Published in final edited form as:

Curr Opin Struct Biol. 2009 February ; 19(1): 110–115. doi:10.1016/j.sbi.2008.12.009.

Understanding and engineering RNA sequence specificity of PUF proteins

Gang Lu, Stephen J. Dolgner, and Traci M. Tanaka Hall*

Laboratory of Structural Biology, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC 27709

Summary

PUF proteins are RNA-binding proteins named for founding members PUMILIO and FBF. Together these proteins represent the range of known RNA recognition properties. PUMILIO is a prototypical PUF protein whose RNA sequence specificity is simple, elegant, and predictable. FBF displays differences in RNA recognition that represent divergence from the prototype. Here we review recent studies that examine the engineering of sequence specificity of PUF proteins and its applications as well as studies that increase our understanding of the natural diversity of RNA recognition by this family of proteins.

Introduction

PUF family proteins are named for founding members *D. melanogaster* PUMILIO and *C. elegans* FBF (fem-3 binding factor) [1,2]. They regulate mRNA expression by binding to specific sequences in the 3' UTR of their target mRNAs and recruiting other proteins that promote mRNA degradation and translational repression. One of the first targets identified for PUF proteins was *hunchback* (*hb*) mRNA, which is recognized by PUMILIO in *Drosophila* and contains tandem copies of the Nanos Response Element (NRE) [3]. The NRE sequences are necessary for repression of *hb* mRNA and establishment of anterior-posterior polarity of the embryo. The core sequence of the *hb* NRE (5'-U₁G₂U₃A₄-U/C₅-A₆U₇A₈-3') serves as a model for PUF protein target sites.

The hallmark of all PUF proteins is a characteristic sequence-specific, RNA-binding domain, known as the Pumilio homology domain (PUM-HD) or PUF domain. Crystallographic studies have shown that the PUM-HD comprises eight tandem copies of α -helical PUF repeats flanked by two imperfect pseudo-repeats, one at each terminus, which adopt a crescent shape [4,5] (Figure 1a). The consensus RNA recognition sequences of PUF proteins begin with 5'-UGUR (where R represents a purine), which we designate positions 1–4 [6–10]. Following this sequence are somewhat variable 3' sequences that may contain some conserved elements across PUF proteins [2].

Structures of the PUM-HD of human PUMILIO 1 (PUM1) bound to *hb* NRE RNA have revealed a general mechanism of modular RNA recognition by PUF repeats [11]. RNA binds to the inner concave surface of PUM1. Each PUF repeat recognizes a single RNA base through

*Correspondence: hall4@niehs.nih.gov, phone: 1-919-541-1017, fax: 1-919-316-4617.

Other authors' e-mail addresses: Gang Lu, lug@niehs.nih.gov; Stephen Dolgner, dolgn001@mc.duke.edu

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

three conserved side chains, two that make hydrogen bond or van der Waals interactions with the edge of an RNA base and a third side chain that stacks with the same base and/or the preceding base. The RNA runs 'antiparallel' to the protein such that nucleotides 1 to 8 are recognized individually by PUF repeats 8 to 1, respectively (Figure 1a). This simple 1 PUF repeat:1 RNA base recognition by PUM1 represents RNA recognition by a prototypical PUF protein.

Design of PUF protein sequence specificity

The crystal structures of PUM1:RNA complexes revealed that particular sets of side chains in a repeat recognize specific bases (Figure 1b). Such an elegant RNA recognition pattern suggested that the sequence specificity of PUM1 could be altered by site-directed mutagenesis. Following proof-of-concept experiments that had been done in collaboration with Phillip Zamore's laboratory [11], our laboratory designed seven mutant proteins using this recognition code that predictably and preferentially bound to their designed target sequences [12]. It was only necessary to change the two residues in a repeat that interact with the Watson-Crick edge of the base. These designed PUM1 proteins bind tightly to their cognate RNAs with binding affinities ranging from 0.051 to 18 nM, while wild-type PUM1 binds to *hb* NRE RNA with a K_D of 0.48 nM. Thus many of the designed PUM1 proteins bind as tightly as wild-type PUM1. The ability to design the sequence specificity of PUF proteins presents opportunities to apply this technology to study other biological questions.

