Citation
Wallis, C.P. and Richman, T.R. and Filipovska, A. and Rackham, O. 2018. Tighter Ligand Binding Can Compensate for Impaired Stability of an RNA-Binding Protein. ACS Chemical Biology. 13 (6): pp. 1499-1505. http://doi.org/10.1021/acschembio.8b00424

1 **Content Type:** Letter

2

3 Tighter ligand binding can compensate for impaired stability 4 of an RNA-binding protein

5

6 Christopher P. Wallis[a], Tara R. Richman[a], Aleksandra Filipovska[a,b], Oliver Rackham[a,b,1]

7

8

9 [a] Harry Perkins Institute of Medical Research and Centre for Medical Research, The University of
10 Western Australia, Nedlands 6009, Australia.
11 [b] School of Molecular and Chemical Sciences, The University of Western Australia, Crawley 6009,
12 Australia.

13

14 [1] To whom correspondence may be addressed. Email: oliver.rackham@uwa.edu.au.

15

16 **Key words:** RNA-binding proteins; protein stability; synthetic biology; RNA-protein interactions;
17 protein engineering; protein epistasis

18 **ABSTRACT:**

19 It has been widely shown that ligand-binding residues, by virtue of their orientation, charge and
20 solvent exposure, often have a net destabilising effect on proteins that is offset by stability conferring
21 residues elsewhere in the protein. This structure-function trade-off can constrain possible adaptive
22 evolutionary changes of function and may hamper protein engineering efforts to design proteins with
23 new functions. Here, we present evidence from a large randomised mutant library screen that in the
24 case of PUF RNA-binding proteins, this structural relationship may be inverted and that active-site
25 mutations that increase protein activity are also able to compensate for impaired stability. We show
26 that certain mutations in RNA-protein binding residues are not necessarily destabilising and that
27 increased ligand-binding can rescue an insoluble, unstable PUF protein. We hypothesise that these
28 mutations re-stabilise the protein via thermodynamic coupling of protein folding and RNA binding.

29 Protein stability is a crucial biophysical property that dictates the concentration of functional, folded
30 protein,[1] a protein's propensity for aggregation,[2, 3] and its half-life which in turn effects its regulation.[4]
31 These properties are therefore important factors in disease,[5, 6] evolutionary fitness,[7, 8] and medical and
32 industrial biotechnology.[9] Proteins typically exhibit low general thermostability and must maintain
33 this in order to function. Proteins with lower stability are more prone to misfolding, aggregation and
34 degradation,[10] whereas proteins with excess stability tend to be more rigid, have lowered activity, and
35 can be misregulated.[11] The stability range of a protein is partly defined by interactions between core
36 residues, which tend to be hydrophobic and stability conferring, and surface and active-site residues,
37 which tend to be charged or polar, and destabilising.[12, 13] Therefore, there is an internal trade-off
38 between stability conferring, core residues, and destabilising active-site residues. In evolutionary
39 terms, the trade-off implies that change of function mutations in active-site residues are destabilising
40 and therefore must be compensated for by additional mutations in stabilising core residues.[14]
41 Consequently, stability or stabilising epistasis, is a key constraint on the evolution of new function.[15
42 16] This relationship between protein stability and function has been simulated computationally in
43 globular proteins and appears to hold over a range of sizes, folds and functions, including enzymes
44 and ligand-binding proteins.[17] However, experimental evidence is still required to confirm this for
45 many proteins including those with unusual structures such as the sequence-specific RNA-binding
46 proteins (RBPs) of the Pumilio/FBF (PUF) protein family.[18]

47 PUF proteins are an important family of RNA-binding proteins involved in controlling the
48 spatial, temporal and functional activities of diverse RNA populations.[19 20] PUF proteins bind specific
49 sequences embedded in the 5′- and 3′-UTR regions of target mRNAs in a modular fashion by virtue
50 of an array of consecutive repeats that each bind a defined nucleotide.[18] It is currently unclear how
51 the modular, repeated structure of PUF affects their robustness to mutation, or to what extent the
52 stability-function dynamic discussed above constrains the adaptability of PUF proteins. Here, we
53 explore the functional-stability constraints on PUF proteins by examining mutations that rescue the
54 function of an insoluble, unstable PUF variant. Our work suggests that, unlike most other ligand-
55 binding and nucleic acid-binding proteins, PUF proteins can be restabilised by active-site mutations
56 that increase RNA-binding activity by associating with their RNA targets.

57 To explore which PUF residues are required for protein stability and how compensatory
58 mutations effect stability and ligand binding we constructed an unstable PUF protein mutant. This
59 variant, that we name PUFmut, was created by substituting residues F882, F905 and Q913 in the PUF
60 domain from the human PUM1 protein with glycine (Figure 1A,B). Loss of these large buttressing
61 residues introduces a cavity that causes loss of function and stability (Figure 1C,D). A library of
62 randomised PUFmut mutants was then created using error-prone PCR, which was optimised to
63 introduce approximately 2-4 mutations per open reading frame (ORF). The randomised library (2.2

64 x 10^6 members) was then transformed into *Saccharomyces cerevisiae* strain Mav204K along with a
65 plasmid expressing its cognate Nanos response element (NRE) RNA sequence (5′-UGUAUAUA-3′).
66 The PUFmut yeast library was then plated on media lacking histidine to select for functional RNA-
67 protein interactions that activate the yeast three-hybrid reporter (Figure 1E).

68 We analysed 157 colonies that survived the selection by sequencing and restriction digestion
69 to identify mutations that restored PUF stability and function (Supporting Table S1). Information on
70 mutations was collated and mapped onto the PUF crystal structure (Figure 2A).[21] The identified
71 mutants from the screen could be predominantly separated in two classes: (i) those that likely repaired
72 the cavity formed by the PUFmut mutations and (ii) those that were distant from this pocket in the
73 PUF structure. Of particular interest, a single mutation, N1080Y that stacks between U6 and G7 bases
74 in the NRE RNA, dominated the selected clones and was present in the majority of the surviving
75 colonies. Given the unexpected predominance of mutations in residues close to the RNA-protein
76 interface, we choose those ≤5 Å from the RNA ligand for further analysis (Figure 2B-E).

