SUPPLEMENTAL MATERIAL

Drosophila Bag-of-marbles directly interacts with the CAF40 subunit of the CCR4-NOT complex to elicit repression of mRNA targets

Annamaria Sgromo, Tobias Raisch, Charlotte Backhaus, Csilla Keskeny, Vikram Alva, Oliver Weichenrieder and Elisa Izaurralde

Supplemental Table S1. Constructs and mutants used in this study.

Name	**Bag of marbles** (Uniprot P22745)	**Comment**
Bam	λN-HA-Bam 1–442	
	SBP-Bam 1–442	
	MS2-HA-Bam 1–442	
	MBP-Bam 1–442	
Bam-N	λN-HA-Bam 1–140	
	GFP-Bam 1–140	
	MS2-HA-Bam 1–140	
	MBP-Bam 1–140	
Bam-C	λN-HA-Bam 141–442	
	GFP-Bam 141–442	
	MBP-Bam 141–442	
CBM	λN-HA-GST-Bam 13–36	
	SBP-MBP-Bam 13–36	
	MS2-HA-MBP-Bam 13–36	
	MBP-Bam 13–36	
	His_6-NusA-Bam 13–36	
ΔCBM	λN-HA-Bam Δ13–36	Disrupts CAF40 binding
	MBP-Bam Δ13–36	Disrupts CAF40 binding
L17E	λN-HA-Bam L17E	Disrupts CAF40 binding
	MBP-Bam 1–442 L17E	Disrupts CAF40 binding
M24E	λN-HA-Bam M24E	Disrupts CAF40 binding
	MBP-Bam 1–442 M24E	Disrupts CAF40 binding
L28E	λN-HA-Bam L28E	Disrupts CAF40 binding
V32E	λN-HA-Bam V32E	Disrupts CAF40 binding
4xMut	λN-HA-Bam L17E, M24E, L28E, V32E	Disrupts CAF40 binding
	SBP-Bam L17E, M24E, L28E, V32E	Disrupts CAF40 binding
	MS2-HA-Bam L17E, M24E, L28E, V32E	Disrupts CAF40 binding

Name	***Hs* NOT1** (Uniprot A5YKK6)	**Comment**
NOT1 N	*Hs* NOT1 1–1000	
NOT1 MIF4G	His_6-*Hs* NOT1 1093–1317	MIF4G-like domain
NOT1 CN9BD	MBP-*Hs* NOT1 1351–1588	CNOT9-binding domain
NOT1 MIF4G-C	MBP-*Hs* NOT1 1607–1815	Predicted MIF4G-like domain
NOT1 SHD	MBP-*Hs* NOT1 1833–2361	Superfamily homology domain

Name	***Hs* NOT2** (Uniprot Q9NZN8)	**Comment**
NOT2-C	MBP-*Hs* NOT2 350–540	

Name	***Hs* NOT3** (Uniprot O75175)	**Comment**
NOT3-N	MBP-*Hs* NOT3 2–212	
NOT3-C	His_6-*Hs* NOT3 607–748	

Name	***Hs* CCR4a** (Uniprot Q9ULM6)	**Comment**
CCR4a full-length	MBP-*Hs* CCR4a	

Name	***Hs* CAF1** (Uniprot Q9UIV1)	**Comment**
CAF1 full-length	MBP-*Hs* CAF1	

Name	***Hs* CAF40** (Uniprot Q92600)	**Comment**
CAF40-ARM wt	His_6-*Hs* CAF40 19–285	
	GST-*Hs* CAF40 19–285	
CAF40 wt	SBP-MBP-CAF40 1–299	
V181E	SBP-MBP-CAF40 1–299 V181E	

Name	***Hs* NOT10** (Uniprot Q9H9A5)	**Comment**
NOT10 TPR	*Hs* NOT10 25–707	

Name	***Hs* NOT11** (Uniprot Q9UKZ1)	**Comment**
NOT11-C	*Hs* NOT11 257–498-His_6	

Name	***Dm* CAF40 (1–304)** (Uniprot Q7JVP2)	**Comment**
CAF40 wt	λN-HA-CAF40 1–304	
	GFP-CAF40 1–304	dsRNA resistant
CAF40-ARM wt	His_6-CAF40 25–291	
V186E	GFP-CAF40 V186E	Disrupts CBM binding; dsRNA resistant
	His_6-CAF40 25–291 V186E	Disrupts CBM binding
2xMut	His_6-CAF40 25–291 Y139E, G146E	Disrupts CBM binding