Since PUM1 can be designed to recognize a given RNA sequence, it can be tethered to additional protein domains to direct them to a selected RNA target. Such an approach has been used to tether effector domains or fluorescent modules to MS2 coat protein, but this requires the insertion of MS2 hairpin sequences in the target RNA [13–15]. Using designed PUM1 as the RNA-binding module, endogenous RNAs can be targeted. Ozawa, *et al.* showed that PUM1 could be used to track an endogenous RNA in live cells [16]. They investigated localization of mitochondrial RNA encoding NADH dehydrogenase subunit 6 (ND6). To introduce higher specificity than an 8-base sequence, they used two designed PUM1 modules to target two 8-base sequences separated by a 5-base linker. Each PUM1 module was fused to a non-fluorescent complementary portion of green fluorescent protein (GFP) or Venus yellow fluorescent protein (Figure 2a). When both proteins bound to target RNA, GFP or Venus was reconstituted, producing fluorescence. Slow dissociation of the reconstituted protein after target RNA degradation can be overcome by monitoring fluorescence recovery after photobleaching.

Opperman *et al.* demonstrated another use of the design of PUF protein specificity in their studies of *C. elegans* PUF protein target recognition [10]. They had discovered that two of the worm PUF proteins appeared to recognize different length core sequences. PUF-8 recognizes a consensus sequence similar to the 8-base *hb* NRE, while FBF, one of the best studied PUF proteins, prefers a 9-base sequence similar to the *hb* NRE, but with an additional base between the 4th and 5th bases in the NRE. Since both proteins contain 8 repeats, how does FBF accommodate this additional base? Using a clever application of the PUF protein RNA recognition code, Opperman, *et al.* sought to distinguish between two likely models: the additional base could be accommodated in the middle of the FBF repeats with the 5' and 3' bases binding to repeats 8 and 1, respectively, or the first 8 bases of RNA could bind with the familiar 1 repeat:1 RNA base pattern and the 9th base could be recognized in some different way (Figure 2b).

To this end, they altered the base specificity of selected repeats, which would result in characteristic sequence specificity depending on which of the models is correct. They probed binding of the mutated proteins to various sequences by yeast 3-hybrid screening, and the

results were clear that the extra base in the FBF recognition sequence is accommodated in the middle of the FBF repeats (Figure 2b). Thus nature has evolved its own design of FBF to alter RNA specificity.

Natural modulation of RNA specificity of PUF proteins

Studies of other PUF proteins provide further evidence that nature has modified PUF proteins to broaden the range of sequences that can be recognized. Very recently, Stumpf, *et al.* have shown that the RNA specificity of the closely-related *C. elegans* PUF-5 and PUF-6 proteins encompasses a core 10-nt sequence beginning with a 5' UGU sequence and ending with a 3' UGU sequence [17]. PUF-5 and PUF-6 both contain eight repeats and so these PUF proteins do not use the simple 1 PUF repeat binding to 1 RNA base as observed in the PUM1 structures.

Evidence of similar diversity of RNA recognition exists with *S. cerevisiae* PUF proteins. Yeast express six PUF proteins, PUF1 (also called JSN1), PUF2, PUF3, PUF4, PUF5 (also called MPT5), and PUF6, which appear to regulate different sets of target mRNAs [7,18]. Genomic screens have identified consensus RNA recognition sequences for PUF3, PUF4, and PUF5 that begin with "UGUA" and end with "UA". Although each of the three proteins contains eight PUF repeats, PUF3, PUF4, and PUF5 recognize 8-nt, 9-nt, and 10-nt RNA sequences, respectively, with varying sequences between the 5' and 3' conserved sequence motifs. Thus PUF4 and PUF5 proteins also do not recognize their RNA target sequences with the prototypical pattern of 1 PUF repeat:1 RNA base.

Recent structural and biochemical studies provide insight into how PUF protein specificity is broadened to accommodate these longer sequences. The crystal structure of yeast PUF4 in complex with its natural RNA target sequence found in the 3' UTR of *HO* endonuclease mRNA (5'-U₁G₂U₃A₄U₅A₆U₇U₈A₉-3'), reveals how a 9-nt RNA sequence can be recognized by eight PUF repeats [19]. Furthermore, crystal structures of PUM1 in complex with two different noncognate RNA sequences, *cycB*_{reverse} (5'-U₁U₂U₃A₄A₅U₆G₇U₈U₉-3') and the *HO* endonuclease PUF5 binding site (5'-U₁U₁G₂U₃A₄A₅U₆A₇U₈U₉A₁₀-3'), were reported by Gupta, *et al.* [20] and show how PUM1 can bind to non-consensus sequences, including the accommodation of a 9th nucleotide by a mechanism similar to PUF4.

