Supporting Information

Wang et al. 10.1073/pnas.0812076106

SI Methods

Protein Expression and Purification. The RNA-binding domain of FBF-2 (amino acids 121–632) was expressed as a fusion protein with GST in *Escherichia coli* as described elsewhere (1). After releasing FBF-2 by TEV protease cleavage, the protein was stabilized by addition of target RNA. This protein/RNA complex was refractory to crystallization, so we incubated the complex (diluted to 0.5 mg/mL) with subtilisin (1 μg/mL) at 4 °C for 4 h. This limited proteolysis produced a stable fragment that was identified by mass spectrometry as residues 164–575.

A cDNA encoding the minimal FBF-2 RNA-binding domain (residues 164–575) was cloned into the vector pGEX-6P1 (GE Healthcare), which encodes an N-terminal GST tag followed by a TEV cleavage site. The vector was transformed into *E. coli* strain BL21 star (DE3) (Invitrogen), and cultures were grown at 37 °C in Luria–Bertani medium supplemented with 100 μg/mL ampicillin until the OD_{600} reached 0.6–0.8. Fusion protein expression was induced by addition of 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and incubation at 25 °C for 16–20 h. Cell pellets were frozen at −20 °C. On thawing, a cell pellet from each liter of culture was resuspended into 25-mL lysis buffer [50 mM Tris (pH 7.5), 250 mM NaCl, 250 mM $(NH_4)_2SO_4$, 5 mM β-mercaptoethanol (β-ME), 1 mM EDTA]. After sonication and centrifugation of the lysate, the soluble fraction was mixed with 1 mL of glutathione resin (Sigma) per liter of culture in a 50-mL conical tube rotating at 4 °C for 3–4 h. The mixture was then transferred into a 25-mL disposable column. The beads were washed with 100 column volumes of lysis buffer, 25 column volumes of an ATP wash buffer [20 mM Tris (pH 8.5), 10 mM ATP, 5 mM $MgCl_2$], and another 100 column volumes of the lysis buffer. The GST–TEV–FBF-2 fusion protein was eluted with 8 column volumes of an elution buffer [50 mM Tris (pH 8.0), 150 mM NaCl, 10 mM reduced glutathione) and cleaved by addition of 1:100 (vol/vol) recombinant TEV protease (1 mg/mL) at 4 °C overnight. FBF-2 was then separated from the GST tag and TEV with a Hi-Trap Heparin column, loading in buffer A and eluting with a gradient from 0–100% buffer B [buffer A: 20 mM Tris (pH 7.5), 5 mM β-ME, 1 mM EDTA; buffer B: 20 mM Tris (pH 7.5), 5 mM β-ME, 1 mM EDTA, 1 M NaCl, 1 M $(NH_4)_2SO_4$]. FBF-2 eluted at about 25% of buffer B. The protein was concentrated to ≈1 mg/mL and mixed with a slight excess of the desired 9-nt RNA at a molar ratio of 1:1.1. The protein–RNA mixture was incubated at 4 °C overnight for optimal binding and then purified by passing through a Superdex 200 gel filtration column [running buffer: 50 mM Tris (pH 7.5), 200 mM NaCl, 5 mM DTT, 1 mM EDTA]. The purified protein/RNA complex was exchanged into a final buffer [10 mM Tris (pH 7.5), 50 mM NaCl, 5 mM DTT] and concentrated to OD_{280} = 2.5 or 7.5 (approximate protein concentration of 1.7 or 5.0 mg/mL, respectively) using Amicon Ultra-15 concentrators (10-kDa cutoff). Selenomethionine (SeMet)-labeled FBF-2 was expressed in the methionine auxotrophic strain B834 (Novagen). Protein purification was as described for the WT, except that 5–10 mM DTT was used in place of 5 mM β-ME.

GST–FBF-2 and GST–PUF-8 were prepared as described above but without TEV cleavage. Following glutathione affinity purification, the fusion proteins were purified with HiTrap Heparin and Superdex 200 gel filtration columns. GST-chimera 1 was prepared similarly; however, after glutathione affinity purification, the fusion protein was purified with a HiTrap Q HP column [GE Healthcare; buffer A: 20 mM Tris (pH 8.0), 5 mM β-ME, 1 mM EDTA; buffer B: 20 mM Tris (pH 8.0), 5 mM β-ME, 1 mM EDTA, 1 M NaCl, 1 M $(NH_4)_2SO_4$] and a Superdex 200 gel filtration column.

Crystal Preparation. Crystals of the FBF-2/RNA complexes typically grew by hanging drop vapor diffusion to 0.2–0.5 μm overnight at room temperature using either crystallization solution 1 [100 mM Bicine (pH 9.0), 2% (vol/vol) 1,4-Dioxane, 10% (wt/vol) polyethylene glycol 20,000, 10 mM $MgCl_2$] at a 1:1 (vol/vol) ratio with a protein/RNA complex at a high protein concentration (OD_{280} = 7.5) or crystallization solution 2 [100 mM Tris (pH 7.5), 10% (wt/vol) polyethylene glycol 8,000, 8% (vol/vol) ethylene glycol] at a 1:1 (vol/vol) ratio with a protein/RNA complex at a lower protein concentration (OD_{280} = 2.5). Crystals of FBF-2/FBE and SeMet FBF-2/FBE were grown using crystallization solution 1, and 300 μL of paraffin oil was added above a 400-μL well solution to decrease the diffusion rate. All other crystals were grown using crystallization solution 2. Shortly before data collection, crystals were transferred into a series of modified well solutions containing 5%, 10%, and 20% (vol/vol) glycerol, incubating in each solution for 5 min. Crystals were frozen by flash-cooling in liquid nitrogen.

Data Collection and Processing. All diffraction data were collected from crystals at 100 K using a synchrotron X-radiation source (SER-CAT beamline 22-ID or 22-BM at the Advanced Photon Source, Argonne National Laboratory). Data for the SeMet crystals were collected at a wavelength of 0.97937 Å, and data for the native data sets were collected at a wavelength of 1.0 Å. Data were indexed and scaled with HKL2000 (2), and converted to structure factors using SCALEPACK2MTZ from the CCP4i software package (3). Data collection and processing statistics are shown in Table S1.

Structure Determination and Refinement. The initial structure of an FBF-2/RNA complex (FBF-2/FBE) was phased using the SAD method. All 14 selenium sites were identified by SOLVE (4). After density modification by RESOLVE (5), 377 residues (91%) were built automatically by ARP/wARP (6), and the longest chain comprised 298 residues. R_{work} and R_{free} were 0.206 and 0.265, respectively. All 9 nucleotides could be identified in the initial experimental electron density map. After refinement with CNS (7), the partial (91%) SeMet FBF-2 structure was used to calculate phases for the native FBF-2/FBE dataset. After manual rebuilding using O (8) and refinement with CNS (simulated annealing, grouped and individual temperature factor refinements as well as energy minimization), the FBF-2/FBE structure comprised residues L168–S567 and nucleotides U1–A9. An N-terminal glycine residue encoded by the TEV cleavage site, N-terminal residues SNNV (164–167), and C-terminal residues THPIYELQ (578–575) were not included in the structure because of poor electron density at the N- and C-termini. Phenix.Refine (9) was used for addition of water molecules and Translation/Libration/Screw (TLS) refinement. Crystal structures of the 5 other FBF-2/RNA complexes (FBF-2/*gld-1* FBEa, FBF-2/*gld-1* FBEa G4A mutant, FBF-2/*fem-3* PME, FBF-2/*fog-1* FBEa, and FBF-2/*gld-1* FBEb) were determined by molecular replacement using the FBF-2 coordinates from the FBF-2/FBE complex as the search model. Simulated annealing (2,500 K) was used to reduce model bias. Model building and structural refinement were performed as for the FBF-2/FBE complex. Refinement statistics are shown in Table S1. For all structures, all φ–ψ torsion angles are within allowed regions of the Ram-

achandran plot and ≥98% are in the most favored regions. All superpositions were calculated using SUPERIMPOSE from the CCP4i software package (3). The change in angle between the N- and C-terminal halves of FBF-2 relative to PUM1 was calculated using DynDom (10). Figures were prepared with PyMol (11).

