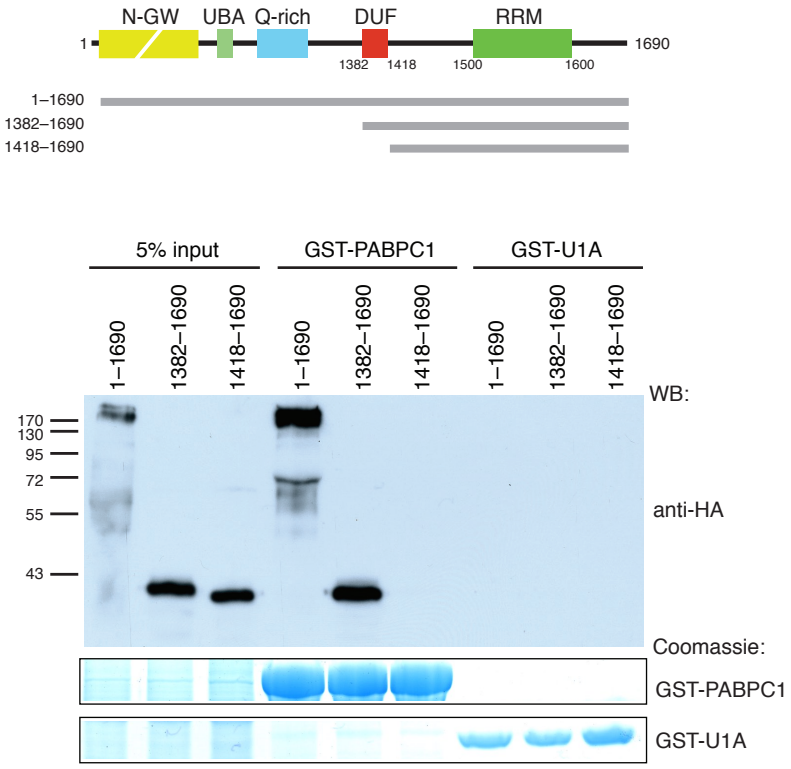
Supplementary Material for:

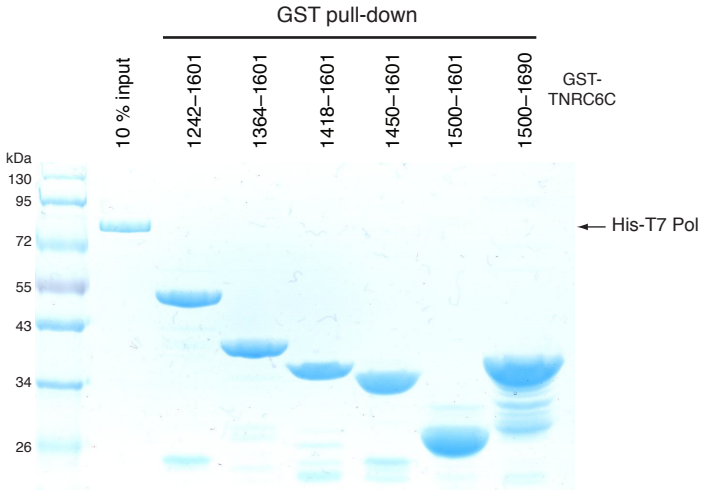
Structural insights into the human GW182-PABC interaction in microRNA-mediated deadenylation

Martin Jinek, Marc R. Fabian, Scott M. Coyle, Nahum Sonenberg & Jennifer A. Doudna

