

Identification of Endogenous Sequences Interacting with METTL3/METTL14 RNA Methyltransferase

Tamaki Endoh,^[b,c] Yiwei Ling,^[d] Shujiro Okuda,^[d] and Miki Imanishi^{*[a]}

- [a] Dr. M. Imanishi
Institute for Chemical Research
Kyoto University
Gokasho, Uji, Kyoto 611-0011, Japan
E-mail: imiki@scl.kyoto-u.ac.jp
- [b] Dr. T. Endoh
Graduate School of Frontiers of Innovative Research in Science and Technology (FIRST)
Konan University
7-1-20 Minatojima-Minamimachi, Chuo-ku, Kobe 650-0047, Japan
- [c] Dr. T. Endoh
Frontier Institute for Biomolecular Engineering Research (FIBER)
Konan University
7-1-20 Minatojima-Minamimachi, Chuo-ku, Kobe 650-0047, Japan
- [d] Dr. Y. Ling and Prof. S. Okuda
Medical AI Center, School of Medicine
Niigata University
Niigata City, Niigata, 951-8514, Japan

Supporting Information for this article is provided via a link at the end of the document.

Abstract: *N*6-methyladenosine (*m*⁶A) is the most abundant RNA modification in mRNA and regulates various biological processes. The RNA-binding properties of *m*⁶A writer proteins play an important role in determining RNA modification sites. METTL3 and METTL14 form the core of the *m*⁶A writer complex, with METTL3 as the catalytic methyltransferase and METTL14 as the RNA-binding scaffold. Thus far, the comprehensive RNA binding properties of METTL3/14 remain unknown. Using RNA-capturing microsphere particles (R-CAMPs), immobilizing RNA fragments derived from endogenous RNAs of human lung carcinoma cells, we isolated RNA fragments that interacted with the METTL3/14 methyltransferase domain. Bioinformatics analysis revealed that the pool of isolated sequences contained significantly more regions with the potential to form RNA G-quadruplexes (rG4s) than the randomly extracted RNA sequence pool and that the (GGA) repeat sequences were most enriched. CD spectroscopy, gel mobility shift assays, and methylation experiments demonstrated that METTL3/14 binds to RNA sequences containing (GGA) repeats that form rG4 structure much stronger than RNAs that do not form rG4s. This study shows the potential of RNA G-quadruplex structures as a modulator of epitranscriptomic modification of *m*⁶As.

*N*6-methyladenosine (*m*⁶A) is the most abundant modification in mRNA and regulates various physiological processes, such as cell differentiation^[1] and carcinogenesis.^[2] METTL3 and METTL14, the core of the writer complex for *m*⁶A, form a stable heterodimer, METTL3/14, in which METTL3 plays a catalytic role as a methyltransferase and METTL14 acts as a scaffold for RNA binding.^[3-5] The writer complex predominantly localizes in the nucleus and installs a methyl moiety to the *N*⁶ position of adenosine in a consensus "DRACH (D=A/G/U, R=A/G, H=A/C/U)" sequence through interaction with histone H4K36me3, RNA polymerase II, and various RNA binding proteins.^[6, 7] However, only a part of DRACH sequences is methylated.^[8, 9] Recent studies have provided a key mechanism of *m*⁶A deposition, in which Exon Junction Complexes exclude *m*⁶A deposition from splice-site-proximal RNA regions, explaining the methylation reaction that occurs co-transcriptionally in the nucleus.^[10-12] However, viral RNAs that enter from outside cells and are present only in the cytoplasm are also known to be methylated by METTL3/14,^[13] indicating the possibility of different *m*⁶A deposition pathways from co-transcriptional installation. Therefore, it is important to analyze the alternative mechanisms of RNA methylation in the cytoplasm. Because METTL3/14 can methylate RNAs *in vitro* without the supportive proteins responsible for nuclear localization, a direct interaction between

Introduction

METTL3/14 and RNA may be one of the most important factors for m⁶A deposition in the cytoplasm.

RNA G-quadruplexes (rG4s) are non-canonical four-stranded RNA structures formed in G-rich sequences, 5'-G_nL_xG_nL_xG_nL_xG_n-3' (n ≥ 2, 1 ≤ x ≤ 20, L = A, G, C or U), in which four or more runs of consecutive guanines (Gs) are separated by loops. rG4 is stabilized by π-π interactions between the G-tetrads formed by Hoogsteen hydrogen bonds and coordination with potassium ions.^[14] rG4 regulates gene expression at the post-transcriptional level, such as translation, alternative splicing, and polyadenylation, by blocking or enhancing interactions with RNA-binding proteins.^[15] Repeated rG4s can cause neurodegenerative diseases by forming aggregates engulfing RNA-interacting proteins.^[16, 17] Thus, the interactions between rG4 and rG4-binding proteins are important in various biological processes. However, to date, knowledge of the roles of rG4 structures and rG4-binding proteins is limited, especially from the viewpoint of epitranscriptomic RNA modifications.

We have previously shown that the RNA-binding domain of METTL14, a core RNA methyltransferase complex component, has the potential to bind to rG4 structures. In addition, the catalytic domain of METTL3/14 methylated adenosine close to rG4 *in vitro*.^[18] According to a previous report, m⁶A positions are well-localized with G-quadruplex-forming sequences in some viral RNAs, including ZIKV and HIV.^[19] These results imply that the interaction between rG4 and METTL3/14 is involved in m⁶A deposition, which regulates the epitranscriptomic chemical modification of RNAs. However, the interaction between METTL3/14 and rG4 has only been demonstrated using model sequences that form stable rG4s comprising three stacked G-tetrads.^[18] With taking account for the structural variability of rG4s depending on the length of loops and their nucleobase compositions, endogenous rG4s that potentially interact with METTL3/14 remain unknown.

To identify possible endogenous RNAs that directly interact with METTL3/14, we used RNA-capturing microsphere particles (R-CAMPs), each of which immobilizes a unique sequence of DNA and RNA on its surface.^[20-24] Using a protein of interest with a hexahistidine tag (His₆-tag), which can be conjugated with fluorophore-labeled nickel-nitrilotriacetic acid (NTA), R-CAMPs interacting with the labeled protein were efficiently isolated by particle sorting.^[22] In the current study, we isolated RNA fragments that interacted with the MTD3/14 heterodimer, which contains the methyltransferase domains (MTDs) of METTL3 and METTL14 and the arginine-glycine-rich RNA binding domain of METTL14 (Figure S1). We achieved this using R-CAMPs that immobilize RNA fragments derived from the total RNA of human lung carcinoma A549 cells. rG4-forming sequences were the most enriched sequences that interacted with MTD3/14. This implies a strong relationship between rG4 structures and epitranscriptome formation.

