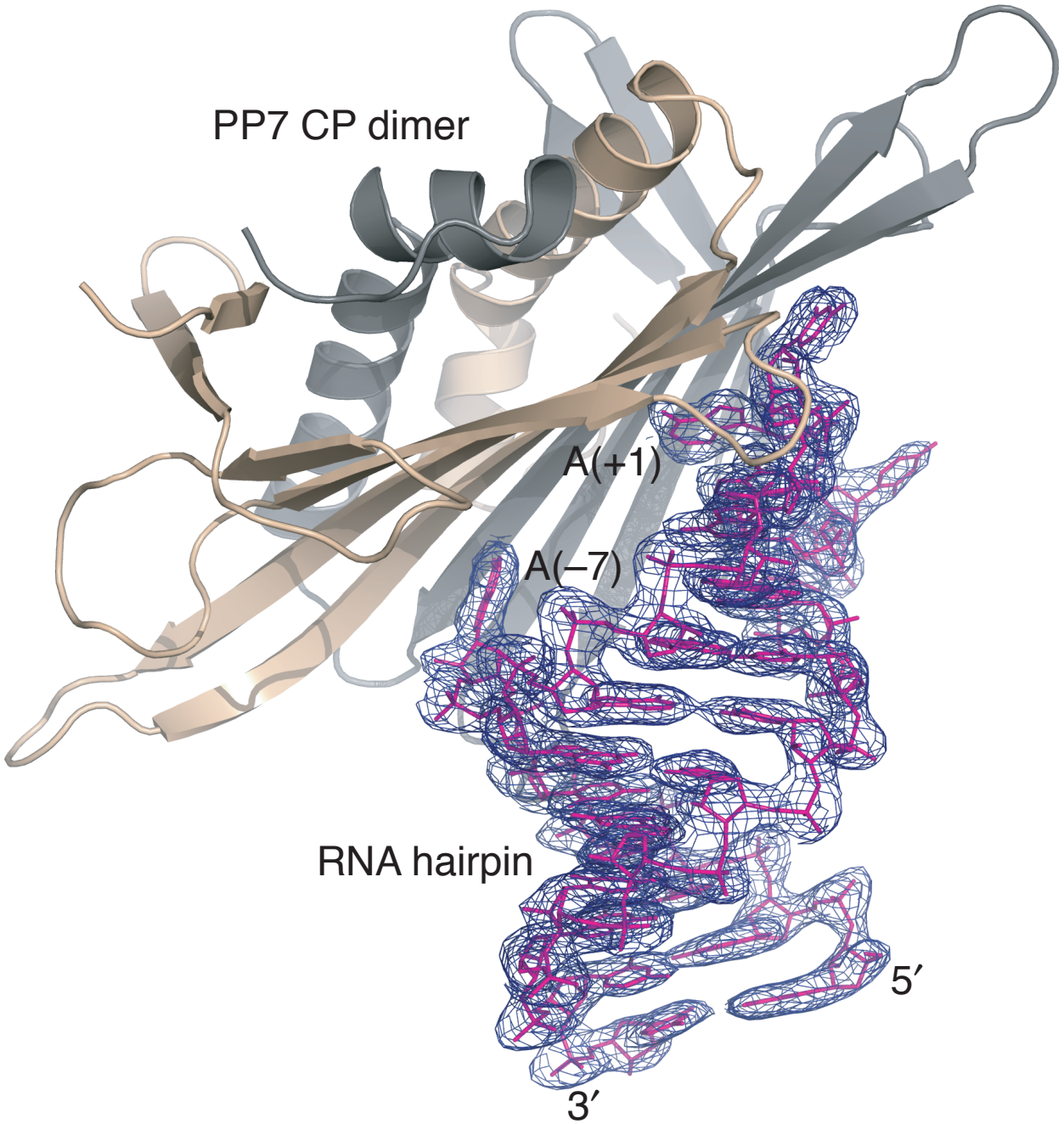
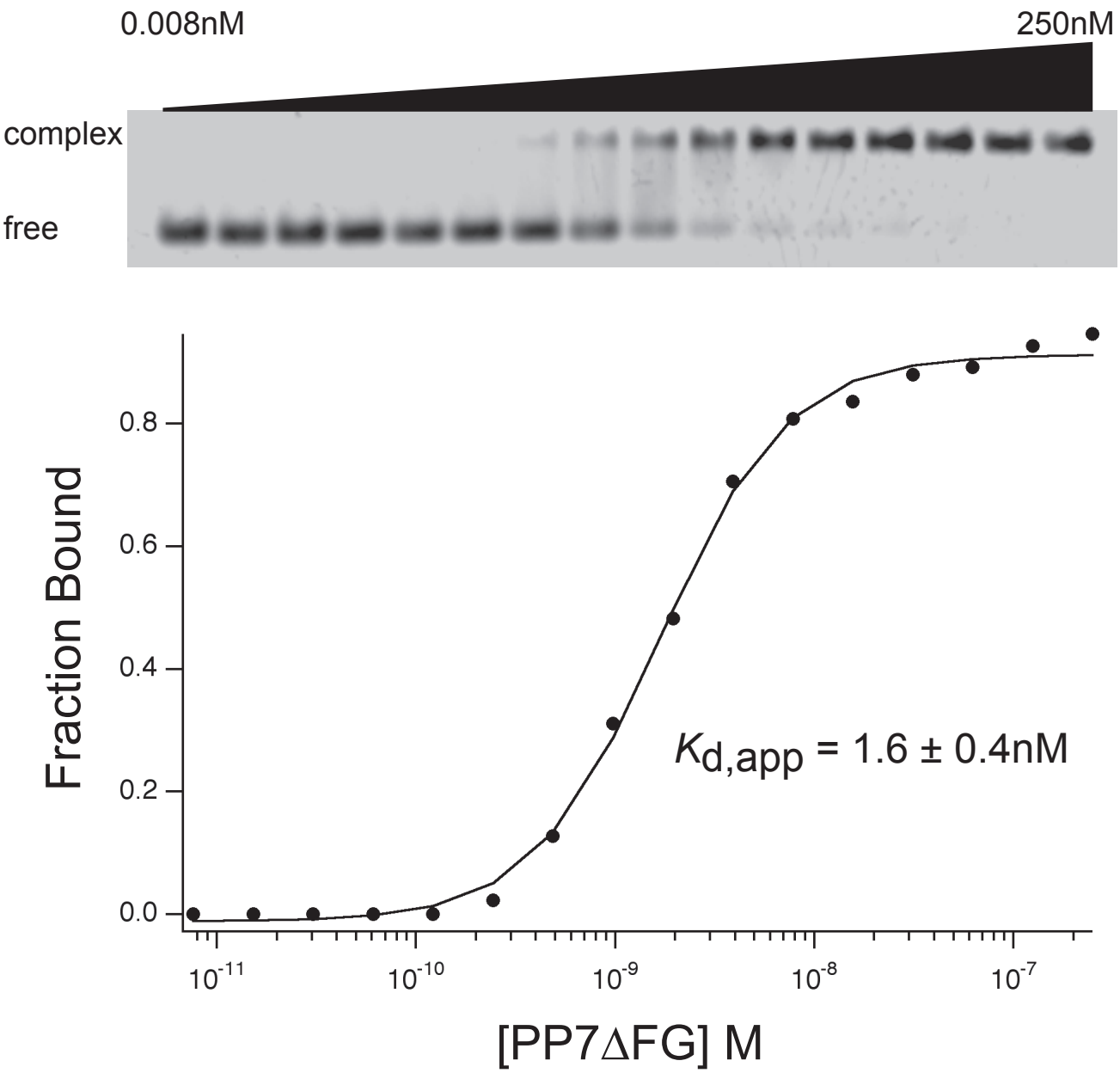