How PUF4 accommodates an extra base, U₇

Many features of the structure of PUF4 and its interaction with its target RNA are similar to what was seen in the structures of human and fly PUMILIO proteins. PUF4 comprises eight tandem α -helical PUF repeats flanked by additional short sequences at both termini, as in PUM1 and *D. melanogaster* PUMILIO. The overall curvature of PUF4 is less than that of PUM1, apparently due to different angles between repeats 3 and 4, which may provide a more extended RNA-binding surface favoring a 9-nt sequence over a prototypical 8-nt sequence. In the structure of the PUF4:*HO* RNA complex, repeats 8 to 5 of PUF4 recognize the conserved 5'-UGUA sequence, repeat 3 recognizes A₆, and repeats 2 and 1 recognize a 3'-UA sequence, like the interaction between PUM1 and the same sequences in NRE RNA. A notable difference in RNA binding is that in repeat 5 of PUM1 the side chain of Arg1008 is sandwiched between A₄ and C₅/U₅, however in PUF4, the equivalent side chain of Cys724 is too short to stack with A₄ (Figure 3a). The shorter side chain of Cys724 allows U₅ to stack directly with A₄, with no hydrogen bonds observed between repeat 4 of PUF4 and U₅.

The extra base in the PUF4 consensus sequence is U₇, which appears to be turned away from the RNA-binding surface and does not interact directly with PUF4 (Figure 3a). Genomic screens indicated that potential target RNA sequences of PUF4 have U, A or C, but rarely G, at the seventh position [7]. By determining the equilibrium binding constants of PUF4 for 9-nt recognition sequences with U, A, or C at the seventh position, Miller, *et al.* found that binding

is about two-fold weaker when the 7th base is a G, which is equivalent to the decrease in PUF4 binding affinity for an 8-nt target RNA vs. 9-nt.

Engineering PUF proteins to modulate recognition of 8- vs 9-nt sequences

With the crystal structure of PUF4, we have a picture of how one PUF protein binds to a 9-base target sequence, but how well do we understand what modifications produce the change in specificity and will the same mechanisms be used for other ‘divergent’ PUF proteins? The first piece of information to guide our understanding comes from Opperman, *et al.* [10]. By creating chimeric proteins they showed that a 74-residue fragment of FBF including parts of repeats 4 and 5 could replace the equivalent region of PUF-8 to create a protein with 9-base target specificity. Thus the specificity can be switched by changing protein elements near the inserted base.

Given the success of Opperman, *et al.* and PUF protein design in general, we attempted to engineer PUF4 to recognize an 8-nt PUF3-BS (binding site)/*hb* NRE RNA sequence instead of 9-nt. We created a T650C/C724R double mutant of PUF4, which changed two residues in PUF4 to what is found in PUF3 or PUM1. This mutant protein showed a two-fold preference for an 8-nt PUF3-BS/*hb* NRE RNA recognition sequence (5'-U₁G₂U₃A₄U₅A₆U₇A₈-3') over the 9-nt PUF4 recognition sequence. Furthermore, a crystal structure of this mutant bound with the 8-nt PUF3-BS RNA showed that the longer side chain of Arg724, which replaced Cys724, stacks between bases A₄ and U₅, and Cys650 contacts the 2' OH of A₆ through van der Waals interaction, as seen with PUM1 (Figure 3b). Thus this structure-guided site directed mutagenesis approach allowed conversion of the specificity of PUF4. The overall degree of curvature of PUF4 did not change in the mutant protein, and perhaps the somewhat flatter and elongated RNA-binding surface is not best suited for an 8-nt sequence. To fully optimize the switch in specificity from 9-nt to 8-nt or vice versa may require understanding what changes alter the overall curvature of PUF proteins.

