nature
structural &
molecular biology

A single spacer nucleotide determines the specificities of two mRNA regulatory proteins

Laura Opperman, Brad Hook, Mia DeFino, David S Bernstein & Marvin Wickens

Regulation of messenger RNA is crucial in many contexts, including development, memory and cell growth. The 3′ untranslated region is a rich repository of regulatory elements that bind proteins and microRNAs. Here we focus on PUF proteins, an important family of mRNA regulatory proteins crucial in stem-cell proliferation, pattern formation and synaptic plasticity. We show that two *Caenorhabditis elegans* PUF proteins, FBF and PUF-8, differ in RNA-binding specificity. FBF requires the presence of a single 'extra' nucleotide in the middle of an eight-nucleotide site, whereas PUF-8 requires its absence. A discrete protein segment is responsible for the difference. We propose that a structural distortion in the central region of FBF imposes the requirement for the additional nucleotide and that this mode of PUF specificity may be common. We suggest that new specificities can be designed and selected using the PUF scaffold.

The importance of control over mRNA in animal cells has emerged unequivocally in the last decade. mRNA processing, transport, translation and stability all can be controlled to determine how much protein is produced from a gene, in which cell and at what time in development[1–7]. mRNAs that have entered the cytoplasm often are regulated by specific proteins and microRNAs. These regulatory molecules determine which mRNAs will be controlled and in what way. Whereas base pairing dictates the specificity of microRNAs, recognition by regulatory proteins is idiosyncratic. Here, we identify the specificity determinants of two regulatory proteins in the PUF protein family. Members of this family bind selectively to elements in the 3′ untranslated regions (3′ UTRs) of specific mRNAs and so repress specific genes at the mRNA level[8].

C. elegans fem-3 mRNA binding factor (FBF) and PUF-8 are distant relatives in the Pumilio/FBF (PUF) protein family[8]. PUF proteins have eight characteristic repeats, each of which contains approximately 40 residues and forms a repeat structure consisting of three α-helices (**Fig. 1a**)[9–11]. In the cocrystal of human Pumilio bound to RNA, the α-helices that contact RNA in the eight repeats are arranged like rungs on a ladder on one face of the protein; an elongated RNA lies on this surface, and each of its bases is recognized predominantly by interactions with side chains of the α-helix opposite (**Fig. 1a**)[11]. This structure probably shows common features of PUF-RNA interactions, including the 'one nucleotide/one α-helix' pattern, but does not reveal the basis of *in vivo* specificity. The residues that contact the RNA in the cocrystal are highly conserved among all PUF proteins[8,11], yet different PUF proteins bind different mRNAs (see for example refs. 12,13) and so regulate distinct pathways[8,12]. Using *C. elegans* FBF and PUF-8 as representative examples, we set out to determine the basis of PUF protein specificity.

RESULTS

Different specificities

C. elegans PUF-8 is a close relative of *Drosophila melanogaster* and vertebrate Pumilio; FBF diverges from this group[8]. Two proteins, FBF-1 and FBF-2, are 91% identical at the amino acid level and have overlapping but distinct functions; they are collectively referred to as FBF[14–17]. Both FBF proteins have eight putative RNA-recognition helices in the typical arrangement and contain the key conserved residues required for RNA binding by Pumilio (**Supplementary Fig. 1** online)[8–11].

To compare the binding specificity of PUF-8 to those of FBF-1 and FBF-2, we first used the yeast three-hybrid system (**Fig. 1b**)[18]. In this system, the interaction of FBF with an RNA in a yeast cell leads to activation of the reporter genes *HIS3* and *LacZ*. The level of expression of the reporter genes is directly related to the affinity of the interaction[19]. We introduced plasmids encoding PUF–activation domain (PUF-AD) fusion proteins and their target RNA sequences and monitored expression of *HIS3* (**Fig. 1c**). Cells carrying each combination of RNA and protein grew equally well on medium that selected only for the presence of the plasmids and not for expression of *HIS3* (**Fig. 1c**, +His). In contrast, on medium that selected for *HIS3* expression, only those cells that contained a cognate combination of PUF protein and RNA site grew (**Fig 1c**, –His). We conclude that FBF bound its cognate RNA site, the FBF-binding element (FBE), but not the Nanos response element (NRE); conversely, PUF-8 bound the NRE but not the FBE (**Fig. 1c**). Several other RNA targets of FBF, including FBEs in *gld-1, fem-3, fog-1* and *fbf* mRNAs, bound FBF but not PUF-8 (data not shown); similarly, many RNAs that bound PUF-8 did not bind FBF (data not shown)[14–17].

Department of Biochemistry, 433 Babcock Drive, University of Wisconsin, Madison, Wisconsin 53706, USA. Correspondence should be addressed to M.W. (wickens@biochem.wisc.edu).

Received 2 August; accepted 28 September; published online 23 October 2005; doi:10.1038/nsmb1010

