# Supporting Information

## Luo and Tong 10.1073/pnas.1412742111

### SI Materials and Methods

**Protein Expression and Purification.** The YTH domains of methylated RNA-binding protein 1 (MRB1) proteins from *Zygosaccharomyces rouxii* (National Center for Biotechnology Information accession code XP_002498076), *Saccharomyces cerevisiae* (NP_010662), and *Kluyveromyces lactis* (XP_454058) (amino acids 166–329, 143–306, and 101–264, respectively) were amplified by PCR from genomic DNA (ATCC). The fragments were cloned into a modified pET-28a vector with an N-terminal 6-His tag without any cleavage site. Mutants of ZrMRB1 were generated by using the QuikChange kit (Stratagene), and their identity was verified by DNA sequencing.

All proteins were overexpressed in *Escherichia coli* BL21 Star (DE3) strain (Novagen). The cells were induced by 0.4 mM isopropyl β-D-1-thiogalactopyranoside for 16 h at 20 °C for wild-type protein and 16 °C for mutants. The harvested cells were resuspended in a buffer containing 50 mM phosphate at pH 7.6, 500 mM NaCl, and 10 mM imidazole and lysed by sonication. Cell lysates were centrifuged at 25,000 × *g* for 30 min at 4 °C before being incubated with nickel beads (Qiagen). After 1 h, beads were transferred to plastic gravity flow columns (Bio-Rad) and washed extensively with a buffer containing 50 mM phosphate at pH 7.6, 500 mM NaCl, and 20 mM imidazole. Proteins were eluted with a buffer containing 50 mM phosphate at pH 7.6, 500 mM NaCl, 250 mM imidazole, and 5% (vol/vol) glycerol, and then directly loaded onto a gel filtration column (Sephacryl S300; GE Healthcare) equilibrated in a buffer containing 5 mM Hepes at pH 7.6, 500 mM NaCl, and 2 mM DTT. The peak fractions were collected and concentrated to 20 mg/mL for wild type and 7 mg/mL for mutants at room temperature and stored at –80 °C. Concentrations of proteins were determined by absorption at 280 nm wavelength (NanoDrop 2000c) and taking into account their theoretical extinction coefficients.

The selenomethionyl ZrMRB1 protein was prepared by the methionine metabolism pathway inhibition method by using M9 medium (1). Protein purification followed the same protocol as the native protein. Protein was concentrated to 5 mg/mL at room temperature and stored at –80 °C.

**RNA Synthesis.** Two RNAs with the sequences 5′-GG($m^6$A)CA-3′ and 5′-AGG($m^6$A)CAU-3′ were chemically synthesized by Dharmacon (GE Healthcare). For electrophoretic mobility shift assay (EMSA) studies, another RNA with a 6-FAM moiety at the 5′ end of the 7-mer oligo was also obtained. The RNAs were all deprotected and desalted by the manufacturer. They were then dissolved in buffer containing 25 mM Hepes at pH 7.6, 250 mM NaCl, 2 mM $MgCl_2$, and 1 mM DTT.

**Protein–RNA Complex Purification.** ZrMRB1 YTH domain was mixed with the 7-mer RNA solution with molar ratio of 1:1 and incubated at room temperature for 20 min. The mixture then was loaded onto a gel filtration column (Sephacryl S300; GE Healthcare) equilibrated with a buffer containing 5 mM Hepes at pH 7.6, 250 mM NaCl, and 2 mM DTT. Fractions corresponding to the complex were collected and concentrated to 4 mg/mL at room temperature.

**Crystallization.** Small crystals appeared 5 d after trays were set up by mixing 1 μL of protein solution with 1 μL of well solution (0.8 M ammonium dihydrogen phosphate) by using sitting-drop method at 4 °C. High-quality crystals were obtained by microseeding using the same well solution under paraffin oil at 4 °C. Crystals were transferred to well solution supplemented with 25% (vol/vol) glycerol as cryo-protectant before being flash frozen in liquid nitrogen. The crystals belong to space group $P6_122$, and there are six complexes in the asymmetric unit.