PUM1 can accommodate an extra nucleotide, U₆

Gupta, *et al.* [20] took another approach to study the plasticity of PUF protein RNA recognition by studying PUM1 binding to non-cognate RNA sequences. Can the prototypical PUF protein PUM1 accommodate varied RNA sequences? The answer appears to be yes, at least *in vitro*. With biochemical binding assays, Gupta, *et al.* showed that PUM1 binds to two non-cognate sequences with relatively high affinity compared with binding to the PUMILIO target sequence in *D. melanogaster* cyclinB mRNA. Crystal structures of PUM1 in complex with the two non-cognate RNA sequences, cycB_{reverse} and *HO* PUF5-BS, reveal that for both sequences, the U₆ base is flipped away from the RNA-binding surface and not contacted by the protein (Figure 3c). Bases 5 and 7 on either side of U₆ interact with repeats 4 and 3, respectively, similar to interaction with C₅/U₅ and A₆ in the PUM1:*hb* NRE RNA complexes. Thus a prototypical PUF protein can also bind to RNA sequences with an additional base, similar to the way PUF4 binds to its natural RNA target. It is unclear if such sequences are recognized by PUM1 *in vivo*, but the ability to optimize target recognition by flipping out a base suggests a mechanism for evolution of PUF proteins that recognize longer target RNA sequences.

PUM1 recognizes non-cognate bases

The non-cognate sequences in the structures of PUM1 with cycB_{reverse} and *HO* PUF5-BS also present RNA bases to some repeats that do not match the optimal ‘code’ used to design specificity of PUM1. Repeats 1, 3, 4, and 7 of PUM1 bind to non-cognate bases in structures with cycB_{reverse} or *HO* PUF5-BS RNAs. In each case one or both of the side chains that would typically recognize the edge of the RNA base, contact the non-cognate base. For example, repeat 1 bears Gln867 and Ser863 that select for A₈ in the *hb* NRE sequence. In the two non-

cognate sequences, repeat 1 instead binds U₉ through two hydrogen bonds between Gln867 and the base. Repeat 4 uses Gln975 to bind to U₅ or C₅ in *hb* NRE RNA, but instead binds to A₅ in *cycB_{reverse}* and *HO* PUF5-BS RNAs. These tolerances of alternative bases give some insight into the question of how PUF proteins appear to recognize target RNAs with a wider range of sequences than suggested by the PUF protein RNA recognition code. Recently MAP kinase mRNAs have been identified as targets of human PUM2 protein, which is very similar in sequence to PUM1 in the RNA-binding domain [21]. The target sequence in *erk2* mRNA differs from the *hb* NRE with a C at the 8th position rather than A, and the sequence in the p38 target RNAs contain A or G at the 5th position, two positions that show flexibility in the PUM1:non-cognate RNA structures.

Perhaps the most surprising accommodation is a U at position 2 in the *cycB_{reverse}* RNA. This substitution results in a change from the highly conserved 5'-UGUR sequence to 5'-UUUA in the *cycB_{reverse}* RNA, and substitution of a U at this position in *hb* NRE RNA results in >100-fold weaker binding by PUM1 [12]. Repeat 7 of PUM1 interacts with U₂ in *cycB_{reverse}* RNA with fewer contacts than seen with G₂ in the canonical RNA recognition sequence. The base of U₂ does not have well defined electron density, which indicates U₂ is somewhat mobile. The decreased interaction with U₂ and increased mobility of the base may explain the weaker binding affinity when G₂ is replaced by U₂.

Conclusions

As our knowledge of PUF protein RNA recognition properties increases, we can better appreciate how this RNA-binding module is used to recognize a wider range of sequences than anticipated from the first structures of PUM1. The plasticity of individual repeats and the ability to accommodate additional nucleotides by flipping them away from the RNA-binding surface help to explain why the RNA-binding residues might be highly conserved between PUF proteins even when the natural RNA sequences they recognize are not as well conserved. The design of RNA sequence specificity of PUM1 has now been applied to image endogenous RNA in living cells and leads the way to approach opportunities to apply engineered PUF domains for other purposes, such as regulating alternative splicing or tethering domains that direct endogenous mRNA localization, stability or translation.

Acknowledgments

We are grateful to our colleagues for critical reading of this manuscript and P. Cacioppo of the NIEHS Arts and Photography group for graphics design. This work was supported by the Intramural Research Program of the National Institutes of Health, National Institute of Environmental Health Sciences.