Electrophoretic Mobility Shift Assays. Equilibrium dissociation constants were determined by electrophoretic mobility shift assay. RNA oligos were purchased from Dharmacon, Inc. and radiolabeled using ^{32}P-γ-ATP (PerkinElmer Life Sciences) and T4 polynucleotide kinase (New England Biolabs). For binding reactions, serially diluted protein samples were mixed with 100 pM radiolabeled RNA oligo for ≈24 h at 4 °C in a buffer containing 10 mM Hepes (pH 7.4), 50 mM KCl, 50 mM $(NH_4)_2SO_4$, 2 mM DTT, 1 mM EDTA, 0.02% (vol/vol) Tween-20, 0.1 mg/mL BSA, and 0.2 mg/mL yeast tRNA. After incubation, 4 μL of 15% (vol/vol) Ficoll 400 was added to each 20-μL reaction immediately before loading, and 7 μL of each sample was loaded on a prerun nondenaturing 10% (wt/vol) polyacrylamide gel in 0.5× TBE. Gels were run at 100 V at 4 °C for 1 h. After drying, the gels were exposed to storage phosphor screens (GE Healthcare) for 2 days and scanned using a Molecular Dynamics Typhoon phosphorimaging system (GE Healthcare). The intensities of bands were measured with ImageQuant (Amersham). The apparent dissociation constants were calculated using IgorPro (WaveMetrics), assuming a Hill coefficient of 1. The dissociation constants and SEs reported here were from at least 3 independent experiments. Statistical analyses were performed with Graphpad Prism (Graphpad LLC). A representative experiment is shown in Fig. S7. The conditions for the assay were established by testing the effect on K_d of varying incubation time (2–48 h), RNA concentration (30–1,000 pM), and time after addition of loading buffer (5–25 min). No significant difference in K_d was detected under these conditions ($P \geq 0.5$). Determination of the percentage of active protein was performed as described previously (12). FBF-2 was 85–98% active, GST–FBF-2 was 45–77% active, GST–PUF-8 was 49–70% active, and GST-chimera 1 was 66–77% active. K_d values in Table 1 and Fig. 5*D* were adjusted based on % active protein.

DNA Constructs Used in the Yeast 3-Hybrid System. FBF-1 (amino acids 121–614; GenBank accession no. NM_062815), FBF-2 (amino acids 121–632; GenBank accession no. NM_062819), and PUF-8 (amino acids 143–535; GenBank accession no. NM_063122) were used both in the 3-hybrid analysis and as templates for creating the chimeric proteins. The portions of FBF and PUF-8 (listed N-terminal to C-terminal) used in the chimeric proteins are as follows: chimera 1 (PUF-8 M143–V320; FBF-1 I326–V399; PUF-8 H380–H535), chimera 2 (PUF-8 M143–H366; FBF-1 L371–V399; PUF-8 H380–H535), chimera 3 (PUF-8 M143–V320; FBF-1 I326–K370; PUF-8 C367–H535), chimera 4 (PUF-8 M143–I325; FBF-2 A333–F342; PUF-8 I336–H535), and chimera 5 (PUF-8 M143–T346; FBF-2 L353–G362; PUF-8 C357–H535). Proteins were expressed in yeast using the pACT2 plasmid. For chimeras 1, 2, and 3, the fragments were blunt-end ligated together. Ligated fragments were selected via PCR and subsequently cloned into the EcoRI and NcoI sites of pACT2. For chimeras 4 and 5, site-directed mutagenesis was used to create the chimeric protein. The single residue changes (Fig. 5*B*, yellow triangles) were generated in FBF-2 (amino acids 121–632) via site-directed mutagenesis. DNA oligonucleotides designed to express various RNA sequences were cloned into the XmaI/SmaI and SphI sites of pIIIa MS2–2.

Yeast 3-Hybrid Assays. Gal4-activation domain fusion proteins were expressed from either pACT or pACT2 plasmids. RNA–protein interactions were analyzed using the B-glo assay described by Hook et al. (13) with the following modification: 100 μL of saturated cultures (36–48 h of growth) was diluted into 4-mL selective media and allowed to grow 2–2.5 h to reach an OD_{660} of 0.1–0.2. For data analysis, the relative light units (RLUs) were adjusted to an OD_{660} of 0.1 and divided by the sample volume to give RLUs per microliter. The values reported are an average of 4 separate experiments.

1. Bernstein D, Hook B, Hajarnavis A, Opperman L, Wickens M (2005) Binding specificity and mRNA targets of a C. elegans PUF protein, FBF-1. *RNA* 11:447–458.
2. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307–326.
3. CCP4 (Collaborative Computational Project N) (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr D* 50:760–763.
4. Terwilliger TC, Berendzen J (1999) Automated MAD and MIR structure solution. *Acta Crystallogr D* 55:849–861.
5. Terwilliger TC (2000) Maximum-likelihood density modification. *Acta Crystallogr D* 56:965–972.
6. Perrakis A, Morris R, Lamzin VS (1999) Automated protein model building combined with iterative structure refinement. *Nat Struct Biol* 6:458–463.
7. Brünger AT, et al. (1998) Crystallography and NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D* 54:905–921.
8. Jones TA, Zou J-Y, Cowan SW, Kjeldgaard M (1991) Improved methods for building protein models in electron-density maps and the location of errors in these models. *Acta Crystallogr A* 47:110–119.
9. Adams PD, et al. (2002) PHENIX: Building new software for automated crystallographic structure determination. *Acta Cryst* D58:1948–1954.
10. Hayward S, Berendsen HJ (1998) Systematic analysis of domain motions in proteins from conformational change: New results on citrate synthase and T4 lysozyme. *Proteins* 30:144–154.
11. DeLano W (2002) *The PyMOL Molecular Graphics System* (DeLano Scientific, San Carlos, CA).
12. Cheong CG, Hall TM (2006) Engineering RNA sequence specificity of Pumilio repeats. *Proc Natl Acad Sci USA* 103:13635–13639.
13. Hook B, Bernstein D, Zhang B, Wickens M (2005) RNA-protein interactions in the yeast three-hybrid system: Affinity, sensitivity, and enhanced library screening. *RNA* 11:227–233.

