Cell Reports, Volume *23*

Supplemental Information

An RNA-Binding Multimer Specifies Nematode Sperm Fate

Scott T. Aoki, Douglas F. Porter, Aman Prasad, Marvin Wickens, Craig A. Bingman, and Judith Kimble

Supplemental Information

Supplemental Experimental Procedures

Biochemistry and crystallography

Protein expression and purification
Full length FOG-3 1-263 was amplified from *C. elegans* N2 cDNA with primers that included a six-histidine tag, stop codon and 12 nucleotides suitable for annealing with ligation independent cloning (LIC) (Aslanidis and de Jong, 1990). Mixed stage N2 cDNA was generated by reverse transcription (SuperScript® II Reverse Transcriptase, Thermo Fisher; Waltham, MA) with oligo-dT (Invitrogen; Carlsbad, CA). The FOG-3 PCR product was cloned into a pET21a (EMD Millipore; Billerica, MA) bacterial expression plasmid by LIC. From this plasmid, histidine-tagged FOG-3 1-238 was amplified by PCR and inserted into pET21a using LIC. For FOG-3 1-137, an *E. coli* codon-optimized sequence for histidine-tagged FOG-3 1-137 was ordered as a gBlock (Integrated DNA Technologies (IDT); Coralville, IA) and cloned into pET21a by LIC. All further changes to the pET21a histidine-tagged FOG-3 1-137 were by Gibson assembly cloning (Gibson, 2009). Residues 1-137 gave good yields, but solubility and stability remained issues at high concentration. We improved solubility by retaining the histidine tag and mutating non-conserved residues to amino acids in related FOG-3 orthologs (H47N and C117A) (**Figure S1**). For chemical crosslinking, we modified this optimized, histidine-tagged FOG-3 1-137 expression plasmid to include mutations (R22K, L64K, I112K, R82K) and extended the C-terminus three amino acids (138-140) to permit histidine tag cleavage (not used). This lysine substitution FOG-3 plasmid was later modified to include mutations R14K or E126K. Final plasmids used for protein expression listed (below).

Expression plasmids were transformed into Rosetta™2(DE3) cells (EMD Millipore; Billerica, MA) and grown in LB (MP Biomedicals; Santa Ana, CA) for 5 hours at 37°C until A600 = ~0.8. The culture was then induced with 0.1 mM IPTG (MP Biomedicals; Santa Ana, CA) and grown at 16°C for 16-20 hours prior to collection, centrifugation, washing and freezing in liquid nitrogen. Cells were defrosted on ice and reconstituted in lysis buffer (20 mM $NaPO_4$ pH 7.4, 300 mM NaCl, 10 mM imidazole, 5 mM β-mercaptoethanol) with cOmplete protease inhibitors (Roche; Indianapolis, IN). Cells were lysed with a French Press, centrifuged (3,220 *x g* and 10,000 *x g*) to remove unlysed cells and precipitate, and incubated with Nickel-NTA beads (Thermo Fisher; Waltham, MA) for 2 hours at 4°C with rocking. Beads were washed with lysis buffer and eluted with an imidazole step gradient (imidazole at 20, 40, 60, 80, 100, 250 mM) in elution buffer (20 mM $NaPO_4$ pH 7.4, 300 mM NaCl, 5 mM β-mercaptoethanol). Protein used for biochemical experiments was dialyzed in FOG-3 buffer (20 mM HEPES pH 7.0, 50 mM NaCl, 0.5 mM (tris(2-carboxyethyl)phosphine) (TCEP)), while protein for crystallization was dialyzed in crystallization buffer (20 mM HEPES pH 7.0, 100 mM $MgS0_4$, 0.5 mM TCEP). Samples were concentrated with Amicon Ultra-4 3000 MW concentrators (EMD Millipore; Billerica, MA) and run on a Superdex 200 (GE Healthcare; Pittsburgh, PA). Recombinant protein was again concentrated with Amicon Ultra-4 3000 MW concentrators and protein concentration estimated by A280.

Plasmids
pJK2068: pET21a vector backbone, *C. elegans* FOG-3 1-238::six histidines
pJK2069: pET21a vector backbone, *E. coli* codon-optimized *C. elegans* FOG-3 1-137 H47N C117A::six histidines
pJK2070: pET21a vector backbone, *E. coli* codon-optimized *C. elegans* FOG-3 1-140 H47N C117A, lysine substitutions (R22K, I112K, L64K, R82K)::six histidines
pJK2071: pET21a vector backbone, *E. coli* codon-optimized *C. elegans* FOG-3 1-140 R14K H47N C117A, lysine substitutions (R22K, I112K, L64K, R82K)::six histidines
pJK2072: pET21a vector backbone, *E. coli* codon-optimized *C. elegans* FOG-3 1-140 H47N C117A E126K, lysine substitutions (R22K, I112K, L64K, R82K)::six histidines
pJK1910: pDD162 vector backbone, CRISPR-Cas9 sgRNA targeting *fog*-3 (caatcagtccccgagtacg)
pJK1925: pDD162 vector backbone, CRISPR-Cas9 sgRNA targeting *fog-3* (ggttctgaccacgtactcg)

Crystallization, data collection and refinement
Initial crystallization trials proved unfruitful. Thinking that a cofactor was missing, we performed a thermal folding assay (Ericsson et al., 2006) with various additives (see below). Magnesium and sulfate improved protein thermostability (**Figure S2C**) so we reasoned that they might aid stability during crystallization. Crystallization conditions were screened with sitting drop trays set up using the Mosquito (TTP Labtech; Cambridge, MA). We obtained crystals using recombinant FOG-3 (1-137 H47N C117A) with an intact histidine tag and incubating our

trays at 4°C. After 3 weeks, rhomboid crystals were observed in conditions A (0.1 M sodium citrate pH 5.6, 10% (vol/vol) isopropanol, 10% (wt/vol) PEG 4000) and B (0.1 M magnesium acetate, 0.1 M sodium citrate pH 5.6, 8% (wt/vol) PEG 10000). UV scanning with a UVEX-M (280 nm excitation, 350 nm emission; JANSi; Seattle, WA) identified these to be protein crystals. Both conditions were reproducible. We were able to collect complete datasets from the crystals grown directly from the condition B screening trays. Phasing was accomplished with molecular replacement using Phaser (McCoy et al., 2007) and a human Tob homolog (PDB ID: 2Z15) as a starting model. Model building and refinement were done in Phenix (Adams et al., 2010) and Coot (Emsley and Cowtan, 2004). Water molecules were first modeled by Phenix before being checked manually. Three densities were too big to be water molecules. We could model one of the densities with sulfate. Two densities were observed in the solvent-accessible area adjacent to residues 52-56 in both copies in the ASU. Density is observed at FoFc contour levels past 6 σ. We attempted modeling of acetate (too small) and citrate (too large), both molecules that were present in the crystallization conditions, but the fit was unsatisfactory. Thus, the final uploaded model does not account for these two large densities. Analyses of protein assemblies, dimer interactions and free energy estimations were done in PISA (Krissinel and Henrick, 2007). Structural images were generated in Pymol (The PyMOL Molecular Graphics System, Version 2.0,1 Schrödinger, LLC).

The FOG-3 missense mutations were modeled in Coot (Emsley and Cowtan, 2004). Their disruption of FOG-3 folding, dimerization or multimerization was inferred based on mutant residue disruption of hydrogen bonds, salt bridges and hydrophobic packing (steric hindrance) in the structure. The lysine substitution sites were chosen based on whether the locations were close enough for intra-dimer (R22K, I112K) and dimer-dimer (L64K, R82K) BS3 crosslinking, sequence variability (a lack of sequence conservation, **Figure S1**) and their expected tolerance for a lysine mutation. R22 was conserved in FOG-3 paralogs, but lysine was easily modeled. Two mutations changed hydrophobic residues, but lysines at these positions were found in other FOG-3 orthologs and could be modeled.