References

Papers of particular interest, published within the period of review, have been highlighted as:

* of special interest

** of outstanding interest

1. Wharton RP, Aggarwal AK. mRNA Regulation by Puf Domain Proteins. *Sci STKE* 2006;2006:pe37. [PubMed: 17003467]
2. Wickens M, Bernstein DS, Kimble J, Parker R. A PUF family portrait: 3'UTR regulation as a way of life. *Trends Genet* 2002;18:150–157. [PubMed: 11858839]
3. Murata Y, Wharton RP. Binding of pumilio to maternal hunchback mRNA is required for posterior patterning in *Drosophila* embryos. *Cell* 1995;80:747–756. [PubMed: 7889568]
4. Edwards TA, Pyle SE, Wharton RP, Aggarwal AK. Structure of Pumilio reveals similarity between RNA and peptide binding motifs. *Cell* 2001;105:281–289. [PubMed: 11336677]

5. Wang X, Zamore PD, Hall TM. Crystal structure of a Pumilio homology domain. *Mol Cell* 2001;7:855–865. [PubMed: 11336708]
6. Bernstein D, Hook B, Hajarnavis A, Opperman L, Wickens M. Binding specificity and mRNA targets of a *C. elegans* PUF protein, FBF-1. *RNA* 2005;11:447–458. [PubMed: 15769874]
- 7*. Gerber AP, Herschlag D, Brown PO. Extensive association of functionally and cytologically related mRNAs with Puf family RNA-binding proteins in yeast. *PLoS Biol* 2004;2:342–354. Genomic screening reveals mRNA target specificity of yeast PUF proteins.
8. Gerber AP, Luschig S, Krasnow MA, Brown PO, Herschlag D. Genome-wide identification of mRNAs associated with the translational regulator PUMILIO in *Drosophila melanogaster*. *Proc Natl Acad Sci U S A* 2006;103:4487–4492. [PubMed: 16537387]
- 9*. Morris AR, Mukherjee N, Keene JD. Ribonomic analysis of human Pum1 reveals cis-trans conservation across species despite evolution of diverse mRNA target sets. *Mol Cell Biol* 2008;28:4093–4103. [PubMed: 18411299] Genome-wide study demonstrates that PUM1 regulates mRNAs encoding functionally related regulatory proteins.
- 10*. Opperman L, Hook B, DeFino M, Bernstein DS, Wickens M. A single spacer nucleotide determines the specificities of two mRNA regulatory proteins. *Nature Struct Mol Biol* 2005;12:945–951. [PubMed: 16244662] Biochemical analysis of RNA-binding specificity of *C. elegans* FBF and PUF-8 reveals unique specificities. Chimeric proteins demonstrate the minimal region of FBF to transfer its specificity to PUF-8.
11. Wang X, McLachlan J, Zamore PD, Hall TM. Modular recognition of RNA by a human pumilio-homology domain. *Cell* 2002;110:501–512. [PubMed: 12202039]
- 12*. Cheong CG, Hall TM. Engineering RNA sequence specificity of Pumilio repeats. *Proc Natl Acad Sci U S A* 2006;103:13635–13639. [PubMed: 16954190] Analysis of seven mutant PUM1 proteins with designed RNA-binding specificity.
13. Bertrand E, Chartrand P, Schaefer M, Shenoy SM, Singer RH, Long RM. Localization of ASH1 mRNA particles in living yeast. *Mol Cell* 1998;2:437–445. [PubMed: 9809065]
14. Collier JM, Gray NK, Wickens MP. mRNA stabilization by poly(A) binding protein is independent of poly(A) and requires translation. *Genes Dev* 1998;12:3226–3235. [PubMed: 9784497]
15. Graveley BR, Hertel KJ, Maniatis T. A systematic analysis of the factors that determine the strength of pre-mRNA splicing enhancers. *EMBO J* 1998;17:6747–6756. [PubMed: 9822617]
- 16*. Ozawa T, Natori Y, Sato M, Umezawa Y. Imaging dynamics of endogenous mitochondrial RNA in single living cells. *Nature Methods* 2007;4:413–419. [PubMed: 17401370] First use of designed PUF proteins for RNA tracking.
- 17*. Stumpf CR, Kimble J, Wickens M. A *Caenorhabditis elegans* PUF protein family with distinct RNA binding specificity. *RNA* 2008;14:1550–1557. [PubMed: 18579869] Analysis of RNA-binding specificity of *C. elegans* PUF-5/6 family demonstrates a unique 10-nt core consensus sequence with 5' and 3' UGU motifs.
18. Gu W, Deng Y, Zenklusen D, Singer RH. A new yeast PUF family protein, Puf6p, represses ASH1 mRNA translation and is required for its localization. *Genes Dev* 2004;18:1452–1465. [PubMed: 15198983]
- 19*. Miller MT, Higgin JJ, Hall TM. Basis of altered RNA-binding specificity by PUF proteins revealed by crystal structures of yeast Puf4p. *Nature Struct Mol Biol* 2008;15:397–402. [PubMed: 18327269] Crystal structures of yeast PUF4 in complex with a 9-nt target RNA and a mutant PUF4 protein with 8-nt RNA specificity.
- 20*. Gupta YK, Nair DT, Wharton RP, Aggarwal AK. Structures of human Pumilio with noncognate RNAs reveal molecular mechanisms for binding promiscuity. *Structure* 2008;16:549–557. [PubMed: 18328718] Unexpected finding that PUM1 can accommodate a 9-nt RNA sequence by flipping a base away from the RNA-binding surface and using non-cognate protein:RNA interactions.
- 21*. Lee MH, Hook B, Pan G, Kershner AM, Merritt C, Seydoux G, Thomson JA, Wickens M, Kimble J. Conserved regulation of MAP kinase expression by PUF RNA-binding proteins. *PLoS Genet* 2007;3:e233. [PubMed: 18166083] Demonstration that *C. elegans* and human PUF proteins regulate MAP kinase mRNAs working together with MAP kinase phosphatase to restrict MAP kinase activity.