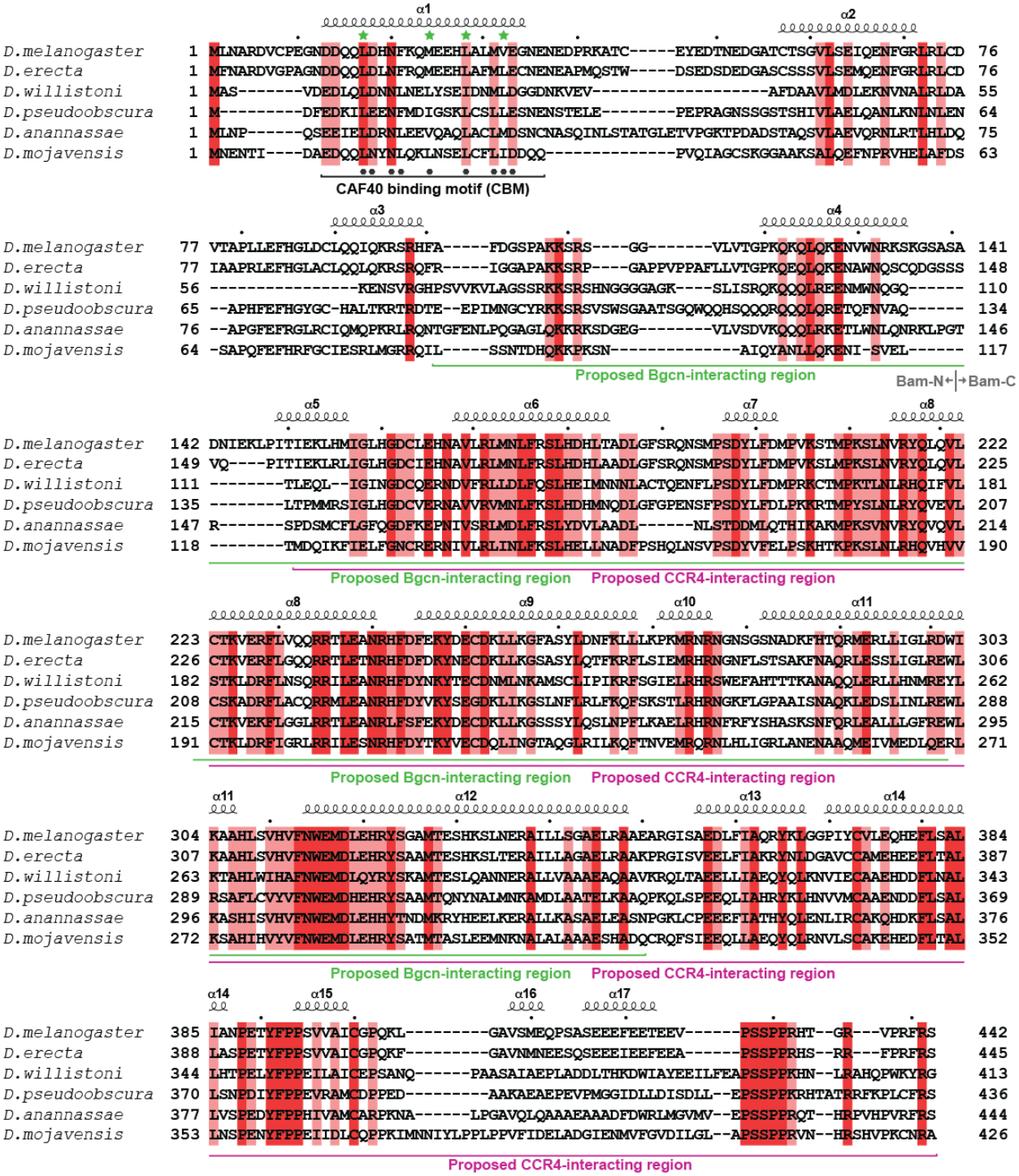
Name	***Dm* NOT1** (Uniprot A8DY81)	**Comment**
NOT1	λN-HA-NOT1	
NOT1-CN9BD	λN-HA-NOT1 1467–1719	CAF40-binding domain
	MBP-NOT1 1468–1719	CAF40-binding domain

Supplemental Table S2. Antibodies used in this study.