FOG-3 protease cleavage

Comparison of FOG-3 sequences from several *Caenorhabditid* species reveals further nematode-specific conservation that extends ~20 amino acids past the predicted Tob/BTG fold. We used recombinant *C. elegans* FOG-3 protein and proteases to identify a single domain spanning the canonical Tob/BTG fold and a nematode-specific extension. Recombinant FOG-3 1-238 with a C-terminal histidine tag was incubated with trypsin (Sigma-Aldrich; St. Louis, MO) and elastase (Sigma-Aldrich; St. Louis, MO) at room temperature prior to SDS-PAGE. The gel was stained with Coomassie to visualize cleavage products. The proteases generated ~15 kDa protected fragments (**Figure S2B**). Samples were also cleaved with trypsin or elastase for 45 minutes at room temperature (~20°C) and submitted for in-solution mass spectrometry (University of Wisconsin-Madison Biotechnology Center). The mass spectrometry fragment that most closely matched the SDS-PAGE band mapped to residues 1-135 for trypsin and 1-142 for elastase. Both fragments included the predicted Tob/BTG fold plus the nematode-specific extension (**Figure 1B**). FOG-3 1-137 exhibited a broad elution peak at higher versus lower concentrations (**Figure S2A**) that could be attributed to different dimer versus monomer states.

Protein folding assay

The protein folding assay followed published protocols (Ericsson et al., 2006). Briefly, recombinant FOG-3 1-137 with histidine tag was incubated with 90x concentrated SYPRO orange (5000x stock, Invitrogen; Carlsbad, CA) in FOG-3 buffer. 18 µl of the protein-dye mix was mixed with 2 µl Additive Screen (Hampton Research; Aliso Viejo, CA) and heated in a 7500 Real-Time PCR System thermocycler (Applied Biosystems; Foster City, CA) at 0.1°C/s from 20°C to 70°C while monitoring A405. SYPRO orange dye bound to unfolded protein. Thus, FOG-3 unfolded at a certain temperature, allowing dye binding and increasing A405 absorbance. The additive was judged as enhancing thermostability based upon the shift in the melting curve to the right, or requiring higher temperatures for signal. This assay was performed twice with similar results.

Negative-stain electron microscopy

Samples were negative stained with Nano-W (Nanoprobes; Yaphank, NY) using the two-step method. A 2 µl droplet of samples was placed on a Pioloform (T. Pella) coated 300 mesh Cu Thin-Bar grid (EMS; Hatfield, PA), coating side down. The excess was wicked with filter paper and allowed to barely dry. A 2 µl droplet of Nano-W was applied, wicked again with clean new filter paper, and allowed to dry. The sample was viewed on a Philips CM120 transmission electron microscope at 80 kV and documented with a SIS (Olympus / Soft Imaging Systems; Münster, Germany) MegaView III digital camera.

Protein crosslinking

Bis[sulfosuccinimidyl] suberate (BS3) (Thermo Scientific; Waltham, MA) was diluted in crosslinking buffer (20 mM HEPES pH 7.0, 150 mM NaCl) and added to recombinant protein for a final concentration of 0.5 mg/ml protein and 0.5, 0.25, 0.125, 0.0625, 0.03125 mM BS3 crosslinker. Buffer alone was added as a negative control. After 30 minutes at room temperature (~20°C), the reaction was quenched with 1M Tris pH 6.8 (50 mM final concentration). Samples were analyzed by SDS-PAGE and coomassie staining. Experiments were performed twice in these conditions with similar results.

***C. elegans* E126K alleles**

CRISPR-Cas9 gene editing of the endogenous *fog-3* gene was achieved using a *dpy-10* roller co-injection strategy (Arribere et al., 2014). Briefly, an sgRNA construct containing the U6 promoter and sgRNA scaffold from pDD162 (Dickinson et al., 2013) along with the targeting sequences caatcagtccccgagtacg (pJK1910) and ggttctgaccacgtactcg (pJK1925) was cloned into the *Xma*I site of pUC19 using one step isothermal DNA assembly. The repair template was a ssDNA oligo (Integrated DNA Technologies (IDT); Coralville, IA) that inserted an E126K mutation and removed an *Ava*I restriction site. See Table (below) for CRISPR-Cas9 target sequences and repair oligos used. Injections were performed in young N2 hermaphrodite *C. elegans,* using either 1) *fog-3* sgRNA plasmids, *dpy-10* sgRNA plasmid, *fog-3* E126K repair template, and Cas9 plasmid as described (Arribere et al., 2014), or 2) *fog-3* crRNA, *dpy-10* crRNA, *fog-3* E126K repair template and recombinant Cas9 protein as described (Paix et al., 2015) (see Table below for reagent sequences), and F1 rollers were screened for the desired mutation by PCR and *Ava*I digest. Alleles were recovered from separate injected animals and therefore represent independent editing events. We verified the *fog-3* mutations by Sanger sequencing. Homozygous mutants had a Fog phenotype and thus could only produce oocytes. These worms were outcrossed twice with N2 before crossing with JK2739 containing balancer *hT2[qIs48](I;III)*.

CRISPR-Cas9 guide RNAs and repair oligos

Name	Type	Strain targeted	Sequence
CRISPR-Cas9 guide RNAs:			
fog-3 5 crRNA	CRISPR-Cas9 RNA	N2	target sequence: taatactgggaaattaaaaa
fog-3 crRNA 8	CRISPR-Cas9 RNA	N2	target sequence: cggttctgaccacgtactcg
his-58 tbb-2 crRNA 1	CRISPR-Cas9 RNA	JA1515	target sequence: aggatcttgcatTTACTTGC
single stranded DNA repair oligos:			
fog-3 E126K repair	ssDNA repair oligo	N2	ataaaaatactttaaatttcatttttccagctaccaatcagtccccAagtaTgtTgtcCgaaccgctgcaatccgcgcggagccttgctcgaatcttgg
fog-3 FLAG3x repair	ssDNA repair oligo	N2	attcaagcatcaacgaccaaatgagatattctccccgtGGAGGATCCGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAGttttaatttcccagtattagaatctcaattatcataccgt
fog-3 lambda repair	ssDNA repair oligo	N2	tcaacgaccaaatgagatattctccccgtATGGACGCCCAAACCCGCCGCCGCGAGCGCCGTGCCGAGAAGCAAGCCCAATGGAAGGCCGCCAACGGAGGATCCGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAGttttaatttcccagtattagaatctcaatt
his-58 tbb-2 boxB repair	ssDNA repair oligo	JA1515	CAA GGC CGT CAC CAA GTA CAC TTC TAG CAA GTA AAT GCA AGA TCC AAC TAC TAA ACT GAT TCC TGG GCC CTG AAG AAG GGC CCC TCG ACT AAG TCC AAC TAC TAA ACT GGG CCC TGA AGA AGG GCC CAT ATA GGG CCC TGA AGA AGG GCC CTA TCG AGG ATA TTA TCT CGA CTT TCA AGC ATT CCC TTC TTC TCT ATC AC

Reporter design and application

The goal was to develop a reporter that could take advantage of the lambda/boxb mRNA tethering system (Baron-Benhamou et al., 2004) and a germline fluorescent expression reporter. However, the published boxb-containing reporters and our own MosSCI-generated reporters could not be detected, most likely due to weak signal in the

germline region expressing FOG-3. We therefore modified a worm strain with a robust germline fluorescent reporter. Using CRISPR-Cas9 gene editing (Paix et al., 2015), three boxB hairpins (3xboxB) were inserted into *weSi2* (JA1515) (Zeiser et al., 2011), a transgene reporter expressing a GFP-tagged histone under the ubiquitous *mex-5* germline promoter and the *tbb-2* 3'UTR. 3xboxB was a modified design based on previous mammalian 3'UTR boxB reporters (Wang et al., 2011). Inserts were screened by a phenotype-based co-injection marker (Arribere et al., 2014) and by PCR for potential inserts. Candidates were homozygosed and insertions sequenced for correct repair. After outcrossing, we confirmed that the allele-containing strains still expressed nuclear GFP throughout the adult germline. For the FOG-3 tagged alleles, either 3xFLAG or λN22::3xFLAG was inserted into *fog-3* at its C-terminus, one amino acid (R262) away from the end. Insertions were screened and sequenced as described for the reporter. Worms with correct insertions were homozygous fertile and could be maintained as cross-fertile male-hermaphrodite lines, strong evidence for fully functional FOG-3. These worms were crossed into the reporter and imaged as described. See accompanying Table (above) for a summary of the CRISPR-Cas9 target sequences and repair oligos used.

