A divergent Pumilio repeat protein family for pre-rRNA processing and mRNA localization

Chen Qiu[a], Kathleen L. McCann[b], Robert N. Wine[a], Susan J. Baserga[b,c,d,1], and Traci M. Tanaka Hall[a,1]

[a]Epigenetics and Stem Cell Biology Laboratory, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC 27709; and Departments of [b]Genetics, [c]Molecular Biophysics and Biochemistry, and [d]Therapeutic Radiology, Yale University School of Medicine, New Haven, CT 06520

Edited by David Baker, University of Washington, Seattle, WA, and approved November 19, 2014 (received for review April 25, 2014)

Pumilio/feminization of XX and XO animals (fem)-3 mRNA-binding factor (PUF) proteins bind sequence specifically to mRNA targets using a single-stranded RNA-binding domain comprising eight Pumilio (PUM) repeats. PUM repeats have now been identified in proteins that function in pre-rRNA processing, including human Puf-A and yeast Puf6. This is a role not previously ascribed to PUF proteins. Here we present crystal structures of human Puf-A that reveal a class of nucleic acid-binding proteins with 11 PUM repeats arranged in an "L"-like shape. In contrast to classical PUF proteins, Puf-A forms sequence-independent interactions with DNA or RNA, mediated by conserved basic residues. We demonstrate that equivalent basic residues in yeast Puf6 are important for RNA binding, pre-rRNA processing, and mRNA localization. Thus, PUM repeats can be assembled into alternative folds that bind to structured nucleic acids in addition to forming canonical eight-repeat crescent-shaped RNA-binding domains found in classical PUF proteins.

Puf-A | crystal structure | Puf6 | ribosome biogenesis | mRNA localization

RNA-binding proteins have evolved to perform biological functions requiring recognition of a variety of RNA ligands, from specific sequence motifs to structural shapes or combinations of sequence and structural features. Classical Pumilio/fem-3 mRNA-binding factor (PUF) proteins, named for *Drosophila melanogaster* Pumilio and *Caenorhabditis elegans* FBF (fem-3 mRNA-binding factor), are evolutionarily conserved in eukaryotes and regulate mRNA stability and translation in embryonic development, germ-line stem cell maintenance, and neurogenesis (1–3). Crystal structures of the characteristic ∼40-kDa RNA-binding domain, known as the Pumilio Homology Domain (PUM-HD) or PUF domain, from fly, human, mouse, yeast, and worm PUF proteins reveal eight α-helical PUM repeats of ∼36 aa each, arranged in a crescent shape (4–10). Single-stranded target RNA binds to the inner concave surface of the protein with the 5′ end of the RNA bound to the C terminus of the PUM-HD. The classical PUF protein, human Pumilio1 (PUM1), uses conserved side chains in its eight repeats to recognize eight RNA bases (4). Structural studies thus far have revealed only PUF proteins with eight PUM repeats.

New protein families with PUM repeats have emerged with the increasing availability of sequence data. One family includes human Puf-A (also known as KIAA0020) and its yeast ortholog, Puf6. Another includes yeast nucleolar protein 9 (Nop9) and its ortholog, human NOP9 (also known as C14orf21). Some of the known cellular functions of the Puf-A/Puf6 and Nop9 families differ from the mRNA regulatory function of classical PUF proteins. For example, Puf-A/Puf6 and Nop9 proteins are localized to the nucleolus, in contrast to the cytoplasmic localization of classical PUF proteins, and both yeast Puf6 and Nop9 are involved in ribosome biogenesis (11–14). Yeast Puf6 also binds to asymmetric synthesis of homothallic switching endonuclease (HO) 1 (ASH1) mRNA and represses its translation until it is localized at the bud tip of daughter cells, where Ash1 protein is asymmetrically segregated and inhibits the expression of HO endonuclease to prevent mating-type switching in the daughter cell (15). In addition to these functional differences, it is unclear how these new PUM repeat proteins would interact with target RNA. For example, only six PUM repeats are predicted in Puf-A and Puf6, and their RNA base-interacting residues are poorly conserved.

Vertebrate Puf-A functions appear to be important for diseases and embryonic development, but more knowledge is needed to connect vertebrate morbidities with molecular mechanisms. Human Puf-A changes localization from predominantly nucleolar to nuclear when cells are treated with transcriptional or topoisomerase inhibitors (14). It is overexpressed in breast cancer cells, with higher levels in more advanced stages (16). A peptide derived from human Puf-A residues 289–297 (RTLDKVLEV) has been classified as minor histocompatibility antigen HA-8 (17), which is associated with an increased risk of graft-versus-host disease (18, 19). Zebrafish Puf-A is involved in the development of eyes and primordial germ cells (20).

To examine the structural and functional relationship between Puf-A/Puf6 proteins and classical PUF proteins, we determined crystal structures of Puf-A. These structures reveal a new protein fold with 11 PUM repeats in an L-like shape, despite only six PUM repeats predicted by amino acid sequence. We show that Puf-A and Puf6 possess nucleic acid binding properties different

Significance

RNA regulation occurs at many levels including processing to mature forms, subcellular localization, and translation. RNA-binding proteins are crucial to direct and regulate these processes. Pumilio/feminization of XX and XO animals (fem)-3 mRNA-binding factor (PUF) proteins are RNA-binding proteins formed from eight α-helical repeats [Pumilio (PUM) repeats] that recognize specific mRNA sequences. Previous structural studies revealed characteristic curved structures and sequence specificity unique to these classical PUF proteins. We show here that PUM repeats also form different folds with 11 PUM repeats. Moreover, these proteins, exemplified by human Puf-A and yeast Puf6 proteins, recognize double-stranded RNA or DNA without sequence specificity. Interestingly, Puf-A and Puf6 PUM repeats lack specificity for RNA bases yet use residues at conserved positions on topologically equivalent protein surfaces for new nucleic acid recognition modes.

