Methyltransferase-like protein 16 binds the 3′-terminal triple helix of MALAT1 long noncoding RNA

Jessica A. Brown[a,1], Charles G. Kinzig[a], Suzanne J. DeGregorio[a], and Joan A. Steitz[a,1]

[a]Department of Molecular Biophysics and Biochemistry, Howard Hughes Medical Institute, Yale University School of Medicine, New Haven, CT 06536

Contributed by Joan A. Steitz, October 24, 2016 (sent for review September 2, 2016; reviewed by Brenda L. Bass and Eric M. Phizicky)

Metastasis-associated lung adenocarcinoma transcript 1 (MALAT1), a cancer-promoting long noncoding RNA, accumulates in cells by using a 3′-triple-helical RNA stability element for nuclear expression (ENE). The ENE, a stem-loop structure containing a U-rich internal loop, interacts with a downstream A-rich tract (ENE+A) to form a blunt-ended triple helix composed of nine U•A-U triples interrupted by a C•G-C triple and C-G doublet. This unique structure prompted us to explore the possibility of protein binding. Native gel-shift assays revealed a shift in radiolabeled MALAT1 ENE+A RNA upon addition of HEK293T cell lysate. Competitive gel-shift assays suggested that protein binding depends not only on the triple-helical structure but also its nucleotide composition. Selection from the lysate using a biotinylated-RNA probe followed by mass spectrometry identified methyltransferase-like protein 16 (METTL16), a putative RNA methyltransferase, as an interacting protein of the MALAT1 ENE+A. Gel-shift assays confirmed the METTL16–MALAT1 ENE+A interaction in vitro: Binding was observed with recombinant METTL16, but diminished in lysate depleted of METTL16, and a supershift was detected after adding anti-METTL16 antibody. Importantly, RNA immunoprecipitation after in vivo UV cross-linking and an in situ proximity ligation assay for RNA–protein interactions confirmed an association between METTL16 and MALAT1 in cells. METTL16 is an abundant (~5 × 10^5 molecules per cell) nuclear protein in HeLa cells. Its identification as a triple-stranded RNA binding protein supports the formation of RNA triple helices inside cells and suggests the existence of a class of triple-stranded RNA binding proteins, which may enable the discovery of additional cellular RNA triple helices.

RNA binding protein | RNA triple helix | noncoding RNA

A triple-helical RNA stability element, known as an element for nuclear expression (ENE), protects viral and cellular long noncoding (lnc)RNAs as well as viral genomic RNAs from degradation (1–8). To protect RNA, a U-rich internal loop in the ENE stem-loop structure associates with a downstream 3′-poly(A) tail or a genomically encoded 3′-A-rich tract to form a triple helix (ENE+A) composed of mostly U•A-U major-groove base triples with adjacent A-minor base triples (Fig. 1*A*) (3, 6).

The first ENE was discovered in the highly abundant polyadenylated nuclear (PAN) lncRNA expressed by the Kaposi's sarcoma-associated herpesvirus (KSHV) (1). Subsequent studies confirmed similar ENEs in two abundant cellular lncRNAs: metastasis-associated lung adenocarcinoma transcript 1 (MALAT1) and multiple endocrine neoplasia beta (MENβ) RNA (5, 7). These cellular ENEs exhibit two distinct features not observed in other ENEs: (*i*) both ENEs interact with a 3′-terminal genomically encoded A-rich tract rather than a posttranscriptionally added polyA tail and (*ii*) both ENEs form blunt-ended bipartite triple helices composed of U•A-U base triples interrupted by a C^+•G-C base triple and C-G base pair (Fig. 1*A*). These similar, yet unique, triple-helical ENE+A structures prompted us to explore the possibility of protein binding, because ribonucleoprotein (RNP) complexes mediate a myriad of cellular processes (9). Importantly, most RNA-binding proteins (RBP) recognize single- or double-stranded (ss or ds) RNA (10, 11); a triple-stranded RBP remains to be discovered.

In this study, we examined the possibility that proteins in HEK293T lysate selectively bind the ENE+A structure from MALAT1, MENβ, and KSHV PAN RNAs. We discovered that methyltransferase-like protein 16 (METTL16), a conserved putative RNA methyltransferase, interacts specifically with the MALAT1 triple helix. This discovery argues that the triple-helical structure of MALAT1 forms inside cells and suggests there may be a class of triple-stranded RBPs.

Results

A Protein in HEK293T Lysate Specifically Recognizes the Triple-Helical Structure and Nucleotide Composition of the MALAT1 ENE+A. To search for ENE+A-interacting proteins, we incubated 5′-radiolabeled MALAT1, MENβ, and KSHV PAN ENE+A RNAs (Fig. 1*A*) in the presence of increasing amounts of HEK293T (Fig. 1*B*) or HeLa (Fig. 1*C*) whole-cell lysate. RNPs were separated from unbound RNA by using a native electrophoretic mobility shift assay (EMSA). Interestingly, three different bands, labeled as RNP1–3, were observed for the MALAT1 ENE+A RNA; the greatest shift corresponded to RNP3 (Fig. 1 *B* and *C*). In contrast, little to no binding was observed for the MENβ and KSHV PAN ENE+A RNAs. The binding profiles were similar for both HEK293T and HeLa cell lysates; therefore, we conclude RNP3 is likely not cell-type specific.

Next, we examined whether protein binding was specific for the MALAT1 triple helix or another region of the ENE+A, such as the ssRNA or dsRNA regions also present in the construct used. Ten different unlabeled MALAT1 ENE+A RNA competitors (Fig. 1*D*), including wild type (WT), were individually added to

Significance

RNA triple helices were deduced to form in vitro almost 60 years ago, yet only three examples from eukaryotic cellular RNAs have been structurally validated. The longest triple helix, the MALAT1 (metastasis-associated lung adenocarcinoma transcript 1) ENE+A (element for nuclear expression with a downstream A-rich tract), presents an opportunity to investigate the biological roles of these enigmatic structures. We have discovered that the MALAT1 triple helix is specifically recognized and bound by METTL16 (methyltransferase-like protein 16). There are two important implications: (*i*) the MALAT1 triple helix is a bona fide structure in the cellular environment and (*ii*) there may exist an undiscovered class of triple-stranded RNA binding proteins. METTL16's antiproliferative role in *Caenorhabditis elegans* suggests that the METTL16–MALAT1 complex may contribute to the oncogenic activity of MALAT1.