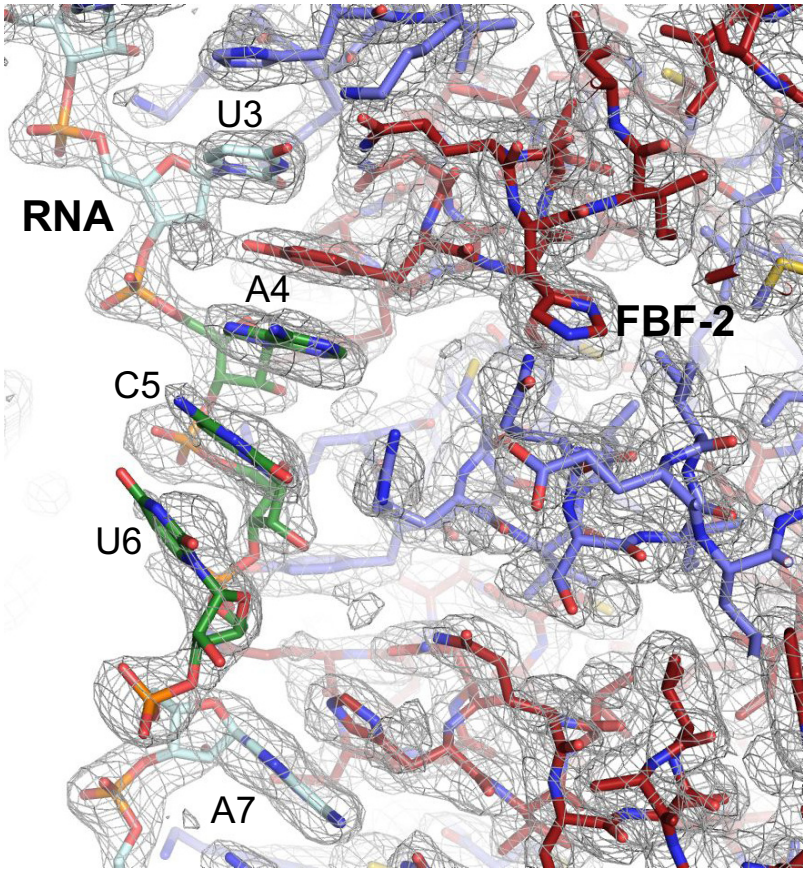

Fig. S1. Experimental SAD electron density map contoured at 1σ. The final refined coordinates of the FBF-2/FBE RNA complex are superimposed. FBF-2 is colored with the carbon atoms to match the repeats in Fig. 1*A*. RNA is shown with cyan carbon atoms (bases 1–3 and 7–9) and green carbon atoms (bases 4–6).

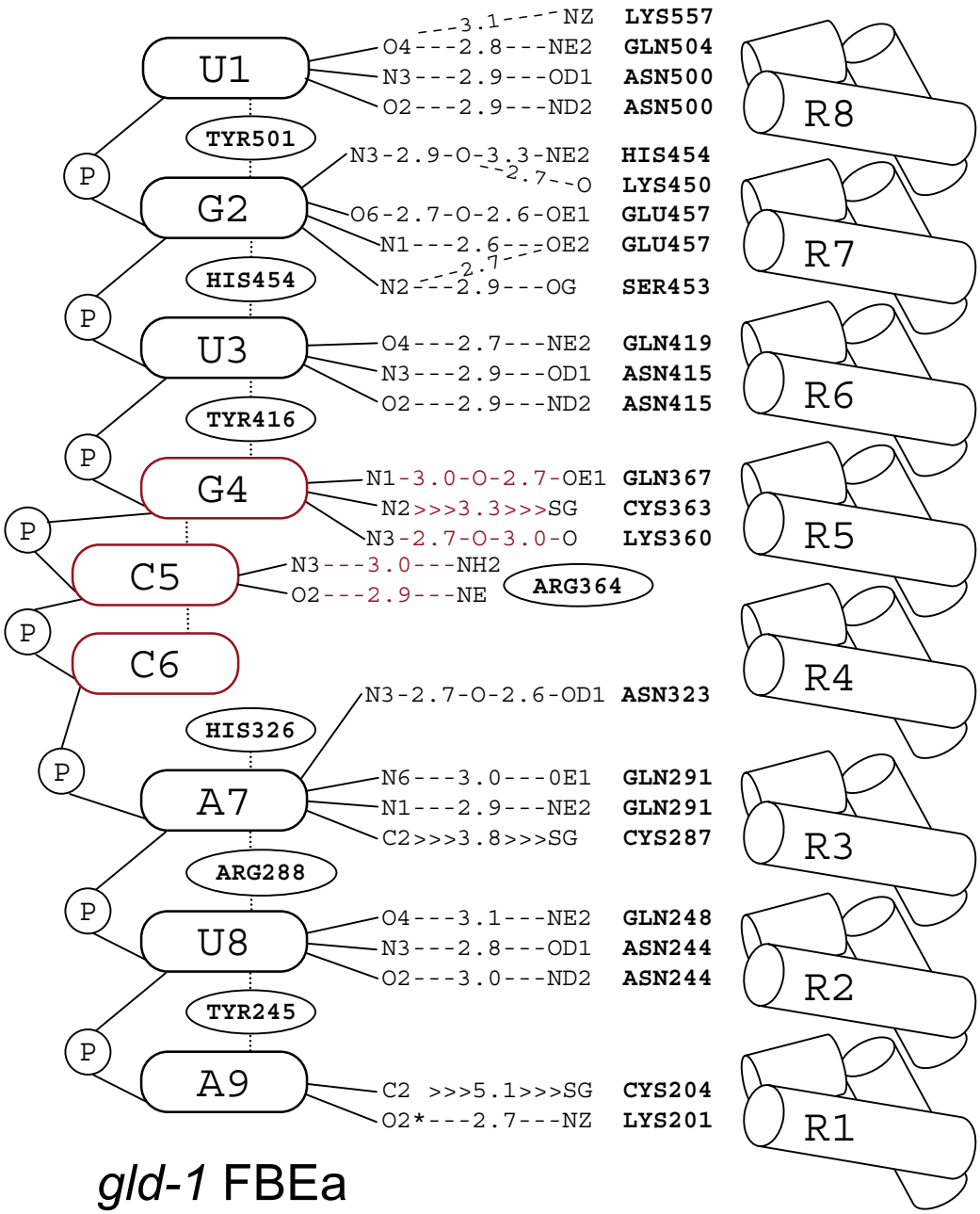

Fig. S2. Schematic representations of the interactions between FBF-2 and *gld-1* FBEa (*A*), *fem-3* PME (*B*), *gld-1* FBEa G4A mutant (*C*), *gld-1* FBEb (*D*), *fog-1* FBEa (*E*), and FBE (*F*) RNAs. Interactions unique to FBF-2 are colored red. Distances between interacting atoms (Å) are indicated. Water-mediated contacts are indicated by dashes with an intervening ''O''; Van der Waals contacts are indicated by arrowheads (>).