**Data Collection, Structure Determination, and Refinement.** A single-wavelength anomalous diffraction dataset at 2.7 Å resolution was collected at the peak absorption wavelength of selenium (0.9790 Å) at the X25 beamline of the National Synchrotron Light Source, using a Pilatus 6M detector. The diffraction images were processed with the HKL package (2). The structure was solved with the program pipeline autoSHARP (3), which includes Solomon for density modification (4) and ARP/wARP for automated model building (5). The overall figure of merit before density modification was 0.31. Manual model building was carried out with Coot (6) and the structure model was refined with Phenix (7). Translation/libration/screw parameters were not used in the refinement. The maximum-likelihood based estimated error in the atomic coordinates is 0.3 Å. The crystallographic information is summarized in Table S1.

**EMSA.** Twenty picomoles of FAM-labeled RNA was mixed with increasing concentrations of wild-type YTH proteins in a buffer containing 25 mM Mes (pH 6.4) and 300 mM NaCl in a total volume of 10 μL and incubated at room temperature of 15 min. For comparison of mutants with wild-type protein, 20 pmol of FAM-labeled RNA and proteins at a final concentration of 4 μM were incubated in 10 μL of total volume. For competition assays, 20 pmol of FAM-labeled RNA and 25 pmol of protein were mixed first and incubated at room temperature for 10 min, and then increasing concentrations of unlabeled 5- or 7-mer RNA were added into the reaction to a final volume of 10 μL and incubated for 30 min at room temperature. The electrophoresis was performed with 0.65% agarose gel at 4 °C in running buffer containing 40 mM Hepes, 0.114% acetic acid (vol/vol), 1 mM EDTA, and the pH was adjusted to 7.0 with sodium hydroxide. The gel was visualized by using a Typhoon FLA 7000 (GE Healthcare) using a method for FAM (Laser 473 nm, 520 nm filter).

**Isothermal Titration Calorimetry.** Isothermal titration calorimetry experiments were performed with a MicroCal Auto-iTC200 microcalorimeter (GE Healthcare) at 25 °C. The sample cell contained 200 μL of ZrMRB1 YTH domain at 2 μM concentration. The protein stock solution at 1 mM concentration was diluted to 2 μM with a buffer containing 50 mM Mes at pH 6.3, 300 mM NaCl, and 0.01% (vol/vol) Nonidet P-40 right before titration. FAM-labeled 7-mer RNA was dissolved in the same buffer at a concentration of 20 μM and used as a titrant in the syringe. Titration was measured by 19 injections of 2 μL. The dilution effect of RNA was measured by injecting into the buffer without protein. The integrated heat data after subtraction of the dilution effect were analyzed by using a one-site model in Origin following the manufacturer's instructions. The binding parameters $\Delta H$ (reaction enthalpy change), $K_a$ (binding constant) and $n$ (bound RNA per YTH domain) were allowed to change in the fitting.


1. Doublié S, et al. (1996) Crystallization and preliminary X-ray analysis of the 9 kDa protein of the mouse signal recognition particle and the selenomethionyl-SRP9. *FEBS Lett* 384(3):219–221.
2. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307–326.
3. Vonrhein C, Blanc E, Roversi P, Bricogne G (2007) Automated structure solution with autoSHARP. *Methods Mol Biol* 364:215–230.
4. Abrahams JP, Leslie AGW (1996) Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta Crystallogr D Biol Crystallogr* 52(Pt 1): 30–42.

5. Perrakis A, Morris R, Lamzin VS (1999) Automated protein model building combined with iterative structure refinement. *Nat Struct Biol* 6(5):458–463.
6. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126–2132.
7. Adams PD, et al. (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr* 58(Pt 11):1948–1954.