Author contributions: J.A.B. and J.A.S. designed research; J.A.B., C.G.K., and S.J.D. performed research; J.A.B., C.G.K., and J.A.S. analyzed data; and J.A.B. and J.A.S. wrote the paper.

Reviewers: B.L.B., University of Utah School of Medicine; and E.M.P., University of Rochester.

The authors declare no conflict of interest.

[1]To whom correspondence may be addressed. Email: jbrown33@nd.edu. or joan.steitz@yale.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1614759113/-/DCSupplemental.

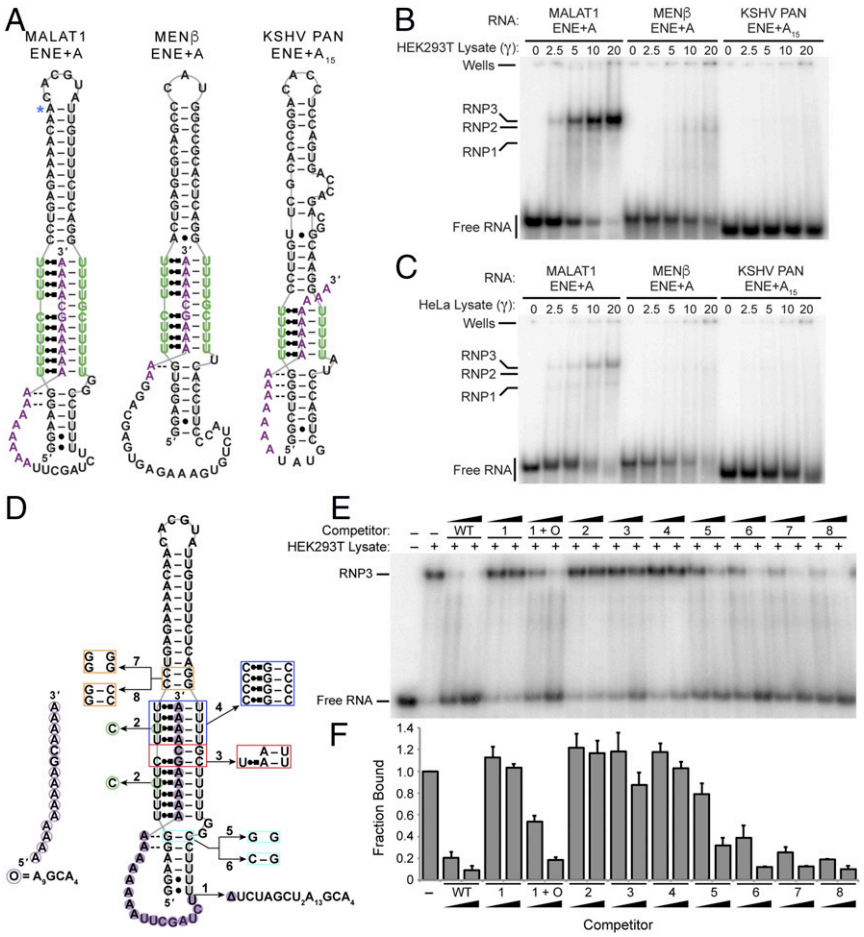

Fig. 1. EMSAs reveal assembly of a MALAT1 ENE+A RNP and protein binding dependent on triple helix formation. (*A*) Schematic diagrams for the ENE+A structures from MALAT1, MENβ, and KSHV PAN RNAs. Nucleotides in the U-rich internal loop are green, and nucleotides in the A-rich tract are purple. Nucleotide interactions are represented as follows: single dash for Watson–Crick base pair, black circle for noncanonical base pair, double dash for A-minor interaction, and Leontis–Westhof's circle-square notation for Hoogsteen interaction. The blue asterisk marks A8290. (*B* and *C*) EMSAs were performed by adding increasing amounts of HEK293T (*B*) or HeLa (*C*) cell lysate to a 5′-[^{32}P]–labeled MALAT1, MENβ, or KSHV PAN ENE+A RNA. Observed RNPs are arbitrarily labeled RNP1–3 to the left. (*D*) Schematic diagram of the mutant MALAT1 ENE+A RNAs used as unlabeled competitors in *E*. Note, the 5′-UCUAGCU$_2$A$_{13}$GCA$_4$-3′ sequence is deleted (Δ) in Mut1. (*E*) Competitive gel-shift assays were performed by incubating HEK293T lysate and 5′-[^{32}P]–labeled MALAT1 ENE+A RNA in the presence of 10- or 50-fold molar excess of the appropriate unlabeled competitor RNA. (*F*) The relative amount of RNP3 formation was quantified for each competitor, with RNP3 in the absence of a competitor set arbitrarily to 1. Fraction bound is the average of at least three independent experiments; error bars represent SD.

the binding reaction mixture in either 10- or 50-fold molar excess over 5′-radiolabeled MALAT1 ENE+A RNA. WT MALAT1 ENE+A competitor RNA reduced RNP3 formation by ∼80–91% (Fig. 1 *E* and *F*). In contrast, Mut1 RNA, which is only the ENE stem-loop structure based on UV thermal denaturation assays (5), did not compete. However, combining Mut1 RNA with oligonucleotide (O in Fig. 1*D*) A_9GCA_4, which forms a triple helix based on UV thermal denaturation assays (5), led to a reduction in RNP3 formation by up to ∼82%. Additionally, no competition was detected for Mut2 RNA, a competitor having two U-to-C substitutions that disrupt Hoogsteen interactions in the triple helix (Fig. 1 *D*–*F*) as confirmed by previous EMSA studies (5). These results suggest that the triple-helical region of the MALAT1 ENE+A is required for RNP formation.

We were then interested in determining whether recognition of the triple helix depends on nucleotide composition as suggested by the lack of RNP3 formation for MENβ and KSHV PAN ENE+A RNAs (Fig. 1 *A*–*C*). To this end, we mutated the unique C^+•G-C/C-G interruption in the MALAT1 ENE+A to U•A-U/A-U (Mut3) and a stack of four U•A-U triples to C^+•G-C triples (Mut4) (Fig. 1*D*). Although Mut3 and Mut4 likely maintain a triple-helical structure based on their stabilization activity in vivo (5, 12), neither of these RNAs significantly affected RNP3 formation, thereby indicating that protein recognition depends on the sequence and structure of the MALAT1 triple helix.