Imaging

For fluorescent imaging, germlines were extruded, fixed and permeabilized with Triton-X (0.5%) as previously described (Crittenden et al., 2017). Germlines were incubated with primary antibodies to FLAG (M2® (mouse), Sigma; St. Louis, MO) and GFP (Rabbit anti-GFP, Invitrogen; Carlsbad, CA) overnight, stained with fluorophore-labeled secondary antibodies (Alexa 555 Donkey anti-Mouse, Alexa 488 Goat anti-Rabbit; Invitrogen; Carlsbad, CA) and DAPI (Invitrogen; Carlsbad, CA), washed, and mounted in Vectashield (Vector Laboratories; Burlingame, CA). Germlines were imaged by confocal microscopy on a Leica SP8 scanning laser confocal microscope, taking 1 µm slices in sequence. Maximum intensity stack projections were generated with ImageJ (Schindelin et al., 2015) and brightness adjusted with Photoshop (Adobe; San Jose, CA). All images were equally treated in ImageJ and Photoshop, with the exception of the reported DAPI images. Imaging experiments were repeated at least twice with similar results.

iCLIP

In vivo crosslinking and immunoprecipitation (iCLIP) was carried out essentially as described (Huppertz et al., 2014), with modifications to worm growth, crosslinking, lysis, RNase digestion, and data analysis as described.

Nematode culture and UV crosslinking for iCLIP

L1 larvae from *C. elegans* strain JK4871 were obtained by bleaching and synchronizing via standard methods (Stiernagle, 2006). Larvae were plated onto 10 cm OP50 plates (~50,000 per plate) and propagated at 20°C for ~40-46 hours until most worms were at the early L4 stage when FOG-3 expression is greatest. Worms were washed with M9 (42.3 mM Na_2HPO_4, 22 mM KH_2PO_4, 85.6 mM NaCl, 1 mM $MgSO_4$), pooled into groups of 250,000 living worms, placed on a 10 cm NGM agarose plate and liquid removed. Animals were irradiated two times sequentially at 254 nm with 0.9999 J/cm^2 in a XL-1000 UV Crosslinker (Spectrolinker, Thomas Scientific; Swedesboro, NJ). Non-crosslinked samples were incubated at room temperature as a negative control for the radiolabeled gel (**Figure S4E**). For the iCLIP negative control, we performed the pulldown of crosslinked JK4871 worm lysate with beads alone (no antibody). Worms were rinsed from plates with cold M9, washed once, and transferred to a 2 mL Eppendorf tube. The pellet was washed again in freezing buffer (50 mM Tris pH 7.5, 150 mM NaCl, 10% (wt/vol) glycerol, 0.05% (vol/vol) Tween 20) and frozen with liquid nitrogen. Pellets were stored at -80ºC until use.

Lysis and RNA digestion

C. elegans pellets were thawed by adding ice cold lysis buffer (50 mM Tris pH 7.5, 100 mM NaCl, 1% Pierce NP-40, 0.1% SDS, 0.5% sodium deoxycholate, Roche cOmplete EDTA-free Protease Inhibitor Cocktail, Ambion ANTI-RNase) and incubated for 20 minutes at 4ºC with rocking. The thawed pellets were centrifuged at 1,000 x *g*, 4ºC for 1 minute and washed 3 times with ice cold lysis buffer. Lysis buffer was added to the pellet along with a 5 mm Retsch stainless steel ball (Verder Scientific; Newtown, PA). Lysis was performed in the cold room using a Retsch 400 MM mill mixer (Verder Scientific; Newtown, PA). Lysis was completed after three 10-minute cycles at a setting of 30 Hz, with four-minute freeze-thaws after the first and second cycles. Freeze-thaws were performed by immersion in liquid nitrogen for 1 minute, then returning to liquid state by immersion in room temperature water for 4 minutes. Worm lysis was confirmed by observing a small aliquot of final lysate on a dissection scope. The lysate was cleared by centrifugation for 15 minutes at 16,100 x *g*, 4ºC. Protein concentration of the cleared lysate was determined with the Direct Detect spectrometer (EMD Millipore; Billerica, MA). Our pellets containing 250,000 worms yielded ~12 mg/mL of total protein, and we used 10 mg total protein per biological replicate. Double RNase

digestion of protein-RNA complexes was performed as previously described (Spitzer et al., 2014). For the first digestion, which occurred immediately after lysis and just prior to immunoprecipitation, guanosine specific RNase T1 (Thermo Fisher; Waltham, MA) was added to the cleared lysate at a final concentration of 1 Unit/µL. The sample was incubated in a Thermomixer for 15 minutes at 22ºC, 1,100 rpm and then cooled on ice for 5 minutes.

Protein G Dynabeads (Life Technologies; Waltham, MA) were aliquoted to a fresh RNase-free roundbottom tube (USA Scientific; Ocala, FL). The tube was placed on a Dynal magnet (Invitrogen; Carlsbad, CA), the existing buffer was removed, and M2 FLAG antibody (Sigma-Aldrich; St. Louis, MO) added at 20 µg antibody to 3 mg Protein G Dynabeads in PBS-T (PBS pH 7.2 (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO4, 1.8 mM KH_2PO4), 0.02% Tween-20). The beads plus antibody solution was incubated at room temperature on a rotator for 45 minutes. The tube was again placed on the magnet, the antibody solution removed, and the cleared lysate added. Immunoprecipitation was carried out overnight at 4°C. As a negative control, we performed the pulldown of crosslinked JK4871 worm lysate with beads alone (no antibody).

Following immunoprecipitation, the beads were washed as described (Huppertz et al., 2014), with minor modifications. We performed washes in the cold room (~4ºC) with two wash buffers: a high-salt wash buffer (50 mM Tris-HCl pH 7.5, 1 M NaCl, 1 mM EDTA, 1% Pierce NP-40, 0.1% SDS, 0.5% sodium deoxycholate) and PNK buffer (20 mM Tris-HCl pH 7.5, 10 mM MgCl2, 0.2% Tween 20). The second RNase T1 digestion was then performed on the washed beads at a final concentration of 100 Units/µL in PNK buffer. Samples were incubated in a Thermomixer for 15 minutes at 22ºC shaking at 1,100 rpm, cooled on ice for 5 minutes, and then processed through the remaining iCLIP protocol as described (Huppertz et al., 2014). We confirmed that FOG-3::3xFLAG crosslinked to RNA by visualizing 5' radioactively labeled RNA bound to the FOG-3 protein when antibody was present on the beads (**Figure S4E**). We confirmed immunoprecipitation of FOG-3 from experimental versus negative control samples by immunoblot with an M2 FLAG antibody (**Figure S4F**).

Single-end sequencing was performed on an Illumina HiSeq 2000 (University of Wisconsin Biotechnology Center; Madison, WI). The cDNA library of each replicate was prepared with a unique "Rclip" reverse transcription primer (as in Huppertz et al., 2014), which contained a partially randomized sequence (i.e., a "barcode"). The constant portion of the barcode enabled each read to be identified by replicate and allowed for replicate multiplexing. The randomized portion of the barcode allowed for computational filtering of artifacts from individual reads caused by PCR amplification of the cDNAs, such as read duplication. After high-throughput sequencing, the barcode sequence preceded the cDNA sequence and thus could be easily identified and removed prior to read mapping.

iCLIP sequence analysis

Reads (**Table S2**, Tab 1, column B) were aligned to the WS235 genome using STAR (Dobin et al., 2013) and previously described parameters (Kassuhn et al., 2016), except for the parameter --alignEndsType Local (mismatches at the ends of reads are tolerated). Multi-mapping reads were removed, and high-confidence mappings were selected as those with alignment scores of at least 20 (**Table S2**, Tab 1, column C). PCR duplicates were collapsed to unique reads (**Table S2**, Tab 1, column E) using the method described in Weyn-Vanhentenryck et al. (2014). Reads were assigned to genes using HTSeq (Anders et al., 2015). CIMS (crosslinking induced mutation sites) and CITS (crosslink induced truncation sites) analyses were performed as described (Weyn-Vanhentenryck et al., 2014), except we did not require CIMS to reproduce between replicates, and are included in **Table S3**, tab 3. For peak analysis, "clusters" were defined as regions of overlapping reads. Using the reads indicated in **Table S2**, tab 1, column E, all reads within a gene had their position randomized 1000 times to empirically determine a cluster p value as the odds of having a cluster with the given maximum read depth from randomized read positions. This is similar to the local Poisson method (Zisoulis et al., 2010), as the Poisson approximates of read scrambling. A Benjamini-Hochberg (BH) correction for multiple hypothesis testing was then applied at 1% FDR, resulting in the cluster numbers in **Table S2**, tab 2, column F. Finally, only overlapping clusters called independently as significant in at least 2 of the 4 replicates were retained as reproducible clusters, resulting in the cluster numbers in **Table S2**, tab 2, column G. Final clusters for FOG-3 and control samples are given in **Table S3**. While this is a simple method that does not account for background RNA abundance, it resulted in only 6 clusters for the negative control samples, suggesting it is effective at removing background in our datasets. We define peaks as all maxima at least 5 reads deep and at least 5% of the highest peak in the given gene; we counted neighboring peaks as distinct only if signal dropped to 50% or less of the lower peak maxima. Our definition of peaks differs from our definition of clusters, which are regions of continuous read coverage that pass the 1% FDR threshold. Clusters extend until iCLIP coverage drops to zero, thereby containing any number of distinct signal concentrations, and motivating our separate definitions of peaks and clusters.

