



ARTICLE

DOI: 10.1038/s41467-018-05243-z

OPEN

# A potentially abundant junctional RNA motif stabilized by m<sup>6</sup>A and Mg<sup>2+</sup>

Bei Liu<sup>1</sup>, Dawn K. Merriman<sup>2</sup>, Seung H. Choi<sup>1</sup>, Maria A. Schumacher<sup>1</sup>, Raphael Plangger<sup>3</sup>, Christoph Kreutz<sup>3</sup>, Stacy M. Horner <sup>4,5</sup>, Kate D. Meyer<sup>1</sup> & Hashim M. Al-Hashimi <sup>1,2</sup>

*N*<sup>6</sup>-Methyladenosine (m<sup>6</sup>A) is an abundant post-transcriptional RNA modification that influences multiple aspects of gene expression. In addition to recruiting proteins, m<sup>6</sup>A can modulate RNA function by destabilizing base pairing. Here, we show that when neighbored by a 5' bulge, m<sup>6</sup>A stabilizes m<sup>6</sup>A-U base pairs, and global RNA structure by  $\sim 1$  kcal mol<sup>-1</sup>. The bulge most likely provides the flexibility needed to allow optimal stacking between the methyl group and 3' neighbor through a conformation that is stabilized by Mg<sup>2+</sup>. A bias toward this motif can help explain the global impact of methylation on RNA structure in transcriptome-wide studies. While m<sup>6</sup>A embedded in duplex RNA is poorly recognized by the YTH domain reader protein and m<sup>6</sup>A antibodies, both readily recognize m<sup>6</sup>A in this newly identified motif. The results uncover potentially abundant and functional m<sup>6</sup>A motifs that can modulate the epitranscriptomic structure landscape with important implications for the interpretation of transcriptome-wide data.

<sup>1</sup>Department of Biochemistry, Duke University School of Medicine, Durham, NC 27710, USA. <sup>2</sup>Department of Chemistry, Duke University, Durham, NC 27710, USA. <sup>3</sup>Institute of Organic Chemistry and Center for Molecular Biosciences Innsbruck (CMBI), University of Innsbruck, 6020 Innsbruck, Austria. <sup>4</sup>Department of Molecular Genetics and Microbiology, Duke University School of Medicine, Durham, NC 27710, USA. <sup>5</sup>Department of Medicine, Duke University School of Medicine, Durham, NC 27710, USA. Correspondence and requests for materials should be addressed to H.M.A.-H. (email: [hashim.al.hashimi@duke.edu](mailto:hashim.al.hashimi@duke.edu))

**N**<sup>6</sup>-methylated adenosine (m<sup>6</sup>A) is the most abundant internal RNA modification in eukaryotic mRNAs<sup>1–3</sup> and long non-coding RNAs (lncRNA)<sup>4,5</sup> and is also found in viral RNAs<sup>6–8</sup>. The modification is dynamically regulated by methyl transferase<sup>9–11</sup> and demethylase enzymes<sup>12,13</sup> that are linked to human disease<sup>14,15</sup>. It frequently occurs at the 3' and 5' untranslated regions of mRNA where it can influence RNA splicing<sup>16,17</sup>, mRNA nuclear exportation<sup>13,18</sup>, RNA decay<sup>19</sup>, and translation initiation<sup>20,21</sup>. The modification is implicated in a growing number of processes including stem cell fate determination, stress response, DNA damage repair, microRNA biogenesis, lncRNA-mediated transcription repression, viral infection, and in the mechanisms of cancer<sup>22–25</sup>.

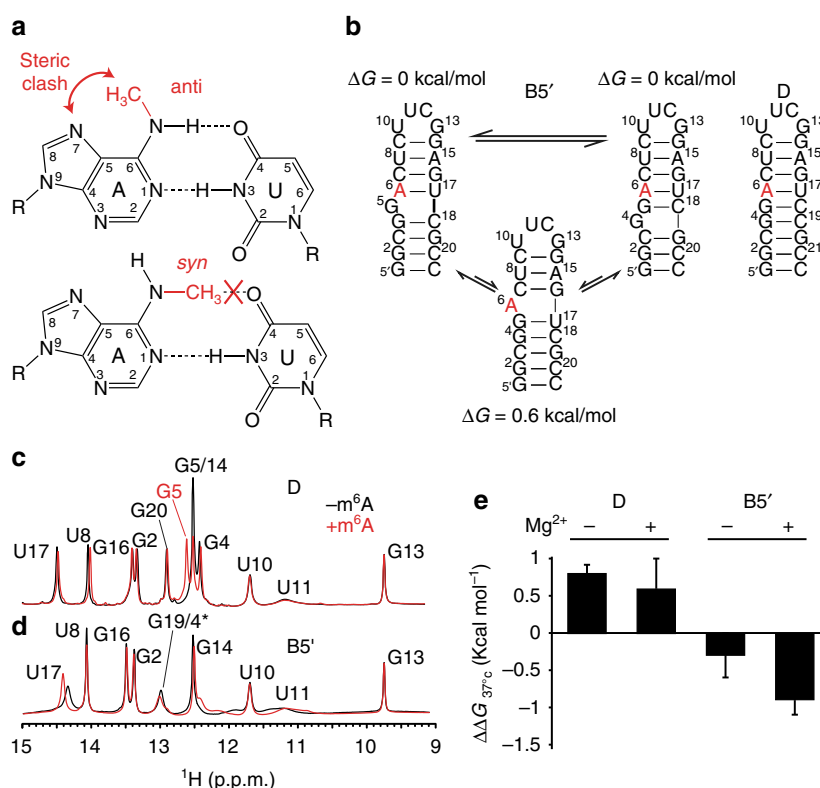
m<sup>6</sup>A is thought to exert its biological effects through two general mechanisms. First, it can help recruit proteins that regulate mRNA fate by influencing splicing<sup>16</sup>, export<sup>26</sup>, decay<sup>19,27</sup>, and translation initiation efficiency<sup>20,27</sup>. Many of these m<sup>6</sup>A reader proteins contain YTH domains, which specifically recognize the methyl group<sup>5</sup>. Second, m<sup>6</sup>A can exert biological effects by modulating RNA structure. In particular, studies have shown that m<sup>6</sup>A destabilizes A–U base pairing and RNA duplexes by 0.5–1.7 kcal mol<sup>−1</sup><sup>28,29</sup>. This destabilization has been proposed to occur due to steric contacts between the N<sup>6</sup>-methyl group and the base<sup>28</sup> (Fig. 1a). Studies have shown that m<sup>6</sup>A can promote RNA melting and thereby enhance binding of single-stranded RNA binding proteins<sup>30–32</sup>. The modification can also disrupt non-canonical A–G base pairs (bps) required for protein binding<sup>33</sup>. Presence of m<sup>6</sup>A in mRNA also impedes tRNA accommodation within the ribosome and translation-elongation most likely by destabilizing A–U bps<sup>34</sup>. Interestingly, while

transcriptome-wide RNA structure mapping data indicate that m<sup>6</sup>A destabilizes RNA structure in vitro and in vivo<sup>28,35</sup>, the destabilization is not observed at the m<sup>6</sup>A nucleotide per se, but rather, at the immediate 5' neighbor which favors single-stranded conformations in methylated RNA<sup>28</sup>. The mechanism by which m<sup>6</sup>A destabilizes its 5'-neighbor is currently unknown.

Fewer studies have examined the potential for m<sup>6</sup>A to stabilize RNA even though there is precedent for this in the literature. Prior studies have shown that when placed at dangling ends of duplexes<sup>28</sup> or in apical loops<sup>29</sup>, m<sup>6</sup>A stabilizes RNA by 0.1–1.2 kcal mol<sup>−1</sup> most likely due to favorable stacking interactions between the methyl group and adjacent bases. Here, by attempting to harness the destabilizing effects of m<sup>6</sup>A in the design of an RNA secondary structural switch, we discovered a bulge motif that is stabilized by m<sup>6</sup>A in a Mg<sup>2+</sup>-dependent manner. This motif is recognized by the YTH domain and m<sup>6</sup>A antibodies and can potentially help explain the global impact of m<sup>6</sup>A on RNA structure observed in transcriptome-wide studies.

## Results

**m<sup>6</sup>A stabilizes a junctional A–U base pair.** Our initial motivation was to examine whether the destabilizing effects of m<sup>6</sup>A on A–U base pairing could be harnessed to induce a change in RNA secondary structure. We designed an RNA hairpin (B5', Fig. 1b) containing the most common<sup>4,5</sup> (GGm<sup>6</sup>ACU) m<sup>6</sup>A consensus sequence (DRACH, where D denotes A, G, or U; R is A or G; and H is A, C or U)<sup>36–38</sup>. This sequence is predicted<sup>39</sup> to fold into two iso-energetic secondary structures in which the m<sup>6</sup>A forms an A–U bp either at the junction next to a bulged guanine or deeper within the upper helix (Fig. 1b). The sequence is also predicted to



**Fig. 1** m<sup>6</sup>A stabilizes junctional A–U base pairs with 5' bulge. **a** The methyl group in m<sup>6</sup>A destabilizes A–U pairing through steric contacts. **b** Design of an m<sup>6</sup>A-dependent RNA structural switch. Secondary structures and free energies computed using the MC-Fold web server<sup>39</sup>. The modified adenine (A6) is highlighted in red. **c, d** Comparison of <sup>1</sup>H NMR spectra of unmodified and m<sup>6</sup>A6 substituted (**c**) D and (**d**) B5' recorded in 15 mM sodium phosphate, 25 mM NaCl, 0.1 mM EDTA and 3 mM Mg<sup>2+</sup> at pH 6.4 and 10 °C. **e** Impact of methylation on the thermal stability of the hairpins with and without 3 mM Mg<sup>2+</sup>. Shown are the differences in the free energy of melting between methylated and unmodified RNA, i.e.,  $\Delta\Delta G = \Delta G(\text{methylated}) - \Delta G(\text{unmodified})$ . The uncertainty reflects the standard deviation from at least three measurements (see Supplementary Table 1)

form an alternative conformation in which the m<sup>6</sup>A is bulged out and in which the 5′ neighboring guanine forms a non-canonical G–U wobble (Fig. 1b). This alternative conformation is predicted to be destabilized by 0.6 kcal mol<sup>-1</sup>, which is within the range of m<sup>6</sup>A destabilization of A–U bps in duplex RNA<sup>28,29</sup>. Thus, we reasoned that m<sup>6</sup>A could destabilize the junctional A–U bp and promote formation of the alternative conformation containing the G–U mismatch, which can be readily detected by 1D <sup>1</sup>H NMR<sup>40</sup>. As a control, we also carried out experiments on a corresponding RNA duplex (denoted D), in which a cytosine nucleotide was included to pair with the bulged guanine (Fig. 1b). Experiments were carried out in the presence of 3 mM Mg<sup>2+</sup>.