Antibody	**Source**	**Catalog Number**	**Dilution**	**Monoclonal/ Polyclonal**
Anti-HA-HRP	Roche	12 013 819 001	1:5,000	Monoclonal
Anti-GFP (for western blotting)	Roche	11 814 460 001	1:2,000	Mouse Monoclonal
Anti-GFP (for immunoprecipitation)	In house			Rabbit Polyclonal
Anti-*Dm* NOT1	Kind gift from E. Wahle	T6199	1:1,000	Rabbit Polyclonal
Anti-*Dm* CAF40	In house		1:1,000	Rabbit Polyclonal
Anti-*Hs* NOT1	In house		1:2,000	Rabbit Polyclonal
Anti-*Hs* NOT2	Bethyl	A302-562A	1:2,000	Rabbit Polyclonal
Anti-*Hs* NOT3	Abcam	Ab55681	1:2,000	Monoclonal
Anti-*Hs* CAF40 (RQCD1)	Proteintech	22503-1-AP	1:1,000	Rabbit Polyclonal
Anti-tubulin	Sigma-Aldrich	T6199	1:10,000	Monoclonal
Anti-*Dm* PABP	In house		1:10,000	Rabbit Polyclonal
Anti-V5	AbD Serotec	MCA1360GA	1:5,000	Monoclonal
Anti-mouse-HRP	GE Healthcare	NA931V	1:10,000	Monoclonal