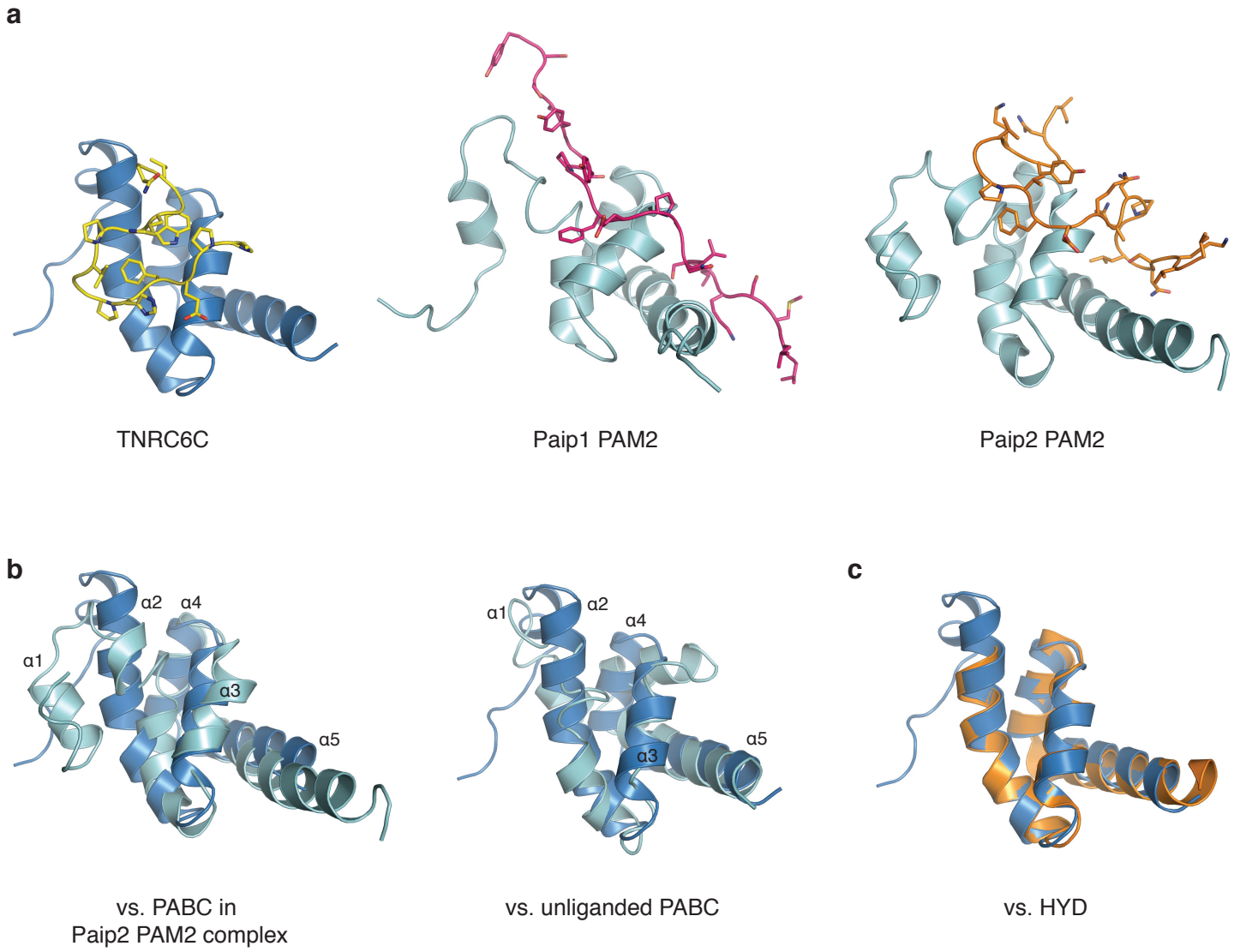
Supplementary Figures 1–6
Supplementary Table 1
Supplementary Methods
Supplementary References