Lastly, we tested whether RNP3 formation requires Watson–Crick base pairing at the duplex–triplex junctions (Mut5-8; Fig. 1*D*). In general, all these mutants reduced RNP3 formation to an extent similar to that of the WT competitor (Fig. 1 *E* and *F*). However, disrupting Watson–Crick base pairing within an A-minor interaction (Mut5) reduced RNP3 formation by a maximum of ∼68% compared with ∼90% for Mut6-8, suggesting that protein recognition is mildly sensitive to duplex formation at that position (Fig. 1 *E* and *F*).

METTL16 Binds to the MALAT1 ENE+A in Vitro. Our next objective was to identify the protein(s) interacting with the MALAT1 ENE+A RNA. We began with a label transfer experiment, whereby uniformly radiolabeled MALAT1 ENE+A RNA was added to HEK293T whole-cell lysate with or without WT or Mut2 competitors (Fig. 1*D*), exposed to UV light, and treated with RNase ONE before loading onto an SDS-denaturing gel (Fig. 2*A*). A protein

with a molecular mass of ~70 kDa was cross-linked and, importantly, WT but not Mut2 competitor RNA reduced the 70-kDa signal, indicating an interaction specific for the triple helix (Fig. 2*A*).

We then proceeded to isolate the MALAT1 ENE+A-interacting protein(s). MALAT1 ENE+A RNA was immobilized as follows: a 24-nucleotide (nt) sequence was added to the 5′ end of ENE+A RNA and annealed to a complementary 3′-biotinylated oligonucleotide, allowing the MALAT1 ENE+A to be coupled to streptavidin-coated beads (Fig. 2*B* and Fig. S1 *A*–*C*). Affinity purification from HEK293T whole-cell lysate (Fig. 2*C*, see *Materials and Methods* for details) revealed a protein of ~70 kDa that was retained with WT MALAT1 ENE+A RNA10 but not mutant RNA11 or the negative control RNA9 (Fig. 2 *B*–*D*). Mass spectrometry identified the major protein in this SDS/PAGE band as METTL16. Other major protein bands were identified as nucleophosmin (NPM), ATP-dependent RNA helicase A (DHX9), interleukin-enhancer binding factor 3 (ILF3), and ILF2 (Fig. 2*D*).

To confirm that METTL16 is a component of the MALAT1 ENE+A RNP (i.e., RNP3), we performed three independent EMSA-based experiments. First, HEK293T lysate or recombinant human METTL16, which was tagged with myc-DDK at its C terminus (METTL16-myc-DDK), was added to 5′-radiolabeled MALAT1 ENE+A RNA in the absence or presence of WT or Mut2 competitors (Fig. 1*D*). The EMSA result showed formation of RNP3, except where WT competitor was added, for both HEK293T lysate and recombinant METTL16, indicating a direct interaction (Fig. 3*A*). Second, RNP3 formation was noticeably lower using HEK293T lysate prepared from cells treated with siRNA against METTL16 versus a negative control (Fig. 3*B*). Third, addition of anti-METTL16 antibody to the binding mixture induced a supershift for RNP3, whereas addition of IgG did not (Fig. 3*C*). Altogether, these results corroborate the presence of METTL16 in RNP3.

Lastly, the presence of NPM, DHX9, ILF3, and ILF2 in RNP3 was evaluated by coimmunoprecipitation of METTL16 from untreated or RNase A-treated HeLa nuclear-enriched lysate followed by Western blot analysis. Fig. S2 shows that NPM1, DHX9, ILF3, and ILF2 interact with METTL16 mostly in an RNA-independent manner. Therefore, we did not continue to study these proteins.

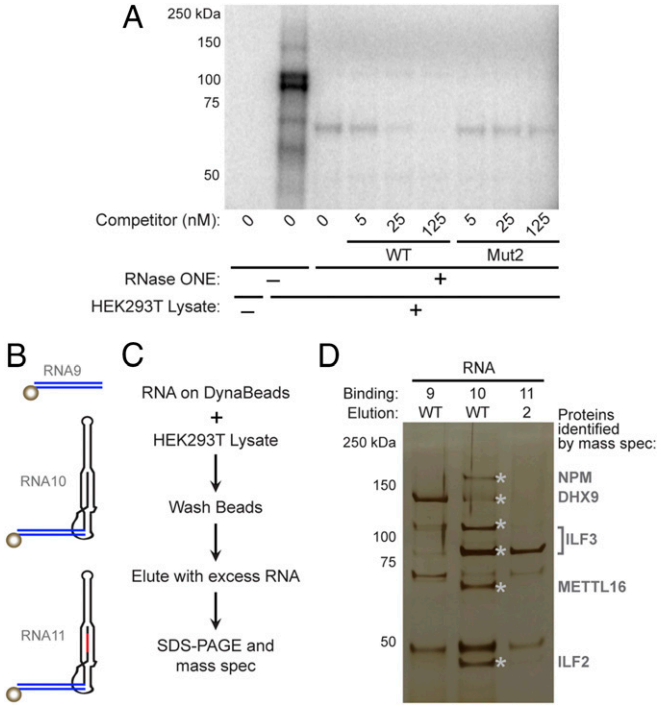

Fig. 2. Identification of a MALAT1 ENE+A-interacting protein. (*A*) An ~70-kDa protein in HEK293T cell lysate UV cross-links to radiolabeled MALAT1 ENE+A RNA. WT or Mut2 competitor RNAs were added to identify protein binding specific for the triple helix. (*B*) Depiction of RNAs 9–11 used in the pulldown scheme in *C*. A 3′-biotinylated oligonucleotide was annealed to its complement (RNA9) or to its complement fused to the 5′ end of a WT (RNA10) or a triplex-disrupting mutant (RNA11) MALAT1 ENE+A RNA (see Fig. S1 *A*–*C* for details). The dsRNA added for pulldown is in blue, the biotin-streptavidin interaction is a brown circle, the ENE+A structure is black, and the mutated region is in red. (*C*) Pulldown strategy used to isolate the MALAT1 ENE+A RNP. Folded RNAs were coupled to streptavidin-coated DynaBeads before incubation in HEK293T cell lysate. Bound proteins were eluted with either the WT or Mut2 MALAT1 ENE+A RNA shown in Fig. 1*D* and subjected to mass spectrometric analysis. (*D*) Eluted proteins were resolved by using SDS/PAGE and silver stained. Protein bands (*) identified by mass spectrometry are listed to the right.