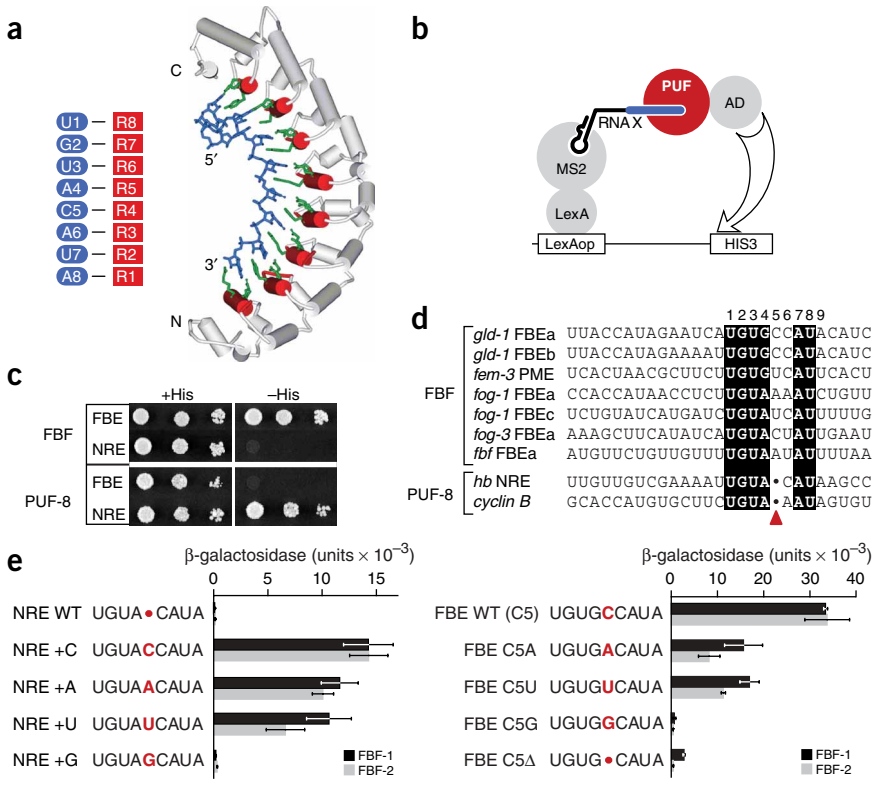


Figure 1 A single nucleotide is a key determinant of the RNA-binding specificity of two PUF proteins. (**a**) Diagram (left) and structure (right, PDB accession code 1M8Y) of the binding site in the human Pumilio1–RNA complex. Eight RNA-recognition helices (R1–R8, red) contain conserved residues (green) that specifically bind the NRE RNA (U1–A8, blue). The diagram at left depicts hydrogen bonding interactions between a base and two residues in an α-helix as a single black line but does not depict stacking interactions or the foldback of nucleotides –1 and –2 evident at the top of the crystal structure. (**b**) The yeast three-hybrid system. A fusion of the LexA DNA binding domain and MS2 coat protein tethers a hybrid RNA upstream of the reporter *HIS3* gene. The *HIS3* gene is transcribed when the PUF-AD fusion protein binds RNA X. (**c**) Results of three-hybrid experiments, showing that FBF-2 interacts specifically with the FBE and PUF-8 with the NRE. Yeast strains were spotted on medium with or without histidine, selecting for the relevant plasmids by omitting leucine and uracil. Cells were spotted in ten-fold serial dilutions, beginning with 10,000 cells on the left. (**d**) Alignment of *C. elegans* 3′ UTRs bound by FBF[15–17] and *Drosophila* 3′ UTRs bound by Pumilio[37–39]. *hb* denotes *hunchback*, which contains the NRE[37,38]. PUF-8 binds the NRE (this paper) and *cyclin B* binding is inferred. The second U of the NRE was changed to C in the three-hybrid system to avoid termination by RNA polymerase III. Black boxes mark the conserved UGUR and AU. (**e**) Activation of a *LacZ* reporter gene in the three-hybrid system (quantified by assaying β-galactosidase activity[19]), showing that binding specificity is switched by a single-nucleotide change in the NRE or FBE. Each bar represents an average from three experiments, with s.d. shown. The RNA sequences are listed, with the varied nucleotide in red.

The key role of a single nucleotide

Sequence alignments of the RNAs used in the binding specificity assay (**Fig. 1c**) and of other RNAs that bind FBF or PUF-8 revealed that the FBF-binding sites and PUF-8–binding sites all contain a UGUR tetranucleotide (R denotes a purine), as is common among all RNAs that bind a PUF protein specifically. Outside the UGUR, conservation is not immediately apparent. However, if the RNAs are aligned with a single-nucleotide gap in the PUF-8 sites, an AU dinucleotide is positioned equivalently in all the targets (**Fig. 1d**, highlighted regions). The sequence alignment suggests that FBF might require an 'extra' nucleotide relative to PUF-8, between UGUR and AU (**Fig. 1d**, red arrowhead).

To test this idea, we first analyzed a series of RNAs in which C, A, U or G was inserted into position 5 of the NRE (**Fig. 1e**). To quantify activation of the *LacZ* reporter gene, we assayed β-galactosidase activity. The NRE did not bind FBF (**Fig. 1c**,**e**). However, an NRE into which either C, A or U had been inserted bound FBF well (**Fig. 1e**). Insertion of G did not result in binding. Similarly, FBF bound its natural site, the FBE, whether the base at the equivalent position was C, A or U, but not if it was G. When the extra nucleotide (position +5) was deleted from the FBE, FBF binding was lost. FBF-1 and FBF-2 behaved identically.

The effect of the single nucleotide on FBF and PUF-8 binding was also observed *in vitro*, using purified recombinant proteins fused to glutathione *S*-transferase (GST) (**Supplementary Fig. 2** online) in an electrophoretic mobility shift assay (EMSA; **Fig. 2**). The insertion of a single nucleotide into a PUF-8–binding element enhanced the binding