77 Many of the selected mutants contained additional mutations in their ORFs, therefore to
78 investigate the effect of individual mutations in isolation, each mutation classified as being in close
79 proximity to RNA was then separately introduced into both the wild-type (PUFwt) and PUFmut
80 backgrounds using site-directed mutagenesis. Yeast three-hybrid survival assays were then performed
81 using each site-directed mutant clone (Figure 3). We used PUFwt, EGFP and PUFmut ORFs as
82 controls. PUFwt, which binds NRE RNA showed no growth defects at all 3-AT concentrations,
83 PUFmut and EGFP showed highly restricted, or no growth at all 3-AT concentrations. The six RNA-
84 protein interface mutations when introduced into the wild-type background (H861P, F897Y, Q931H,
85 M932K, C1007S and Q1040R) were comparable to PUFwt except F897L, which showed marginally
86 less growth at all 3-AT concentrations, and both K976R and N1080Y which showed slightly higher
87 growth at higher (≥0.5 mM) 3-AT concentrations. Increased growth under these conditions suggests
88 a lower Kd for these PUF variants and the NRE RNA compared to PUFwt. In the PUFmut background
89 four mutants, F897Y, Q931H, M932K and N1080Y have growth profiles comparable to PUFwt, and
90 are therefore able to rescue the PUFmut. All other mutants show significant growth defects at all 3-
91 AT concentrations, including K976R, which showed a slight growth advantage in a PUFwt genetic
92 background. This finding indicates that many mutations at the RNA-protein interface are able to
93 rescue the PUFmut but not all, and those that cannot rescue the PUF protein might have been selected
94 due to their effects when present in combination with other library mutations.

95 To confirm that compensatory PUFmut mutations were in fact restoring protein stability as
96 well as function, and were not an artefact of the yeast three-hybrid system, we performed
97 immunoblotting on myc-tagged versions of PUFwt, PUFmut, PUFmut-F897Y, PUFmut-Q931H,
98 PUFmut-M932K and PUFmut-N1080Y (which showed growth in PUFmut genetic backgrounds)

99 (Figure 4A). The protein abundance of all four of the tested library mutants, F897Y, Q931H, M932K
100 and N1080Y was increased compared to the PUFmut. F897Y had comparable protein levels to the
101 PUFwt, whereas Q931H, M932K and N1080Y had increased protein abundance. This suggests that
102 these individual residue changes were sufficient to rescue the PUFmut levels to those of the PUFwt
103 protein. Because these mutations were found at the RNA-protein interface it could be possible that
104 binding to RNA within the yeast cells stabilised them or, alternatively, that the mutation stabilised
105 the proteins independent of RNA-binding. To test this we purified the PUFmut-N1080Y mutant
106 protein and tested its thermostability (Supporting Figure S1). We found that this protein was more
107 stable in the presence of its RNA target, indicating that the stabilising effect of this mutation relies
108 on the protein's association with RNA.

109 As well as showing that PUFmut mutants containing active-site, RNA-proximal mutations
110 rescue both RNA-binding activity and protein stability, we wanted to know if these specific mutations
111 increased RNA-binding affinity compared to PUFwt and if these changes correlated with the observed
112 increase in protein stability. To quantify changes in RNA-binding affinity we performed RNA
113 electrophoretic mobility shift assays (REMSAs) on purified library mutants in the PUFwt background
114 (Figure 4B). All purified PUFs bound strongly to the NRE RNA, with M932K, N1080Y, Q931H and
115 F897Y mutants binding more strongly than PUFwt. Our results from the binding assays correspond
116 well with the yeast growth assay, which are known to correlate with Kd.[24] Furthermore, using yeast
117 three-hybrid analyses and REMSAs using RNA targets with single nucleotide mismatches to the PUF
118 binding site we observed that these mutant proteins did not sacrifice their specific RNA recognition
119 to achieve higher affinity binding (Supporting Figure S2).

120 Taken together, our data show that the activity of the impaired, insoluble and inactive PUFmut
121 protein can be rescued by certain mutations in the RNA-protein interface. Although N1080Y has been
122 reported before as increasing binding affinity[23] to our knowledge, this is the first time this mutation
123 and others (F897Y, Q931H and M932K) have been shown to also be linked to changes in protein
124 stability. Residues with both stabilising and catalytic functions have been reported for the bovine
125 pancreatic ribonuclease A (RNase A) enzyme, which also interacts with RNA. Here, Asp121 forms
126 a catalytic dyad via a salt-bridge with His119, which is required for both RNA cleavage and overall
127 conformational stability.[25] Likewise surface-exposed catalytic phenylalanine residues in the
128 RNP1/RNP2 motif of a *Bacillus subtilis* CspB cold-shock protein where also shown to be essential
129 for thermostability.[26] However in these examples it isn't clear to what degree catalysis and stability
130 trade-off with regards to changes in these residues, the studies simply showed they are required for
131 both. Our work shows dual roles for some residues, demonstrating that there is not always a structural
132 and evolutionary trade-off between stability conferring residues and catalytic residues. The trade-off
133 hypothesis[7, 27] suggests that changes to catalytic or active-site residues that result in increased or

