Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Chemical Biology

RNA epigenetics — chemical messages for posttranscriptional gene regulation

Ian A Roundtree and Chuan He

Chemical modifications in cellular RNA are diverse and abundant. Commonly found in ribosomal RNA (rRNA), transfer RNA (tRNA), long-noncoding RNA (lncRNA), and small nuclear (snRNA), these components play various structural and functional roles. Until recently, the roles of chemical modifications within messenger RNA (mRNA) have been understudied. Recent maps of several mRNA modifications have suggested regulatory functions for these marks. This review summarizes recent advances in identifying and understanding biological roles of posttranscriptional mRNA modification, or 'RNA epigenetics', with an emphasis on the most common internal modification of eukaryotic mRNA, N^6-methyladenosine (m^6A). We also discuss YTH proteins as direct mediators of m^6A function and the emerging role of this mark in a new layer of gene expression regulation.

Address
Department of Chemistry, Department of Biochemistry and Molecular Biology, Institute for Biophysical Dynamics and Howard Hughes Medical Institute, The University of Chicago, 929 East 57th Street, Chicago, IL 60637, USA

Corresponding author: He, Chuan (chuanhe@uchicago.edu)

Current Opinion in Chemical Biology 2016, **30**:46–51

This review comes from a themed issue on **Omics**

Edited by **Daniel K Nomura**, **Alan Saghatelian** and **Eranthie Weerapana**

For a complete overview see the Issue and the Editorial

Available online 26th November 2015

http://dx.doi.org/10.1016/j.cbpa.2015.10.024

1367-5931/Published by Elsevier Ltd.

Messenger RNA modifications

Posttranscriptional modifications in RNA were discovered nearly 60 years ago with the identification of pseudouridine (ψ) as an abundant component of tRNA. Since this time over 100 additional modifications of RNA have been documented, including internal modifications within coding transcripts [1,2].

N^6-methyladenosine (m^6A)

The most abundant internal modification in eukaryotic mRNA is methylation at the N^6 position of adenosine, which is present between ~3 times per mRNA on average in mammalian cells [3]. This modification is installed by a multicomponent methyltransferase complex [4,5•,6] and can be reversed by functionally significant demethylases [7••,8••]. Two independent efforts in 2012 mapped the location of m^6A in mRNA using antibody-based affinity capture coupled to high throughput sequencing [9•,10•]. These experiments revealed a previously unknown enrichment for m^6A within coding regions and the 3′ untranslated region, peaking sharply near the stop codon. This distribution, together with noted conservation of the RRACH (R = A,G; H = A,C,U) sequence motif [11,12] between mouse and human methylomes is suggestive of a mark with fundamental importance in RNA biology. Additionally, studies have identified a unique methylation pattern in *Arabidopsis thaliana* [13] and rice [14], revealing additional enrichment near the start codon and the requirement for m^6A in plant development [15]. m^6A is also an abundant component of viral RNA [16] and occurs in bacterial mRNA [17].

Identifying the precise location of m^6A within a transcript remains a challenge. Current methods rely on chemical fragmentation of mRNA to increase resolution, but fail to provide single base information. Site-specific cleavage and radioactive-labeling followed by ligation-assisted extraction and thin-layer chromatography (SCARLET) can deliver base-resolution information on location and modification fraction of m^6A, but is not applicable to high-throughput analysis [18]. Recent antibody-based cross-linking strategies have increased the resolution of m^6A methylomes, and utilized unique mutation signatures to map sites at the individual-nucleotide level [19,20]. Such high-resolution data will enable researchers to observe perturbations of individual m^6A loci in a variety of biological contexts.

m^6A methyltransferases — 'Writers'

m^6A is installed posttranscriptionally by a methyltransferase complex consisting of METTL3 and METTL14, as well as the regulatory subunit WTAP (Wilms' tumor 1-associating protein) and in mammalian cells [4,5•,6]. Both METTL3 and METTL14 are capable of transferring a methyl group from cofactor S-adenosyl methionine (SAM) to GGACU and GGAUU sequences within single stranded and stem-loop RNA *in vitro*, with METTL14 showing the greatest individual activity well short of that observed for the complex. Perturbation of individual subunits each leads to significant decreases in the m^6A abundance in mRNA. Knockdown of METTL3 or METTL14 in mouse embryonic stem cells (mESCs) results in decreased capacity for self-renewal [21]. Knockout models of *mettl3* in mESCs have revealed the requirement for m^6A

methylation in early differentiation processes [22,23]. Each study highlights a failure of mESCs to downregulate pluripotency markers and upregulate transcripts required for differentiation, potentially due to the absence of the m^{6}A-dependent mRNA decay (discussed below). Geula *et al.* also show this phenotype in *mettl14* mESCs, indicating this process requires methylation activity of the multiprotein complex [23]. Despite an essential role of m^{6}A in mammalian development and viability, we currently do not know how the m^{6}A methyltransferase complex is regulated. Recent work has identified the protein interaction network of METTL3 and two distinct subsets of m^{6}A modification as WTAP-dependent or WTAP-independent, adding an additional element of complexity to the m^{6}A epitranscriptome [24]. However, consensus methylation motifs are common to nearly every mRNA transcript, yet only a very small fraction contains methylation. Selectivity of the methyltransferase complex, perhaps driven by guide RNAs or chromatin marks, as well as regulation of its catalytic activity, is an active area of research.