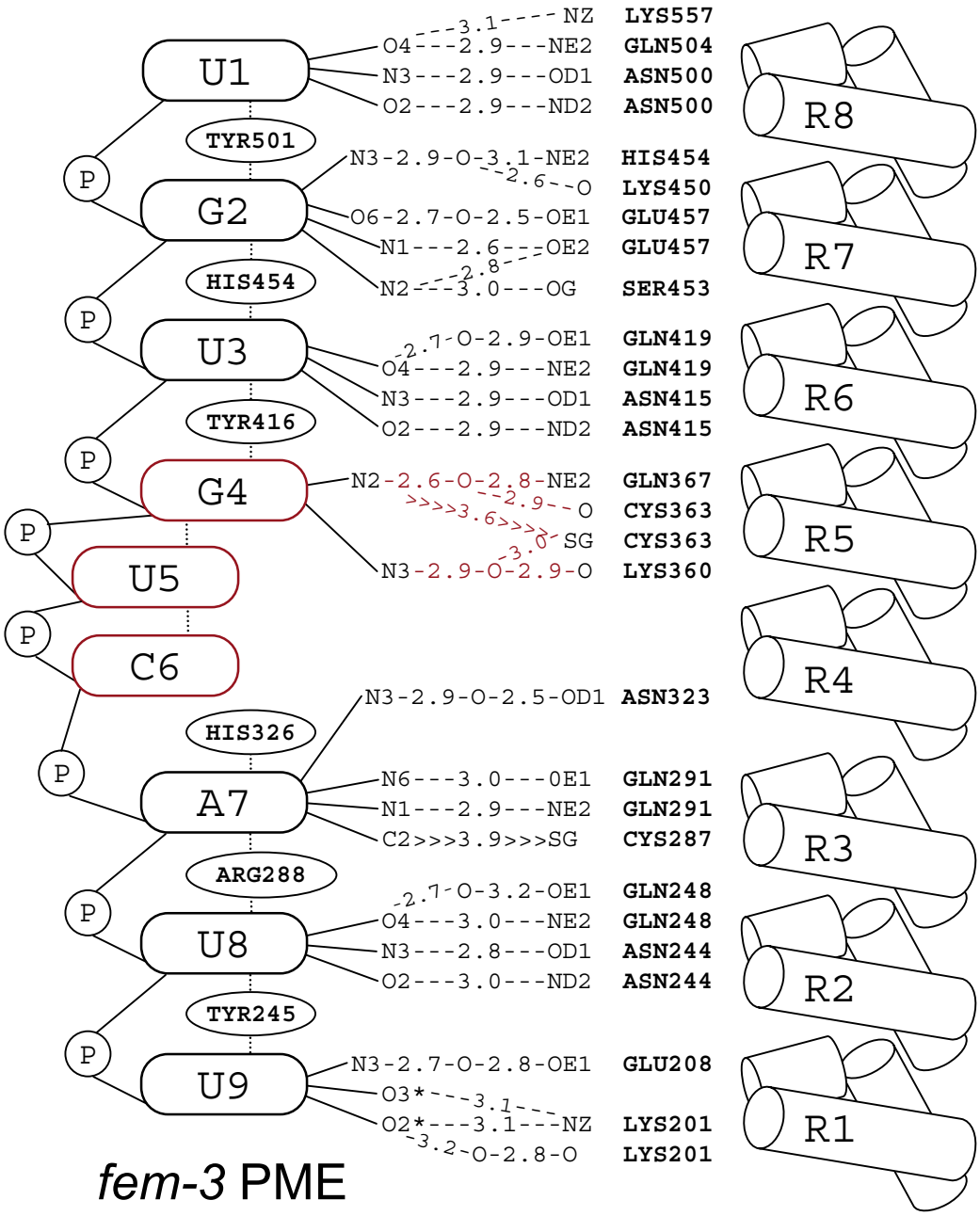

Fig. S2 (continued).

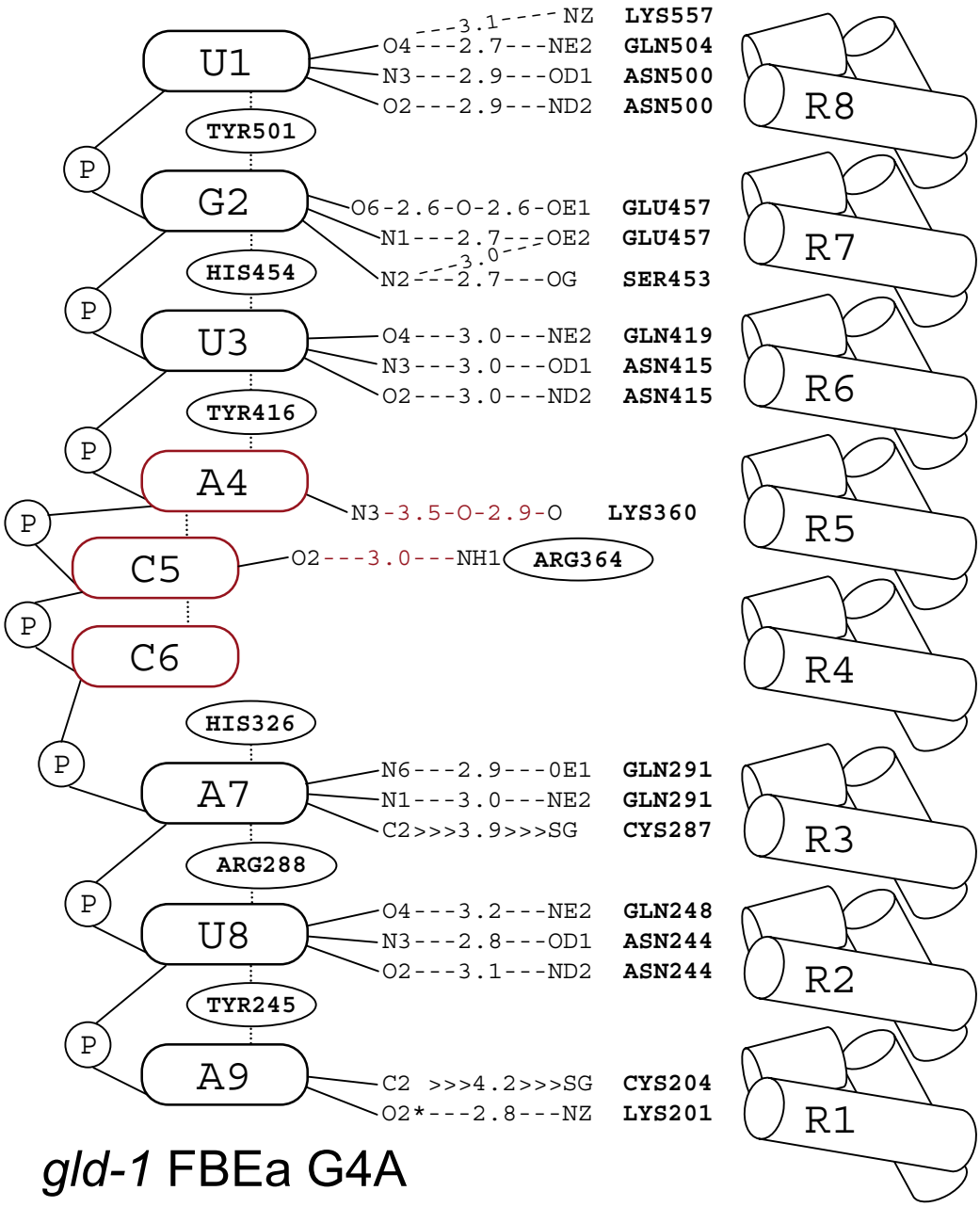

Fig. S2 (continued).

