Supplemental Materials

GIGYF1/2 proteins use auxiliary sequences to selectively bind to 4EHP and repress target mRNA expression

Daniel Peter, Ramona Weber, Felix Sandmeir, Lara Wohlbold, Sigrun Helms, Praveen Bawankar, Eugene Valkov, Cátia Igreja and Elisa Izaurralde

DNA constructs

The plasmids used for the expression of human eIF4E, 4E-BP1 and eIF4G1 (full-length or fragments) in *Escherichia coli* or in human cells have been previously described (Peter et al. 2015a; Grüner et al. 2016). The plasmids for the expression of 4EHP fragments (M1–F234) and (A52–F234) in *E. coli* were obtained by inserting the corresponding cDNA fragments either into the pnYC-NpH (between the XhoI and NheI restriction sites) or the pnYC-CvH (between the XhoI and BamHI restriction sites) vectors that include N- and C-terminal His_6 tags (Diebold et al. 2011), respectively. DNA fragments encoding for GYF1 [residues K33–K52 (C), K33–D71 (C+L+NC) and K33–M103 C+L+NC+A)] and GYF2 [residues K35–K54 (C), K35–Q72 (C+L+NC) and K35–T105 (C+L+NC+A)] were inserted into the NdeI-NheI and NdeI-XbaI restriction sites in the pnEA-NpM vector (Diebold et al. 2011), respectively. These constructs express GYF fragments that are N-terminally fused to an MBP-tag, which is cleavable by the HRV3C protease. A DNA fragment encoding the B1 domain of immunoglobulin-binding protein G (GB1; Cheng and Patel 2004) was inserted C-terminally to the GYF fragments by site-directed mutagenesis using the QuikChange mutagenesis kit (Stratagene).

The plasmids for the expression of V5-streptavidin binding protein (SBP)-tagged and λN-hemagglutinin (HA)-tagged 4EHP in human cells were obtained by inserting the full-length 4EHP cDNA into the XhoI and BamHI sites in the pT7-V5-SBP and pλN-HA-C1 vectors (Kuzuoglu-Ozturk et al. 2016), respectively. The plasmids for the expression of full-length GYF1 and GYF2, which are N-terminally fused to GFP or MS2-HA were obtained by inserting the GYF1 cDNA (XhoI-EcoRI, obtained from the Kazusa DNA Research Institute; sj03926) or the GYF2 cDNA (XhoI-BamHI) into the corresponding sites of the pT7-EGFP-C1 and pT7-MS2-HA-C1 vectors. cDNA fragments encoding for GYF1 (residues M1–C177) and GYF2 (residues M1–P180) were introduced into the XhoI and BamHI restriction sites in

the pT7-EGFP-C1 vector. The cDNA encoding 4E-T (eIF4E-Transporter protein; EIF4ENIF1) was inserted into the HindIII and BamHI restriction sites in the pT7-EGFP-C1 vector. The cDNA encoding TTP (Tristetraprolin, residues M1–P313; Fabian et al. 2013) was inserted between the XhoI and EcoRI restriction sites of the pλN-HA-C1 vector. To generate a reporter containing the ARE-element (pCIneo-R-Luc-ARE-A_{90}-MALAT1), the sequence of the ARE element present in the 3' UTR of the TNF (Tumor Necrosis Factor)-α mRNA was inserted twice into the 3' UTR of the pCIneo-R-Luc parental plasmid by site-directed mutagenesis. A cDNA containing a stretch of 90 A and the MALAT1 sequence was then inserted into the XhoI and NotI restriction sites of the R-Luc-ARE vector. The DNA sequence of the TNF-α ARE is as follows: TTATTTATTATTTATTTATTATTTATTTATTT. All of the mutants used in this study were generated by site-directed mutagenesis using the QuikChange mutagenesis kit (Stratagene). All of the constructs and mutations were confirmed by sequencing and are listed in Supplemental Table S1.

Pulldown and competition assays

In the pulldown assays shown in Fig. 1 and Supplemental Fig. S2 and S7, bacterial lysates expressing recombinant 4EHP-His_6 (residues M1–F234, wild-type and mutants) or purified eIF4E-His_6 (2 μM; 50 μg) were incubated with Ni-NTA beads for 30 min. The immobilized 4EHP and eIF4E proteins were then incubated for 30 min with bacterial lysates expressing GYF1, GYF2 or 4E-BP1 fragments (wild-type and mutants) that were N-terminally tagged with MBP and C-terminally tagged with GB1. Proteins associated with 4EHP or eIF4E were eluted with imidazole and analyzed by SDS-PAGE followed by Coomassie staining.

For the competition assays, purified 4EHP–4EBP1 complexes (2 μM) containing 4EHP (residues A52–F234)-His_6 and 4E-BP1 C+L+NC (residues R50–S83; with a C-terminal GB1

tag) were incubated with Ni-NTA beads for 30 min in 50 mM Na-phosphate (pH 7.0) and 200 mM NaCl. The immobilized complexes were then incubated with equimolar amount of purified, GB1-tagged competitor peptides or with MBP as a negative control. After the specified time points, the beads were pelleted and washed three times in the same buffer. Proteins bound to the Ni-NTA beads were eluted with the same buffer containing 500 mM imidazole and analyzed by SDS-PAGE followed by Coomassie staining. The amount of 4E-BP1 bound to 4EHP was quantified using the ImageJ software and normalized to 4EHP levels present at each time point. These values were set to 100 in the presence of MBP. Data points from at least three independent experiments were plotted and the resulting fitting curves were determined using the Levenberg-Marquardt algorithm for single exponential decay functions. The R^2 values associated with the fitting of the exponential decay curves were between 0.82 and 0.96.

ITC analysis

For the ITC measurements, the GB1-stabilized GYF1/2 peptides (wild-type and mutants) and 4E-BP1 peptides were purified as previously described for the other 4E-BPs (Igreja et al. 2014; Peter et al. 2015a,b). The 4EHP protein (residues A52–F234; wild-type and mutants) used in the ITC measurements was expressed with an N-terminal His_6 tag and purified from cleared cell lysates using a nickel column (HisTrap HP 5 ml, GE Healthcare). The His_6 tag was cleaved by HRV3C protease overnight at 4°C. The protein was further purified using a heparin column (HiTrap Heparin HP 5 ml, GE Healthcare) and a final purification on a Superdex 75 column (GE Healthcare). The 4EHP-GYF2 complex (GYF2 residues K35–T105) used for measuring the affinity for m^7GpppG cap analog was purified as described for the 4EHP–GYF2 complex used for crystallization. All of the proteins used in the ITC

measurements were stored at -80°C in a buffer consisting of 20 mM Na-phosphate (pH 7.0) and 200 mM NaCl.

The ITC experiments were performed using a VP-ITC microcalorimeter (MicroCal) at 20°C as described previously (Igreja et al. 2014; Peter et al. 2015a,b). A solution containing either 4EHP (residues A52–F234, wild-type, S99N mutant and dimerization mutant, 1-5 μM) or eIF4E (residues K36–V217, 5 μM) in a calorimetric cell was titrated with tenfold concentrated solutions of GB1-stabilized peptides that were dissolved in the same buffer (20 mM Na-phosphate (pH 7.0) and 150 mM NaCl). The following peptides were used: GYF1 (C, residues K33–K52, 50 μM; C+L+NC, residues K33–D71, 20 μM; C+L+NC+A wild type or dimerization mutant, residues K33–M103, 10 μM), GYF2 (C, residues K35–K54, 50 μM; C+L+NC, residues K35–Q72, 20 μM; C+L+NC+A wild type or dimerization mutant, residues K35–T105, 10 μM) and 4E-BP1 (C+L+NC, residues T50–S83, 50 μM). The affinity for the m^7GpppG cap analog was measured in a buffer containing 20 mM HEPES (pH 7.5) and 200 mM NaCl by titrating a solution of m^7GpppG (400 μM; New England Biolabs) into a solution of 4EHP (residues A52–F234, 40 μM) or 4EHP in complex with GIGYF2 (residues K35–T105, 40 μM) diluted in the same buffer.

The titration experiments consisted of an initial injection of 2 μl followed by 28 injections of 10 μl at 240 s intervals. Each binding experiment was repeated three times. Correction for dilution heating and mixing was achieved by subtracting the final baseline, which consisted of small peaks of similar size. The thermodynamic parameters were estimated using a one-site binding model (Origin version 7.0), whereby the data points for the first injection were removed from the analysis (Mizoue and Tellinghuisen 2004). Because the protein concentration used in these measurements is low (1 μM for 4EHP in the calorimetric cell), dimerization of the 4EHP–GYF1/2 complexes is unlikely to occur and thus it does not contribute to the measured binding constants. Accordingly, GYF2 and 4EHP dimerization

mutants still display a binding affinity similar to that observed for the complexes containing the wild type proteins (Supplemental Table S2 and Fig. S7).

Crystallization

Crystals of 4EHP (residues A52–F234) in complex with GYF1 (residues K33–M103) were obtained at 18°C using the hanging-drop vapor diffusion method two days after mixing the protein solution (16 mg/ml; 1 μl) with the crystallization solution (1 μl) containing 20% PEG 3350 in 0.2 M potassium nitrate. Crystals of 4EHP (residues A52–F234) bound to GYF2 (residues A35–T105) were obtained at 18°C using the hanging-drop vapor diffusion method one day after mixing the protein solution (16 mg/ml, 1 μl) with the crystallization solution (1 μl) containing 0.1 M sodium citrate (pH 5.0), 0.1 M magnesium chloride and 12% PEG 4000.. Crystals of 4EHP (residues A52–F234) in complex with GYF2 (residues A35–Q72) were obtained at 18°C using the hanging-drop vapor diffusion method. Crystals grew in one day after mixing the protein solution (16 mg/ml, 1 μl) with the crystallization solution (1 μl) containing in 0.1 M sodium acetate (pH 4.6) and 0.6 M diammonium phosphate. All of the crystals containing GYF peptides were soaked in mother liquor supplemented with 10–15% glycerol for cryoprotection before flash-freezing in liquid nitrogen.