The m<sup>6</sup>A substitution resulted in small perturbations in NMR spectra of the control duplex D (Fig. 1c). We observed the U17–H3 imino resonance of the m<sup>6</sup>A6 partner (Fig. 1c), which also has a nuclear Overhauser effect (NOE) cross peak with the amino group of m<sup>6</sup>A6 (Supplementary Fig. 1). This indicates that m<sup>6</sup>A6 forms a m<sup>6</sup>A6–U17 Watson–Crick bp stabilized by two H-bonds, as described previously for other RNA duplexes containing m<sup>6</sup>A<sup>28</sup>. Two NOE cross peaks with comparable intensities were observed between the m<sup>6</sup>A6 methyl proton and the base m<sup>6</sup>A6–H8 and m<sup>6</sup>A6–H2 (Supplementary Fig. 1). This indicates that the amino group either deviates from a perfectly *anti* conformation or rapidly (on the NMR chemical shift timescale) exchanges between the *anti* and *syn* conformations (see Supplementary Note).

The NMR spectra show that unmodified B5′ folds into a conformation with the predicted junctional A6–U17 Watson–Crick bp, bulged G5, and with G4 either forming a weak junctional bp or a bulge (Fig. 1d and Supplementary Fig. 2a, c). The m<sup>6</sup>A6 substitution did not lead to the predicted transition to the alternative conformation, as we did not observe the upfield shifted (10–12 ppm) imino resonances characteristic of G–U wobbles (Fig. 1d and Supplementary Fig. 2a). On the contrary, the substitution sharpened the U17 imino resonance, consistent with stabilization of the junctional Watson–Crick m<sup>6</sup>A6–U17 bp (Fig. 1d and Supplementary Fig. 2a). This together with uninterrupted NOE distance-based connectivity between U17 and G16 indicate that the methylation stabilizes the helical structure (Supplementary Fig. 3b). As in D, the NMR data indicate that while m<sup>6</sup>A6 forms a m<sup>6</sup>A6–U17 Watson–Crick bp, the methyl group deviates from a perfectly *anti* conformation (Supplementary Fig. 1 and see Supplementary Note). Therefore, counter to our predictions, and to the behavior observed in duplexes, m<sup>6</sup>A locally stabilizes the junctional A–U bp in B5′.

### m<sup>6</sup>A globally stabilizes B5′ in a Mg<sup>2+</sup> dependent manner.

While the NMR data indicate that the m<sup>6</sup>A modification locally stabilizes the junctional A–U bp, this provides little information regarding the effect of the modification on overall RNA stability. We therefore used UV melting experiments to examine the impact of the m<sup>6</sup>A6 on RNA stability. Consistent with prior studies<sup>28,29</sup>, m<sup>6</sup>A destabilized D by 0.6 ± 0.4 kcal mol<sup>-1</sup> at 37 °C. In stark contrast, m<sup>6</sup>A stabilized B5′ by 0.9 ± 0.2 kcal mol<sup>-1</sup> at the same temperature (Fig. 1e).

To further examine the basis for m<sup>6</sup>A-dependent stabilization, we carried out UV melting and NMR experiments in the absence of Mg<sup>2+</sup>. The modification destabilized D by a slightly greater amount (~0.8 kcal mol<sup>-1</sup>) in the absence of Mg<sup>2+</sup> (Fig. 1e). This was accompanied by broadening of the imino resonances in the absence of Mg<sup>2+</sup> (Supplementary Fig. 3d), consistent with the destabilization observed using UV. In contrast, in the case of B5′, the degree of m<sup>6</sup>A stabilization decreased significantly from 0.9 ± 0.2 kcal mol<sup>-1</sup> to 0.3 ± 0.3 kcal mol<sup>-1</sup> in the absence of Mg<sup>2+</sup> (Fig. 1e). NMR spectra, including data for site-specifically labeled

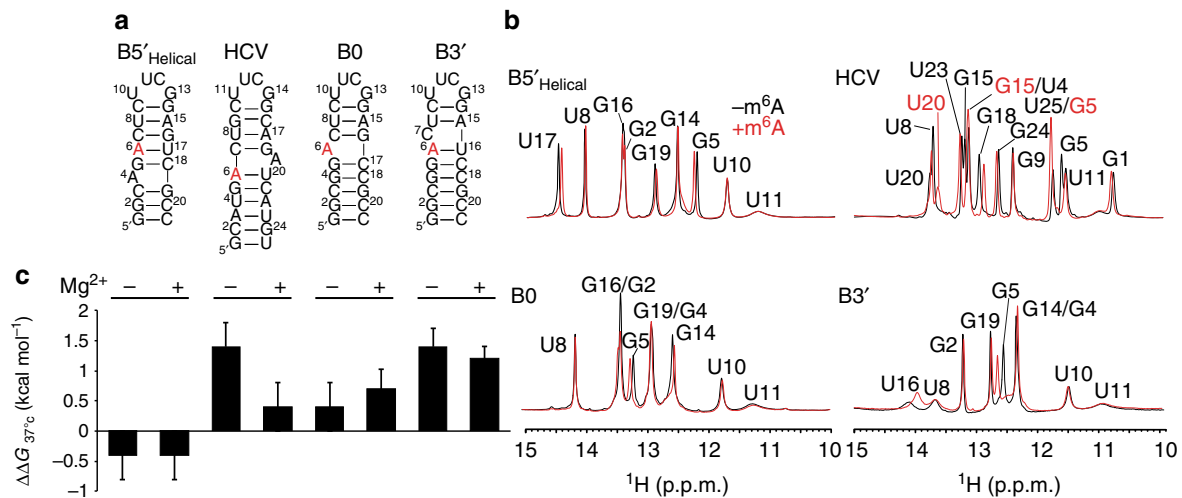
B5′ (Supplementary Fig. 2a), indicate that in the absence of Mg<sup>2+</sup>, unmodified B5′ folds into many distinct and slowly (on the NMR chemical shift timescale) exchanging conformations. The U17 imino resonance in methylated B5′ is significantly broadened, consistent with local melting of the m<sup>6</sup>A6–U17 bp in the absence of Mg<sup>2+</sup> (Supplementary Fig. 2a and Fig. 3d). Here, the NOE data indicate that the methyl group adopts an *anti* conformation (Supplementary Fig. 1) and that G5 forms a G5–U17 mismatch (Supplementary Fig. 2a). Thus, the addition of Mg<sup>2+</sup> induces a conformational switch relative to the structure without Mg<sup>2+</sup> by stabilizing the junctional m<sup>6</sup>A6–U17 bp with a 5′ neighboring G5 bulge.

**Structural requirements for m<sup>6</sup>A stabilization.** Previous studies have shown that when placed at dangling ends, m<sup>6</sup>A stabilizes RNA duplexes by ~0.1–1.2 kcal mol<sup>-1</sup><sup>28,29</sup>. This has been attributed to favorable stacking and hydrophobic shielding of the methyl group. Interestingly, this stabilization is greater when m<sup>6</sup>A is placed at the 5′ (~1.2 kcal mol<sup>-1</sup>) versus 3′ (~0.1–0.8 kcal mol<sup>-1</sup>)<sup>28,29</sup> end of the duplex probably due to stronger stacking with the 3′ neighbor. Based on the thermodynamic parameters obtained from the UV experiments (Supplementary Table 1), the observed m<sup>6</sup>A-mediated stabilization of B5′ is driven by more favorable enthalpy, which is consistent with formation of favorable structural interactions. In particular, the 5′ neighboring guanine bulge could provide m<sup>6</sup>A enough “wobble room” to form m<sup>6</sup>A–U WC bps in which m<sup>6</sup>A optimally stacks with the 3′ neighbor. This may require an unusual backbone conformation at the bulge, which is stabilized by Mg<sup>2+</sup>. Such conformations with optimal stacking may not be easily accommodated within the more rigid duplex interior, resulting in a net destabilization of the duplex.

To test the structural requirements for m<sup>6</sup>A stabilization, we varied the position and structural context of m<sup>6</sup>A (Fig. 2a) using three variants of the B5′ RNA (B5′<sub>helical</sub>, B0, and B3). We also examined a model RNA (HCV) derived from a naturally occurring m<sup>6</sup>A site identified in the 3′ UTR of the hepatitis C virus (HCV) genome<sup>8</sup>. In HCV, the m<sup>6</sup>A site is at the junction but on the strand opposite to the bulge. We performed NMR analysis in the presence of at 25 mM NaCl and 3 mM Mg<sup>2+</sup> and confirmed that for all variants the methylated and unmodified constructs fold into their predicted secondary structures (Fig. 2b and Supplementary Fig. 3c).