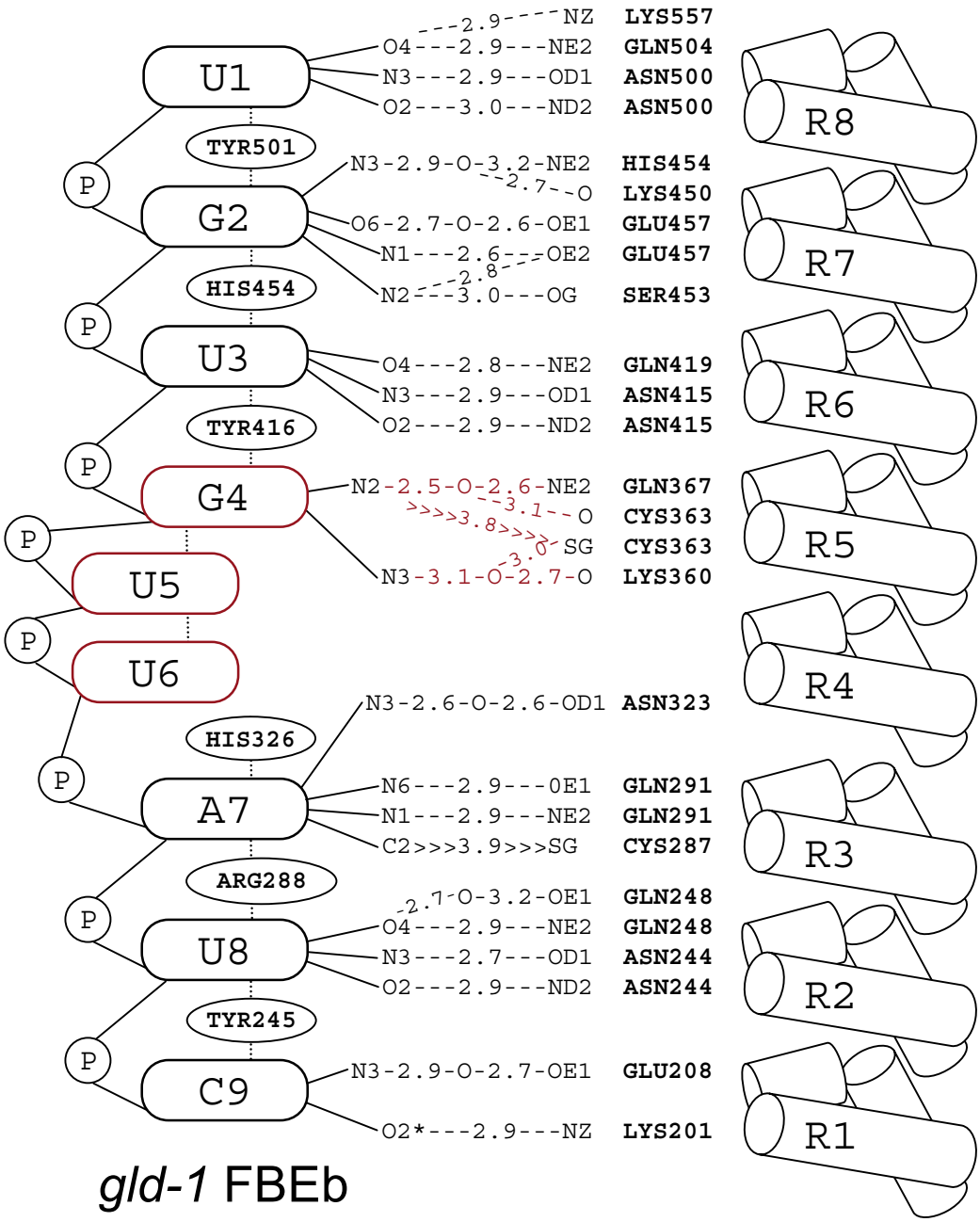

Fig. S2 (continued).

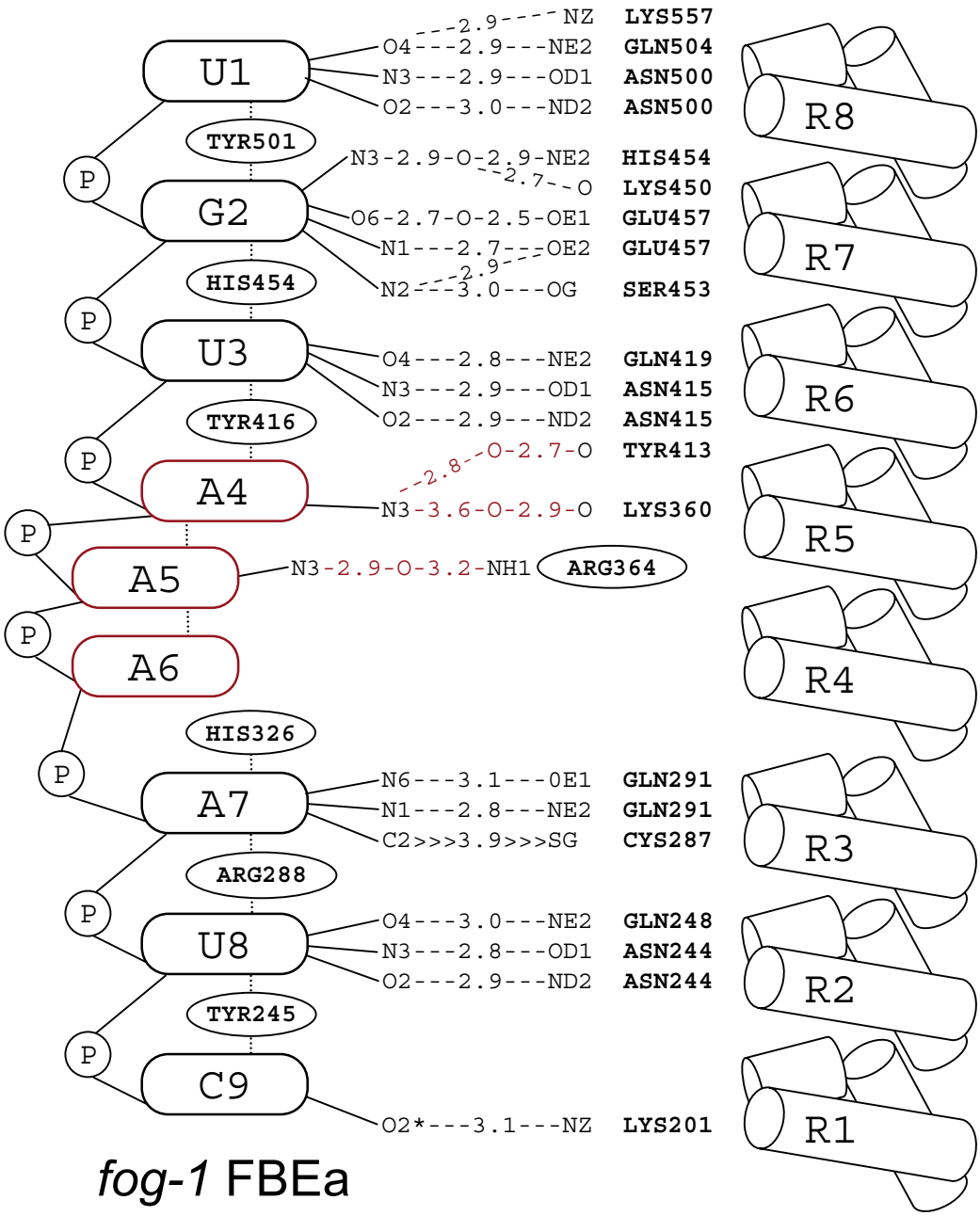

Fig. S2 (continued).