To compare our results with previous FOG-3 RIP results (Noble et al., 2016), we calculated overlap with the top 722 FOG-3 targets, and evaluated significance by Fisher's exact test. To determine whether FOG-3 targets were associated with oogenesis, spermatogenesis, or mitosis, we used the method described previously (Noble et al., 2016), with significance evaluated by Fisher's exact test. Figures depicting iCLIP results (**Figure S4M,N**) were generated using Matplotlib (Hunter, 2007) and Python scripts available at https://github.com/dfporter.

We could not find enriched sequence motifs at FOG-3 binding sties. Instead, we sought motifs enriched near the binding sites and found enrichment of a CUCAC motif (**Figure S4P,** p value < 1.8×10^{-229}). CUCA is part of the GLD-1/STAR signature motif (Ryder et al., 2004). GLD-1 regulates germline sex determination, but it can promote either the sperm or oocyte fate (Francis et al., 1995).

Raw sequence files of all replicates are available through Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/; GEO accession GSE76521).

SUPPLEMENTAL REFERENCES

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., *et al.* (2010). *PHENIX*: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr *66*, 213-221.

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166-169.

Arribere, J.A., Bell, R.T., Fu, B.X., Artiles, K.L., Hartman, P.S., and Fire, A.Z. (2014). Efficient marker-free recovery of custom genetic modifications with CRISPR/Cas9 in *Caenorhabditis elegans*. Genetics *198*, 837-846.

Aslanidis, C., and de Jong, P.J. (1990). Ligation-independent cloning of PCR products (LIC-PCR). Nucleic Acids Res *18*, 6069-6074.

Dickinson, D.J., Ward, J.D., Reiner, D.J., and Goldstein, B. (2013). Engineering the *Caenorhabditis elegans* genome using Cas9-triggered homologous recombination. Nat Methods *10*, 1028-1034.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21.

Emsley, P., and Cowtan, K. (2004). *Coot*: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr *60*, 2126-2132.

Ericsson, U.B., Hallberg, B.M., Detitta, G.T., Dekker, N., and Nordlund, P. (2006). Thermofluor-based high-throughput stability optimization of proteins for structural studies. Anal Biochem *357*, 289-298.

Francis, R., Barton, M.K., Kimble, J., and Schedl, T. (1995). *gld-1*, a tumor suppressor gene required for oocyte development in *Caenorhabditis elegans*. Genetics *139*, 579-606.

Gibson, D.G. (2009). Synthesis of DNA fragments in yeast by one-step assembly of overlapping oligonucleotides. Nucleic Acids Res *37*, 6984-6990.

Hunter, J.D. (2007). Matplotlib: A 2D graphics environment. Computing In Science & Engineering *9*, 90-95.

Kassuhn, W., Ohler, U., and Drewe, P. (2016). Cseq-Simulator: A data simulator for CLIP-Seq experiments. Pac Symp Biocomput *21*, 433-444.

Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. J Mol Biol *372*, 774-797.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). *Phaser* crystallographic software. J Appl Crystallogr *40*, 658-674.

Paix, A., Folkmann, A., Rasoloson, D., and Seydoux, G. (2015). High efficiency, homology-directed genome editing in *Caenorhabditis elegans* using CRISPR-Cas9 ribonucleoprotein complexes. Genetics *201*, 47-54.

Ryder, S.P., Frater, L.A., Abramovitz, D.L., Goodwin, E.B., and Williamson, J.R. (2004). RNA target specificity of the STAR/GSG domain post-transcriptional regulatory protein GLD-1. Nat Struct Mol Biol *11*, 20-28.

Schindelin, J., Rueden, C.T., Hiner, M.C., and Eliceiri, K.W. (2015). The ImageJ ecosystem: An open platform for biomedical image analysis. Mol Reprod Dev *82*, 518-529.

Spitzer, J., Hafner, M., Landthaler, M., Ascano, M., Farazi, T., Wardle, G., Nusbaum, J., Khorshid, M., Burger, L., Zavolan, M., *et al.* (2014). PAR-CLIP (Photoactivatable Ribonucleoside-Enhanced Crosslinking and Immunoprecipitation): a step-by-step protocol to the transcriptome-wide identification of binding sites of RNA-binding proteins. Methods Enzymol *539*, 113-161.

Stiernagle, T. (2006). Maintenance of *C. elegans*. In WormBook, T.*C.e.*R. Community, ed. (WormBook).

Wang, K.C., Yang, Y.W., Liu, B., Sanyal, A., Corces-Zimmerman, R., Chen, Y., Lajoie, B.R., Protacio, A., Flynn, R.A., Gupta, R.A., *et al.* (2011). A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. Nature *472*, 120-124.

Weyn-Vanhentenryck, S.M., Mele, A., Yan, Q., Sun, S., Farny, N., Zhang, Z., Xue, C., Herre, M., Silver, P.A., Zhang, M.Q.*, et al.* (2014). HITS-CLIP and integrative modeling define the Rbfox splicing-regulatory network linked to brain development and autism. Cell Rep *6*, 1139-1152.
Zeiser, E., Frøkjær-Jensen, C., Jorgensen, E., and Ahringer, J. (2011). MosSCI and gateway compatible plasmid toolkit for constitutive and inducible expression of transgenes in the *C. elegans* germline. PLoS One *6*, e20082.
Zisoulis, D.G., Lovci, M.T., Wilbert, M.L., Hutt, K.R., Liang, T.Y., Pasquinelli, A.E., and Yeo, G.W. (2010). Comprehensive discovery of endogenous Argonaute binding sites in *Caenorhabditis elegans*. Nat Struct Mol Biol *17*, 173-179.