METTL16 Is an Abundant Nuclear Protein That Interacts with MALAT1 in Vivo. The preceding experiments were performed in cell lysate, making it possible that METTL16 is not a physiologically relevant binding partner of MALAT1. To address this concern, we determined the cellular localization of human METTL16; MALAT1 is known to reside specifically in nuclear speckles (13). In immunofluorescence experiments, anti-METTL16 stained the entire nucleoplasm of HeLa cells, although nucleolar staining varied from cell to cell (Fig. 4*A*). Because MALAT1 is a relatively abundant nuclear RNA, we measured the cellular levels of METTL16 and MALAT1 in both HeLa and HEK293T cells. Quantitative Northern blot analysis estimated ~3,500 and ~600 molecules of MALAT1 in HeLa and HEK293T cells, respectively (Fig. S3*A*), whereas quantitative Western blot analysis of METTL16 showed ~500,000 and ~750,000 molecules of METTL16 in HeLa and HEK293T cells, respectively (Fig. S3*B*). Thus, there is more than adequate METTL16 to theoretically associate with MALAT1 in vivo, indicating that a METTL16–MALAT1 interaction in cells is possible.

To confirm that METTL16 associates with MALAT1 in cells, we pursued several orthogonal approaches. First, we used UV cross-linking of intact HeLa cells and immunopurification of endogenous METTL16 from nuclear-enriched lysate (Fig. 4*B*) and analyzed the coprecipitated RNAs by reverse transcription and quantitative PCR (CLIP-qPCR) (Fig. 4*C*). Our CLIP-qPCR results show an enrichment for MALAT1, but not for MENβ or the predominantly nuclear cancer-associated lncRNA HOTAIR, in the METTL16 IP compared with IgG. Because METTL16 is predicted to be an rRNA methyltransferase (14) and is relatively abundant, we checked its association with 5.8S, 18S, and 28S rRNAs. The METTL16 IP samples indeed showed enrichment of all three rRNAs compared with the IgG control but no strong preference for a single rRNA species (Fig. 4*C*).

Although our CLIP-qPCR data support an interaction between METTL16 and MALAT1, they do not reveal whether this interaction occurs with the triple-helical ENE+A or another region of the ~8-knt MALAT1 lncRNA. To address this question, we repeated our CLIP-qPCR experiment using HEK293T cells transfected with intronless β-globin reporter (βΔ1,2) constructs whose transcripts terminate with either the MALAT1 or MENβ ENE+A structure (Fig. 4 *D* and *E*). The βΔ1,2 reporter mRNAs were expressed at similar levels based on RT-qPCR data (Fig. 4*F*, *Left*). Importantly, the βΔ1,2-MALAT1 ENE+A reporter was enriched ~sixfold over the βΔ1,2-MENβ ENE+A reporter in the METTL16 IP samples, consistent with METTL16 interacting directly with the MALAT1 triple helix in vivo (Fig. 4*F*, *Right*).

To corroborate the in vivo interaction, we visualized the METTL16-MALAT1 ENE+A RNP in HeLa cells by using an in situ RNA-proximity ligation assay (RNA-PLA) (15). RNA-PLA uses a cDNA oligonucleotide to recognize a target RNA and an antibody to recognize the protein factor of the putative RNP.

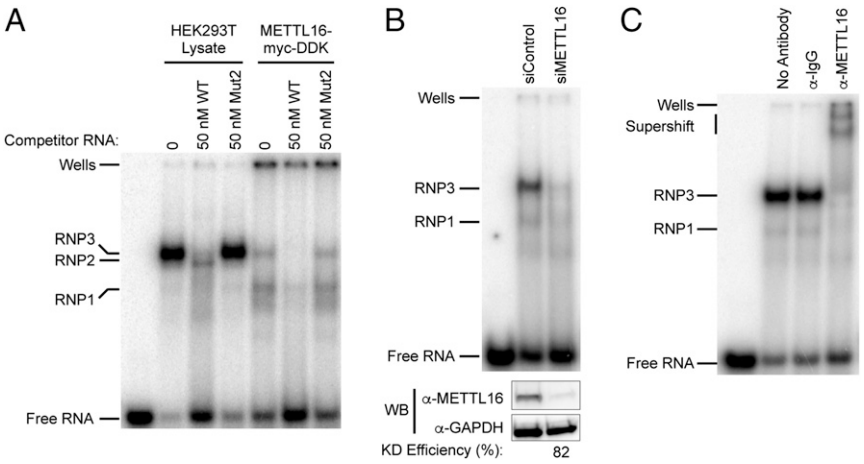

Fig. 3. METTL16 binds to MALAT1 ENE+A RNA in vitro. (*A*) Native gel-shift assays were performed by incubating HEK293T lysate or recombinant C-terminal myc-DDK–tagged human METTL16 with 5′-[^{32}P]–labeled MALAT1 ENE+A RNA (1 nM; Fig. 1*A*) in the absence or presence of 50 nM WT or Mut2 (Fig. 1*D*) competitor RNA. (*B*) Cell lysate was prepared from HEK293T cells treated with either negative control siRNA (siControl) or siRNA against METTL16 (siMETTL16). Western blots (*Lower*) confirmed knockdown; GAPDH was used as a loading control. The siControl or siMETTL16 lysate was then incubated with 5′-[^{32}P]–labeled MALAT1 ENE+A RNA (1 nM) before separation on native PAGE (*Upper*). (*C*) A supershift assay was performed by adding IgG control or anti-METTL16 antibodies to a binding mixture containing HEK293T lysate and 5′-[^{32}P]–labeled MALAT1 ENE+A RNA.

Stable and transient RNA–protein interactions are visualized by fluorescence per the standard PLA protocol (16). Here, we used a MALAT1-antisense DNA probe that recognizes an accessible 45-nt sequence immediately 5′ to the ENE+A, whereas the anti-METTL16 antibody recognizes endogenous METTL16. A PLA signal was detected in nuclei only for the MALAT1 antisense–anti-METTL16 probe combination (Fig. S4). No strong PLA signal was detected for MALAT1 sense–anti-METTL16 or MALAT1 antisense–anti-NPM1 probe combinations (Fig. S4). We conclude that METTL16 interacts directly with the MALAT1 triple helix inside cells.