Author contributions: C.Q., K.L.M., R.N.W., S.J.B., and T.M.T.H. designed research; C.Q., K.L.M., and R.N.W. performed research; C.Q., K.L.M., R.N.W., S.J.B., and T.M.T.H. analyzed data; and C.Q., K.L.M., R.N.W., S.J.B., and T.M.T.H. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates and structure factors for the crystal structures have been deposited in the Research Collaboratory for Structural Bioinformatics (RCSB) Protein Data Bank (www.rcsb.org) [accession nos. 4WZR (for Puf-A alone), and 4WZW (for the Puf-A:DNA complex)].

[1]To whom correspondence may be addressed. Email: hall4@niehs.nih.gov or susan.baserga@yale.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1407634112/-/DCSupplemental.

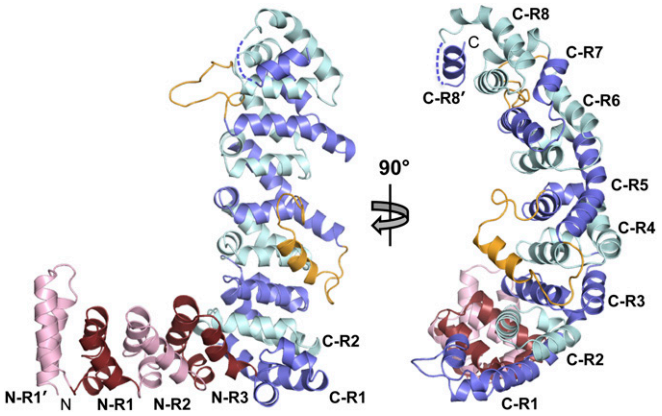

Fig. 1. Puf-A/Puf6 proteins represent a new PUM repeat fold. Ribbon diagram of a crystal structure of human Puf-A (residues 131–646). PUM repeats are colored alternately red and pink in the *N*-terminal domain (*N*-R1–*N*-R3) and light blue and pale cyan in the C-terminal domain (C-R1–C-R8). *N*- and C-terminal pseudorepeats are indicated (*N*-R1′ and C-R8′, respectively). Insertions in repeats C-R5 and C-R7 are colored gold.

from classical PUF proteins. Puf-A and Puf6 are more promiscuous and bind to double- or single-stranded RNA or DNA without sequence specificity, in contrast to classical PUF proteins, like PUM1, which bind to RNA bases with designable specificity (4, 21–23). We further demonstrate that conserved basic surfaces in and near the N-terminal PUM repeats of Puf6 are required for nucleic acid binding, pre-rRNA processing, and *ASH1* mRNA localization.

Results

Eleven PUM Repeats Form an L-Shaped Human Puf-A Protein. We determined a 2.2-Å resolution crystal structure of human Puf-A (Table S1). The structure revealed a new nucleic acid binding fold, related structurally to that found in the classical PUF proteins. Puf-A is composed of two subdomains of PUM repeats that form a right angle (Fig. 1). The C-terminal subdomain (residues 278–646) retains the curved assembly with eight PUM repeats (C-R1 to C-R8) and a C-terminal pseudorepeat (C-R8′) seen in human PUM1 (Fig. 2*A*). The *N*-terminal subdomain (residues 131–277) includes three additional PUM repeats (*N*-R1 to *N*-R3) flanked by an *N*-terminal pseudorepeat (*N*-R1′). The protein construct for crystallization lacked the nucleolar localization sequence and residues preceding it (residues 1–123). Eight of the 11 PUM repeats in Puf-A are structurally similar to prototypical PUM repeats (rmsd, 1.0–1.7 Å over 36 Cα atoms), whereas three repeats diverge in structure from prototypical repeats (C-R1, C-R5, and C-R7) (Fig. S1). Only six repeats were predicted based on amino acid sequence: *N*-R1 to *N*-R3 and C-R3 to C-R5 (24).

Although the C-terminal region of Puf-A is structurally similar to that of PUM1, its function appears distinct. For classical PUF proteins, a five-residue motif in the α2 helix of each PUM repeat defines the specificity of RNA base recognition (Fig. 2 *B* and *C* and Fig. S2) (4, 21). These RNA-interacting residues are highly conserved among classical PUF proteins but are different in Puf-A. For example, repeats C-R2 through C-R5 contain a leucine or a methionine at the fifth position. These hydrophobic residues are buried beneath a long insertion in repeat C-R5 between helices α2 and α3 (Fig. 2*D*), making it impossible for Puf-A to interact with RNA as classical PUF proteins do. Nevertheless, 12 of the 33 residues in the "RNA-binding" motifs are basic residues (Fig. 2*B*). Given these structural differences, we asked whether Puf-A binds RNA but adopts a different RNA-binding mode.

Puf-A Binds to Structured Nucleic Acids Using Conserved Basic Residues. Three distinct patches of positive electrostatic potential on the surface of Puf-A appear well suited to interact with negatively charged nucleic acids (labeled 1, 2, and 3 in Fig. S3). We obtained crystals of Puf-A with dsDNA, although we attempted crystallization with dsRNA, ssRNA, stem-loop RNA, and ssDNA. We

