ACS SyntheticBiology

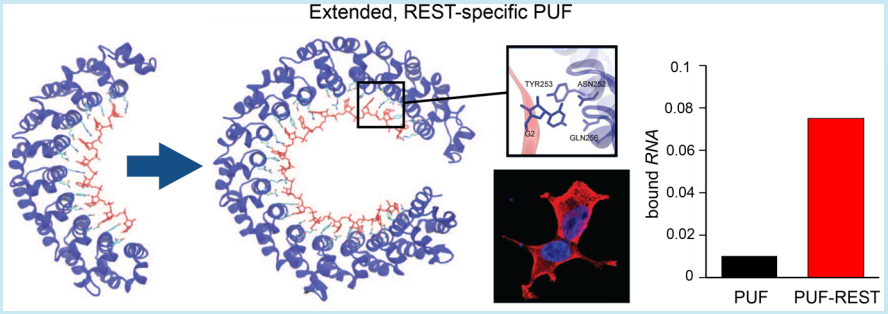
pubs.acs.org/synthbio Research Article

Engineering REST-Specific Synthetic PUF Proteins to Control Neuronal Gene Expression: A Combined Experimental and Computational Study

Stefania Criscuolo,[#] Mahad Gatti Iou,[#] Assunta Merolla,[#] Luca Maragliano,[*] Fabrizia Cesca,[*] and Fabio Benfenati[*]

Cite This: *ACS Synth. Biol.* 2020, 9, 2039–2054 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information



ABSTRACT: Regulation of gene transcription is an essential mechanism for differentiation and adaptation of organisms. A key actor in this regulation process is the repressor element 1 (RE1)-silencing transcription factor (REST), a transcriptional repressor that controls more than 2000 putative target genes, most of which are neuron-specific. With the purpose of modulating REST expression, we exploited synthetic, *ad hoc* designed, RNA binding proteins (RBPs) able to specifically target and dock to REST mRNA. Among the various families of RBPs, we focused on the Pumilio and FBF (PUF) proteins, present in all eukaryotic organisms and controlling a variety of cellular functions. Here, a combined experimental and computational approach was used to design and test 8- and 16-repeat PUF proteins specific for REST mRNA. We explored the conformational properties and atomic features of the PUF-RNA recognition code by Molecular Dynamics simulations. Biochemical assays revealed that the 8- and 16-repeat PUF-based variants specifically bind the endogenous REST mRNA without affecting its translational regulation. The data also indicate a key role of stacking residues in determining the binding specificity. The newly characterized REST-specific PUF-based constructs act as excellent RNA-binding modules and represent a versatile and functional platform to specifically target REST mRNA and modulate its endogenous expression.

KEYWORDS: *gene transcription, RNA binding proteins (RBPs), molecular dynamics simulation, computational modeling, free energy calculations*

Neuronal maturation occurs through an intricate network of signaling molecules and epigenetic modifications that leads to the acquisition and maintenance of neuronal identity. A key actor in this regulation process is the repressor element 1 (RE1)-silencing transcription factor (REST), also known as neuron-restrictive silencer factor (NRSF). REST is a member of the Kruppel-type zinc finger transcription factor family, and represses transcription by binding a DNA sequence of 21−23 bp, called neuron-restrictive silencer element (NRSE, also known as RE1).[1,2] Genome-wide chromatin immunoprecipitation (ChIP) and bioinformatics studies have revealed more than 2000 putative REST target genes in human embryonic stem cells (ESC) and ESC-derived neurons.[3] Due to its involvement in several cellular processes, even subtle alterations of REST levels and/or activity may give rise to severe pathologies.[4] High REST expression levels have been associated with medulloblastoma and glioblastoma,[5,6] cerebral ischemia,[7] Down syndrome,[8] and Huntington's disease,[9] while low REST expression has been correlated to Alzheimer's disease,[10] cardiac hypertrophy,[11] and prostate, breast, and small cell lung cancers.[12−14] Given the major impact of REST dysregulation in neuronal pathologies, various strategies have been developed to modulate its action,[15,16] in the attempt to

Received: February 27, 2020
Published: July 17, 2020

restore the physiological levels of expression of REST target genes.

Recently, we developed an optogenetic system able to decrease REST activity in a dynamic and reversible way, by inhibiting the assembly of the REST complex onto the promoter of its targets genes, resulting in an increment of their transcription.[17] An alternative approach to control REST function is to directly modulate its expression levels. To this aim, here we exploit synthetic, *ad hoc* designed RNA binding proteins (RBPs) able to specifically anchor REST mRNA.

RBPs play an essential role in every aspect of RNA biology. The association with their targets determines the cellular localization, lifetime, processing, and translational rate of coding and noncoding RNAs.[18] In recent years, researchers have focused on investigating the structural basis for RNA recognition by RBPs, with the aim of designing tools with tailored sequence specificity to alter gene expression for biomedical applications.[19] Many RBPs are composed of modular repeats of individual domains that, combined in various arrangements, generate versatile proteins capable to bind RNA sequences with high affinity and specificity.[20] Among the various families of RBPs, Pumilio and FBF (PUF) proteins are of particular interest in the field of synthetic biology.[19]

PUF proteins are named after Pumilio in *Drosophila melanogaster* (DmPUM) and fem-3mRNA-binding factor (FBF) in *Caenorhabditis elegans*; however, PUF homology proteins have been found in all eukaryotic organisms, where they regulate diverse processes including stem cell maintenance, organelle biogenesis, oogenesis, neuron function, and memory formation.[21] PUF proteins are versatile regulators that employ multiple mechanisms to regulate their mRNA targets. For example, DmPUM inhibits mRNA translation by competing with the translational initiator factor eIF4E;[22] similarly, yeast PUF5 acts as a post-transcriptional repressor by binding a specific recognition sequence in the 3′UTR of the target mRNA, shortening the poly(A) tail and consequently influencing both RNA stability and translation.[23] Conversely, PUM triggers transcriptional activation of *Xenopus laevis* cyclin B mRNA by stabilizing the cofactor CPEB on the transcript.[24] PUF proteins can also play a role in mRNA localization, as shown for yeast PUF3 and PUF6.[25]

Structural studies have revealed that PUF domains are typically composed of 8 imperfectly repeated 36 amino acid motifs (PUF repeats), flanked by conserved sequences, which form a sequence-specific, single-stranded (ss) RNA-binding domain.[26] Each PUF repeat interacts with a single RNA base and is composed by three α-helices, where five residues in the second helix are fundamental for the binding. This motif displays a conserved pattern of amino acids at positions 12-13-X-X-16 in each repeat of 36 amino acids. Residues at position 12 and 16 are called "edge-recognizing" for their ability to form hydrogen bonds with the nucleobase of the target RNA; the residue at position 13 is referred to as "stacking", because it establishes stacking interactions with the nucleobase and other residues in the forward and backward repetitions;[27] finally, X's represent any hydrophobic amino acid. While the importance of the edge-recognizing residues in the PUF/RNA recognition mechanism is broadly accepted, the specific role of the stacking residues is less precisely defined; it may be different among different repeats, and should be determined on a case-by-case basis.[28,29]

The elucidation of the above-described RNA recognition code of PUF proteins allowed the engineering of synthetic RBPs for targeting specific mRNAs, thus paving the way to direct manipulation of the transcriptome. Such constructs have been exploited in the regulation of a broad range of biological processes to efficiently control mRNA translation from synthetic as well as endogenous target constructs. However, their application to modulate the expression of endogenous mRNAs presents additional challenges such as minimization of off-targeting and maximization of binding affinity to achieve appreciable biological outcomes.[30−34] One possible way to address the first issue is to engineer extended PUF domains capable of binding RNA tracts longer than the canonical 8-nt, which was achieved by following different strategies.[34,35] Among these, in the work of Filipovska *et al.*,[36] an 8-repeat PUF domain was inserted between repeats 5 and 6 of a copy of the same protein sequence, obtaining a 16-repeat PUF protein (PUF16) able to efficiently bind a 16-nt RNA sequence.

In this work, a combined experimental and computational approach was used to design and test 8- and 16-repeat PUF proteins specific to REST mRNA. Computationally, we employed Molecular Dynamics (MD) simulations to study the conformational properties of PUF8 and a modeled PUF16 in solution, with and without bound RNA, and to explore the atomic features of the PUF-RNA recognition code in native and mutated systems. Experimentally, we engineered a number of PUF constructs designed to recognize REST-specific 8- and 16-ribonucleotide sequences. Electrophoretic mobility shift assay (EMSA) and cross-linking immunoprecipitation (CLIP) revealed an important role of stacking residues in determining binding specificity. Finally, we demonstrated that the binding of PUF proteins *per se* does not affect the translational regulation of endogenous REST mRNA. Thus, the PUF-based constructs presented in this work are configured as excellent RNA-binding modules suitable to be assembled in multi-protein scaffolds to influence REST mRNA stability and protein translation.

RESULTS

Design of Synthetic PUF Proteins Specific for REST mRNA. To design synthetic PUF proteins specific for REST mRNA, we followed a two-step process, whereby we first identified suitable sequences on REST mRNA and subsequently engineered the corresponding REST-specific PUF proteins.

Identification of REST-Specific RNA Sequences. We started from the well-studied PUMILIO-Homology Domain from Human Pumilio1 (HsPUM-HD), which recognizes a *consensus* Nanos Response Element (NRE) sequence 5′-UGUA**X**AUA-3′, with **X** indicating U or C bases.[37,38] The crystal structure of HsPUM-HD bound to the *D. melanogaster* hunchback (hb) NRE sequence (5′-UGUAUAUA-3′) identifies a prototypical model for modular PUF:RNA recognition, where each base is paired to a single protein repeat, for a total of eight blocks.[39] Binding occurs in an antiparallel fashion, resulting in the protein N- and C-terms paired to 3′ and 5′ RNA, respectively. Although structural studies showed minor differences in the protein:RNA binding modes when varying the fifth base of the consensus sequence, binding affinities were not affected.[39] In this work, the wild type, 8-repeat HsPUM-HD PUF protein is referred to as PUF8wt. When designing the REST specific PUF domain, we first aimed at minimizing perturbations to the original protein structure. Thus, we adopted a conservative