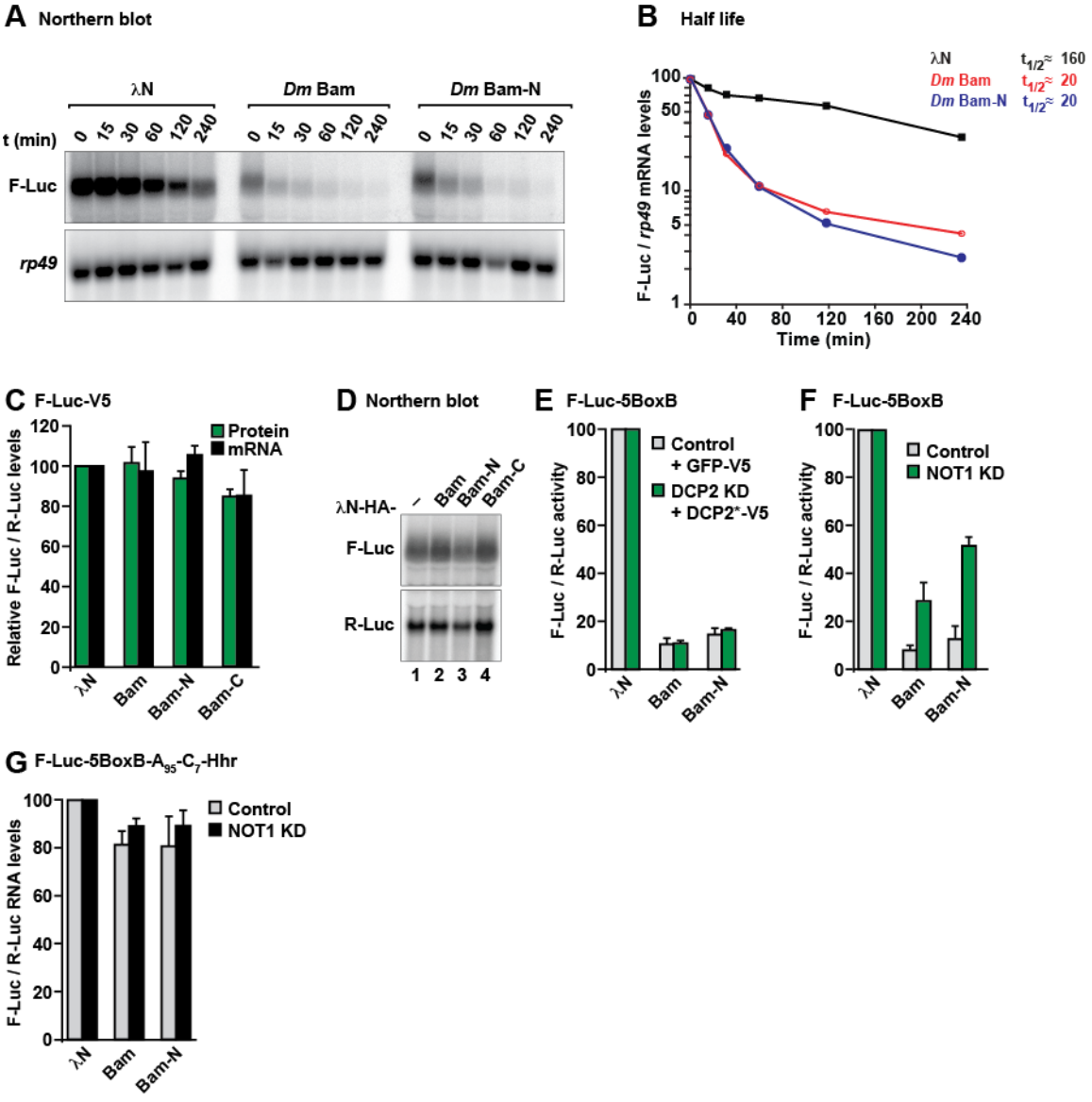
SUPPLEMENTAL FIGURES

Supplemental Figure S1



Supplemental Figure S1. Sequence alignment of *Drosophila* Bam. The secondary structure elements, as predicted by PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred/), are indicated in black. Residues conserved in all aligned sequences are shown with a dark red background, and residues with >70% similarity are highlighted in light red; conservation scores were calculated using the SCORECONS webserver (Valdar 2002). The CAF40-binding motif (CBM) is indicated. Black dots indicate residues in the CBM that directly contact CAF40. Green asterisks indicate residues mutated in this study.