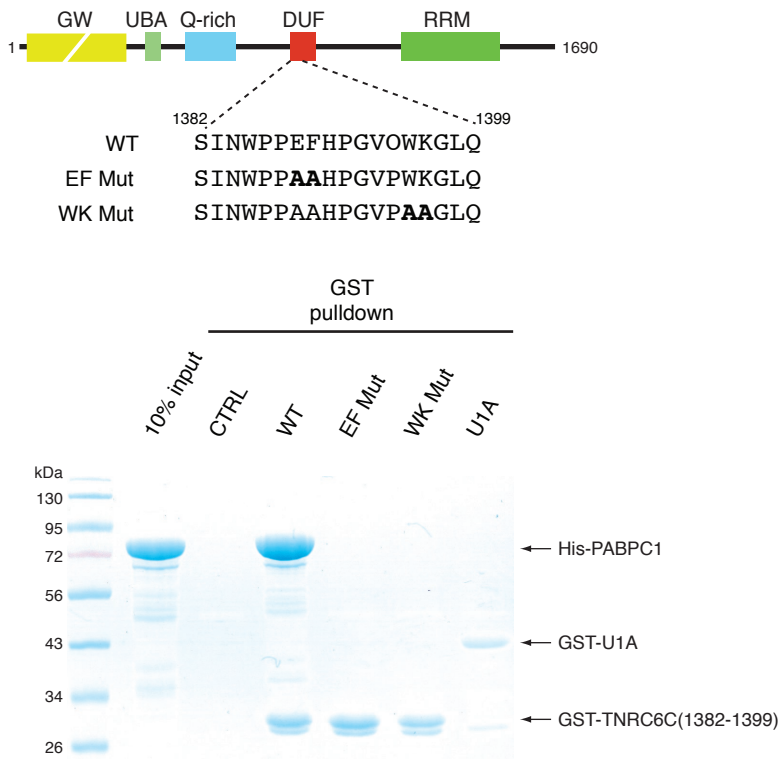
Supplementary Figure 1. The DUF domain of TNRC6C is required for stable interaction with PABPC1. Top: Schematic representation of HA epitope-tagged human TNRC6C fragments expressed in HEK293T cells. Bottom: Extracts from cells transiently expressing the indicated constructs were incubated with recombinant GST-PABPC1 or GST-U1A (negative control) immobilized on glutathione sepharose. Precipitated proteins were resolved by SDS-PAGE and analyzed by immunoblotting using an anti-HA antibody and staining with Coomassie blue.



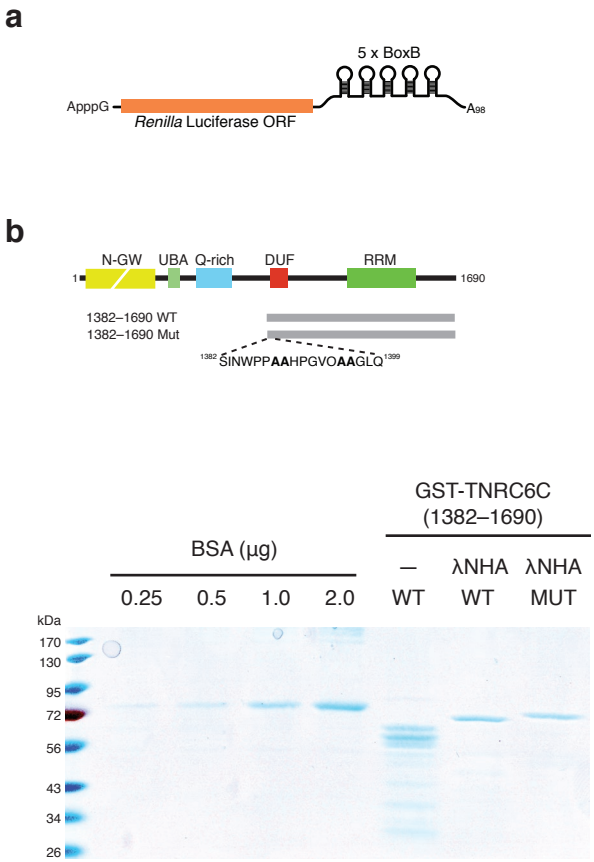
Supplementary Figure 2. Control pulldown assay of GST-TNRC6C fragments incubated with recombinant His-tagged T7 RNA polymerase. Precipitated proteins were resolved by SDS-PAGE and analyzed by staining with Coomassie blue.