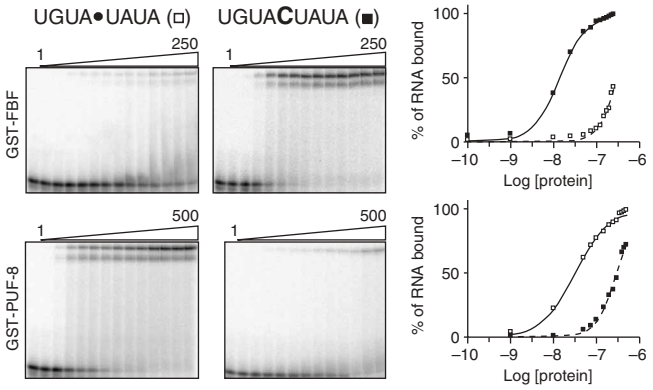


Figure 2 A single nucleotide determines RNA-binding specificity *in vitro*. Results of EMSAs with purified recombinant GST–FBF-2 and GST–PUF-8 (**Supplementary Fig. 2**). From left to right, GST–PUF-8 protein concentrations were 0, 1, 10, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450 and 500 nM; the 350 nM point was not included for the complex of GST–PUF-8 with UGUA**C**UAUA. From left to right, GST-FBF-2 protein concentrations were 0, 0.1, 1, 10, 25, 50, 75, 100, 125, 150, 175, 200, 225 and 250 nM. RNA sequences (derived from RNA 5 in **Fig. 3b**) are indicated with the varied nucleotide in bold, larger text. Two complexes of slightly different electrophoretic mobilities were observed with both proteins. All complexes were sequence-specific, as they required the UGU sequence (not shown).

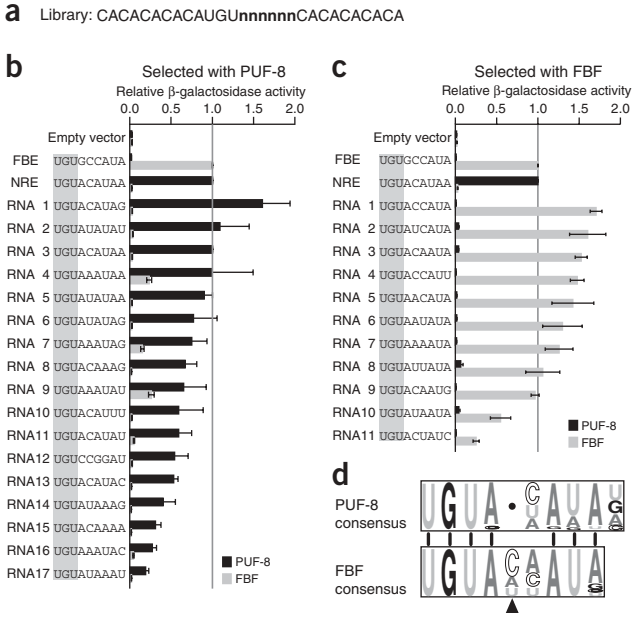
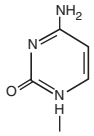
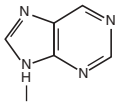
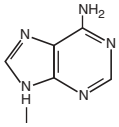
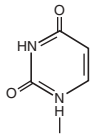
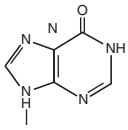
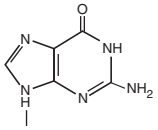


Figure 3 RNA selection in yeast reveals that FBF requires a single 'extra' nucleotide. (**a**) Shared sequence of the library of RNAs containing six randomized positions (bold n) prepared as described in ref. 13. RNAs interacting with PUF-8 were identified by activation of *HIS3* and *LacZ* genes (see Methods). (**b**) Binding of RNAs identified in the screen with PUF-8, as indicated by β-galactosidase activity in the three-hybrid assay with FBF-2–AD (gray, relative to FBE) and PUF-8–AD (black, relative to NRE)[19]. Each bar represents the average of three or four experiments, with s.d. shown. The sequence of each bound RNA in the critical region is indicated. (**c**) Binding of RNAs identified in the screen with FBF-1–AD, as in **b**. FBF-binding RNAs reported in ref. 13 bind only FBF-2 strongly. (**d**) Consensus sequences obtained in the screen, which differ for FBF-1 and PUF-8 by the presence or absence of a single nucleotide. Letter height is proportional to the percent occurrence at that position among all bound RNAs.

of purified recombinant GST-FBF 20-fold, from an apparent K_d of 300 ± 30 nM to 15 ± 5 nM (**Fig. 2**). GST–PUF-8 preferred an element without an extra nucleotide by ten-fold, showing K_ds with and without the nucleotide of 300 ± 15 nM and 30 ± 5 nM, respectively (**Fig. 2**). Similar results were obtained using other PUF-8–binding elements with or without the extra nucleotide (data not shown). Notably, the presence of a single nucleotide affected binding substantially even though the identity of the preferred base was flexible.

To compare the specificities of PUF-8 and FBF independently, we performed RNA selection experiments in yeast using the three-hybrid system (**Fig. 3**). We prepared a library of RNAs containing UGU followed by six randomized nucleotides; CA dinucleotide repeats flanked these core sequences (**Fig. 3a**). The library was cotransformed with plasmids encoding either PUF-8–AD or FBF-AD. RNAs that bound each PUF protein were identified by *HIS3* selection. Seventeen unique RNAs bound PUF-8 (**Fig. 3b**, black bars). None of the 17 bound FBF strongly; 3 of them (4, 7 and 9) bound FBF weakly (**Fig. 3b**, gray bars). Eleven unique RNAs were identified that bound FBF, as described in ref. 13 (**Fig. 3c**, gray bars). None of these RNAs bound PUF-8 appreciably (**Fig. 3c**, black bars).

Two consensus sequences were assembled from these data (**Fig. 3d**). UGUA and AUA were preferred by both proteins and were necessary for high-affinity binding, but the spacing between the two elements differed. FBF-binding sites contained an extra nucleotide relative to PUF-8 sites, located between the UGUA and AUA sequences. A, C or U was allowed in this 'extra' position, but G was excluded. These data emphasize that the presence or absence of the single base was the crucial determinant of RNA-binding specificity.