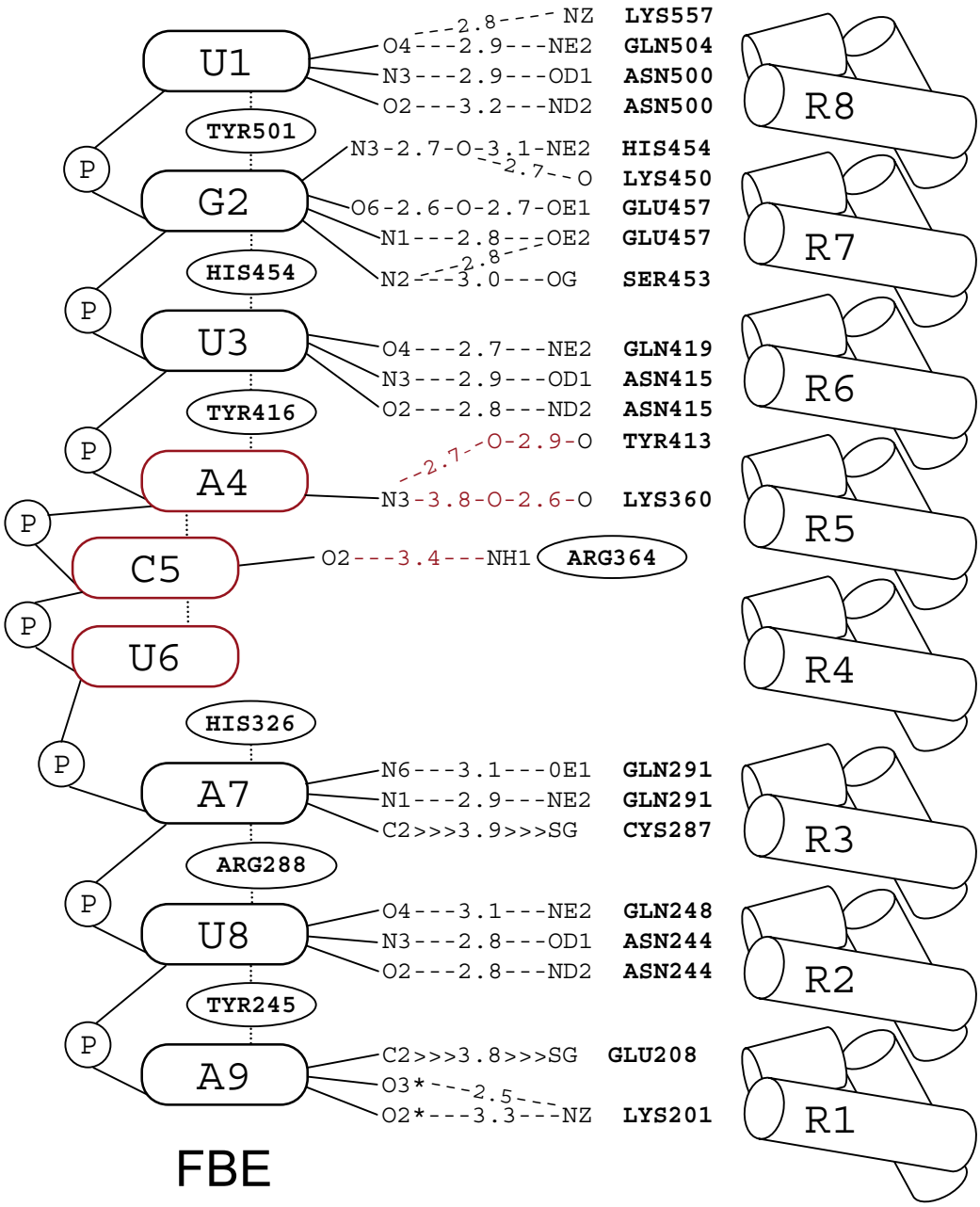

Fig. S2 (continued).

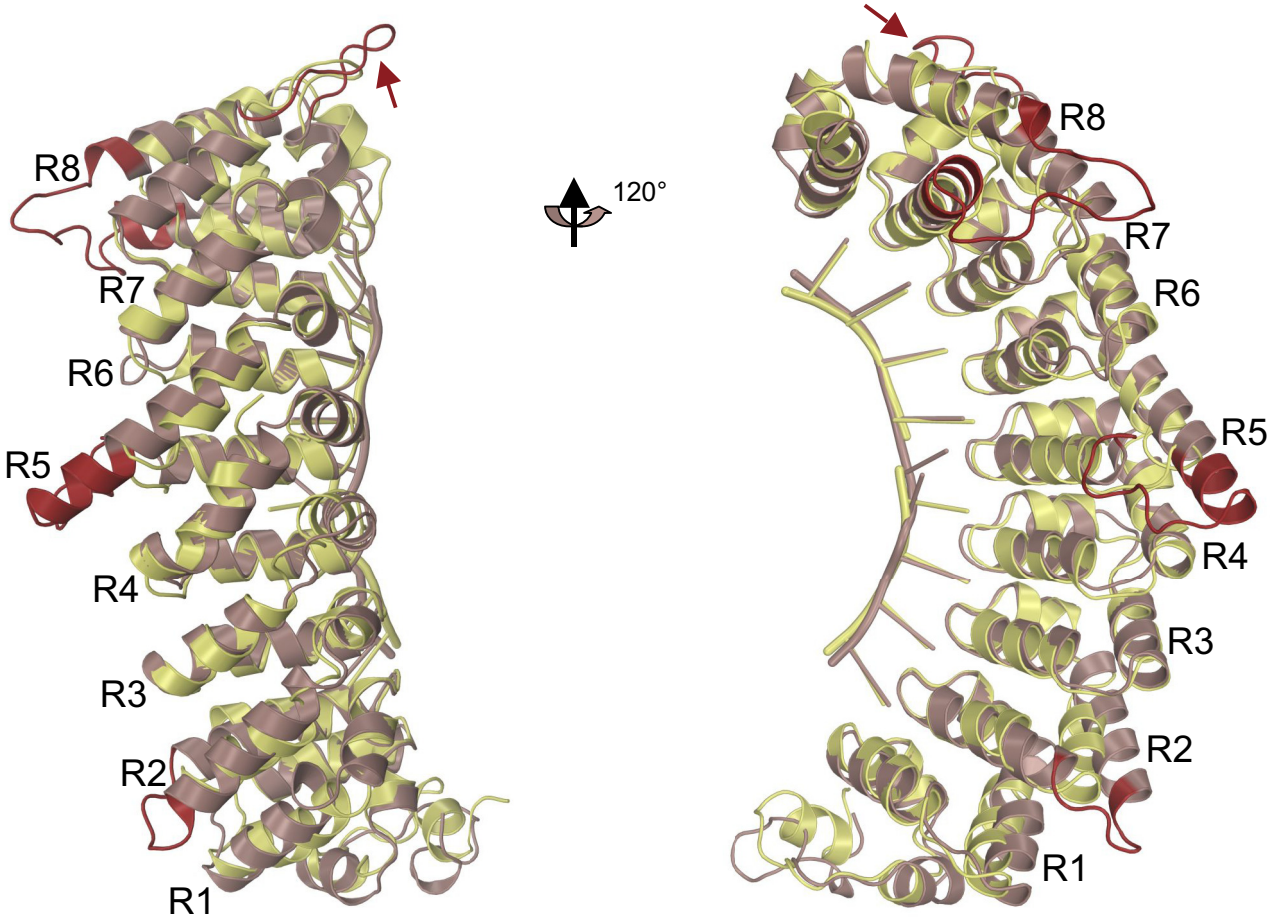

Fig. S3. Superposition of the ribbon diagrams of FBF-2 (mauve) and PUM1 (yellow). The N- and C-terminal halves of PUM1 (repeats 1–4 and 5–8, respectively) are individually superimposed on the equivalent portions of FBF-2. The molecules on the right are rotated 120° about the *y* axis relative to the molecules on the left. The longer helix (third helix in repeat 5) and longer loop regions between the second and third helices (repeats 2, 5, and 8) in FBF-2 are shown in red. The extended loop between repeats 7 and 8 is also shown in red and indicated by an arrow.