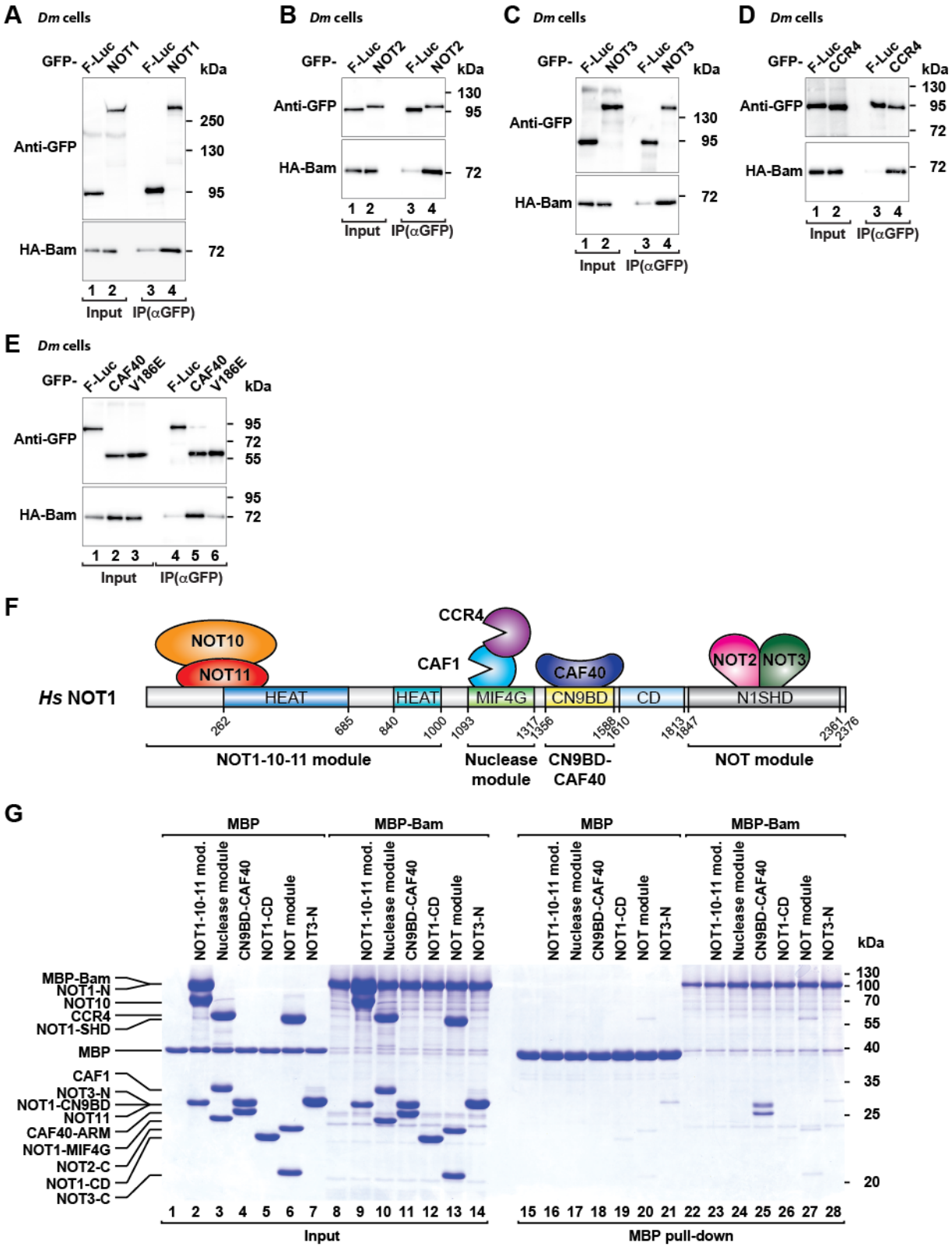
Supplemental Figure S2



Supplemental Figure S2. Bam promotes mRNA degradation. (*A*) Representative northern blot showing the decay of the F-Luc-5BoxB mRNA in S2 cells expressing λN-HA or λN-HA-tagged Bam or the Bam-N fragment after inhibition of transcription by actinomycin D. (*B*) F-Luc mRNA levels were normalized to those of the *rp49* mRNA and plotted against time. (*C,D*) Tethering assay using the F-Luc reporter lacking BoxB sites and λN-HA-tagged Bam (full-length or the indicated fragments) in S2 cells. The samples were analyzed as described in Figure 1B–D.

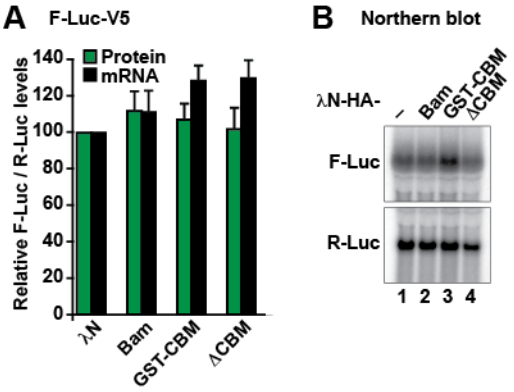
The corresponding experiment with the F-Luc-5BoxB reporter is shown in Figure 1B. (**E**) Normalized F-Luc activity values corresponding to the experiment described in Figure 2A and B. The F-Luc-5BoxB activity was normalized to that of the R-Luc transfection control and set to 100% in cells expressing the λN-HA peptide. The grey and green bars represent the normalized F-Luc-5BoxB activity in control cells expressing GFP-V5 and in DCP2-depleted cells expressing GFP-DCP2*-V5, respectively. (*F*) Normalized F-Luc-5BoxB activity values corresponding to the experiment described in Figure 2D and E. (G) Normalized F-Luc-5BoxB-A_{95}-C_7-HhR reporter mRNA levels corresponding to the experiment described in Figure 2F and G.