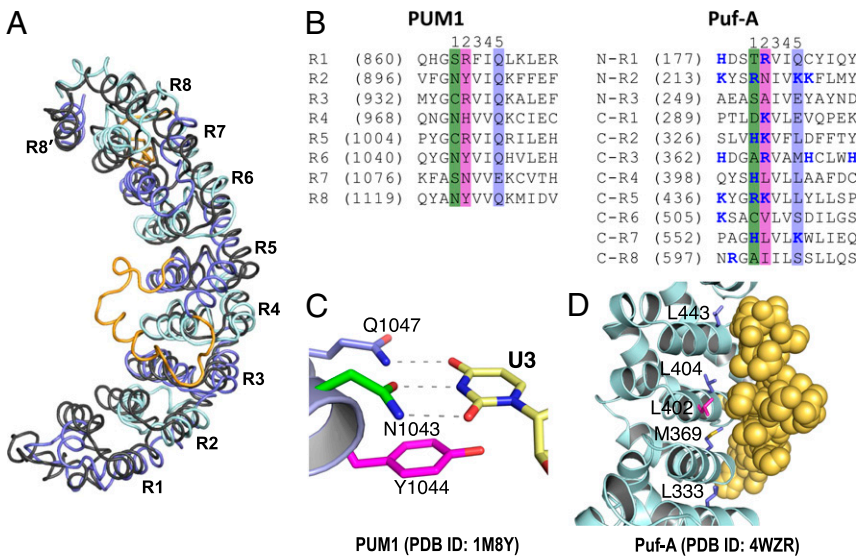

Fig. 2. (*A*) Superposition of the C-terminal domain of Puf-A with human PUM1. The Cα trace of Puf-A is colored as in Fig. 1, and PUM1 is colored gray (rmsd, 2.84 Å over 268 Cα atoms). (*B*) Alignment of α2 helix amino acid sequences of PUM1 and Puf-A. The five-residue sequences that recognize RNA in PUM1 are numbered 1–5 above the sequences. Residues in PUM1 that recognize the edges of bases (first and fifth positions) are highlighted green and blue, respectively, whereas residues that stack with RNA bases (second position) are highlighted magenta. Equivalent positions in Puf-A are indicated. Basic residues within or near the α2 helices are shown in blue. (*C*) Representative sequence-specific interaction of a PUM repeat with an RNA base. Interaction of residues in PUM1 repeat 6 with base U3 is shown. Side chains are colored as in *B*. Dotted lines indicate hydrogen bond interactions. (*D*) Hydrophobic side chains in repeats C-R2 to C-R5 are buried beneath a long insertion in repeat C-R5. Side chains in repeats C-R2 to C-R5 are shown in stick representation, colored as in *B*. Atoms in the long insertion are shown as gold space-filling spheres.

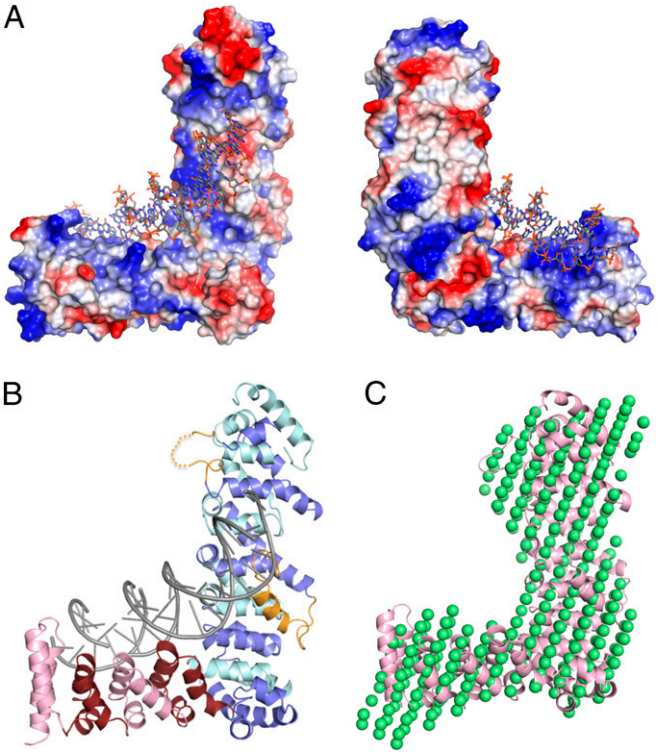

Fig. 3. Puf-A/Puf6 family proteins interact with structured nucleic acids. (*A*) Surface representation of a crystal structure of human Puf-A in complex with double-stranded DNA. The molecular surface of Puf-A is colored by electrostatic potential and shown with 14 bp of DNA (modeled as poly G:poly C), which formed a continuous helix through the crystals. DNA is shown as a stick representation with atoms colored by element (carbon, gray; nitrogen, blue; oxygen, red; phosphorus, orange). (*B*) Ribbon diagram of Puf-A in complex with DNA. Puf-A is colored as in Fig. 1, and the DNA is shown as a cartoon representation. The protein in the structure of the complex is very similar to the structure of Puf-A alone, with an rmsd of 1.03 Å over 495 Cα atoms. (*C*) Ab initio SAXS envelope model for *S. cerevisiae* Puf6 (green spheres) superimposed with ribbon diagram of Puf-A (pink).

determined a 3.0-Å resolution crystal structure of Puf-A with 14 bp of DNA visible (Fig. 3 *A* and *B*). The electron density map shows nonideal B-form DNA extending through the crystals, although the density was too weak to model the correct DNA sequence (Fig. S4*A*). The DNA binds along the concave faces of the *N*- and C-terminal subdomains of the protein. Protein:DNA interactions visible in the crystal structure are sequence-independent, because they involve the positively charged residues of Puf-A in basic surface patches 1 and 2 and the negatively charged phosphate backbone of the DNA (Fig. S4*B*).

Puf-A, like other PUM repeat proteins, appears to interact with nucleic acids using amino acid side chains from the α2 helices of PUM repeats. In classical PUF proteins like PUM1, protein:RNA interactions are dominated by recognition of bases through the conserved RNA-binding motifs in the α2 helices (Fig. 2*B*). In addition, residues at the *N* termini of the α2 helices of PUM1 and other classical PUF proteins make van der Waals contacts with the ribose bases of the single-stranded RNA (Fig. S4*C*, *Right*). In Puf-A, many basic residues within or near the PUM1 RNA-binding motifs contact the phosphate backbone of DNA (Fig. S4*B*). Moreover, conserved aromatic residues at the *N* termini of the α2 helices of Puf-A repeats C-R4 and C-R5 are near the major groove of the DNA, suggesting they may be involved in nucleic acid interaction (Fig. S4*C*, *Left*). A unique feature of Puf-A is a long inserted loop in repeat C-R5 that contains an α helix whose N terminus is near the minor groove of the DNA (Fig. 3*B*), suggesting the potential to form interactions with target nucleic acids.