Discussion

RNPs participate in key cellular processes, and their dysfunction underlies many human diseases, including cancer (9). Proteins recognize their RNA targets in a sequence- and/or a structure-dependent manner, including single-, double-, or even four-stranded RNAs (10, 11, 17). In this work, we discovered that METTL16 interacts with the MALAT1 triple helix in vitro (Figs. 2–3) and in vivo (Fig. 4 and Fig. S4). Interestingly, this interaction depends on both the triple-helical structure and sequence of the MALAT1 ENE+A (Fig. 1 *D–F*). METTL16 is a protein shown to recognize an RNA triple helix inside cells. This discovery, coupled with our X-ray crystal structure of the MALAT1 ENE+A (6), provides persuasive evidence that the MALAT1 triple helix, and likely other RNA triple helices, exists in the cellular environment.

The function of the METTL16-MALAT1 ENE+A RNP is unknown. In this regard, studies of two METTL16-related proteins, METT-10 from *Caenorhabditis elegans* (38% identity) and YbiN from *Escherichia coli* (31% identity) (14, 18, 19), are relevant. Both METTL16 (Fig. 4*A*) and METT-10 (17) are nuclear, with the localization of METT-10 facilitated by a nuclear localization sequence as well as protein–protein interactions with dynein light chain 1 (DYNLL1). In *C. elegans*, the known functions of METT-10 include inhibition of germ cell proliferation and promotion of mitotic cell cycle division (18). Such functions in cell growth may underlie the cancer-promoting activities (i.e., cell proliferation, invasion, and metastasis) of MALAT1 (20-23). Future experiments are needed to address this possible molecular link to cancer.

The METTL16-MALAT1 ENE+A RNP may contain additional proteins, such as DYNLL1 (19) or those identified in our RNA pulldown experiment: NPM, DHX9, ILF3, or ILF2 (Fig. 2 *B–D*). Our METTL16 coimmunoprecipitation experiment did not provide support for NPM, DHX9, ILF3, or ILF2 having a strong RNA-dependent interaction with METTL16 (Fig. S2). Nonetheless, it is intriguing that DHX9, an NTP-dependent DNA and RNA helicase, was identified. DHX9 reportedly interacts with G quadruplexes and antiparallel DNA triple helices (24). Thus, DHX9 may contribute to the unwinding of the MALAT1 triple helix, possibly the first step in the 3′-5′ decay of MALAT1. Identifying the METTL16 interactome is an exciting next step.

METTL16 is predicted to be a rRNA methyltransferase because YbiN in *E. coli* methylates the N^6 position of A1618 in 23S rRNA (14). This putative function raises the question of whether the MALAT1 ENE+A is a substrate for METTL16. Thus far, a relatively small N^6-methyladenosine signal has been detected near A8290 in the MALAT1 ENE+A by using m6A individual-nucleotide-resolution (mi)CLIP methodology (25); however, A8290 resides in the upper stem loop and not the triple helix (Fig. 1*A*). Additional experiments are needed to establish the cellular substrates of METTL16 and the type of modification it catalyzes, if it is in fact catalytically competent.

Because METTL16 is present at ~5–7.5 × 10^5 copies per cell (Fig. S3*B*), it could potentially interact with and/or methylate a large number of cellular RNAs, including its predicted rRNA target. Our METTL16 CLIP-qPCR results show interactions with 5.8S, 18S, and 28S rRNAs (Fig. 4*C*). It is unclear whether METTL16 interacts with rRNA precursors, mature rRNAs, and/or the ribosome (26). Regardless, we infer that these METTL16–rRNA interactions are unlikely to occur via a triple helix. Rather, we propose that an oligomeric form of METTL16 uniquely recognizes the MALAT1 triple helix. This speculative model is supported by the two bands observed in our EMSA results using METTL16-myc-DDK (Fig. 3*B*) and by the oligomerization activity detected for METT-10 in *C. elegans* (19). Altogether, it will be interesting to complete the biochemical characterization of METTL16, especially if it is a player in N^6-methyladenosine biology and the epitranscriptome.

In summary, with the discovery of METTL16 as its protein-binding partner, the MALAT1 triple helix continues to expand our knowledge about the cellular roles of naturally occurring RNA triple helices. METTL16 may be the founding member of a previously unknown class of triple-stranded RBPs. A 3D structure of the METTL16-MALAT1 ENE+A RNP therefore has the potential