Supplemental Figure S3



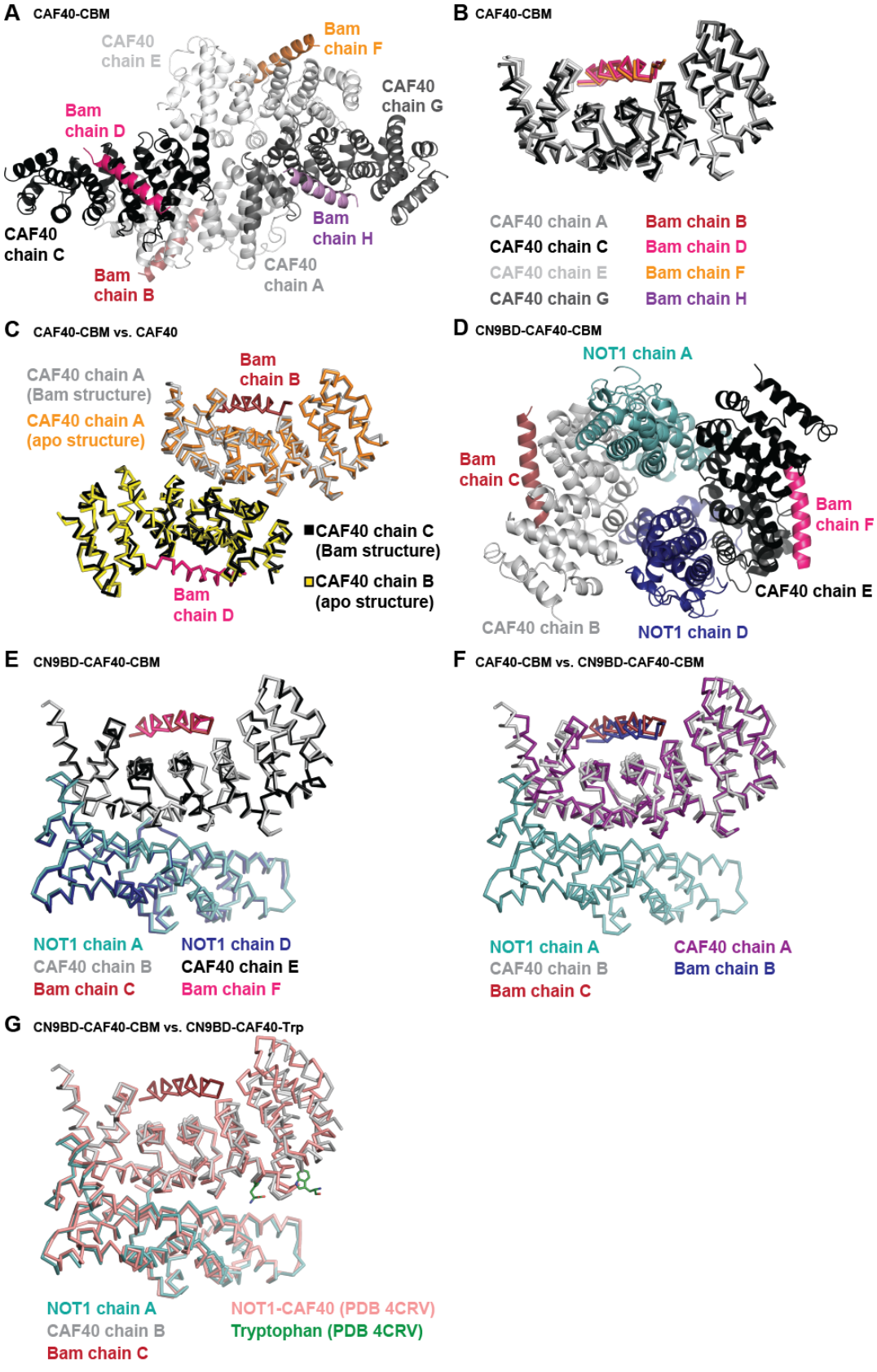
Supplemental Figure S3. Bam interacts with the CCR4-NOT complex. (*A*-*E*) Coimmunoprecipitation assays showing the interaction of HA-tagged Bam with the indicated GFP-tagged CCR4-NOT subunits in S2 cell lysates treated with RNase A. In all panels, GFP-F-Luc served as a negative control. Inputs (1% for the HA-tagged proteins and 3% for the GFP-tagged proteins) and immunoprecipitates (30% for the HA-tagged proteins and 10% for the GFP-tagged proteins) were analyzed by western blotting. Protein size markers are shown on the right in each panel. (*F*) Schematic representation of the *Hs* CCR4-NOT complex. NOT1 contains two HEAT repeat domains (shown in blue and petrol), a MIF4G domain composed of HEAT repeats (green), a three-helix bundle domain (CN9BD, yellow), a connector domain (CD, light blue) and a NOT1 superfamily homology domain (SHD, gray), which also consists of HEAT repeats. The additional subunits of the complex are shown at their binding positions on NOT1. (*G*) *In vitro* MBP pull-down assay testing the interaction of MBP-tagged full-length Bam with the indicated *Hs* CCR4-NOT subcomplexes. MBP served as a negative control.