https://dx.doi.org/10.1021/acssynbio.0c00119
ACS Synth. Biol. 2020, 9, 2039−2054

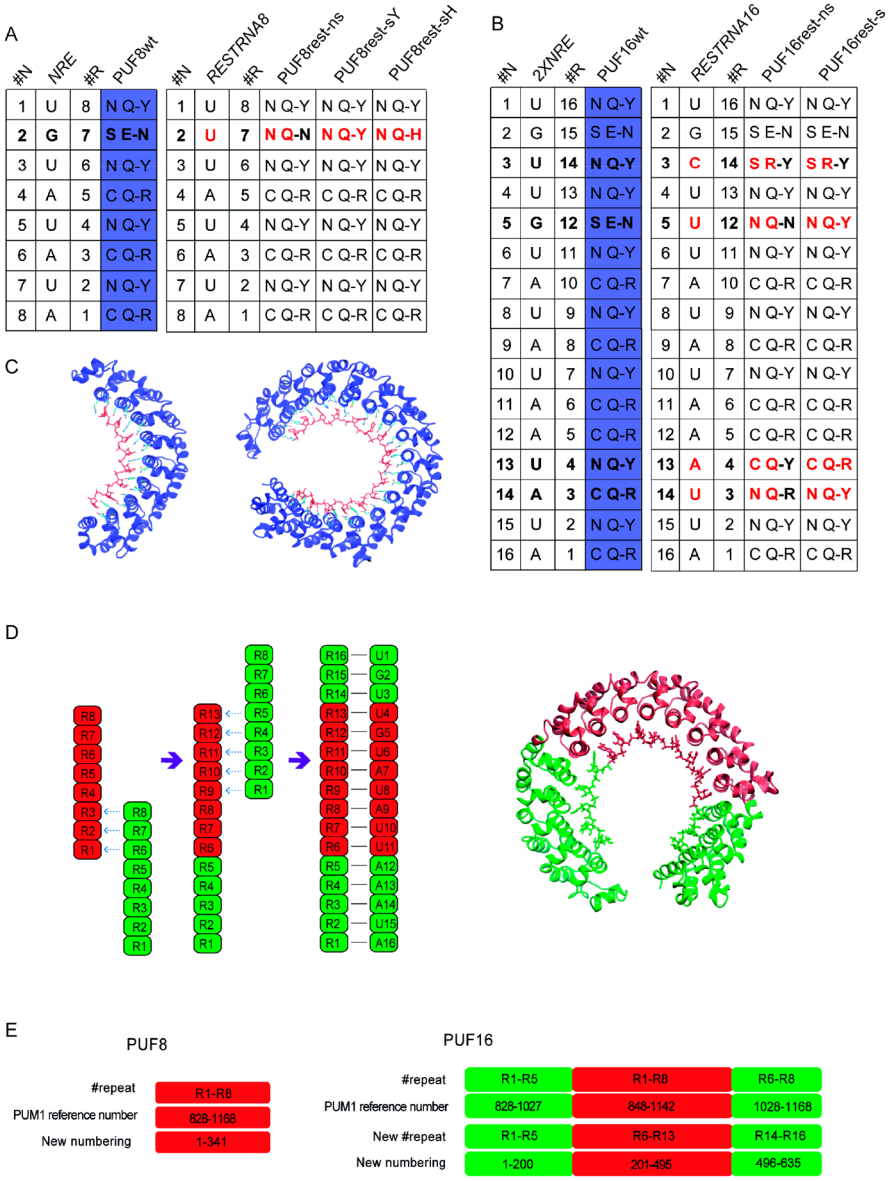


Figure 1. Sequences and structural depiction of the PUF-RNA systems used in this study. The tables report, for each repetition (#R): the starting RNA sequence with nucleotide numbering (#N), the corresponding PUFwt amino acid sequence, the mutated REST-specific RNA sequence and the mutated REST-specific protein sequences. Note that protein N- and C-termini are paired to 3′ and 5′ RNA, respectively. (A) Eight-repeat PUF systems. The *RESTRNA8* sequence differs from *NRE* by one ribonucleotide (uracil, highlighted in red) at R7. The mutations performed to obtain the three REST-specific PUF8 proteins are represented in red (see main text for details). (B) Sixteen-repeat PUF systems. The *RESTRNA16* sequence differs from 2X*NRE* in four ribonucleotides corresponding to R3, R4, R12, and R14 (in red). The mutations performed to obtain the REST-specific PUF16 proteins (PUF16rest-ns and PUF16rest-s) are highlighted in red. (C) Left: PUF8wt crystal structure bound to the *NRE* sequence (PDB code: 3Q0P). The protein backbone is represented in blue and the ribonucleotide sequence in red; protein residues involved in the recognition code are highlighted in cyan. Right: Our 3D structural model of the extended PUF16wt, built from 2 PUF8wt domains, in complex with *2XNRE*. (D) Schematic representation of the strategy used to model the 3D structure of PUF16wt. Two identical 8-repeat PUF domains, colored in red and green, were fused together to produce the extended model, represented on the right. Dashed blue arrows represent the alignment operation used to preserve the relative orientation of fused repeats. The same strategy was used to model the *2XNRE* sequence. (E) Definition of the sequence numbering used in this work. To simplify the counting, we renamed the amino acid sequences of our constructs starting from one. The PUM1 Uniprot entry was used as the reference notation for both PUF8 and PUF16.

approach and searched for a possible PUF target sequence in REST mRNA 3′UTR that minimized the number of mutations required to achieve specific binding and was not overlapping with, or in close proximity to, miRNA sites (Table S1). In this way, we identified an 8-nt sequence that diverges from the original PUF target by only one base, *i.e.*, 5′-UUUAUAUA-3′, which we call here *RESTRNA8* (Figure S1). We also created an extended PUF protein composed of 16 RNA-binding repeats, referred to as PUF16wt. This PUF is formed by two HsPUM-HD, fused according to the strategy of Filipovska *et al.*,[36] and should bind a double *NRE* sequence 5′-UGUUG-UAUAUAAUAUA-3′, which we call *2XNRE*. The REST-specific PUF16 target was built on a sequence spanning the 8 nts of the *RESTRNA8* plus three additional nt at the 5′ and five

https://dx.doi.org/10.1021/acssynbio.0c00119
ACS Synth. Biol. 2020, 9, 2039−2054

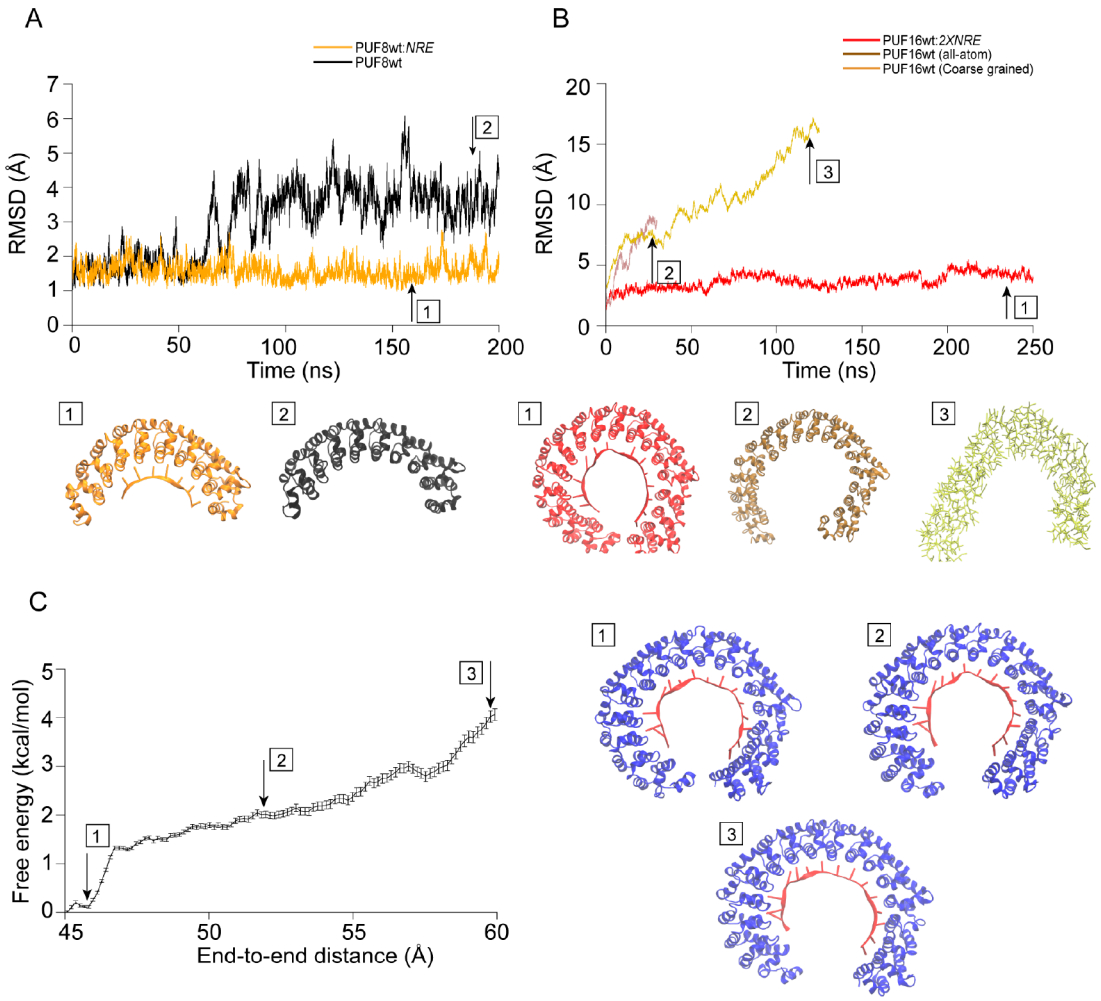


Figure 2. RNA plays a fundamental role in maintaining PUF structure. (A) RMSD profiles of protein backbone atom from MD simulations ofPUF8wt:*NRE* (orange) and unbound PUF8wt (black). Snapshots extracted from the simulated trajectories at selected time frames are indicated by numbers 1−2. (B) RMSD profiles of protein backbone atom from MD simulations of PUF16wt:*2XNRE* (red) and unbound PUF16wt, all-atom (brown) and CG (ochre). Snapshots extracted from the simulated trajectories at selected time frames are indicated by numbers 1−3. (C) Free energy profile associated with progressive increases in the distance between the terminals of the extended PUF domain, *i.e.*, the COMs of the first three (R1−R3) and the last three (R14−R16) protein repeats. Snapshots extracted from the simulated trajectories at selected distance values are indicated by numbers 1−3.

at the 3′, thus obtaining a 16 nt sequence (5′-UGCUU-UAUAUAAAUUA-3′), which we called *RESTRNA16*, and that differs from *2XNRE* at four positions. Additionally, we tested the possibility of engineering PUF proteins for any RNA sequence, irrespective of the similarity to the original *NRE* site, thus releasing the minimal perturbation requirement. To this purpose, we identified an *NRE*-independent 16 nt sequence in REST 3′UTR (5′-ATTGGCTTAGTAAATT-3′) characterized by a minimum overlap with other mRNA sequences of the mouse transcriptome. We refer to this structure as *RESTRNA16−2.0* (Figures S1 and S2).

Engineering REST-Specific PUF Protein Sequences. We modified PUF8wt and PUF16wt through site-directed mutagenesis to obtain proteins able to bind the selected REST sequences. We indicate these proteins as PUF8rest and PUF16rest. In order to elucidate the importance of the stacking residues in the PUF/RNA recognition mechanism, we engineered two variants of synthetic PUF proteins: (i) mutated only in the edge-recognizing residues (position 12 and 16), referred to as "ns" (not mutated in stacking residues); and (ii) mutated in all edge-recognizing and stacking amino acid residues (position 12, 13, and 16), referred to as "s" (mutated in stacking residues). More specifically, the PUF8rest-ns protein was obtained by mutating the edge recognizing residues of repeat R7 in PUF8wt. As for the stacking mutated constructs, we tested the first and second most frequent residues in R7, *i.e.*, Tyr and His, obtaining the proteins PUF8rest-sY and PUF8rest-sH (Figure 1A). Following the same approach, we mutated the residues in R3, R4, R12 and R14 of PUF16wt, obtaining PUF16rest-ns and PUF16rest-s (here, only the most frequent residue in each position was used) (Figure 1B). We also engineered a PUF16rest-2.0 specific for binding the *RESTRNA16*-2.0 sequence. In this case, we changed the edge-recognizing and stacking residues in all repeats binding the RNA nts different from the original *2XNRE* sequence (Figure S2A).

In summary, we engineered a total of six REST-specific PUF proteins: three to bind *RESTRNA8*, (PUF8rest-ns, PUF8rest-sY, PUF8rest-sH), two to bind *RESTRNA16* (PUF16rest-ns and PUF16rest-s), and one for *RESTRNA16−2.0* (PUF16rest-2.0). PUF8wt and PUF16wt with their cognate sequences were used as controls throughout. In this work, protein-RNA

https://dx.doi.org/10.1021/acssynbio.0c00119
ACS Synth. Biol. 2020, 9, 2039−2054

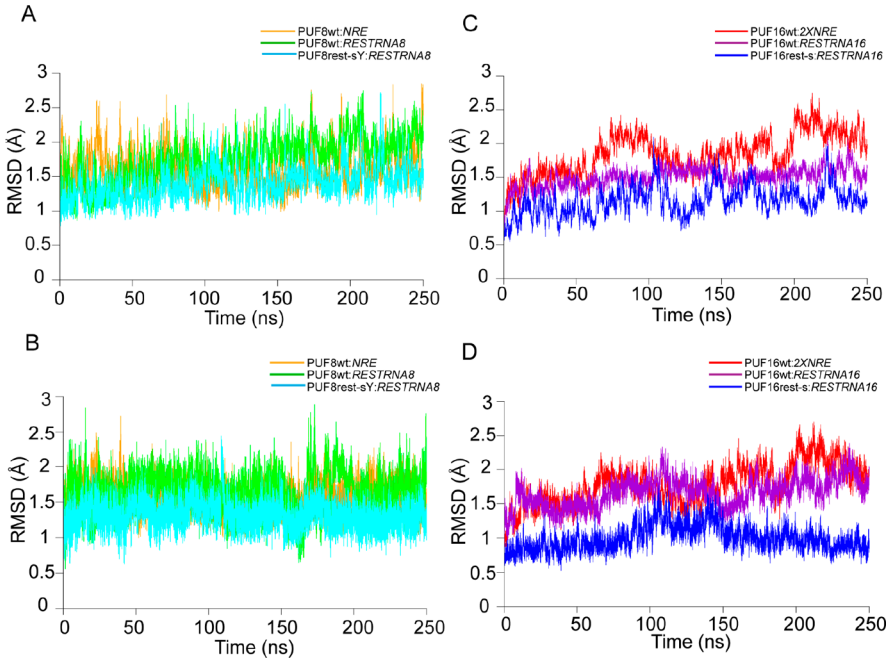


Figure 3. RNA-bound 8- and 16-repeat PUF proteins show stable protein and RNA main-chain conformations. RMSD profiles of protein (A) and RNA (B) backbone atoms from MD simulations of the three 8-repeat systems (PUF8wt:*NRE*, PUF8wt:*RESTRNA8*, PUF8rest-sY:*RESTRNA8*). RMSD profiles of protein (C) and RNA (D) backbone atoms from MD simulations of the three 16-repeat systems (PUF16wt:*2XNRE*, PUF16wt:*RESTRNA16*, PUF16rest-s:*RESTRNA16*).

systems are indicated as PROTEIN:*RNA*. All the final protein and RNA sequences are reported in Figure 1. The position of the selected sequences within mouse REST mRNA 3′UTR (NCBI #NM_011263.2) is reported in Figure S1.

Structural Modeling. To investigate the structural details of PUF/RNA systems, we resorted to computational structural modeling and MD simulations. While for PUF8wt we could rely on a crystallographic structure (PDB code 3Q0P; Figure 1C, left), no experimentally derived structural information is available for the 16-repeat PUF. Hence, we designed a structural model by exploiting the modular structure of the PUF domain (Figure 1D,E, more details are provided in the Materials and Methods section). The resulting model extends the curved structure of the PUF8, resulting in a horseshoe-like conformation (Figure 1C, right).