Figure 1.

Prototypical PUF protein RNA recognition by PUM1. (a) Crystal structure of PUM1 in complex with *hb* NRE RNA. Repeats are colored alternately blue and yellow and side chains that contact the RNA are shown. RNA and PUM1 side chains are colored by atom type (blue, nitrogen; red, oxygen; grey, carbon; orange, phosphorous; yellow, sulfate). Side chains that contact the edges of bases are shown with blue or yellow carbon atoms, and those that form stacking interactions are shown with magenta carbon atoms. (b) PUM1 RNA recognition code. Schematic representation of recognition of uracil, adenosine and guanine by PUM1 is shown. Recognition of cytosine cannot yet be engineered, but a single cytosine can be tolerated at the 5th position in a recognition sequence. Dotted lines indicate hydrogen bonds,))) indicates van der Waals contacts, and |||| indicates stacking interactions.

Figure 2.

Applications of designed PUF proteins. (a) Tracking of endogenous RNA by fusing designed PUM1 with fluorescent protein (GFP). An N-terminal fragment of GFP (green) is fused at the N-terminus of PUM1 (PUM1-N) and a C-terminal fragment of GFP (yellow) is fused at the C-terminus of a second molecule of PUM1 (PUM1-C). Binding of both molecules to the target RNA allows reconstitution of GFP fluorescence. (b) Understanding FBF RNA recognition by design of repeat specificity. Schematic representations of two possible models of FBF:9-nt RNA interaction are shown at the top of the figure. Side chains that likely contact the RNA bases are shown. Model 1 shows the 5th base (yellow) accommodated in the center of the protein while Model 2 shows the 9th base (yellow) accommodated near the N terminus of the

protein. Model 2 requires non-canonical repeat:base interactions, some of which were observed in structures of PUM1 by Gupta, *et al.* [20]. One example of how site-directed mutagenesis was used to probe FBF RNA interaction is shown in the bottom figure. Opperman, *et al.* [10] mutated repeat 2 to recognize a guanine (orange) and found that the mutant protein preferred the sequence A₇G₈A₉ (bracketed sequence) as predicted by Model 1 (Model 2 predicts G₇U₈A₉, which was not bound by the mutant FBF).

Figure 3.

Accommodation of an 'extra' RNA base by PUF4 and PUM1. (a) Crystal structure of PUF4 in complex with *HO* target RNA reveals that the 7th base is flipped away from the RNA-binding surface. This "flipping-out" of U₇ is stabilized by Arg651 of repeat 3, which stacks between bases A₆ and U₈ and directly contacts the backbone ribose of U₇. (b) Crystal structure of PUF4 T650C/C724R mutant demonstrates that Arg724 stacks between bases 4 and 5. (c) Crystal structure of PUM1 in complex with *HO* PUF5-BS RNA. The 6th base is flipped away from the RNA-binding surface to best accommodate the sequence. Tyr1005 in repeat 5 and Gln968 in repeat 4 stabilize the displaced U₆ through van der Waals interactions with its phosphate and ribose groups, respectively. Atoms are colored as in Figure 1a.