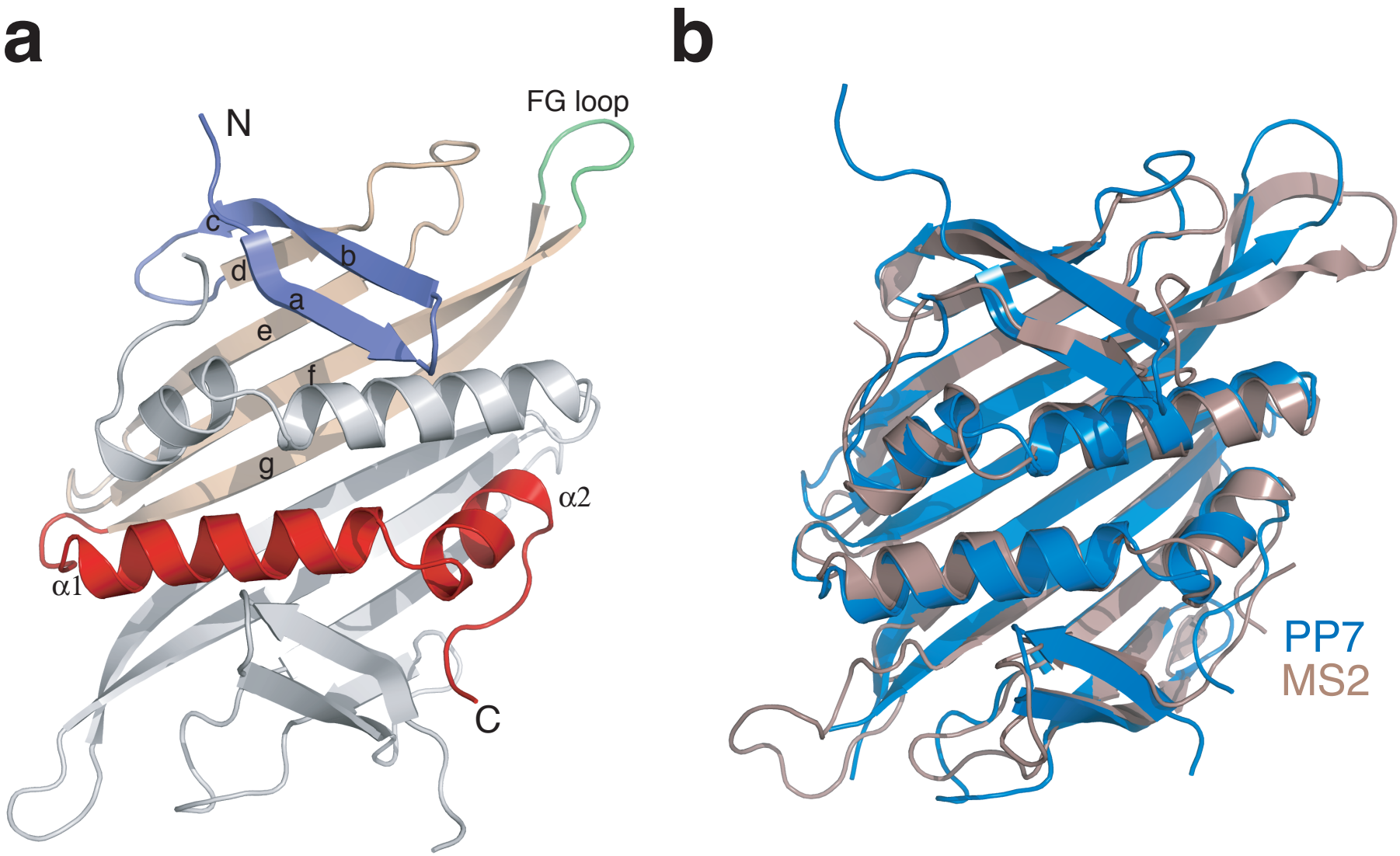