m^{6}A demethylases — 'Erasers'

The discovery of m^{6}A demethylating enzymes showed that mRNA methylation is a reversible process *in vivo*, further indicating a role in gene regulation [25]. Two members of the Fe(II)-dependent and 2-oxoglutarate-dependent oxygenase superfamily, FTO and ALKBH5, exhibit catalytic activity *in vitro* and *in vivo*. FTO (fat mass and obesity-associated protein) oxidizes m^{6}A to A through N^6-hydroxymethyladenosine (hm^{6}A) and N^6-formyladenosine (f^{6}A) intermediates and is sensitive to secondary structure [7••,26]. A recent study showed that FTO regulates adipogenesis by affecting splicing patterns of preadipocyte differentiation markers, suggesting a mechanism by which m^{6}A influences metabolism [27]. ALKBH5 (alkylation repair homologue protein 5) also shows highest activity towards ssRNA *in vitro*, and notable sequence preference for m^{6}A consensus methylation motifs. Knockdown of ALKBH5 affects mRNA processing and export in HeLa cells, though the mechanism of this function remains unclear. Mice deficient in *alkbh5* show increased levels of m^{6}A in isolated organs and display defective spermatogenesis. Furthermore, transcriptome analysis of mRNA from knockout mice testis shows differential expression of over 1500 genes, 127 of which are spermatogenesis-related [8••]. As with m^{6}A methyltransferases, mechanisms of regulation and selectivity for these demethylases are unknown.

m^{6}A effector proteins — 'Readers'

RNA modifications influence protein binding behavior, generating specificity for 'reader' proteins and deterring potential 'anti-reader' proteins. m^{6}A 'reader' proteins of the YTH family were identified by affinity chromatography using methylated RNA as bait [9•]. Mammals encode five YTH proteins: YTH Domain Family (YTHDF) proteins 1, 2 and 3, and YTH Domain Containing (YTHDC) proteins 1 and 2 [28]. To date, four of these proteins have been shown to exhibit m^{6}A selectivity *in vitro* and *in vivo* [29••,30]. A crystal structure of the YTH domain of YTHDF2 as well as that of YTHDC1 bound to methylated RNA have been solved, revealing a conserved hydrophobic binding pocket specific for m^{6}A [30–32]. These m^{6}A 'reader' proteins of the YTH family provide a direct mechanistic link between mRNA methylation and transcript fate.

Recently, two m^{6}A binding proteins have been functionally characterized: YTHDF1 and YTHDF2 [29••,33••]. While all methyltransferase and demethylase components are nuclear, these two proteins reside strictly within the cytoplasm. As expected of m^{6}A 'readers', high-resolution mapping of transcript binding sites reveals a preference for GGACU sequence motifs in mRNA with high overlap with sites of m^{6}A methylation. YTHDF1 enhances translation efficiency of methylated mRNAs by interacting with initiation factors, enhancing ribosome loading of its targets, and promoting translation initiation. YTHDF2 promotes mRNA decay of methylated transcripts by localizing targets to cytoplasmic processing bodies, reducing the half-life of the protein-coding message. Together, these effector proteins are capable of facilitating robust gene expression within a confined time frame and regulate a dynamic set of genes involved in many levels of genetic regulation (Figure 1).

YTHDF3 and YTHDC1 are also known to bind with specificity for the m^{6}A modification in mRNA. YTHDF3 is similar in sequence to YTHDF1 and YTHDF2 and also localized to the cytoplasm but may play unique roles in tissue-specific processes that have yet to be thoroughly explored. Unlike the YTHDF proteins, YTHDC1 is strictly nuclear, and has previously been reported to influence splice-site selection of several mRNAs [28]. YTHDC2 contains a highly conserved YTH domain, but binding specificity and molecular function of this potential 'reader' have yet to be elucidated.

Additional functions of m^{6}A

RNA methylation has been associated with a broad set of biological functions, few of which are currently understood in mechanistic detail. Methylation at the N^6 position of adenosine destabilizes RNA duplexes by 0.5-1.7 kcal/mol, and conversely favors single-stranded conformation when located in unpaired positions due to increased base stacking interactions [34]. RNA secondary structure mapping shows increased icSHAPE reactivity at m^{6}A sites consistent with reduced secondary structure, and may be used to predict sites of methylation [35]. To this end, m^{6}A may regulate biological processes dependent on RNA structure, bypassing the need for a direct 'reader' mechanism. This is indeed apparent for hnRNPC, which binds to U-rich sites adjacent to 'm^{6}A structural switches' and regulates pre-mRNA splicing