RNA		Apparent K_d (nM)	K_{rel}
Cytosine		25 ± 5	1.0
Purine ribonucleoside		65 ± 10	2.6
Adenine		95 ± 15	3.8
Uracil		95 ± 5	3.8
Abasic		100 ± 10	4.0
Inosine		130 ± 20	5.2
Guanine		390 ± 60	15.6

Exclusion of guanine

The fact that G was not allowed in the extra position (**Figs. 1** and **3**) might mean that G could not make base-specific contacts that are possible with A, C or U; alternatively, G could be excluded because it interfered with recognition of one or more bases and/or the ribose-phosphate backbone. To distinguish between these possibilities, we analyzed the binding of RNAs carrying distinct substitutions at position +5 using gel-shift assays (**Fig. 4**). The natural FBE in *gld-1* mRNA carries a cytosine at this position and yielded an apparent K_d of 25 nM with purified FBF. Substitution of position +5 with a purine ribonucleoside that lacked extracyclic groups or with adenine or uracil increased the K_d two- to four-fold. These data imply that the contribution of hydrogen bonding to binding energy is small. The absence of

Figure 4 Effect of substitutions at the 'extra' nucleotide position on binding. Chemically synthesized RNAs with the indicated bases substituted at position +5 (**Fig. 1d**) were analyzed in EMSAs with recombinant, purified GST–FBF-2. Each K_d is an average of two experiments. K_{rel} is the ratio of the K_d to that of the construct having +5C.

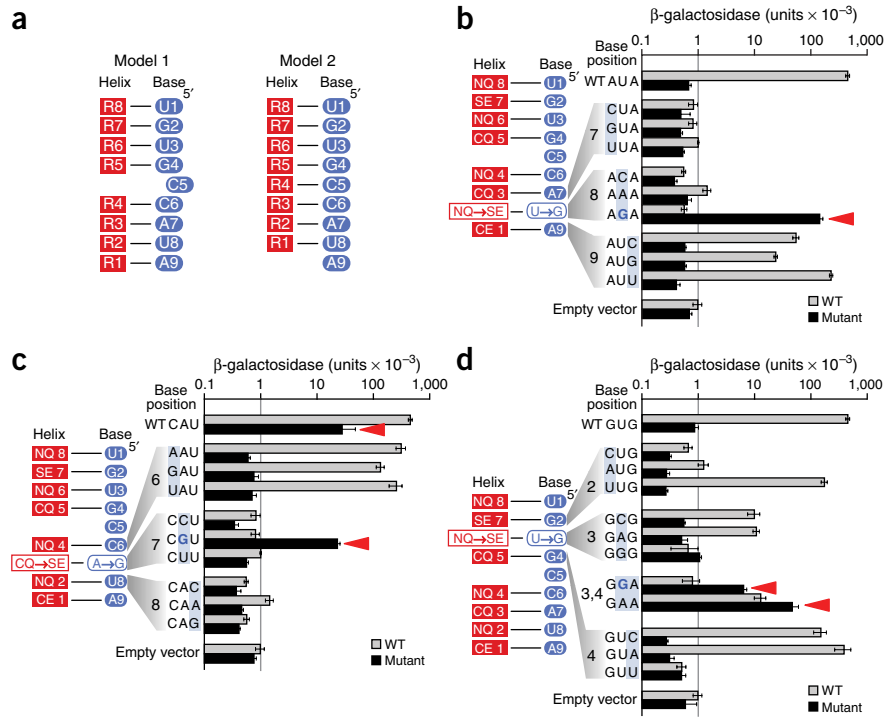


Figure 5 Alignment of PUF repeats and RNA. (**a**) Two models illustrating how the RNA-recognition helices of each repeat (R1–R8, red) could interact with the eight nucleotides in the core binding site (A9–U1, blue). Hydrogen bonding interactions are black lines; stacking interactions are not shown. (**b–d**) Binding of mutant repeats 2, 3 and 6. Left, schematic of binding, with compensatory protein and RNA mutations unshaded. Right, binding as measured by β-galactosidase activity in the three-hybrid assay for wild-type (FBF-2) and mutant proteins[19]. Key new interactions are indicated by red arrowheads. All four bases (three mutants and wild-type) were tested at each position, as indicated. Each bar represents the average of three experiments, with s.d. shown.

a base at position +5 also modestly decreased binding affinity, indicating that a base is not essential for FBF binding. Substitution with G at position +5 yielded an RNA that binds with 15-fold less affinity than C, corroborating the three-hybrid data. The binding of G was three-fold lower than that of inosine, implying that the extracyclic amino group of G interferes with binding. We infer that poor binding may be due in part to interference by G, though it is formally possible that the G-containing RNAs adopt a different structure. Regardless, it is notable that the effects of G substitution (15-fold) and deletion of the nucleotide (20-fold; **Fig. 2**) are comparable.

Alignment of RNA and protein

In the structure of human Pumilio bound to the NRE, three residues in each of eight successive repeats together recognize eight successive single bases[11]. Sequence comparisons suggest that PUF-8 and FBF contain these eight repeats, which are characteristic of the PUF protein family (**Supplementary Fig. 1**)[8,14]. The repeats of FBF could be aligned with its RNA binding site in either of two ways (**Fig. 5a**). In model 1, the extra nucleotide is a 'spacer,' and repeats on either side contact the RNA as they do in human Pumilio. In model 2, the eight repeats are aligned with eight consecutive bases, and the extra base is not recognized by the corresponding protein repeat.

To distinguish between these alternatives, we aligned the repeats in FBF to the RNA sequence empirically. To do so, we generated compensatory mutations between the protein and RNA (**Fig. 5b–d**). In the human Pumilio–RNA complex, three conserved residues recognize each base. Two are involved in hydrogen bonding and the other in stacking interactions. By switching these key residues from one α-helix to another, we could switch the base specificity of the helix, as had been done with one repeat of human Pumilio[11]. We then identified which base in the RNA had to be switched to bind the mutant protein. For example, in the RNA-recognition helix of repeat 7, serine and glutamate interact specifically with guanine. By substituting serine and glutamate at analogous positions in another repeat, we could determine which base that repeat recognized. We used this strategy to rationally design altered specificities in repeats 2, 3 and 6 of FBF and to align their RNA-recognition helices to the RNA (**Fig. 5b–d**).