RNA Plays a Fundamental Role in Preserving PUF Curved Structures. As a first step, we used MD simulations to explore the structural features of the protein-RNA complexes formed by native PUF domains and their cognate RNA sequences. We simulated PUF8wt:*NRE* and PUF16wt:*2XNRE* in water with and without the respective RNA molecules. We employed all-atom MD simulations for bound PUF8wt:*NRE*, unbound PUF8wt and bound PUF16wt:*2XNRE* and both all-atom and CG simulations for isolated PUF16wt. Considering the size of the solvated PUF16 systems (exceeding 200 000 atoms), we used a CG model to reach longer time-scales, while consuming less computational time.

Figure 2A shows protein RMSD profiles for the PUF8wt:*NRE* complex (orange line) and for the isolated PUF8wt (black line). When bound to RNA, the protein maintained the typical arc-like shape throughout the simulation (RMSD values plateau around 1.5 Å), while in the unbound state we observed a deformation of the structure toward a less curved conformation, causing RMSD values to reach 4.5−5 Å. Two snapshots extracted at times indicated along the different trajectories are also shown. Figure 2B shows simulation results for the PUF16wt:*2XNRE* system (red line) and for the isolated PUF16wt, CG and all-atom model (ochre and brown line, respectively). Similar to the 8-repeat system, the RNA-bound protein maintained a stable curved conformation corresponding to a plateau in RMSD values, while removal of RNA induced a wide opening of the protein observed in both all-atom and CG simulations with rapidly increasing RMSD. Snapshots extracted along the different trajectories are also shown.

To further investigate the stability and reliability of our structural model of RNA-bound PUF16wt:*2XNRE*, we used all-atom US simulations to calculate the free energy associated with widening the protein-RNA structure. As starting conformation, we took an equilibrated structure from the RNA-bound PUF16wt:*2XNRE* trajectory. The distance between the COMs of repeats R1−R3 and R14−R16 was used as CV. The free energy profile was reconstructed using the WHAM method. Results are reported in Figure 2C. Starting at a distance of 45 Å, corresponding to the equilibrated horseshoe-like structure, the free energy increases rapidly, revealing that the less curved conformations are not favored. Snapshots of structures corresponding to various distances between terminals are reported on the right. From these results, we conclude that the conformation corresponding to a distance of 45 Å is the most probable.

To summarize, the results of the above-described simulations allow us to conclude that binding to the corresponding RNA sequence has a fundamental role in maintaining the curved structure of both PUF8wt:*NRE* and PUF16wt:*2XNRE* systems.

PUF8rest-sY and PUF16rest-s Preferentially Interact with Their Specific Target Sequences on REST mRNA. In order to investigate how the protein-RNA recognition code

https://dx.doi.org/10.1021/acssynbio.0c00119
ACS Synth. Biol. 2020, 9, 2039−2054

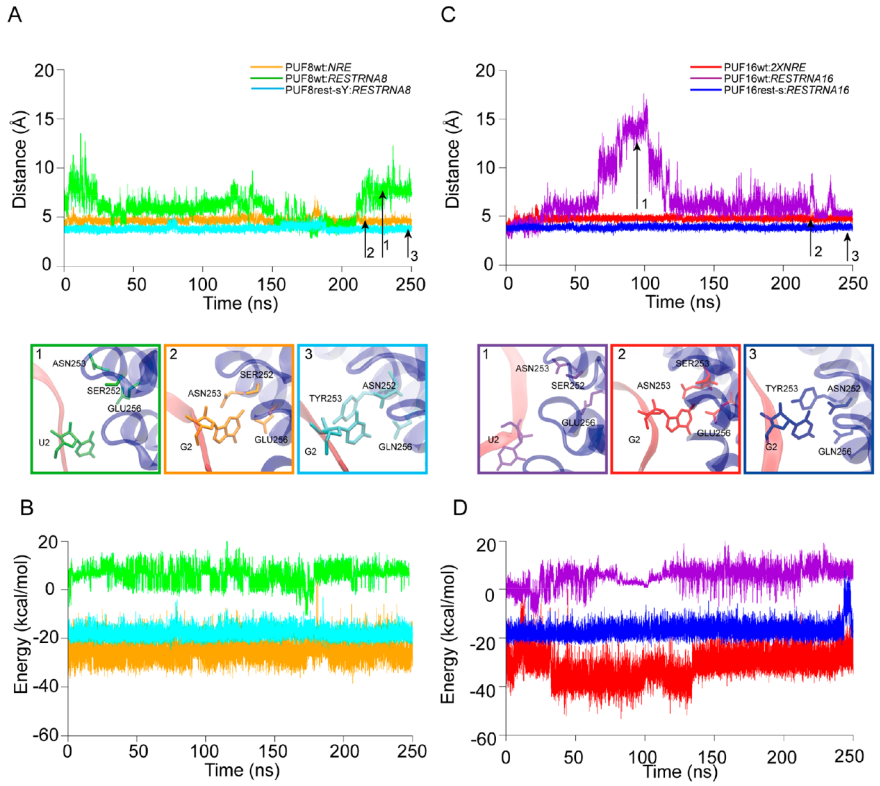


Figure 4. PUF8rest-sY and PUF16rest-s preferentially interact with their specific target sequences on REST mRNA. (A,B) Time series of distances (A) and nonbonding interacting energies (B; electrostatic and van der Waals) between protein and RNA atoms at the R7−N2 interface from MD simulations of the 8-repeat systems. Panels 1−3 represent snapshots extracted from the simulated trajectories at the time frames indicated by the corresponding numbers. (C,D) Time series of distances (C) and nonbonding interacting energies (D) between protein and RNA atoms at the R12−N5 interface from MD simulations of the 16-repeat systems. Panels 1−3 represent snapshots extracted from the simulated trajectories.

modulates protein-nt interactions, we then simulated the systems representing different combinations of PUF proteins and RNA sequences. We analyzed the matching partners PUF8wt:*NRE*, PUF8rest-sY:*RESTRNA8*, PUF16wt:*2XNRE* and PUF16rest-s:*RESTRNA16*. In addition, we created the two mismatched systems PUF8wt:*RESTRNA8* and PUF16wt:*RESTRNA16*. Figure 3 shows RMSD values for protein backbone (Figure 3A,C) and RNA (Figure 3B,D) atoms in the simulated systems. All proteins and RNA molecules show stable backbone conformations over the observed time-scales, as revealed by RMSD plateaus. This means that the perturbation induced in the mismatched systems by mutating only few residues does not impact the global structure of the complex over the limited time-scale of the simulations. Interestingly, however, the REST-specific constructs PUF8rest-sY:*RESTRNA8* and PUF16rest-s:*RESTRNA16* (Figure 3A,B, cyan line, and Figure 3C,D, blue line) showed smaller deviations from the initial structure, indicating more stably interacting protein-RNA complexes.

Since no major structural changes were noticed in all systems, to clarify the role of specific residues in determining their stability, we studied the structural details of the PUF-RNA interaction surfaces at the atomic level. We thus analyzed distances and nonbonding interaction energies (*i.e.*, electrostatic and van der Waals) between COMs of groups of atoms belonging to all interacting protein-RNA repeats. For each analyzed repeat, the first COM was defined using the side chains of the edge recognizing and stacking residues and the second using the nucleobase. Results for the R7−N2 interface in the 8-repeat constructs and the R12−N5 interface for the 16-repeat constructs are reported in Figure 4A,B and Figure 4C,D, respectively. We recall that, at these repeats, the mutations from the native to the REST-specific sequence were Guanine to Uracil for the RNA and SNxxE to NYxxQ for the protein chain. In addition, the mismatched systems PUF8wt:*RESTRNA8* and PUF16wt:*RESTRNA16* present a suboptimal interface according to the consensus recognition code, with a SNxxE protein sequence facing the Uracil ribonucleotide. Plots corresponding to other sites in the 8- and 16-repeat systems are reported in Figure S3.

Among the 8-repeat systems, the mismatched PUF8wt:*RESTRNA8* showed the highest fluctuations, both in distances and nonbonding energy values (Figure 4A,B, green lines). The distance profile increased at the beginning and the end of the simulated trajectory, reaching over 10 Å. At these time intervals, we observed a base-flipping event, indicating a substantial loss of interaction between the nucleobase and the facing protein residues due to the destabilizing effect of the mutation in the protein-RNA surface at that repeat (Figure 4A, green, snapshot 1). Conversely, the matching systems PUF8wt:*NRE* and PUF8rest-sY:*RESTRNA8* showed stationary distance profiles around 5 Å (Figure 4A, orange and cyan lines, respectively), maintaining "crystal-like" contacts in the protein-RNA interface with atomic positions close to those experimentally determined (Figure 4A, snapshots 2−3). This observation was further confirmed by the analysis of the nonbonding energy terms between the protein and RNA groups. Indeed, the mismatched PUF8wt:*RESTRNA8* system showed mostly positive (*i.e.*, unfavorable) interaction energy values (Figure 4B, green line), while the matching

https://dx.doi.org/10.1021/acssynbio.0c00119
ACS Synth. Biol. 2020, 9, 2039−2054

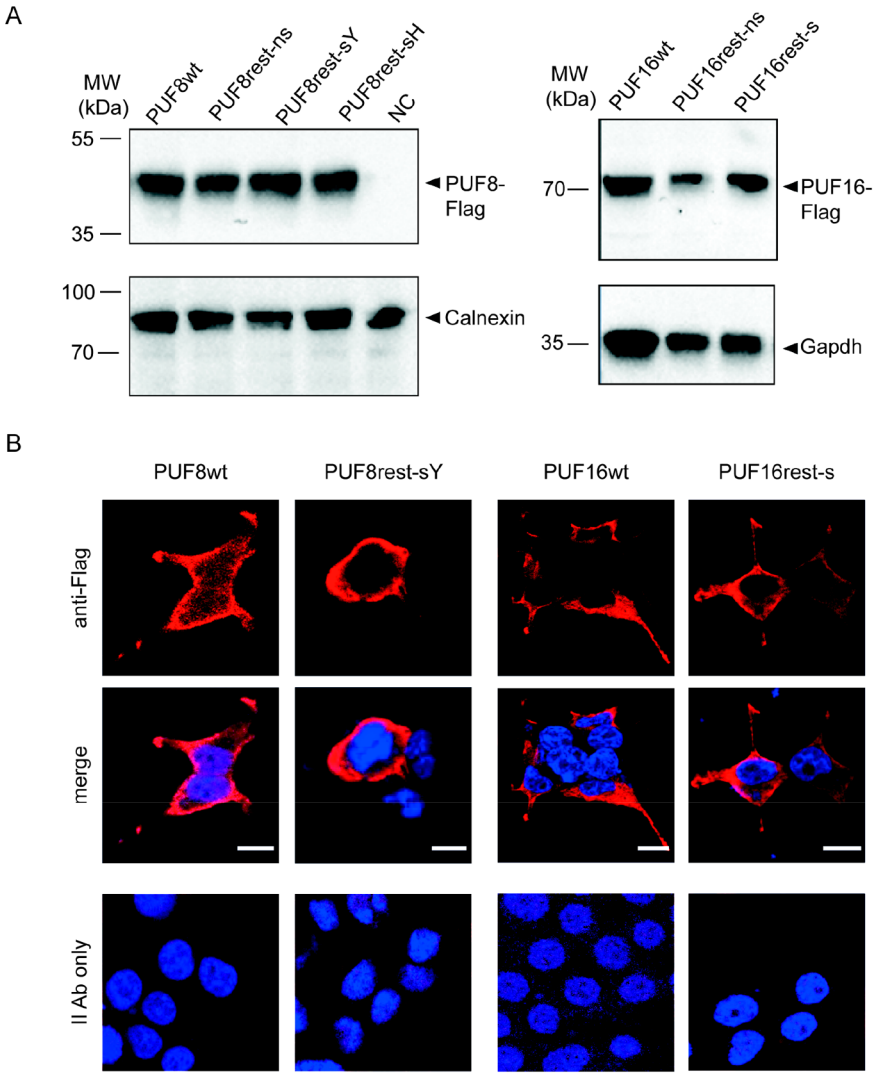


Figure 5. REST-specific PUF proteins are expressed and localized to the cytoplasm in HEK293T cells. (A) HEK293T cells were transfected with the indicated PUF constructs. NC: negative control (cells transfected with the empty Flag vector). Protein expression was analyzed by Western blotting using anti-Flag antibodies (expected MWs: 8-repeat PUF constructs, 47 kDa; 16-repeat PUF constructs, 75 kDa) and antibodies for the housekeeping proteins calnexin and Gapdh. (B) HEK293T cells transfected with the indicated constructs were processed for indirect immunofluorescence using anti-Flag antibodies (red) to detect PUF constructs and Hoechst (blue) to visualize cell nuclei. The overlay images (merge) reveal the cytosolic localization of the constructs. The specificity of the Flag staining was confirmed by analysis of cells processed with fluorescent secondary antibodies, but omitting primary antibodies (bottom panel). For all constructs, the transfection efficiency was ∼50%, with 100% of transfected cells showing cytoplasmic staining. Scale bars: 10 μm.

PUF8wt:*NRE* and PUF8rest-sY:*RESTRNA8* systems remained stable at negative energies (Figure 4B, orange and cyan lines, respectively).