Crystals of 4EHP (residues A52–F234) in complex with 4E-BP1 (residues T50–S83) were obtained at 18°C using the sitting-drop vapor diffusion method. The crystals grew three days after mixing the protein solution (16.5 mg/ml; 0.1 μl) with the crystallization solution (0.1 μl) containing 0.1 M sodium acetate (pH 4.6) and 1.7 M sodium formate. The crystals were cryoprotected in mother liquor supplemented with 3.5 M sodium formate and flash-frozen in liquid nitrogen.

Data collection and structure determination

The data for all the crystals were collected at 100K on a PILATUS 6M detector at the PXII beamline at the Swiss Light Source. Diffraction data were processed with XDS and scaled using XSCALE (Kabsch 2010). The phases were obtained by molecular replacement using PHASER (McCoy et al., 2007). For the 4EHP–GYF2 (residues K35–T105; C+L+NC+A) complex, the structure of human 4EHP (PDB 2JGB; Rosettani et al. 2007) was used as a search model with an asymmetric unit containing two copies of the model. To solve the structures of the 4EHP–GYF2 (residues K35–Q72; C+L+NC) and 4EHP–4E-BP1 (residues T50–S83; C+L+NC) complexes, two copies of 4EHP from the 4EHP–GYF2 (C+L+NC+A) complex were used as a search model. In the case of the 4EHP–GYF1 (residues K33–M103; C+L+NC+A) complex, four copies of the 4EHP–GYF2 (C+L+NC+A) complex were used as a search model. To minimize model bias, the molecular replacement solutions were used to rebuild the initial models using the PHENIX AutoBuild wizard (Terwilliger et al., 2008). To complete the structure, iterative cycles of model building and refinement were performed with COOT (Emsley et al. 2010) and PHENIX (Afonine et al. 2012), respectively. The GYF2 (C+L+NC+A and C+L+NC) and 4E-BP1 (C+L+NC) peptide chains were manually built into the difference density in COOT and further refined with PHENIX. In the final refinement rounds for the 4EHP–GYF1 (C+L+NC+A) and 4EHP–GYF2 (C+L+NC+A) complexes, translation/libration/screw (TLS) parameters were refined for the peptide chains in addition to the individual B-factors; in the case of the GYF1 complex, non-crystallographic symmetry (NCS) torsional restraints were also used in refinement.

The stereochemical properties for all of the structures were verified with MOLPROBITY (Chen et al. 2010), and structural images were prepared with PyMOL (http://www.pymol.org). The diffraction data and refinement statistics are summarized in Table 1.

Small-angle X-ray scattering (SAXS)

SAXS experiments were conducted at the SWING beamline at the SOLEIL synchrotron. Data collection for the 4EHP–GYF2 complexes was performed in-line with size exclusion chromatography (Superdex 200 Increase 5/150 GL, GE Healthcare) using an Agilent HPLC system in a buffer containing 20 mM HEPES pH 7.5, 200 mM NaCl, and 1 mM TCEP [Tris(2-carboxyethyl)phosphine]. The scattering data were collected at 1 s exposures using an Aviex charge-coupled device detector at a sample-detector distance of 1798 mm and a wavelength of 1.033 Å. Data reduction to absolute units, frame averaging and buffer subtraction were performed using the FOXTROT software (Xenocs, France). Theoretical scattering curves and fitting to the experimental SAXS data was performed using the FoXS software (Schneidman-Duhovny et al. 2013). To ensure protein stability during SAXS data collection, all the 4EHP–GYF2 complexes were measured with a 1.5x molar excess of m^7GpppG cap analog (New England Biolabs) in the protein samples. Therefore, the coordinates of the structures used during the fitting procedures were adjusted such that the 4EHP cap-binding loops were fixed in the bound conformation including the cap analog, which was based on the structure of the m^7GTP-bound 4EHP (PDB 2JGB; Rosettani et al. 2007).

Generation of GYF1/2-null cell line

Two sgRNAs targeting GYF1 and two sgRNAs targeting GYF2 were designed using the DNA 2.0 (ATUM) or CHOPCHOP (http://chopchop.cbu.uib.no) online tools and cloned into the pSpCas9(BB)-2A-Puro (PX459) vector [a gift from F. Zhang, Addgene plasmid 48139; (Ran et al., 2013)]. HEK293T cells were transfected with the sgRNA-Cas9 vectors and selected with puromycin (3 μg ml^{-1}) to obtain stable GYF1/2 knockout cells. To obtain clonal cell lines, single cells were distributed in 96-well plates using serial dilutions. Genomic

DNAs from single clones were isolated using the Wizard SV Genomic DNA Purification System (Promega) and the targeted GYF1 and GYF2 loci were PCR amplified and sequenced to confirm gene editing. For GYF1 we observed two frameshift mutations in exon 7 (4 bp deletion and a 8 bp deletion together with a C->T mutation) targeted by sgGYF1-a and two indels in exon 16 (insertion of 43 bp or 93 bp) produced by sgGYF1-b. These mutations changed the GYF1 reading frame after the respective targeted site and introduced premature STOP codons. One frameshift mutation (16 bp deletion in the first exon removing the start codon) was detected for the GYF2 locus (targeted by sgGYF2-a and sgGYF2-b). This deletion was caused by sgGYF2-b. In contrast, sgGYF2-a did not target the genomic locus as the sequence around this target site is wild-type. The knockouts of GYF1/2 were further confirmed by western blotting. For the GYF2 gene we observe low levels of truncated protein fragments that are consistent with translation initiation at internal AUGs (Figure 6B, lane 4). Taking the GYF2 sequence and the position of the mutations into account, these truncated forms lack the 4EHP-binding region and the expression levels are approximately 10% of wild-type levels. The following guide sequences were used: sgGYF1-a: 5'-GCCAGCGGTCGCCGTCTCGC-3'; sgGYF1-b: GACAAGGACCGGCTCATCGT-3'; sgGYF2-a: 5'- ATTTTGAAAACTCACCATTC-3'; sgGYF2-b: 5'-AATACGGAAAAGAATGGCAG

Supplemental Table S1. Mutants and constructs used in this study.

Protein	Name of the construct	Fragments / mutations	Binding site / motif
***Hs* 4EHP** (isoform 1) O60573-1	4EHP	Full-length (1–245)	
	4EHP ΔC-term	1–234	Δ235–245
	4EHP truncated	52–234	Δ1–51 & 235–245
	W-A	W95A	Dorsal surface
	IM-AA	I85A, M101A	Lateral surface
	WIM-AAA	W95A, I85A, M101A	Dorsal + lateral surface
	RE-LL	R103L, E149L	Auxiliary surface
	WRE-ALL	W95A, R103L, E149L	Dorsal + auxiliary surface
	S99N	S99N	Dorsal surface
	Cap mutant (cap*)	W124A	Cap-binding pocket
	D* (dimer mutant)	Q159S, M161D, R202E	Dimer interface
***Hs* eIF4E** (isoform 1) P06730-1	4E	Full-length (1–217)	
	4E trunc	36–217	
	Cap mutant (cap*)	W102A	Cap-binding pocket
***Hs* GIGYF1** O75420	GYF1	Full-length (1–1035)	
	C+L+NC+A	33–103	Complete 4EHP-binding region
	C+L+NC	33–71	Bipartite 4EHP-binding region
	C	33–52	Canonical 4EHP-binding region
	C*	Y39A, Y41A, M46A, L47A	Canonical
	NC*	L60D, F65D, V68D	Non-canonical
	A1*	P76D, L77A	Auxiliary site 1
	A2*	E86A, N95F	Auxiliary site 2
	NC+A1*	L60D, F65D, V68D, P76D, L77A	Non-canonical + auxiliary site 1
	NC+A2+3*	L60D, F65D, V68D, E86A, N95F	Non-canonical + auxiliary site 2
	1-177		N-terminus
	D* (dimer mutant)	E44A, E45F, Q87A	Dimer interface
***Hs* GIGYF2** (isoform 1) Q6Y7W6-1	GYF2	Full-length (1–1299)	
	C+L+NC+A	35–105	Complete 4EHP-binding region
	C+L+NC	35–72	Bipartite 4EHP-binding region
	C	35–54	Canonical 4EHP-binding region
	C*	Y41A, Y43A, M48A, L49A	Canonical motif
	NC*	L62D, F67D, I70D	Non-canonical
	A1*	P78D, L79A	Auxiliary site 1
	A2*	E88A, N96F	Auxiliary site 2
	NC+A1*	L62D, F67D, I70D, P78D, L79A	Non-canonical + auxiliary site 1
	NC+A2+3*	L62D, F67D, I70D, E88A, N97F	Non-canonical + auxiliary site 2
	1-180		N-terminus
	D* (dimer mutant)	E46A, E47F, Q89A	Dimer interface
***Hs* 4E-BP1** Q13541	4E-BP1	Full-length (1–118)	
	4E-BP1 C+L+NC	50–83	eIF4E-binding region
	C*	Y54A, L59A	Canonical motif
	NC*	L75A, V81A	Non-canonical
	C+NC*	Y54A, L59A, L75A, V81A	Canonical+ non-canonical
***Hs* 4E-T** Q9NRA8	4E-T	Full-length (1–985)	
***Hs* TTP** (1–326) P26651	TTP ΔNIM	1–313	Δ314-326, deletion of the NOT1 interacting motif (NIM)

Supplemental Table S2. Thermodynamic parameters for the interaction of 4EHP and eIF4E with the indicated peptides.