To align the RNA-recognition helix of repeat 2 with the RNA, we changed Asn-Gln to Ser-Glu and then tested whether this substitution altered base specificity at positions 7, 8 or 9 of the RNA. We predicted that the mutant protein would require a G at either position 8 (model 1) or position 7 (model 2). The Asn-Gln to Ser-Glu mutant of repeat 2 (**Fig. 5b**) was tested with ten different RNAs, representing all four bases at positions 7, 8 and 9 (**Fig. 5b**; position 9 was included as a control). Wild-type FBF protein was analyzed in parallel. The wild-type protein requires an A at position 7 and a U at position 8, and it tolerates all four bases at position 9. In contrast, the mutant protein required a G at position 8 and did not bind the wild-type RNA sequence or any other mutant RNA (**Fig. 5b**). The mutant protein bound the G8 RNA nearly as well as the wild-type protein binds wild-type RNA. Thus, changing Asn-Gln to Ser-Glu in repeat 2 changes the specificity at nucleotide 8 from U to G. On the basis of these data, we position the RNA-recognition helix of repeat 2 opposite nucleotide 8.

Analogous results were obtained for repeat 3 (**Fig. 5c**). We replaced Cys-Gln in repeat 3 with Ser-Glu, which was predicted to change the specificity from A to G. Indeed, the protein carrying Ser-Glu in repeat 3 bound a G at position 7, whereas the wild-type protein did not. The mutant repeat also still bound A, as does the wild-type protein.

In repeat 6, Asn-Gln was changed to Ser-Glu, and this substitution was predicted to change the specificity from U to G (**Fig. 5d**). Indeed, the variant protein did not bind the wild-type sequence but did bind RNAs with a G at position 3. The next position, 4, also had to be changed to an A to observe this switch from U to G. The variant protein accepted an A at position 3 when followed by an A, as does the wild-type protein. The coupling of nucleotide-identity requirements at positions 3 and 4 may reflect an important role of the stacking residue of repeat 6 in determining RNA-binding specificity.

Together, these data can be used to align the RNA-recognition helices of three repeats to the RNA, and they strongly support the alignment depicted in model 1. The nucleotide at position 5 is a spacer whose base identity is flexible; RNA-recognition helices on either side bind the RNA in the same manner as they do in the human Pumilio-RNA cocrystal and presumably in PUF-8 (ref. 11).

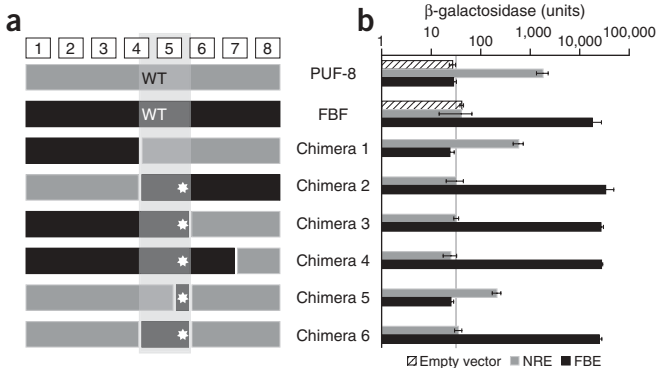


Figure 6 FBF RNA-binding specificity maps to a discrete, central region of the protein. (**a**) Design of chimeric constructs. Numbered boxes, positions of the eight PUF repeats; dark gray, PUF-8 segments; black, FBF-1 segments; WT, wild-type; asterisk, location of the 16-residue insertion in FBF-1; light gray overlay, the FBF-1 region sufficient to confer FBF-1's binding specificity. (**b**) Binding of each chimera to both the NRE (gray bars) and FBE (black bars). Empty vector was tested to establish a baseline for noninteraction (striped bars). Binding was measured by β-galactosidase activity in the three-hybrid system[19]. Each bar is the average of three experiments, with s.d. shown.

A spacer module in the protein

In light of the protein's modularity, we sought to determine whether a single discrete region of FBF imposed the requirement for an extra nucleotide. We prepared a series of chimeric proteins combining portions of FBF and PUF-8 (**Supplementary Fig. 3** online and Methods). A single region of FBF encompassing repeat 5 and flanking residues is responsible for the requirement (**Fig. 6**).

The most telling variant, chimera 6, has 74 residues of FBF grafted into PUF-8; this new protein switched specificity from that of PUF-8 to that of FBF. In those 74 residues, FBF uniquely contains a 16-residue insertion compared to other PUF proteins (**Supplementary Fig. 1**, yellow highlight). However, insertion of those 16 residues alone is not sufficient to switch specificity (**Fig. 6**, chimera 5).

DISCUSSION

We conclude that the different specificities of FBF and PUF-8 are due to the presence or absence, respectively, of a single nucleotide in their binding sites (**Fig. 7a**). Repeat 5 of FBF, plus flanking residues, are sufficient to confer this specificity. We propose that FBF's requirement for an extra nucleotide reflects the need to accommodate a distortion in the central region of FBF relative to PUF-8. The extra 16-residue segment present in FBF but not PUF-8 (or most other PUF proteins) is not on its own sufficient to switch specificity to that of FBF. However, it may be important in the context of the larger region whose insertion does switch specificity; the distortion in FBF may be an abnormal rotational or translational displacement of adjacent helices or repeats imposed by the latter, 74-residue region. Other members of the helical-repeat protein superfamily, such as the HEAT, armadillo and ankyrin repeats, pack on each other in a regular manner to generate superhelical structures[20]. Deviations from the consensus repeat sequence result in nonconsensus packing of helices, as exemplified by the armadillo repeats of β-catenin, the HEAT repeats of the PR65/A subunit of protein phosphatase 2A and the ankyrin repeats of inhibitor κBα (IκBα)[20–24]. These local disruptions of consensus packing are associated with ligand binding and changes in the protein's rigidity[20,21,23–25]. Analogous perturbations in FBF may be crucial for binding to its mRNA targets and multiple protein partners.