Similar results were obtained for the 16-repeat systems (Figure 4C,D). The mismatched system PUF16wt:*RESTRNA16* showed the highest fluctuations in COM distances (Figure 4C, violet line), with a base-flipping event observed at about 90 ns (Figure 4C, snapshot 1). Instead, both PUF16wt:*2XNRE* and PUF16rest-s:*RESTRNA16* profiles (Figure 4C, red and blue line, respectively) were stationary around a plateau at 5 Å (the same value observed for the matching 8-repeat systems). Snapshots from these trajectories showed typical PUF-RNA interface conformations (Figure 4C, snapshots 2 and 3). Moreover, the nonbonding energy analysis revealed that the mismatched PUF16wt:*RESTRNA16* system had the least favorable energy values (Figure 4D, violet line), while the matching PUF16wt:2X*NRE* and PUF16rest-sY:*RESTRNA16* systems remained stable at negative energies (Figure 4D, red and blue lines, respectively).

Summarizing, the results described in this section provide structural evidence that the consensus PUF-RNA recognition code results in preferential interactions of the proteins with their specific target RNA sequences.

PUF8rest-sY and PUF16rest-s Bind Their Cognate Sequences on REST mRNA *in Vitro*. As a first step in the experimental validation of the engineered PUF constructs, we checked whether they are correctly expressed in mammalian cells. To this aim, we transfected HEK293T cells and analyzed the expression and intracellular localization of the various proteins by Western blotting and immunofluorescence. All the recombinant proteins were abundantly expressed (Figure 5A) and localized to the cytosol of transfected cells (Figure 5B;

https://dx.doi.org/10.1021/acssynbio.0c00119
ACS Synth. Biol. 2020, 9, 2039−2054

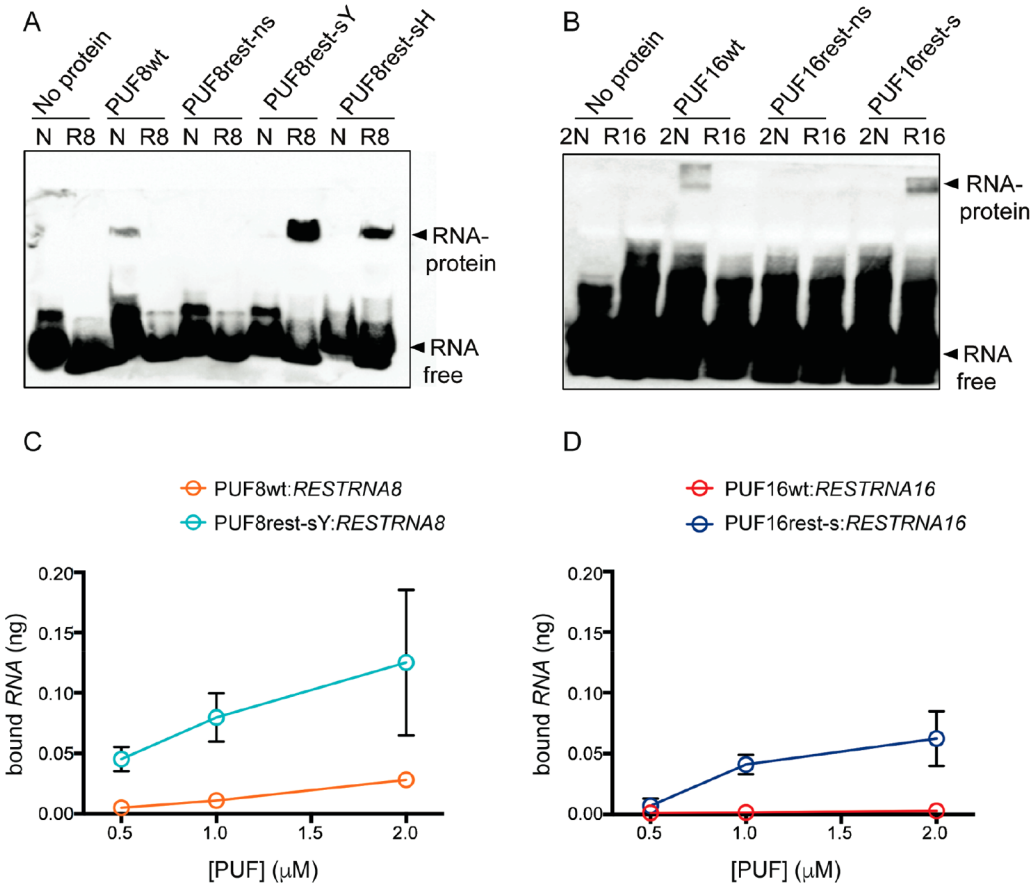


Figure 6. REST-specific PUF proteins specifically bind their cognate REST target sequences. (A,B) PUF constructs' binding specificity was evaluated by EMSA analysis. The indicated 8-repeat PUF proteins were incubated with biotinylated *NRE* (N) and *RESTRNA*8 (R8) ribonucleotide sequences (A). Sixteen-repeat PUF proteins were incubated with biotinylated *2XNRE* (2N) and *RESTRNA16* (R16) ribonucleotide sequences (B). Protein-RNA complexes and unbound RNA are indicated. (C,D) Quantification of the amount of REST RNA (ng) bound to PUF8rests-Y (C) and PUF16rest-s (D) at [PUF] = 0.5, 1.0, and 2.0 μM, compared to their corresponding wild type systems. n = 6 independent experiments, data are shown as means ± SEM.

Figure S2B,C). We subsequently purified the proteins starting from mammalian cells, since mutations may affect the protein solubility and the efficiency of production in bacteria.[40] Proteins isolated from transfected HEK293T cells were analyzed by SDS-PAGE and showed a satisfactory purification rate (Figure S4).

To verify whether the designed PUF constructs were able to specifically bind their cognate RNA sequences, we performed EMSA analysis by incubating purified PUF proteins (1 μM) with biotinylated RNA targets. For the 8-repeat constructs (PUF8rest-ns, PUF8rest-sY, and PUF8rest-sH), the target RNA sequences were *NRE* and *RESTRNA8*. For the 16-repeat constructs, the target RNA sequences were *2XNRE* and *RESTRNA16* for PUF16wt, PUF16rest-ns and PUF16rest-s. For the PUF16rest-2.0 construct, the target sequence was *RESTRNA16−2.0*. We first investigated the binding of each PUF construct to its cognate and noncognate sequences (Figure 6). Notably, we observed the formation of a specific RNA−protein complex between PUF8rest-sY and PUF8rest-sH with *RESTRNA8*, while we did not detect any binding to *NRE*. *Vice versa*, we observed the formation of a specific complex of PUF8wt with *NRE*, but not with *RESTRNA8* (Figure 6A). We did not observe any RNA−protein complex for PUF8rest-ns with any of the target sequences, even when the concentration of PUF8rest-ns used in the assay was raised to 10 μM (Figure S5A). Similar results were obtained when EMSA was performed with the 16-repeat constructs. In this case, we observed the formation of a specific complex of PUF16rest-s with *RESTRNA16*, and of PUF16wt with*2XNRE*, but not with their noncognate sequences (Figure 6B). No RNA−protein complexes were detected for PUF16rest-ns (Figure 6B) and PUF16rest-2.0 (Figure S2D) with any of the sequences, even when the PUF concentration was increased to 10 μM (FigureS5B). The quantification of the binding of PUF8rest-sY to *RESTNRE8* and of PUF16rest-s to *REST-NRE16* is reported in Figure 6C,D: the results show appreciable binding specificity of REST-specific PUF proteins to RESTRNA target sequences, compared to wild type PUFs, as well as a clear dose-related increase in bound RNA at increasing PUF concentrations. Altogether, these results show that proper stacking residues are indeed crucial to confer binding specificity to PUF-based constructs. On the basis of these data, in all subsequent experiments we used the proteins with optimized stacking residues.

PUF8rest-sY and PUF16rest-s Bind Endogenous REST mRNA. Once the binding specificity of the PUF constructs was demonstrated *in vitro*, we investigated whether they are able to bind endogenous REST mRNA using CLIP. N2a cells were transfected with PUF8rest-sY, PUF16rest-s, PUF8wt, and PUF16wt; after 48 h cell lysates were immunoprecipitated with

https://dx.doi.org/10.1021/acssynbio.0c00119
ACS Synth. Biol. 2020, 9, 2039−2054

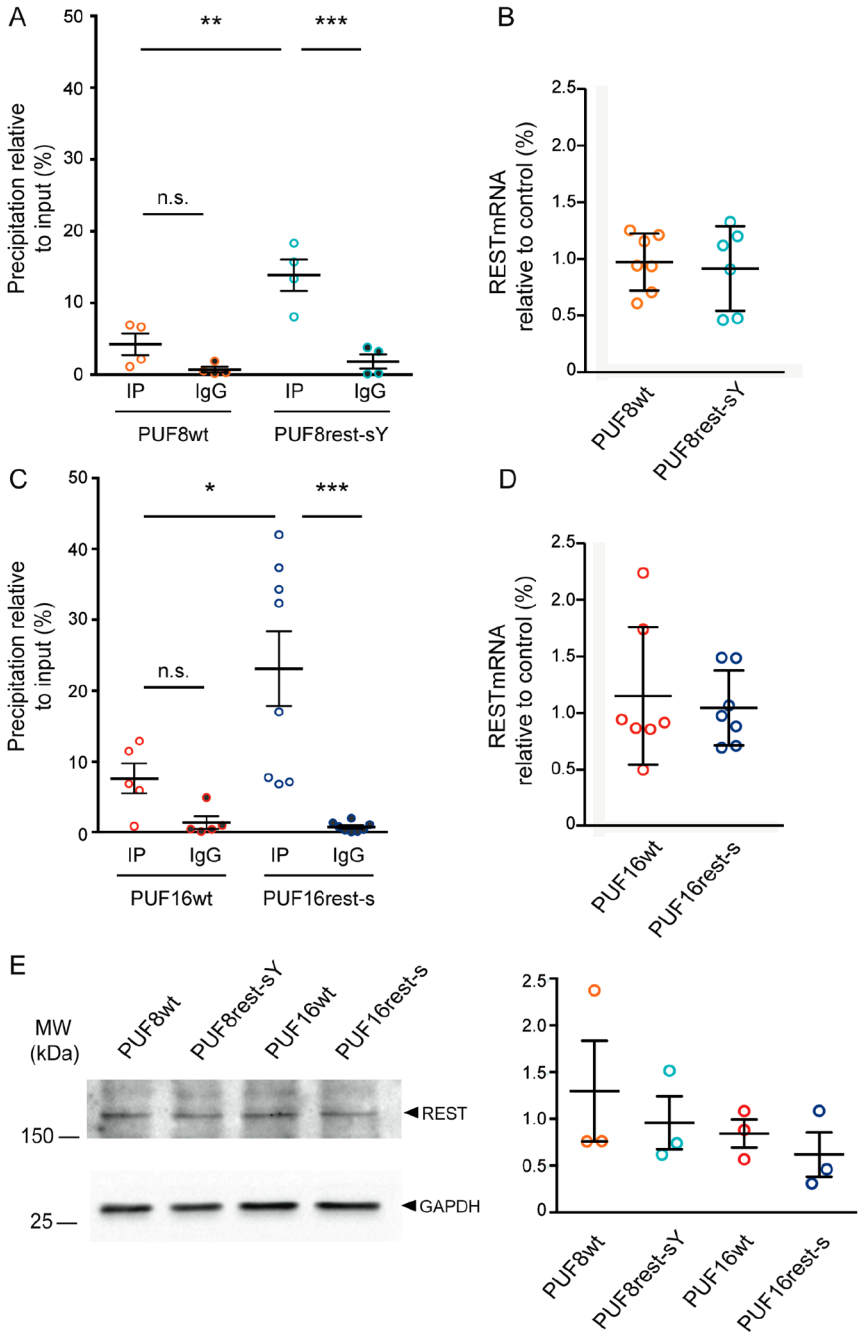


Figure 7. PUF8rest-sY and PUF16rest-s proteins bind endogenous REST mRNA without altering its expression. (A,C) CLIP was performed with anti-Flag and IgG antibodies bound to agarose A beads in N2a cell lysates transfected with the indicated constructs. REST mRNA values were normalized to the input value and plotted as percent specific precipitation. One-way ANOVA/Tukey's tests, $n = 4-8$ independent experiments. n.s., nonsignificant, $p > 0.05$; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$. (B,D) N2a cells were transfected with the indicated constructs. After 48 h, REST mRNA levels were quantified *via* qRT-PCR analysis and expressed as percent of control samples transfected with the empty Flag vector. Gapdh was used as control housekeeping gene. No significant differences were observed, $p > 0.05$, one sample Student's *t*-test, $n = 7$ independent experiments. (E) N2a cells were transfected with the indicated constructs. After 48 h, REST protein levels were quantified by Western blot analysis using REST-specific antibodies. Gapdh was used as loading control. A representative immunoblot is shown on the left, while quantification of 3 independent experiments is reported on the right. No significant differences were observed, $p > 0.05$, one sample Student's *t*-test, $n = 3$ independent experiments. Values from single experiments are plotted, with bars representing means ± SEM.

anti-Flag antibodies. Immunoprecipitated RNA was subsequently analyzed by qRT-PCR with primers specific for REST 3′UTR and Gapdh as control. Both REST-specific proteins (PUF8rest-sY and PUF16rest-s) bound endogenous REST mRNA more strongly than their wild type counterparts (PUF8wt, PUF16wt). As expected, the binding of wild type proteins was not different from that of control samples, immunoisolated with nonspecific IgG antibodies. REST-specific proteins bound REST mRNA with comparable affinity, although PUF16rest-s showed stronger binding (Figure 7A,C). No detectable binding of any of the proteins to Gapdh mRNA was observed, confirming the specificity of our RNA-binding

https://dx.doi.org/10.1021/acssynbio.0c00119
ACS Synth. Biol. 2020, 9, 2039−2054

constructs (Figure S6). These results demonstrate that PUF8rest-sY and PUF16rest-s are capable to specifically bind REST endogenous mRNA.