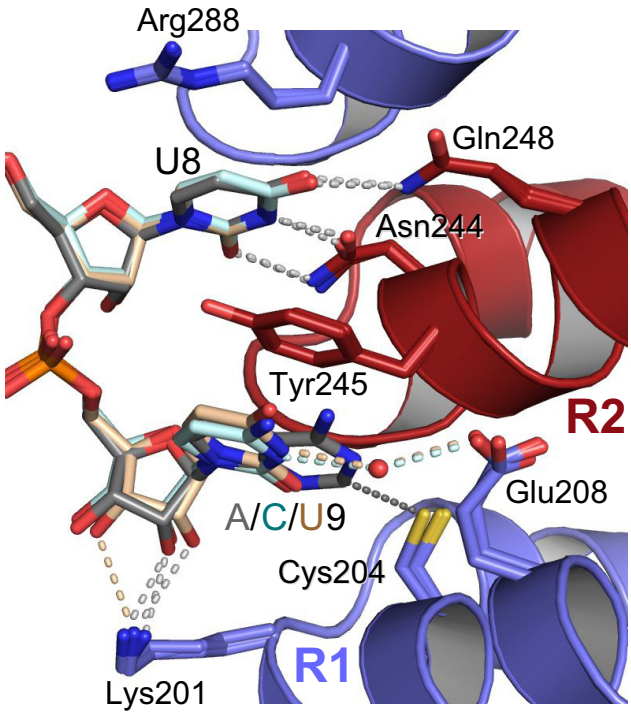

Fig. S4. Recognition of different RNA bases at the ninth position. For comparison, structures of FBF-2 in complex with *gld-1* FBEa (gray, A9), *fog-1* FBEa (cyan, C9) and *fem-3* PME (tan, U9) are superimposed.

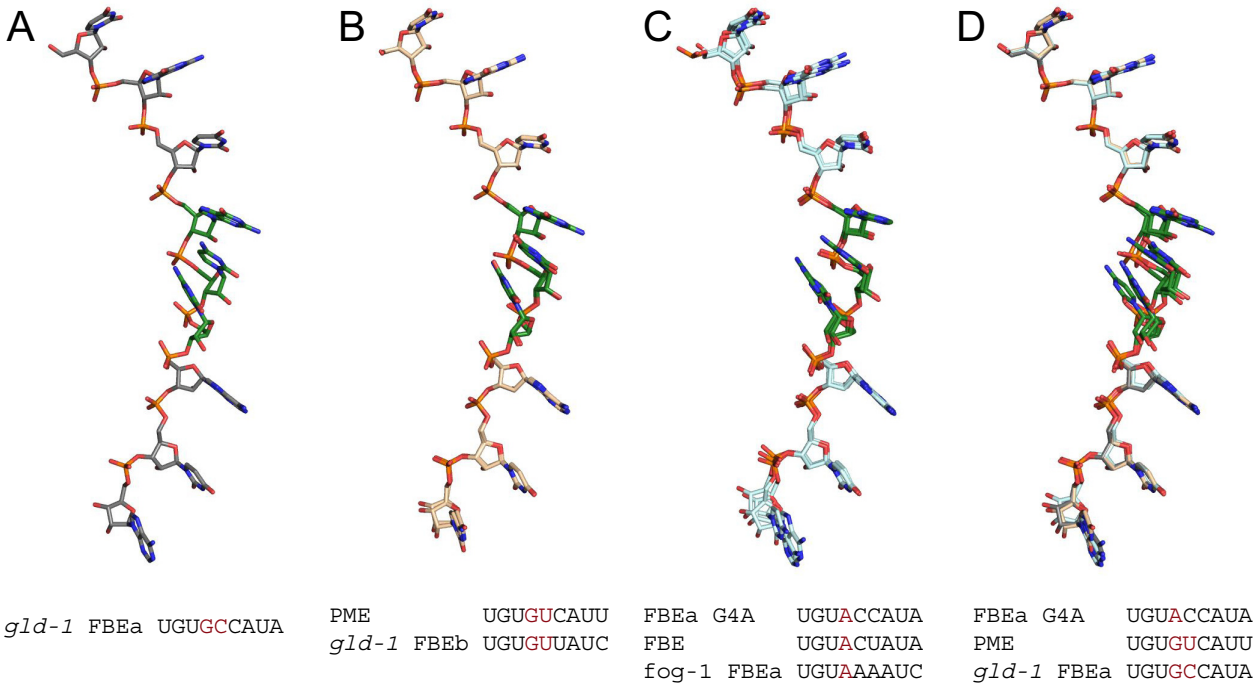

Fig. S5. Three classes of RNA conformation in structures of FBF-2 with target RNAs. The structures of bases 4–6 of FBF-2 target RNAs group into 3 conformational clusters. RNA sequences with a guanosine at the fourth position can be divided into 2 classes with either a cytosine (*A*) or uracil (*B*) at the fifth position, and sequences with an adenine at the fourth position (*C*) form the third class. (*D*) Superposition of the structures of representative members of each class: *gld-1* FBEa RNA (gray), *fem-3* PME RNA (tan), and *gld-1* FBEa G4A mutant RNA (cyan). In the first class (G4C5, as in *gld-1* FBEa), the R364 side chain intimately interacts with C5, forming hydrogen bonds with the N_3 and O_2 atoms on the Watson–Crick edge of C5 (Fig. 3*A*). In the second class (G4U5, as in the *fem-3* PME), R364 is rotated away from the fifth base so that it does not contact U5 (Fig. 3*B*). In the third class (A4C5, as in the FBE consensus and FBEa G4A mutant), R364 is ≈4 Å away from the Watson–Crick edge of the fifth base, positioned to form mainly van der Waals contacts with the base (Fig. 3*C*). A4A5, as found in the *fog-1* FBEa, is structurally similar to the third class.