GYF peptides *vs* 4EHP					
GYF protein	**K_D (nM)**	**ΔH (kcal mol^{-1})**	**-TΔS (kcal mol^{-1})**	**ΔG (kcal mol^{-1})**	**Molar ratio**
GYF1 C (33-52)	360 ± 120	-23 ± 3	14.6	-8.7	1.00 ± 0.01
GYF1 C+L+NC (33-71)	12 ± 2	-30 ± 1	19.3	-10.6	1.00 ± 0.01
GYF1 C+L+NC+A (33-103)	0.4 ± 0.2	-37 ± 4	24.4	-12.6	1.00 ± 0.02
GYF2 C (35-54)	290 ± 160	-22 ± 2	13.5	-8.8	1.00 ± 0.01
GYF2 C+L+NC (35-72)	14 ± 1	-23 ± 2	12.3	-10.6	1.00 ± 0.01
GYF2 C+L+NC+A (35-105)	0.3 ± 0.1	-32 ± 1	19.1	-12.8	1.00 ± 0.01
GYF peptides *vs* 4EHP dimerization mutants					
GYF protein	**K_D (nM)**	**ΔH (kcal mol^{-1})**	**-TΔS (kcal mol^{-1})**	**ΔG (kcal mol^{-1})**	**Molar ratio**
GYF1 C+L+NC+A (33-103) D*	0.4 ± 0.3	-34 ± 1	21.6	-12.7	1.01 ± 0.01
GYF2 C+L+NC+A (35-105) D*	0.5 ± 0.3	-30.8 ± 0.5	18.1	-12.7	1.01 ± 0.02
4EBP1 C+L+NC *vs* eIF4E or 4EHP					
4E molecule	**K_D (nM)**	**ΔH (kcal mol^{-1})**	**-TΔS (kcal mol^{-1})**	**ΔG (kcal mol^{-1})**	**Molar ratio**
eIF4E	5 ± 2	-18 ± 1	6.4	-11.2	1.00 ± 0.01
4EHP	55 ± 14	-16.4 ± 0.8	6.6	-9.8	1.01 ± 0.01
4EHP S99N	4 ± 1	-21 ± 2	9.7	-11.3	1.01 ± 0.01
m^{7}GpppG cap analog *vs* 4EHP or 4EHP-GIGYF2 complex					
Protein	**K_D (μM)**	**ΔH (kcal mol^{-1})**	**-TΔS (kcal mol^{-1})**	**ΔG (kcal mol^{-1})**	**Molar ratio**
4EHP	4 ± 1	-7.3 ± 0.4	0.1 ± 0.5	-7.2	1.01 ± 0.01
4EHP–GYF2 (35–105) complex	6 ± 3	-9 ± 2	1 ± 3	-7.2	1.01 ± 0.01

Note that the presence of the auxiliary sequences increases the entropic penalty (-TΔS) of the interaction between GYF1/2 and 4EHP compared to that of the peptides lacking these sequences [$\Delta(-T\Delta S)^{GYF1}$ = 5.1 kcal/mol^{-1}, $\Delta(-T\Delta S)^{GYF2}$ = 6.8 kcal/mol^{-1}]. One explanation for the increase in the entropic penalty is a higher disorder-to-order transition for the binding of the GYF1/2 C+L+NC+A peptides compared to the C+L+NC peptides. This is supported by the crystal structures in which the auxiliary sequences fold into two α-helices in complex with 4EHP.

Supplemental Table S3. Experimental and theoretical SAXS parameters for different 4EHP–GYF complexes.

Experimental parameters					
4EHP bound to:	**R_g (Guinier) [Å]**	**R_g (real space) [Å]**	**D_{max} [Å]**	**Exp. I(0) [x10^{-2}]**	**Concentration [mg/ml]**
GYF2 C+L+NC	21.1	21.2	71.1	4	10
GYF2 C+L+NC+A	26.1	26.2	90.3	7.9	10
GYF2 C+L+NC+A	25.8	25.8	89.3	3.6	5
GYF2 C+L+NC+A	24.7	24.8	83.7	1.7	2.5
GYF2 C+L+NC+A	23.7	23.7	80.2	0.7	1.25
4EHP D mutant bound to:	**R_g (Guinier) [Å]**	**R_g (real space) [Å]**	**D_{max} [Å]**	**Exp. I(0) [x10^{-2}]**	**Concentration [mg/ml]**
GYF2 C+L+NC+A D* (dim. mutant)	20.6	20.6	73.7	1.3	5

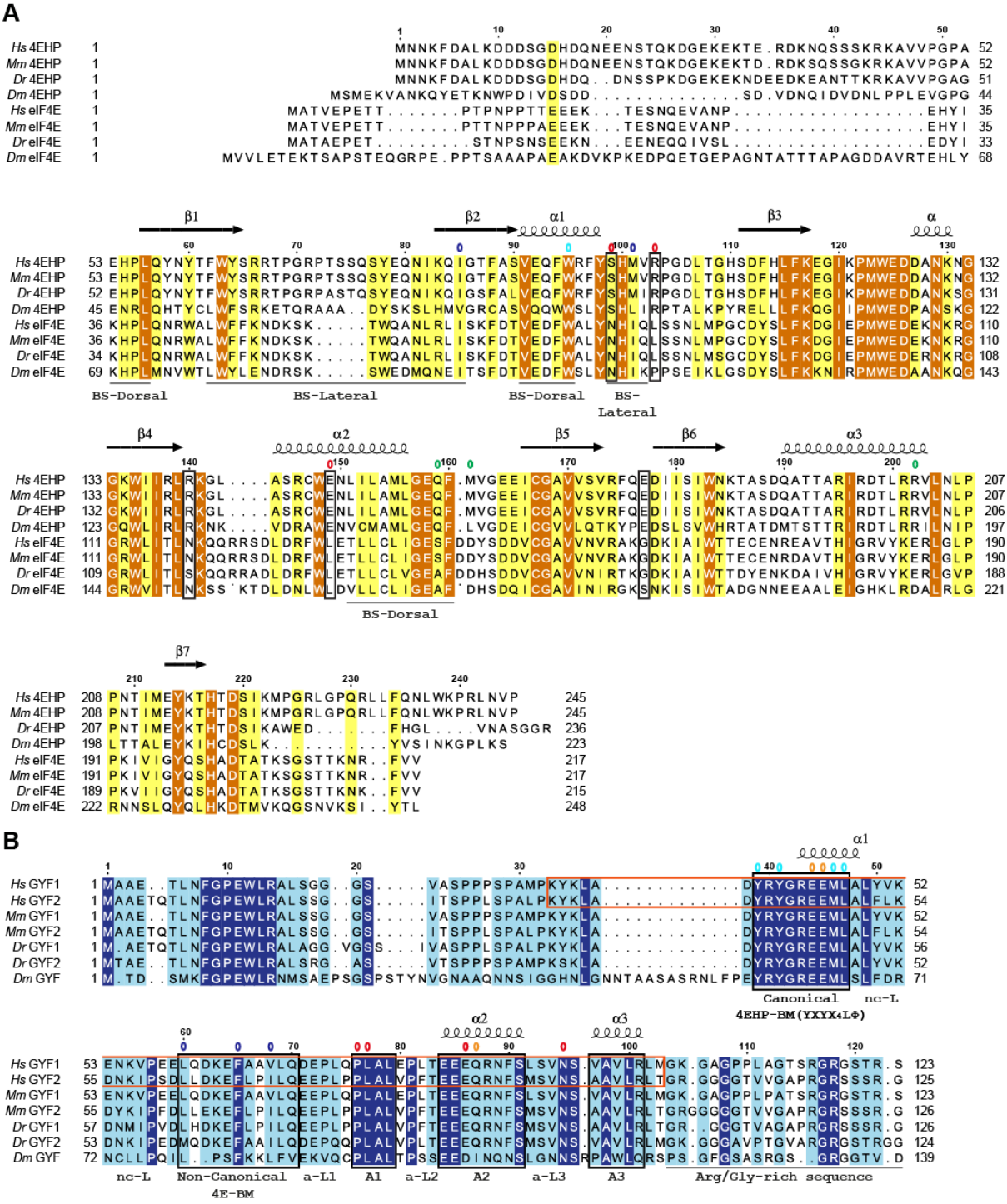
Theoretical parameters			
4EHP–GYF2 C+L+NC+A			
Single complex		**Symmetric dimer**	
R_g (Guinier) [Å]	**D_{max} [Å]**	**R_g (Guinier) [Å]**	**D_{max} [Å]**
17.7	64	24.7	85
4EHP–GYF2 C+L+NC			
Single complex		**Symmetric dimer**	
R_g (Guinier) [Å]	**D_{max} [Å]**	**R_g (Guinier) [Å]**	**D_{max} [Å]**
17.1	64	26.1	90

Supplemental Table S4. Antibodies used in this study.

Antibody	**Source**	**Catalog Number**	**Dilution**	**Monoclonal/ Polyclonal**
Anti-HA-HRP (Western blot)	Roche	12 013 819 001	1:5,000	Monoclonal
Anti-HA (Immunoprecipitation)	Biolegend	MMS-101P	1:1,000	Monoclonal
Anti-*Hs* GYF2	Bethyl laboratories	A303-731A	1:1,000	Rabbit polyclonal
Anti-*Hs* GYF1	Bethyl laboratories	A304-132A-M	1:1,000	Rabbit polyclonal
Anti-*Hs* 4E-T	Abcam	ab95030	1:2,000	Rabbit polyclonal
Anti-*Hs* 4EHP	In house		1:200	Rabbit polyclonal
Anti-*Hs* eIF4E	Bethyl laboratories	A301-154A	1:2,000	Rabbit polyclonal
Anti-*Hs* 4E-BP1	Cell Signaling Technology	9452	1:1,000	Rabbit polyclonal
Anti-GFP	In house		IP	Rabbit polyclonal
Anti-GFP	Roche	11814460001	1:2,000	Monoclonal
Anti-rabbit-HRP	GE Healthcare	NA934V	1:10,000	Polyclonal
Anti-mouse-HRP	GE Healthcare	RPN4201	1:10,000	Polyclonal
Anti-V5	QED Bioscience Inc.	18870	1:5,000	Rabbit polyclonal
Anti-V5	LSBio LifeSpan BioSciences, Inc.	LS-C57305	1:5,000	Monoclonal
Anti-tubulin	Sigma Aldrich	T6199	1:10,000	Monoclonal