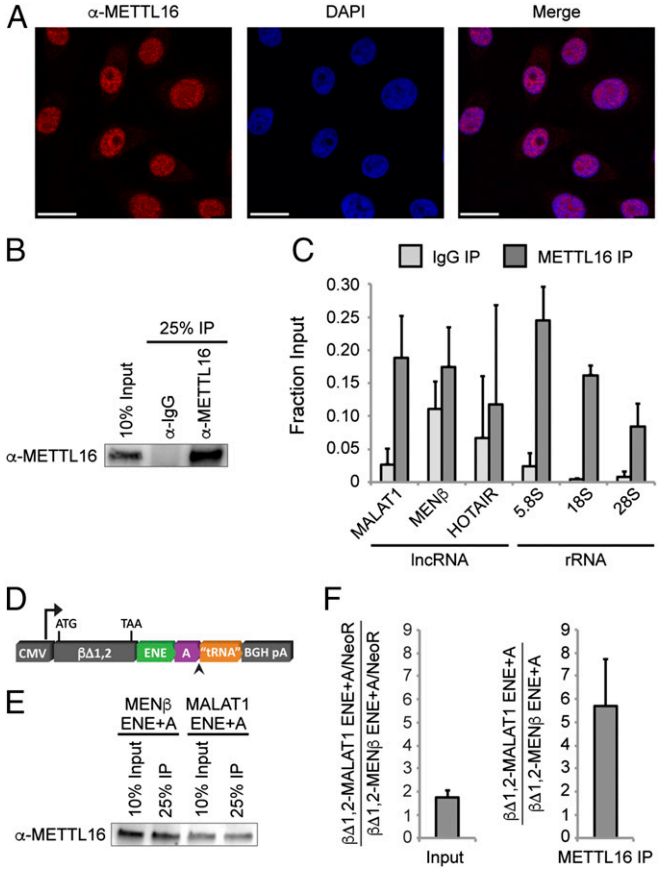

Fig. 4. METTL16 associates with MALAT1 in vivo. (*A*) Subcellular localization of METTL16 in HeLa cells using immunofluorescence. Nuclei were stained with DAPI. (Scale bars: 20 μm.) (*B*) HeLa nuclear-enriched lysates were subjected to anti-METTL16 or control IgG selection and subsequently verified by Western blot. (*C*) RT-qPCR quantified the amounts of endogenous RNAs relative to input for both IgG (light gray bars) and METTL16 (dark gray bars) IPs. Error bars represent SD from three biological replicates. (*D*) Schematic diagram of the intronless β-globin (βΔ1,2) plasmid constructs with the ENE (green), A-rich tract (purple), and tRNA-like sequence (orange, representing mascRNA or menRNA) from MALAT1 or MENβ. β-globin expression is driven by the CMV promoter, the RNase P cleavage site is indicated by an arrowhead, and BGH pA is the bovine growth hormone polyadenylation signal. (*E*) HEK293T cells were transfected with plasmids expressing βΔ1,2-MALAT1 ENE+A or βΔ1,2-MENβ ENE+A reporter mRNA, and METTL16 was immunoprecipitated from prepared nuclear-enriched lysates. (*F*) RT-qPCR quantified the expression level of the βΔ1,2 reporter mRNA (normalized to the NeoR transfection control) in the inputs for βΔ1,2-MALAT1 ENE+A relative to the βΔ1,2-MENβ ENE+A reporter mRNA (*Left*). (*Right*) RT-qPCR quantified the fold enrichment in the METTL16 IPs for the βΔ1,2-MALAT1 ENE+A relative to the βΔ1,2-MENβ ENE+A reporter mRNA.

to uncover a triple-stranded RNA binding domain or motif. Most importantly, the METTL16–MALAT1 ENE+A complex may enable the future discovery of other triple-stranded RNA binding proteins and their associated RNAs.

Materials and Methods

Preparation of RNA and Oligonucleotides. RNA oligonucleotides 5′-A_9GCA_4-3′ (Dharmacon), 5′-phosphorylated UUCACAGUGGCUAAGUUCCGC-3′ biotin via C6 linker (IDT), and 5′-GGAACUUAGCCACUGUGGAAG-3′ (IDT) were prepared as described (5, 27). For EMSA, the MALAT1 (nucleotides 8263–8355), MENβ (nucleotides 22651–22743), and KSHV PAN (nucleotides 895–964) ENEs were transcribed by his-tagged T7 RNA polymerase using a PCR-generated template, and the RNA product was processed as described (5) except that the 5′-triphosphate moiety was removed by using calf intestinal alkaline phosphatase (New England Biolabs) per the manufacturer's protocol. Competitor RNAs (Fig. 1*D*) and MALAT1 RNAs used for affinity purification (Fig. S1 *B* and *C*) had nucleotide changes incorporated into the chemically synthesized oligonucleotides (IDT or Keck Oligo Synthesis Resource) used for PCR; RNAs were prepared as described (5). To measure MALAT1 molecules per cell, MALAT1 RNA standards (no. 1, nucleotides 7818–7916; no. 2, nucleotides 8121–8355) were transcribed and gel purified as described (5). To enhance transcription for the above RNAs, the templates contained a 5′-GG at the start of transcription. A 3′-A_{15} tail was added to serve as the polyA tail of KSHV PAN ENE. PLA oligonucleotide probes 5′-GACCCTACTGAAGAGCATTGGAGATCAGCTTCCGCTAAGATGCTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATATGACAGAACTAGACACTCTT-3′ (antisense MALAT1) and 5′-AGCATCTTAGCGGAAGCTGATCTCCAATGCTCTTCAGTAGGGTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATATGACAGAACTAGACACTCTT-3′ (sense MALAT1) were purchased from Keck Oligo Synthesis Resource (Yale University) and gel purified before use.

Cell Culture, Preparation of Native Lysate, and Transfections. HEK293T and HeLa JW36 cells were grown in Dulbecco's modified Eagle medium (Life Technologies) supplemented with 10% (vol/vol) FBS, 2 mM L-glutamine, and 1× penicillin/streptomycin solution at 37 °C and 5% CO_2. Native whole-cell lysate was prepared as described (28) except that the total protein concentration was measured by using a Bradford assay (Pierce) with BSA as a standard before freezing aliquots. For METTL16 knockdown, 50 nM negative control siRNA (Life Technologies AM4611) or siRNA against METTL16 (Life Technologies AM16708) was transfected into 3×10^5 HEK293T cells in 12-well plates by using Lipofectamine RNAiMAX according to the manufacturer's instructions. After 48 h, cells were transfected with siRNA a second time and expanded into six-well plates. Cells were harvested 48 h after the second transfection, and lysate was prepared as described above.

EMSA. RNAs were 5′-[^{32}P]-labeled by using γ-[^{32}P]ATP (PerkinElmer) and T4 PNK (New England Biolabs) per the manufacturer's protocol. Free γ-[^{32}P]ATP was removed by passing the reaction mixture through a Microspin G-50 column (GE Healthcare). Immediately before setting up the binding reaction, RNA was heated at 95 °C for 1 min, snap-cooled on ice for 2 min, and allowed to equilibrate at room temperature for at least 30 min. HEK293T or HeLa whole-cell lysate (~1 μg/μL unless indicated otherwise) was added to a binding mixture containing 1 nM 5′-[^{32}P]-labeled RNA in buffer B [25 mM Tris pH 7.5, 150 mM KCl, 1 mM $MgCl_2$, 0.5 mg/mL carrier RNA, 5% (vol/vol) glycerol, 0.5 U/μL RNase inhibitor (Roche), and 1 mM DTT]. The binding reaction was incubated at room temperature for 30 min before loading onto a 5 or 6% native Tris-borate gel with 1 mM $MgCl_2$ in the gel and running buffer. After electrophoresis for 2.5 h at 150 V and room temperature, the gels were dried, exposed to a phosphorimager screen overnight, and scanned by using a Storm 860 (GE Healthcare). Amounts of competitor RNAs (10 or 50 nM) are shown in the figures or legends and were added to the radiolabeled RNA before adding lysate (~1 μg/μL) or recombinant human METTL16-myc-DDK (~50 ng/μL). ImageQuant (Molecular Dynamics) was used to quantitate the radiolabeled RNA signal. Recombinant human METTL16-myc-DDK (OriGene) was isolated from HEK293T cells. For supershift assays, 50 ng/μL of either anti-rabbit IgG control (Santa Cruz Biotechnology sc-2027) or anti-METTL16 antibody (Bethyl Laboratories A304-192A) were added to a reaction mixture devoid of DTT after the 30-min incubation period and incubated for an additional 60 min before gel separation.