134 altered activity must lower net stability unless compensated for by stability conferring changes
135 elsewhere in the protein. This idea has been widely explored both experimentally and computationally
136 in a variety of globular proteins and enzymes. These studies suggest that the trade-off holds over a
137 range of sizes, folds and functions.[17] For example, active-site mutations that increase the catalytic
138 ability of antibiotics like TEM-1-lactamase are generally destabilising.[28] Conversely, active-site
139 mutations of AmpC β-lactamase that increase stability are concomitant with decreased activity.[29] This
140 has also been observed in other proteins such as Barnase,[30] Barstar,[31] and T4 lysozymes.[32] Increases
141 in evolvability observed in TEM-1-lactamase and P450 have likewise been explained by additional
142 strongly stabilising mutations (called global suppressors) that have a buffering effect on change of
143 function mutations and thus accommodate losses in stability.[33, 34]

144 Our work shows that, contrary to the examples listed above, PUF proteins can evolve via
145 mutations that both increase stability and increase function. Indeed, it has been shown that new-
146 function mutations are not necessarily more destabilising than the average mutation, but only more
147 destabilising than neutral mutations.[17] In addition, it should be noted that in this case the wild-type
148 activity of the PUF protein was regained, not the acquisition of a new function. The exact mechanism
149 by which the mutations at the PUF-RNA interface restore activity is currently unclear, but it may be
150 connected to the intimate relationship PUF proteins have with RNA. We suggest that RNA has a
151 buffering effect on PUF proteins that expands the acceptable ΔG stability range and thereby allows
152 them to tolerate function-altering mutations and therefore be adaptable without compromising
153 stability. Work on the mutational buffering effects of chaperone proteins provides an analogous
154 model. Here, over expression of various chaperones increased the tolerance of target proteins to
155 destabilising non-synonymous mutations, increasing their evolvability for new functions.[35] [36] [37]
156 Interestingly, these authors also showed that the degree of mutational buffering was related to the
157 strength of the binding interaction between the chaperone and its target. Increased RNA-protein
158 affinity shown in Q931H, M932K and N1080Y PUF variants may have a similar effect. These
159 mutations may increase binding of partially folded PUF intermediates to RNA, with the RNA acting
160 as a kind of molecular scaffold which may orientate certain key PUF residues more favourably or
161 stabilise intermediate states allowing a greater proportion of protein to properly fold. Alternatively,
162 the RNA scaffolding may provide a new, lower energy folding pathway, a pathway unavailable to
163 PUF proteins that bind RNA with lower affinity.

164 Work on the overall increased thermostability afforded to proteins bound to nucleic acids also
165 supports the notion that increased binding in PUF proteins is restabilising rather than destabilising.
166 For example, RNA aptamers inhibited the unfolding of bound tetanus toxoid proteins exposed to
167 water.[38] Also, rRNA has been linked to improved protein folding, for example, lactate dehydrogenase,
168 carbonic anhydrase lysozyme and malate dehydrogenase all showed improved folding in the presence

169 of the V domain of 23S rRNA.[39, 40 41 42] Heat denatured EcoRI and luciferase have also been shown
170 to increase refolding yields when incubated with mitochondrial 12S and 16S rRNA.[43] Finally, a
171 variety of RNA and DNA molecules of various lengths have been shown to function as efficient
172 chaperones *in vitro*, suppressing the aggregation of citrate synthase more than 300-fold better than
173 GroEL and cooperating with other proteins to promote protein refolding.[44]

174 Interestingly, when we examined natural PUF family sequences we found that the RNA
175 affinity-increasing mutations identified in our laboratory selections were also represented in many
176 natural PUF proteins (Supporting Figure S3). This observation corroborates previous work showing
177 that compensatory mutations can readily identified as the consensus in alignments.[34] These findings
178 highlight the ability of natural evolution to discover function enhancing mutations and,
179 simultaneously, the power of laboratory evolution to recapitulate this on a practical timescale. Some
180 families of RNA-binding proteins, such as RNA-binding viral proteins and pentatricopeptide repeat
181 (PPR) proteins are known to have generally lower overall stability.[45 46] These proteins also have very
182 high affinities for RNA. Whether these properties have resulted from an ancestral acquisition of
183 mutations that increase their RNA-binding affinities to compensate for mutations that reduced their
184 stabilities remains an intriguing avenue for further research.

185 **METHODS**

186 **Plasmid and library construction.** For yeast three-hybrid and library selection PUFwt, PUFmut and
187 EGFP ORFs were cloned into the yeast hybrid plasmid pGAD-RC using AscI and NotI restriction
188 sites. NRE RNA (5′-UGUAUAUA-3′) was expressed from either pIIIA-MS2-2-p3HR2-NRE[47] or
189 pGBK-RC-NRE, pGBK-RC-NRE was made by amplifying NRE from pIIIA-MS2-2-p3HR2-NRE
190 and cloning into pGBK-RC using AscI and NgoMIV restriction sites. A random PUFmut mutant
191 library was created with error-prone PCR using the GeneMorph II Random Mutagenesis kit (Agilent
192 Technologies) according to the manufacturer's instructions. Parameters such as template
193 concentration and cycling conditions were also optimised according to the manufacturer's guidelines
194 to introduce approximately 0-4.5 mutations per kilobase of amplified DNA. PCR generated mutated
195 PUFmut ORFs were cloned into pGAD-RC using AscI and NotI restriction sites and transformed into
196 *E. coli* strain DH10B using electroporation. The *E. coli* library was then grown overnight in 100 ml
197 of LB broth at 37°C with shaking. The following day plasmid DNA was extracted and purified using
198 a QIAGEN HiSpeed Plasmid Maxi kit. Purified PUFmut plasmid DNA was then transformed into
199 the *S. cerevisiae* strain MaV204K (*MATα, trp1–901, leu2–3, 112, his3Δ200, ade2–101Δ::kanMX,*
200 *gal4Δ, gal80Δ, SPAL10::URA3, UASGAL1::HIS3, GAL1::lacZ*)[48] containing NRE RNA expression
201 plasmids using the lithium acetate method.[49] Yeast cells were grown in synthetic complete (SC) media
202 lacking leucine, tryptophan and histidine at 30°C to select for functional RNA-protein binding
203 variants. We found that a single mutation, N1080Y, dominated the selected clones and to identify this
204 mutation rapidly in the pool of candidates and find other mutations in the PUFmut ORF we designed
205 a simple second round of selection. The second selection consisted of using the formation of an AluI
206 restriction site (AG<u>C</u>T) in the N1080Y clones, which was PCR amplified directly from yeast colonies,
207 digested with AluI and examined by agarose gel electrophoresis. Clones containing the AG<u>C</u>T
208 mutation appeared as two bands, clones without AGCT appeared as a single band and were thus
209 distinguished without sequencing. Clones lacking the N1080Y mutation were then sequenced to
210 identify mutations responsible for restoration of function.