Puf-A and Puf6 Bind Sequence Independently to RNA or DNA. To probe structure-function questions, we turned to the yeast ortholog Puf6 and identified a basic surface in the N-terminal subdomain of Puf6 that is important for in vitro RNA binding. Puf6 and Puf-A share ∼24% sequence identity. We performed small-angle X-ray scattering (SAXS) analysis of Puf6 and through ab initio modeling confirmed that Puf6 is structurally similar to Puf-A (Fig. 3*C*). We therefore used the structure of Puf-A to guide analyses of yeast Puf6's established functions in pre-rRNA processing and *ASH1* mRNA localization. We assessed the nucleic acid-binding ability of Puf6 and Puf-A using fluorescence polarization assays with representative RNA or DNA. Both Puf6 and Puf-A bound without apparent specificity to single- or double-stranded RNA or DNA, and binding was dramatically reduced when the salt concentration was raised from 50 mM to 150 mM or tRNA was added (Table S2 and Fig. S5). This general nucleic acid-binding activity is distinct from sequence-specific binding to single-stranded RNA by classical PUF proteins, such as human PUM2 (Table S2). Using the Puf-A:DNA crystal structure as a guide, we mutated five groups of conserved basic residues to alanine in Puf6 (Fig. 4*A* and Figs. S2 and S3). Mutants M1a/M1b, M2, and M3a/M3b correspond to basic surface patches 1–3, respectively. Mutant M3a includes residues in basic patch 3 that do not contact DNA in the crystal structure. In vitro binding to single- or double-stranded RNA was reduced for mutants M1a and M1b but was not affected for mutant M2, M3a, or M3b (Table 1). Thus, basic patch 1 is most important for in vitro RNA binding.

Conserved Basic Residues on Puf6 Are Required for Pre-rRNA Processing and *ASH1* mRNA Localization. We tested whether mutation of the conserved basic amino acids in Puf6 interferes with pre-rRNA processing in the yeast *S. cerevisiae* and found that basic patch 1a residues are important for 7S pre-rRNA processing. Previously, it has been shown that yeast lacking Puf6 accumulate the 35S, 27S and 7S pre-rRNAs (11), all intermediates that result from the cleavage events that produce the large ribosomal subunit (LSU; Fig. 4*B*) (25). We created a yeast strain where the endogenous Puf6 can be conditionally and acutely depleted by growth in glucose (Fig. 4*C*). Either the unmutated (wild-type or WT) or each of the Puf6 mutations in Fig. 4*A* was ectopically expressed from a plasmid (p414GPD). A strain bearing the empty vector (EV) was included as a negative control. The endogenous Puf6 was tagged with a triple-HA epitope, and the plasmid-expressed Puf6 was tagged with a triple-FLAG epitope. Western blotting indicated that after growth of this strain in glucose for 24 h, the endogenous Puf6 was reduced to undetectable levels (Fig. 4*C*). Similarly, each plasmid-expressed Puf6 was expressed to comparable levels after 72 h at 17 °C (Fig. 4*C*). We analyzed the in vivo effects on pre-rRNA processing of the five sets of Puf6 mutations (Fig. 4*A*) by assessing yeast growth rate and the levels of intermediate pre-rRNAs by Northern blotting (Fig. S6 and Fig. 4*D*). We monitored growth at 17 °C, as was done previously for Puf6 (11), because defects in ribosome assembly often result in cold sensitivity (26). Our results indicate that Puf6-depleted yeast (EV) grew more slowly than yeast expressing WT Puf6 (Fig. S6*A*). Mutant M1a showed mildly reduced cell growth compared with WT, whereas the other mutants grew similarly to WT. In Northern blot analysis for pre-rRNA processing defects, mutant M1a displayed a statistically significant defect in 7S pre-rRNA processing, which was comparable to Puf6-depleted yeast (EV; $P < 0.0001$) (Fig. 4*D*). Unlike previous work (11), no strain exhibited defects in 27S pre-rRNA processing. This may be due to differential effects on pre-rRNA processing when Puf6 is acutely depleted, as we have done here, compared with when the *PUF6* gene is deleted, as was done previously. This may be an example of