Figure S1

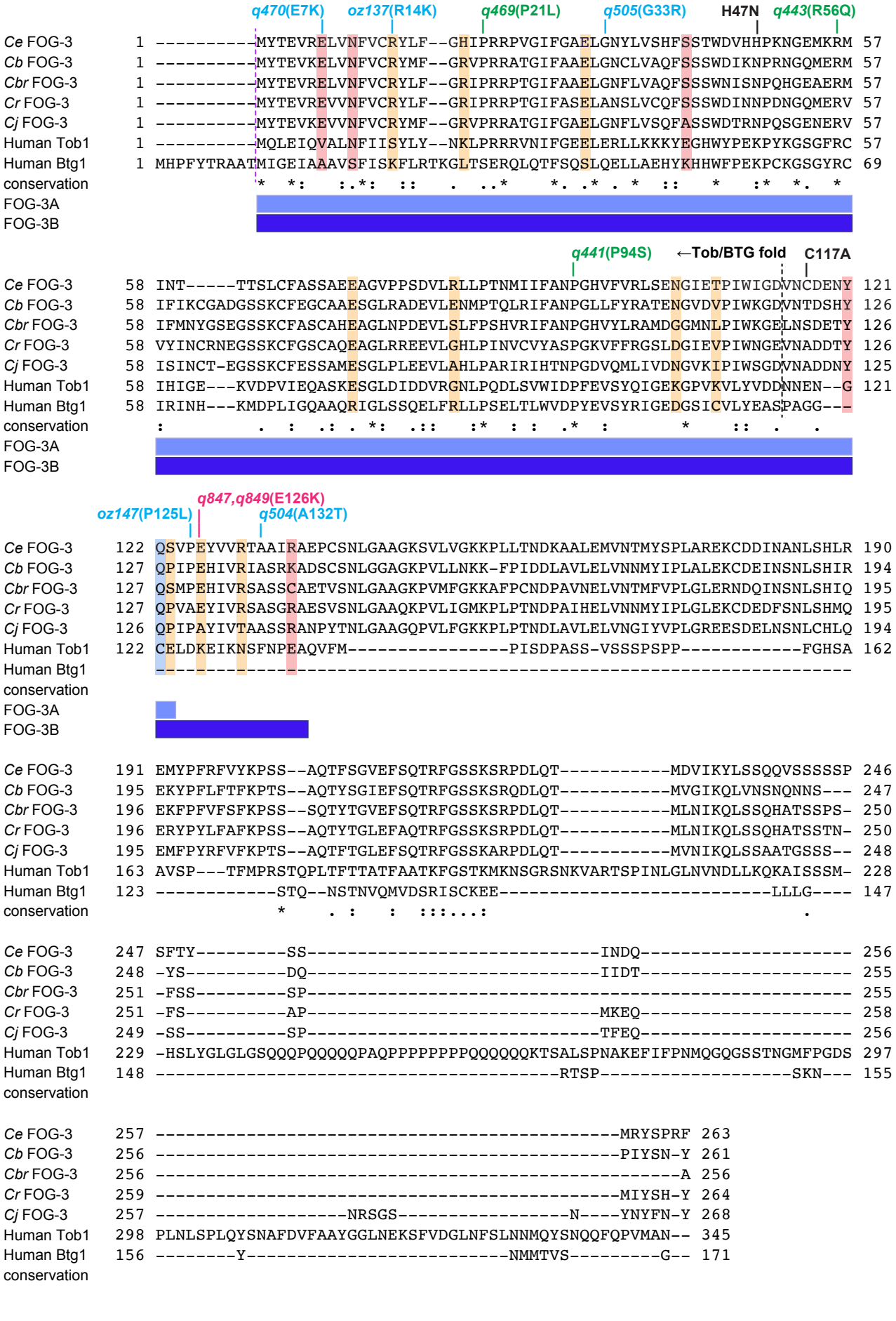
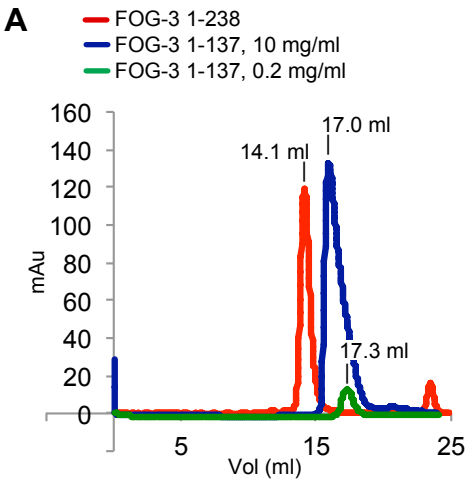
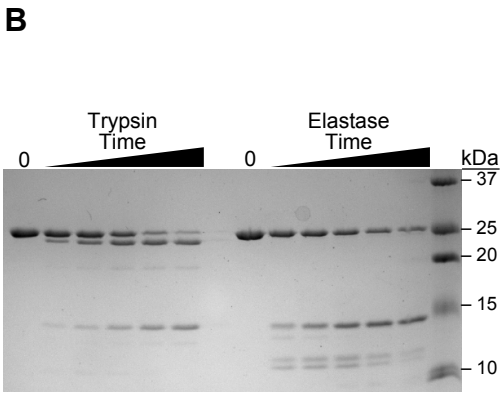
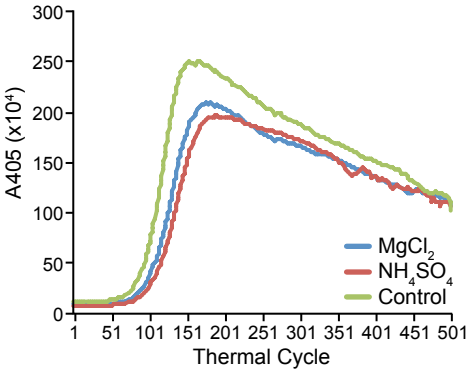

q470(E7K)
oz137(R14K)
q469(P21L)
q505(G33R)
H47N
q443(R56Q)
Ce FOG-3 1 ----------MYTEVRELVNFVCRYLF--GHIPRRPVGIFGAELGNYLVSHFSSTWDVHHPKNGEMKRM 57
Cb FOG-3 1 ----------MYTEVKELVNFVCRYMF--GRVPRRATGIFAAELGNCLVAQFSSSWDIKNPRNGQMERM 57
Cbr FOG-3 1 ----------MYTEVRELVNFVCRYLF--GRIPRRPTGIFAAELGNFLVAQFSSSWNISNPQHGEAERM 57
Cr FOG-3 1 ----------MYTEVREVVNFVCRYLF--GRIPRRPTGIFASELANSLVCQFSSSWDINNPDNGQMERV 57
Cj FOG-3 1 ----------MYTEVKEVVNFVCRYMF--GRVPRRATGIFGAELGNFLVSQFASSWDTRNPQSGENERV 57
Human Tob1 1 ----------MQLEIQVALNFIISYLY--NKLPRRRVNIFGEELERLLKKKYEGHWYPEKPYKGSGFRC 57
Human Btg1 1 MHPFYTRAATMIGEIAAAVSFISKFLRTKGLTSERQLQTFSQSLQELLAEHYKHHWFPEKPCKGSGYRC 69
conservation
FOG-3A
FOG-3B
q441(P94S)
←Tob/BTG fold
C117A
Ce FOG-3 58 INT-----TTSLCFASSAEEAGVPPSDVLRLLPTNMIIFANPGHVFVRLSENGIETPIWIGDVNCDENY 121
Cb FOG-3 58 IFIKCGADGSSKCFEGCAAESGLRADEVLENMPTQLRIFANPGLLFYRATENGVDVPIWKGDVNTDSHY 126
Cbr FOG-3 58 IFMNYGSEGSSKCFASCAHEAGLNPDEVLSLFPSHVRIFANPGHVYLRAMDGGMNLPIWKGELNSDETY 126
Cr FOG-3 58 VYINCRNEGSSKCFGSCAQEAGLRREEVLGHLPINVCVYASPGKVFFRGSLDGIEVPIWNGEVNADDTY 126
Cj FOG-3 58 ISINCT-EGSSKCFESSAMESGLPLEEVLAHLPARIRIHTNPGDVQMLIVDNGVKIPIWSGDVNADDNY 125
Human Tob1 58 IHIGE---KVDPVIEQASKESGLDIDDVRGNLPQDLSVWIDPFEVSYQIGEKGPVKVLYVDDNNEN--G 121
Human Btg1 58 IRINH---KMDPLIGQAAQRIGLSSQELFRLLPSELTLWVDPYEVSYRIGEDGSICVLYEASPAGG---
conservation
FOG-3A
FOG-3B
q847,q849(E126K)
oz147(P125L)
q504(A132T)
Ce FOG-3 122 QSVPEYVVRTAAIRAEPCSNLGAAGKSVLVGKKPLLTNDKAALEMVNTMYSPLAREKCDDINANLSHLR 190
Cb FOG-3 127 QPIPEHIVRIASRKADSCSNLGGAGKPVLLNKK-FPIDDLAVLELVNNMYIPLALEKCDEINSNLSHIR 194
Cbr FOG-3 127 QSMPEHIVRSASSCAETVSNLGAAGKPVMFGKKAFPCNDPAVNELVNTMFVPLGLERNDQINSNLSHIQ 195
Cr FOG-3 127 QPVAEYIVRSASGRAESVSNLGAAQKPVLIGMKPLPTNDPAIHELVNNMYIPLGLEKCDEDFSNLSHMQ 195
Cj FOG-3 126 QPIPAYIVTAASSRANPYTNLGAAGQPVLFGKKPLPTNDLAVLELVNGIYVPLGREESDELNSNLCHLQ 194
Human Tob1 122 CELDKEIKNSFNPEAQVFM----------PISDPASS-VSSSPSPP----------FGHSA 162
Human Btg1 ---
conservation
FOG-3A
FOG-3B
Ce FOG-3 191 EMYPFRFVYKPSS--AQTFSGVEFSQTRFGSSKSRPDLQT----------MDVIKYLSSQQVSSSSSP 246
Cb FOG-3 195 EKYPFLFTFKPTS--AQTYSGIEFSQTRFGSSKSRQDLQT----------MVGIKQLVNSNQNNS--- 247
Cbr FOG-3 196 EKFPFVFSFKPSS--SQTYTGVEFSQTRFGSSKSRPDLQT----------MLNIKQLSSQHATSSPS- 250
Cr FOG-3 196 ERYPYLFAFKPSS--AQTYTGLEFAQTRFGSSKSRPDLQT----------MLNIKQLSSQHATSSTN- 250
Cj FOG-3 195 EMFPYRFVFKPTS--AQTFTGLEFSQTRFGSSKARPDLQT----------MVNIKQLSSAATGSSS-- 248
Human Tob1 163 AVSP---TFMPRSTQPLTFTTATFAATKFGSTKMKNSGRSNKVARTSPINLGLNVNDLLKQKAISSSM- 228
Human Btg1 123 ----------STQ--NSTNVQMVDSRISCKEE----------LLLG---- 147
conservation
Ce FOG-3 247 SFTY---------SS---------INDQ---------- 256
Cb FOG-3 248 -YS----------DQ---------IIDT---------- 255
Cbr FOG-3 251 -FSS---------SP---------- 255
Cr FOG-3 251 -FS----------AP---------MKEQ---------- 258
Cj FOG-3 249 -SS----------SP---------TFEQ---------- 256
Human Tob1 229 -HSLYGLGLGSQQQPQQQQQQPAQPPPPPPPPQQQQQQQKTSALSPNAKEFIFPNMQGQGSSTNGMFPGDS 297
Human Btg1 148 ----------RTSP----------SKN--- 155
conservation
Ce FOG-3 257 ----------MRYSPRF 263
Cb FOG-3 256 ----------PIYSN-Y 261
Cbr FOG-3 256 ----------A 256
Cr FOG-3 259 ----------MIYSH-Y 264
Cj FOG-3 257 ----------NRSGS----------N----YNYFN-Y 268
Human Tob1 298 PLNLSPLQYSNAFDVFAAYGGLNEKSFVDGLNFSLNNMQYSNQQFQPVMAN-- 345
Human Btg1 156 --------Y----------NMMTVS---------G-- 171
conservation