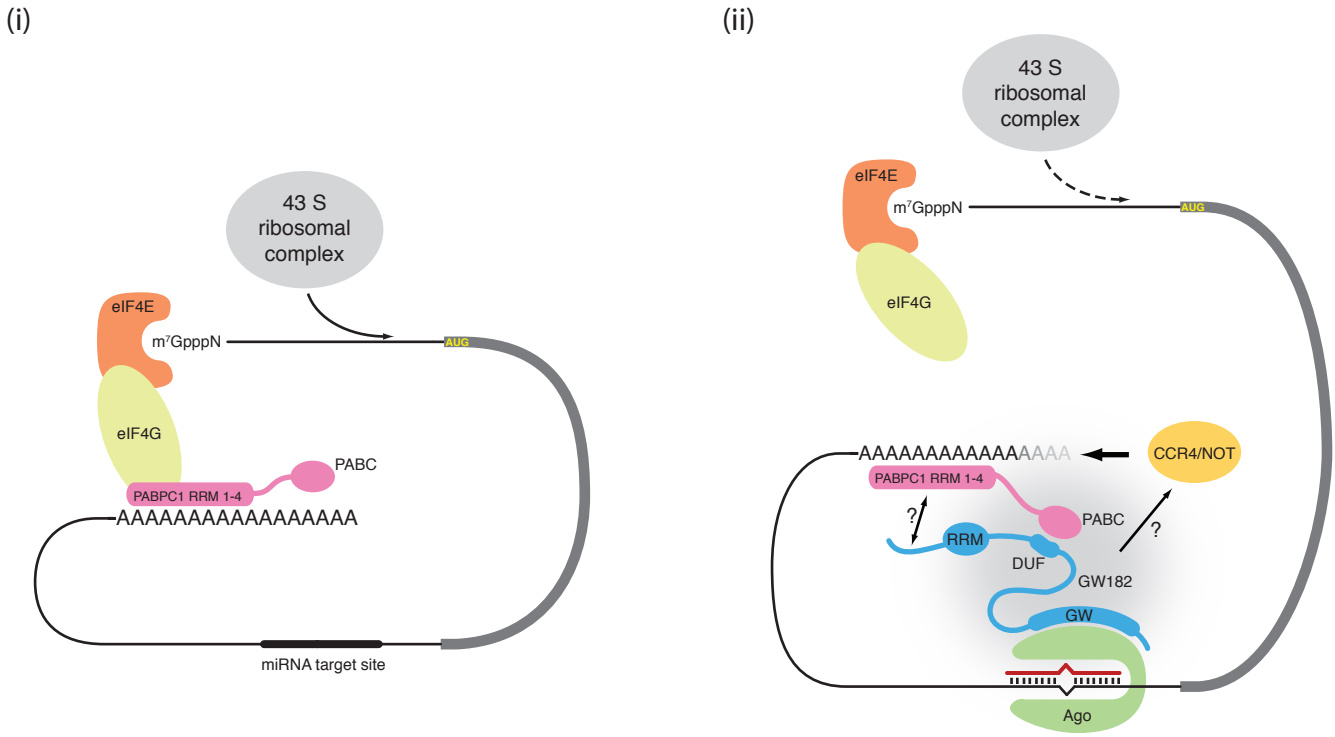
Supplementary Figure 3. Structural comparison of the human PABPC1 C-terminal (PABC) domain in complexes with TNRC6C and PAM2 peptides. **(a)** Ribbon diagrams of the human PABC domain in complexes with TNRC6C DUF domain (left, PABC domain in blue, TNRC6C peptide in yellow), Paip1 PAM2 motif (center, PABC domain in green, Paip2 PAM2 peptide in magenta, PDB code 1JH4) and Paip2 PAM2 motif (right, PABC domain in green, Paip1 PAM2 peptide in orange, PDB code 1JGN). The complexes were superposed using the SSM server[1] and are shown in identical orientations. **(b)** Left: Superposition of the PABC domain as observed in the crystal structure of the TNRC6C-PABC complex (blue) and in the NMR structure of the Paip2 PAM2-PABC complex (green, PDB code 1JGN). Right: Superposition of the PABC domain as observed in the crystal structure of the TNRC6C-PABC complex (blue) and in the NMR structure of the unliganded PABC domain (green, PDB code 1G9L). The structures superpose with root mean square deviations (rmsd) of 3.9 Å and 2.9 Å, respectively, over all Cα atoms, as determined with DALI[2]. **(c)** Superposition of the PABC domain as observed in the crystal structure of the TNRC6C-PABC complex (blue) and with the crystal structure of the PABC domain from the human hyperplastic discs (HYD) protein (orange, PDB code 1I2T). The structures superpose with an rmsd of 0.8 Å over 59 Cα atoms.