R-CAMPs comprise millions of particles that immobilize fragments of both RNA and template DNA. Each particle immobilized the same sequences of RNA and template DNA, and different particles immobilized different sequences derived from the initial library. R-CAMPs were prepared from a fragmented RNA library derived from A549 cells. The particles were isolated by fluorescence-activated cell sorting (FACS) based on their fluorescence signals. The heterodimer of the RNA methyltransferases MTD3/14 was labeled with a fluorophore through an interaction between the His₆-tag at the N-terminus of the MTD14 construct and the NTA unit in NTA-ATTO647N. If RNA fragments immobilized on a particle interact with fluorophore-labeled MTD3/14, the particle should show strong fluorescence signals and be isolated in a collection tube (Figure 1A). Figure 1B shows the fluorescent signals of R-CAMPs mixed with ATTO647N-conjugated MTD3/14. Two control samples were also prepared. The first was R-CAMPs mixed with NTA-ATTO647N alone in the absence of MTD3/14 (no-protein control). The other comprised control particles, which only immobilized template DNA fragments for RNA transcription, mixed with ATTO647N-conjugated MTD3/14 (no-RNA control). The fluorescence histogram of R-CAMPs mixed with ATTO647N-conjugated MTD3/14 shifted toward higher fluorescence than that of the control samples, suggesting that the RNAs immobilized on the particles tended to interact with MTD3/14. The no-RNA control showed slightly higher fluorescence than the no-protein control, suggesting the possibility of a weak interaction between MTD3/14 and DNA duplexes on the particles. R-CAMPs showing strong fluorescence are targets for isolation that immobilize RNA fragments interacting with MTD3/14 with high affinity. When the sorting gate was set at a region with a fluorescence intensity greater than approximately 4000 a.u., 177 particles (0.23% of total particles analyzed) were present in the gate for R-CAMPs mixed with the ATTO647N-conjugated MTD3/14. The number of particles for the no-protein and no-RNA controls was 4 (0.0052% of total particles analyzed) and 9 (0.011% of total particles analyzed), respectively (Figure 1C). Finally, 740 particles in the sorting gate were isolated using the R-CAMPs sample mixed with the ATTO647N-conjugated MTD3/14.

Results and Discussion

Screening of endogenous RNA interacting with MTD3/14

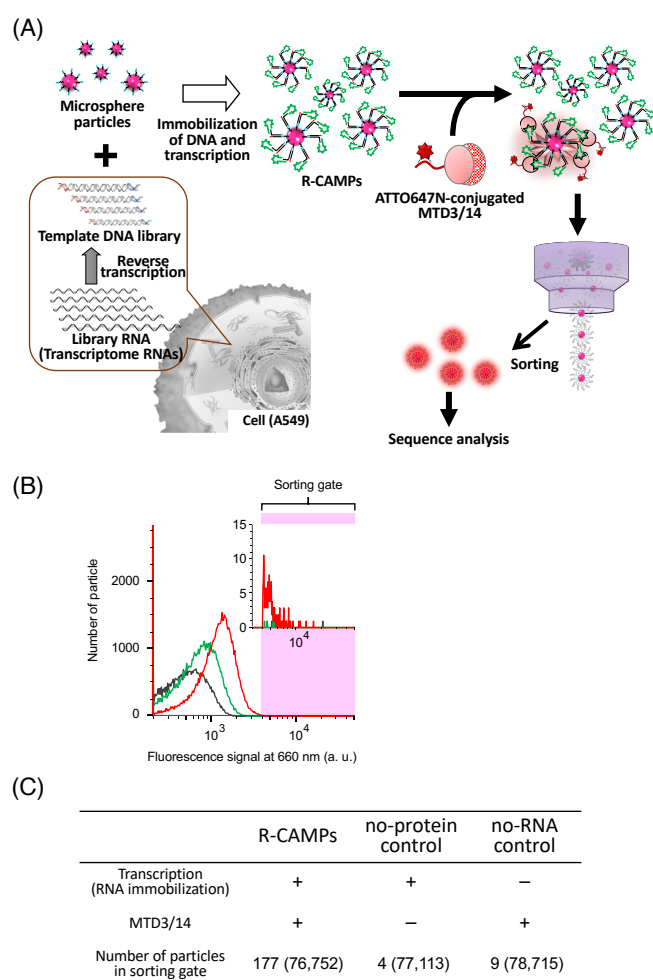


Figure 1. Isolation of RNA fragments interacting with MTD3/14. (A) Process of RNA fragments isolation using R-CAMPs. RNA fragments derived from A549 cells were displayed on microsphere particles to prepare R-CAMPs. After incubating R-CAMPs with ATTO647N-conjugated MTD3/14, R-CAMPs with high fluorescence intensity were sorted and the RNA sequences were analyzed. (B) Distribution of fluorescence signals of the particles. Fluorescence signals of R-CAMPs incubated with ATTO647N-conjugated MTD3/14 (red), no-protein control, in which R-CAMPs were incubated only with NTA-ATTO647N dye without MTD3/14 (black), and no-RNA control particles incubated with ATTO647N-conjugated MTD3/14 (green) were analyzed using a 633 nm excitation laser and a 660 nm emission filter. The distribution of fluorescence signals of particles within a region of the sorting gate is enlarged within the panel. (C) Number of particles within the sorting gate (highlighted in pink in (B)). No-RNA control particles only presented the DNA template sequence but no RNA. Numbers within parentheses represent the total number of particles analyzed.

Analysis of the isolated RNA sequences

The isolated particles contained multiple copies of the template DNA fragments encoding the RNA sequence that interacted with MTD3/14. The sequences were analyzed using next-generation sequencing (NGS) after amplification of the template DNAs. The NGS data were mapped to the reference human transcriptome

(Figure S2A, Table S1). When a hit sequence corresponded to more than one splicing variant, the sequence was mapped to the longest splicing variant. Among approximately 100,000 reads, 76,975 reads were successfully aligned to the reference transcriptome with a mapping identity threshold of 0.95. There were regions in which numerous NGS reads were mapped and those in which only a few reads were mapped (Figure S2A). The number of regions in which more than one read was mapped was higher than 4000, i.e., much higher than the number of isolated particles (Figure S2B). The NGS reads mapped to the same region with a high coverage number were considered derived from isolated particles. In contrast, the regions in which few NGS reads were mapped were considered derived from contamination of DNA fragments released into the solution during the R-CAMPs isolation process. The number of regions quickly decreased when the minimum coverage number was 5 and then gradually decreased with increasing coverage number (Figure S2B). To analyze the general properties of RNAs, which interact with MTD3/14, it is required to reliably extract the RNA fragments obtained based on the interaction of MTD3/14. Here, we treated contiguous regions with a minimum coverage number of 10 or more for subsequent analysis. There were 285 contiguous regions, with an average length of 163 nt (Table S1).

The characteristics of the isolated RNA fragments were analyzed by comparing them to the control RNA pool, the sequences of which were randomly extracted from the transcriptome with the same number and length as the isolated RNA fragments, i.e., 285 fragments with a length of 163 nt (Table S1). Motif enrichment analysis, STREME, which discovers *de-novo* motifs enriched in the input sequences against the separately provided set of control sequences,^[25] indicated the “RGAGGAGGAGGAGGA” (R=G/A) motif as the most enriched sequence in the isolated RNA fragments (Figure 2A). Namely, 132 sites (46.3%) of GGA repeat sequences were found in the 285 isolated RNA fragments, whereas only 10 sites (3.5%) of these sequences were found in the control RNA pool. Consistent with the fact that the (GGA)_n sequence is a possible G-quadruplex-forming sequence,^[26] the number of possible rG4 sequences based on the QGRS mapping algorithm^[27] in the isolated RNA fragments was significantly higher than in the control pool (Figure 2B).