211

212 **Site directed mutagenesis.** Site directed mutagenesis was used to introduce specific mutations
213 identified in the PUFmut library into both PUF-WT and PUFmut genetic backgrounds. Briefly,
214 primers containing SNPs chosen from the PUFmut library were used in enzymatic inverse PCR (EI-
215 PCR) to amplify pGAD-PUFmut and pGAD-PUF-WT, these were then digested and transformed
216 into *E. coli*.

217

218 **Yeast three-hybrid growth assays.** Yeast growth assays were done using the strain YBZ-1 (*MATa,*
219 *ura3-52, leu2-3, 112, his3-200, trp1-1, ade2, LYS2 :: (LexAop)-HIS3, ura3 :: (lexA-op)-lacZ, LexA-*

220 *MS2 coat (N55K)*)[24] or MaV204K. PUF variants were transformed into YBZ-1 and MaV204K and
221 then grown in 5 ml selective synthetic complete media with shaking for 24 h at 30°C. Cells were then
222 diluted to ~3 x 10^7 cells/ ml and spotted (5 µl) in rows onto selective agar containing increasing
223 concentrations of 3-amino-1,2,4-triazole (3-AT). Each row was sequentially diluted 10-fold from the
224 previous. Cells were incubated for 3 days at 30°C and then photographed.
225
226 **Protein purification.** PUFwt, PUFwt-F897Y, PUFwt-Q931H, PUFwt-M932K and PUFwt-N1080Y
227 ORFs were sub-cloned into pETM30 and expressed as fusions to N-terminal His-tag and GST
228 domains in *E. coli* ER2566 cells. Cells were lyzed by sonication in 0.3 M NaCl, 5 mM Imidazole and
229 0.1 mM PMSF. Lysates were then clarified by centrifugation and incubated for 40 min with His-
230 select nickel affinity beads (Sigma) at 4 °C. Protein bound beads were then transferred to BioRad
231 Poly-Prep columns washed with 50 mM Tris-HCl (pH 8.0), 0.3 M NaCl and 10 mM imidazole.
232 Proteins were then eluted with 50 mM Tris-HCl (pH 8.0), 0.3 M NaCl, 250 mM imidazole, before
233 being dialyzed overnight in 25 mM Tris-HCl (pH 8.0), 0.2 M NaCl at 4°C. Protein concentration was
234 determined by the bicichroninic acid (BCA) assay using bovine serum albumin (BSA) as a standard.
235
236 **RNA electrophoretic mobility shift assays.** Purified PUFwt proteins were incubated at room
237 temperature for 30 min with fluorescein labelled NRE RNA (5′-(Fl)AUUGUAUAUA-3′) or NRE-
238 G2C RNA (5′-(Fl)AUUCUAUAUA-3′) oligonucleotides (Dharmacon) in 10 mM HEPES (pH 8.0),
239 1 mM EDTA, 50 mM KCl, 2 mM DTT, 0.1 mg.ml-1 fatty acid-free BSA, and 0.02% Tween-20.
240 Binding reactions were analysed by 10% PAGE in TAE and fluorescence was detected using a
241 Typhoon FLA 9500 biomolecular imager (GE).
242
243 **Thermal shift assays.** Thermal scanning (25 to 95°C at 1°C/min) was performed using the HRM
244 channel of a Rotor-Gene Q real-time PCR machine (Qiagen). PUFmut-N1080Y was purified as
245 described above but eluted from the His-select nickel affinity beads using TEV protease in 50 mM
246 Tris-HCl pH 8.0, 300 mM NaCl, 0.5 mM EDTA, 1 mM DTT, in order to produce tag-free protein.
247 All assay experiments were performed in triplicate in a final volume of 22 µl. Each sample contained
248 16 µl of protein solution, 4 µl of 20 µM NRE RNA (5′-AUUGUAUAUA-3′) in DEPC-treated water
249 or DEPC-treated water only, and 2 µl 200x SYPRO Orange dye (Invitrogen).
250
251 **Immunoblotting.** PUFwt, PUFmut, PUFmut-F897Y, PUFmut-Q931H, PUFmut-M932K and
252 PUFmut-N1080Y proteins were expressed with C-terminal 9xmyc-tags and detected using anti-myc
253 antibodies (9E10, ThermoFisher Scientific, diluted 1:5000). Mouse anti-rabbit secondary antibodies

254 were used and immunoblots were visualised using an Odyssey infrared imaging system (Li-COR
255 Biosciences).
256
257 **ASSOCIATED CONTENT**
258 **Supporting Information** includes Supporting Table S1 and Supporting Figures (Figures S1-S3).
259 This material is available free of charge via the internet at http://pubs.acs.org.
260
261 **ACKNOWLEDGEMENTS**
262 This work was supported by fellowships and grants from the Australian Research Council
263 (DP140104111, DP170103000 and DP180101656 to A.F. and O.R.), the National Health and Medical
264 Research Council (APP1058442 to A.F.), the Cancer Council Western Australia (to O.R.), the Raine
265 Medical Research Foundation (to T.R.), and a CSIRO Future Science Fellowship in Synthetic
266 Biology (to T.R.). C.P.W. was supported by a UWA Postgraduate Scholarship and a Colliers
267 International PhD Scholarship. We thank J. Karanicolas, A. Bazzoli and J. Budiardjo for prediction
268 of the cavity forming mutation, M. Wickens for the YBZ1 yeast strain, T. Ito for the MaV204K yeast
269 strain as well as pGAD-RC and pGBK-RC plasmids, J. Ermer for assistance with immunoblotting
270 experiments, A.-M. Shearwood for assistance with REMSA experiments, and the EMBL Protein
271 Expression and Purification Facility for the pETM30 plasmid.
272
273 **COMPETING FINANCIAL INTERESTS**
274 The authors declare no conflict of interest.