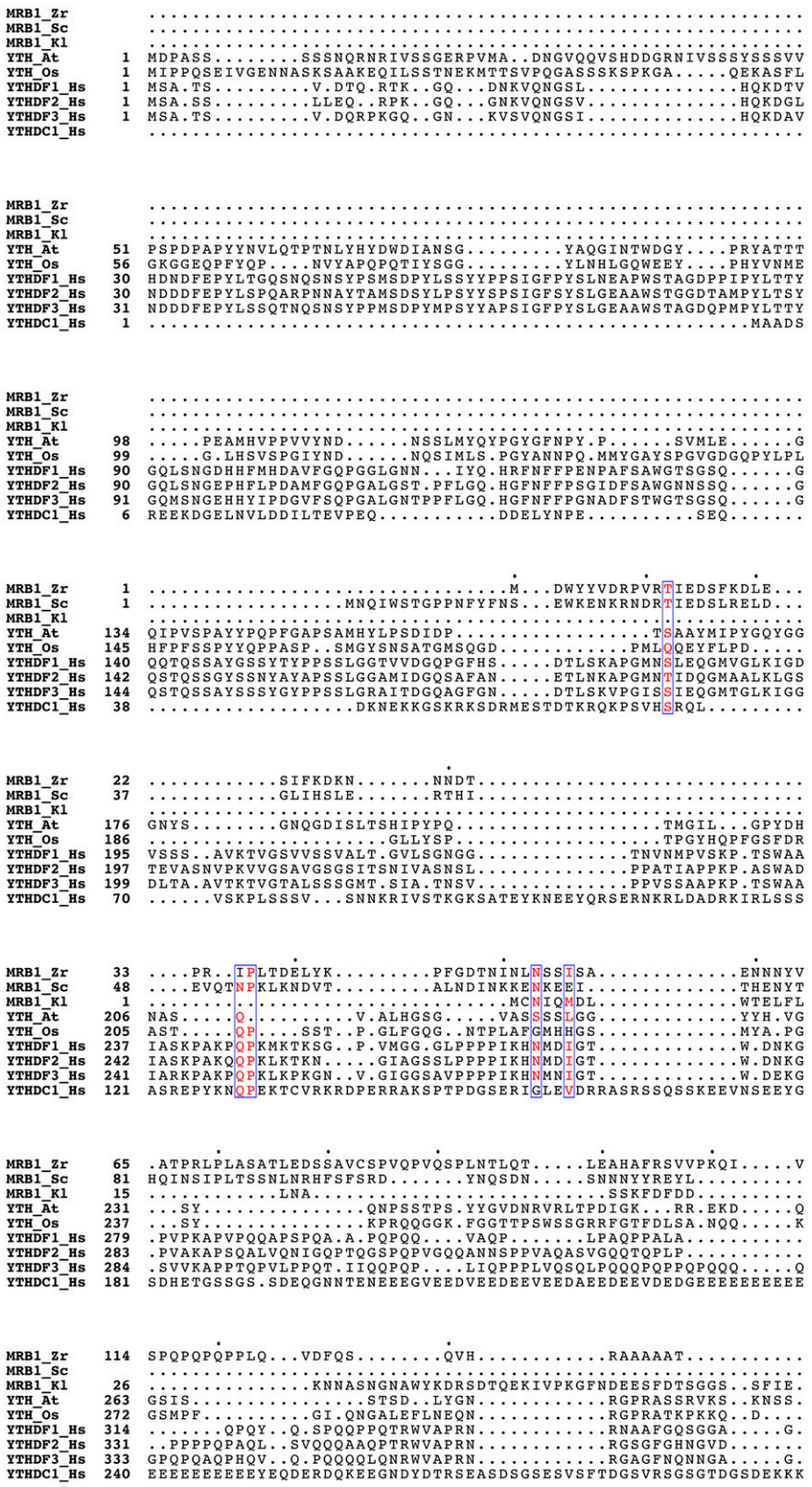


**Fig. S1.** (Continued)

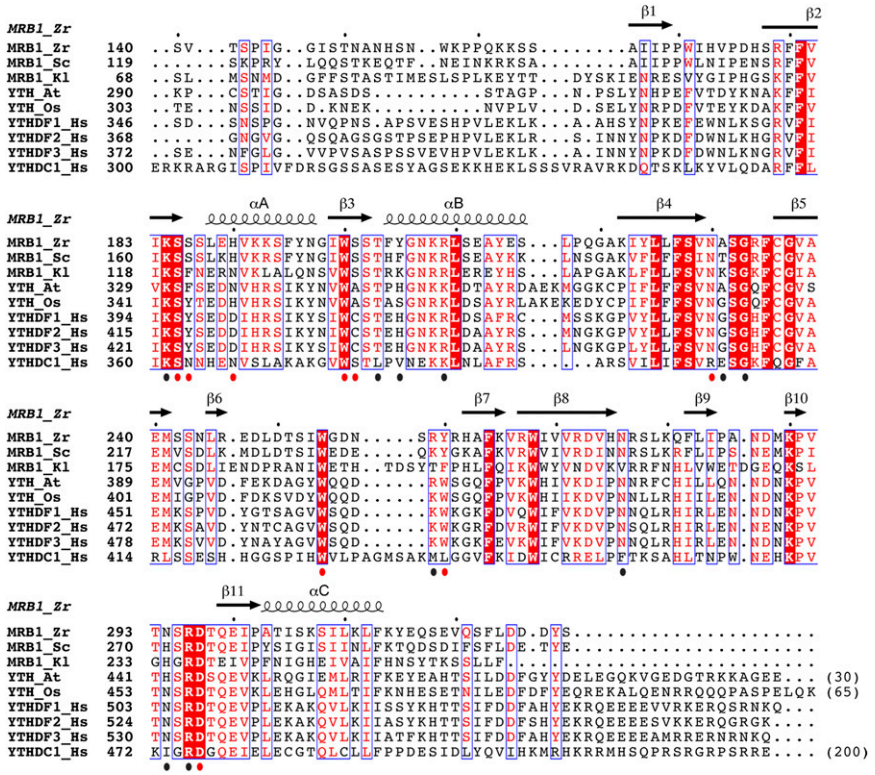

**Fig. S1.** Sequence alignment of *Z. rouxii* (Zr), *S. cerevisiae* (Sc), and *K. lactis* (Kl) MRB1, *A. thaliana* (At) and *Oryza sativa* (Os) YTH proteins, and human (Hs) YTHDF1-3 and YTHDC1. The secondary structure elements in the structure of the ZrMRB1 YTH domain are indicated. Residues that contact the m$^6$A residue are indicated with the red dots below the alignment, and those that contact the rest of the RNA are indicated with the black dots. The N-terminal segments of the proteins are not conserved. Additional C-terminal residues in three sequences are indicated with the numbers in parentheses. Produced with ESPript (1).

1. Gouet P, Courcelle E, Stuart DI, Métoz F (1999) ESPript: Analysis of multiple sequence alignments in PostScript. *Bioinformatics* 15(4):305–308.