Previous PAR-CLIP-seq demonstrated that RNAs containing the methylation consensus sequence DRACH co-precipitated with METTL3 and METTL14.^[6] Despite the significant difference in the number of potential rG4-forming sequences, the number of DRACH consensus sequences was 19.7 and 19.6 per 1 kb for the isolated RNA fragments and control RNA pool, respectively (Figure 2C). Statistically, the DRACH sequence motif occurs per every 57 nucleotides, i.e., 17.6 per 1 kb. Thus, the results shown in Figure 2C indicate that the frequency of DRACH was unbiased through the isolation of RNA fragments with MTD3/14, suggesting that the interaction between MTD3/14 and DRACH was not dominant in RNA isolation. Because the abundance of RNA fragments in the R-CAMPs library depends on the levels of each RNA transcripts in cells, it would be difficult to isolate RNA region, where the expression level is extremely low in the original RNA library. Regardless, these results suggest that

RESEARCH ARTICLE

the MTD3/14 heterodimer is more likely to be isolated through interaction with G-quadruplex-forming sequences.

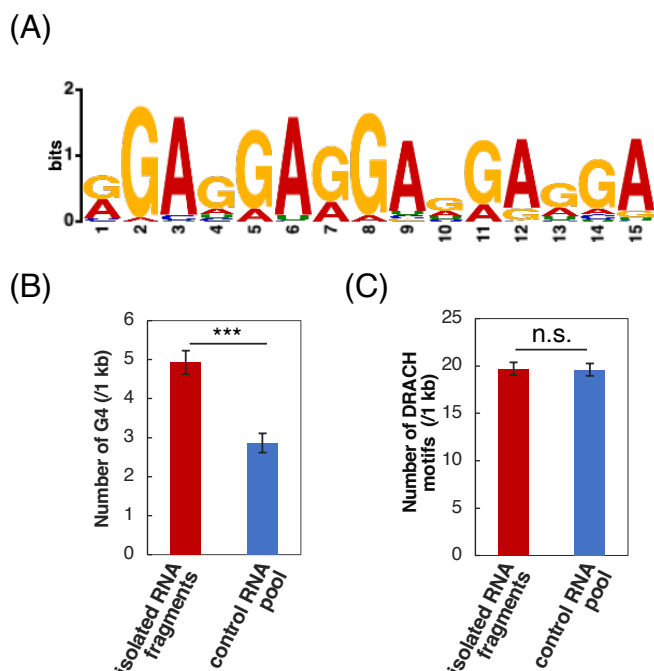


Figure 2. Analysis of isolated RNA sequences interacting with MTD3/14 and the control RNA pool. (A) Sequence motif enriched in the isolated RNA sequences. Motif enrichment analysis was performed using STREME.^[25] (B)(C) The number of potential G-quadruplex forming sequences (B) or DRACH sequences (C) in the isolated RNA sequences (red) or control RNA pool (blue). The numbers were standardized per 1 kb. Data are shown as mean $\pm$ SEM (n=285). ***; $p < 0.01$, n.s.; not significant (Student's t-test).

RNA binding and methylation activity of MTD3/14 to the isolated RNA fragments

The RNA-binding and methylation abilities of MTD3/14 on the isolated RNA fragments were examined *in vitro*. The short RNA oligonucleotide derived from the isolated fragment of calreticulin (CALR) mRNA (CALR RNA oligo) was prepared with modification of FAM at its 5' end (Figure 3A). The sequence was selected according to the following criteria: (1) the possible rG4 forming region was close to the possible m⁶A site, based on the human m⁶A modification database, RMBase v2.0,^[28] and (2) the methylation site was within a DRACA sequence so that the evaluation method of adenosine methylation using the ACA-specific endoribonuclease MazF could be applied.^[29] The CD spectra of the CALR RNA oligo demonstrated a positive peak around 265 nm and a negative peak around 240 nm, typical of parallel G4 structures,^[30] in the presence of potassium ions, but not lithium ions (Figure 3B). This indicated that the CALR RNA oligo, including multiple GGA sequences, formed a parallel G4 structure with two tetrads in the presence of potassium ions. The binding affinity of MTD3/14 for the CALR RNA oligo was approximately 2.7-fold higher in the presence of potassium ions

than in the presence of lithium ions (Figure 3C). We tested another RNA oligonucleotide derived from the isolated fragment of apurinic/apyrimidinic endodeoxyribonuclease 1 (APEX1) mRNA (APEX1 RNA oligo). The APEX1 RNA oligo also has the potential to form a G-quadruplex consisting of two tetrads. Similar to the CALR RNA oligo, the CD spectra of the APEX1 RNA oligo indicated potassium ion-dependent parallel G4 RNA structure formation (Figure S3A, B). The binding affinity of MTD3/14 for the APEX1 RNA oligo was higher in the presence of potassium ions than in the presence of lithium ions (Figure S3C). These results indicate that MTD3/14 recognizes the two-tetrad G4 RNA structures derived from endogenous mRNA sequences.

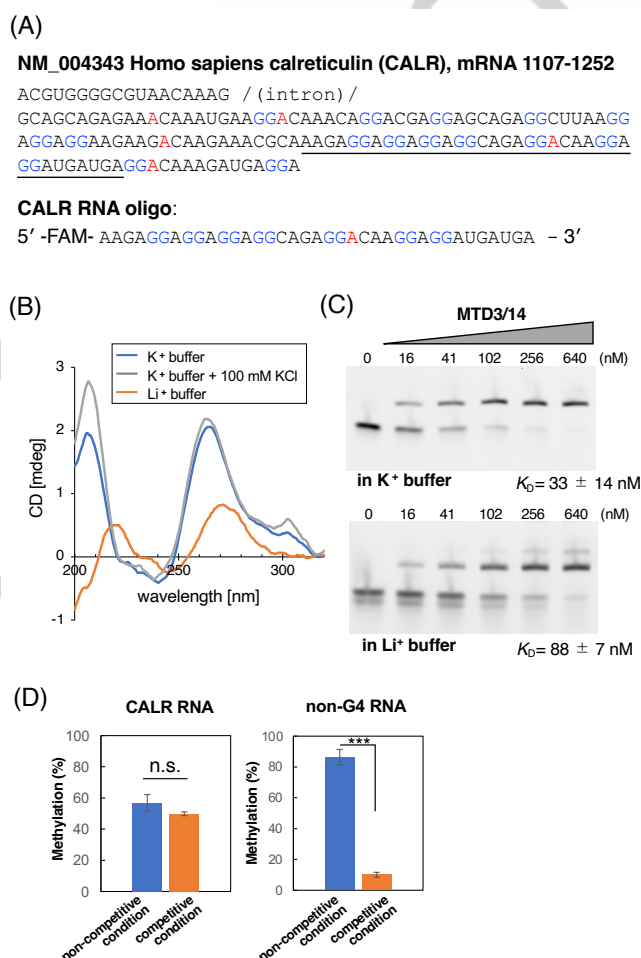


Figure 3. RNA binding ability of MTD3/14 to the isolated RNA fragment. (A) Sequence of the isolated RNA fragment in CALR-derived mRNA and that of the CALR RNA oligo used for *in vitro* experiments. Consecutive guanines potentially involved in G-quadruplex formation and adenine in the consensus DRACH sequence are highlighted in blue and red, respectively. (B) CD spectra of the CALR RNA oligo under the indicated buffer conditions. (C) Results of the electrophoretic mobility shift assay for MTD3/14 and the CALR RNA oligo in potassium or lithium buffer. (D) RNA methylation ratio of CALR RNA oligo (left) or non-G4 RNA oligo (right) under non-competitive (blue) or competitive

(orange) conditions. Data are shown as mean $\pm$ SD. ***; $p < 0.01$, n.s.; not significant (Student's t-test).