PP7 CP dimer
A(+1)
A(–7)
RNA hairpin
5′
3′

Supplemental Figure 1. Sample electron density of RNA hairpin. Cartoon showing PP7ΔFG CP (wheat) bound to RNA hairpin (violet) showing 2Fo-Fc electron density (blue) calculated at 1σ.


0.008nM
250nM
complex
free
Fraction Bound
0.0
0.2
0.4
0.6
0.8
10^{-11}
10^{-10}
10^{-9}
10^{-8}
10^{-7}
[PP7ΔFG] M
$K_{d,app}$ = 1.6 ± 0.4nM

Supplemental Figure 2. PP7ΔFG CP binds its translational operator with high affinity. Representative electrophoretic mobility shift assay shown with free 27nt RNA hairpin and RNA-protein complex labeled. Filled triangle in top panel represents PP7ΔFG CP concentration increasing by a factor of two per lane. A plot of the fraction of RNA bound as a function of PP7ΔFG concentration and its fit to the Hill equation is presented for the representative data. The $K_{d,app}$ based upon three independent measurements is 1.26±0.38nM.


a
N
FG loop
c
b
d
a
e
f
g
α2
α1
C
b
PP7
MS2

Supplemental Figure 3. Structure of PP7ΔFG CP dimer and its superposition with MS2 W82R CP. (a) Cartoon of PP7ΔFG CP dimer showing N-terminal β-hairpin (blue), five stranded β-sheet (wheat), truncated FG loop (green) and C-terminal α-helices (red) of one protomer with the other protomer shown in grey. (b) PP7ΔFG and MS2 W82R (1MSC) CPs have similar folds and superpose with an r.m.s.d. of 2.3Å for all C_α atoms.