Decreasing the conformational freedom available to m<sup>6</sup>A by moving the m<sup>6</sup>A–U bp one base pair deeper into the upper helix (B5′<sub>helical</sub>, Fig. 2a) decreased the m<sup>6</sup>A dependent stabilization from 0.9 ± 0.2 kcal mol<sup>-1</sup> to 0.4 ± 0.4 kcal mol<sup>-1</sup> (Fig. 2c). Decreasing the conformational freedom by placing m<sup>6</sup>A opposite rather than immediately adjacent to the bulge in HCV resulted in a net destabilization by 0.4 ± 0.4 kcal mol<sup>-1</sup> (Fig. 2c). NMR spectra show that in both cases, the modification does not locally stabilize the m<sup>6</sup>A–U bp again possibly due to the lack of conformational freedom needed for optimal stacking (Fig. 2b).

Disrupting stacking with the 3′ neighbor through placement of a bulge 3′ to m<sup>6</sup>A in B3′ also resulted in a net destabilization by ~1.2 kcal mol<sup>-1</sup> (Fig. 2c). Interestingly, although m<sup>6</sup>A globally destabilizes B3′, it did locally stabilize the m<sup>6</sup>A6–U16 bp as evidenced by sharpening of the U16 imino resonance (Fig. 2b). Finally, placement of m<sup>6</sup>A in a bulge position destabilized B0 by ~0.7 ± 0.3 kcal mol<sup>-1</sup> (Fig. 2c). Here, the NMR spectra (Supplementary Fig. 1) indicate that the methyl group adopts a *syn* conformation, consistent with prior studies of single-stranded RNA<sup>41</sup>, and that the modification induces the flipping out of the bulge adenine (Supplementary Fig. 4b). Loss of intra-helical stacking with the adenine bulge could explain the m<sup>6</sup>A-induced



**Fig. 2** Structural requirements for m<sup>6</sup>A stabilization of junctional A-U base pairs. **a** Hairpin constructs that vary the secondary structural context of m<sup>6</sup>A. Secondary structures are predicted using MC-fold<sup>39</sup> and supported by NMR data (see Supplementary Fig. 3c). The HCV sequence is derived from the base of stem loop 1 (SL1) in the 3' end of HCV genome (nt 9633-9642 linked by UUCG to nt 9668-9678) from the genotype 2A JFH1 strain (GeneBank accession: AB047639 [https://www.ncbi.nlm.nih.gov/nuccore/AB047639]). **b** Comparison of <sup>1</sup>H NMR spectra of unmodified and m<sup>6</sup>A substituted hairpin constructs recorded in 15 mM sodium phosphate, 25 mM NaCl, 0.1 mM EDTA and 3 mM Mg<sup>2+</sup> at pH 6.4 and 10 °C. **c** Impact of methylation on the thermal stability of the hairpins with and without 3 mM Mg<sup>2+</sup>. Shown are the differences in the free energy of melting between methylated and unmodified RNA, i.e., ΔΔG = ΔG(methylated) - ΔG(unmodified). The uncertainty reflects the standard deviation from at least three measurements (see Supplementary Table 1)

destabilization of B0. This result shows that the destabilizing effects of m<sup>6</sup>A are not limited to bps in duplexes but can extend to bulges as well.

Similar results were obtained in the absence of Mg<sup>2+</sup> with the exception of HCV for which the destabilization was more significant in the absence (1.4 ± 0.4 kcal mol<sup>-1</sup>) versus presence (0.4 ± 0.4 kcal mol<sup>-1</sup>) of Mg<sup>2+</sup> (Fig. 2c). Together, these studies indicate that the observed m<sup>6</sup>A dependent stabilization requires a specific motif in which an m<sup>6</sup>A-U bp is neighbored by a 5' guanine bulge.

#### Abundance of junctional m<sup>6</sup>A base pairs transcriptome-wide.

Prior transcriptome-wide nuclease mapping studies of methylated RNA from human cells revealed a clear transition in the average RNA structure at the DRACH motif upon methylation<sup>28</sup>. In the methylated RNA, the purine nucleotide immediately 5' to m<sup>6</sup>A had a much higher probability of being single-stranded whereas the methylated adenine had a slightly higher probability of being double-stranded<sup>28</sup>. The two nucleotides immediately 3' to m<sup>6</sup>A showed a strong tendency to be paired. Similar results were observed with *in vitro* and *in vivo* SHAPE experiments<sup>35</sup> although the m<sup>6</sup>A site showed slightly higher reactivity, indicating a greater tendency to be unpaired. These results indicate that methylation thermodynamically biases m<sup>6</sup>A in cellular mRNAs to be located between single-stranded unpaired RNA and adjacent to helices<sup>28</sup>.

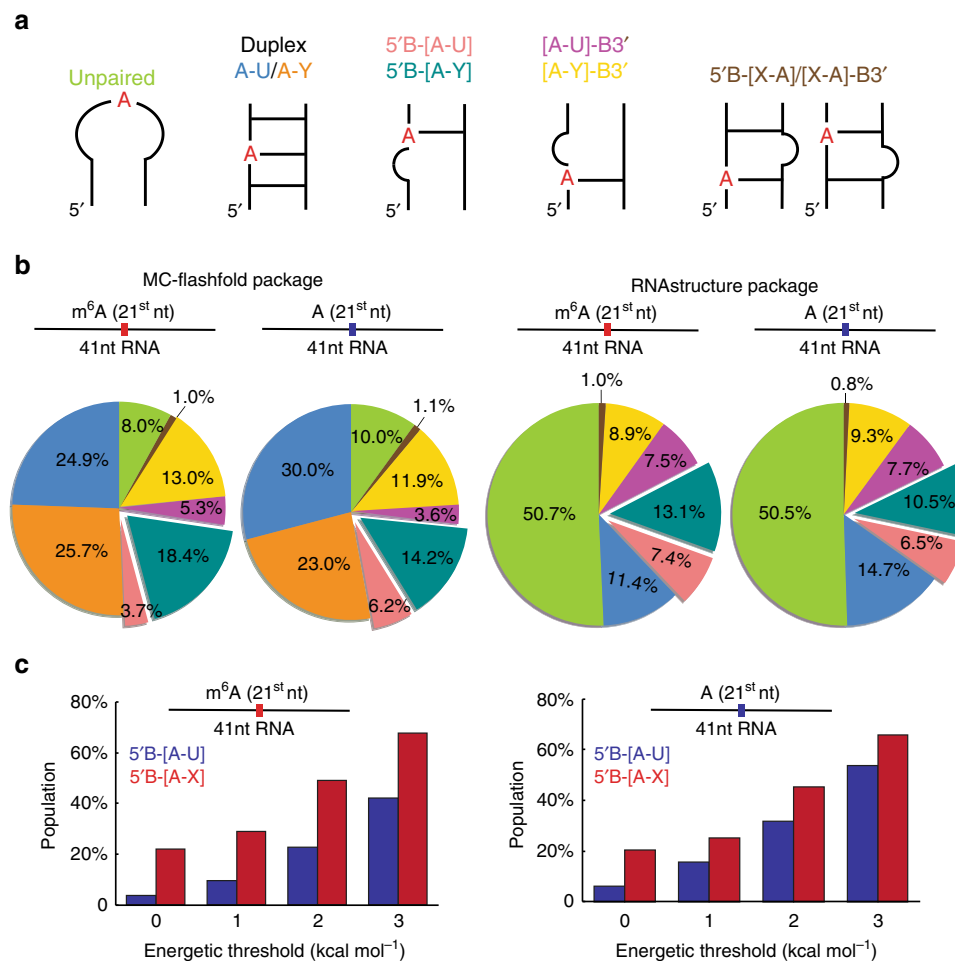
Remarkably, the m<sup>6</sup>A stabilized B5' motif identified in this work captures the unique conformational signatures induced by m<sup>6</sup>A in transcriptome-wide studies. The motif partially protects m<sup>6</sup>A by positioning it at a junction, exposes the 5' guanine bulge, while the 3' neighbors are helical. The abundance of the 5'B-[m<sup>6</sup>A-X] (where B is any bulge nucleotide and X is any nucleotide) motif could increase significantly on methylation because of two combined effects. On the one hand, m<sup>6</sup>A destabilizes bps in duplexes as well as bulges by as much as ~1.7 kcal mol<sup>-1</sup>. On the other hand, it can stabilize the 5'B-[m<sup>6</sup>A-U] motif by ~1 kcal mol<sup>-1</sup>. The combined contribution of ~3 kcal mol<sup>-1</sup> could significantly bias the RNA folding landscape

away from structures in which m<sup>6</sup>A is embedded within duplexes and bulges toward 5'B-[m<sup>6</sup>A-U] and perhaps other 5'B-[m<sup>6</sup>A-X] motifs and to a degree that help account for the observed changes in transcriptome-wide structure-mapping data. Such a strong bias is required to explain the changes in RNA conformation induced by methylation especially since m<sup>6</sup>A occurs sub-stoichiometrically in most RNAs<sup>5,42</sup>.

To test this hypothesis, we asked what fraction of the 140,574 m<sup>6</sup>A sites that have been mapped throughout the human transcriptome<sup>43</sup> are predicted to fold into the 5'B-[m<sup>6</sup>A-U] or related 5'B-[m<sup>6</sup>A-X] motifs as the energetically most favorable secondary structure. We then asked what additional sites are predicted to fold into the 5'B-[m<sup>6</sup>A-U] or 5'B-[m<sup>6</sup>A-X] motifs as higher energy conformations that are within the ~3 kcal mol<sup>-1</sup> energetic threshold. These conformations could be remodeled by m<sup>6</sup>A such to adopt the 5'B-[m<sup>6</sup>A-U] or 5'B-[m<sup>6</sup>A-X] motifs as the most energetically stable conformation. In particular, we used MC-Flashfold<sup>39</sup> to predict secondary structures for 41 nt fragments chosen such that the m<sup>6</sup>A residue was positioned in the middle. Predicted RNA structures were classified according to the secondary structural context and position of m<sup>6</sup>A (Fig. 3a). As a control, secondary structures were also predicted for 140,574 randomly selected RNA sequences from the same transcriptome that do not contain the m<sup>6</sup>A consensus sequence.