Preferable methylation of adenosines in the consensus GGACA sequences close to the rG4 structures in the CALR RNA oligo by MTD3/14 was confirmed under competitive conditions. Non-G4 RNA oligo (5'-UGAGUGUUAGGGACAUGAGUGUGAGUG-3') was used as the control. Under non-competitive conditions, in which only one of the RNA substrates (CALR oligo or non-G4 oligo) was present with MTD3/14 in the reaction mixture for methylation, approximately 60% and 80% of CALR and non-G4 oligos were methylated, respectively. The methylation level of the CALR oligo was maintained even under competitive conditions, in which both of the RNA substrates along with yeast-extracted tRNA were present with MTD3/14 in the reaction mixture. In contrast, the methylation level of the non-G4 RNA oligo significantly decreased under competitive conditions (Figure 3D). The APEX1 RNA oligo was also methylated by MTD3/14 under competitive conditions (Figure S3D). In contrast to our previous study, in which only stable model-rG4 sequences with three G-tetrads were used,^[18] the current results demonstrated that MTD3/14 binds to naturally occurring rG4 sequences with two-tetrads and exhibits catalytic activity in the vicinity. Since multiple pathways for mRNA methylation by METTL3/14 have been demonstrated,^[6, 7, 10-12, 31] the physiological significance of adenosine methylation near rG4 structures in these mRNAs remains challenging to elucidate. However, by considering the dynamic behaviors of rG4s, including those influenced by methylation,^[32] our results support the idea that recognition of the rG4 structure of MTD3/14 is one of the key pathways involved in the modulation of adenosine methylation and subsequent gene expression.

Conclusions

Using R-CAMPs, we demonstrated that MTD3/14, responsible for the RNA binding and methyltransferase activity of METTL3/14, preferentially interacted with G-quadruplex-forming RNA sequences within various RNA sequences derived from human lung carcinoma cells. Since sequences rich in GGA trinucleotides were well-conserved in the isolated RNA fragments, MTD3/14 has the potential to interact with G-quadruplexes comprising two tetrads. Additionally, adenosines located in proximity to G-quadruplex-forming sequences are preferably methylated in the presence of non-specific RNA molecules. Although *in cellulo* experiments are required, the rG4 binding preference of METTL3/14 may contribute to the colocalization of m⁶A and rG4, for example, in ZIKV genomic RNA, which undergoes methylation in the cytoplasm. These results provide new insights into a possible mechanism of epitranscriptome formation: direct binding of METTL3/14 to rG4 and subsequent adenosine methylation.

Experimental Section

Protein preparation: MTD3/14 heterodimer was prepared according to a previous study^[18] using the MTD3/MTD14-RGG expression vector (Strep

tag II-METTL3(255-580)/hexa-histidine tag-METTL14(111-456) in the pETDuet vector), except that the StrepTrapXT column (Cytiva) was used instead of the StrepTrapHT column (Cytiva). For the StrepTrap XT column, the following buffers were used: binding and wash buffer, 10 mM potassium phosphate, and 100 mM KCl (pH 7.5); elution buffer, 10 mM potassium phosphate, 100 mM KCl, and 50 mM biotin (pH 7.5). The purified proteins were kept on ice and used for experiments on the same day.

Preparation of RNA-capturing microsphere particles: RNA-capturing microsphere particles (R-CAMPs) were prepared according to previous reports.^[20, 22, 24] Briefly, total RNA of A549 cells was obtained using an RNeasy Mini Kit (Qiagen) and a fragmented cDNA library was synthesized from the total RNA using SMARTer Stranded Total RNA-Seq Kit v2 - Pico Input Mammalian (Takara Bio) according to the manufacturer's protocol. The cDNA library was amplified by PCR using sense (5'-CCATCTCATCCCTGCGTGTCTCCGACTCAGTAATACGACTCACTATA GGGTTCAGACGTGTGCTCTTCCGAT-3') and antisense (5'-CCTCTCTATGGGCAGTCGGTGATACACTCTTCCCTACACGACGC-3') primers and purified using a QIAquick PCR Purification Kit (Qiagen) followed by size selection using an E-Gel SizeSelect II Agarose Gel (Thermo Fisher Scientific) to obtain fragments at approximately 300 to 350 base pairs. Single copies of DNA derived from the library were amplified and immobilized on the microsphere particles by emulsion PCR using the Ion PGM Hi-Q View OT2 Kit (Thermo Fisher Scientific) and Ion OneTouch 2 system (Thermo Fisher Scientific), according to the manufacturer's protocol. Double stranded template DNAs for RNA transcription were synthesized on the microsphere particles by primer extension using a capture primer, ATCGGAAGAGACACGTCGAACCC-link-CAGTAATACGACTCACTATAGGG, in which "-link-" indicates oligo ethylene glycol linker. R-CAMPs were prepared by capturing nascent RNAs on a single-stranded tail that remained on the capture primer during transcription. This process ensured the presence of DNA and RNA with the same sequence on the same particle. The buffer solution, in which R-CAMPs were dispersed, was replaced with 50 mM MES-LiOH (pH 7.0), 100 mM KCl, 0.5 mM MgCl₂, and 0.01% Tween-20, and kept at -80 °C until the selection. No-RNA control particles, which did not immobilize RNA, were prepared using the same procedure, except for the transcription reaction.

Isolation of R-CAMPs interacting with MTD3/14: R-CAMPs and non-RNA control particles were dissolved in a buffer containing 25 mM Tris-HCl (pH 7.5), 100 mM KCl, and 0.01% Tween 20, and kept at room temperature for at least 30 min to form RNA structures on the particles. Purified MTD3/14 was added to the particles at a concentration of 100 nM with 25 μ M ZnCl₂, 10 μ M S-adenosyl-methionine, 100 nM NTA-ATTO647N (Sigma), and 50 ng/ μ L yeast tRNA, and samples were further incubated for at least 30 min. A sample without MTD3/14 (no-protein control) was also prepared. The sample solution contained 1 mM MES and 0.01 mM MgCl₂ derived from a stock solution of R-CAMPs. R-CAMPs that immobilize RNA interacting with MTD3/14 were selected according to a previous report.^[22] Fluorescence signals of R-CAMPs and control particles at 660 nm were analyzed using FACSaria II (BD Biosciences) with a 633-nm excitation laser. Single particles were distinguished by the fluorescence signal of the particles themselves at 585 nm upon 488-nm excitation. From the R-CAMPs sample mixed with ATTO647N-conjugated MTD3/14, particles showing a strong fluorescence signal at 660 nm were sorted into a sample tube containing 10 μ L of H₂O.