FBF directly binds several proteins, including NOS-3 (ref. 26), CPB-1 (ref. 27) and GLD-3 (ref. 28). Each of these three proteins is an RNA-binding protein in its own right. The identity of the spacer base could determine which partner joins an FBF–RNA complex, adding another layer of regulation *in vivo*. Analogous interactions have been observed in other RNA–protein complexes: for example, four nucleotides upstream of the core Pumilio-binding site determine whether the Nanos protein binds the Pumilio–RNA complex[29]. The binding of NusA to the N protein–RNA complex may be analogous, in that it depends on a base in a pentaloop that is not essential for the prior binding of N[30–32].

The localized-distortion model of FBF accommodates our finding that although the presence of a spacer nucleotide is crucial, the identity of its base is flexible. On the basis of sequence conservation, compensatory mutants and protein chimeras, we deduce that in both FBF and PUF-8, RNA-recognition helices of repeats 1–3 bind AUA and those of repeats 6–8 bind UGU. The identity of the spacer nucleotide is flexible, implying that the protein accepts a variety of different combinations of stacking interactions and potential hydrogen bond donors and acceptors from the base; however, some combinations are more acceptable than others. We suggest the spacer nucleotide is needed for FBF to make contacts from multiple other repeats in the one-helix/one-nucleotide mode seen with human Pumilio1 (ref. 11). The affected contacts may be nearby, in the RNA-recognition helix of the nearest repeat, for example. Alternatively, the spacer could be required to align the RNA properly at multiple repeats. In an extreme scenario, the spacer could position the 5′ and 3′ halves of the RNA to interact simultaneously with the two halves of FBF (**Fig. 7b**).

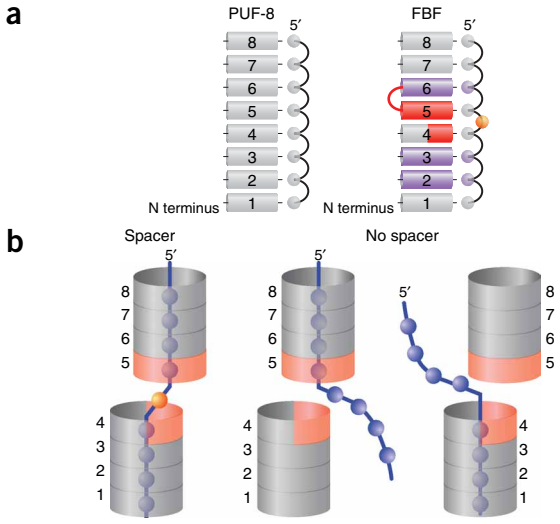


Figure 7 Model of FBF and PUF-8 interaction with RNA. (**a**) Protein-RNA contacts. Cylinders, RNA-recognition α-helices; balls, bases; red, the region of FBF sufficient to confer RNA-binding specificity; red loop, the 16-residue insertion; purple, the three repeats (2, 3 and 6) mapped to specific nucleotides by compensatory mutation studies (**Fig. 5**); gold ball, the spacer nucleotide. (**b**) A local distortion in FBF imposes the requirement for a spacer nucleotide. The diagram depicts the extreme model of the consequences of the local distortion in FBF, in which multiple nucleotides cannot contact RNA-recognition helices in a spacerless RNA. Coloration is as in **a**, except that RNA is blue.

Without the spacer, only repeats 1–4 or only 5–8 might interact with RNA, but the two groups of repeats might not be positioned to interact simultaneously with a spacerless RNA. Mechanisms for recognition of DNA by certain proteins may be instructive: for example, phage 434 repressor requires central noncontacted base pairs within the operator to position two monomers appropriately[33,34]. Direct structural analysis of FBF is needed to test this model, but this has been hampered by difficulties in obtaining concentrated, soluble protein.

The role of localized distortions and spacer nucleotides in determining PUF specificity may be common. For example, in *Saccharomyces cerevisiae*, the preferred binding sites of Puf3 (ref. 12,35) and those inferred for Puf4 and Puf5 (refs. 12,36) can be explained by allowing spacer nucleotides at two different positions while maintaining the same amino acid contacts with UGU and AUA. Effectively, yeast Puf3, 4 and 5 may identify distinct and/or overlapping targets by demanding 0, 1 or 2 spacer nucleotides in different positions of the RNA-binding element. Studies of compensatory mutants between the RNAs and proteins should test this model decisively.

Our data support and extend the earlier observations of ref. 11 and emphasize the extraordinary modularity of PUF proteins and PUF–RNA complexes. That modularity, implied by sequence alignments[14,37] and evident in the structures of free and RNA-bound Pumilio[9–11], here is shown to be crucial in determining the specificities of PUF proteins for RNA. Moreover, variant proteins with predictably altered specificity were created by rationally designed alterations (this paper and ref. 11). Insertion of a relatively short region of FBF into PUF-8 imposed the requirement for a spacer nucleotide. We suggest that new proteins based on the PUF scaffold, with new and diverse specificities, could be designed, selected and exploited *in vivo* to target new RNAs.

METHODS

Yeast three-hybrid assays. Three-hybrid assays were performed as described in the yeast strain YBZ-1 (refs. 18,19). Gal4–activation domain fusion proteins were expressed from pACT and pACT2 plasmids. For PUF-8 (**Fig. 3**), ~900,000 transformants were screened on medium lacking histidine and containing 1 mM 3-aminotriazole. From these, 32 unique clones were recovered and 17 bound PUF-8. All bound RNAs were recovered in *Escherichia coli* and reintroduced into naive yeast to confirm the interaction and specificity. Screen parameters for FBF-1 are described in ref. 13.