Approximately 4 and 22% of the m<sup>6</sup>A sites are predicted to fold into the 5'B-[m<sup>6</sup>A-U] or 5'B-[m<sup>6</sup>A-X] motif respectively as the most energetically favorable secondary structure (Fig. 3b). Similar abundances were obtained for the control unmodified sequences (Fig. 3b), indicating that sequences selected for methylation do not have an enhanced propensity to fold into the 5'B-[A-X] motif relative to random sequences. Similar results were also obtained when using a different RNA structure prediction program (RNAstructure<sup>44</sup>) and when varying the position of m<sup>6</sup>A or length of the sequences subjected to structure prediction (Fig. 3b and Supplementary Fig. 5a).

A much larger fraction of m<sup>6</sup>A sequences (~42 and ~68% for 5'B-[m<sup>6</sup>A-U] and 5'B-[m<sup>6</sup>A-X], respectively) are predicted to fold into the junctional motif with energies < 3 kcal mol<sup>-1</sup> relative



**Fig. 3** Predicted secondary structures of m<sup>6</sup>A containing RNA sequences. **a** Predicted RNA secondary structures were classified into the following motifs: Junctional Watson-Crick m<sup>6</sup>A-U or mismatch m<sup>6</sup>A-Y (Y denotes A or C or G.) with the bulge located 5' (5'B-[A-U] and 5'B-[A-Y]) or 3' ([A-U]-B3' and [A-Y]-B3') to m<sup>6</sup>A; or with the bulge located 5' or 3' (5'B-[X-A]/[X-A]-B3', X denotes A or C or U or G) to m<sup>6</sup>A partner nucleotide. **b** Distribution of the lowest energy predicted structures for 140,574 m<sup>6</sup>A sites identified using individual-nucleotide-resolution cross-linking and immunoprecipitation (miCLIP)<sup>43</sup> in the human transcriptome (hg19) were predicted using RNAstructure<sup>44</sup> and MC-Flashfold<sup>39</sup>. Secondary structures were predicted for RNA sequences of 41-nt long with m<sup>6</sup>A (or A) in the middle. Also shown are corresponding predictions for 140,574 unmodified adenine sites selected randomly from the same human transcriptome. **c** Population of m<sup>6</sup>A and A sites predicted (MC-Flashfold) to fold into the 5'B-[A-U] and 5'B-[A-X] conformation within different energetic thresholds relative to the lowest energy structure

to the lowest energy structure (Fig. 3c). Similar results were obtained when predicting structures using RNAstructure<sup>44</sup> (Supplementary Fig. 5b). These higher energy structures could become the most stable structures upon methylation. While a similar increase in abundance is observed for the control unmodified sequences (Fig. 3c), indicating that the potential to form 5'B-[A-X] motif is not sequence dependent, these sequences are less likely to be methylated and to experience the energetic bias. Consequently only 20% of random A sites are expected to fold into 5'B-[A-X] motif as the minimum free energy structure. Therefore, while further experiments are needed to assess how m<sup>6</sup>A stabilizes / destabilizes various motifs in different sequence contexts, these results indicate that m<sup>6</sup>A could strongly bias RNA folding toward 5'B-[m<sup>6</sup>A-X] motifs, and thus help explain the m<sup>6</sup>A induced RNA conformational changes observed in transcriptome-wide structure mapping data.

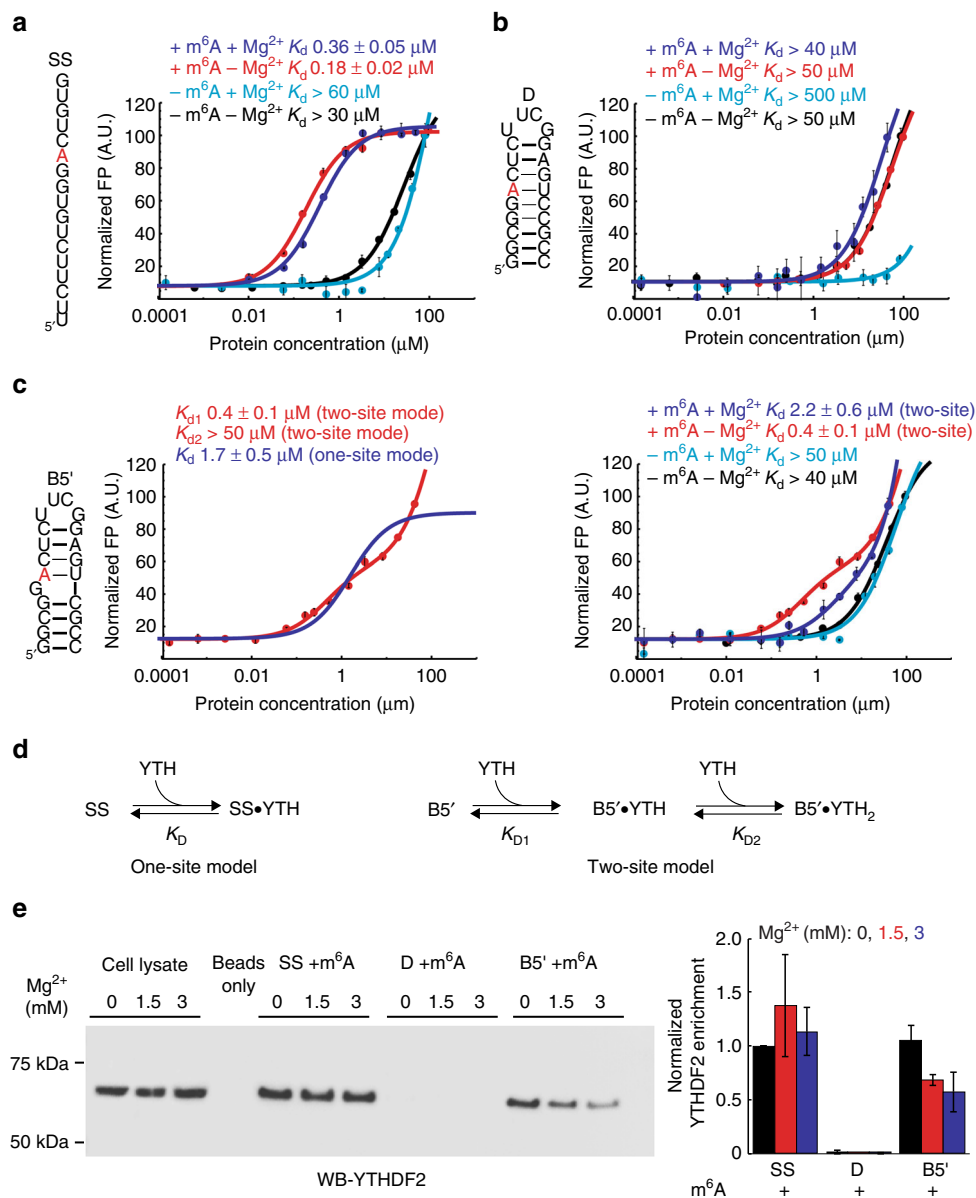
**m<sup>6</sup>A is recognized in 5'G-[m<sup>6</sup>A-U] but not in duplexes.** To examine the functional significance of the 5'G-[m<sup>6</sup>A-U] motif, we examined whether it can be recognized by the YTH domain m<sup>6</sup>A reader (residues 380-579) from the human YTHDF2

(NP\_057342.2) protein<sup>45</sup>. Binding was measured using a fluorescence polarization (FP) assay employing fluorescein tagged RNA<sup>46</sup>. Experiments were initially carried out in the absence of Mg<sup>2+</sup> to allow comparison with previous studies<sup>45,47-50</sup>.

As a positive control, the YTH domain binds to single-stranded RNA (SS) containing m<sup>6</sup>A and the consensus sequence (GGm<sup>6</sup>ACU) with  $K_D = 0.18 \pm 0.02 \mu\text{M}$  (Fig. 4a) in good agreement with previously reported values (0.2-2  $\mu\text{M}$ )<sup>45,47-49</sup>. The binding affinity is diminished at least ~100-fold ( $K_D > 30 \mu\text{M}$ ) for unmodified SS (Fig. 4a), in good agreement with 10-100-fold weaker affinity reported previously<sup>45,47-49</sup>.

Surprisingly, the YTH domain binds weakly ( $K_D > 50 \mu\text{M}$ ) to both unmodified and methylated duplex RNA (D) (Fig. 4b). In both cases, the affinity is comparable to that measured for unmodified SS (Fig. 4b). This indicates that m<sup>6</sup>A sites in helical m<sup>6</sup>A-U Watson-Crick bps are not accessible for recognition by the YTH domain. This result is significant considering that helical A-U Watson-Crick bps occur frequently in the structures of RNA.