PUF protein binding may affect the stability and translational activity of their target mRNAs through various mechanisms.[27] To analyze mRNA stability upon PUF binding, we transfected N2a cells with PUF8rest-sY, PUF16rest-s, PUF8wt, and PUF16wt and performed qRT-PCR and WB analysis to measure REST mRNA and protein levels, respectively. We did not detect any significant change in REST mRNA and protein, suggesting that the PUF constructs do not affect REST mRNA and protein stability (Figure 7B,D,E).

We attempted to fuse the translational activator poly(A) polymerase Germ Line Development 2 (GLD2)[41] to PUF8wt and PUF8rest-sY to see whether docking GLD2 would increase REST mRNA stability (Figure S7). All constructs were expressed in transfected HEK293T cells (Figure S7A). We performed luciferase assays on HEK293T cells cotransfected with the various PUF constructs and a luciferase reporter vector containing the *NRE* or *RESTRNA8* target sequence within the luciferase mRNA. Interestingly, the values obtained with PUF8wt-GLD2 were higher than those of PUF8wt in the presence of the cognate *NRE* sequence, indicating that GLD2 may indeed act by stabilizing mRNA and increasing luciferase translation, thus supporting the validity of our approach. Unfortunately, no increase in the Renilla/luciferase ratio was detected in cells cotransfected with PUF8rest-sY-GLD2 and the cognate *RESTRNA8* sequence, compared to cells cotransfected with PUF8rest-sY (Figure S7B). Similar results were obtained by measuring the levels of endogenous REST mRNA in cells transfected with GLD2, PUF8rest-sY or PUF8rest-sY-GLD2: no increase of REST mRNA was observed in the presence of the GLD2 fusion protein (Figure S7C). Thus, in our case, fusion of the cytoplasmic poly(A) polymerase GLD2 to the REST-specific PUF8 did not apparently alter either the stability or the translation efficiency of REST mRNA.

DISCUSSION

In this work, we have used the versatile PUF system to engineer RNA-binding proteins specific for the mRNA of the transcriptional repressor REST. To this purpose, we employed PUF proteins bearing 8 and 16 modular repeats. The relationship between length and specificity has a key role in PUF/RNA binding and hence in the design of synthetic PUF modules. Since longer PUF proteins have more binding sites, they are expected to show greater affinity for their cognate targets. Experimental evidence, however, suggests that the mechanism is system-dependent.[42] In addition, the 1:1 relation between PUF repeats and bound ribonucleotides is not universally respected.[43] For example, the Pumilio repeat-containing protein Nop9 is composed by 10 Pumilio repeats and binds specifically a 11-nt sequence of 18S rRNA by establishing interactions only with 8 bases, leaving the other three unbound and flipped out in the middle of the sequence.[44] Similarly, APUM23, a nucleolar 10-repeat PUF protein of *Arabidopsis*, binds an 11-nt target site, with one repeat (R7) not involved in RNA binding.[45] The number of PUF repeats is also crucial to determine the biological function of the PUF-containing protein, as recently demonstrated.[46] The PUF16 system employed here was adopted from the one rationally designed by Filipovska and colleagues, which was characterized by high binding affinity and selectivity for its 16 nt RNA target.[36] Indeed, among the various possible strategies to engineer a 16-repeat protein, the design from Filipovska, with two 8-repeat PUFs nested one inside the other, was one of the best performing.[34] Despite several works addressing longer PUFs, the structural mechanisms determining specificity seem to be highly specific for each protein. Thus, computational and biochemical studies have to be performed for each studied system. In particular, structural information about the 16-repeat PUF is still lacking. Although our structural model and the synthetic PUF16 were both generated by fusing together modules of the experimentally determined PUF8 protein, we do not currently have a direct measure of how close the modeled structure is to the one assumed by the protein. Interestingly, however, it was proven that two PUF8 domains bind a double NRE sequence in a head-to-tail, almost circular conformation, strikingly similar to our designed PUF16 model.[47]

Our structural modeling and simulation data show that RNA-bound PUF8wt and PUF16wt are structurally stable and maintain a curved conformation. When RNA is removed, both proteins partially loose their curvature, with PUF16wt showing the largest modification. The conformational change we observe in PUF8wt upon RNA removal is larger than the experimentally determined difference between the isolated and the RNA-bound protein (RMSD of 1.2 Å).[37] However, the crystal structure of the protein alone shows weak densities and high B factors (in particular at the C terminal region),[37,48] suggesting conformational heterogeneity. Moreover, crystals of HsPUM-HD bound to NRE1−14 and NRE2−10 differ by a RMSD of 1.1 Å, indicating that the bound protein can undergo conformational fluctuations (termed "breathing" motions in ref 37). More generally, because of experimental conditions such as high salt and molecular concentration, low temperatures, and crystal packing, X-ray crystallography might underestimate the conformational fluctuations that a protein can experience in physiological conditions. At variance, MD simulations performed in solution at the physiological state allow a broader sampling of conformations, often deviating significantly from the crystal structure. Summarizing, our simulations reveal protein structural fluctuations consistent with the experimentally suggested breathing motions and indicate that these are more pronounced when the RNA is not present.

Given the absence of structural studies on the extended PUF domains, it was essential to validate extensively our PUF16 computational model. To this goal, we employed all-atom and CG standard MD simulations and free energy calculations. Results reveal that the RNA strand has a fundamental role in maintaining the horseshoe-like conformation of the complex, and that such conformation is more energetically favorable than those with lower curvatures, obtained by progressively increasing the distance between the end points. The effect of RNA binding on the curved conformation of PUF proteins has been inferred before. Conformational changes are observed as a consequence of RNA binding, although the direction of the transformation seems to be system dependent. A decrease of the overall curvature is found in an engineered nine-repeats PUF after binding of the cognate 9-nt RNA.[42] Conversely, in ref 44 the crystal structure of the PUF-like, 10-repeat Nop9 protein was obtained in RNA-bound and free states, and a more compact protein conformation was observed in the presence of the ribonucleotide sequence. Finally, it is worth noticing that not all PUF domains show RNA-dependent

https://dx.doi.org/10.1021/acssynbio.0c00119
ACS Synth. Biol. 2020, 9, 2039−2054

conformational changes. For instance, it was reported that the yeast PUF4 does not change conformation upon binding its cognate target.[43]

Dissecting the structural details of PUF/RNA interactions is of crucial importance to understand how these proteins work. To this goal, we simulated different combinations of PUF proteins and RNA sequences: the wild-type constructs (PUF8wt:*NRE* and PUF16wt:*2XNRE*), the optimized variants for the REST transcript (PUF8rest-s:*RESTRNA8* and PUF16rest-s:*RESTRNA16*), and the mismatched variants (PUF8wt:*RESTRNA8* and PUF16wt:*RESTRNA16*). By looking at RMSDs over time, we observed for all systems globally stable structures on the limited time-scale of the trajectories (250 ns). These results imply that the REST-targeting structures are as sound as the native, PUF8wt:*NRE* one, and that the perturbation induced in the mismatched systems does not impact the global structure of the complex, at least over the considered time window. To investigate in more detail the effects of mutations on PUF/RNA interactions, we looked at distances and nonbonding energy (*i.e.*, electrostatic and van der Waals) between all protein−nucleotide interacting repeats. The mismatched systems systematically showed higher (*i.e.*, weaker) interaction energies and mutual distances at the mutated or neighboring repeats. In some cases, the weaker interaction resulted in flipping of the nucleobase. We believe that these local structural changes at the PUF/RNA interface are early events of a process leading to unbinding at longer time scales, not accessed in our simulations, but that we verified experimentally. Interestingly, we observed more marked effects in the PUF16 constructs. For example, in PUF16wt:*RESTRNA16*, the R4−A13 interaction is weakened enough to observe base-flipping throughout most of the simulated trajectory. The higher destabilization of the PUF/RNA surface when multiple mutations are introduced clearly suggests a tolerance beyond which the binding could be severely affected. This has been addressed by ref 32, who showed that more than one mismatch is sufficient to drastically reduce the binding between the PUF and the RNA.

In addition, although many works[32,33,49] proved that PUF acquires the desired specificity by changing only the amino acids at positions 12 and 16, our experimental results, in line with the work of Koh *et al.*,[28] demonstrate that PUF proteins mutated in residues at positions 12−16 and 13 are the only mutant proteins that specifically bind REST mRNA. This underlies the importance of changing the stacking residues in the mutation design, in view of their strong contribution to specificity and binding affinity for RNA sequences.

Interestingly, however, for our PUF16rest-2.0 system, where we changed all 9 repeats (including the stacking ones) to respect the recognition code, while maintaining the same PUF scaffold, we were not able to detect any significant binding to REST RNA sequence. It is possible that the mutations may alter the secondary and tertiary structure of the PUF scaffold in a way that impairs RNA binding, and/or the RNA target sequence may have a 3D structure different from the consensus *NRE*, not accessible to the synthesized PUF protein. Thus, while it is important to respect the recognition code when creating new target-specific PUF constructs, experimental validation is always necessary since several factors may affect binding specificity.[35]

For what concerns the cell biology experiments, a first point to make is that all the engineered constructs were successfully expressed in cell lines and efficiently purified therefrom, despite the relatively high number of mutations introduced, which could have impacted on protein expression.[40] The results of the EMSA experiments demonstrated that only the PUF constructs mutated in edge-on and stacking residues are able to bind the selected REST sequence, and moreover, that the introduction of tyrosine as stacking amino acid confers higher binding efficacy than that observed with histidine, in agreement with recent literature[46] and strengthening the notion that stacking residues are crucial for determining binding specificity.[42] Interestingly, both our 8- and 16-repeat REST-specific proteins were characterized by higher binding to their target RNA than that observed for the wild type systems, while in the literature the affinity of mutant PUF proteins greatly vary with respect to their wild type counterparts.[35,36,46] Importantly, we confirmed these results by performing CLIP experiments, showing that the REST-specific PUFs are able to bind the endogenous REST mRNA. The CLIP methodology suffers of some recognized drawbacks, including low specificity and high variability, which are also affecting our experiments. Although CLIP alone would not be sufficient to draw general conclusions about binding kinetics and affinity of PUFs for their cognate RNAs, in association with the results of EMSA experiments provides reliable evidence of the relative specificity of REST-specific PUFs for REST mRNA. Thus, we provide the first evidence that also the 16 repeat PUF protein can be mutated according to the code in order to specifically bind endogenous sequences. Overall, our computational and experimental work proves that synthetic 8- and 16-repeat PUF domains, engineered by mutating and replicating the 8-repeat PUF, can be synthesized and expressed, and that they are able to bind differentially to RNA sequences, depending on whether the protein-nucleotide recognition code is respected or not. Moreover, we provide an original structural model of a PUF16 domain refined in atom detail, which we have assessed *via* unbiased MD and free energy calculations.

Binding of PUF proteins to their target mRNA could in principle degrade it, stabilize it or leave it unaffected.[25,50,51] An important problem to consider when engineering PUF proteins is how RNA structure may affect binding. It is indeed reasonable to assume that more structured regions can hinder target sequences, and when those are not available, high-quality predictions can be obtained.[52] There are however notable examples where PUF binding substantially changes the structure of the RNA,[53] so that direct experimental testing is always necessary. In addition, PUF binding may also interfere with the physiological post-translational mechanisms controlling mRNA stability and protein translation, triggering nonsense-mediated decay (NMD).[54] We tried to avoid altering the physiological regulation by miRNAs by selecting target regions that do not overlap with predicted miRNA sites. As far as NMD mechanisms are concerned, at present no experimental evidence about this type of post-translational control on REST mRNA is available, and our functional studies do not support this hypothesis, as both PUF8rest-sY and PUF16rest-s do not alter REST endogenous mRNA and protein levels, suggesting that they do not have any intrinsic effect on REST mRNA stability. This is relevant, as these constructs could be used as scaffolds to anchor effector proteins to REST mRNA to modulate its stability and/or translation rate, as shown for the polyA polymerase GLD-2[41,55,56] and for the eukaryotic translation initiation factor 4E (eIF4E).[32] In our case, fusion of GLD2 to the REST-specific

https://dx.doi.org/10.1021/acssynbio.0c00119
ACS Synth. Biol. 2020, 9, 2039−2054

PUF8 did not apparently alter the stability nor the translation efficiency of REST mRNA. This evidence suggests that the selected PUF docking site may not be in the ideal position for an efficient action of GLD2. Whether this is due to inefficient poly adenylation by GLD-2 or to compensatory mechanisms is still matter of investigation. Notwithstanding these results, our REST-specific PUF-based constructs represent a versatile and functional platform to specifically target REST mRNA and modulate its endogenous expression. Since alterations of REST levels and activity have been associated with a high number of pathologies,[4] the possibility to manipulate its function paves the way to the design of novel therapeutic strategies addressing REST-associated diseases.