DNA constructs used in the three-hybrid system. The FBF and PUF-8 truncations we analyzed contained the eight PUF repeats and flanking residues; this corresponds to residues 121–614 of FBF-1 (GenBank accession number NM_062815), 121–632 of FBF-2 (NM_062819) and 127–519 of PUF-8 (NM_063122). The FBF-2 site-directed mutants (**Fig. 5**) were created by the QuikChange method (Stratagene). Amino acid changes were as follows: for mutant repeat 2, N244S and Q248E; for mutant repeat 3, C287S and Q291E; for mutant repeat 6, N415S and Q419E. The portions of FBF-1 and PUF-8 used in the chimeras (**Fig. 6**) were as follows: chimera 1, residues 121–325 of FBF-1 and 305–519 of PUF-8; chimera 2, 326–614 of FBF-1 and 127–304 of PUF-8; chimera 3, 121–399 of FBF-1 and 364–519 of PUF-8; chimera 4, 121–453 of FBF-1 and 416–519 of PUF-8; chimera 5, 371–399 of FBF-1 and 127–350 and 364–519 of PUF-8; chimera 6, 326–399 of FBF-1 and 127–304 and 364–519 of PUF-8. For each chimera, the fragments were blunt-end ligated together and then were cloned into the EcoRI and NcoI sites of pACT2. DNA oligonucleotides designed to express various RNA sequences were cloned into the XmaI and SphI sites of pIIIa MS2-2.

Recombinant PUF proteins. FBF-2 (residues 121–632) and PUF-8 (127–519) GST fusion proteins were purified as described in ref. 13. Protein was eluted from the glutathione-Sepharose (Amersham) without cleavage and dialyzed as described in ref. 13. Protein purity was estimated by SDS gel electrophoresis (>90% purity) and concentration measured by a Bradford assay.

Electrophoretic mobility shift assays. We combined 100 fmol ^{32}P 5′ end–labeled RNA oligonucleotides (IDT and Dharmacon) with a range of protein concentrations in 10 mM HEPES (pH 7.4), 1 mM EDTA, 50 mM KCl, 2 mM DTT, 0.1 mg ml^{-1} BSA, 0.02% (v/v) Tween-20 and 0.2 mg ml^{-1} yeast total RNA (Sigma) for 30 min at room temperature. Loading dye (2.5 µl of 10% (v/v) Ficoll 400 K, 5% (v/v) DMSO) was added to each 10 µl reaction immediately before loading on a prerun nondenaturing polyacrylamide gel (6% (w/v) 29:1 acrylamide/bisacrylamide, 0.5× TBE). Gels were run at 200 V at 4 °C for 3–5 h. The fraction of bound RNA was determined using ImageQuant (Amersham). The apparent K_d, the concentration of protein at which half-maximal binding occurs, was calculated using GraphPad Prism 4.

Accession codes. BIND identifiers (http://bind.ca): 335741, 335742, 335743, 335744, 335745, 335746, 335747, 335748, 335749 and 335750.

Note: Supplementary information is available on the Nature Structural & Molecular Biology website.

ACKNOWLEDGMENTS

We thank members of the Kimble and Wickens laboratories for their suggestions, advice and comments on the manuscript, and we appreciate the suggestions of J. Kimble and S. Butcher. Figures were prepared by the University of Wisconsin Biochemistry Media Center. This work was supported by US National Institutes of Health grants to M.W. L.O. was supported by a US National Institutes of Health Molecular Biology Training Grant.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Published online at http://www.nature.com/nsmb/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Sonenberg, N., Hershey, J.W.B. & Mathews, M. (eds.) *Translational Control* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA, 1996).
2. Kuersten, S. & Goodwin, E.B. The power of the 3′ UTR: translational control and development. *Nat. Rev. Genet.* **4**, 626–637 (2003).
3. Sachs, A.B., Sarnow, P. & Hentze, M.W. Starting at the beginning, middle, and end: translation initiation in eukaryotes. *Cell* **89**, 831–838 (1997).
4. van Hoof, A. & Parker, R. Messenger RNA degradation: beginning at the end. *Curr. Biol.* **12**, R285–R287 (2002).
5. Lim, L.P. *et al.* Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* **433**, 769–773 (2005).
6. Grabowski, P.J. & Black, D.L. Alternative RNA splicing in the nervous system. *Prog. Neurobiol.* **65**, 289–308 (2001).
7. Wickens, M., Kimble, J. & Strickland, S. Translational control of developmental decisions. in *Translational Control* (eds. Sonenberg, N., Hershey, J.W.B. & Mathews, M.) 411–450 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA, 1996).
8. Wickens, M., Bernstein, D.S., Kimble, J. & Parker, R.A. PUF family portrait: 3′UTR regulation as a way of life. *Trends Genet.* **18**, 150–157 (2002).
9. Edwards, T.A., Pyle, S.E., Wharton, R.P. & Aggarwal, A.K. Structure of Pumilio reveals similarity between RNA and peptide binding motifs. *Cell* **105**, 281–289 (2001).
10. Wang, X., Zamore, P.D. & Hall, T.M. Crystal structure of a Pumilio homology domain. *Mol. Cell* **7**, 855–865 (2001).
11. Wang, X., McLachlan, J., Zamore, P.D. & Hall, T.M. Modular recognition of RNA by a human pumilio-homology domain. *Cell* **110**, 501–512 (2002).
12. Gerber, A.P., Herschlag, D. & Brown, P.O. Extensive association of functionally and cytotopically related mRNAs with Puf family RNA-binding proteins in yeast. *PLoS Biol.* **2**, E79 (2004).
13. Bernstein, D., Hook, B., Hajarnavis, A., Opperman, L. & Wickens, M. Binding specificity and mRNA targets of a *C. elegans* PUF protein, FBF-1. *RNA* **11**, 447–458 (2005).
14. Zhang, B. *et al.* A conserved RNA-binding protein that regulates sexual fates in the *C. elegans* hermaphrodite germ line. *Nature* **390**, 477–484 (1997).
15. Crittenden, S.L. *et al.* A conserved RNA-binding protein controls germline stem cells in *Caenorhabditis elegans*. *Nature* **417**, 660–663 (2002).
16. Lamont, L.B., Crittenden, S.L., Bernstein, D., Wickens, M. & Kimble, J. FBF-1 and FBF-2 regulate the size of the mitotic region in the *C. elegans* germline. *Dev. Cell* **7**, 697–707 (2004).
17. Thompson, B.E. *et al.* Dose-dependent control of proliferation and sperm specification by FOG-1/CPEB. *Development* **132**, 3471–3481 (2005).
18. Bernstein, D.S., Buter, N., Stumpf, C. & Wickens, M. Analyzing mRNA-protein complexes using a yeast three-hybrid system. *Methods* **26**, 123–141 (2002).
19. Hook, B., Bernstein, D., Zhang, B. & Wickens, M. RNA-protein interactions in the yeast three-hybrid system: affinity, sensitivity, and enhanced library screening. *RNA* **11**, 227–233 (2005).
20. Groves, M.R. & Barford, D. Topological characteristics of helical repeat proteins. *Curr. Opin. Struct. Biol.* **9**, 383–389 (1999).