In contrast, the YTH domain does bind tightly to methylated B5' (but not its unmodified counterpart;  $K_D > 40 \mu\text{M}$ ) (Fig. 4c). The binding curve deviates from a simple two-state model and

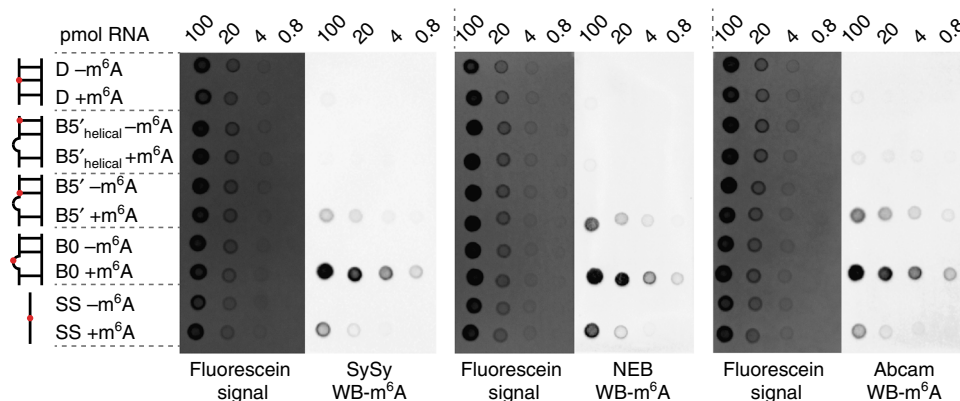


**Fig. 4** YTH recognizes m<sup>6</sup>A in B5' but not D. **a–c** Binding curves for YTH domain with **a** SS, **b** D, and **c** B5'. **d** One-site and independent two-site binding models used for curve fitting. Data is shown for unmodified and methylated RNA with and without 3 mM Mg<sup>2+</sup>. Uncertainty reflects the standard deviation from three measurements. (**c**, left) Fits to the data assuming single (blue) versus two independent (red) binding sites (see Methods). **e** In vitro RNA pulldown assay using 5'biotinylated RNA constructs with m<sup>6</sup>A. Showing YTHDF2 binds preferentially to m<sup>6</sup>A in unpaired context or at the junction, and Mg<sup>2+</sup> decreases the enrichment of YTHDF2 by methylated B5'. Blots shown are representative of results from three experiments. Uncropped blots are shown in Supplementary Fig. 6. The semi-quantitative YTHDF2 enrichment analysis was performed using ImageJ<sup>56</sup>. The uncertainty represents the standard deviation from three experiments

the data is better fit to a model that assumes two independent binding sites (Fig. 4d). A similar behavior was previously observed for YTH domain binding to methylated ssRNA<sup>45</sup>. One of two  $K_D$ s obtained from this analysis ( $K_{D1} = 0.3 \pm 0.1 \mu\text{M}$ ) is consistent with high affinity binding as observed for methylated SS while the other ( $K_{D2} > 50 \mu\text{M}$ ) is consistent with weak binding as observed with unmodified RNA. It is possible that the YTH domain specifically binds to m<sup>6</sup>A in B5' with high affinity but also weakly and non-specifically binds to other parts of the RNA (e.g. duplex or apical loop region). Alternatively, the YTH domain could bind to m<sup>6</sup>A in two different RNA conformations (e.g. helical versus more single-stranded). These results indicate that in contrast to duplexes, m<sup>6</sup>A in 5'B-[m<sup>6</sup>A-U] is recognized by the YTH domain.

Based on the NMR data, Mg<sup>2+</sup> stabilizes the junctional m<sup>6</sup>A-U bp in B5' potentially making m<sup>6</sup>A6 less accessible for recognition. Indeed, the binding affinity to methylated B5' decreased by ~6-fold ( $K_{D1} 0.4 \pm 0.1 \mu\text{M}$  versus  $2.2 \pm 0.6 \mu\text{M}$ , Fig. 4c) in the presence of 3 mM Mg<sup>2+</sup> but remained ~10 fold tighter than that of unmodified SS or B5' and ~20-fold tighter than methylated D. By comparison, the binding affinity of YTH domain to methylated or unmodified SS and D as well as unmodified B5' was only slightly (by ~1-fold) weakened in the presence of Mg<sup>2+</sup> (Fig. 4a, b).

We confirmed the above results for the full-length YTHDF2 protein using in vitro RNA pulldown experiments (Fig. 4e). Here, biotinylated RNA baits were incubated with HEK293T cell lysates, and RNA:protein complexes were purified with



**Fig. 5**  $m^6A$  antibodies recognize  $m^6A$  in B5' but not D. Dot blot assays assessing the antibody specificity for  $m^6A$  in different secondary structure contexts. Increasing amounts of 5'fluorescein labeled RNAs with and without  $m^6A$  were spotted onto the membrane. The quantity of RNA was visualized by the fluorescein signal. Three different  $m^6A$  antibodies (SySy polyclonal, NEB monoclonal, Abcam polyclonal) were tested. Red dots in the schematic secondary structures represent the  $m^6A$  sites. The SS construct used in dot blot assays is 5'-GGCGGm<sup>6</sup>ACUC-3'. Blots shown are representative of results from three experiments

streptavidin agarose. The bound proteins were then eluted and subjected to SDS-PAGE, followed by western blot to detect YTHDF2. In agreement with results from the FP assay, in the absence of  $Mg^{2+}$ , modified B5' enriches YTHDF2 protein by an amount comparable to modified SS. While similar enrichments were observed in the absence and presence of  $Mg^{2+}$  for SS, the enrichment with B5' decreased 1-fold in the presence of  $Mg^{2+}$ . In sharp contrast, YTHDF2 was barely enriched by methylated D in the absence and presence of  $Mg^{2+}$ .

Similar results were also obtained when examining how well  $m^6A$  is recognized by three  $m^6A$  antibodies (Abcam polyclonal antibody, SySy polyclonal antibody and NEB monoclonal antibody) that are used in transcriptome-wide studies. As expected, all unmodified RNA constructs were not well recognized by the antibodies in dot blot assays (Fig. 5). Both methylated SS and B0 exhibited strong signals, suggesting that the antibodies can recognize unpaired  $m^6A$  (Fig. 5 and Supplementary Fig. 5c). Once again, the methylated D and methylated B5'<sub>helical</sub> were poorly recognized by all three antibodies. This was also the case for a 9-mer duplex lacking the stabilizing UUCG apical loop (Supplementary Fig. 5c). In contrast, the junctional  $m^6A$ -U bp in B5' was well recognized by all three antibodies.

Taken together, these results indicate the  $m^6A$  within the 5'G-[ $m^6A$ -U] motif is well recognized by YTH reader proteins as well as commonly used  $m^6A$  antibodies, whereas  $m^6A$ -U bps embedded in duplexes are not.

## Discussion

A growing number of studies indicate that  $m^6A$  can exert biological effects by modulating RNA structure. So far, studies have primarily focused on the destabilization of base pairing due to steric collisions with the methyl group<sup>28,29,31,33</sup>. However,  $m^6A$  has also been shown to stabilize duplexes when placed at dangling ends and within apical loops<sup>28,29</sup>. Our results identify a 5' bulge motif that is predicted to be abundant and stabilized by  $m^6A$  in a  $Mg^{2+}$  dependent manner. The 5' bulge most likely provides  $m^6A$  the conformational freedom needed to allow the methyl group to optimally stack with the 3' neighbor while maintaining Watson-Crick  $m^6A$ -U pairing.  $Mg^{2+}$  may help to stabilize the resulting backbone conformation at the bulge by neutralizing electrostatic repulsion. In this regard, it is noteworthy that  $Mg^{2+}$  preferentially binds to guanine and adenine nucleotides<sup>51</sup>, thus the naturally occurring DRACH motif may enhance the propensity for the 5' neighbor to bulge out. This motif with base-paired  $m^6A$  stacking

on its 3' neighbor and with a 5'-neighboring bulged nucleotide can potentially form in a variety of contexts, including bulges, apical loops, and possibly other higher order structures. Further studies are needed to assess the robustness of the observed stabilization across these different structural and sequence contexts.

While biophysical studies show that  $m^6A$  destabilizes pairing of adenine nucleotides in vitro, these results were at odds with in vitro transcriptome-wide nuclease mapping<sup>28</sup> and SHAPE<sup>35</sup> data showing that methylation primarily acts to increase the single-stranded character of the purine nucleotide immediately 5' to the  $m^6A$  site, while the adenine nucleotide itself was less affected<sup>28</sup>. These data were obtained devoid of cellular proteins and the conformational differences likely originate directly from the impact of  $m^6A$  on RNA structure. A bias toward the junctional  $m^6A$  bps flanked by a 5' bulge following methylation could help explain the increase in single-stranded character observed specifically for the purine nucleotide 5' to the  $m^6A$  bulge as well as the increase in the double-stranded character observed for the  $m^6A$  nucleotide using transcriptome-wide nuclease mapping experiments<sup>28</sup>. This bias could be significant especially when considering that  $m^6A$  selectively stabilizes the 5'B-[A-U] motif but destabilizes other common motifs such as duplexes and bulges. It should be noted, however, that in vivo SHAPE data<sup>35</sup> suggest that the impact of methylation on RNA structure can be more complex, possibly due to interactions with proteins.

Our results show that  $m^6A$  is effectively recognized by both YTH domain reader proteins and  $m^6A$  antibodies when located in the 5'G-[A-U] motif. The motif could therefore also serve functional roles recruiting YTH domains and is likely to be identified in current methods used to map  $m^6A$  transcriptome-wide. In contrast, our results indicate that both YTH domain reader proteins and  $m^6A$  antibodies do not recognize  $m^6A$ -U bps embedded within duplexes. Current methods for  $m^6A$  mapping attempt to remove these structural biases through nuclease digestion and heating of the RNA prior to immunoprecipitation. However, the RNA fragments are still 30–130 nt long<sup>43</sup> and are subjected to cooling on ice following heating, allowing for re-annealing of the RNA. Indeed, we found that  $m^6A$  sites buried in A-U bps within duplexes were poorly recognized by the antibodies even when increasing the temperature and/or duration of heating with or without 8 M urea. This raises the possibility that there are  $m^6A$  sites buried within duplexes that have evaded detection using current transcriptome-wide methods. This could potentially confound the interpretation of  $m^6A$  induced conformational changes based on transcriptome-wide structure

mapping data since motifs in which m<sup>6</sup>A is accessible such as 5'G-[A-U] could be over-represented compared with those in which m<sup>6</sup>A is not accessible. Additional studies are needed to examine whether these structure-specific biases in m<sup>6</sup>A recognition also occur during immunoprecipitation experiments with free-floating RNA used in the transcriptome-wide mapping of m<sup>6</sup>A sites.