Supplemental Table 1. Data collection and refinement statistics

	PP7ΔFG	PP7ΔFG-RNA
Data collection		
Space group	H3	C2
Cell dimensions		
a, *b*, *c* (Å)	155.0, 155.0, 31.6	175.0, 145.4, 109.7
α, β, γ (°)	90, 90, 120	90, 122.9, 90
Resolution (Å)	38.75-1.60 (1.66-1.60)	33.80-2.44 (2.54-2.44)
R_{sym} or R_{merge}	0.032 (0.475)	0.065 (0.450)
$I / \sigma I$	36.8 (2.1)	18.7 (2.3)
Completeness (%)	100.0 (100.0)	98.6 (87.6)
Redundancy	4.3 (3.9)	4.3 (4.0)
Refinement		
Resolution (Å)	38.75-1.60	33.80-2.44
No. reflections	35,670	82,206
R_{work} / R_{free}	15.7/19.7	17.5/23.6
No. atoms		
Protein	2,005	11,222
RNA	NA	3,198
Water	352	512
B-factors		
Protein	28.76	53.55
RNA	NA	77.96
Water	40.00	46.35
R.m.s. deviations		
Bond lengths (Å)	0.012	0.011
Bond angles (°)	1.51	1.59

*Values in parentheses are for highest-resolution shell.

Structural basis for the co-evolution of a viral RNA-protein complex

Jeffrey A. Chao, Yury Patskovsky, Steven C. Almo, Robert H. Singer

Supplemental Data

Materials and Methods

Protein and RNA preparation

The pP7CT plasmid containing the full length PP7 coat protein gene was provided by D. Peabody (Albuquerque, New Mexico). Similar to the MS2 CP, the PP7 CP forms an obligate dimer that is the building block of the bacteriophage capsid and also responsible for translational repression of the replicase by binding an RNA stem-loop that contains the replicase start codon[1]. These two functions can be separated for the MS2 CP because the FG loop, located between the sixth and seventh β-strands, that stabilizes capsid formation is not involved in RNA recognition[2-4]. A PP7ΔFG truncation was constructed by PCR using the structure of the PP7 capsid as a guide that replaced loop residues 67-74 with a BspEI restriction site (Ser-Gly) [5]. The excised region includes the two cysteines that form inter-molecular disulphide bonds and covalently link the dimers at the capsid fivefold and quasi-sixfold axes[5].

This construct was cloned into pET22HT as an N-terminal His_6-tag followed by a TEV cleavage site[6]. Recombinant protein was expressed in *E. coli* strain BL21(DE3) grown at 37°C by induction with 1mM IPTG at A600 = 0.7. Cell pellets were resuspended in lysis buffer (50mM sodium phosphate, 1M NaCl, 10mM imidazole, 10mM β-mercaptoethanol, pH 8.0) containing one Complete EDTA-free protease inhibitor tablet (Roche) and lysed by sonication. Cell debris was removed by centrifugation at 15,000g for 20 minutes at 4°C and the soluble His_6-tagged protein was incubated with HIS-Select affinity gel (Sigma) equilibrated with lysis buffer for 1 hour at 4°C. The affinity gel was extensively washed with lysis buffer and protein was eluted with lysis buffer supplemented with 250mM imidazole. Elution fractions containing

His_6- PP7ΔFG were pooled and dialyzed at 4°C against 25mM HEPES (pH 7.0), 25mM NaCl, 1mM DTT, 1mM EDTA. During dialysis, the His_6-tag was removed by incubation with TEV protease leaving a residual Gly-Gly-Ser at the N-terminus of the PP7ΔFG construct. PP7ΔFG was further purified by cation exchange chromatography using a HiTrap SP HP column (GE). The resulting protein was dialyzed against 25mM HEPES (pH 7.0), 25mM NaCl, 1mM EDTA and concentrated to 1.3mM. Protein concentration was calculated by measuring the absorbance at 280nm using an extinction coefficient of 9970 M^{-1} cm^{-1} (0.733 mg ml^{-1}) determined using the ProtParam server[7].