Supplementary Figure 4. Conserved aromatic residues in the TNRC6C DUF domain are required for interaction with PABPC1. GST pulldown assay of GST-fused TNRC6C minimal DUF domain peptide (TNRC6C residues 1382-1399) and its mutants with full-length His-tagged PABPC1. GST-fused U1A protein served as negative control.



Supplementary Figure 5. Tethering TNRC6C fragments to reporter RNAs in the Krebs extract. **(a)** Schematic representation of the reporter RNA construct. The RNA is ApppG capped and carries a 98-nt poly(A) tail. **(b)** TNRC6C constructs used in the tethering assay. The constructs feature an N-terminal GST-tag and a C-terminal FLAG epitope. The proteins were expressed in *E. coli* and purified by glutathione and anti-FLAG antibody affinity chromatographic steps. Equal amounts of control, WT and MUT constructs were analyzed by SDS-PAGE and stained with Coomassie blue. BSA serves as a quantification standard.



Supplementary Figure 6. A hypothetical model of GW182-PABPC1 interactions in miRNA-mediated silencing. (i) In the absence of miRNA-mediated silencing, interaction between the translational initiation factor eIF4G and PABPC1 circularizes the mRNA, which stimulates translational initiation by enhancing the recruitment of the 43S ribosomal complex to the 5' UTR of the mRNA. (ii) Cognate miRNA directs the assembly of the miRISC (consisting of the miRNA, an Argonaute protein and GW182) on the 3' UTR of the mRNA. GW182 binding to PABPC1 through the DUF-PABC interface, and possibly through secondary interactions involving the RRM-flanking sequences, helps to sequester the poly(A) tail to the vicinity of the miRISC. Disruption of the PABPC1-eIF4G interaction through competition with GW182 inhibits translational initiation due to reduced 43S complex recruitment. The sequestration of the poly(A) tail through the PABPC1-GW182 interactions juxtaposes the poly(A) tail against the miRISC-associated CCR4/NOT deadenylase complex and promotes mRNA deadenylation. It is presently unclear whether the deadenylase complex is directly recruited by miRISC, although PABPC1 has been implicated in the process[3].

Supplementary Table 1. Data collection and refinement statistics.

	Native	SeMet peak
Data collection		
X-ray source	ALS 8.2.1	ALS 8.3.1
Space group	$P2_1$	$P2_1$
Cell dimensions		
a, b, c (Å)	37.1, 72.2, 37.0	37.0, 72.3, 37.1
α, β, γ (°)	90.0, 119.8, 90.0	90.0, 120.1, 90.0
Wavelength (Å)	0.999954	0.979491
Resolution (Å)*	50–1.5 (1.6–1.5)	50–1.5 (1.6–1.5)
R_{sym} (%)*	5.2 (46.2)	4.7 (24.8)
$I/\sigma I$*	17.0 (2.9)	16.8 (5.0)
Completeness (%)*	99.3 (99.4)	97.5 (97.0)
Redundancy*	3.6 (3.7)	2.4 (2.4)
Refinement		
Resolution (Å)	50 - 1.5	
No. reflections	99486	
R_{work} / R_{free}	0.164 / 0.187	
No. atoms		
Protein	1444	
Ions	10	
Water	217	
B-factors		
Protein	17.7	
Ions	23.4	
Water	31.7	
R.m.s. deviations		
Bond lengths (Å)	0.006	
Bond angles (°)	1.0	