MATERIALS AND METHODS

Materials. All biochemical reagents and drugs were from Sigma-Aldrich (Milano, IT) and Promega (Milano, IT), unless otherwise specified. Tissue culture reagents and media were from Gibco (Life Technologies Corp, Monza, Italy) or Sigma-Aldrich. List of antibodies used: rabbit polyclonal anti-GAPDH (#SAB3500247; Sigma-Aldrich), rabbit monoclonal anti-FLAG (#F7425; Sigma-Aldrich), rabbit polyclonal anticalnexin (#10286; Abcam, Cambridge, UK), rabbit polyclonal anti-β tubulin (#T2200 Sigma-Aldrich). Horseradish peroxidase (HRP)-conjugated secondary antibodies for Western blot analysis were stabilized goat antirabbit IgG (H+L), peroxidase conjugated (#32460; Thermo Scientific).

Structural Models of the PUF Domains. Models of wild type and mutant 8-repeat PUF systems were generated starting from the crystal structure of PUMILIO-homology domain from Human PUMILIO1 (PDB code: 3Q0P) in complex with the Drosophila hunchback (hb) NRE1mRNA.[39] For the 16-repeat systems, since experimental structural data about extended PUF domains were not available, we built the models from scratch, following as reference the cloning strategy described in Filipovska *et al.*[36] In this procedure, a 16-repeat PUF domain is formed by extending an 8-repeat domain with a second identical domain, attaching five (R1−R5) and three (R6−R8) repeats of the second PUF protein at the N- and C-terminal of the first protein, respectively, after removing its flanking regions. The protein sequences were cut based on the Uniprot structure domain section of the 8-repeat PUF (http://www.uniprot.org/uniprot/Q14671#structure). To determine the relative orientation of the attached repeats, we exploited the modular structure of the PUF protein by superimposing the backbone of redundant segments. Specifically, to fuse the R1−R5 segment of the second PUF to the N-terminal of the first PUF we superimposed R6−R8 of the second PUF to R1−R3 of the first PUF. We then deleted the R6−R8 segment of the second PUF and attached the remaining R1−R5 segment to the first PUF. The same strategy was adopted to fuse the R6−R8 segment of the second PUF and also to generate the extended bound RNA molecule (*2XNRE*). The PSFgen plugin of the VMD[57] program was used to generate the full structural models.

Molecular Dynamics Simulations. The starting conformation of the PUF8 system was obtained from the PDB structure 3Q0P, while that of the PUF16 was generated as described in the previous section. All necessary protein and RNA mutations were performed using the VMD program by deleting the original side-chain atoms and letting VMD generate the positions of the new atoms. When overlaps were observed, the structure was optimized by hand based on local alignment with similar crystal structures. All systems were solvated in explicit water using the TIP3P water model[58] with the total charge neutralized by the addition of Na^+ and Cl^- ions at their physiological concentration. The total number of atoms, including water and ions, was about 130 000 for PUF8 systems and 290 000 for PUF16 systems. Periodic Boundary Conditions were used to avoid finite-size effects, and long-range electrostatic interactions were treated with the Particle Mesh Ewald method.[59] For each system, the internal energy was minimized *via* 1000 steps of steepest descent and then 2 ns of equilibration were performed at constant pressure and temperature (NPT ensemble) with pressure 1 atm and temperature 300 K.[60] A time step of 2 fs was used. Finally, production runs were carried out in the constant volume and temperature NVT ensemble, for a total of 250 ns per system, with the exception of the isolated, PUF16wt, simulated for 30 ns. All simulations were performed with the NAMD program[61] and using the Amber force field ff14SB[62] for proteins and ff9bsc0 with glycosidic dihedral χ modification (χ_{OL3}) for RNA molecules.[63,64] To evaluate overall structural stability of each system we monitored the atomic root mean squared displacement (RMSD) over time. A coarse grained (CG) model of the isolated PUF16wt was obtained using the CHARMM-GUI Web server.[65−67] The Martini v2.2 force field[68] was used to model the protein residues and water molecules. The resulting total number of particles was about 58 000. PBC conditions were used, PME were disabled, and the dielectric constant was set to 15.0. The structure was minimized for 10 000 steps and equilibrated for 25 ns in a NPT ensemble with pressure at 1 atm and temperature at 300 K. Finally, production run was carried out in a NVT ensemble with a time step of 5 fs. A 125 ns trajectory was produced. All analyses were carried out with VMD, and distances and energy contributions were calculated using Tcl scripts.

Free Energy Calculations. To calculate the conformational free energy of the extended PUF structure, we employed the Umbrella Sampling (US) method using the NAMD Colvar internal module.[69] As collective variable (CV), we chose the distance $d(\mathbf{r})$ between the centers of mass (COMs) of the first three (R1−R3) and the last three (R14−R16) repeats of the extended PUF domain. We set up a set of 30 different simulations each biased by a harmonic potential $1/2k(d(\mathbf{r}) - d_i)^2$ to restrain the distance at 30 consecutive values d_i from 45 to 60 Å by intervals of 0.5 Å. Each window trajectory was run for 5 ns. The force constant for the harmonic restraint was 5 kcal/mol^2, and the free energy profile was reconstructed using the WHAM algorithm with a tolerance value of 10^{-6}.[70]

PUF Vectors and Cloning Strategy. The plasmids coding PUF8wt and PUF16wt were kindly provided by Dr. Rackham (The University of Western Australia, Perth, Australia). PUF8rest-ns, PUF8rest-s, PUF8rest-sh and PUF16rest have been obtained starting from PCR amplification of PUF8wt and PUF16wt, using Pfu DNA polymerase (#M7745; Promega) with specific mutagenesis primers. PCR conditions were as follows: 95 °C, 5 min; 95 °C, 30 s; 55 °C, 30 s; 72 °C, 17 min; 30 cycles; 72 °C, 5 min. PCR products were digested using the DpnI enzyme (Promega) and transformed into DH5α cells. Positive colonies were verified by DNA sequencing. To avoid recombination events in PUF16rest, the mutations in REP3, 4, and 14 were performed on PUF8wt while the mutation in REP12 on PUF16wt. Mutated PUF16wt and PUF8wt were then digested with the SacI enzyme (Promega) and ligated by using the T4 DNA

https://dx.doi.org/10.1021/acssynbio.0c00119
ACS Synth. Biol. 2020, 9, 2039−2054

Ligase (Promega). PUF16rest-s was produced by Biomatik Corporation (Cambridge, ON, Canada). Mutated constructs were then amplified using the Pfu DNA polymerase (#M7745; Promega) with primers flanked by restriction sites for NotI and *Bam*HI. PCR-amplified sequences were subsequently cloned in NotI /*Bam*HI (Promega) digested CMV_3Xflag vector, kindly provided by Dr. di Bernardo (Telethon Institute of Genetics and Medicine, TIGEM, Naples, Italy). The sequences of the mutagenesis primers, primers for cloning, and annealed oligos are provided in SI Materials and Methods. On the basis of the known RNA recognition code of PUF proteins,[37] we replaced the serine at position 12 and the glutamic acid at position 16 of PUF8wt with an asparagine and a glutamine, respectively (SE to NQ), obtaining PUF8rest-ns. We then replaced also the asparagine at position 13 with a tyrosine (N to Y) obtaining PUF8rest-s. Moreover, according to the expanded recognition code in artificial PUF scaffolds,[71] we also replaced the same asparagine with a histidine (N to H), which is the second most frequent amino acid after tyrosine in repeat R7. In this way, we finally obtained the sequence PUF8rest-sH, shown in Figure 1A. Following the same approach used for PUF8wt, we mutated the edge-recognizing amino acid residues at position 12 and 16 of repeats R3, R4, R12, and R14 of PUF16wt, obtaining PUF16rest-ns. Specifically, in repeat R3 we replaced cysteine with asparagine (C to N), in repeat R4 asparagine with cysteine (N to C), in repeat R12 serine and glutamic acid with asparagine and glutamine (SE to NQ), and in repeat R14 asparagine and glutamine with serine and arginine (NQ to SR). Similar to the 8-repeat PUF constructs, we also modified the stacking residues, creating the PUF16rest-s protein, in which we replaced arginine with tyrosine (R to Y) in repeat R3, tyrosine with arginine (Y to R) in repeat R4, and asparagine with tyrosine (N to Y) in repeat R12 (Figure 1B).

To obtain the fusion of REST-specific PUF proteins with the poly(A) polymerase GLD-2, the sequence of GLD-2 (kindly provided by Dr. Wickens, Department of Biochemistry, University of Wisconsin, Madison) was amplified using the Pfu DNA polymerase (#M7745; Promega) with primers flanked by restriction sites for *Bam*HI (Promega) and subsequently ligated by using the T4 DNA Ligase (Promega) with the *Bam*HI-digested vector CMV_3Xflag. PUF8wt and PUF8rest-sYwere amplified with primers flanked by restriction sites for NotI and XbaI (Promega). The PCR-amplified sequences were subsequently cloned in NotI/XbaI CMV_GLD2_3Xflag digested vector.

Mammalian Cell Cultures. Murine Neuro2a (N2a) neuroblastoma cells and human HEK293T cells were cultured in DMEM (#11965−092; Gibco) supplemented with 10% (v/v) FBS, glutamine (2 mM), and antibiotics, in a humidified 5% CO_2 atmosphere at 37 °C. Expression vectors were transiently transfected into cultured cells using Lipofectamine 2000 (Life Technologies) following standard procedures.

Protein Extraction and Western Blotting. Total protein lysates were obtained from cells lysed on ice for 20 min in RIPA buffer [20 mM Tris·HCl; pH 7.4; 1% Triton X-100; 10% (v/v) glycerol; 150 mM NaCl; 1% phenylmethylsulfonyl fluoride (PMSF); protease inhibitor mixture: Complete protease inhibitor mixture tablets (#04693116001; Roche Applied Science, Monza, IT)]. The final protein concentration was quantified by using the bicinchoninic acid (BCA) protein assay kit (#23225; Pierce Biotech, Monza, IT). Proteins were separated using precast 10% NuPAGE Novex Bis-Tris Gels (Life Technologies) and transferred to a nitrocellulose membrane. After incubation with primary antibodies, membranes were incubated with HRP-conjugated secondary antibodies and revealed by autoradiography using the Super-Signal West Pico Chemiluminescent Substrate (#34077; Thermo Scientific).

Immunofluorescence and Confocal Microscopy. Cells were fixed with 4% (wt/vol) paraformaldehyde and 20% (wt/vol) sucrose in PBS for 15 min at room temperature (RT) and permeabilized with 0.1%. Triton X-100 in PBS for 5 min at RT. Samples were blocked for 30 min in immunofluorescence buffer [2% (wt/vol) BSA, 10% (v/v) goat serum in PBS]. Primary and secondary antibodies were diluted in immunofluorescence buffer and incubated for 45 min at RT. Coverslips were mounted using ProLong antifade (#P36931; Life Technologies) and imaged by confocal microscopy. Confocal fluorescence images were obtained using a Leica SP5 confocal microscope with a 40× and 20× objective and analyzed with the Leica LAS AF software.

RNA Preparation and qRT-PCR. Total cellular RNA was extracted using the RNeasy Mini Kit (#74104; Qiagen), and isolated RNA was subjected to DNase I (#9PIM610; Promega) treatment. cDNA was synthesized starting from 0.5 μg of RNA according to the High-Capacity cDNA Reverse Transcription Kit manual (#4368814; Applied Biosystems) and used for qRT-PCR. qPCR was performed using the Power SYBR Green PCR Master Mix (#4309155; Applied Biosystems). The complete list of qRT-PCR primers is available in SI Materials and Methods.

Purification of Flag-Tagged Proteins. Cytosolic protein lysates (300 μg) were purified with 30 μL of Anti-FLAG M2Magnetic Beads (Sigma-Aldrich) prewashed with PBS/Tween 0.1%, and incubated in the following solution: 50 mM Tris-HCl, 150 mM NaCl, pH 7.4; in a total volume of 200 μL for 2 h RT. After incubation, beads were washed five times with PBS/Tween 0.1%, and bound proteins were eluted by incubation with 3× FLAG peptides (# F4799; Sigma-Aldrich) at the final concentration of 150 ng/μL, for 45 min at 4 °C. An aliquot of the total eluted volume (6%) was resolved on a denaturing 10% polyacrylamide gel and stained by using the Pierce Silver Stain Kit (#24612; Thermo Fisher Scientific) to verify the purity of the purified material. The final protein concentration was quantified by using the BCA protein assay kit (#23225; Pierce Biotech).