Cross-Linking and Affinity Purification of Proteins from Extract. For in vitro UV cross-linking studies, a binding reaction was set up as described for gel-shift assays except that 200,000 cpm of uniformly radiolabeled MALAT1 ENE+A RNA was added. After 30 min, reactions were transferred to ice, irradiated 1 cm from the UV light for 20 min inside a Stratalinker, and digested with 5 U of RNase ONE (Promega) for 30 min at 30 °C. Samples were resolved on a 4–12% NuPage Novex Bis-Tris protein gel; gels were dried, exposed to a phosphorimager screen overnight, and scanned by using a Storm 860.

For affinity purification of proteins, RNAs were annealed to a 3′-biotinylated oligonucleotide (1:1 ratio) by heating the RNA mixture at 95 °C for 2 min, snap-cooled on ice for 2 min, and allowed to equilibrate at 4 °C for 4.5 h overnight. RNA was then immobilized on MyOne C1 streptavidin DynaBeads for 1 h at room temperature or overnight at 4 °C. Washed beads were then combined with 450 μL of precleared HEK293T whole-cell lysate in buffer B and rotated for 30–75 min at room temperature. Beads were washed three times in buffer B at 250 mM KCl and transferred to a clean tube. Proteins were eluted from beads by adding 200 pmol of WT or mutant RNA in 20 mM Tris pH 7.5, 30 mM KCl, 1 mM $MgCl_2$, 1 mM DTT, and EDTA-free cOmplete protease inhibitor mixture (Roche). After rotating the sample at room temperature for 4 h, supernatant was loaded onto a 4–12% NuPage Novex Bis-Tris protein gel.

Proteins were visualized by using Pierce Silver Stain for Mass Spectrometry Kit per the protocol. Excised bands were destained before submission for mass spectrometric analysis (Keck MS & Proteomics Resource, Yale University).

Coimmunoprecipitation and Western Blot Analyses. Three 10-cm plates of ∼80% confluent HeLa cells were rinsed with PBS, and the nuclear fraction was isolated by resuspending cells in 500 μL of a 0.5% Igepal 630-PBS solution supplemented with 1 mM PMSF and EDTA-free cOmplete protease inhibitor mixture (Roche). After a 5-min incubation on ice, the lysis efficiency was confirmed by inspecting nuclei under a microscope. Nuclei were pelleted by centrifugation for 5 min at 1,000 × *g* (4 °C) and washed before resuspension in 500 μL of NET-2 buffer. Nuclei were sonicated at 30% for 3 × 5 s with a 60-s pause on ice. Cellular debris was removed via centrifugation for 15 min at 15,000 × *g* and 4 °C. One-third of the nuclear-enriched lysate was treated with 3 μg of protease-free RNase A (ThermoFisher) for 15 min at 37 °C, and then 8 μg of anti-rabbit IgG (Santa Cruz Biotechnology sc-2027) or anti-METTL16 antibody (Abcam 185990) was used for each immunoprecipitation reaction along with 80 μL of Dynabeads protein A (Life Technologies). Beads were washed five times with NET-2 buffer at 250 mM KCl. The following antibody dilutions were used for Western blot analysis: METTL16 (Abcam 128528, 1:1,000), DHX9 (Bethyl Laboratories A300-855A-T, 1:750), ILF3 (Bethyl Laboratories A303-120A-T, 1:1,000), ILF2, (Bethyl Laboratories A303-147A-T, 1:1,000), NPM1 (Bethyl Laboratories A302-402A-T, 1:1,000), GAPDH (Cell Signaling Technology 14C10, 1:1,000), and Clean-Blot IP Detection Reagent HRP (ThermoFisher, 1:400).

Immunofluorescence. HeLa cells were fixed onto chamber slides with 4% (vol/vol) paraformaldehyde in PBS for 30 min on ice, permeabilized with 0.2% Triton X-100 in PBS for 10 min on ice, and blocked with 5 mg/mL BSA in PBS for 1 h on ice. METTL16 was visualized with an anti-METTL16 rabbit polyclonal antibody (Sigma, HPA020352, 1:200) followed by anti-rabbit Alexa Fluor 594 antibody (Invitrogen). Nuclei were stained with DAPI. Images were collected on a Leica TCS SP5 confocal microscope (Yale Center for Cellular and Molecular Imaging).

CLIP-qPCR. One 15-cm plate of HeLa cells at ∼80% confluency was rinsed with PBS, transferred to a bed of ice, and UV irradiated 2 × 400 mJ·cm^2 by using a Stratalinker. The nuclear-enriched fraction was prepared as described above for the coimmunoprecipitation experiment except that nuclei were resuspended in RIPA buffer. The immunoprecipitation was performed as described above except that the RNase A treatment was omitted. Beads were washed five times with RIPA buffer at 850 mM KCl. Immunoprecipitation was verified by Western blot using anti-mouse METTL16 antibody (Abcam, 128528, 1:1,000). For qPCR, RNA samples were treated with Proteinase K (Amresco), isolated using TRIzol (Life Technologies), treated with RQ1 DNase (Promega), and cDNA was synthesized with SuperScript III RT (Invitrogen) and random hexamer primers (Invitrogen). Next, cDNA was added to FastStart Essential DNA SYBR Green Master Mix and analyzed by using a LightCycler Real-Time PCR System (Roche). Each run included three technical replicates for the following: MALAT1 (forward, 5′-GATCTAGCACAGACCCTTCAC-3′; reverse, 5′-CGACACCATCGTTACCTTGA-3′), MENβ, (forward, 5′-GTGTCCACAGGTCTTAGATTCC-3′; reverse, 5′-TCTGTGTAGTAGGGTGGGATAG-3′), HOTAIR, (forward, 5′-GGTAGAAAAAGCAACCACGAAGC-3′; reverse, 5′-ACATAAACCTCTGTCTGTGAGTGCC-3′), 5.8S (forward, 5′-GGTGGATCACTCGGCTCGT-3′; reverse, 5′-GCAAGTGCGTTCGAAGTGTC-3′), 18S (forward, 5′-CCCAGTAAGTGCGGGTCATAA-3′; reverse, 5′-GATCCGAGGGCCTCACTAAC-3′), and 28S (forward, 5′-TCATCAGACCCCAGAAAAGG-3′; reverse, 5′-GATTCGGCAGGTGAGTTGTT-3′).