* Values in parentheses denote highest resolution shell

SUPPLEMENTARY METHODS

His-PABP and GST-PABP expression and purification

The sequence encoding human PABP1 (residues 2-636) was PCR-amplified from a commercially available cDNA (Open Biosystems IMAGE clone 3940309) and subcloned into pETM-14 (His-PABPC1) or pGEX6p1 (GST-PABPC1). Both constructs were expressed in *Escherichia coli* strain BL21 at 18^{o}C for 16 hours after induction with 0.2 mM IPTG. GST-PABPC1 expressing cells were lysed in 20 mM HEPES 7.5, 500 mM KCl, 1 mM TCEP, in a homogenizer (Avestin). Clarified lysate was loaded onto a glutathione sepharose 4B (GE Life Sciences) column. The column was washed with 60 ml lysis buffer (minus protease inhibitors) and bound protein was eluted with 100 mM HEPES pH 7.5, 100 mM KCl, 1 mM TCEP. The protein was dialyzed overnight against 20 mM HEPES pH 7.5, 100 mM KCl, 1 mM EDTA, 1 mM TCEP, 10% glycerol, concentrated, flash-frozen in liquid nitrogen and stored at -80^{o}C. His-PABPC1 expressing cells were lysed in 20 mM Tris-Cl pH 8.5, 500 mM NaCl, 5 mM imidazole (supplemented with protease inhibitor cocktail from Roche). Clarified lysate was loaded onto a Talon Co^{2+}-affinity column (Clontech). The column was washed with 50 ml of 20 mM Tris-Cl pH 8.5, 500 mM NaCl, 2 M urea, followed by washing with 50 ml of lysis buffer. Bound His-PABPC1 was eluted in 20 mM Tris-Cl, 150 mM NaCl, 100 mM imidazole (pH 8.0) and dialyzed overnight against 20 mM HEPES 7.5, 100 mM KCl. To remove co-purifying nucleic acid, His-PABP was further purified on a 5 ml HiTrap Heparin column (GE Life Sciences) and eluted with a linear gradient of 100 mM – 2 M KCl. Eluted protein was dialyzed overnight against 20 mM HEPES pH 7.5, 100 mM KCl, 1 mM TCEP, 10% glycerol. The dialyzed protein was concentrated to ~ 2 mg.ml^{-1}, flash-frozen in liquid nitrogen and stored at -80^{o}C.

DNA constructs and transfections of HEK293T cells. For expression in mammalian cells, full-length human TNRC6C and its fragments were PCR-amplified and subcloned into pcDNA3.1 in frame with an upstream sequence

encoding a HA epitope, a λN-peptide and a FLAG epitope. Transfections were performed in 10-cm dishes seeded with $2x10^6$ cells using Fugene HD transfection reagent (Roche) using 3 µg of plasmid DNA per transfection.

GST pull-downs from HEK293T cell lysates. HEK293T cells were collected 48 hrs post transfection, washed twice with PBS and resuspended in 1.5 ml of assay buffer (20 mM HEPES 7.5, 100 mM KCl, 2 mM $MgCl_2$, 1 mM EDTA, 1mM DTT, 0.1% (v/v) Triton X-100) supplemented with protease inhibitors (Roche). Cells were lysed by rocking at 4°C for 10 minutes, followed by sonication (15 one-second bursts) and lysates were clarified by centrifuging twice at 16,100 g for 10 minutes. 30 µg of recombinant GST-PABP1 or GST-U1A were added to 500 µl of the clarified lysate and the mixture was incubated on a Nutator together with 20 µl (packed volume) of glutathione sepharose beads. The beads were washed five times with assay buffer and bound proteins were eluted by heating the beads in 1x SDS-PAGE sample loading buffer for 5 min at 95°C. Proteins were separated by SDS-PAGE and transferred to polyvinylidene fluoride membranes. Membranes were blocked in PBS containing 4% fat-free milk powder and probed with HRP-conjugated anti-HA antibody (Roche) used at 1:500 dilution in PBS. Blots were developed using the Western Lightning chemiluminiscence reagent (Perkin-Elmer) as recommended by the manufacturer.

GST pull-down experiments using recombinant proteins. GST-fused fragments of human TNRC6C were subcloned into pGEX6p1 using BamHI and NotI restriction sites. The proteins were expressed in 0.75 l cultures of *Escherichia coli* strain BL21 at 18°C for 16 hours after induction with 0.2 mM IPTG. Cell pellets were resuspended in 10 ml of 20 mM HEPES 7.5, 500 mM KCl, 1 mM TCEP, supplemented with protease inhibitors (Roche) and flash-frozen. For pulldown experiments, 1 ml aliquots of cell suspension were lysed by sonication (30 one-second bursts) and clarified by centrifuging twice at 16,100 g for 10 min. Clarified lysates were incubated with 100 µl of glutathione sepharose 4B beads for 1hr at 4°C. The beads were washed with 30 ml of lysis buffer in a