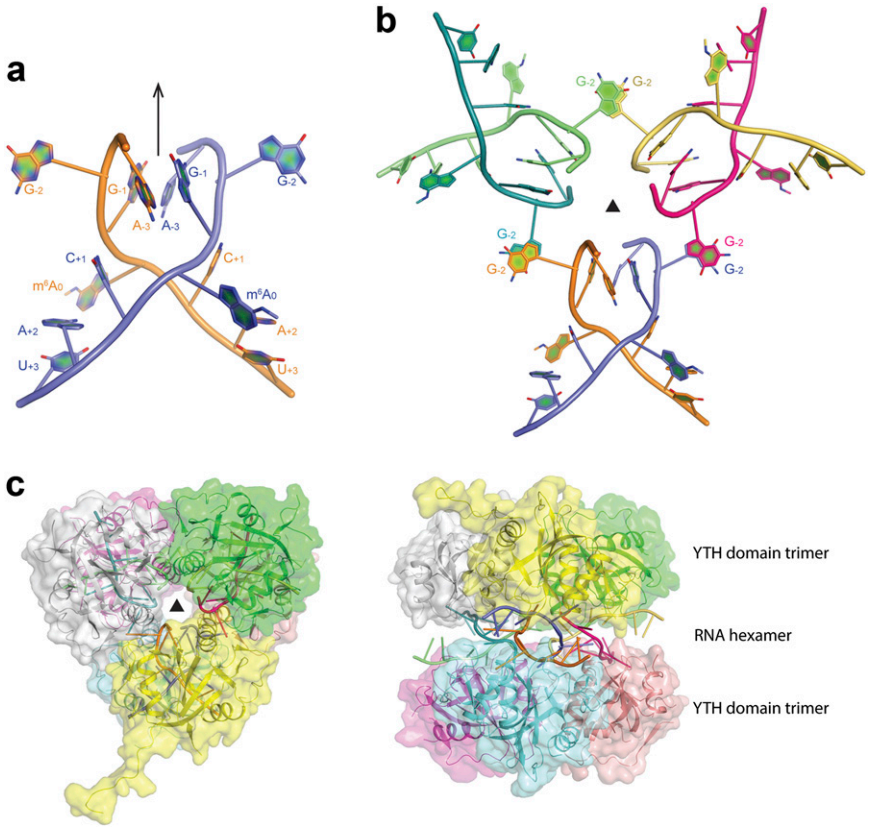


**Fig. S2.** (*A*) Schematic drawing of a dimer of the $m^6A$ RNA. $A_{-3}$ of one monomer is stacked with $G_{-1}$ of the other monomer. The twofold axis is vertical, indicated with the arrow. (*B*) A hexamer of the $m^6A$ RNA, formed primarily by the stacking of the $G_{-2}$ bases of neighboring molecules. The threefold symmetry axis of the hexamer is indicated with the black triangle. (*C*) Two views of the $m^6A$ RNA hexamer in complex with six YTH domains, shown with transparent surfaces. The RNA hexamer is located in the middle of the assembly, with three YTH domains on each face. Produced with the program PyMOL (www.pymol.org).

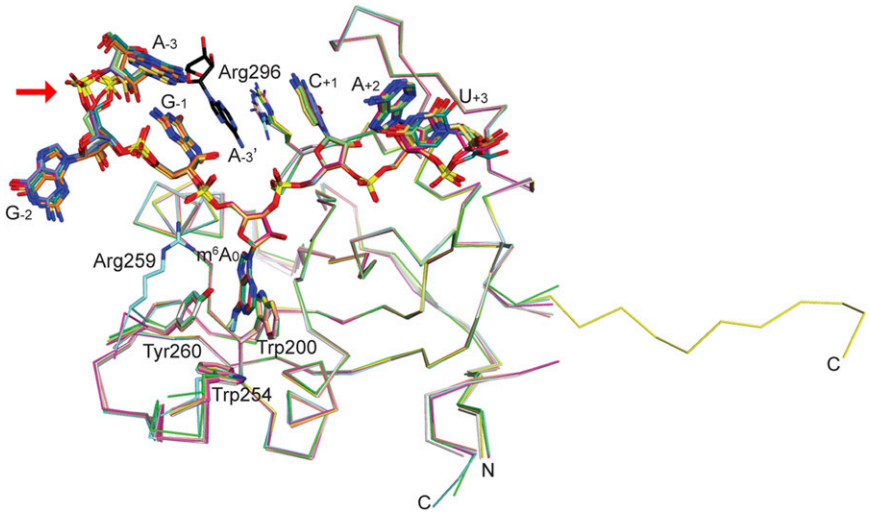


**Fig. S3.** Overlay of the structures of the six YTH domain–RNA complexes in the asymmetric unit. The proteins are shown as Cα traces, and their overall structures are remarkably similar. The C-terminal segment of the monomer in yellow is swapped with its crystallographic symmetry-mate, but it maintains equivalent interactions with the rest of the YTH domain. The RNAs are shown as stick models. The Arg296 side chain has conformational differences among the six monomers, possibly because it is also stacked with the $A_{-3}$ base from the other monomer of the RNA dimer (black). The Arg259 side chain has good density in only one of the six monomers. The different conformations for the 5′ phosphate group of the $G_{-2}$ residue is indicated with the red arrow.