Preparation of DNA template for sequence analysis of isolated RNAs: Sample tube containing sorted R-CAMPs was sonicated twice with 10-sec pulse sonication and incubated at 95 °C for 15 min. DNAs in the solution were filtered with 0.2 μ m membrane filter (Millipore) and amplified by PCR using A primer (5'-CCATCTCATCCCTGCGTGTCTCCGACTCAG-3') and B-trP1 primer (5'-CTATCCCTGTGTGCCTTGGCAGTCTCAGCCTCTCTATGGGCAGTC GGTGAT-3'). DNA amplification was stopped at the logarithmic amplification phase, which was preliminarily determined using an

Mx3000P qPCR system (Agilent Technologies). The DNA sample was treated with 100 U of Exonuclease I (New England Biolabs) at 37 °C for 30 min and purified using a QIAquick PCR Purification Kit (Qiagen). The concentration of the purified DNA sample was quantified using the Mx3000P qPCR system with A primer and B-trP1 primer by comparison with known concentrations of a standard DNA fragment.^[20]

Analysis of isolated RNAs: Sequences of the purified DNAs were analyzed using the Ion PGM System (Thermo Fisher Scientific) according to a standard protocol. Obtained single-end reads were trimmed using 5' adaptor (5' TAATACGACTCACTATAGGTTTCAGACGTGTGCTCTTCCGATCT-3') and 3' adaptor (5' AGATCGGAAGAGCGTCGTGTAGGAAAGAGTGT-3') by CLC Genomics Workbench (Qiagen). Reads lacking the 3' adapter region were discarded, and only those likely to represent full-length sequences of isolated RNA fragment were retained. Exported FASTQ data, in which over 99.8% of bases had a Phred quality score $\geq$ Q20 with the peak mean Phred score at 29, were mapped to the reference human genome GRCh37 (hg19), which contains 72,567 mRNA sequences in autosomes and the X chromosome, using Minimap2 (ver. 2.24-r1122). Mapped reads with an identity threshold of 0.95 and exceeding 50 nucleotides in length were used for coverage analysis to differentiate fragments contaminated by the liberation of DNAs from the particles into the sample solution. Regions exhibiting significantly high mapping coverage were used for further analysis because the RNA fragments were selected based on their interaction with MTD3/14.

Annotation of possible rG4 and m⁶A sites within isolated RNAs: To extract possible rG4 loci, a typical G4 prediction tool, the QGRS algorithm,^[27] was used. Additionally, m⁶A sites were extracted from the RMBase database (version 2.0)^[28] and matched to the sequences of the selected RNA fragments.

CD measurements: Synthesized RNA oligonucleotides (Japan Bio Service) were dissolved in 10 mM potassium phosphate (pH 7.5), 10 mM potassium phosphate with additional 100 mM KCl, or 10 mM lithium phosphate (pH 7.5), heated at 95 °C for 5 min followed by gradual cooling to room temperature, and kept on ice until use. RNA oligonucleotides (2 μ M) were measured using a J-820-L CD spectrometer (JASCO), using a 0.1 cm path length cell, with a scanning speed of 100 nm/min, and a response time of 0.5 s over a wavelength range of 200–300 nm.

Electrophoretic mobility shift assay: 5'FAM-labeled RNA oligonucleotides were annealed before use, as described above. MTD3/14 was incubated with 5 nM RNA oligo in binding buffer [10 mM potassium phosphate (pH 7.5), 25 μ M ZnCl₂, 0.5 mM DTT, 0.005% Tween 20, and 2.5% ficoll] in the presence of 15 ng/ μ L of yeast tRNA (Thermo Fisher Scientific) at 4 °C for 1 h. For the assays in the presence of Li⁺, 10 mM lithium phosphate (pH 7.5) was used instead of potassium phosphate. Each sample was loaded onto a native-polyacrylamide gel and electrophoresed in 0.5 X TBE buffer at 4 °C.

RNA methylation assay: 500 nM MTD3/14 was mixed with FAM-labeled RNA oligo (CALR RNA oligo or APEX1 RNA oligo) or TAMRA-labeled Non-G4 RNA, respectively, in the methylation buffer [10 mM potassium phosphate (pH 7.5), 25 μ M ZnCl₂, 10 μ M S-adenosylmethionine (SAM), 0.5 mM DTT, 0.005% Tween 20, and 2.5% ficoll]. For competitive methylation assays, MTD3/14 was mixed with RNA oligo in the methylation buffer with 15 ng/ μ L yeast tRNA. The reaction mixture was incubated at 25 °C for 30 min and then heated at 95 °C for 3 min. The samples were subjected to MazF reaction at 37 °C for 30 min and loaded onto a denaturing polyacrylamide gel. The FAM or TAMRA-labeled RNAs were visualized using Amercham Typhoon (GE healthcare) and the fraction of uncleaved RNA (methylated RNA) was measured using the ImageQuant TL software (Cytiva).

Supporting Information

Supporting Information is available.

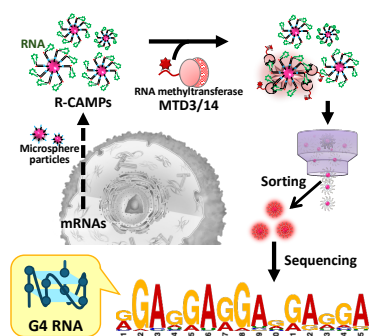
Acknowledgements

This work was supported by a Grant-in-Aid for Transformative Research Areas (B), "Dimensional Response Genome" (JSPS KAKENHI 21H05110 to M.I. and 21H05108 to T.E. and Y.L.). The authors thank Dr. Hisae Tateishi-Karimata for constructive discussions. This study was also supported by the Konan New Century Strategic Research Project. The authors declare no conflicts of interest.