Supplemental Figures

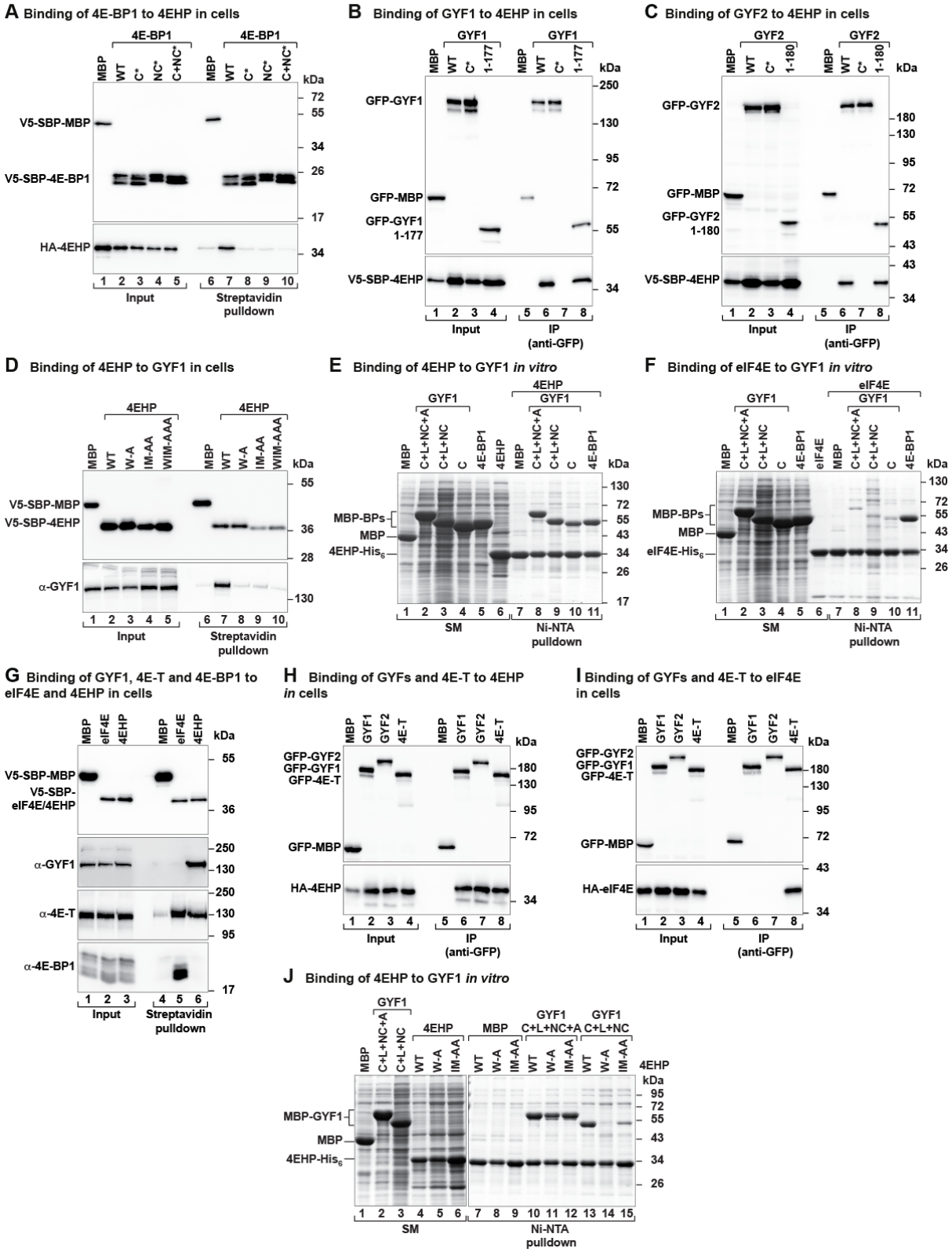
Figure S1



Supplemental Figure S1. Sequence alignments. In all aligned sequences, residues with >70% similarity are shown with a light color background and conserved residues are highlighted with a darker background and printed in white. Secondary structure elements are indicated above the sequences for 4EHP and GYF1 and are based on the structures presented

in this study. (*A*) Sequence alignment of 4EHP and eIF4E orthologous proteins from *Homo sapiens* (*Hs*), *Mus musculus* (*Mm*), *Danio rerio* (*Dr*) and *Drosophila melanogaster* (*Dm*). Residues highlighted in black boxes are specific for 4EHP and are relevant for the interactions described in this study. The dorsal and lateral binding surfaces (BS) of 4EHP are indicated by a line below the sequences. Residues that were mutated in this study are indicated by open circles colored as follows: cyan (dorsal surface), blue (lateral surface), red (4EHP specific residues) and green (dimerization). (*B*) Sequence alignment of GYF proteins. The canonical (C), non-canonical (NC) and auxiliary (A1, A2, A3) sequences are boxed in black. The GYF1/2 sequences visible in the crystal structures are indicated with a red box. Only a short stretch of the Arg/Gly-rich sequence adjacent to the auxiliary motif is shown and underlined. The species are as in A. Open circles above the alignment indicate the residues mutated in this study and are colored as follows: cyan (canonical), blue (non-canonical), red (auxiliary) and orange (dimerization).

Figure S2

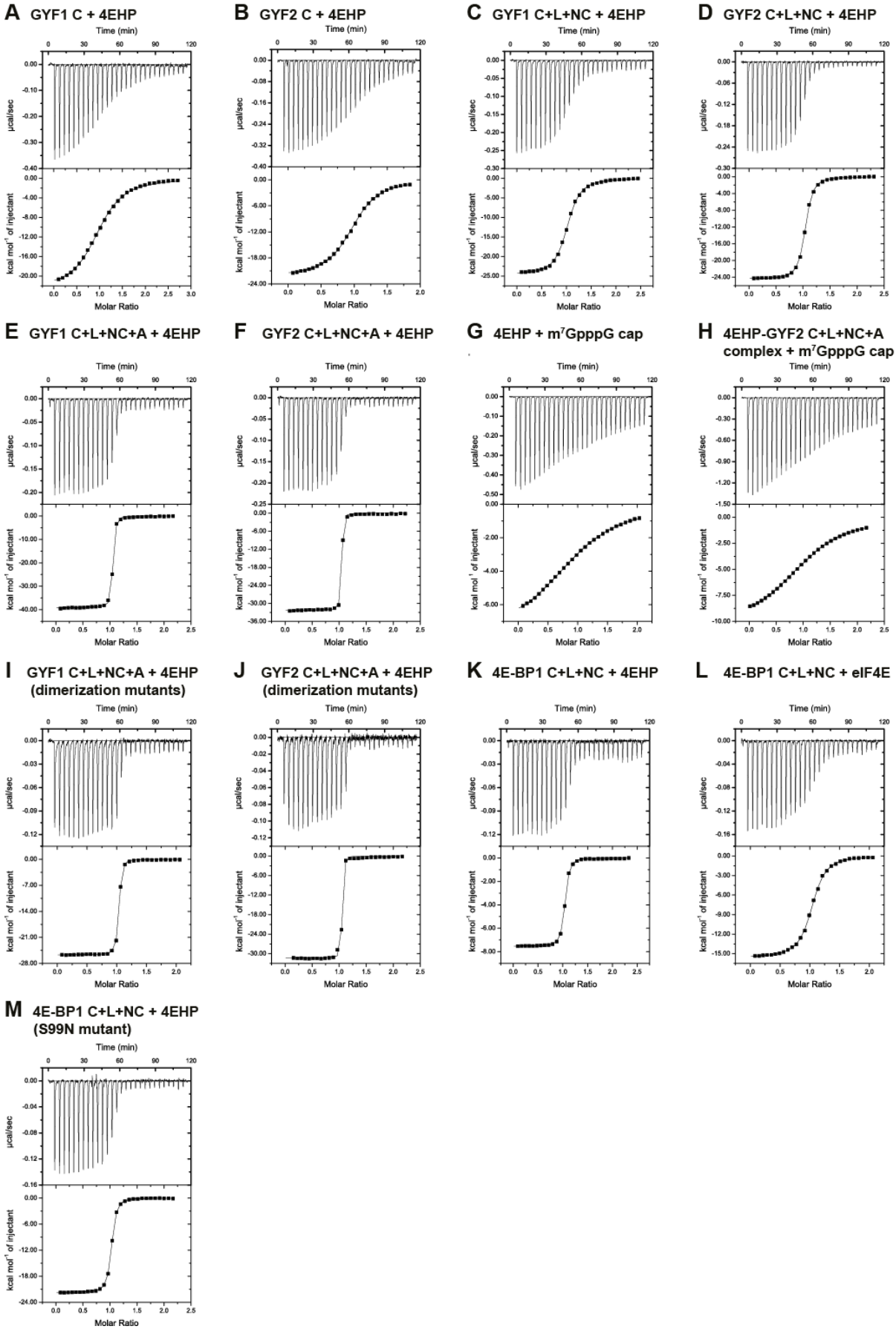


Supplemental Figure S2. Interaction of GYF1, GYF2 and 4E-BP1 with 4EHP. (*A*) The interaction of HA-4EHP with V5-SBP-4E-BP1 (wild-type or the indicated mutants) was tested in HEK293T cell lysates. The proteins were pulled down using streptavidin-coated

beads. V5-SBP-MBP served as negative control. The inputs (1.5%) and bound fractions (3% for the V5-proteins and 5% for HA-4EHP) were analyzed by western blotting using anti-HA and anti-V5 antibodies. (*B*) The interaction of GFP-GYF1 [either full-length, canonical mutant (C*) or N-terminal fragment (residues 1–177)] with V5-SBP-4EHP was analyzed by immunoprecipitation assay in HEK293T cells using anti-GFP antibodies. GFP-MBP served as negative control. The input samples (1.5%) and the immunoprecipitates (10%) were analyzed by western blotting using anti-V5 and anti-GFP antibodies. GYF1 residues 1–177 bound to 4EHP to a similar extent as the full-length protein, indicating that this protein fragment contains the principal 4EHP-binding region of the protein. (*C*) The interaction of GFP-GYF2 [either full-length, canonical mutant (C*) or N-terminal fragment (residues 1–180)] with V5-SBP-4EHP was analyzed as described in B. GYF2 residues 1–180 bound to 4EHP to a similar extent as the full-length protein. (*D*) Western blot showing the interaction of V5-SBP-4EHP (wild-type or the indicated mutants) with endogenous GYF1. The proteins were pulled down using streptavidin-coated beads. V5-SBP-MBP served as negative control. The inputs (1.5% for the V5-tagged proteins and 3% for GYF1) and bound fractions (3% for V5-tagged proteins and 35% for GYF1) were analyzed by western blotting using anti-V5 and anti-GYF1 antibodies. (*E, F*) Ni-NTA pulldown assays showing the interactions of GYF1 fragments (C+L+NC+A, C+L+NC and C) with 4EHP-His_6 (E) or eIF4E-His_6 (F). The eIF4E-binding region of 4E-BP1 (C+L+NC) binds similarly to both 4EHP and eIF4E, whereas GYF1 associates preferentially with 4EHP. The GYF1 and 4E-BP1 peptides contain an N-terminal MBP-tag and a C-terminal GB1 tag. The starting material (4% for the GYF1 fragments, 6% for 4EHP and recombinant eIF4E) and bound fractions (10% and 15% in panels E and F, respectively) were analyzed by SDS-PAGE followed by Coomassie blue staining. MBP served as a negative control. (*G*) The interaction of V5-SBP-tagged eIF4E or 4EHP proteins with endogenous GYF1, 4E-T and 4E-BP1 was analyzed in HEK293T cell