Supplemental Figure S4



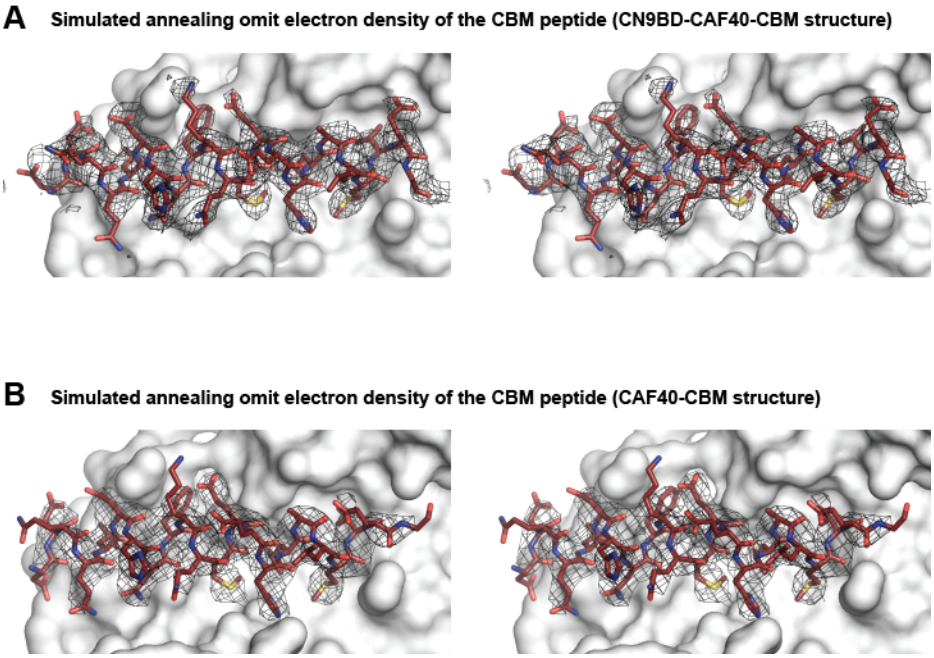
Supplemental Figure S4. Bam requires binding to the mRNA target to induce degradation. (*A,B*) Tethering assay using the F-Luc reporter lacking BoxB sites and λN-HA-tagged Bam (full-length or the indicated fragments) in S2 cells. The samples were analyzed as described in Figure 1B–D. The corresponding experiment with the F-Luc-5BoxB reporter is shown in Figure 3E and F.