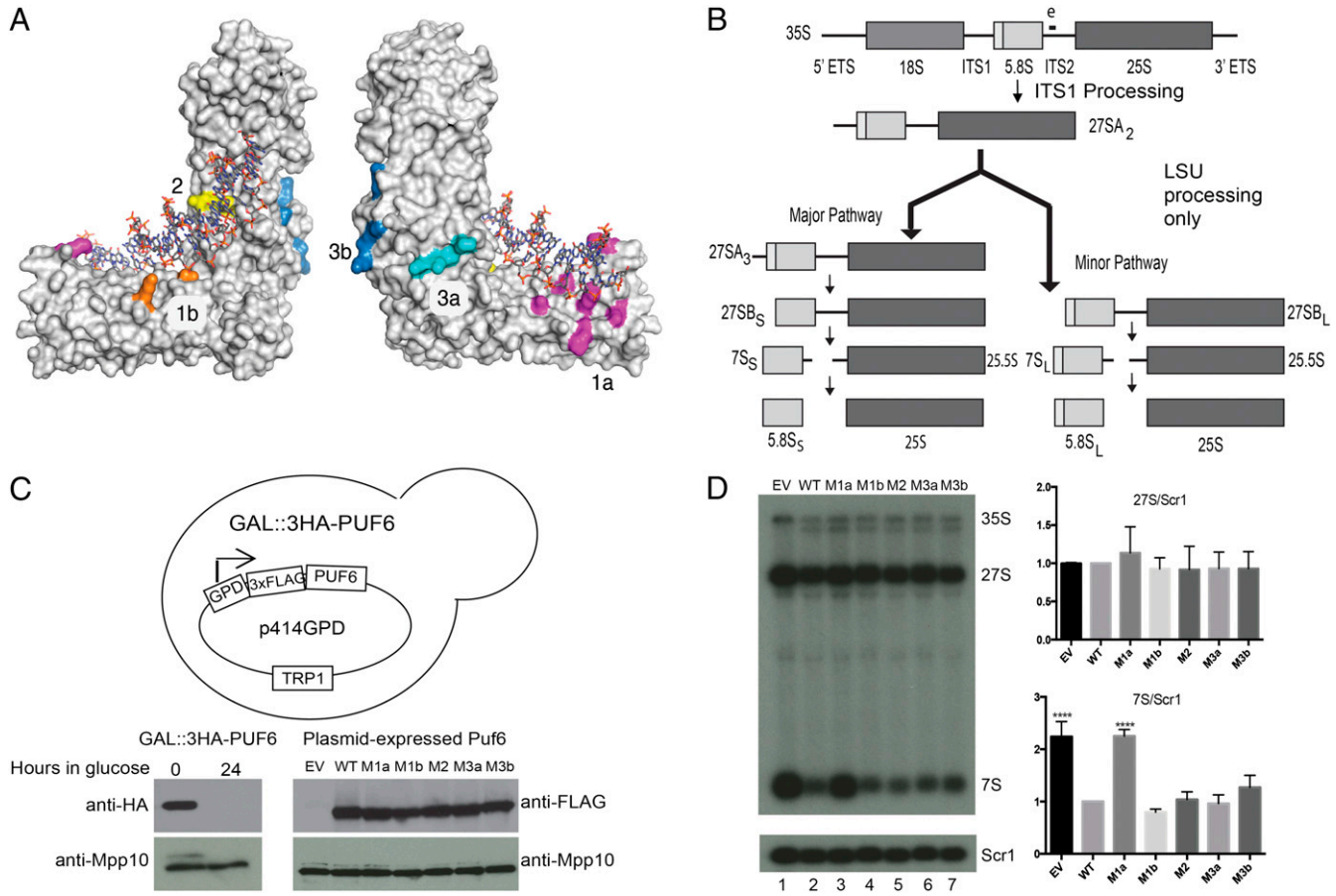

Fig. 4. Structure-guided mutations of Puf6 basic residues disrupt growth and pre-rRNA processing. (*A*) Location of mutation groups on the surface of Puf-A. Five sets of mutations of conserved basic residues to alanine in yeast Puf6 were created (mutant M1a, K139, K148, H177, R181, R216, and K221; mutant M1b, K228 and R268; mutant M2, K436 and K440; mutant M3a, K381, K385, and K388; mutant M3b, K472, K473, R479, and R480), and the locations of the equivalent residues (mutant M1a, K128, K137, H168, R172, K207, and K212; mutant M1b, R219 and R259; mutant M2, K427 and R431; mutant M3a, K372, K376, and K379; mutant M3b, K465, K466, R472, and H473) are shown on a molecular surface representation of the crystal structure of human Puf-A in complex with DNA (stick representation colored by element). (*B*) Diagram focused on the LSU pre-rRNA processing steps in yeast. The pre-rRNA is transcribed as a 35S polycistronic precursor and is processed through a number of cleavage events to remove the external transcribed spacers (5′ and 3′ ETS) and internal transcribed spacers (ITS1 and 2) to produce the mature 18S, 5.8S, and 25S rRNAs. Oligonucleotide probe e in ITS2 was used for Northern blotting (indicated). This probe detects all 27S and 7S pre-rRNAs. (*C*) Schematic of the yeast strain used for testing Puf6 mutants. Endogenous Puf6 was placed under the control of the inducible *GAL* promoter, and FLAG-tagged WT or mutant Puf6 was constitutively expressed from the p414GPD plasmid. A strain bearing the EV was also included. Total yeast extract was separated on an SDS/PAGE gel and analyzed by Western blotting with either anti-HA, anti-FLAG, or anti-Mpp10 antibodies. Mpp10 levels serve as a loading control. (*D*) Mutation of conserved basic residues in patch 1a of Puf6 disrupts processing of the 7S precursor to the 5.8S rRNA. Total RNA was extracted from yeast bearing WT, EV, or mutant Puf6 after depletion of endogenous Puf6 in glucose for 72 h at 17 °C and detected on a Northern blot with oligonucleotide probe e in ITS2. The intensity of the bands was quantified, and the ratios of 27S or 7S to the loading control RNA, Scr1, were calculated. Bar graphs created in GraphPad PRISM plot the averages from three replicate experiments, calculated with error bars representing the SEM. The significance of the ratios of 27S/Scr1 and 7S/Scr1 of each Puf6 mutant or of Puf6 depleted yeast (EV) compared with WT was evaluated using one-way ANOVA. ****$P < 0.0001$.

how yeast adapt to chronic defects in ribosome biogenesis (27). These results indicate that the conserved basic residues in mutant group M1a on the inner concave surface of the N-terminal region of Puf6 are important for 7S pre-rRNA processing.

Similarly, we found that conserved basic residues on the surface of Puf6 containing mutant group M1a are important for localization of *ASH1* mRNA in *S. cerevisiae*. To assess a second in vivo function of Puf6, we scored *ASH1* mRNA localization at late anaphase by in situ hybridization (Fig. 5), using the yeast strain and WT, EV, and mutant plasmids created to test pre-rRNA processing defects (Fig. 4*C*). We evaluated effects on *ASH1* mRNA localization after growth in glucose at 30 °C to deplete endogenous Puf6 but to avoid pre-rRNA processing defects (no growth defects were observed upon Puf6 depletion at 30 °C) (Fig. S6*B*). We found significantly different *ASH1* localization patterns between yeast expressing WT Puf6 and those depleted of Puf6 (EV; $P < 0.0001$). Yeast expressing WT Puf6 had a greater proportion of daughter cells with *ASH1* mRNA at the bud tip compared with yeast depleted of Puf6 (55% WT vs. 14% EV). Expression of Puf6 mutants M1a and M3a showed the strongest defects in *ASH1* localization (~23% daughter cells with bud tip localization; Fig. 5), with *ASH1* localization patterns

Table 1. In vitro binding of Puf6 and Puf6 mutants to representative RNAs (K_d, nM)

Protein	ssRNA	dsRNA	Stem-loop RNA 1	Stem-loop RNA 2
Puf6 WT	42.5 ± 6.7	34.0 ± 0.2	27.7 ± 3.1	20.4 ± 2.2
M1a	162 ± 35	192 ± 18	119 ± 5.7	93.0 ± 10.9
M1b	112 ± 13	124 ± 13	54.6 ± 5.5	41.7 ± 1.2
M2	49.8 ± 1.6	27.4 ± 1.9	27.5 ± 3.3	25.3 ± 0.6
M3a	31.4 ± 0.4	29.6 ± 2.5	27.4 ± 2.5	20.1 ± 3.1
M3b	35.1 ± 0.2	33.6 ± 3.1	26.7 ± 2.5	22.8 ± 2.7