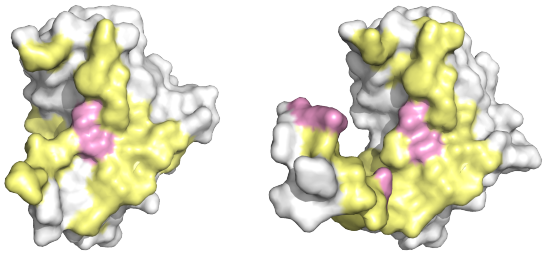
Figure S1. Alignment of FOG-3 ortholog sequences, related to Figure 1. Amino acid sequence alignment of FOG-3 orthologs, including human Tob and BTG proteins. Nematode orthologs include FOG-3 from *C. elegans* (*Ce*), *C. briggsae* (*Cb*), *C. brenneri* (*Cbr*), *C. remanei* (*Cr*), and *C. japonica* (*Cj*). Alignment by T-Coffee (Di Tommaso et al., 2011). Conservation noted by identity (*) plus high (:) or moderate (.) similarity. Missense alleles are labeled with their amino acid changes (Chen et al., 2000; this work); allele labels are colored to mark conservation among most metazoan orthologs (green), conservation among most nematode orthologs (blue) and a mutation generated in this study (magenta). Boundary of the canonical Tob/BTG fold is marked with a dashed line; extents of dimer subunits are shown below, including subunit A (light blue) and subunit B (dark blue). Amino acids highlighted indicate dimer subunit-subunit contacts (red), dimer-dimer contacts (orange), and both (blue); these contacts include both hydrogen bonds and salt bridges.

Figure S2









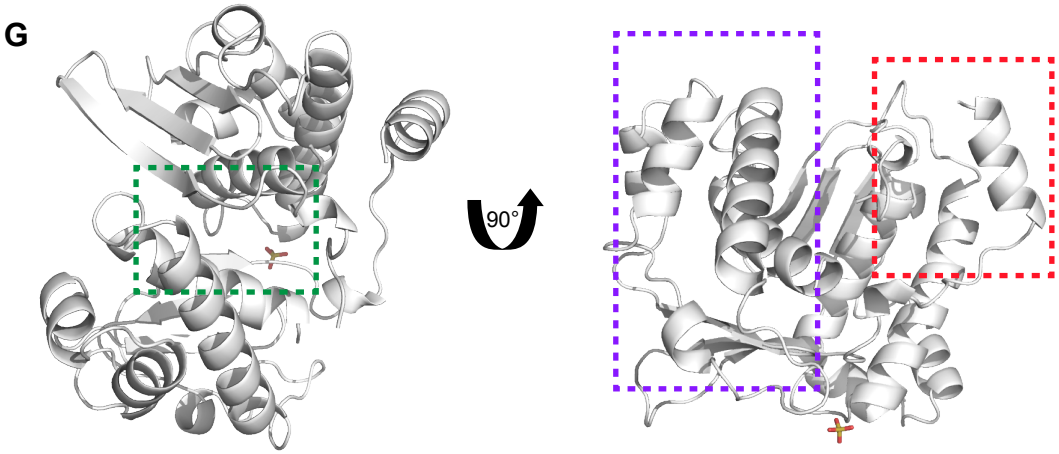
E

Dimer subunit: H-bonds

A	B
E7	N10
N10	E7
S41	R135
Y121	Q122

F

Mutation	Conservation	Site in structure	Proposed defect
P21L	All Tob/BTG folds	Tob/BTG fold	Tob/BTG Folding
R56Q	All Tob/BTG folds	Tob/BTG fold	Tob/BTG Folding
P94S	All Tob/BTG folds	Tob/BTG fold	Tob/BTG Folding
E7K	FOG-3 Tob/BTG fold	Dimer interface	Dimerization
G33K	FOG-3 Tob/BTG fold	Dimer interface	Dimerization
R14K	FOG-3 Tob/BTG fold	Dimer interface	Dimer/polymer
P125L	FOG-3 extension	Linker-helix	Dimer/polymer
A132T	FOG-3 extension	Linker-helix	Dimer/polymer
E126K	FOG-3 extension	Linker-helix	Polymerization



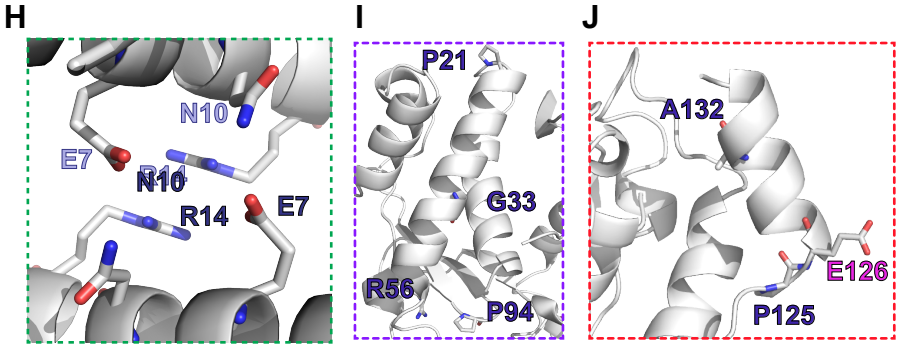
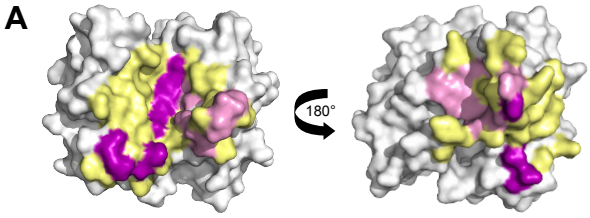


Figure S2. FOG-3 biochemical characterization, dimer subunit protein contacts and missense mutants, related to Figure 1. (**A**) Size exclusion chromatography elution profile of recombinant FOG-3 protein. Red, amino acids 1-238 with its histidine tag at 10 mg/ml; blue, amino acids 1-137 with histidine tag at 10 mg/ml; green, amino acids 1-137 with histidine tag at 0.2 mg/ml. Position of max peaks labeled. A280 milliabsorbance units, mAU. (**B**) Protease mapping the FOG-3 Tob/BTG-containing domain. FOG-3 1-238 was incubated with either trypsin or elastase, and samples were collected over time. Incubation with either protease produced a cleavage product of ~15 kDa. Protein incubated without protease labeled as "0." (**C**) Thermal-folding assay reveals domain stabilization with magnesium and sulfate. See Supplemental Experimental Procedures and main text for further details. (**D**) Contact surface between FOG-3 dimer subunits. Residues making hydrogen bonds (pink) and other contacting residues (yellow) based on distance. (**E**) Table of subunit-subunit hydrogen bonds. (**F**) Summary of FOG-3 missense mutations. (**G**) FOG-3 dimer with boxed regions enlarged in **H-J**. (**H-J**) Residues from FOG-3A are colored light blue and those from FOG-3B are dark blue. (**H**) Sites of *fog-3* missense mutants at the Tob/BTG domain dimer interface. (**I**) Sites of *fog-3* missense mutants conserved across FOG-3 orthologs and human Tob proteins. (**J**) Sites of *fog-3* missense mutants in linker-helix extension, including one generated in this study (magenta).