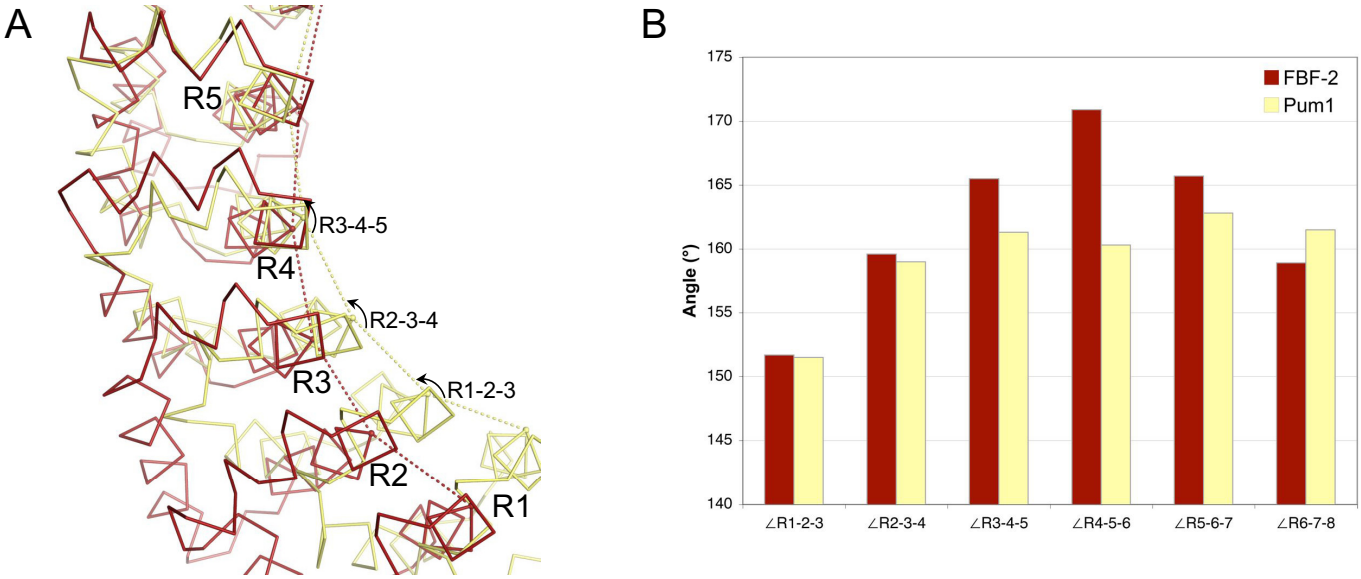

Fig. S6. FBF-2 affinity for *gld-1* FBEa RNA target. (*A*) Representative electrophoretic mobility shift assay of FBF-2 binding to *gld-1* FBEa RNA sequence (UGUGCCAUA). (*B*) Representative analysis of binding data for FBF-2 binding to *gld-1* FBEa RNA sequence.

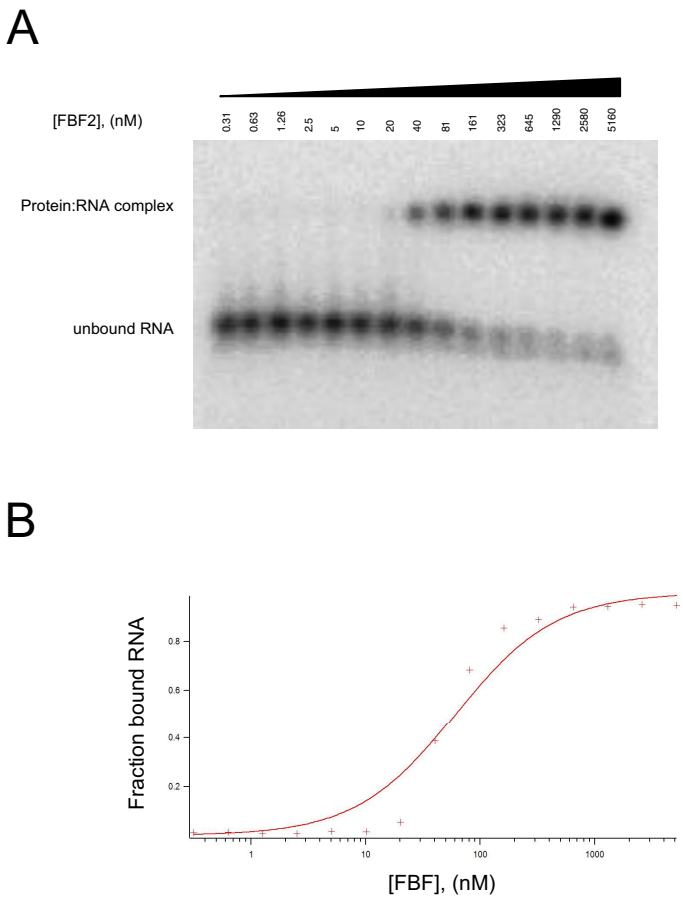

Fig. S7. Curvature of the RNA-binding surface of FBF-2 is flatter than that of PUM1. (*A*) CA traces of FBF-2 (red) and PUM1 (yellow) aligned over repeats 5–8. Angles formed between the RNA-binding helices of FBF-2 and PUM1 are defined by calculating the angle between the CA atoms of residues that would typically form base-stacking interactions (small spheres). (*B*) Plot of the repeat-to-repeat angles of FBF-2 and PUM1.

Table S1. Data collection, phasing, and refinement statistics

RNA	FBE (Se-SAD)	FBE	*gld-1* FBEa	*fem-3* PME	*gld-1* FBEb	FBEa G4A mutant	*fog-1* FBEa
Data collection							
Space group	P61	P61	P61	P61	P61	P61	P61
Cell dimensions							
a, *b*, *c*; Å	98.266, 98.266, 102.714	98.806, 98.806, 103.309	97.226, 97.226, 101.275	96.675, 96.675, 101.542	96.620, 96.620, 101.678	97.048, 97.048, 101.703	97.173, 97.173, 102.305
α , β , γ ; °	90.0, 90.0, 120.0	90.0, 90.0, 120.0	90.0, 90.0, 120.0	90.0, 90.0, 120.0	90.0, 90.0, 120.0	90.0, 90.0, 120.0	90.0, 90.0, 120.0
Resolution, Å	50–2.1 (2.14–2.10)	50–2.2 (2.24–2.20)	50–2.3 (2.34–2.30)	50–2.0 (2.03–2.00)	50–1.9 (1.93–1.90)	50–2.4 (2.44–2.40)	50–2.2 (2.25–2.21)
R_{sym}	0.182 (0.369)	0.097 (0.789)	0.112 (0.75)	0.118 (0.838)	0.083 (0.807)	0.128 (0.895)	0.117 (0.895)
I / σI	20.6 (10.6)	21.4 (2.9)	22.0 (4.4)	23.2 (3.7)	28.4 (2.9)	18.2 (3.7)	19.7 (2.6)
Completeness, %	100 (100)	99.9 (100)	100 (100)	99.9 (100)	99.6 (93.5)	99.8 (99.4)	99.8 (96.8)
Redundancy, %	22.7 (19.0)	8.2 (8.0)	11.5 (11.5)	11.5 (11.4)	11.2 (8.1)	8.7 (8.6)	11.4 (9.4)
Refinement							
Resolution, Å		2.2	2.3	2.0	1.9	2.4	2.2
No. reflections		332,626	280,441	423,127	471,989	184,949	313,274
R_{work}/R_{free}		0.159/0.203	0.169/0.218	0.154/0.193	0.161/0.200	0.162/0.232	0.176/0.216
No. atoms							
Protein		3,183	3,183	3,183	3,183	3,141	3,132
Ligand/ion		189	187	185	185	186	188
Water		239	94	475	387	181	124
B-factors							
Protein		40.5	53.7	27.4	30.9	40.5	47.3
Ligand/ion		46.4	66.0	34.1	45.0	49.9	63.1
Water		42.0	48.9	41.5	41.5	43.4	43.1
rmsd							
Bond lengths, Å		0.004	0.006	0.004	0.004	0.006	0.004
Bond angles, °		0.602	0.678	0.590	0.616	0.691	0.617

Values in parentheses are for the highest-resolution shell.