Our results indicate that while m<sup>6</sup>A destabilizes duplex A-U bps and disrupts base stacking in certain motifs, the structure landscape of the epitranscriptome is also likely punctuated by motifs such as 5'G-[m<sup>6</sup>A-U] that are stabilized by m<sup>6</sup>A. This motif presents a readable m<sup>6</sup>A site in a structured region of RNA. While the biological significance of this motif, including whether it is recognized by specific RNA-binding proteins, requires further investigation, our findings expand the mechanisms by which m<sup>6</sup>A can modulate RNA structure and should facilitate the interpretation of transcriptome-wide data.

## Methods

**Sample preparation.** Modified (N<sup>6</sup>-methylated adenosine, 5'-biotinylated RNA and 5'-fluorescein labeled RNA) and unmodified RNA oligonucleotides were synthesized using a MerMade 6 Oligo Synthesizer employing 2'-tBDSilyl protected phosphoramidite (ChemGenes) on 1 μmol standard synthesis columns (1000 Å) (BioAutomation). n-acetyl protected rC, rA, and rG phosphoramidites were used to avoid incomplete deprotection when using isobutryl protected rG and benzoyl protected rA phosphoramidites. m<sup>6</sup>A phosphoramidite was purchased from ChemGenes, biotin phosphoramidites, and fluorescein phosphoramidites were purchased from Glen Research. The <sup>15</sup>N1-labeled guanosine and the <sup>15</sup>N3-labeled uridine phosphoramidites were synthesized according to a published procedure<sup>52</sup>. ssRNA oligonucleotides were synthesized with the option to leave the final 5'-protecting group (4,4'-dimethoxytrityl (DMT)) on for 2'O deprotection and cartridge purification. UUGC-capped hairpin RNAs were synthesized with DMT group off for DMT-off 2'O deprotection and PAGE purification, since regular 2'O deprotection method might cause incomplete deprotection for UUGC-capped samples. Synthesized oligonucleotides were cleaved from the 1 μmol column using 1 mL ammonia methylamine (1:1 ratio of 30% ammonium hydroxide and 30% methylamine) followed by 2-hour incubation at room temperature to allow base deprotection. The solution was then air-dried and dissolved in 115 μL DMSO, 60 μL TEA, and 75 μL TEA-3HF for regular 2'O deprotection, or in 100 μL DMSO and 125 μL TEA-3HF for DMT-off 2'O deprotection, followed by 2.5 h incubation at 65 °C. Regular 2'O deprotected samples were then quenched with Glen-Pak RNA quenching buffer and loaded onto Glen-Pak RNA cartridges (Glen Research Corporation) for purification using the online protocol (<http://www.glenresearch.com/>). Samples were then ethanol precipitated, air dried, dissolved in water and buffer exchanged or diluted into the desired buffer (15 mM sodium phosphate, 25 mM NaCl, 0.1 mM EDTA, 10% D<sub>2</sub>O, pH 6.4 with or without 3 mM Mg<sup>2+</sup>). DMT-off 2'O deprotected samples were directly ethanol precipitated, and purified using 20% (w/v) denaturing PAGE, and electroeluted into 20 mM Tris buffer, pH 8, and subsequently ethanol precipitated. The samples were then dissolved in water (50 μM for hairpin and 200–500 μM for duplex), annealed by heating at 95 °C for 10 min, and cooled on ice for 30 min (for hairpin) or at room temperature for 2 h (for duplex). RNA samples were buffer exchanged at least three times using a centrifugal concentrator (EMD Millipore) into the desired buffer. The purity of methylated and unmodified oligonucleotides was verified using NMR to monitor resonances from deprotecting groups and for B5' using liquid chromatography-mass spectrometry LC/MS (Novatia).

**NMR spectroscopy.** All NMR experiments were collected on a 600 MHz Bruker NMR spectrometer equipped with an HCN cryogenic probe. Data were processed and analyzed using NMRpipe<sup>53</sup> and SPARKY (T.D. Goddard and D.G. Kneller, SPARKY 3, University of California, San Francisco), respectively. Resonances were assigned using 2D HSQC, HMQC, <sup>1</sup>H-<sup>1</sup>H NOESY (mixing times of 100 ms and 150 ms), and HCN experiments in the absence of Mg<sup>2+</sup>.

**UV melting.** Thermal melting experiments were conducted on a PerkinElmer Lambda 25 UV/VIS spectrometer with a RTP 6 Peltier Temperature Programmer and a PCB 1500 Water Peltier System. All RNA samples were buffer exchanged at least three times with a centrifugal concentrator (EMD Millipore) to desired buffers (15 mM sodium phosphate, 25 mM NaCl, 0.1 mM EDTA, pH 6.4 with or without 3 mM Mg<sup>2+</sup>), followed by direct dilution to 3 μM with the same buffer. At least three measurements were carried out for each RNA with a sample volume of 400 μL in a Teflon-stoppered 1 cm path length quartz cell. The absorbance at 260 nm was monitored while the temperature was varied between 15 and 95 °C.

Thermodynamic parameters from UV melting experiments were fitted using nonlinear model fitting in Mathematica 10.0 (Wolfram Research) such that melting

temperature ( $T_m$ ) and enthalpy ( $\Delta H$ ) for duplex and hairpin association were obtained by the fitting to Eqs. (1) and (2), respectively<sup>54</sup>,

$$f_{\text{duplex}} = \frac{1 + 4e^{\frac{(\frac{1}{T_m} - \frac{1}{T_m^0})\Delta H}{R}} - \sqrt{1 + 8e^{\frac{(\frac{1}{T_m} - \frac{1}{T_m^0})\Delta H}{R}}}}{4e^{\frac{(\frac{1}{T_m} - \frac{1}{T_m^0})\Delta H}{R}}} \quad (1)$$

$$f_{\text{hairpin}} = \frac{e^{\frac{(\frac{1}{T_m} - \frac{1}{T_m^0})\Delta H}{R}}}{1 + e^{\frac{(\frac{1}{T_m} - \frac{1}{T_m^0})\Delta H}{R}}} \quad (2)$$

where  $T$  is the temperature (K),  $R$  is the gas constant (kcal mol<sup>-1</sup>), and  $f$  is the fraction of folded duplex or hairpin for Eqs. (1) and (2), respectively.  $\Delta G$  and  $\Delta S$  were calculated from Eqs. (3) and (4) for duplex association and hairpin folding respectively.

$$\Delta S = \frac{\Delta H}{T_m} - R \ln\left(\frac{C_T}{2}\right); \Delta G = \Delta H - T\Delta S \quad (3)$$

$$\Delta S = \frac{\Delta H}{T_m}; \Delta G = \Delta H - T\Delta S \quad (4)$$

where  $C_T$  is the total concentration of RNA. The uncertainty in  $T_m$ ,  $\Delta H$ ,  $\Delta G$ , and  $\Delta S$  was obtained based on the standard deviation in triplicate measurements. The destabilization or stabilization effects of m<sup>6</sup>A was calculated using the following equations:

$$\Delta\Delta G = \Delta G_{m^6A} - \Delta G_{\text{unmodified}} \quad (5)$$

$$\Delta\Delta H = \Delta H_{m^6A} - \Delta H_{\text{unmodified}} \quad (6)$$

$$\Delta\Delta S = \Delta S_{m^6A} - \Delta S_{\text{unmodified}} \quad (7)$$

**Secondary structure prediction.** The programs MC-Flashfold package<sup>39</sup> and RNAstructure package<sup>44</sup> were used to predict the secondary structure of 140,574 sequences containing m<sup>6</sup>A that were identified using single-nucleotide transcriptome-wide m<sup>6</sup>A mapping of human genome build (hg19)<sup>43</sup>. Another 140,574 control sequences that do not contain the DRACH motif and that were randomly selected from the same transcriptome were also analyzed. In-house Python scripts were used to perform the analysis. The sequences were 41 nt long with m<sup>6</sup>A positioned in the middle. Analysis was repeated when varying the position and length of the m<sup>6</sup>A site (31-nt long sequences with m<sup>6</sup>A site at the eleventh or twentyfirst position). The lowest energy predicted structures were classified into different categories: duplex, unpaired, 5'B-[A-U], 5'B-[A-Y] (Y denotes A or C or G), [A-U]-B3', [A-Y]-B3' and 5'B-[X-A]/[X-A]-B3' (X denotes A or C or G or U). 5'-B-[A-U] and 5'B-[A-Y] are motifs containing junctional A-U or A-Y bp with unpaired residues at 5' side of A. Similarly, [A-U]-B3', [A-Y]-B3' are motifs containing junctional A-U or A-Y bps with unpaired residues at 3' side of A. 5'B-[X-A]/[X-A]-B3' is the motif with junctional A-X bps that are on the opposite strand of unpaired residues. With the exception of G-U bps, the RNAstructure package does not predict mismatches. To calculate the population of 5'B-[A-Y] and [A-Y]-B3', any A-Y site that was immediately adjacent to a canonical Watson-Crick bp was counted as a mismatch. The MC-Flashfold package does predict mismatches allowing further classification of the duplex category into duplex (A-U) and duplex (A-Y). MC-Flashfold was used to predict the secondary structures setting the maximum energy difference of output structures to minimum free energy (MFE) of 0, 1, 2 or 3 kcal mol<sup>-1</sup>. For RNAstructure package, the maximum percent energy difference between predicted structures and MFE was set to 10, 20, or 30%. The population of 5'B-[A-U] or 5'B-[A-X] was determined to be the number of sequences predicted to form 5'B-[A-U] or 5'B-[A-X] in the output structures divided by the total number of sequences (140,574).