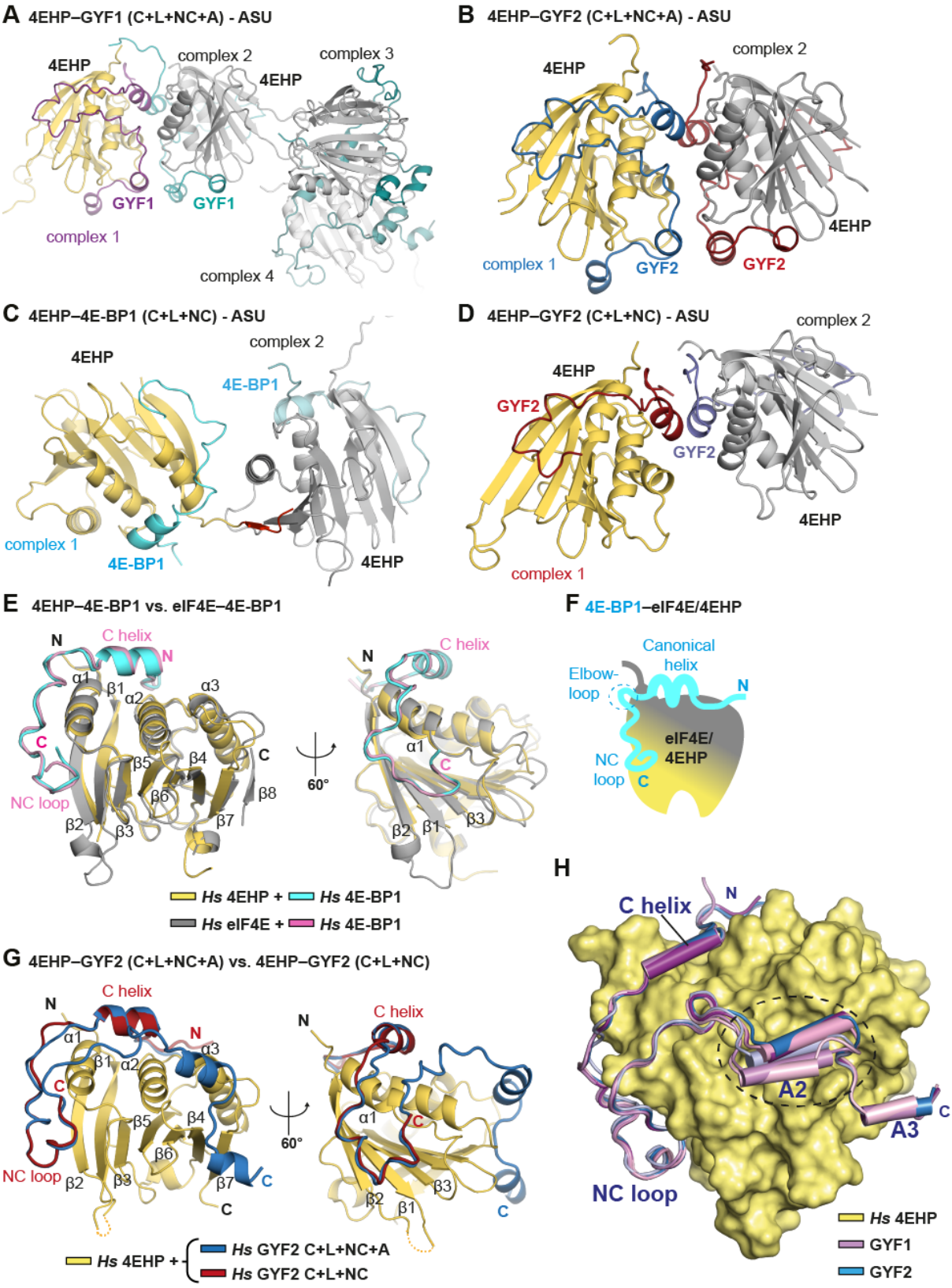
lysates. The proteins were pulled down using streptavidin-coated beads. Inputs (1.5%) and bound fractions (30% for 4E-BP1, GYF1 and 4E-T and 5% for the V5-SBP-tagged proteins) were analyzed by western blotting using anti-V5, anti-4E-BP1, anti-4E-T and anti-GYF1 antibodies. (*H, I*) Western blot analysis showing the interaction of GFP-tagged GYF1, GYF2 and 4E-T with HA-4EHP (H) or HA-eIF4E (I) in HEK293T cells. The proteins were immunoprecipitated using anti-GFP antibodies. The inputs (0.75% for GFP-tagged proteins and 0.5% for the HA-tagged proteins) and immunoprecipitates (15% for GFP-tagged proteins and 25% for HA-tagged proteins) were analyzed using anti-GFP and anti-HA antibodies, respectively. (*J*) Ni-NTA pulldown assay showing the interaction of 4EHP-His_6 (wild-type, W-A and IM-AA mutants) with GYF1 fragments with or without the auxiliary region (C+L+NC+A vs. C+L+NC). MBP served as a negative control. The starting material (4%) and bound fractions (9%) were analyzed by SDS-PAGE followed by Coomassie blue staining.

Figure S3


A GYF1 C + 4EHP
B GYF2 C + 4EHP
C GYF1 C+L+NC + 4EHP
D GYF2 C+L+NC + 4EHP
E GYF1 C+L+NC+A + 4EHP
F GYF2 C+L+NC+A + 4EHP
G 4EHP + m7GpppG cap
H 4EHP-GYF2 C+L+NC+A complex + m7GpppG cap
I GYF1 C+L+NC+A + 4EHP (dimerization mutants)
J GYF2 C+L+NC+A + 4EHP (dimerization mutants)
K 4E-BP1 C+L+NC + 4EHP
L 4E-BP1 C+L+NC + eIF4E
M 4E-BP1 C+L+NC + 4EHP (S99N mutant)
Time (min)
µcal/sec
kcal mol-1 of injectant
Molar Ratio

Supplemental Figure S3. Calorimetric titration data for the interaction of 4EHP with peptides derived from GYF1, GYF2 and 4E-BP1 or with m^7GpppG cap analog. (*A–F*) Isothermal titration calorimetry (ITC) profiles for the interaction of 4EHP (residues 52–234) with the following peptides: (A) GYF1 C; (B) GYF2 C; (C) GYF1 C+L+NC; (D) GYF2 C+L+NC; (E) GYF1 C+L+NC+A; (F) GYF2 C+L+NC+A. (*G*) ITC profile for the binding of 4EHP (residues 52-234) to m^7GpppG cap analog. (*H*) ITC profile for the binding of 4EHP-GYF2 C+L+NC+A complex to m^7GpppG cap analog. (*I, J*) ITC profiles for the interaction of 4EHP dimerization mutant with GYF1 and GYF2 (C+L+NC+A) dimerization mutant peptides. (*K - M*) ITC profiles for the interaction of 4E-BP1 (residues 50–83, C+L+NC) with the following proteins: (K) wild-type 4EHP (residues 52–234); (L) wild-type eIF4E (residues 36–217); (M) 4EHP (residues 52–234) S99N mutant. The thermodynamic parameters are shown in Table S2. Upper panels show raw data in (μcal sec^{-1}) and lower panels represent the integration of heat changes associated with each injection (kcal mol^{-1} of injectant). Data was fit using a one-site binding model.