RNAs used for crystallization and binding studies were prepared by chemical synthesis (Dharmacon) and purified by denaturing PAGE, deprotected, lyophilized, and stored according to the manufacturer's protocol. RNAs were resuspended in 25mM HEPES (pH 7.0), 25mM NaCl, 1mM EDTA. Annealing of the RNAs into hairpins was performed by incubation at 90ºC for 1 min followed by snap cooling of the RNA on ice for 30 minutes.

Crystallization and structure determination of PP7ΔFG construct

The free PP7ΔFG protein was crystallized using hanging-drop vapor diffusion at 22°C by mixing equal volumes of the protein and reservoir solution (1.1 M Na Malonate pH 7.0, 0.1 M HEPES pH 7.0, 0.5% v/v Jeffamine ED-2001 pH 7.0). Crystals were cryoprotected by soaking them in the reservoir solution supplemented with 25% glycerol before flash cooling in liquid nitrogen. Data were collected to 1.60Å resolution from a single crystal at the National Synchotron Light Source X29a beamline (Brookhaven, NY) at a wavelength of 1.008Å. The diffraction data were indexed, integrated and scaled with HKL2000[8].

The structure of PP7ΔFG was determined by molecular replacement with Phaser using a monomer from the previously determined structure of PP7 bacteriophage capsid (1DWN) as a search model[5,9]. The crystal contained two copies of the monomer corresponding to the expected dimer in the asymmetric unit. Rounds of refinement and model building were carried out with REFMAC5 and Coot[10,11]. Non-crystallographic symmetry (NCS) restraints were applied during initial stages of refinement. Protein stereochemistry was checked using Molprobity[12]. The final model contains residues –3-127 (chain A) and 1-65:76-127 (chain B) using the residue numbering described for 1DWN with the GGS cloning artifact and initial methionine numbered as –3-0.[5]. The non-native Ser and Gly generated during truncation of the FG loop are numbered as 67 and 68.

Crystallization and structure determination of PP7ΔFG-RNA complex construct

The PP7ΔFG-RNA complex was formed by equilibrating the protein and RNA hairpin (5′-GGCACAGAAGAUAUGGCUUCGUGCC-3′) at a 2:1.1 molar ratio on ice for 30 minutes prior to crystallization trials. The final concentration of the complex used for crystallization was approximately 250μM. The number of base pairs in the lower helix of the RNA stem-loop was systematically varied and only the RNA construct exactly 25nts in length resulted in crystals of the complex. Crystals were obtained using sitting-drop vapor diffusion at 22°C by mixing equal volumes of the complex and reservoir solution (15% PEG 3350, 0.1M Na Citrate pH 5.5, 10% Hampton Crystal Screen 2 #25 (0.1M MES pH 6.5, 0.01M Cobaltous chloride hexahydrate, 1.8M Ammonium sulfate). Crystals were cryoprotected by soaking them in the reservoir solution supplemented with 30% glycerol before flash cooling in liquid nitrogen. Data

were collected to 2.44Å resolution from a single crystal as part of the "Express Crystallography" service at the Advanced Photon Source SGX-CAT beamline (Argonne, IL) at a wavelength of 0.9797Å. The diffraction data were indexed, integrated and scaled with HKL2000[8].

The structure of the PP7ΔFG-RNA complex was determined by molecular replacement with Phaser using the PP7ΔFG dimer as a search model[9]. The crystal contained six copies of the RNA-protein complex in the asymmetric unit. Initial electron density for the RNA after molecular replacement was of excellent quality and allowed unambiguous placement of all 25 nucleotides using X-fit[13]. The RNA stems were found to stack on the two-fold axis explaining the RNA length dependence of crystallization. Early rounds of simulated-annealing refinement and model building were performed using CNS and X-fit[13,14]. After all six RNA stem-loops were built, refinement and model building was continued using REFMAC5 and Coot[10,11]. Tight NCS restraints were applied initially and were relaxed during iterative rounds of refinement. TLS groups were defined for all twelve PP7ΔFG monomers and the six RNA stem-loops. Protein stereochemistry and RNA geometry was checked using Molprobity[12]. The final model contains residues –1-127 (chain A), 1-127 (chain B), 1-25 (chain C), 0-127 (chain D), 0-127 (chain E), 1-25 (chain F), -1-127 (chain G), 1-21:24-127 (chain H), 1-25 (chain I), -1-127 (chain J), -1-65:67-127 (chain K), 1-25 (chain L), 3-127 (chain M), 1-21:24-127 (chain N), 1-25 (chain O), 0-127 (chain P), 1-21:23-64:68-127 (chain Q) and 1-25 (chain R).