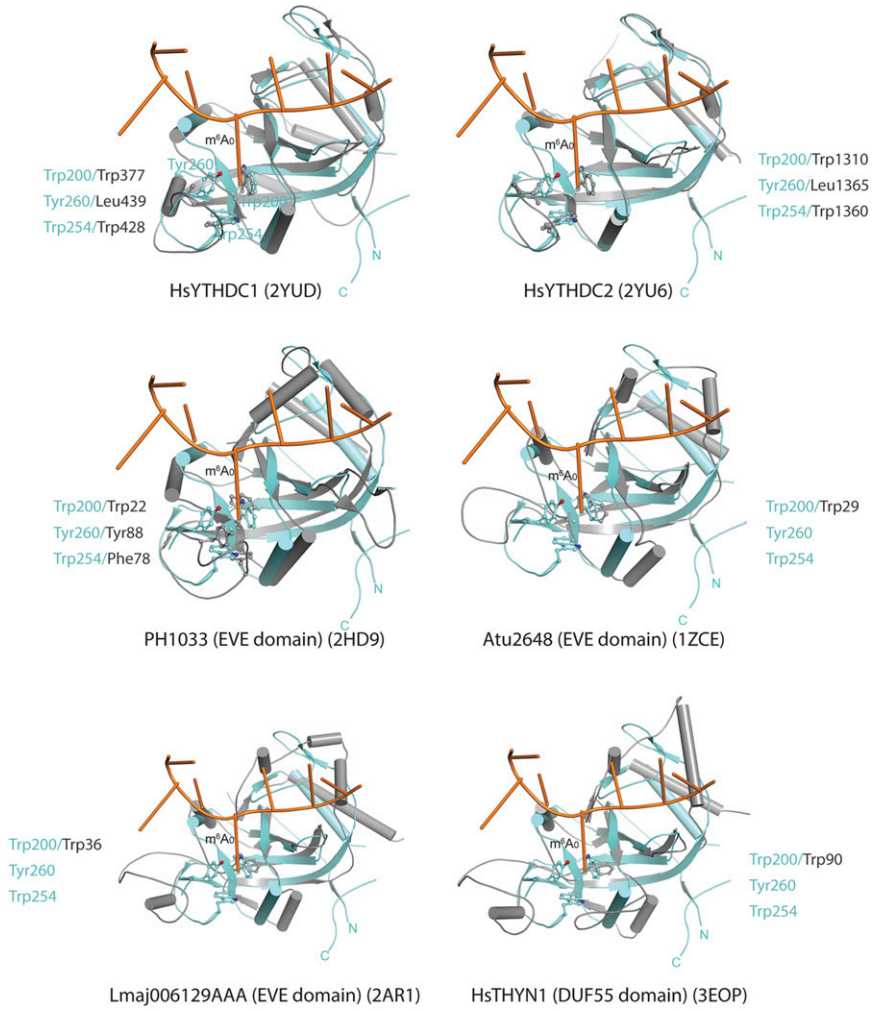


**Fig. S4.** Overlay of the ZrMRB1 YTH domain structure (cyan) with other homologous structures (gray) (1–4). The $m^6A$ RNA is shown in orange. The PDB accession codes are given in parentheses. Residues that are equivalent to the aromatic cage of the ZrMRB1 YTH domain are shown as ball-and-stick models. THYN1, DUF55 domain of thymocyte nuclear protein 1.


1. Sugahara M, Asada Y, Morikawa Y, Kageyama Y, Kunishima N (2008) Nucleant-mediated protein crystallization with the application of microporous synthetic zeolites. *Acta Crystallogr D Biol Crystallogr* 64(Pt 6):686–695.
2. Arakaki T, et al. (2006) Structure of Lmaj006129AAA, a hypothetical protein from Leishmania major. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 62(Pt 3):175–179.
3. Yu F, et al. (2009) Determining the DUF55-domain structure of human thymocyte nuclear protein 1 from crystals partially twinned by tetartohedry. *Acta Crystallogr D Biol Crystallogr* 65(Pt 3):212–219.
4. Bertonati C, et al. (2009) Structural genomics reveals EVE as a new ASCH/PUA-related domain. *Proteins* 75(3):760–773.