B

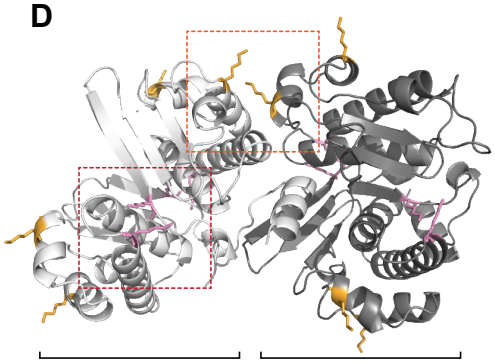
Dimer-dimer H-bonds

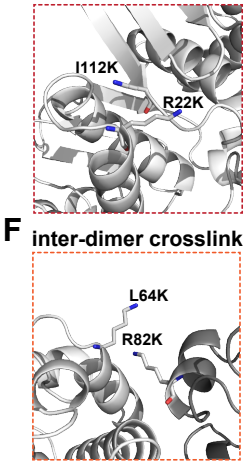
Dimer 1	Dimer 2
A:F11	B:R130
A:E103	A:S123
A:N104	A:D118
A:N104	A:Y121
A:N104	A:Q122
A:G105	A:Q122
A:G105	B:Q122
A:E107	B:Q122
A:E107	A:V124
A:T108	B:V124

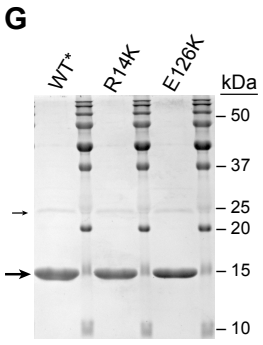
C

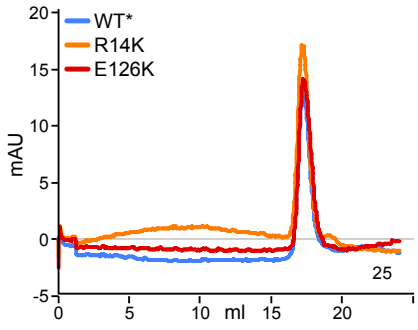
Dimer-dimer salt bridges

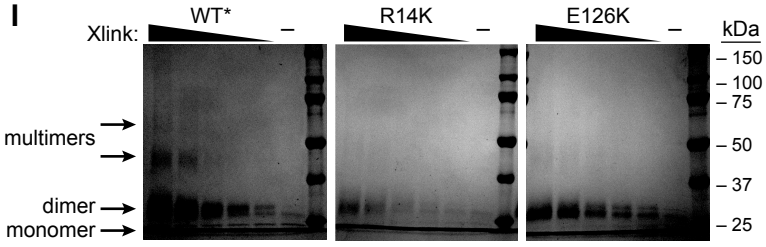
Dimer 1	Dimer 2
A:R14	B:E126
B:R14	B:E126
B:H19	A:E31
B:R82	A:E72

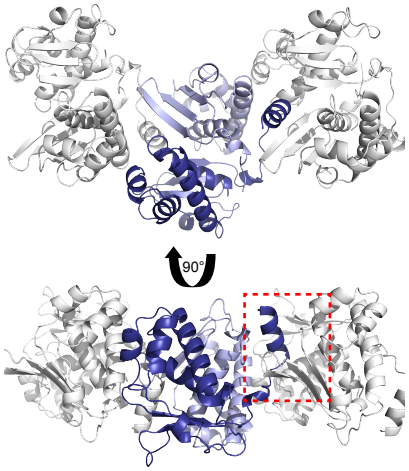


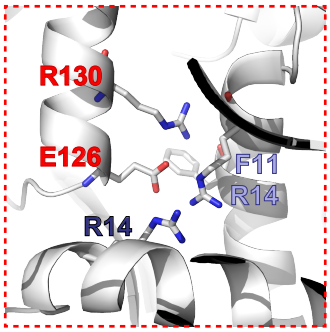






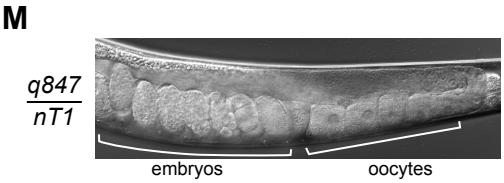






L

genotype	% fertile	% Fog	n
fog-3(q847)/nT1	100	0	105
fog-3(q847)/(q847)	0	100	106
fog-3(q849)/nT1	100	0	107
fog-3(q849)/(q849)	0	100	110



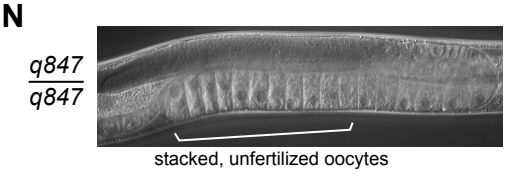
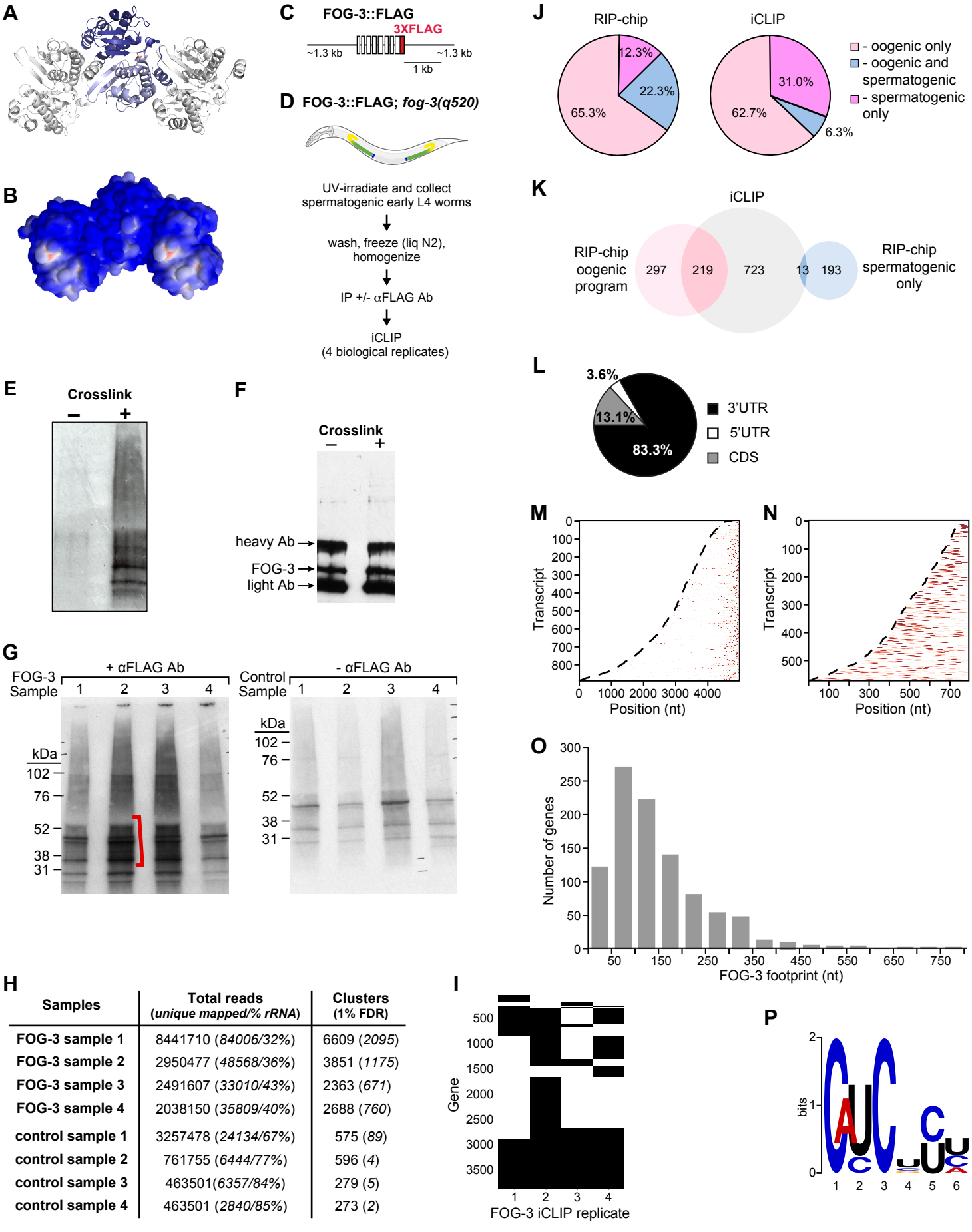