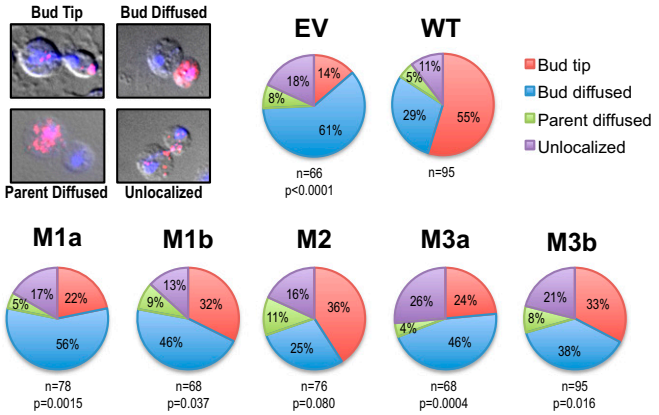

Fig. 5. Structure-guided mutations of Puf6 basic residues disrupt *ASH1* mRNA localization to the daughter cell bud tip. *ASH1* mRNA localization in parent–daughter pairs was classified into four categories (*Top Left*, representative images). The *GAL::3HA-PUF6* strain was transformed with EV or vectors expressing WT or mutant Puf6. After shifting from galactose to glucose medium for 18–20 h at 30 °C to deplete 3HA-Puf6 expression, parent–daughter bud pairs in late anaphase displaying *ASH1* mRNA signal were identified (*n*, the number of pairs counted) and classified by *ASH1* mRNA localization pattern, and the proportions of each pattern are shown on pie charts. The distributions were evaluated using Fisher's exact test, and *P* values versus 3xFLAG-WT control are indicated.

significantly different from WT ($P = 0.0015$ and 0.0004 for M1a and M3a, respectively) and not significantly different from the EV control ($P = 0.63$ and 0.19, respectively). The corresponding basic patches are adjacent to each other, in and near the *N*-terminal PUM repeats of the Puf-A crystal structures (Fig. 4*A*). Yeast expressing mutants M1b and M3b displayed intermediate defects in *ASH1* localization (~33% daughter cells with bud tip localization). The localization patterns for mutants M1b and M3b were significantly different from WT ($0.01 < P < 0.05$), and M1b was also not significantly different from EV ($P = 0.07$). However, M3b was significantly different from EV ($P = 0.015$), consistent with the intermediate *ASH1* localization defect. Basic patch 2 is not important for *ASH1* bud tip localization: The *ASH1* localization patterns for yeast expressing mutant M2 were not significantly different from WT (Fig. 4*A*; $P = 0.08$). Together these data indicate that the protein surface containing basic patches 1a and 3a is important for in vivo *ASH1* mRNA localization.

Discussion

Our crystal structures and mutational analyses of Puf-A/Puf6 proteins identify a family of atypical PUM repeat-containing proteins. In contrast to the sequence-specific PUF proteins represented by the namesake Pumilio and FBF proteins, Puf-A/Puf6 proteins appear to bind to nucleic acids without sequence specificity. Although the C-terminal repeats bear resemblance to the curved eight-repeat PUF proteins, basic residues on the concave surface typically used for RNA interaction are not important for Puf-A/Puf6 RNA interaction. Instead, conserved basic residues on the inner concave surface of the N-terminal extension of Puf6 are necessary for nucleic acid interaction, pre-rRNA processing, and *ASH1* mRNA localization.

Our experiments suggest that the Puf-A/Puf6 proteins bind equally well and without apparent sequence specificity to single- or double-stranded RNA or DNA, leading us to speculate that Puf-A/Puf6 protein specificity for its RNA targets (e.g., pre-rRNA and *ASH1* mRNA for Puf6) requires binding partners. For *ASH1* mRNA, switching deficient 5p (Swi5p)-dependent HO expression protein 2 (She2) binds cotranscriptionally, and then Puf6 and localization of *ASH1* mRNA protein 1 (Loc1) join the ribonucleoprotein (RNP) complex as the RNA passes through the nucleolus. She2 associates in vivo with Puf6 or Loc1 in an RNA-independent manner; thus, She2 or Loc1 may enhance Puf6 specificity. However, no interaction between Puf6 and She2 or Loc1 was detected in vitro with purified components (28, 29). This inconsistency may indicate that other factors are necessary for Puf6 to associate with the She2/Loc1 complex on *ASH1* mRNA. Such factors may bridge the interaction between She2/Loc1 and Puf6 or perhaps remodel the RNP so that Puf6 can bind. Our mutational analyses showed that conserved basic surfaces in and near the *N*-terminal region of Puf6 are important for in vitro RNA-binding, pre-rRNA processing, and *ASH1* mRNA localization. The representative nucleic acids in our in vitro binding assays could be missing sequence or structural elements that are recognized by basic patch 3. Alternatively, rather than interacting with RNA, basic surface 3 may interact with proteins in complexes specific for *ASH1* mRNA localization.

Many α-helical repeat protein families form adaptable scaffolds for ligand binding with varying numbers of modular repeats (30–32), and this work demonstrates that PUM repeats form structures with more than eight repeats and with nucleic acid binding properties different from classical PUF proteins like PUM1. Puf-A/Puf6 proteins bind RNA or DNA regardless of sequence, whereas PUM1 binds specific RNA sequences. Although the proteins have different functions, the overall structure of Puf-A is reminiscent of the size and shape of the Huntingtin/elongation factor-3/protein phosphatase 2A/target of rapamycin 1 (HEAT) repeat-containing α-subunit of the clathrin adaptor complex, AP-2 (33), and nucleic acid interaction is similar to the enfolding of DNA by transcription terminator MTERF1 (34) (Fig. 6). These similarities reflect the utility of α-helical repeats to bind protein or nucleic acid ligands. Our structures of Puf-A follow this theme and establish the expanded use of PUM repeats for nucleic acid recognition.