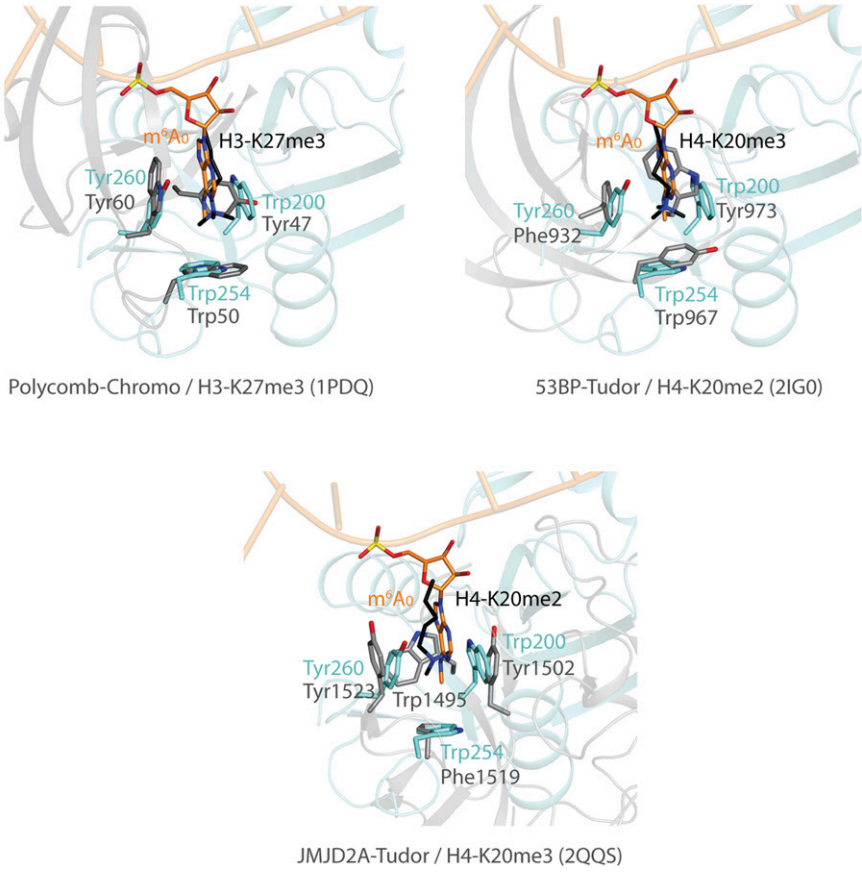


**Fig. S5.** Overlay of the ZrMRB1 YTH domain (cyan) in complex with the $m^6A$ RNA (orange) with the structures of the aromatic cage (gray) for recognizing methylated lysine (black) in polycomb chromo domain (1), 53BP tudor domain (2), and JMJD2A tudor domain (3). The overlay was done manually, based on the side chains in the aromatic cage. JMJD2A has an additional Trp residue (Trp1495) in the cage. The PDB entry codes are given in parentheses.

1. Fischle W, et al. (2003) Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. *Genes Dev* 17(15):1870–1881.
2. Botuyan MV, et al. (2006) Structural basis for the methylation state-specific recognition of histone H4-K20 by 53BP1 and Crb2 in DNA repair. *Cell* 127(7):1361–1373.
3. Lee J, Thompson JR, Botuyan MV, Mer G (2008) Distinct binding modes specify the recognition of methylated histones H3K4 and H4K20 by JMJD2A-tudor. *Nat Struct Mol Biol* 15(1):109–111.