Figure S4

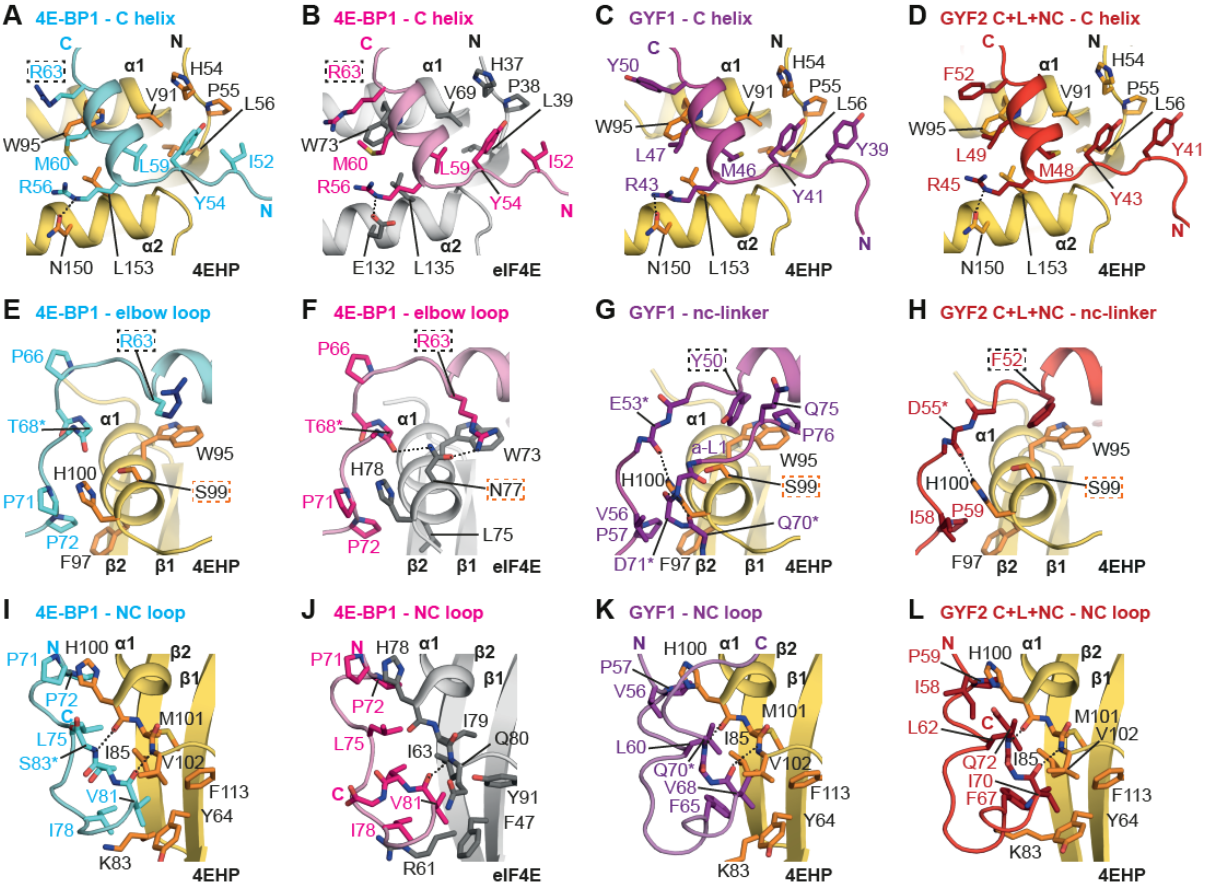


Supplemental Figure S4. Structures of 4EHP bound to GYF1, GYF2 and 4EBP1. (*A*) Cartoon representation showing the asymmetric unit (ASU) of the 4EHP–GYF1 crystal

form. The ASU contains four 4EHP–GYF1 complexes. In complex 1, which was used for representation, GYF1 is colored in purple and 4EHP is colored in yellow; in the other complexes of the ASU GYF1 is colored in cyan and 4EHP in grey. (*B*) Cartoon representation showing the ASU of the 4EHP–GYF2 crystal form. The ASU contains two 4EHP–GYF2 complexes. In complex 1, GYF2 is colored in blue and 4EHP in yellow. In complex 2, GYF2 is colored in red and 4EHP in grey. (*C*) ASU of the 4EHP–4E-BP1 (C+L+NC) crystal form. There are two 4EHP–4E-BP1complexes in the ASU. In complex 1, 4E-BP1 is colored in cyan and 4EHP is colored in yellow. In complex 2, 4EHP is colored in grey. The N-terminal portion of the 4EHP molecule from complex 1 is colored in red and contains residues from the expression tag, which mediate contacts to complex 2. (*D*) Cartoon representation showing the ASU of the 4EHP–GYF2 (C+L+NC) crystal structure. The ASU contains two 4EHP–GYF2 complexes. In complex 1, GYF2 is colored in red and 4EHP yellow. In complex 2, GYF2 is colored in blue and 4EHP in grey. The structural arrangement of the two complexes that lack the GYF2 auxiliary sequences appears similar to the dimeric arrangement of the complexes containing the auxiliary sequences (panel B). (*E*) Superposition of the structure of 4E-BP1 (cyan) bound to 4EHP (yellow) to the structure of 4E-BP1 (magenta) bound to eIF4E (grey; PDB: 4UED, Peter et al. 2015a). Selected secondary structural elements in 4EHP are label in black. The structures superpose with an RMSD of 0.41 Å over 194 Cα atoms. (*F*) Schematic representation of eIF4E and 4EHP bound to 4E-BP1. (*G*) Superposition of the structure of 4EHP bound to GYF2 C+L+NC+A (blue) with the structure of 4EHP bound to GYF2 C+L+NC (red) peptides. Selected secondary structural elements in 4EHP are label in black. The structures superpose with an RMSD of 0.38 Å over 207 Cα atoms. (*H*) Overlay of all complex structures of 4EHP bound to GYF1 and GYF2 peptides to illustrate the conformational flexibility of helix α2 (A2), which is circled with a black dashed line. The surface of 4EHP is shown in pale yellow and

the GYF peptides are colored in purple and blue for GYF1 and GYF2, respectively. Helical secondary elements (canonical helix, helices α2 and α3) are represented as cylinders.

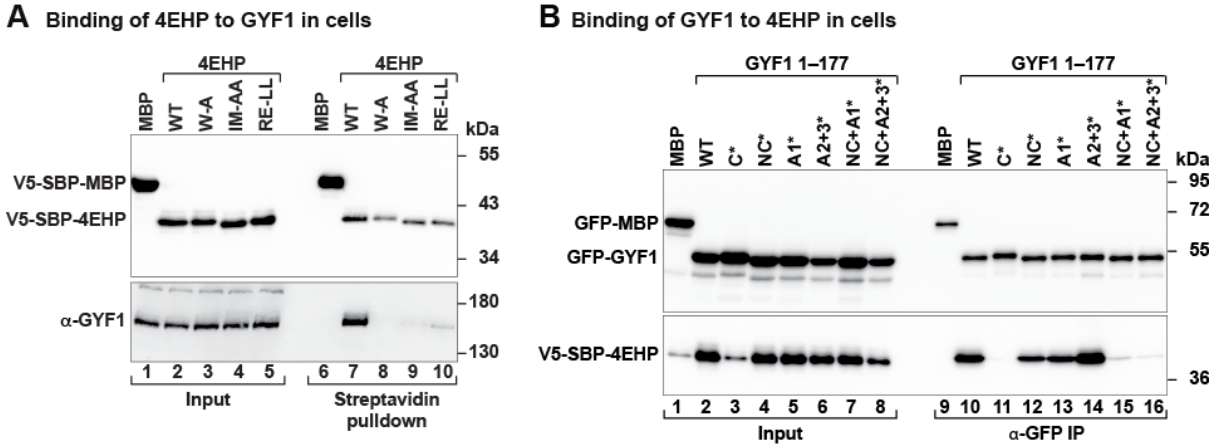
Figure S5



Supplemental Figure S5. Interactions of GYF1, GYF2 and 4E-BP1 with 4EHP and eIF4E. (*A–D*) Close-up views of the interaction between the dorsal surface of 4EHP (A,C,D) or eIF4E (B) and the canonical helices of 4E-BP1 (A, B; Peter et al. 2015a), GYF1 C+L+NC+A (C) and GYF2 C+L+NC (D). Selected residues are shown as sticks. Selected secondary structure elements are labeled in black for 4EHP or eIF4E and in color for the interacting partners. Residue $R63^{4E\text{-}BP1}$ is colored in dark blue following the Cγ atom in A and highlighted by a black dashed box in A and B. (*E–H*) Close-up views of the interaction between 4EHP (E,G,H) or eIF4E (F) and the non-canonical linkers of 4E-BP1 (E, F; Peter et al. 2015a), GYF1 C+L+NC+A (G) and GYF2 C+L+NC (H). Selected residues are shown as

sticks. Residue R63$^{4E\text{-}BP1}$ is colored in dark blue following the Cγ atom in E and highlighted by a black dashed box in E and F. The corresponding residues in GYF1 (Y50) or GYF2 (F52) are also highlighted by a black dashed box. For visual clarity, only backbone atoms are shown for the residues labeled with an asterisk. The residues N77 in eIF4E and S99 in 4EHP are highlighted with orange dashed boxes. (*I–L*) Close-up views of the interaction between the lateral surface of 4EHP (I,K,L) or eIF4E (J) and the non-canonical loops of 4E-BP1 (I, J; Peter et al. 2015a), GYF1 C+L+NC+A (K) and GYF2 C+L+NC (L).