275 **REFRENCES**

276 1. Yue, P., Li, Z., and Moult, J. (2005) Loss of protein structure stability as a major causative factor
277 in monogenic disease, *J Mol Biol 353*, 459-473.
278 2. Chiti, F., Taddei, N., Bucciantini, M., White, P., Ramponi, G., and Dobson, C. M. (2000)
279 Mutational analysis of the propensity for amyloid formation by a globular protein, *EMBO J*
280 *19*, 1441-1449.
281 3. Wang, W., Nema, S., and Teagarden, D. (2010) Protein aggregation - pathways and influencing
282 factors, *Int J Pharm 390*, 89-99.
283 4. Varshavsky, A. (2005) Regulated protein degradation, *Trends Biochem Sci 30*, 283-286.
284 5. Gregersen, N., Bross, P., Vang, S., and Christensen, J. H. (2006) Protein misfolding and human
285 disease, *Annu Rev Genomics Hum Genet 7*, 103-124.
286 6. Hartl, F. U. (2017) Protein Misfolding Diseases, *Annu Rev Biochem 86*, 21-26.
287 7. DePristo, M. A., Weinreich, D. M., and Hartl, D. L. (2005) Missense meanderings in sequence
288 space: a biophysical view of protein evolution, *Nat Rev Genet 6*, 678-687.
289 8. Zeldovich, K. B., Chen, P., and Shakhnovich, E. I. (2007) Protein stability imposes limits on
290 organism complexity and speed of molecular evolution, *Proc. Natl. Acad. Sci 104*, 16152-
291 16157.
292 9. Li, Y., and Cirino, P. C. (2014) Recent advances in engineering proteins for biocatalysis,
293 *Biotechnol Bioeng 111*, 1273-1287.
294 10. Pakula, A. A., and Sauer, R. T. (1989) Genetic analysis of protein stability and function, *Annu*
295 *Rev Genet 23*, 289-310.
296 11. Somero, G. N. (1995) Proteins and temperature, *Annu Rev Physiol 57*, 43-68.
297 12. Kauzmann, W. (1959) Some factors in the interpretation of protein denaturation, *Adv Protein*
298 *Chem 14*, 1-63.
299 13. Dill, K. A. (1990) Dominant forces in protein folding, *Biochem 29*, 7133-7155.
300 14. Starr, T. N., and Thornton, J. W. (2016) Epistasis in protein evolution, *Protein Sci 25*, 1204-1218.
301 15. Olabode, A. S., Kandathil, S. M., Lovell, S. C., and Robertson, D. L. (2017) Adaptive HIV-1
302 evolutionary trajectories are constrained by protein stability, *Virus Evol 3*, vex019.
303 16. Gong, L. I., Suchard, M. A., and Bloom, J. D. (2013) Stability-mediated epistasis constrains the
304 evolution of an influenza protein, *Elife 2*, e00631.
305 17. Tokuriki, N., Stricher, F., Schymkowitz, J., Serrano, L., and Tawfik, D. S. (2007) The stability
306 effects of protein mutations appear to be universally distributed, *J Mol Biol 369*, 1318-1332.
307 18. Filipovska, A., and Rackham, O. (2012) Modular recognition of nucleic acids by PUF, TALE and
308 PPR proteins, *Mol Biosyst 8*, 699-708.

309 19. Wickens, M., Bernstein, D. S., Kimble, J., and Parker, R. (2002) A PUF family portrait: 3′ UTR
310 regulation as a way of life, *Trends Genet 18*, 150-157.
311 20. Quenault, T., Lithgow, T., and Traven, A. (2011) PUF proteins: repression, activation and mRNA
312 localization, *Trends Cell Biol 21*, 104-112.
313 21. Wang, X., Zamore, P. D., and Hall, T. M. T. (2001) Crystal structure of a Pumilio homology
314 domain, *Mol Cell 7*, 855-865.
315 22. Gupta, Y. K., Lee, T. H., Edwards, T. A., Escalante, C. R., Kadyrova, L. Y., Wharton, R. P., and
316 Aggarwal, A. K. (2009) Co-occupancy of two Pumilio molecules on a single hunchback NRE,
317 *RNA 15*, 1029-1035.
318 23. Dong, S., Wang, Y., Cassidy-Amstutz, C., Lu, G., Bigler, R., Jezyk, M. R., Li, C., Hall, T. M. T.,
319 and Wang, Z. (2011) Specific and modular binding code for cytosine recognition in
320 Pumilio/FBF (PUF) RNA-binding domains, *J Biol Chem 286*, 26732-26742.
321 24. Hook, B., Bernstein, D., Zhang, B., and Wickens, M. (2005) RNA–protein interactions in the
322 yeast three-hybrid system: Affinity, sensitivity, and enhanced library screening, *RNA 11*, 227-
323 233.
324 25. Quirk, D. J., Park, C., Thompson, J. E., and Raines, R. T. (1998) His···Asp Catalytic Dyad of
325 Ribonuclease A: Conformational Stability of the Wild-Type, D121N, D121A, and H119A
326 Enzymes, *Biochem 37*, 17958-17964.
327 26. Schindler, T., Perl, D., Graumann, P., Sieber, V., Marahiel, M. A., and Schmid, F. X. (1998)
328 Surface-exposed phenylalanines in the RNP1/RNP2 motif stabilize the cold-shock protein
329 CspB from Bacillus subtilis, *Proteins 30*, 401-406.
330 27. Wang, X., Minasov, G., and Shoichet, B. K. (2002) Evolution of an antibiotic resistance enzyme
331 constrained by stability and activity trade-offs, *J Mol Biol 320*, 85-95.
332 28. Thomas, V. L., McReynolds, A. C., and Shoichet, B. K. (2010) Structural bases for stability–
333 function tradeoffs in antibiotic resistance, *J Mol Biol 396*, 47-59.
334 29. Beadle, B. M., and Shoichet, B. K. (2002) Structural bases of stability–function tradeoffs in
335 enzymes, *J Mol Biol 321*, 285-296.
336 30. Meiering, E. M., Serrano, L., and Fersht, A. R. (1992) Effect of active site residues in barnase on
337 activity and stability, *J Mol Biol 225*, 585-589.
338 31. Schreiber, G., Buckle, A. M., and Fersht, A. R. (1994) Stability and function: two constraints in
339 the evolution of barstar and other proteins, *Structure 2*, 945-951.
340 32. Shoichet, B. K., Baase, W. A., Kuroki, R., and Matthews, B. W. (1995) A relationship between
341 protein stability and protein function, *Proc. Natl. Acad. Sci 92*, 452-456.
342 33. Bloom, J. D., Labthavikul, S. T., Otey, C. R., and Arnold, F. H. (2006) Protein stability promotes
343 evolvability, *Proc. Natl. Acad. Sci 103*, 5869-5874.