RNA Electrophoretic Mobility Shift Assay (EMSA). 3′ biotinylated RNA sequences (*NRE*: CCUGUAUAUAAGU; *RESTRNA8*: CCUUUAUAUAAGU; *2XNRE*: CCUGUUGUAUAUAAUAUAAGU; *RESTRNA16*: CCUGCUUUAUAUAAAUUAAGU; PUFrest16(2.0): CCAUUGGCUUAGUAAAUUAGU) were chemically synthesized (Sigma-Aldrich). To perform EMSAs, the LightShift Chemiluminescent RNA EMSA Kit (#20158; Thermo Fisher Scientific) was used. For each gel-shift reaction, a total of 10 nM of biotin-labeled probe was dissolved in the Binding Buffer provided, supplemented with 50% glycerol, 2 μg tRNA, 4 μg BSA, 50 mM KCl, 2 mM DTT, 0.02% Tween, 1 mM EDTA, and incubated with 1 μM-10 μM of purified proteins. The reaction mixture was incubated for 40 min at RT and resolved in a nondenaturing 10% polyacrylamide gel 0.5× TBE (1 h at 4 °C) that had been pre-electrophoresed for 60 min. RNA−protein complexes were transferred to nylon membranes and cross-linked for 13 min using a commercial UV-light cross-linking instrument (254 nm bulbs).

https://dx.doi.org/10.1021/acssynbio.0c00119
ACS Synth. Biol. 2020, 9, 2039−2054

Cross-Linking RNA Immunoprecipitation (CLIP). Transfected N2a cells (1×10^6) were subjected to UV cross-linking at 4000×100 $\mu J/cm^2$. Cells were subsequently harvested in buffer A [20 mM Tris-HCl (pH 8), 10 mM NaCl, 3 mM $MgCl_2$, 0.1% NP40, 10% Glycerol, 0.2 mM EDTA, 1 mM DTT, protease inhibitor mixture tablets (#04693116001; Roche), RNase inhibitor RiboLock (#EO0381; Thermo Scientific)] and centrifuged at 2500 *g* at 4 °C for 5 min. Agarose beads protein A salmon sperm (#16−157; Millipore) were precleaned with CLIP buffer (1XPBS, 0.1% SDS, 0.5% NP-40) and incubated for 1 h with the lysates plus RIPA buffer (10 mM Tris-HCl (pH 7.5), 1 mM EDTA, 140 mM NaCl, 0.5 mM EGTA, 1% TritonX-100, 0.1% SDS, 0.1% Na-deoxycolate) to a final volume of 1 mL. After centrifugation at 2500 *g* for 5 min, the supernatant was divided in two equal parts, to incubate with 5 μg antiflag rabbit (#F7425; SIGMA) and 2.5 μg antirabbit IgG antibodies (PP54B; Millipore) at 4 °C overnight. For each sample, 50 μL were kept as input RNA for real time detection, and 20 μL as input protein for WB detection. Precleaned beads were incubated with the antibody-protein-RNA complex at 4 °C for 2 h. Beads were washed with CLIP buffer and high salt wash buffer [5XPBS without Mg^{2+}/Ca^{2+}(#D1408; Sigma-Aldrich), 0.1% SDS, 0.5% NP-40], and resuspended in RIPA buffer. Proteinase K (#P6556; Sigma-Aldrich) was added to the samples and incubated 45 min at 45 °C before proceeding to RNA extraction by using the miRNeasy kit (#217004; Qiagen). Half of the total RNA extracted volume was used for cDNA synthesis and qRT-PCR. The primers used in qRT-PCR experiments are listed in SI Materials and Methods.

Luciferase Assays. Reporter plasmids for luciferase assays were produced starting from annealed oligos: 1 μg sense and 1 μg antisense oligos for the *NRE* and *RESTRNA8* sequences were resuspended in the following solution: 10 mM Tris-HCl (pH7.5), 0.1 M NaCl, 1 mM EDTA, and incubated in a T100 Thermocycler (Biorad) as follows: 95 °C, 4 min; 70 °C, 10 min; decrease temperature to 4 °C (0.1 °C/min). Annealed oligos were inserted *via* standard ligation procedures in the digested XhoI/NotI psiCHECK Vector 2.0, provided by Dr. Contestabile (Istituto Italiano di Tecnologia, Genoa, Italy). All constructs were verified by DNA sequencing. Reporter and expression vectors were transiently cotransfected into cultured cells using Lipofectamine 2000 (Invitrogen). Control samples were cotransfected with the empty vector corresponding to the effector plasmids. Luciferase activity was assayed after 48 h by using the Dual-Luciferase reporter assay system (Promega).

Statistical Analysis. Results are expressed as means ± SEM throughout. Data were analyzed by unpaired Student's *t*-test when two groups were compared, or one-way ANOVA followed by Tukey's multiple comparison test when more than two groups were analyzed.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acssynbio.0c00119.

Figure S1: *RESTRNA8*, *RESTRNA16*, *RESTRNA16−2.0* sequences are highlighted on REST 3′UTR; Figure S2: Characterization of the PUF16rest-2.0 construct; Figure S3: Nonbonded interactions energies and distances for the R8-N1 interface in the three 8-repeat constructs; Figure S4: Purification of PUF constructs; Figure S5: Evaluation of binding specificity of PUF8rest-ns and PUF16rest-ns; Figure S6: Cross-link RNA immunoprecipitation (CLIP) of the REST-PUF constructs mutated in the stacking residues for endogenous Gapdh mRNA, and Immunoblot of immunoprecipitated complexes revealed with antiflag antibodies; Figure S7: Modulation of the stability of REST mRNA through fusion of REST-specific PUF proteins and GLD2; Table S1: microRNA sites prediction analysis; Supporting Methods: Primer list (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Luca Maragliano − *Center for Synaptic Neuroscience and Technology, Istituto Italiano di Tecnologia, Genova 16132, Italy; IRCCS Ospedale Policlinico San Martino, Genova 16132, Italy;* orcid.org/0000-0002-5705-6967; Phone: +39-0105558382; Email: luca.maragliano@iit.it; Fax: +39-0105558491

Fabrizia Cesca − *Center for Synaptic Neuroscience and Technology, Istituto Italiano di Tecnologia, Genova 16132, Italy; Department of Life Sciences, University of Trieste, Trieste 34127, Italy;* orcid.org/0000-0003-2190-6314; Phone: +39-0405588727; Email: fabrizia.cesca@iit.it; Fax: +39-0105558491

Fabio Benfenati − *Center for Synaptic Neuroscience and Technology, Istituto Italiano di Tecnologia, Genova 16132, Italy; IRCCS Ospedale Policlinico San Martino, Genova 16132, Italy;* orcid.org/0000-0002-0653-8368; Phone: +39-01071781434; Email: fabio.benfenati@iit.it; Fax: +39-0105558491

Authors

Stefania Criscuolo − *Center for Synaptic Neuroscience and Technology, Istituto Italiano di Tecnologia, Genova 16132, Italy*

Mahad Gatti Iou − *Center for Synaptic Neuroscience and Technology, Istituto Italiano di Tecnologia, Genova 16132, Italy*

Assunta Merolla − *Center for Synaptic Neuroscience and Technology, Istituto Italiano di Tecnologia, Genova 16132, Italy; University of Genova, Genova 16132, Italy*

Complete contact information is available at: https://pubs.acs.org/10.1021/acssynbio.0c00119

Author Contributions

#S.C., M.G.I., and A.M. should be regarded as joint first authors.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Dr. Rackham (The University of Western Australia, Perth, Australia) for providing the plasmids coding PUF8wt and PUF16wt, Dr. di Bernardo (Telethon Institute of Genetics and Medicine, TIGEM, Naples, Italy) for the CMV_3Xflag vector, Dr. Contestabile (Istituto Italiano di Tecnologia, IIT, Genoa, Italy) for psiCHECK Vector 2.0, Dr. Irene Bozzoni (La Sapienza University, Rome, Italy) for technical help with the CLIP experiments. We acknowledge the IIT Platform "CompuNet" for CPU time and a PRACE (Partnership for Advanced Computing in Europe) infrastructure Award DECI-8 (to L.M.) allowing the access to resource Monte Rosa at Swiss National Supercomputing Centre. This work was

supported by the European Union Integrating Project DESIRE (Development and Epilepsy—Strategies for Innovative Research to improve diagnosis, prevention and treatment in children with difficult to treat Epilepsy; FP7-HEALTH-2013-INNOVATION-1 Grant 602531 to F.B.).

REFERENCES

(1) Schoenherr, C. J., and Anderson, D. J. (1995) The neuron-restrictive silencer factor (NRSF): a coordinate repressor of multiple neuron-specific genes. *Science 267* (5202), 1360−3.

(2) Chong, J. A., Tapia-Ramirez, J., Kim, S., Toledo-Aral, J. J., Zheng, Y., Boutros, M. C., Altshuller, Y. M., Frohman, M. A., Kraner, S. D., and Mandel, G. (1995) REST: a mammalian silencer protein that restricts sodium channel gene expression to neurons. *Cell (Cambridge, MA, U. S.) 80* (6), 949−57.

(3) Satoh, J., Kawana, N., and Yamamoto, Y. (2013) ChIP-Seq Data Mining: Remarkable Differences in NRSF/REST Target Genes between Human ESC and ESC-Derived Neurons. *Bioinf. Biol. Insights 7*, 357−368.

(4) Baldelli, P., and Meldolesi, J. (2015) The Transcription Repressor REST in Adult Neurons: Physiology, Pathology, and Diseases(1,2,3). *eNeuro 2* (4), ENEURO.0010-15.2015.

(5) Taylor, P., Fangusaro, J., Rajaram, V., Goldman, S., Helenowski, I. B., MacDonald, T., Hasselblatt, M., Riedemann, L., Laureano, A., Cooper, L., and Gopalakrishnan, V. (2012) REST is a novel prognostic factor and therapeutic target for medulloblastoma. *Mol. Cancer Ther. 11* (8), 1713−1723.

(6) Conti, L., Crisafulli, L., Caldera, V., Tortoreto, M., Brilli, E., Conforti, P., Zunino, F., Magrassi, L., Schiffer, D., and Cattaneo, E. (2012) REST controls self-renewal and tumorigenic competence of human glioblastoma cells. *PLoS One 7* (6), e38486.

(7) Calderone, A., Jover, T., Noh, K. M., Tanaka, H., Yokota, H., Lin, Y., Grooms, S. Y., Regis, R., Bennett, M. V., and Zukin, R. S. (2003) Ischemic insults derepress the gene silencer REST in neurons destined to die. *J. Neurosci. 23* (6), 2112−21.

(8) Lepagnol-Bestel, A. M., Zvara, A., Maussion, G., Quignon, F., Ngimbous, B., Ramoz, N., Imbeaud, S., Loe-Mie, Y., Benihoud, K., Agier, N., Salin, P. A., Cardona, A., Khung-Savatovsky, S., Kallunki, P., Delabar, J. M., Puskas, L. G., Delacroix, H., Aggerbeck, L., Delezoide, A. L., Delattre, O., Gorwood, P., Moalic, J. M., and Simonneau, M. (2009) DYRK1A interacts with the REST/NRSF-SWI/SNF chromatin remodelling complex to deregulate gene clusters involved in the neuronal phenotypic traits of Down syndrome. *Hum. Mol. Genet. 18* (8), 1405−14.

(9) Zuccato, C., Tartari, M., Crotti, A., Goffredo, D., Valenza, M., Conti, L., Cataudella, T., Leavitt, B. R., Hayden, M. R., Timmusk, T., Rigamonti, D., and Cattaneo, E. (2003) Huntingtin interacts with REST/NRSF to modulate the transcription of NRSE-controlled neuronal genes. *Nat. Genet. 35* (1), 76−83.

(10) Lu, T., Aron, L., Zullo, J., Pan, Y., Kim, H., Chen, Y., Yang, T. H., Kim, H. M., Drake, D., Liu, X. S., Bennett, D. A., Colaiacovo, M. P., and Yankner, B. A. (2014) REST and stress resistance in ageing and Alzheimer's disease. *Nature (London, U. K.) 507* (7493), 448−54.

(11) Bingham, A. J., Ooi, L., Kozera, L., White, E., and Wood, I. C. (2007) The repressor element 1-silencing transcription factor regulates heart-specific gene expression using multiple chromatin-modifying complexes. *Mol. Cell. Biol. 27* (11), 4082−92.

(12) Shimojo, M., Shudo, Y., Ikeda, M., Kobashi, T., and Ito, S. (2013) The small cell lung cancer-specific isoform of RE1-silencing transcription factor (REST) is regulated by neural-specific Ser/Arg repeat-related protein of 100 kDa (nSR100). *Mol. Cancer Res. 11* (10), 1258−68.

(13) Wagoner, M. P., Gunsalus, K. T., Schoenike, B., Richardson, A. L., Friedl, A., and Roopra, A. (2010) The transcription factor REST is lost in aggressive breast cancer. *PLoS Genet. 6* (6), e1000979.

(14) Liang, H., Studach, L., Hullinger, R. L., Xie, J., and Andrisani, O. M. (2014) Down-regulation of RE-1 silencing transcription factor (REST) in advanced prostate cancer by hypoxia-induced miR-106b ~ 25. *Exp. Cell Res. 320* (2), 188−99.

(15) Soldati, C., Bithell, A., Conforti, P., Cattaneo, E., and Buckley, N. J. (2011) Rescue of gene expression by modified REST decoy oligonucleotides in a cellular model of Huntington's disease. *J. Neurochem. 116* (3), 415−25.