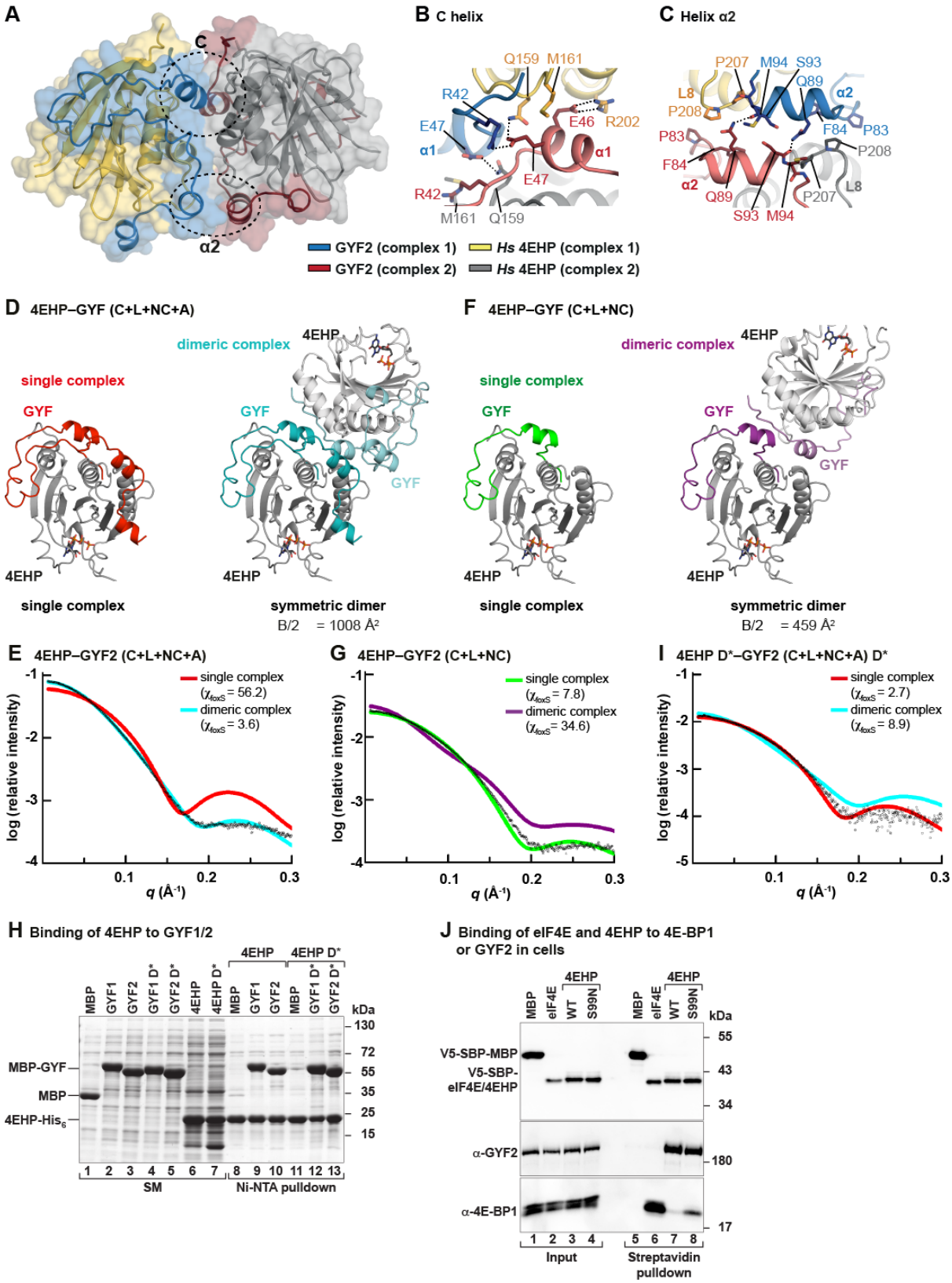
Figure S6



Supplemental Figure S6. Validation of the interfaces observed in the 4EHP–GYF1/2 and 4EHP–4E-BP1 complex structures. (*A*) Western blot analysis showing the interaction of endogenous GYF1 with V5-SBP-tagged 4EHP (WT or the indicated mutants). The proteins were pulled down using streptavidin-coated beads. V5-SBP-MBP served as negative control. The inputs (2.5%) and immunoprecipitates (3% for the V5-tagged proteins and 20% for GYF1) were analyzed by western blotting using anti-V5 and anti-GYF1 antibodies. (*B*) Interaction of GFP-GYF1 N-terminal fragment (residues 1–177; either wild-type or the indicated mutants) with V5-SBP-tagged full-length 4EHP. The proteins were immunoprecipitated from HEK293T cell lysates using anti-GFP antibodies. GFP-MBP served as negative control. The inputs (1.5% for the GFP-tagged proteins and 0.5% for V5-

SBP-4EHP) and immunoprecipitates (7.5% for the GFP-tagged proteins and 30% for V5-SBP-4EHP) were analyzed by western blotting using anti-GFP and anti-V5 antibodies.

Figure S7

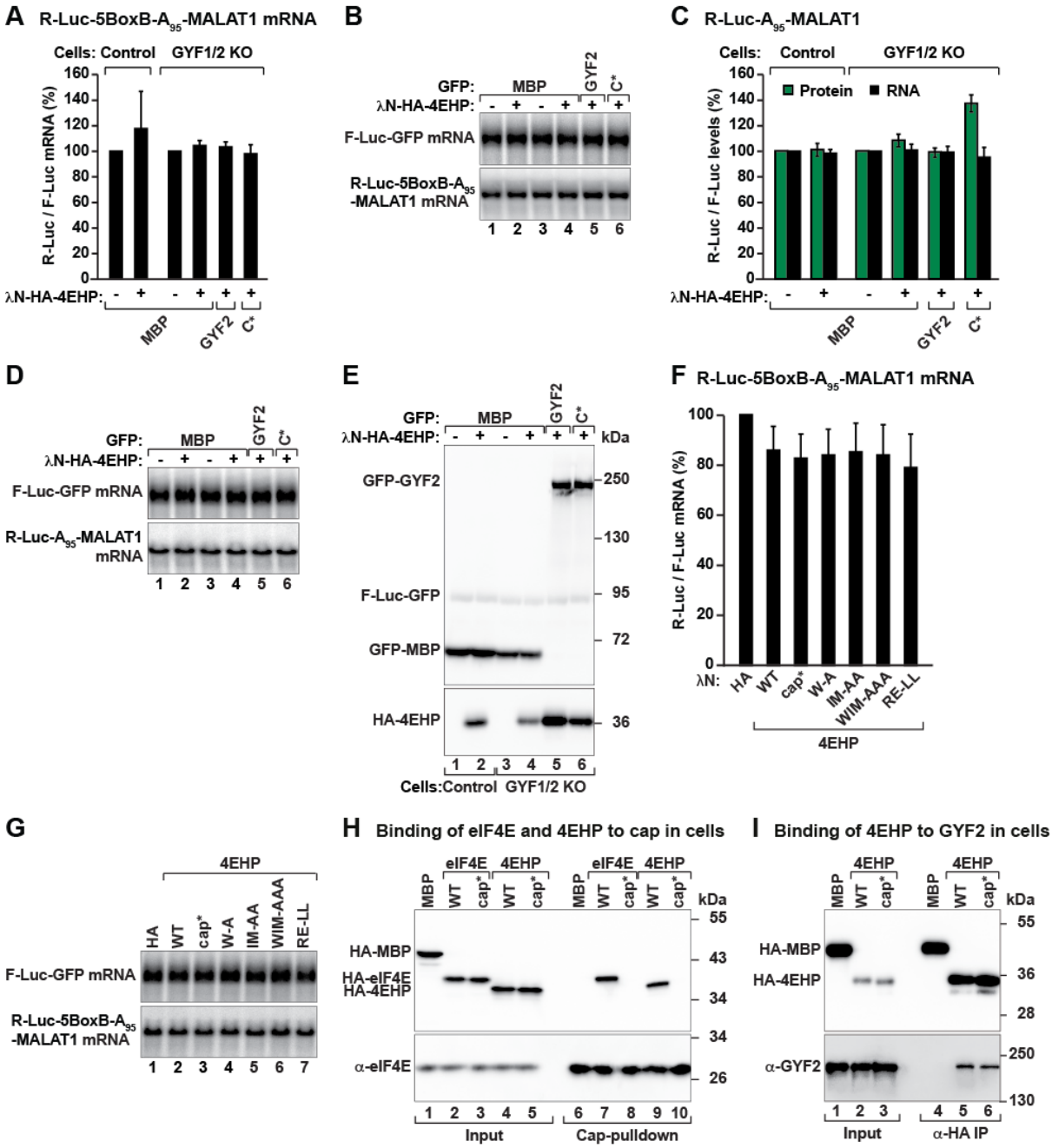


Supplemental Figure S7. The 4EHP–GYF1/2 complexes form dimers in solution. (*A*) Representation of the dimeric arrangement of the 4EHP–GYF2 complexes in the asymmetric

unit of the crystal. In this dimeric arrangement, the two 4EHP–GYF2 complexes adopt a two-fold rotational symmetry with their dorsal surfaces facing each other. The interface of this arrangement involves: I) the canonical helix of GYF2, which is in contact with the neighboring 4EHP molecule and is extended by the canonical helix of the GYF2 molecule of the neighboring complex, and II) the auxiliary region of GYF2, where helix α2 of one complex faces, in an antiparallel fashion, helix α2 of the GYF2 and loop 8 of the 4EHP present in the neighboring complex. The contacts between the two complexes are highlighted with dashed black circles. In complex 1, GYF2 is colored in blue and 4EHP in yellow. In complex 2, GYF2 is colored in red and 4EHP in grey. (*B, C*) Close-up views on the dimeric interface involving the canonical helices (C helix, panel B) and the auxiliary helix 2 (helix α2, panel C). Selected residues are shown as sticks and colored as in panel A. GYF2 residues E46 and E47 within the canonical motifs of interacting GYF2 molecules contact the guanidinium group of $R202^{4EHP}$ and the side chain of $Q159^{4EHP}$ on neighboring 4EHP molecules. $M161^{4EHP}$ contacts the aliphatic portion of $E46^{GYF2}$ in the canonical helix of the neighboring complex. Pro residues in 4EHP loop L8 (P207 and P208) are facing residues proximal to the helix α2 of GYF2 from the neighboring complex (P83, F84, $Q89^{GYF2}$). (*D, F*) Crystallographic models of the 4EHP–GYF2 complexes. The radius of gyration (R_G) and the maximum particle size (D_{max}) were calculated using Scatter and are summarized in Table S3. In the case of the dimeric assemblies, the dimer interface (B/2) was calculated using PISA from the CCP4 package and is indicated below the structures. 4EHP is shown in grey. The GYF2 (C+L+NC+A) peptide is colored in red and cyan in the single and dimeric arrangements, respectively. The GYF2 (C+L+NC) peptide is colored in green and purple in the single and dimeric arrangements, respectively. (*E, G, I*) Small-angle X-ray scattering profiles comparing single and dimeric arrangements of the 4EHP–GYF2 complexes with the experimental data. The data are plotted with the logarithmic scattering intensity on the y-axis

and the scattering angle *q* on the x-axis. Experimental scattering data of the complexes in solution are shown as open black circles and the fits for the single and dimeric arrangements are shown as a line colored as indicated on the right. The goodness-of-fit χ values, calculated using FoXS, are indicated for each fit. (*H*) Ni-NTA pulldown assay showing the interaction of 4EHP-His_6 (M1–F234, wild-type and dimerization mutant) with MBP-tagged GYF1 and GYF2 proteins [(wild-type and dimerization mutant (D*)]. MBP served as a negative control. The input (07% for MBP-tagged proteins and 2% for 4EHP) and bound fractions (9%) samples were analyzed by SDS-PAGE followed by Coomassie blue staining. (*J*) Analysis of the interaction of V5-SBP-tagged eIF4E, 4EHP and the 4EHP S99N mutant with endogenous GYF2 and 4E-BP1 proteins in HEK293T cell lysates. The proteins were pulled down using streptavidin-coated beads. V5-SBP-MBP served as negative control. Input samples (1% for 4E-BP1 and GYF2 and 1.5% for the V5-SBP-tagged proteins) and bound fractions (20% for 4E-BP1 and GYF2 and 5% for V5-SBP-tagged proteins) were analyzed by western blotting using anti-V5, anti-GYF2 and anti-4E-BP1 antibodies.