Figure S3. Dimer-dimer subunit protein contacts and supplemental biochemical analyses, related to Figure 2. (**A**) Interacting residues between FOG-3 dimers. Salt bridges (magenta), hydrogen bonds (pink), and contacting residues (yellow). (**B**) Table of predicted dimer-dimer hydrogen bonds. (**C**) Table of predicted dimer-dimer salt bridges. (**D-F**) Lysine substitution of recombinant FOG-3 generates residues permitting intra- and inter-dimer crosslinks. Location of lysine substitutions in the context of two FOG-3 dimers (**D**), which are enlarged in **E** (red box) and **F** (orange box). **(E)** Lysine substitutions R22K and I112K facilitate an intra-dimer crosslink. **(F)** Lysine substitutions L64K and R82K facilitate an inter-dimer crosslink. (**G**) Coomassie-stained gel of purified FOG-3 recombinant proteins. Left, "wild type" (WT*) FOG-3 (1-140, H47N C117A) with lysine substitutions; middle, missense mutant R14K predicted to abrogate dimer formation; right, missense mutant E126K predicted to abrogate polymer formation. All proteins (10 µg each) ran at ~15 kDa (large arrow). In addition, a minor ~25 kDa contaminant was observed (small arrow). (**H**) Size exclusion chromatography elution profile of WT*, R14K, and E126K recombinant FOG-3 (0.2 mg/ml). A280 milli-absorbance units, mAU. (**I**) Different exposure of **Figure 2D**. Coomassie-stained gel of modified FOG-3 recombinant protein incubated with increasing amounts of BS3 crosslinker. "–" represents no BS3 included. (**J**) Crystal packing of the FOG-3 dimer. FOG-3A and FOG-3B in the asymmetric unit are represented in light and dark blue, respectively. Each dimer buries the helix extension in FOG-3B into the adjacent dimer. The helix extension (red box) is enlarged in **K**. (**K**) Packing of the helix extension of one FOG-3 subunit into the adjacent dimer. Amino acids in subunits colored as in **J**, except for red residues from helix extension of the adjacent dimer. (**L**) Mutation of a key residue in the helix extension sexually transforms the germline (Fog phenotype). Two identical but independently-generated CRISPR-Cas9 alleles (*q847* and *q849*) mutated glutamate 126 to ~~a~~ lysine (E126K). Alleles were placed over a GFP-expressing balancer (*nT1*). Heterozygous (green) and homozygous (non-green) L4 worms were singled and analyzed 3 and 4 days later for fertility and the Fog phenotype. (**M-N**) Representative DIC images of hermaphrodite adults, heterozygous (**M**) or homozygous (**N**) for the E126K mutation. The embryos in heterozygotes demonstrate fertility, while oocytes stacking in homozygotes demonstrate sterility due to lack of sperm and hence lack of embryo production.

Figure S4



H

Samples	Total reads (*unique mapped/% rRNA*)	Clusters (1% FDR)
FOG-3 sample 1	8441710 (*84006/32%*)	6609 (*2095*)
FOG-3 sample 2	2950477 (*48568/36%*)	3851 (*1175*)
FOG-3 sample 3	2491607 (*33010/43%*)	2363 (*671*)
FOG-3 sample 4	2038150 (*35809/40%*)	2688 (*760*)
control sample 1	3257478 (*24134/67%*)	575 (*89*)
control sample 2	761755 (*6444/77%*)	596 (*4*)
control sample 3	463501(*6357/84%*)	279 (*5*)
control sample 4	463501 (*2840/85%*)	273 (*2*)

Figure S4. FOG-3 binds RNA *in vivo,* related to Figure 3 and Experimental Procedures. (**A**) Model of FOG-3 multimer composed of three dimers; subunits colored as in **Figure 1**. (**B**) Electrostatic surface potential of multimer modeled in **A**. Blue, basic; red, acidic. (**C**) Diagram of rescuing, epitope-tagged FOG-3::3xFLAG transgene, adapted from (Noble et al., 2016). (**D**) Outline of FOG-3 iCLIP protocol. See Supplemental Experimental Procedures for details. (**E-F**) FOG-3 crosslinks with RNA *in vivo*. Living worms expressing FOG-3::3xFLAG were UV-crosslinked (+) or mock treated (-), and then immunoprecipitated with αFLAG antibody. Bound sample 5' radiolabeled and run on SDS-PAGE. (**E**) Overnight exposure of radiolabeled samples run on SDS-PAGE. (**F**) Immunoblot of FOG-3::3xFLAG immunoprecipitation samples, visualized with αFLAG antibody. FOG-3 and heavy and light antibody chains labeled. (**G**) Gel analysis of samples used for iCLIP. Each sample was immunoprecipitated with (+) or without (-) αFLAG antibody, radiolabeled and run on the SDS-PAGE gel. The region above the expected size for FOG-3::3xFLAG (red bracket) was used for iCLIP processing and sequencing. Samples include four biological replicates and their paired controls. (**H**) FOG-3 iCLIP reads and cluster (regions of overlapping reads) statistics. Unique mapped reads (middle column) were determined by mapping with STAR, filtering out multimapping reads and low confidence alignments, and collapsing duplicate reads (see Supplemental Experimental Procedures). The fraction of unique mapped reads that mapped to rRNA is given as a percentage. The number of significant clusters at FDR 1% (right column) is highly dependent on FOG-3 purification. See Supplemental Experimental Procedures for further details. (**I**) Targets (black) identified for separate FOG-3 replicates. (**J**) Comparison of FOG-3 targets identified using FOG-3 iCLIP versus FOG-3 RIP-chip (microarray). Targets belonging to the oogenic program include mRNAs found only in oogenic germlines as well as mRNAs found in both oogenic and spermatogenic germlines, as described (Noble et al., 2016). (**K**) Venn diagram of mRNA target overlap between FOG-3 iCLIP (this study) and FOG-3 RIP-chip (FOG-3 IP with microarray analysis of associated RNAs) (Noble et al., 2016). (**L**) Percentages of FOG-3 iCLIP reads in mRNA regions. (**M-N**) FOG-3 binding sites are represented on a heat map, from no signal (white) to strong signal (red). Only genes with annotated 3'UTRs of at least 50 nt were included. (**M**) FOG-3 binding sites are at the 3' end of mRNAs, which are arranged by predicted nucleotide (nt) length from 5' to 3'. Dashed line marks the 5' end. Note prevalence of FOG-3 binding sites at the 3' ends of transcripts. (**N**) Binding sites occur throughout 3'UTRs, which are arranged by predicted 3'UTR nucleotide (nt) length from stop codon (dashed line) to 3' end. (**O**) FOG-3 iCLIP footprint on target transcripts. Coverage includes transcript regions above two reads deep. (**P**) Motif analysis of iCLIP clusters. Analysis and image generated by MEME (Bailey et al., 2015), except T was replaced by U.

Table S1. Data collection and refinement statistics, related to Figure 1.

	C. elegans FOG-3
PDB ID	5TD6
Wavelength (Å…)	0.9537
Resolution range (Å…)	29.22 - 2.034 (2.106 - 2.034)
Space group	P 31 2 1
Unit cell	64.992 64.992 133.57 90 90 120
Total reflections	147208 (14689)
Unique reflections	21617 (2091)
Multiplicity	6.8 (7.0)
Completeness (%)	99.79 (98.18)
Mean I/sigma(I)	24.12 (2.56)
Wilson B-factor	39.54
R-merge	0.0498 (0.7383)
R-meas	0.05399
CC1/2	1 (0.789)
CC*	1 (0.939)
Reflections used for R-free	1993
R-work	0.1769 (0.2255)
R-free	0.2285 (0.2886)
Number of non-hydrogen atoms	2210
macromolecules	2067
ligands	5
water	138

Protein residues	260
RMS(bonds)	0.009
RMS(angles)	1.10
Ramachandran favored (%)	98
Ramachandran allowed (%)	2
Ramachandran outliers (%)	0
Clashscore	3.89
Average B-factor	46.60
macromolecules	46.70
ligands	32.50
solvent	45.50

Statistics for the highest-resolution shell are shown in parentheses.