RNAs detected, RNA modifications represent another layer of biological information that can also be harnessed for diagnostic use.

Disease conditions (e.g., those involving acute kidney cell damage) can lead to the release of modified RNA nucleosides (m\(^{7}\)A) that are readily detected using antibody-based methods, potentially providing a sensitive diagnostic for early tissue injury and cell damage.

High-throughput detection methods based on mass spectrometry can simultaneously detect multiple types of RNA modifications in one small RNA sample, providing new molecular signatures of disease conditions as in the case of metabolic disorders.

High throughput RNA-seq has been used to show that RNA editing events (A-to-I) in small RNAs correlate with the malignancy grade of some tumors, suggesting the potential use of this approach as a diagnostic screen.

High-resolution profiling of small RNA modifications holds great promise for sensitive and specific detection of aberrant physiological conditions and disease as well as for the development of novel therapeutic approaches designed to manipulate specific RNA modifications.
Concluding Remarks: How Many New Bits of Information?

Compared with the accelerating rate of discovery of many types of small RNAs [75], the study of small RNA modifications in biology and disease remains in its infancy. Emerging evidence indicates that RNA modifications are important molecular markers that can alter the biogenesis, function, and informational capacity of small RNAs. New techniques including high-throughput LC-MS/MS and sequencing-based approaches that specifically target modified RNAs using antibodies, enzymatic treatments, or chemical alterations hold great promise for elucidating the role of tRNAs, tRNA derivatives, and other modified small RNAs as signaling and regulatory molecules (see Outstanding Questions (Box 1)). These approaches may prove particularly powerful when used in combination, providing new insights into the physiological roles of modified small RNAs as well as new opportunities for diagnosis of small RNA-associated diseases.

Acknowledgments

This research was partially supported by the NIH/NHGRI (HG006753-02 to T.M.L., P30GM110767-03 to Q.C.) and start-up support from the University of Nevada, Reno School of Medicine (to Q.C.).

References


5. Ping, X.L. et al. (2014) Mammalian WTAP is a regulatory subunit of the RNA m^6^-methyladenosine methyltransferase. Cell Res. 24, 177–189


Outstanding Questions

How evolutionarily conserved are small RNA modifications across different types of small RNAs?

Are there tissue-specific or developmentally regulated small RNA modifications?

Which small RNA modifications are reversible versus irreversible and what are their regulatory machineries?

How do specific modifications alter the structure, stability, and molecular interactions of small RNAs?

Do RNA modifications contribute to the stability and cross-tissue transfer of mobile small RNAs?

Can small RNA modifications be used as general diagnostic markers of disease (metabolic dysfunction, cancer, infections, or others)?