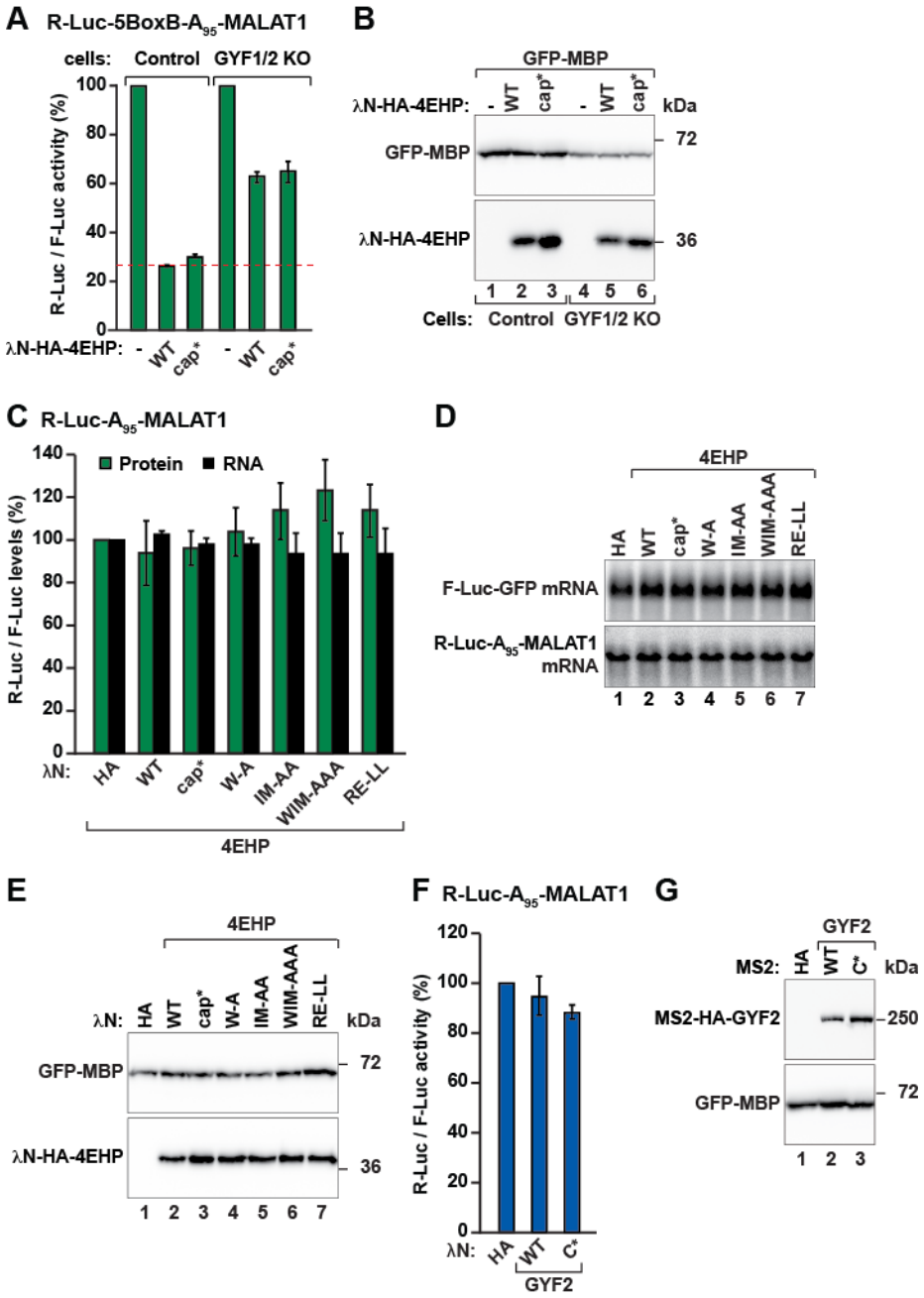
Figure S8



Supplemental Figure S8. 4EHP requires GYF1/2 proteins to repress the expression of bound mRNAs (*A*) A complementation assay using the R-Luc-5BoxB-A$_{95}$-MALAT1 reporter and λN-HA-4EHP (either wild-type or the indicated mutants) was performed in control and GYF1/2-null HEK293T cells expressing GFP-MBP or GFP-GYF2 (wild-type or canonical mutant, C*). A plasmid expressing F-Luc-GFP served as the transfection control. R-Luc activity and mRNA levels were normalized to those of the F-Luc transfection control

and set to 100% in cells expressing the λN-HA peptide. Normalized R-Luc activities are shown in Figure 6A. The panel shows the corresponding normalize R-Luc-5BoxB-A_{95}-MALAT1 mRNA levels. Bars represent the mean values and error bars represent standard deviations from three independent experiments. (*B*) Northern blot of representative RNA samples corresponding to the experiment shown in A and Fig. 6A. (*C*) Complementation assay in WT and GYF1/2-null cells using the R-Luc-A_{95}-MALAT1 reporter lacking the BoxB hairpins. A plasmid expressing F-Luc-GFP was used as a transfection control. R-Luc activity and mRNA levels were normalized to those of the F-Luc transfection control and set to 100% in cells expressing the λN-HA peptide. (*D*) Northern blot of representative RNA samples corresponding to the experiment shown in C. (*E*) Western blot analysis showing the equivalent expression of the λN-HA-4EHP and GYF2 proteins used in the complementation shown in C and D. (*F*) Normalized R-Luc-5BoxB-A_{95}-MALAT1 mRNA levels corresponding to the experiment shown in Fig. 6D,E. (*G*) Northern blot of representative RNA samples corresponding to the experiment shown in F. (*H*) Lysates from HEK293T cells expressing HA-tagged eIF4E or 4EHP (wild-type or cap mutant, cap*) were pulled down with m^7GTP-sepharose beads. Endogenous eIF4E served as positive control. Inputs (0.75% for the HA-tagged proteins and 1% for endogenous eIF4E) and bound fractions (15% for the HA-tagged proteins and 5% for endogenous eIF4E) were analyzed by Western blot using anti-HA and anti-eIF4E antibodies. (*I*) Interaction of HA-tagged 4EHP (wild-type or cap mutant, cap*) with endogenous GYF2 in HEK293T cells. HA-tagged MBP served as a negative control. Inputs (0.37% for the HA-tagged proteins and 0.75% for endogenous GYF2) and immunoprecipitates (15% for the HA-tagged proteins and 20% for endogenous GYF2) were analyzed by Western blot using anti-HA and anti-GYF2 antibodies.

Figure S9



Supplemental Figure S9. 4EHP requires GYF1/2 proteins to repress translation of bound mRNAs (*A*) A complementation assay using the R-Luc-5BoxB-A_{95}-MALAT1 reporter and λN-HA-4EHP [either wild-type or cap mutant (cap*, W124A)] was performed in control and GYF1/2-null HEK293T cells expressing GFP-MBP or GFP-GYF2 (wild-type or canonical mutant). A plasmid expressing F-Luc-GFP served as the transfection control. For each cell type, R-Luc activity was normalized to that of the F-Luc transfection control and set

to 100% in cells expressing the λN-HA peptide. Samples were analyzed as described in Fig. 6A. (*B*) Western blot showing similar expression of the proteins used in A. (*C*) Tethering assay using the R-Luc-A_{95}-MALAT1 reporter and λN-HA-4EHP (wild-type or mutants) in HEK293T cells. Samples were analyzed as described in Fig. 6A. (*D*) Northern blot of representative RNA samples corresponding to the experiment shown in C. (*E*) Western blot showing the equivalent expression of the λN-HA-4EHP proteins used in the tethering assay shown in D.

(*E*) Tethering assay using the R-Luc-A_{95}-MALAT1 reporter lacking MS2 binding sites and MS2-HA-GYF2 (wild-type or canonical mutant). Samples were analyzed as described in Fig. 6F. The corresponding assay with the reporter containing the MS2 binding sites is shown in Fig. 6F. (*G*) Western blot analysis showing the expression of the GFP-GYF2 proteins used in the tethering assays shown in F.

Supplemental references

Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH, Adams PD. 2012. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* **68:** 352–367.

Chen VB, Arendall WB, 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. 2010. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66:** 12–21.

Cheng Y, Patel DJ. 2004. An efficient system for small protein expression and refolding. *Biochem Biophys Res Commun* **317:** 401–405.

Diebold ML, Fribourg S, Koch M, Metzger T, Romier C. 2011. Deciphering correct strategies for multiprotein complex assembly by co-expression: application to complexes as large as the histone octamer. *J Struct Biol* **175:** 178–188.

Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66:** 486–501.

Fabian MR, Frank F, Rouya C, Siddiqui N, Lai WS, Karetnikov A, Blackshear PJ, Nagar B, Sonenberg N. 2013. Structural basis for the recruitment of the human CCR4-NOT deadenylase complex by tristetraprolin. *Nat Struct Mol Biol* **20**: 735-739.

Kabsch W. 2010. Xds. *Acta Crystallogr D Biol Crystallogr* **66:** 125–132.

Kuzuoglu-Ozturk D, Bhandari D, Huntzinger E, Fauser M, Helms S, Izaurralde E. 2016. miRISC and the CCR4-NOT complex silence mRNA targets independently of 43S ribosomal scanning. *EMBO J* **35:** 1186–1203.

McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. 2007. Phaser crystallographic software. *J Appl Crystallogr* **40:** 658–674.

Mizoue LS, Tellinghuisen J. 2004. The role of backlash in the "first injection anomaly" in isothermal titration calorimetry. *Anal Biochem* **326:** 125–127.

Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. 2013. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* **8:** 2281–2308.

Schneidman-Duhovny D, Hammel M, Tainer JA, Sali A. 2013. Accurate SAXS profile computation and its assessment by contrast variation experiments. *Biophys J* **105:** 962–974.

Terwilliger TC, Grosse-Kunstleve RW, Afonine PV, Moriarty NW, Zwart PH, Hung LW, Read RJ, Adams PD. 2008. Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr D Biol Crystallogr* **64:** 61–69.