344 34. Bershtein, S., Goldin, K., and Tawfik, D. S. (2008) Intense neutral drifts yield robust and
345 evolvable consensus proteins, *J Mol Biol 379*, 1029-1044.
346 35. Aguilar-Rodríguez, J., Sabater-Muñoz, B., Montagud-Martínez, R., Berlanga, V., Alvarez-Ponce,
347 D., Wagner, A., and Fares, M. A. (2016) The Molecular Chaperone DnaK Is a Source of
348 Mutational Robustness, *Genome Biol Evol 8*, 2979-2991.
349 36. Fares, M. A., Ruiz-González, M. X., Moya, A., Elena, S. F., and Barrio, E. (2002) Endosymbiotic
350 bacteria: groEL buffers against deleterious mutations, *Nature 417*, 398-398.
351 37. Maisnier-Patin, S., Roth, J. R., Fredriksson, A., Nystrom, T., Berg, O. G., and Andersson, D. I.
352 (2005) Genomic buffering mitigates the effects of deleterious mutations in bacteria, *Nat Genet*
353 *37*, 1376-1379.
354 38. Jain, N. K., Jetani, H. C., and Roy, I. (2013) Nucleic acid aptamers as stabilizers of proteins: The
355 stability of tetanus toxoid, *Pharm Res 30*, 1871-1882.
356 39. Chattopadhyay, S., Das, B., and Dasgupta, C. (1996) Reactivation of denatured proteins by 23S
357 ribosomal RNA: role of domain V, *Proc. Natl. Acad. Sci 93*, 8284-8287.
358 40. Samanta, D., Mukhopadhyay, D., Chowdhury, S., Ghosh, J., Pal, S., Basu, A., Bhattacharya, A.,
359 Das, A., Das, D., and DasGupta, C. (2008) Protein folding by domain V of Escherichia coli
360 23S rRNA: specificity of RNA-protein interactions, *J Bacteriol 190*, 3344-3352.
361 41. Banerjee, D., and Sanyal, S. (2014) Protein Folding Activity of the Ribosome (PFAR)—A Target
362 for Antiprion Compounds, *Viruses 6*, 3907-3924.
363 42. Voisset, C., Saupe, S. J., and Blondel, M. (2011) The various facets of the protein-folding activity
364 of the ribosome, *Biotechnol J. 6*, 668-673.
365 43. Sulijoadikusumo, I., Horikoshi, N., and Usheva, A. (2001) Another function for the mitochondrial
366 ribosomal RNA: protein folding, *Biochem 40*, 11559-11564.
367 44. Docter, B. E., Horowitz, S., Gray, M. J., Jakob, U., and Bardwell, J. C. (2016) Do nucleic acids
368 moonlight as molecular chaperones?, *Nucleic Acids Res. 44*, 4835-4845.
369 45. Chen, P., and Shakhnovich, E. I. (2010) Thermal adaptation of viruses and bacteria, *Biophys J*
370 *98*, 1109-1118.
371 46. Coquille, S., Filipovska, A., Chia, T., Rajappa, L., Lingford, J. P., Razif, M. F., Thore, S., and
372 Rackham, O. (2014) An artificial PPR scaffold for programmable RNA recognition, *Nat*
373 *Commun 5*.
374 47. Filipovska, A., Razif, M. F., Nygård, K. K., and Rackham, O. (2011) A universal code for RNA
375 recognition by PUF proteins, *Nat Chem Biol 7*, 425-427.
376 48. Ito, T., Tashiro, K., Muta, S., Ozawa, R., Chiba, T., Nishizawa, M., Yamamoto, K., Kuhara, S.,
377 and Sakaki, Y. (2000) Toward a protein–protein interaction map of the budding yeast: a

378 comprehensive system to examine two-hybrid interactions in all possible combinations
379 between the yeast proteins, *Proc. Natl. Acad. Sci 97*, 1143-1147.
380 49. Gietz, R. D., and Woods, R. A. (2002) Transformation of yeast by lithium acetate/single-stranded
381 carrier DNA/polyethylene glycol method, *Methods Enzymol 350*, 87-96.
382