Keywords: RNA methylation, G-quadruplex RNA, RNA screening, protein-RNA interaction

- [1] P. J. Batista, B. Molinie, J. Wang, K. Qu, J. Zhang, L. Li, D. M. Bouley, E. Lujan, B. Haddad, K. Daneshvar, A. C. Carter, R. A. Flynn, C. Zhou, K.-S. Lim, P. Dedon, M. Wernig, A. C. Mullen, Y. Xing, C. C. Giallourakis, H. Y. Chang, "m⁶A RNA Modification Controls Cell Fate Transition in Mammalian Embryonic Stem Cells" *Cell Stem Cell* **2014**, *15*, 707-719.
- [2] X. Deng, R. Su, H. Weng, H. Huang, Z. Li, J. Chen, "RNA N⁶-Methyladenosine Modification in Cancers: Current Status and Perspectives" *Cell Res.* **2018**, *28*, 507-517.
- [3] P. Śledź, M. Jinek, "Structural Insights into the Molecular Mechanism of the m⁶A Writer Complex" *eLife* **2016**, *5*, e18434.
- [4] P. Wang, K. A. Doxtader, Y. Nam, "Structural Basis for Cooperative Function of Mettl3 and Mettl14 Methyltransferases" *Mol. Cell* **2016**, *63*, 306-317.
- [5] X. Wang, J. Feng, Y. Xue, Z. Y. Guan, D. L. Zhang, Z. Liu, Z. Gong, Q. Wang, J. B. Huang, C. Tang, T. T. Zou, P. Yin, "Structural Basis of N⁶-Adenosine Methylation by the METTL3-METTL14 Complex" *Nature* **2016**, *534*, 575-578.
- [6] J. Liu, Y. Yue, D. Han, X. Wang, Y. Fu, L. Zhang, G. Jia, M. Yu, Z. Lu, X. Deng, Q. Dai, W. Chen, C. He, "A METTL3-METTL14 Complex Mediates Mammalian Nuclear RNA N⁶-Adenosine Methylation" *Nat. Chem. Biol.* **2014**, *10*, 93-95.
- [7] P. C. He, C. He, "m⁶A RNA Methylation: from Mechanisms to Therapeutic Potential" *EMBO J.* **2021**, *40*, e105977.
- [8] D. Dominissini, S. Moshitch-Moshkovitz, S. Schwartz, M. Salmon-Divon, L. Ungar, S. Osenberg, K. Cesarkas, J. Jacob-Hirsch, N. Amariglio, M. Kupiec, R. Sorek, G. Rechavi, "Topology of the Human and Mouse m⁶A RNA Methylomes Revealed by m⁶A-Seq" *Nature* **2012**, *485*, 201-206.
- [9] K. D. Meyer, Y. Saletore, P. Zumbo, O. Elemento, C. E. Mason, S. R. Jaffrey, "Comprehensive Analysis of mRNA Methylation Reveals Enrichment in 3' UTRs and near Stop Codons" *Cell* **2012**, *149*, 1635-1646.
- [10] Z. Luo, Q. Ma, S. Sun, N. Li, H. Wang, Z. Ying, S. Ke, "Exon-intron Boundary Inhibits m⁶A Deposition, Enabling m⁶A Distribution Hallmark, Longer mRNA Half-Life and Flexible Protein Coding" *Nat. Commun.* **2023**, *14*, 4172.
- [11] A. Uzonyi, D. Dierks, R. Nir, O. S. Kwon, U. Toth, I. Barbosa, C. Burel, A. Brandis, W. Rossmanith, H. Le Hir, B. Slobodin, S. Schwartz, "Exclusion of m⁶A from Splice-site Proximal Regions by the Exon Junction Complex Dictates m⁶A Topologies and mRNA Stability" *Mol. Cell* **2023**, *83*, 237-251.
- [12] X. Yang, R. Triboulet, Q. Liu, E. Sendinc, R. I. Gregory, "Exon Junction Complex Shapes the m⁶A Epitranscriptome" *Nat. Commun.* **2022**, *13*, 7904.
- [13] G. Lichinchi, B. S. Zhao, Y. Wu, Z. Lu, Y. Qin, C. He, T. M. Rana, "Dynamics of Human and Viral RNA Methylation during Zika Virus Infection" *Cell Host Microbe* **2016**, *20*, 666-673.
- [14] D. Bhattacharyya, G. Mirihana Arachchilage, S. Basu, "Metal Cations in G-Quadruplex Folding and Stability" *Front. Chem.* **2016**, *4*, 38.
- [15] M. M. Fay, S. M. Lyons, P. Ivanov, "RNA G-Quadruplexes in Biology: Principles and Molecular Mechanisms" *J. Mol. Biol.* **2017**, *429*, 2127-2147.
- [16] A. Cammas, S. Millevoi, "RNA G-Quadruplexes: Emerging Mechanisms in Disease" *Nucleic Acids Res.* **2017**, *45*, 1584-1595.

- [17] E. Wang, R. Thombre, Y. Shah, R. Latanich, J. Wang, "G-Quadruplexes as Pathogenic Drivers in Neurodegenerative Disorders" *Nucleic Acids Res.* **2021**, *49*, 4816-4830.
- [18] A. Yoshida, T. Oyoshi, A. Suda, S. Futaki, M. Imanishi, "Recognition of G-quadruplex RNA by a Crucial RNA Methyltransferase Component, METTL14" *Nucleic Acids Res.* **2022**, *50*, 449-457.
- [19] A. M. Fleming, N. L. B. Nguyen, C. J. Burrows, "Colocalization of m(6)A and G-Quadruplex-Forming Sequences in Viral RNA (HIV, Zika, Hepatitis B, and SV40) Suggests Topological Control of Adenosine N-6-Methylation" *Acs Cent. Sci.* **2019**, *5*, 218-228.
- [20] T. Endoh, T. Ohyama, N. Sugimoto, "RNA-Capturing Microsphere Particles (R-CAMPs) for Optimization of Functional Aptamers" *Small* **2019**, *15*, e1805062.
- [21] T. Endoh, N. Sugimoto, "Signaling Aptamer Optimization through Selection Using RNA-Capturing Microsphere Particles" *Anal. Chem.* **2020**, *92*, 7955-7963.
- [22] T. Endoh, S. Takahashi, N. Sugimoto, "Endogenous G-Quadruplex-Forming RNAs Inhibit the Activity of SARS-CoV-2 RNA Polymerase" *Chem. Commun.* **2023**, *59*, 872-875.
- [23] T. Endoh, J. H. Tan, S. B. Chen, N. Sugimoto, "Cladogenetic Orthogonal Light-Up Aptamers for Simultaneous Detection of Multiple Small Molecules in Cells" *Anal. Chem.* **2023**, *95*, 976-985.
- [24] S. Satpathi, T. Endoh, P. Podbevsek, J. Plavec, N. Sugimoto, "Transcriptome Screening Followed by Integrated Physicochemical and Structural Analyses for Investigating RNA-Mediated Berberine Activity" *Nucleic Acids Res.* **2021**, *49*, 8449-8461.
- [25] T. L. Bailey, "STREME: Accurate and Versatile Sequence Motif Discovery" *Bioinformatics* **2021**, *37*, 2834-2840.
- [26] T. Mashima, A. Matsugami, F. Nishikawa, S. Nishikawa, M. Katahira, "Unique Quadruplex Structure and Interaction of an RNA Aptamer against Bovine Prion Protein" *Nucleic Acids Res.* **2009**, *37*, 6249-6258.
- [27] O. Kikin, L. D'Antonio, P. S. Bagga, "QGRS Mapper: a Web-Based Server for Predicting G-Quadruplexes in Nucleotide Sequences" *Nucleic Acids Res.* **2006**, *34*, W676-682.
- [28] J. J. Xuan, W. J. Sun, P. H. Lin, K. R. Zhou, S. Liu, L. L. Zheng, L. H. Qu, J. H. Yang, "RMBase v2.0: Deciphering the Map of RNA Modifications from Epitranscriptome Sequencing Data" *Nucleic Acids Res.* **2018**, *46*, D327-D334.
- [29] M. Imanishi, S. Tsuji, A. Suda, S. Futaki, "Detection of N⁶-Methyladenosine Based on the Methyl-Sensitivity of MazF RNA Endonuclease" *Chem. Commun.* **2017**, *53*, 12930-12933.
- [30] S. Xu, Q. Li, J. Xiang, Q. Yang, H. Sun, A. Guan, L. Wang, Y. Liu, L. Yu, Y. Shi, H. Chen, Y. Tang, "Directly Lighting up RNA G-Quadruplexes from Test Tubes to Living Human Cells" *Nucleic Acids Res.* **2015**, *43*, 9575-9586.
- [31] M. Imanishi, "Mechanisms and Strategies for Determining m⁶A RNA Modification Sites by Natural and Engineered m⁶A Effector Proteins" *Chemistry Asian J.* **2022**, *17*, e202200367.
- [32] Y. Iwasaki, Y. Ookuro, K. Iida, K. Nagasawa, W. Yoshida, "Destabilization of DNA and RNA G-Quadruplex Structures Formed by GGA Repeat Due to N⁶-Methyladenine Modification" *Biochem. Biophys. Res. Commun.* **2022**, *597*, 134-139.