Polyprep disposable column (BioRad) and recovered in 1ml of assay buffer (20 mM HEPES 7.5, 100 mM KCl, 2 mM $MgCl_2$, 1 mM EDTA, 1% (v/v) Triton X-100) supplemented with protease inhibitors (Roche). 20 µl of beads were incubated with 40 µg of His-PABP in a total volume of 1 ml of assay buffer for 1 hr at 4^{o}C. The beads were washed five times with assay buffer and bound proteins were eluted by heating the beads in 40 µl of 1x SDS-PAGE sample loading buffer for 5 min at 95^{o}C. 5 µl samples were separated by SDS-PAGE and stained with SimplyBlue SafeStain reagent (Invitrogen). For Western blots, samples were separated by SDS-PAGE and transferred to polyvinylidene fluoride membranes. Membranes were blocked in PBS containing 4% fat-free milk powder and probed with a monoclonal anti-PABP antibody (10E10 from Sigma, used at 1:500 dilution in PBS), followed by an HRP-conjugated anti-mouse IgG secondary antibody (GE Life Sciences, used at 1:2500 dilution in 4% fat-free milk). For control experiments, His-tagged T7 RNA polymerase was prepared as previously described[1]. 50 µg of protein were used per binding reaction.

Expression and purification of the TNRC6C-PABC complex. The TNRC6C DUF domain peptide (corresponding to residues 1382-1399 of human TNRC6C, sequence SINWPPEFHPGVPWKGLQ) was synthesized using solid-state FMOC chemistry, deprotected and lyophilized. The C-terminal domain of human PABPC1 (residues 545-619) was subcloned into the expression vector pGEX6p1 (GE Life Sciences) in frame with the GST tag and a rhinovirus protease 3C cleavage sequence. The resulting protein was expressed in *Escherichia coli* BL21 strain at 18^{o}C by inducing with 0.2 mM IPTG for 16 hours. Cells were resuspended in 20 mM HEPES pH 7.5, 500 mM KCl (supplemented with a protease inhibitor cocktail from Roche) and lysed in a homogenizer (Avestin). The protein was purified from clarified lysate by affinity chromatography on glutathione sepharose (GE Life Sciences) and eluted with 100 mM HEPES pH 7.5, 100 mM KCl, 20 mM reduced glutathione. The eluted fusion protein was cleaved with rhinovirus 3C protease for 16 hours at 4^{o}C. The cleaved mixture was combined with a two-fold molar excess of the DUF domain peptide. The

complex was subsequently separated from cleaved GST by size exclusion chromatography on a Superdex 75 (16/60) column (GE Life Sciences) in 20 mM HEPES pH 7.5, 100 mM KCl and concentrated to 50 mg.ml^{-1}. The composition of the complex was verified by FT-ICR mass spectrometry. Selenomethionine (SeMet)-substitued PABP-C domain was expressed in *E. coli* strain BL21 grown in minimal medium supplemented with selenomethionine. The SeMet-containing complex was reconstituted and purified as described for the native complex and concentrated to 30 mg.ml$^{-1.}$

Crystallization and structure determination. The TNRC6C-PABC complex was crystallized by the hanging drop vapor diffusion method by mixing equal volumes of protein solution and a reservoir solution containing 0.1 M sodium acetate pH 4.6, 0.2 M ammonium sulfate and 30% (w/v) PEG 4000. Plate-like monoclinic crystals containing two copies of the complex in the asymmetric unit grew to a maximum dimension of approximately 200 x 200 x 30 µm within a week. Crystals of SeMet-labeled complex were obtained from 0.1 M sodium acetate pH 4.6, 0.1 M ammonium sulfate and 24% (w/v) PEG 4000. Crystals were cryoprotected by a brief soak in 0.1 M sodium acetate pH 4.6, 0.2 M ammonium sulfate, 25% (w/v) PEG 4000, 20% (v/v) ethylene glycol, and flash-cooled in liquid nitrogen. Native data were collected at beamline 8.2.1 of the Advanced Light Source (ALS, Lawrence Berkeley National Laboratory). Single-wavelength anomalous dispersion (SAD) data were collected at the selenium K-edge (wavelength 0.979491 Å) at beamline 8.3.1 of the ALS. Both native and SAD data were processed with XDS[2]. Analysis of the data using Phenix.xtriage indicated the presence of pseudo-merohedral twinning, with twinning operator (*l*,-k,*h*) and twin fractions in the range of 0.19-0.45. Substructure solution, phasing and density modification were carried out using the PHENIX crystallographic package[3-5], giving acentric figures of merit of 0.60 and 0.75 before and after density modification, respectively. The resulting electron density map was of excellent quality, allowing automatic model building using Arp/warp[6, 7]. The atomic model was completed by iterative rounds of manual building in