(16) Kuwabara, T., Hsieh, J., Nakashima, K., Taira, K., and Gage, F. H. (2004) A small modulatory dsRNA specifies the fate of adult neural stem cells. *Cell (Cambridge, MA, U. S.) 116* (6), 779−93.

(17) Paonessa, F., Criscuolo, S., Sacchetti, S., Amoroso, D., Scarongella, H., Pecoraro Bisogni, F., Carminati, E., Pruzzo, G., Maragliano, L., Cesca, F., and Benfenati, F. (2016) Regulation of neural gene transcription by optogenetic inhibition of the RE1-silencing transcription factor. *Proc. Natl. Acad. Sci. U. S. A. 113* (1), E91−E100.

(18) Goldstrohm, A. C., Hall, T. M. T., and McKenney, K. M. (2018) Post-transcriptional Regulatory Functions of Mammalian Pumilio Proteins. *Trends Genet. 34* (12), 972−990.

(19) Hall, T. M. (2016) De-coding and re-coding RNA recognition by PUF and PPR repeat proteins. *Curr. Opin. Struct. Biol. 36*, 116−21.

(20) Abil, Z., and Zhao, H. (2015) Engineering reprogrammable RNA-binding proteins for study and manipulation of the transcriptome. *Mol. BioSyst. 11* (10), 2658−65.

(21) Kiani, S. J., Taheri, T., Rafati, S., and Samimi-Rad, K. (2017) PUF Proteins: Cellular Functions and Potential Applications. *Curr. Protein Pept. Sci. 18* (3), 250−261.

(22) Cho, P. F., Poulin, F., Cho-Park, Y. A., Cho-Park, I. B., Chicoine, J. D., Lasko, P., and Sonenberg, N. (2005) A new paradigm for translational control: inhibition via 5′-3′ mRNA tethering by Bicoid and the eIF4E cognate 4EHP. *Cell (Cambridge, MA, U. S.) 121* (3), 411−23.

(23) Goldstrohm, A. C., Hook, B. A., Seay, D. J., and Wickens, M. (2006) PUF proteins bind Pop2p to regulate messenger RNAs. *Nat. Struct. Mol. Biol. 13* (6), 533−9.

(24) Pique, M., Lopez, J. M., Foissac, S., Guigo, R., and Mendez, R. (2008) A combinatorial code for CPE-mediated translational control. *Cell (Cambridge, MA, U. S.) 132* (3), 434−48.

(25) Gu, W., Deng, Y., Zenklusen, D., and Singer, R. H. (2004) A new yeast PUF family protein, Puf6p, represses ASH1 mRNA translation and is required for its localization. *Genes Dev. 18* (12), 1452−65.

(26) Edwards, T. A., Pyle, S. E., Wharton, R. P., and Aggarwal, A. K. (2001) Structure of Pumilio reveals similarity between RNA and peptide binding motifs. *Cell (Cambridge, MA, U. S.) 105* (2), 281−9.

(27) Quenault, T., Lithgow, T., and Traven, A. (2011) PUF proteins: repression, activation and mRNA localization. *Trends Cell Biol. 21* (2), 104−12.

(28) Koh, Y. Y., Wang, Y., Qiu, C., Opperman, L., Gross, L., Tanaka Hall, T. M., and Wickens, M. (2011) Stacking interactions in PUF-RNA complexes. *RNA 17* (4), 718−27.

(29) Filipovska, A., and Rackham, O. (2011) Designer RNA-binding proteins: New tools for manipulating the transcriptome. *RNA Biol. 8* (6), 978−83.

(30) Wang, Y., Cheong, C. G., Hall, T. M., and Wang, Z. (2009) Engineering splicing factors with designed specificities. *Nat. Methods 6* (11), 825−30.

(31) Cooke, A., Prigge, A., Opperman, L., and Wickens, M. (2011) Targeted translational regulation using the PUF protein family scaffold. *Proc. Natl. Acad. Sci. U. S. A. 108* (38), 15870−5.

(32) Cao, J., Arha, M., Sudrik, C., Schaffer, D. V., and Kane, R. S. (2014) Bidirectional regulation of mRNA translation in mammalian cells by using PUF domains. *Angew. Chem., Int. Ed. 53* (19), 4900−4.

(33) Zhang, W., Wang, Y., Dong, S., Choudhury, R., Jin, Y., and Wang, Z. (2014) Treatment of type 1 myotonic dystrophy by engineering site-specific RNA endonucleases that target (CUG)(n) repeats. *Mol. Ther. 22* (2), 312−320.

(34) Shinoda, K., Tsuji, S., Futaki, S., and Imanishi, M. (2018) Nested PUF Proteins: Extending Target RNA Elements for Gene Regulation. *ChemBioChem 19* (2), 171−176.

https://dx.doi.org/10.1021/acssynbio.0c00119
ACS Synth. Biol. 2020, 9, 2039−2054

(35) Adamala, K. P., Martin-Alarcon, D. A., and Boyden, E. S. (2016) Programmable RNA-binding protein composed of repeats of a single modular unit. *Proc. Natl. Acad. Sci. U. S. A. 113* (19), E2579−88.
(36) Filipovska, A., Razif, M. F., Nygard, K. K., and Rackham, O. (2011) A universal code for RNA recognition by PUF proteins. *Nat. Chem. Biol. 7* (7), 425−7.
(37) Wang, X., McLachlan, J., Zamore, P. D., and Hall, T. M. (2002) Modular recognition of RNA by a human pumilio-homology domain. *Cell (Cambridge, MA, U. S.) 110* (4), 501−12.
(38) Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M., Jr., Jungkamp, A. C., Munschauer, M., Ulrich, A., Wardle, G. S., Dewell, S., Zavolan, M., and Tuschl, T. (2010) Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell (Cambridge, MA, U. S.) 141* (1), 129−41.
(39) Lu, G., and Hall, T. M. (2011) Alternate modes of cognate RNA recognition by human PUMILIO proteins. *Structure (Oxford, U. K.) 19* (3), 361−7.
(40) Cheong, C. G., and Hall, T. M. (2006) Engineering RNA sequence specificity of Pumilio repeats. *Proc. Natl. Acad. Sci. U. S. A. 103* (37), 13635−9.
(41) Kwak, J. E., Wang, L., Ballantyne, S., Kimble, J., and Wickens, M. (2004) Mammalian GLD-2 homologs are poly(A) polymerases. *Proc. Natl. Acad. Sci. U. S. A. 101* (13), 4407−12.
(42) Zhao, Y. Y., Mao, M. W., Zhang, W. J., Wang, J., Li, H. T., Yang, Y., Wang, Z., and Wu, J. W. (2018) Expanding RNA binding specificity and affinity of engineered PUF domains. *Nucleic Acids Res. 46* (9), 4771−4782.
(43) Miller, M. T., Higgin, J. J., and Hall, T. M. (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.
(44) Wang, B., and Ye, K. (2017) Nop9 binds the central pseudoknot region of 18S rRNA. *Nucleic Acids Res. 45* (6), 3559−3567.
(45) Bao, H., Wang, N., Wang, C., Jiang, Y., Liu, J., Xu, L., Wu, J., and Shi, Y. (2017) Structural basis for the specific recognition of 18S rRNA by APUM23. *Nucleic Acids Res. 45* (20), 12005−12014.
(46) Bhat, V. D., McCann, K. L., Wang, Y., Fonseca, D. R., Shukla, T., Alexander, J. C., Qiu, C., Wickens, M., Lo, T. W., Tanaka Hall, T. M., and Campbell, Z. T. (2019) Engineering a conserved RNA regulatory protein repurposes its biological function in vivo. *eLife*, DOI: 10.7554/eLife.43788.
(47) Gupta, Y. K., Lee, T. H., Edwards, T. A., Escalante, C. R., Kadyrova, L. Y., Wharton, R. P., and Aggarwal, A. K. (2009) Co-occupancy of two Pumilio molecules on a single hunchback NRE. *RNA 15* (6), 1029−35.
(48) Wang, X., Zamore, P. D., and Hall, T. M. (2001) Crystal structure of a Pumilio homology domain. *Mol. Cell 7* (4), 855−65.
(49) Opperman, L., Hook, B., DeFino, M., Bernstein, D. S., and Wickens, M. (2005) A single spacer nucleotide determines the specificities of two mRNA regulatory proteins. *Nat. Struct. Mol. Biol. 12* (11), 945−51.
(50) Miller, M. A., and Olivas, W. M. (2011) Roles of Puf proteins in mRNA degradation and translation. *Wiley Interdiscip Rev: RNA 2* (4), 471−92.
(51) Friend, K., Campbell, Z. T., Cooke, A., Kroll-Conner, P., Wickens, M. P., and Kimble, J. (2012) A conserved PUF-Ago-eEF1A complex attenuates translation elongation. *Nat. Struct. Mol. Biol. 19* (2), 176−83.
(52) Gruber, A. R., Lorenz, R., Bernhart, S. H., Neubock, R., and Hofacker, I. L. (2008) The Vienna RNA websuite. *Nucleic Acids Res. 36*, W70−W74.
(53) Kedde, M., van Kouwenhove, M., Zwart, W., Oude Vrielink, J. A., Elkon, R., and Agami, R. (2010) A Pumilio-induced RNA structure switch in p27−3′ UTR controls miR-221 and miR-222 accessibility. *Nat. Cell Biol. 12* (10), 1014−20.
(54) Hug, N., Longman, D., and Caceres, J. F. (2016) Mechanism and regulation of the nonsense-mediated decay pathway. *Nucleic Acids Res. 44* (4), 1483−95.
(55) Suh, N., Crittenden, S. L., Goldstrohm, A., Hook, B., Thompson, B., Wickens, M., and Kimble, J. (2009) FBF and its dual control of gld-1 expression in the Caenorhabditis elegans germline. *Genetics 181* (4), 1249−60.
(56) Kashiwabara, S., Nakanishi, T., Kimura, M., and Baba, T. (2008) Non-canonical poly(A) polymerase in mammalian gametogenesis. *Biochim. Biophys. Acta, Gene Regul. Mech. 1779* (4), 230−8.
(57) Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: visual molecular dynamics. *J. Mol. Graphics 14* (1), 33−8.
(58) Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys. 79* (2), 926−935.
(59) Darden, T., York, D., and Pedersen, L. (1993) Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys. 98* (12), 10089−10092.
(60) Feller, S. E., Zhang, Y., Pastor, R. W., and Brooks, B. R. (1995) Constant Pressure Molecular Dynamics Simulation: The Langevin Piston Method. *J. Chem. Phys. 103*, 4613−4621.
(61) Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L., and Schulten, K. (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem. 26* (16), 1781−802.
(62) Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., and Simmerling, C. (2015) ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput. 11* (8), 3696−713.
(63) Zgarbova, M., Otyepka, M., Sponer, J., Mladek, A., Banas, P., Cheatham, T. E., 3rd, and Jurecka, P. (2011) Refinement of the Cornell et al. Nucleic Acids Force Field Based on Reference Quantum Chemical Calculations of Glycosidic Torsion Profiles. *J. Chem. Theory Comput. 7* (9), 2886−2902.
(64) Perez, A., Marchan, I., Svozil, D., Sponer, J., Cheatham, T. E., 3rd, Laughton, C. A., and Orozco, M. (2007) Refinement of the AMBER force field for nucleic acids: improving the description of alpha/gamma conformers. *Biophys. J. 92* (11), 3817−29.
(65) Jo, S., Kim, T., Iyer, V. G., and Im, W. (2008) CHARMM-GUI: a web-based graphical user interface for CHARMM. *J. Comput. Chem. 29* (11), 1859−65.
(66) Qi, Y., Ingolfsson, H. I., Cheng, X., Lee, J., Marrink, S. J., and Im, W. (2015) CHARMM-GUI Martini Maker for Coarse-Grained Simulations with the Martini Force Field. *J. Chem. Theory Comput. 11* (9), 4486−94.
(67) Hsu, P. C., Bruininks, B. M. H., Jefferies, D., Cesar Telles de Souza, P., Lee, J., Patel, D. S., Marrink, S. J., Qi, Y., Khalid, S., and Im, W. (2017) CHARMM-GUI Martini Maker for modeling and simulation of complex bacterial membranes with lipopolysaccharides. *J. Comput. Chem. 38* (27), 2354−2363.
(68) de Jong, D. H., Singh, G., Bennett, W. F., Arnarez, C., Wassenaar, T. A., Schafer, L. V., Periole, X., Tieleman, D. P., and Marrink, S. J. (2013) Improved Parameters for the Martini Coarse-Grained Protein Force Field. *J. Chem. Theory Comput. 9* (1), 687−97.
(69) Fiorin, G., Klein, M. L., and Henin, J. (2013) Using Collective Variables to Drive Molecular Dynamics Simulations. *Mol. Phys. 111*, 3345−3362.
(70) Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H., and Kollman, P. A. (1992) The Weighted Histogram Analysis Method for Free- energy Calculations On Biomolecules. I. The method. *J. Comput. Chem. 13*, 1011−1021.
(71) Campbell, Z. T., Valley, C. T., and Wickens, M. (2014) A protein-RNA specificity code enables targeted activation of an endogenous human transcript. *Nat. Struct. Mol. Biol. 21* (8), 732−8.