For experiments using the β-globin reporter system, HEK293T cells were transfected with the previously described βΔ1,2 constructs containing the human MALAT1 ENE+A-rich tract+mascRNA (nucleotides 8254–8424) and human MENβ ENE+A-rich tract+menRNA (nucleotides 22643–22812) (5) by using TransIT-293 (Mirus Corporation) per the manufacturer's protocol. Cells were harvested 24 h after transfection, and CLIP-qPCR was performed as described above. Sequences of qPCR primers are as follows: βΔ1,2 (forward, 5′-CAAGAAAGTGCTCGGTGCCT-3′; reverse, 5′-AATTCTTTGCCAAAGTGATGGG-3′) and NeoR, (forward, 5′-CTTGGGTGGAGAGGCTATTC-3′; reverse, 5′-GTCGGTCTTGACAAAAAGAACC-3′).

ACKNOWLEDGMENTS. We thank Kazio Tycowski and Johanna Withers for critical review of the manuscript, Angela Miccinello for editorial work, and all J.A.S. laboratory members for thoughtful discussions. This work was supported by NIH Grants K99GM111430 (to J.A.B.) and GM26154 and CA16038 (to J.A.S.). J.A.S. is an Investigator of the Howard Hughes Medical Institute.

1. Conrad NK, Steitz JA (2005) A Kaposi's sarcoma virus RNA element that increases the nuclear abundance of intronless transcripts. *EMBO J* 24(10):1831–1841.
2. Conrad NK, Mili S, Marshall EL, Shu MD, Steitz JA (2006) Identification of a rapid mammalian deadenylation-dependent decay pathway and its inhibition by a viral RNA element. *Mol Cell* 24(6):943–953.
3. Mitton-Fry RM, DeGregorio SJ, Wang J, Steitz TA, Steitz JA (2010) Poly(A) tail recognition by a viral RNA element through assembly of a triple helix. *Science* 330(6008):1244–1247.
4. Tycowski KT, Shu MD, Borah S, Shi M, Steitz JA (2012) Conservation of a triple-helix-forming RNA stability element in noncoding and genomic RNAs of diverse viruses. *Cell Reports* 2(1):26–32.
5. Brown JA, Valenstein ML, Yario TA, Tycowski KT, Steitz JA (2012) Formation of triple-helical structures by the 3′-end sequences of MALAT1 and MENβ noncoding RNAs. *Proc Natl Acad Sci USA* 109(47):19202–19207.
6. Brown JA, et al. (2014) Structural insights into the stabilization of MALAT1 noncoding RNA by a bipartite triple helix. *Nat Struct Mol Biol* 21(7):633–640.
7. Wilusz JE, et al. (2012) A triple helix stabilizes the 3′ ends of long noncoding RNAs that lack poly(A) tails. *Genes Dev* 26(21):2392–2407.
8. Tycowski KT, Shu MD, Steitz JA (2016) Myriad triple-helix-forming structures in the transposable element RNAs of plants and fungi. *Cell Reports* 15(6):1266–1276.
9. Lukong KE, Chang KW, Khandjian EW, Richard S (2008) RNA-binding proteins in human genetic disease. *Trends Genet* 24(8):416–425.
10. Lunde BM, Moore C, Varani G (2007) RNA-binding proteins: Modular design for efficient function. *Nat Rev Mol Cell Biol* 8(6):479–490.
11. Gleghorn ML, Maquat LE (2014) 'Black sheep' that don't leave the double-stranded RNA-binding domain fold. *Trends Biochem Sci* 39(7):328–340.
12. Brown JA, Kinzig CG, DeGregorio SJ, Steitz JA (2016) Hoogsteen-position pyrimidines promote the stability and function of the MALAT1 RNA triple helix. *RNA* 22(5):743–749.
13. Hutchinson JN, et al. (2007) A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. *BMC Genomics* 8:39.
14. Sergiev PV, Serebryakova MV, Bogdanov AA, Dontsova OA (2008) The ybiN gene of Escherichia coli encodes adenine-N6 methyltransferase specific for modification of A1618 of 23 S ribosomal RNA, a methylated residue located close to the ribosomal exit tunnel. *J Mol Biol* 375(1):291–300.
15. Xie M, et al. (2015) The host Integrator complex acts in transcription-independent maturation of herpesvirus microRNA 3′ ends. *Genes Dev* 29(14):1552–1564.
16. Söderberg O, et al. (2006) Direct observation of individual endogenous protein complexes in situ by proximity ligation. *Nat Methods* 3(12):995–1000.
17. Vasilyev N, et al. (2015) Crystal structure reveals specific recognition of a G-quadruplex RNA by a β-turn in the RGG motif of FMRP. *Proc Natl Acad Sci USA* 112(39):E5391–E5400.
18. Dorsett M, Westlund B, Schedl T (2009) METT-10, a putative methyltransferase, inhibits germ cell proliferative fate in Caenorhabditis elegans. *Genetics* 183(1):233–247.
19. Dorsett M, Schedl T (2009) A role for dynein in the inhibition of germ cell proliferative fate. *Mol Cell Biol* 29(22):6128–6139.
20. Xu C, Yang M, Tian J, Wang X, Li Z (2011) MALAT-1: A long non-coding RNA and its important 3′ end functional motif in colorectal cancer metastasis. *Int J Oncol* 39(1):169–175.
21. Gutschner T, et al. (2013) The noncoding RNA MALAT1 is a critical regulator of the metastasis phenotype of lung cancer cells. *Cancer Res* 73(3):1180–1189.
22. Gutschner T, Hammerle M, Diederichs S (2013) MALAT1 - a paradigm for long noncoding RNA function in cancer. *J Mol Med* (Berl) 91(7):791–801.
23. Arun G, et al. (2016) Differentiation of mammary tumors and reduction in metastasis upon Malat1 lncRNA loss. *Genes Dev* 30(1):34–51.
24. Lee T, Pelletier J (2016) The biology of DHX9 and its potential as a therapeutic target. *Oncotarget* 7(27):42716–42739.
25. Linder B, et al. (2015) Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nat Methods* 12(8):767–772.
26. Khatter H, Myasnikov AG, Natchiar SK, Klaholz BP (2015) Structure of the human 80S ribosome. *Nature* 520(7549):640–645.
27. Guo YE, Steitz JA (2014) 3′-Biotin-tagged microRNA-27 does not associate with Argonaute proteins in cells. *RNA* 20(7):985–988.
28. Kataoka N, Dreyfuss G (2004) A simple whole cell lysate system for in vitro splicing reveals a stepwise assembly of the exon-exon junction complex. *J Biol Chem* 279(8):7009–7013.