383 **FIGURE LEGENDS**

384

385 **Figure 1.** Modular structure of the wild-type human Pumilio PUF RNA-binding domain (PUFwt) in
386 complex with its cognate RNA target the nanos-response element (NRE), location of the of the
387 PUFmut cavity mutations compared to PUFwt, growth and stability of PUFmut, and PUFmut library
388 preparation and screening. (A) PUF proteins are comprised of 8 repeats (R1-R8) each formed by 3
389 alpha-helix bundles arranged in an extended arc flanked by two related bundles (R1′ & R8′). RNA
390 sits inside the concave surface of the PUF protein in an anti-parallel orientation with each nucleotide
391 (U1-U8) binding a single PUF repeat. Binding is achieved via hydrogen bonding or van der Waals
392 contacts with the Watson-Crick edge of each base by amino acids at position 12 and 16. The
393 recognition of each nucleotide as determined by specific residue combinations at position 12 and 16
394 is such that that cysteine and glutamine bind adenine, asparagine and glutamine bind uracil, and serine
395 and glutamate bind guanine. (B) Location and structural context of mutations introduced into PUF
396 proteins. PUFwt contains two phenylalanine and a single glutamine residue at position 882, 905, 913
397 respectively (relevant residues are labelled and coloured red). Switching of these bulky residues for
398 smaller glycines at these positions (PUFmut) causes the protein to become insoluble and abolishes
399 RNA-binding activity. (C) Yeast transformed with an NRE RNA expressing plasmid and either
400 PUFwt, PUFmut or EGFP expressing plasmids. Growth on selective media is dependent on RNA-
401 protein binding. (D) PUFwt and PUFmut expression and stability shown by immunoblotting using
402 anti-myc antibodies. (E) PUFmut libraries were prepared by amplifying PUFmut ORFs using epPCR,
403 which were then ligated into a GAL4 fusion plasmid and transformed into MaV204K using the
404 lithium acetate method. These strains contain the NRE RNA "bait" plasmid (red in diagram). The
405 PUFmut mutant library was then plated on selective SC-L-T-H media to isolate positive RNA-binding
406 clones. Surviving clones were analysed by PCR followed by AluI digestion and Sanger sequencing.

407

408 **Figure 2.** Mutations identified that rescue the function of PUFmut. (A) Mutations were mapped onto
409 the protein structure and occurrence of each mutation is represented by colour, transition from blue
410 to red represents an increase in abundance. Location of specific mutations that were able to rescue
411 PUFmut activity. (B) F897 sits on the inner surface of R2 and was mutated to Y897, the two residues
412 are structurally similar with tyrosine containing an additional hydroxyl group. (C) Q931 is polar and
413 uncharged, located on the R3 repeat and was mutated to H931, which contains a large positively
414 charged side chain. (D) M932 sits immediately next to Q931 on R3 but is closer to the nearest
415 ribonucleotide (U7). Methionine contains a sulphur group in the side chain and was mutated to lysine
416 (K932), which is positively charged. (E) N1080 is a stacking residue that sits immediately between
417 two pairs of RNA binding residues, N1043 and Q1047 on R6 which bind U3, and S1079 and Q1083

418 on R7 which bind G2. Mutation to Y1080 changes a positively charged side chain to a large
419 hydrophobic side chain.
420
421 **Figure 3.** RNA-binding of mutant PUF proteins is modulated by mutations that rescue PUFmut
422 function. Yeast three-hybrid survival assays on synthetic complete media lacking leucine, uracil and
423 histidine with increasing concentration of 3-AT. The top panels show individual mutants in the
424 PUFwt background, the bottom panels show mutants in a PUFmut background. All yeast cultures
425 where grown for 24 h under identical conditions and OD_{600} normalised so that an equal number of
426 cells were used for each strain; each spot moving left to right represents a 10-fold dilution of the
427 preceding one.
428
429 **Figure 4.** RNA-binding affinity of library mutants in PUFwt backgrounds, and stability of PUFmut
430 library mutants. (A) Protein stability of PUFmut mutants. Immunoblot of total yeast lysate performed
431 using anti-myc antibodies. A mitochondrial protein (mtCO2) was used as a loading control. (B)
432 Purified PUFwt proteins containing RNA-protein interface mutations titrated against fluorescent
433 NRE-RNA probes in an RNA electrophoretic mobility shift assay (REMSA). The left most lanes
434 contained no protein, each lane moving right contains an increasing amount of purified protein.

Figure 1.