**Protein expression and purification.** A gene, codon optimized (Supplementary Table 2) for *Escherichia coli* expression, encoding the YTH domain (residues 380–579) of human YTHDF2 (NP\_057342.2) was subcloned into the pET15b vector (the plasmid containing the codon optimized sequence was purchased from GenScript). The plasmid was transformed into *E. coli* strain C41(DE3). The cells were grown in LB medium containing 50 μg ml<sup>-1</sup> ampicillin at 37 °C to an OD<sub>600</sub> between 0.4 and 0.6. Recombinant proteins were induced by adding isopropyl-β-D-thiogalactopyranoside (IPTG) to 0.5 mM and overexpressed in LB medium at 20 °C overnight. Cells expressing YTH domain were harvested at 4 °C and stored at -80 °C or lysed immediately using a microfluidizer in the buffer containing 20 mM HEPES (pH 7.4), 200 mM NaCl, 1 mM DTT supplemented with protease inhibitors and DNase. Following centrifugation at 17,500 rpm for 3 min at 4 °C, the supernatant was loaded to Ni-NTA affinity column. The eluted protein was further purified by size exchange chromatography (SEC) using a Superdex 75 pg column. The peak fractions were collected and concentrated. The protein purity was



assessed using SDS-PAGE analysis and protein concentration determined using Bradford assays.

**YTH domain binding assays.** Binding experiments were carried out using a previously reported<sup>46</sup> fluorescence polarization (FP) assay at 25 °C and using a PanVera Beacon 2000 instrument (Invitrogen, Madison, WI, USA). 5'-fluorescein-labeled RNAs were synthesized and purified as described in the sample preparation section. Fluorescein labeled samples were dissolved in water and then directly diluted into the desired buffer (20 mM HEPES 50 mM NaCl, 3 mM DTT, pH 8.2, with or without 3 mM Mg<sup>2+</sup>). YTH domain protein was serially diluted into 200 μL of the same binding buffer containing 2 nM 5'-fluorescein-labeled RNA. Fluorescence polarization was measured at an excitation wavelength of 490 nm and an emission wavelength at 530 nm. The total fluorescent intensities did not vary significantly throughout the measurements (i.e.,  $(I_{\max} - I_{\min})/I_{\min} \sim 5\%$ , in which  $I_{\max}$  and  $I_{\min}$  are the maximum and minimum intensities, respectively). The binding curves were fitted to either one-site (Eq. (8)) or two-site<sup>55</sup> (Eq. (9)) binding mode using Mathematica 10.0 (Wolfram Research).

$$A_t = \frac{A - (A - B) \times (R_t + L_t + K_D) - \sqrt{(R_t + L_t + K_D)^2 - 4 \times R_t \times L_t}}{2 \times R_t} \quad (8)$$

$$A_t = \frac{A + \frac{B \times L}{K_{D1}} + \frac{C \times L^2}{K_{D1} \times K_{D2}}}{1 + \frac{L}{K_{D1}} + \frac{L^2}{K_{D1} \times K_{D2}}} \quad (9)$$

$L_t$  is the total concentration of protein,  $R_t$  is the concentration of RNA.  $A$  and  $B$  represent free RNA anisotropy and RNA-protein intrinsic anisotropy of the saturated protein-RNA complex.  $K_D$  is the dissociation constant. In Eq. (9),  $K_{D1}$  and  $K_{D2}$  are dissociation constants of the two binding events respectively.  $A$ ,  $B$ , and  $C$  represent free RNA anisotropy, RNA-Protein and RNA-[Protein]<sub>2</sub> intrinsic anisotropy of the saturated protein-RNA complexes. By titrating excess protein against 2 nM RNA,  $L_{\text{free}}$  remains in excess over concentrations of RNA-Protein and RNA-[Protein]<sub>2</sub> complexes.  $L_{\text{free}}$  is therefore well approximated by  $L_t$  and is referred to simply as  $L$ .

**In vitro RNA pulldown.** Biotinylated RNA samples were synthesized and purified as described in the sample preparation section. HEK293T cells were harvested at ~70–80% confluency, washed with cold PBS and lysed by dounce homogenization in lysis buffer (10 mM NaCl, 2 mM EDTA, 0.5% Triton X-100, 0.5 mM DTT, 10 mM Tris, pH 7.4), and the lysate was then brought to 150 mM KCl and 5% Glycerol (v/v). Mammalian protease inhibitor cocktail (Sigma-Aldrich) and phosphatase inhibitor cocktail (Sigma-Aldrich) were freshly added to the lysis buffer. The following steps were performed as described previously<sup>5</sup>, except that binding buffers containing different Mg<sup>2+</sup> concentrations were used (10 mM Tris pH 7.5, 150 mM KCl, 0.5 mM DTT, 0.05% (v/v) NP-40 and 0 mM, 1.5 mM, or 3 mM MgCl<sub>2</sub>). Proteins were eluted under mild conditions with elution buffer (50 mM Tris, pH 7.5, 200 mM NaCl, 2% SDS (w/v), and 1 mM biotin) at 60 °C, 1100 rpm on a thermal block shaker. The pulldown eluent was loaded on 4–12% polyacrylamide Bis-Tris polyacrylamide gels (ThermoFisher) and then transferred onto PVDF membranes (GE Healthcare, Amersham) using a wet electrophoretic transfer system (BioRad). The membrane was then blocked with 5% nonfat dry milk in 0.1% PBST (0.1% Tween-20 in 1× PBS, pH 7.4). Rabbit polyclonal YTHDF2 antibody (Aviva system biology ARP67917\_P050) was diluted 1:1000 in 0.1% PBST and incubated on the membrane for 1 h at room temperature (15–25 °C) or overnight at 4 °C. Membranes were then washed in 0.1% PBST, and then incubated with HRP-conjugated goat anti-rabbit IgG (Abcam ab6721) at 1:2500 dilution in 0.1% PBST for 1 h at 25 °C. Membranes were then subsequently washed with 0.1% PBST and exposed with enhanced chemiluminescence (ECL; GE Healthcare).

**Dot blot assays.** RNA samples were quantified using UV spectroscopy. Methylated and unmodified RNA samples were spotted onto a nylon membrane (GE healthcare). Fluorescein labeled RNA samples (prepared as described in sample preparation section) were scanned by UVP imaging system to visualize the quantity of RNA spotted onto the membrane. The membrane was then UV-crosslinked and blocked for 1 h in 5% nonfat dry milk in 0.1% PBST (0.1% Tween-20 in 1× PBS, pH 7.4). The following steps were performed essentially as described for RNA pull-down (above). m<sup>6</sup>A antibody (SySy 202003, NEB E1610S, or Abcam ab151230) was diluted 1:1000 in 0.1% PBST and incubated on the membrane for 1 h at room temperature or overnight at 4 °C. Membranes were then washed in 0.1% PBST, incubated with HRP-conjugated goat anti-rabbit IgG (Abcam ab6721) (1:2500) in 0.1% PBST for 1 h at room temperature. The membrane was then washed again in 0.1% PBST, and developed with enhanced chemiluminescence (ECL; GE Healthcare). The SS construct used in dot blot assay is 5'-GGCGGm<sup>6</sup>ACUC-3'.

**Code availability.** The in-house Python scripts used for secondary structure prediction are available on request to the corresponding author.

**Data availability.** The authors declare that the data supporting the findings of this study are available within the article and its Supplementary Information files, or are available upon reasonable requests to the authors.