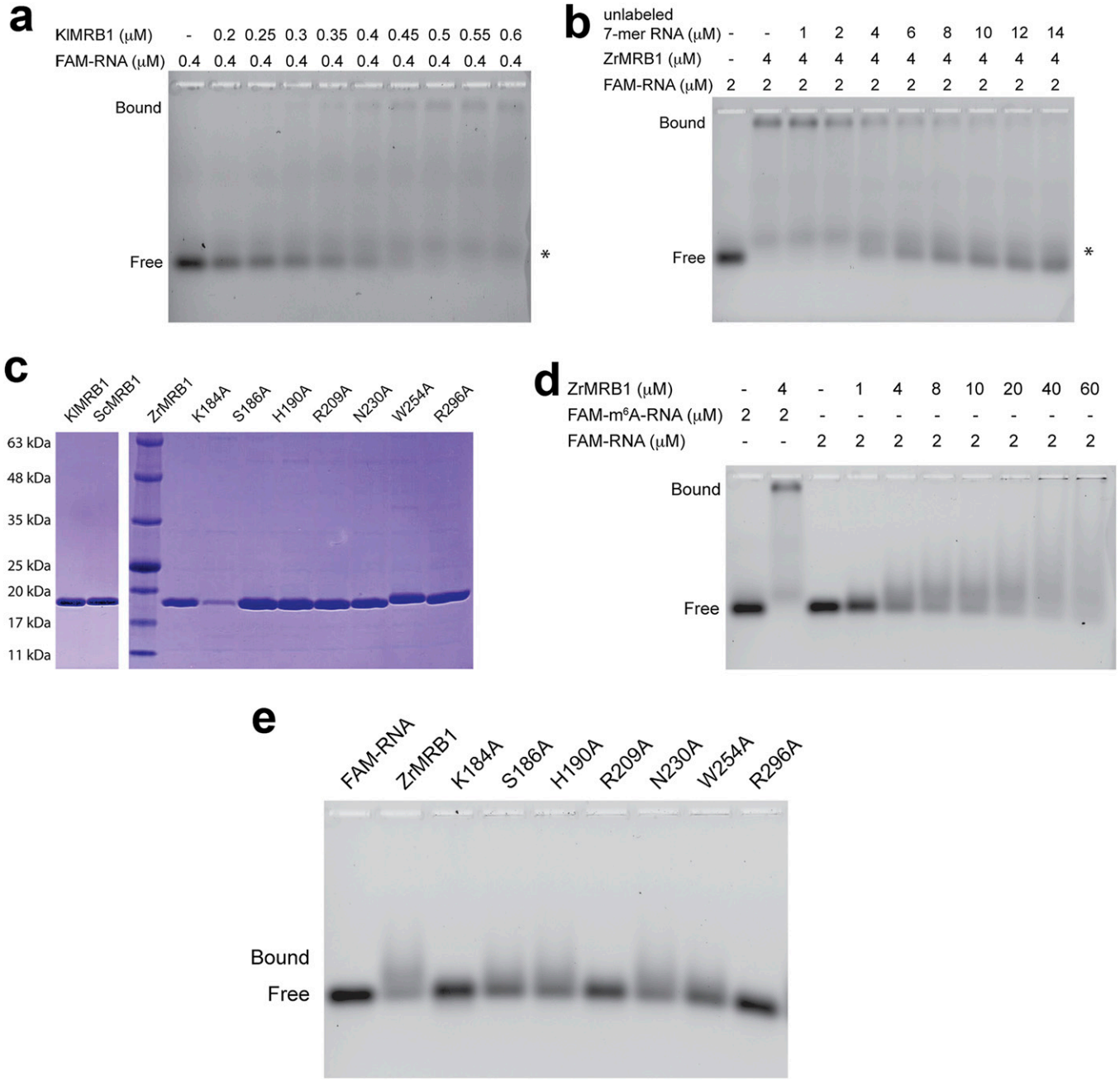


**Fig. S6.** (*A*) Titration experiment showing the association between the KlMRB1 YTH domain and the RNA. (*B*) Competition experiment with increasing concentrations of the unlabeled 7-mer RNA against labeled 7-mer RNA. (*C*) SDS gel of wild-type and mutant YTH domains. (*D*) Titration experiment with an unmethylated RNA with the same sequence as the $m^6A$ RNA. (*E*) EMSA with the unmethylated RNA against wild-type and mutant ZrMRB1 YTH domains (at 6 μM concentration).

**Table S1. Summary of crystallographic information**

| Structure | ZrMRB1 YTH domain–RNA complex |
|---|---|
| Space group | $P6_122$ |
| Cell dimensions | |
| $a = b$, Å | 117.4 |
| $c$, Å | 380.9 |
| Wavelength, Å | 0.9790 |
| Resolution*, Å | 25–2.7 (2.8–2.7) |
| $R_{merge}$, % | 13.7 (49.8) |
| I/σI | 10.9 (3.2) |
| Completeness, % | 94 (95) |
| No. of observations | 186,118 |
| Redundancy | 4.5 (4.6) |
| No. of reflections | 41,335 |
| $R_{work}$, % | 19.9 (27.6) |
| $R_{free}$†, % | 24.5 (36.7) |
| rms deviations | |
| Bond lengths, Å | 0.015 |
| Bond angles, ° | 1.3 |
| No. of atoms | |
| Protein | 7,561 |
| Nucleic acid | 900 |
| Water | 111 |
| Average $B$ value, Å$^2$ | |
| Protein | 34.6 |
| Nucleic acid | 36.4 |
| Water | 28.6 |

*The numbers in parentheses are for the highest resolution shell.
†Five percent of the reflections were selected for free $R$ calculation.