Methods

Detailed methods are available in *SI Methods*.

Protein Expression and Purification. Puf-A and Puf6 proteins were expressed as His_6-SUMO fusion proteins in *Escherichia coli* strain BL21-CodonPlus(DE3)-RIL (Agilent) and purified with Ni^{2+}-NTA resin. After cleavage of the fusion protein with Ulp1 protease, Puf-A or Puf6 was purified by heparin affinity and size exclusion chromatography.

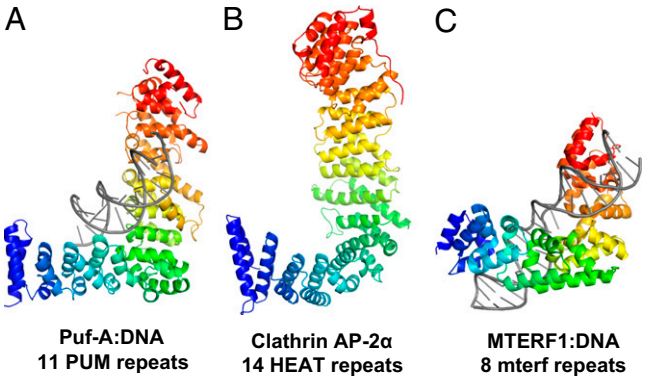

Fig. 6. Comparison of Puf-A structures with other tandem helical repeat proteins. (*A*) Structure of Puf-A complexed with dsDNA. (*B*) Structure of the clathrin adaptor protein AP-2 α-subunit, which contains 14 HEAT repeats [Protein Data Bank (PDB) ID code 2VGL] (33). Although each HEAT repeat comprises two α-helices, the arrangement of repeats 1–11 in AP-2α is similar to that in Puf-A. (*C*) Structure of human mitochondrial transcription terminator MTERF1 in complex with DNA (PDB ID code 3MVA) (34). Similar to the Puf-A:DNA structure, the DNA molecule binds along the concave face of MTERF1. MTERF1:DNA contacts are mediated by a combination of sequence-specific and sequence-independent interactions.

Crystallization, X-Ray Structure Determination, and SAXS. Crystals of Puf-A or a Puf-A:DNA complex were obtained by hanging drop vapor diffusion. X-ray diffraction data were collected at the Southeast Regional Collaborative Access Team (SER-CAT) Beamline 22-ID at the Advanced Photon Source, Argonne National Laboratories. Data statistics are shown in Table S1. The crystal structure of Puf-A was determined by SeMet single wavelength anomalous diffraction and of the Puf-A:DNA complex by molecular replacement. A Puf6 SAXS model was calculated from scattering curves collected with three different concentrations of Puf6 (1.0, 2.3, and 3.2 mg/mL).

RNA-Binding Assays. RNA-binding affinities were determined for Puf-A, Puf6, and PUM2 proteins by fluorescence polarization using 5′-fluorescein–labeled RNA or DNA. Assays were performed in triplicate, and mean K_d values with SEM are shown in Table 1 and Table S2.

Yeast Analyses. A *GAL::3HA-PUF6* strain was created that expresses 3HA-tagged endogenous Puf6 when grown in medium containing galactose but that represses endogenous Puf6 expression when grown in medium containing glucose. The 3xFLAG-tagged unmutated (WT) or mutant Puf6 was expressed from a plasmid in Puf6-depleted yeast. A strain transformed with EV was also included. RNA was harvested after shifting from galactose to glucose medium for 72 h at 17 °C and analyzed on a Northern blot using an oligonucleotide probe complementary to ITS2 of the yeast pre-rRNA. The blot was also probed for Scr1 RNA levels using a complementary oligonucleotide, as in Li et al. (11). The 7S and 27S pre-rRNAs were quantified on a Biorad Personal Molecular Imager, and the ratios of 7S or 27S to the RNA Scr1 were calculated. Each experiment was repeated three times, and averages and error bars (SEM) were calculated using GraphPad PRISM. Significance was determined using one-way ANOVA. Localization of *ASH1* mRNA by in situ hybridization was performed for yeast spheroplasts after shifting from galactose to glucose medium for 18–20 h at 30 °C using the same *GAL::3HA-PUF6* strain and EV, WT, or mutant Puf6 plasmids as for the pre-rRNA processing analyses. Significance of differences in *ASH1* mRNA localization versus WT or EV controls was determined using Fisher's exact test (SAS 9.3, SAS Institute, Inc.).

ACKNOWLEDGMENTS. We thank L. Pedersen and the staff of the Southeast Regional Collaborative Access Team (SER-CAT) beamlines for assistance with X-ray data collection, J. Williams and the National Institute of Environmental Health Sciences (NIEHS) Protein Microcharacterization Facility for mass spectrometric analyses, J. Tucker and A. Janoshazi of the NIEHS Fluorescence Microscopy Imaging Center for imaging and data analysis support, G. Kissling for statistical analyses, and A. Clark for discussions regarding yeast growth and spheroplasting. This work was supported in part by the Intramural Research Program of the National Institutes of Health, National Institute of Environmental Health Sciences and National Institutes of Health Grant GM52581 (to S.J.B.). The Advanced Photon Source used for this study is supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract W-31-109-Eng-38. The Advanced Light Source used for this study is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the US Department of Energy under Contract DE-AC02-05CH11231.