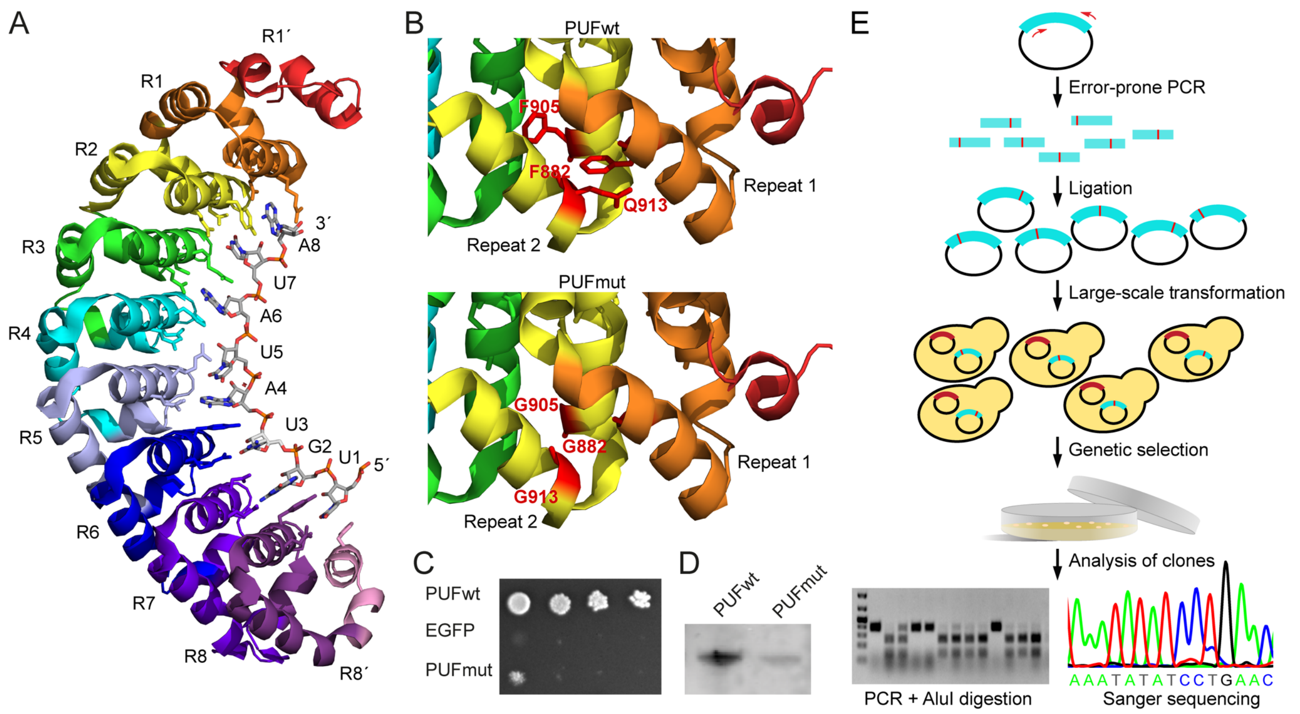

Figure 2.

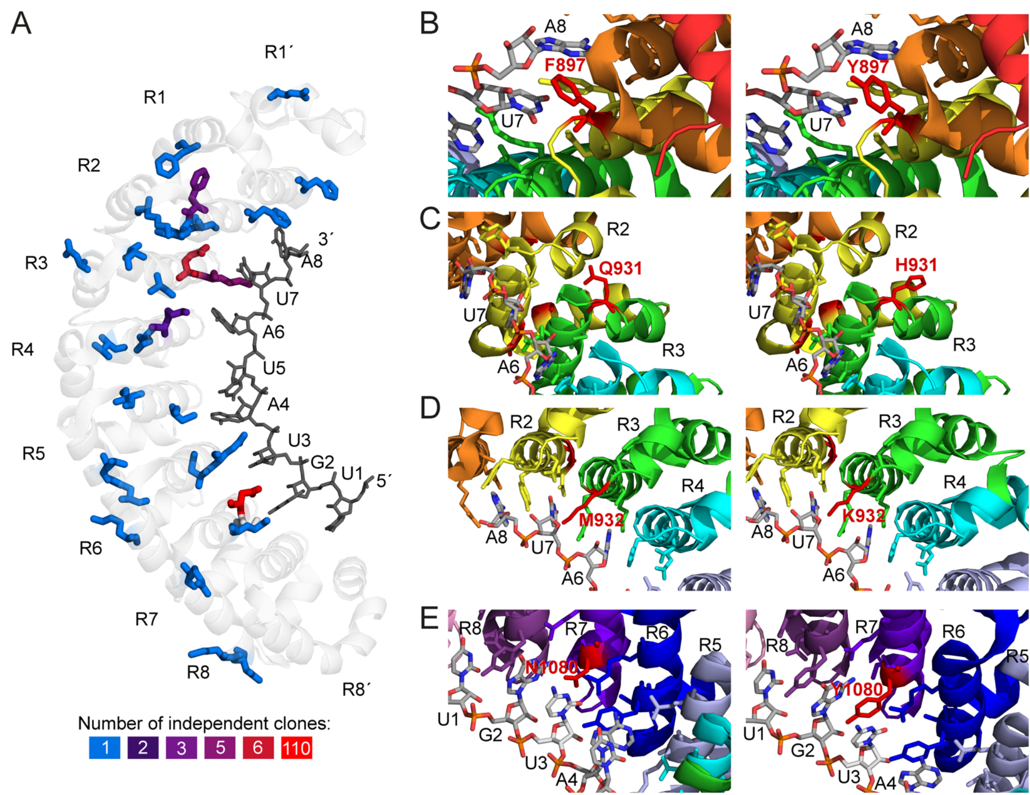

Figure 3.

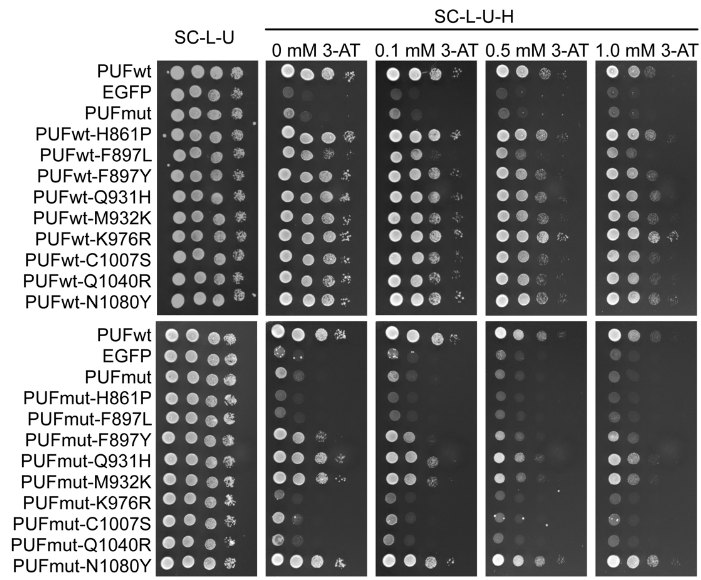

Figure 4.