Supporting Information

Identification of endogenous sequences interacting with METTL3/METTL14 RNA methyltransferase

Tamaki Endoh,^[b,c] Yiwei Ling,^[d] Shujiro Okuda,^[d] and Miki Imanishi*^[a]

[a] M. Imanishi

Institute for Chemical Research, Kyoto University

Gokasho, Uji, Kyoto 611-0011, Japan

E-mail: imiki@scl.kyoto-u.ac.jp

[b] T. Endoh

Graduate School of Frontiers of Innovative Research in Science and Technology (FIRST),

Konan University

7-1-20 Minatojima-Minamimachi, Chuo-ku, Kobe 650-0047, Japan

[c] T. Endoh

Frontier Institute for Biomolecular Engineering Research (FIBER), Konan University

7-1-20 Minatojima-Minamimachi, Chuo-ku, Kobe 650-0047, Japan

[d] Y. Ling and Prof. S. Okuda

Medical AI Center, School of Medicine, Niigata University

Niigata City, Niigata, 951-8514, Japan

MTD3/14 heterodimer

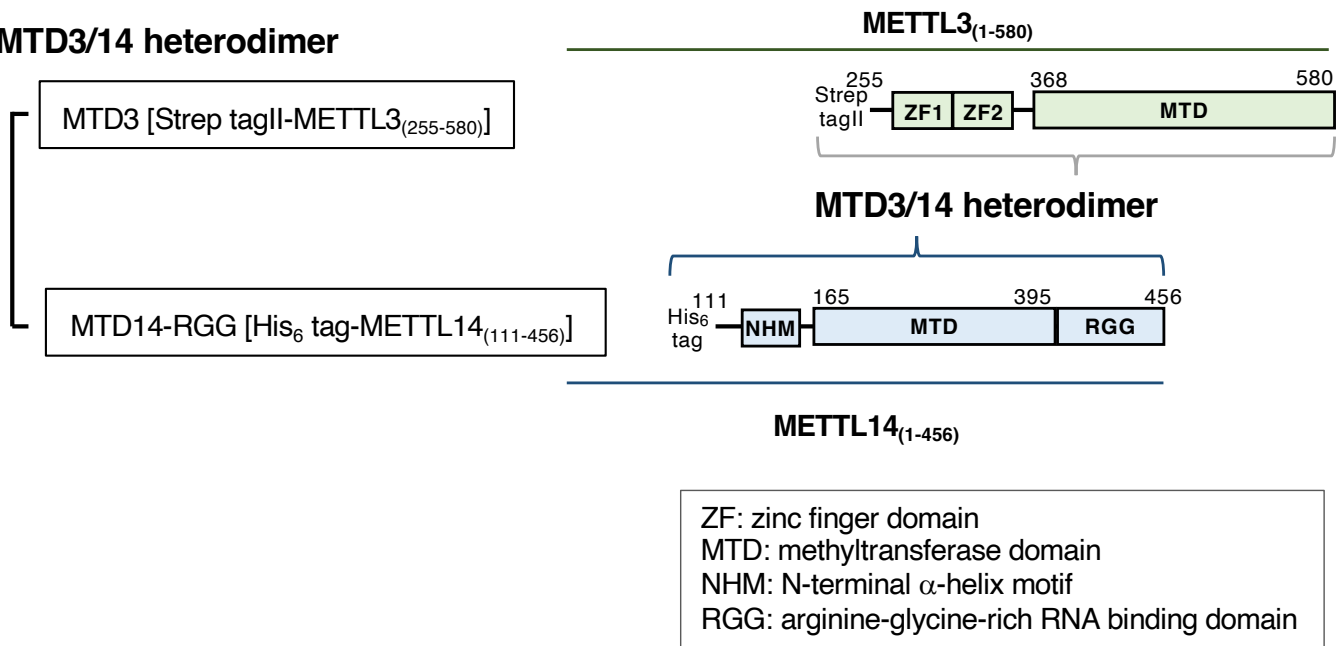
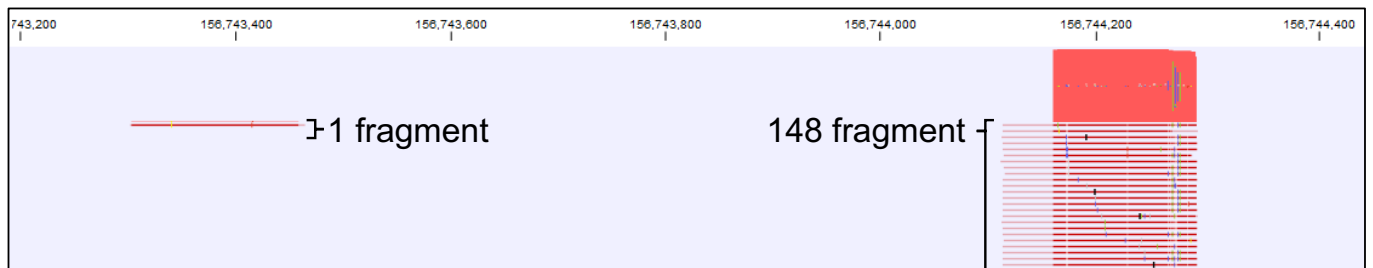


Figure S1. Domain structure of the MTD3/14 heterodimer. MTD3/14 used in this study contains METTL3(255-580) with strep tag II and METTL14(111-456) with His₆-tag. Methyltransferase domains (MTDs) of METTL3 and METTL14 form a stable heterodimer, making it possible to prepare MTD3/14 heterodimers through consecutive two-step purification using affinity columns for Strep-tag II and His₆-tag. ^[19]

(A)



(B)