1. Wang Y, Wang Z, Hall TMT (2013) Engineered proteins with Pumilio/fem-3 mRNA binding factor scaffold to manipulate RNA metabolism. *FEBS J* 280(16):3755–3767.
2. Wharton RP, Aggarwal AK (2006) mRNA regulation by Puf domain proteins. *Sci STKE* 2006(354):pe37.
3. Wickens M, Bernstein DS, Kimble J, Parker R (2002) A PUF family portrait: 3′UTR regulation as a way of life. *Trends Genet* 18(3):150–157.
4. Wang X, McLachlan J, Zamore PD, Hall TMT (2002) Modular recognition of RNA by a human Pumilio-homology domain. *Cell* 110(4):501–512.
5. Jenkins HT, Baker-Wilding R, Edwards TA (2009) Structure and RNA binding of the mouse Pumilio-2 Puf domain. *J Struct Biol* 167(3):271–276.
6. Zhu D, Stumpf CR, Krahn JM, Wickens M, Hall TMT (2009) A 5′ cytosine binding pocket in Puf3p specifies regulation of mitochondrial mRNAs. *Proc Natl Acad Sci USA* 106(48): 20192–20197.
7. Wang Y, Opperman L, Wickens M, Hall TMT (2009) Structural basis for specific recognition of multiple mRNA targets by a PUF regulatory protein. *Proc Natl Acad Sci USA* 106(48):20186–20191.
8. Edwards TA, Pyle SE, Wharton RP, Aggarwal AK (2001) Structure of Pumilio reveals similarity between RNA and peptide binding motifs. *Cell* 105(2):281–289.
9. Wang X, Zamore PD, Hall TMT (2001) Crystal structure of a Pumilio homology domain. *Mol Cell* 7(4):855–865.
10. Miller MT, Higgin JJ, Hall TMT (2008) Basis of altered RNA-binding specificity by PUF proteins revealed by crystal structures of yeast Puf4p. *Nat Struct Mol Biol* 15(4): 397–402.
11. Li Z, et al. (2009) Rational extension of the ribosome biogenesis pathway using network-guided genetics. *PLoS Biol* 7(10):e1000213.
12. Tafforeau L, et al. (2013) The complexity of human ribosome biogenesis revealed by systematic nucleolar screening of pre-rRNA processing factors. *Mol Cell* 51(4):539–551.
13. Thomson E, Rappsilber J, Tollervey D (2007) Nop9 is an RNA binding protein present in pre-40S ribosomes and required for 18S rRNA synthesis in yeast. *RNA* 13(12):2165–2174.
14. Chang HY, et al. (2011) hPuf-A/KIAA0020 modulates PARP-1 cleavage upon genotoxic stress. *Cancer Res* 71(3):1126–1134.
15. Gu W, Deng Y, Zenklusen D, Singer RH (2004) A new yeast PUF family protein, Puf6p, represses *ASH1* mRNA translation and is required for its localization. *Genes Dev* 18(12):1452–1465.
16. Fan CC, et al. (2013) Upregulated hPuf-A promotes breast cancer tumorigenesis. *Tumour Biol* 34(5):2557–2564.
17. Brickner AG, et al. (2001) The immunogenicity of a new human minor histocompatibility antigen results from differential antigen processing. *J Exp Med* 193(2):195–206.
18. Akatsuka Y, et al. (2003) Disparity for a newly identified minor histocompatibility antigen, HA-8, correlates with acute graft-versus-host disease after haematopoietic stem cell transplantation from an HLA-identical sibling. *Br J Haematol* 123(4):671–675.
19. Turpeinen H, et al. (2013) Minor histocompatibility antigens as determinants for graft-versus-host disease after allogeneic haematopoietic stem cell transplantation. *Int J Immunogenet* 40(6):495–501.
20. Kuo MW, et al. (2009) A novel puf-A gene predicted from evolutionary analysis is involved in the development of eyes and primordial germ-cells. *PLoS ONE* 4(3):e4980.
21. Cheong CG, Hall TMT (2006) Engineering RNA sequence specificity of Pumilio repeats. *Proc Natl Acad Sci USA* 103(37):13635–13639.
22. Dong S, et al. (2011) Specific and modular binding code for cytosine recognition in Pumilio/FBF (PUF) RNA-binding domains. *J Biol Chem* 286(30):26732–26742.
23. Filipovska A, Razif MF, Nygård KK, Rackham O (2011) A universal code for RNA recognition by PUF proteins. *Nat Chem Biol* 7(7):425–427.
24. Anonymous; UniProt Consortium (2014) Activities at the Universal Protein Resource (UniProt). *Nucleic Acids Res* 42(Database issue):D191–D198.
25. Woolford JL, Jr, Baserga SJ (2013) Ribosome biogenesis in the yeast *Saccharomyces cerevisiae*. *Genetics* 195(3):643–681.
26. Guthrie C, Nashimoto H, Nomura M (1969) Structure and function of E. coli ribosomes. 8. Cold-sensitive mutants defective in ribosome assembly. *Proc Natl Acad Sci USA* 63(2):384–391.
27. Sulima SO, et al. (2014) Bypass of the pre-60S ribosomal quality control as a pathway to oncogenesis. *Proc Natl Acad Sci USA* 111(15):5640–5645.
28. Niedner A, Müller M, Moorthy BT, Jansen RP, Niessing D (2013) Role of Loc1p in assembly and reorganization of nuclear *ASH1* messenger ribonucleoprotein particles in yeast. *Proc Natl Acad Sci USA* 110(52):E5049–E5058.
29. Müller M, et al. (2011) A cytoplasmic complex mediates specific mRNA recognition and localization in yeast. *PLoS Biol* 9(4):e1000611.
30. Rubinson EH, Eichman BF (2012) Nucleic acid recognition by tandem helical repeats. *Curr Opin Struct Biol* 22(1):101–109.
31. Tewari R, Bailes E, Bunting KA, Coates JC (2010) Armadillo-repeat protein functions: Questions for little creatures. *Trends Cell Biol* 20(8):470–481.
32. Xu W, Kimelman D (2007) Mechanistic insights from structural studies of beta-catenin and its binding partners. *J Cell Sci* 120(Pt 19):3337–3344.
33. Collins BM, McCoy AJ, Kent HM, Evans PR, Owen DJ (2002) Molecular architecture and functional model of the endocytic AP2 complex. *Cell* 109(4):523–535.
34. Yakubovskaya E, Mejia E, Byrnes J, Hambardjieva E, Garcia-Diaz M (2010) Helix unwinding and base flipping enable human MTERF1 to terminate mitochondrial transcription. *Cell* 141(6):982–993.