21. Huber, A.H., Nelson, W.J. & Weis, W.I. Three-dimensional structure of the armadillo repeat region of beta-catenin. *Cell* **90**, 871–882 (1997).
22. Groves, M.R., Hanlon, N., Turowski, P., Hemmings, B.A. & Barford, D. The structure of the protein phosphatase 2A PR65/A subunit reveals the conformation of its 15 tandemly repeated HEAT motifs. *Cell* **96**, 99–110 (1999).
23. Huxford, T., Huang, D.B., Malek, S. & Ghosh, G. The crystal structure of the IkappaBalpha/NF-kappaB complex reveals mechanisms of NF-kappaB inactivation. *Cell* **95**, 759–770 (1998).
24. Jacobs, M.D. & Harrison, S.C. Structure of an IkappaBalpha/NF-kappaB complex. *Cell* **95**, 749–758 (1998).
25. Graham, T.A., Weaver, C., Mao, F., Kimelman, D. & Xu, W. Crystal structure of a beta-catenin/Tcf complex. *Cell* **103**, 885–896 (2000).
26. Kraemer, B. *et al.* NANOS-3 and FBF proteins physically interact to control the sperm-oocyte switch in *Caenorhabditis elegans*. *Curr. Biol.* **9**, 1009–1018 (1999).
27. Luitjens, C., Gallegos, M., Kraemer, B., Kimble, J. & Wickens, M. CPEB proteins control two key steps in spermatogenesis in *C. elegans*. *Genes Dev.* **14**, 2596–2609 (2000).
28. Eckmann, C.R., Kraemer, B., Wickens, M. & Kimble, J. GLD-3, a bicaudal-C homolog that inhibits FBF to control germline sex determination in *C. elegans*. *Dev. Cell* **3**, 697–710 (2002).
29. Sonoda, J. & Wharton, R.P. Recruitment of Nanos to hunchback mRNA by Pumilio. *Genes Dev.* **13**, 2704–2712 (1999).
30. Chattopadhyay, S., Garcia-Mena, J., DeVito, J., Wolska, K. & Das, A. Bipartite function of a small RNA hairpin in transcription antitermination in bacteriophage lambda. *Proc. Natl. Acad. Sci. USA* **92**, 4061–4065 (1995).
31. Mogridge, J., Mah, T.F. & Greenblatt, J. A protein-RNA interaction network facilitates the template-independent cooperative assembly on RNA polymerase of a stable antitermination complex containing the lambda N protein. *Genes Dev.* **9**, 2831–2845 (1995).
32. Legault, P., Li, J., Mogridge, J., Kay, L.E. & Greenblatt, J. NMR structure of the bacteriophage lambda N peptide/boxB RNA complex: recognition of a GNRA fold by an arginine-rich motif. *Cell* **93**, 289–299 (1998).
33. Koudelka, G.B., Harrison, S.C. & Ptashne, M. Effect of non-contacted bases on the affinity of 434 operator for 434 repressor and Cro. *Nature* **326**, 886–888 (1987).
34. Koudelka, G.B., Harbury, P., Harrison, S.C. & Ptashne, M. DNA twisting and the affinity of bacteriophage 434 operator for bacteriophage 434 repressor. *Proc. Natl. Acad. Sci. USA* **85**, 4633–4637 (1988).
35. Olivas, W. & Parker, R. The Puf3 protein is a transcript-specific regulator of mRNA degradation in yeast. *EMBO J.* **19**, 6602–6611 (2000).
36. Tadauchi, T., Matsumoto, K., Herskowitz, I. & Irie, K. Post-transcriptional regulation through the HO 3′-UTR by Mpt5, a yeast homolog of Pumilio and FBF. *EMBO J.* **20**, 552–561 (2001).
37. Zamore, P.D., Williamson, J.R. & Lehmann, R. The Pumilio protein binds RNA through a conserved domain that defines a new class of RNA-binding proteins. *RNA* **3**, 1421–1433 (1997).
38. Wharton, R.P. & Struhl, G. RNA regulatory elements mediate control of *Drosophila* body pattern by the posterior morphogen nanos. *Cell* **67**, 955–967 (1991).
39. Sonoda, J. & Wharton, R.P. *Drosophila* Brain Tumor is a translational repressor. *Genes Dev.* **15**, 762–773 (2001).