Received: 2 March 2018 Accepted: 2 May 2018

Published online: 17 July 2018

## References

- Li, X., Xiong, X. & Yi, C. Epitranscriptome sequencing technologies: decoding RNA modifications. *Nat. Methods* **14**, 23–31 (2016).
- Rottman, F., Shatkin, A. J. & Perry, R. P. Sequences containing methylated nucleotides at the 5' termini of messenger RNAs: possible implications for processing. *Cell* **3**, 197–199 (1974).
- Desrosiers, R., Friderici, K. & Rottman, F. Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells. *Proc. Natl Acad. Sci. USA* **71**, 3971–3975 (1974).
- Meyer, K. D. et al. Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell* **149**, 1635–1646 (2012).
- Dominissini, D. et al. Topology of the human and mouse m<sup>6</sup>A RNA methylomes revealed by m<sup>6</sup>A-seq. *Nature* **485**, 201–206 (2012).
- Lichinchi, G. et al. Dynamics of the human and viral m<sup>6</sup>A RNA methylomes during HIV-1 infection of T cells. *Nat. Microbiol.* **1**, 16011 (2016).
- Kennedy, E. M. et al. Posttranscriptional m<sup>6</sup>A editing of HIV-1 mRNAs enhances viral gene expression. *Cell Host Microbe* **19**, 675–685 (2016).
- Gokhale, N. S. et al. N<sup>6</sup>-methyladenosine in flaviviridae viral RNA genomes regulates infection. *Cell Host Microbe* **20**, 654–665 (2016).
- Bokar, J. A., Rath-Shambaugh, M. E., Ludwiczak, R., Narayan, P. & Rottman, F. Characterization and partial purification of mRNA N<sup>6</sup>-adenosine methyltransferase from HeLa cell nuclei. Internal mRNA methylation requires a multisubunit complex. *J. Biol. Chem.* **269**, 17697–17704 (1994).
- Liu, J. et al. A METTL3-METTL14 complex mediates mammalian nuclear RNA N<sup>6</sup>-adenosine methylation. *Nat. Chem. Biol.* **10**, 93–95 (2014).
- Wang, Y. et al. m<sup>6</sup>-methyladenosine modification destabilizes developmental regulators in embryonic stem cells. *Nat. Cell Biol.* **16**, 191–198 (2014).
- Jia, G. et al. N<sup>6</sup>-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat. Chem. Biol.* **7**, 885–887 (2011).
- Zheng, G. et al. ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol. Cell* **49**, 18–29 (2013).
- Lin, S., Choe, J., Du, P., Triboulet, R. & Gregory, R. I. The m<sup>6</sup>A methyltransferase METTL3 promotes translation in human cancer cells. *Mol. Cell* **62**, 335–345 (2016).
- Zhang, C. et al. Hypoxia induces the breast cancer stem cell phenotype by HIF-dependent and ALKBH5-mediated m<sup>6</sup>A-demethylation of NANOG mRNA. *Proc. Natl Acad. Sci. USA* **113**, E2047–E2056 (2016).
- Xiao, W. et al. Nuclear m<sup>6</sup>A reader YTHDC1 regulates mRNA splicing. *Mol. Cell* **61**, 507–519 (2016).
- Zhao, X. et al. FTO-dependent demethylation of N<sup>6</sup>-methyladenosine regulates mRNA splicing and is required for adipogenesis. *Cell Res.* **24**, 1403–1419 (2014).
- Fustin, J. M. et al. RNA-methylation-dependent RNA processing controls the speed of the circadian clock. *Cell* **155**, 793–806 (2013).
- Wang, X. et al. N<sup>6</sup>-methyladenosine-dependent regulation of messenger RNA stability. *Nature* **505**, 117–120 (2014).
- Wang, X. et al. N<sup>6</sup>-methyladenosine modulates messenger RNA translation efficiency. *Cell* **161**, 1388–1399 (2015).
- Meyer, K. D. et al. 5' UTR m<sup>6</sup>A promotes cap-independent translation. *Cell* **163**, 999–1010 (2015).
- Zhao, B. S., Roundtree, I. A. & He, C. Post-transcriptional gene regulation by mRNA modifications. *Nat. Rev. Mol. Cell Biol.* **18**, 31–42 (2017).
- Roundtree, I. A., Evans, M. E., Pan, T. & He, C. Dynamic RNA modifications in gene expression regulation. *Cell* **169**, 1187–1200 (2017).
- Meyer, K. D. & Jaffrey, S. R. Rethinking m<sup>6</sup>A readers, writers, and erasers. *Annu. Rev. Cell. Dev. Biol.* **33**, 319–342 (2017).
- Deng, X., Su, R., Feng, X., Wei, M. & Chen, J. Role of N<sup>6</sup>-methyladenosine modification in cancer. *Curr. Opin. Genet. Dev.* **48**, 1–7 (2017).
- Roundtree, I. A. et al. YTHDC1 mediates nuclear export of N<sup>6</sup>-methyladenosine methylated mRNAs. *eLife* **6**, e13111 (2017).
- Shi, H. et al. YTHDF3 facilitates translation and decay of N<sup>6</sup>-methyladenosine-modified RNA. *Cell Res.* **27**, 315–328 (2017).
- Roost, C. et al. Structure and thermodynamics of N<sup>6</sup>-methyladenosine in RNA: a spring-loaded base modification. *J. Am. Chem. Soc.* **137**, 2107–2115 (2015).
- Kierzek, E. & Kierzek, R. The thermodynamic stability of RNA duplexes and hairpins containing N<sup>6</sup>-alkyladenosines and 2-methylthio-N<sup>6</sup>-alkyladenosines. *Nucleic Acids Res* **31**, 4472–4480 (2003).

30. Zhou, K. I. et al.  $N^6$ -methyladenosine modification in a long noncoding RNA hairpin predisposes its conformation to protein binding. *J. Mol. Biol.* **428**, 822–833 (2016).
31. Liu, N. et al.  $N^6$ -methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature* **518**, 560–564 (2015).
32. Liu, N. et al.  $N^6$ -methyladenosine alters RNA structure to regulate binding of a low-complexity protein. *Nucleic Acids Res.* **45**, 6051–6063 (2017).
33. Huang, L., Ashraf, S., Wang, J. & Lilley, D. M. Control of box C/D snoRNP assembly by  $N^6$ -methylation of adenine. *EMBO Rep.* **18**, 1631–1645 (2017).
34. Choi, J. et al.  $N^6$ -methyladenosine in mRNA disrupts tRNA selection and translation-elongation dynamics. *Nat. Struct. Mol. Biol.* **23**, 110–115 (2016).
35. Spitale, R. C. et al. Structural imprints in vivo decode RNA regulatory mechanisms. *Nature* **519**, 486–490 (2015).
36. Kane, S. E. & Beemon, K. Precise localization of  $m^6A$  in Rous sarcoma virus RNA reveals clustering of methylation sites: implications for RNA processing. *Mol. Cell Biol.* **5**, 2298–2306 (1985).
37. Narayan, P. & Rottman, F. M. An in vitro system for accurate methylation of internal adenosine residues in messenger RNA. *Science* **242**, 1159–1162 (1988).
38. Csepány, T., Lin, A., Baldick, C. J. Jr. & Beemon, K. Sequence specificity of mRNA  $N^6$ -adenosine methyltransferase. *J. Biol. Chem.* **265**, 20117–20122 (1990).
39. Parisien, M. & Major, F. The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data. *Nature* **452**, 51–55 (2008).
40. Geerdes, H. A. & Hilbers, C. W. Ring current shifts in GU base pairs. *FEBS Lett.* **107**, 125–128 (1979).
41. Engel, J. D. & von Hippel, P. H. Effects of methylation on the stability of nucleic acid conformations: studies at the monomer level. *Biochemistry* **13**, 4143–4158 (1974).
42. Horowitz, S., Horowitz, A., Nilsen, T. W., Munns, T. W. & Rottman, F. M. Mapping of  $N^6$ -methyladenosine residues in bovine prolactin mRNA. *Proc. Natl Acad. Sci. USA* **81**, 5667–5671 (1984).
43. Linder, B. et al. Single-nucleotide-resolution mapping of  $m^6A$  and  $m^6A_m$  throughout the transcriptome. *Nat. Methods* **12**, 767–772 (2015).
44. Reuter, J. S. & Mathews, D. H. RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinforma.* **11**, 129 (2010).
45. Zhu, T. et al. Crystal structure of the YTH domain of YTHDF2 reveals mechanism for recognition of  $N^6$ -methyladenosine. *Cell Res.* **24**, 1493–1496 (2014).
46. Rossi, A. M. & Taylor, C. W. Analysis of protein-ligand interactions by fluorescence polarization. *Nat. Protoc.* **6**, 365–387 (2011).
47. Xu, C. et al. Structural basis for selective binding of  $m^6A$  RNA by the YTHDC1 YTH domain. *Nat. Chem. Biol.* **10**, 927–929 (2014).
48. Luo, S. & Tong, L. Molecular basis for the recognition of methylated adenines in RNA by the eukaryotic YTH domain. *Proc. Natl Acad. Sci. USA* **111**, 13834–13839 (2014).
49. Li, F., Zhao, D., Wu, J. & Shi, Y. Structure of the YTH domain of human YTHDF2 in complex with an  $m^6A$  mononucleotide reveals an aromatic cage for  $m^6A$  recognition. *Cell Res.* **24**, 1490–1492 (2014).
50. Zhang, Z. et al. The YTH domain is a novel RNA binding domain. *J. Biol. Chem.* **285**, 14701–14710 (2010).
51. Zheng, H., Shabalín, I. G., Handing, K. B., Bujnicki, J. M. & Minor, W. Magnesium-binding architectures in RNA crystal structures: validation, binding preferences, classification and motif detection. *Nucleic Acids Res.* **43**, 3789–3801 (2015).
52. Neuner, S., Santner, T., Kreutz, C. & Micura, R. The “speedy” synthesis of atom-specific  $^{15}N$  imino/amido-labeled RNA. *Chemistry* **21**, 11634–11643 (2015).
53. Delaglio, F. et al. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).
54. Bloomfield, V. A., Crothers, D. M. & Tinoco, I. *Nucleic Acids: Structures, Properties, and Functions* (University Science Books, Sausalito, CA, 2000).
55. Wilson, G. M., Sun, Y., Lu, H. & Brewer, G. Assembly of AUF1 oligomers on U-rich RNA targets by sequential dimer association. *J. Biol. Chem.* **274**, 33374–33381 (1999).
56. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH image to imageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675 (2012).

### Acknowledgements

We thank Nicole Orlovsky, Atul Rangadurai, and Mary Clay, members of the Al-Hashimi laboratory, and Dr. Tony Mustoe (UNC-Chapel Hill) for assistance and critical comments on the manuscript. We thank Dr. Richard Brennan for providing the UV-Vis spectrophotometer. We thank New England Biolabs (NEB) for providing the monoclonal  $m^6A$  antibody as a gift to S.M.H. We acknowledge technical support and resources from the Duke Magnetic Resonance Spectroscopy Center. This work was supported by the US National Institutes of Health (P01GM0066275 to H.M.A., R01AI125416 to S.M.H. and R00MH104712 to K.D.M.) and Burroughs Welcome Fund (to S.M.H.). The research reported in this article was performed by the Duke University faculty, research associate and was funded by US National Institute of Health contract to H.M.A.

### Author contributions

B.L., D.K.M. and H.M.A. conceived the project and experimental design. B.L. and D.K.M. prepared NMR samples and performed NMR experiments and analyzed NMR data. R.P. and C.K. prepared  $^{15}N$ -N1/3 labeled Guanosine and Uridine phosphoramidite. B.L. performed UV melting experiments, transcriptome-wide RNA secondary structure prediction, and dot blot assays. B.L. performed the YTH domain protein binding assays, with assistance from M.A.S. B.L. and S.H.C. performed the in vitro RNA pulldown experiments. H.M.A. and B.L. wrote the manuscript with critical input from D.K.M., K. D.M., M.A.S., S.M.H., and S.H.C.

### Additional information

**Supplementary Information** accompanies this paper at <https://doi.org/10.1038/s41467-018-05243-z>.

**Competing interests:** H.M.A. is an advisor to and holds an ownership interest in Nymirum, an RNA-based drug discovery company. The remaining authors declare no competing interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018