COOT[8] and refinement in Phenix.refine[3] against the native dataset with the lowest twin fraction (0.12). The final model has a free-*R* factor 18.7% and a working-*R* factor of 16.3% (**Supplementary Table 1**), with excellent stereochemistry. A Ramachandran plot analysis using PROCHECK[9] indicates that 99.4% of all residues fall in the most favored region of the Ramachandran plot, with the remaining 0.06% residues in the additionally allowed region. One copy of the complex in the asymmetric unit contains residues 545-616 of PABPC1 and residues 1385-1399 of TNRC6C, while the second copy of the complex contains PABP residues 545-615 and TNRC6C residues 1386-1399. All molecular drawings were generated using PyMol (http://pymol.sourceforge.net).

In vitro **deadenylation assays.** Wild-type and mutant GST-λNHA-TNRC6C(1382-1690)-FLAG were expressed in Rosetta-2(DE3) *E. coli* cells (EMD Biosciences) and purified by two sequential affinity chromatography steps, first over glutathione Sepharose 4B resin (GE Life Sciences), followed by M2-FLAG affinity resin (Sigma). A plasmid encoding RL-5BoxB-pA was generated by digesting pRL-5BoxB[10] with XbaI and NotI restriction enzymes and subcloning the fragment into a RLpA-derived vector containing XbaI and NotI sites between the luciferase stop codon and the 98-nucleotide poly(A) tail. ^{32}P-radiolabeled 5-BoxB-pA transcripts were generated, and *in vitro* deadenylation assays were carried out as previously published[11] in the presence or absence of 172 nM wild-type or mutant GST-λNHA-TNRC6C(1382-1690)-FLAG proteins or a control wild-type construct lacking the λNHA tag.

SUPPLEMENTARY REFERENCES

1. He, B. *et al.* Rapid mutagenesis and purification of phage RNA polymerases. *Protein Expr Purif* **9**, 142-151 (1997).
2. Kabsch, W. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Crystallogr.* **26**, 795-800 (1993).
3. Adams, P.D. *et al.* Recent developments in the PHENIX software for automated crystallographic structure determination. *J Synchrotron Radiat* **11**, 53-55 (2004).
4. Adams, P.D. *et al.* PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr* **58**, 1948-1954 (2002).
5. Zwart, P.H. *et al.* Automated structure solution with the PHENIX suite. *Methods Mol Biol* **426**, 419-435 (2008).
6. Cohen, S.X. *et al.* ARP/wARP and molecular replacement: the next generation. *Acta Crystallogr D Biol Crystallogr* **64**, 49-60 (2008).
7. Perrakis, A., Sixma, T.K., Wilson, K.S. & Lamzin, V.S. wARP: improvement and extension of crystallographic phases by weighted averaging of multiple-refined dummy atomic models. *Acta Crystallogr D Biol Crystallogr* **53**, 448-455 (1997).
8. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132 (2004).
9. Laskowski, R.A., MacArthur, M.W., Moss, D.S. & Thornton, J.M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283-291 (1993).
10. Pillai, R.S., Artus, C.G. & Filipowicz, W. Tethering of human Ago proteins to mRNA mimics the miRNA-mediated repression of protein synthesis. *RNA* **10**, 1518-1525 (2004).
11. Fabian, M. *et al.* Mammalian miRNA RISC Recruits CAF1 and PABP to Affect PABP-Dependent Deadenylation. *Mol Cell* (2009).