Electrophoretic mobility shift assay

The complex between PP7ΔFG and a 27nt RNA hairpin (5′-GCGCACAGAAGAUAUGGCUUCGUGCGC-3′) was monitored using an electrophoretic gel mobility shift assay. Fluorescein 5-thiosemicarbazide was used to 3′-end label the hairpin using a method adapted from Reines and Cantor[15,16]. The RNA hairpin (100 pM) was equilibrated with a two-fold serial dilution of PP7ΔFG (250 nM to 0.008 nM). The complex was equilibrated in buffer containing 10mM Tris pH 7.5, 25 mM NaCl, 0.1 mM EDTA, 0.01 mg ml^{-1} tRNA, 50 μg ml^{-1} heparin and 0.01% (v/v) IGEPAL CA630 for 2 hours. Prior to loading, a 6X loading dye (30% (v/v) glycerol and 0.01% (w/v) bromocresol green was added to each binding reaction. Protein-RNA complexes were resolved from free RNA by native polyacrylamide gel electrophoresis (6% (w/v) 29:1 acrylamide/Bis-acrylamide, 0.5x TBE) at 22°C. Gels were run at 120V for 35 minutes and then scanned using a fluor-imager (GE, Typhoon 9400) with a blue laser at 488 nm for excitation. The fraction of bound RNA hairpin was determined and fit to a modified version of the Hill equation as previously described[17].

References

1. Witherell, G.W., Gott, J.M. & Uhlenbeck, O.C. Specific interaction between RNA phage coat proteins and RNA. *Prog Nucleic Acid Res Mol Biol* 40, 185-220 (1991).
2. Peabody, D.S. & Ely, K.R. Control of translational repression by protein-protein interactions. *Nucleic Acids Res* 20, 1649-55 (1992).
3. Ni, C.Z. et al. Crystal structure of the MS2 coat protein dimer: implications for RNA binding and virus assembly. *Structure* 3, 255-63 (1995).
4. LeCuyer, K.A., Behlen, L.S. & Uhlenbeck, O.C. Mutants of the bacteriophage MS2 coat protein that alter its cooperative binding to RNA. *Biochemistry* 34, 10600-6 (1995).
5. Tars, K., Fridborg, K., Bundule, M. & Liljas, L. The three-dimensional structure of bacteriophage PP7 from Pseudomonas aeruginosa at 3.7-A resolution. *Virology* 272, 331-7 (2000).

6. Chao, J.A. et al. Dual modes of RNA-silencing suppression by Flock House virus protein B2. *Nat Struct Mol Biol* 12, 952-7 (2005).
7. Gasteiger, E. et al. (eds.). *Protein Identification and Analysis Tools on the ExPASy Server*, 571-607 (Humana Press, 2005).
8. Otwinowski, Z. & Minor, W. Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods in Enzymology* 276, 307-26 (1997).
9. Read, R.J. Pushing the boundaries of molecular replacement with maximum likelihood. *Acta Crystallogr D Biol Crystallogr* 57, 1373-82 (2001).
10. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60, 2126-32 (2004).
11. Winn, M.D., Isupov, M.N. & Murshudov, G.N. Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr D Biol Crystallogr* 57, 122-33 (2001).
12. Davis, I.W. et al. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* (2007).
13. McRee, D.E. XtalView/Xfit--A versatile program for manipulating atomic coordinates and electron density. *J Struct Biol* 125, 156-65. (1999).
14. Brunger, A.T. et al. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* 54, 905-21. (1998).
15. Reines, S.A. & Cantor, C.R. New fluorescent hydrazide reagents for the oxidized 3'-terminus of RNA. *Nucleic Acids Res* 1, 767-86 (1974).
16. Pagano, J.M., Farley, B.M., McCoig, L.M. & Ryder, S.P. Molecular basis of RNA recognition by the embryonic polarity determinant MEX-5. *J Biol Chem* 282, 8883-94 (2007).
17. Recht, M.I. & Williamson, J.R. Central domain assembly: thermodynamics and kinetics of S6 and S18 binding to an S15-RNA complex. *J Mol Biol* 313, 35-48 (2001).