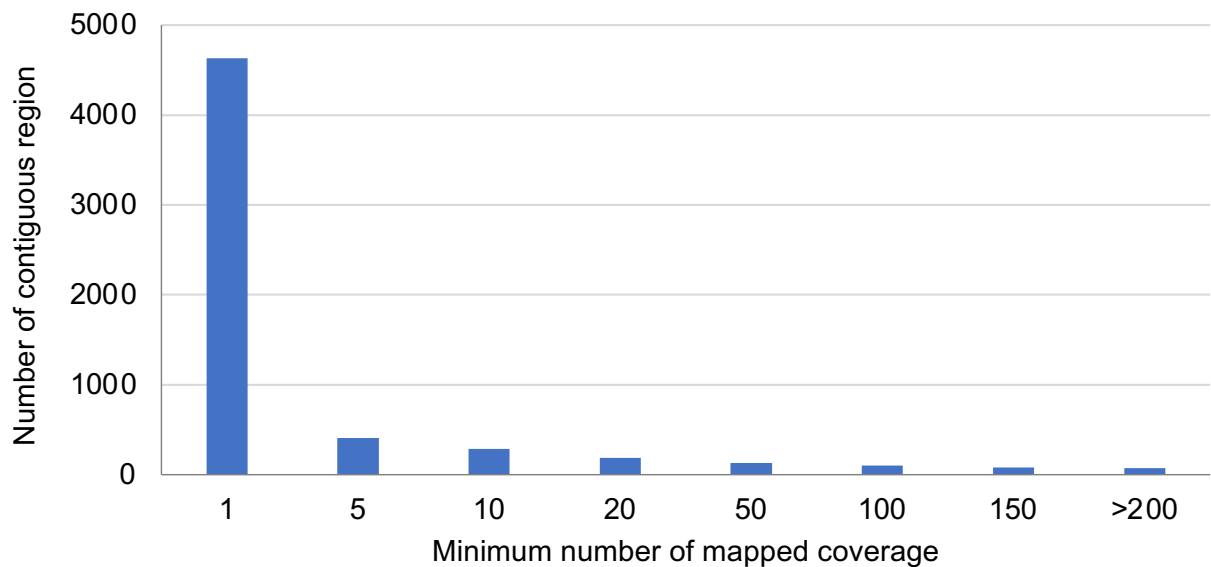


Figure S2. Mapping of NGS data (A) Example of mapped image of NGS data. A region on right side, where has 148 mapped fragments, is considered to RNA encoding region isolated by R-CAMPs interacted with MTD3/14. Another region on left side, where has only 1 mapped fragment, is considered to contaminated fragment during R-CAMPs isolation. (B) Number of contiguous region mapped on reference transcriptome. NGS data was mapped on human transcriptome. Contiguous regions, where the sequenced fragments were mapped more than the minimum number indicated, were calculated.

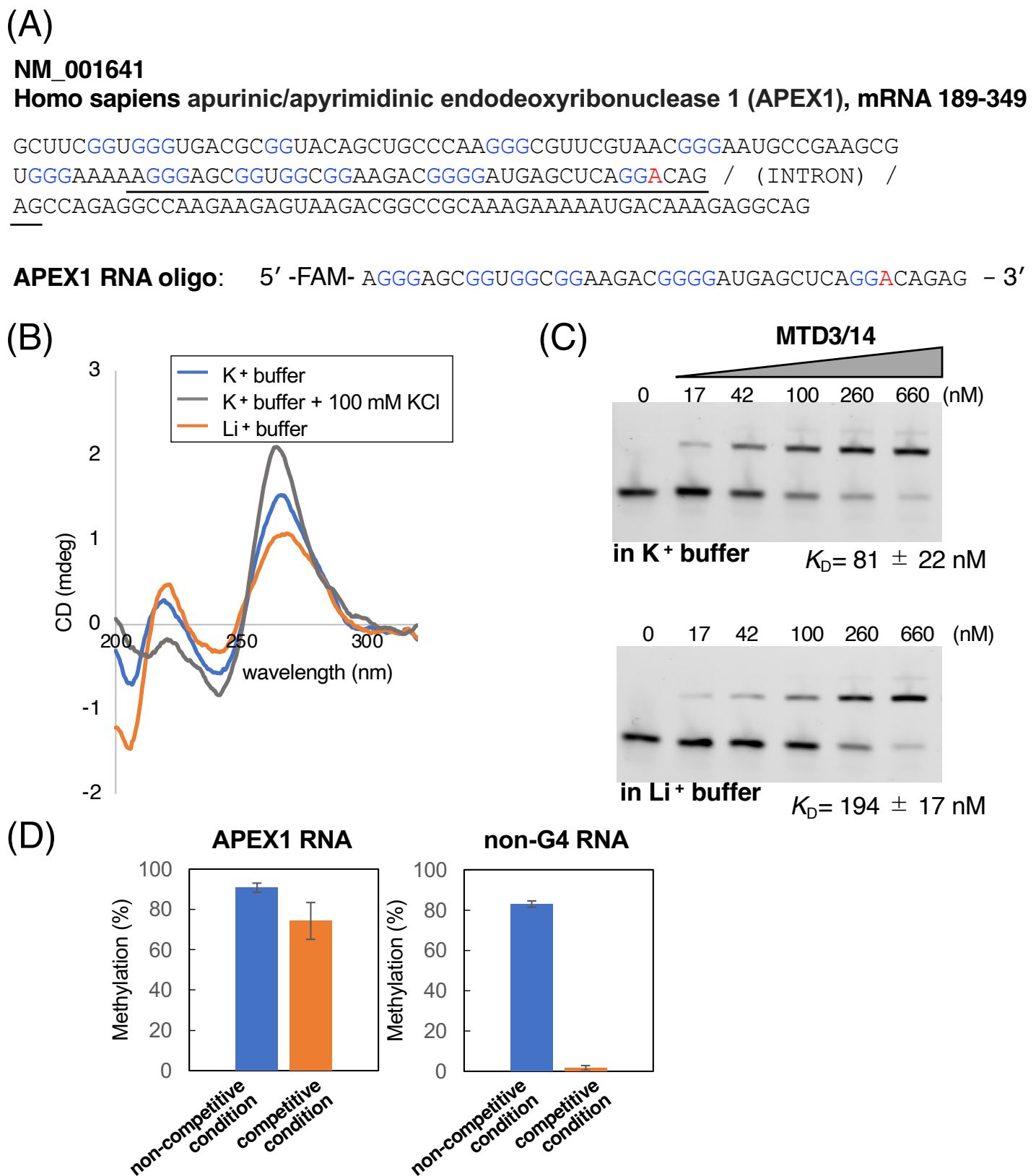


Figure S3. RNA binding ability of MTD3/MTD14 to the isolated RNA fragment. (A) Sequence of isolated RNA fragment in APEX1-derived mRNA and that of APEX1 RNA oligo used for in vitro experiments. Consecutive guanines potentially involved in G-quadruplex formation and adenine in the consensus DRACH sequence were highlighted in blue and red, respectively. (B) CD spectra of the APEX1 RNA oligo under indicated buffer conditions. (C) The results of EMSA of MTD3/14 to APEX1 RNA oligo in potassium or lithium buffer. (D) RNA methylation ratio of APEX1 RNA oligo (left) or non-G4 RNA oligo (right) under non-competitive (blue) or competitive (orange) conditions.