Supplemental Figure S5



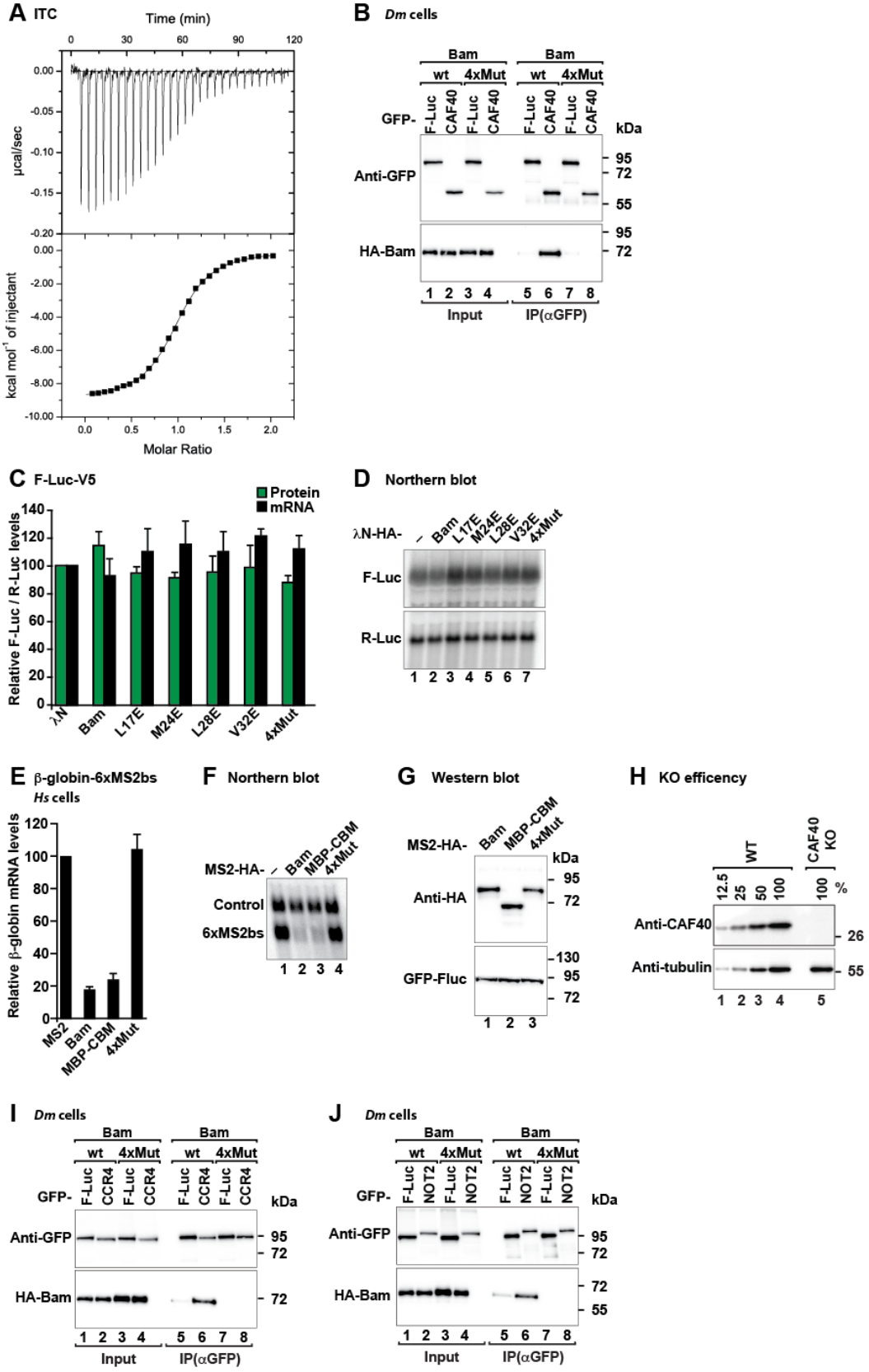
Supplemental Figure S5. Crystal structure of the Bam CBM bound to CAF40 and the CN9BD–CAF40 module. (*A*) Crystal packing of the CAF40–Bam CBM complex. The four copies of CAF40 (chains A, C, E and G) are shown in different shades of gray, the four Bam CBM peptides (chains B, D, F and H) in different colors. (*B*) Superposition of the four CAF40-Bam CBM complexes in the asymmetric unit in ribbon representation. Colors are as in (A). (*C*) Superposition of a CAF40 homodimer (orange and yellow, PDB 2FV2; Garces et al. 2007), with a CAF40 homodimer bound to the Bam CBM (chains A–D, colors are as in (A)). (*D*) Crystal packing of the NOT1 CN9BD-CAF40-Bam CBM complex in cartoon representation. The NOT1 CN9BD is shown in cyan and blue (chains A and D, respectively), CAF40 in gray and black (chains B and E), and Bam in red and pink (chains C and F). (*E*) Superposition of the two NOT1 CN9BD-CAF40-Bam complexes in the asymmetric unit. Colors are as in (D). (*F*) Superposition of the CAF40-Bam complex and NOT1 CN9BD-CAF40-Bam complex structures. (*G*) Superposition of the NOT1 CN9BD-CAF40-Bam CBM complex with the NOT1 CN9BD-CAF40 complex with bound tryptophan (PDB 4CRV) (Chen et al. 2014).

Supplemental Figure S6



Supplemental Figure S6. Simulated annealing electron density of the Bam CBM peptide. (*A*) Stereo view showing the $2F_O$-F_C simulated annealing composite omit map surrounding the CN9BD-CAF40-bound CBM peptide contoured at 1.0 σ. This map was generated with Phenix.Composite_omit_map (Afonine et al. 2012) using the final refined CN9BD-CAF40-Bam model. (*B*) Stereo view showing the $2F_O$-F_C simulated annealing composite omit map surrounding the CAF40-bound CBM peptide contoured at 1.0 σ. This map was generated with Phenix.Composite_omit_map (Afonine et al. 2012) using the final refined CAF40-Bam model.

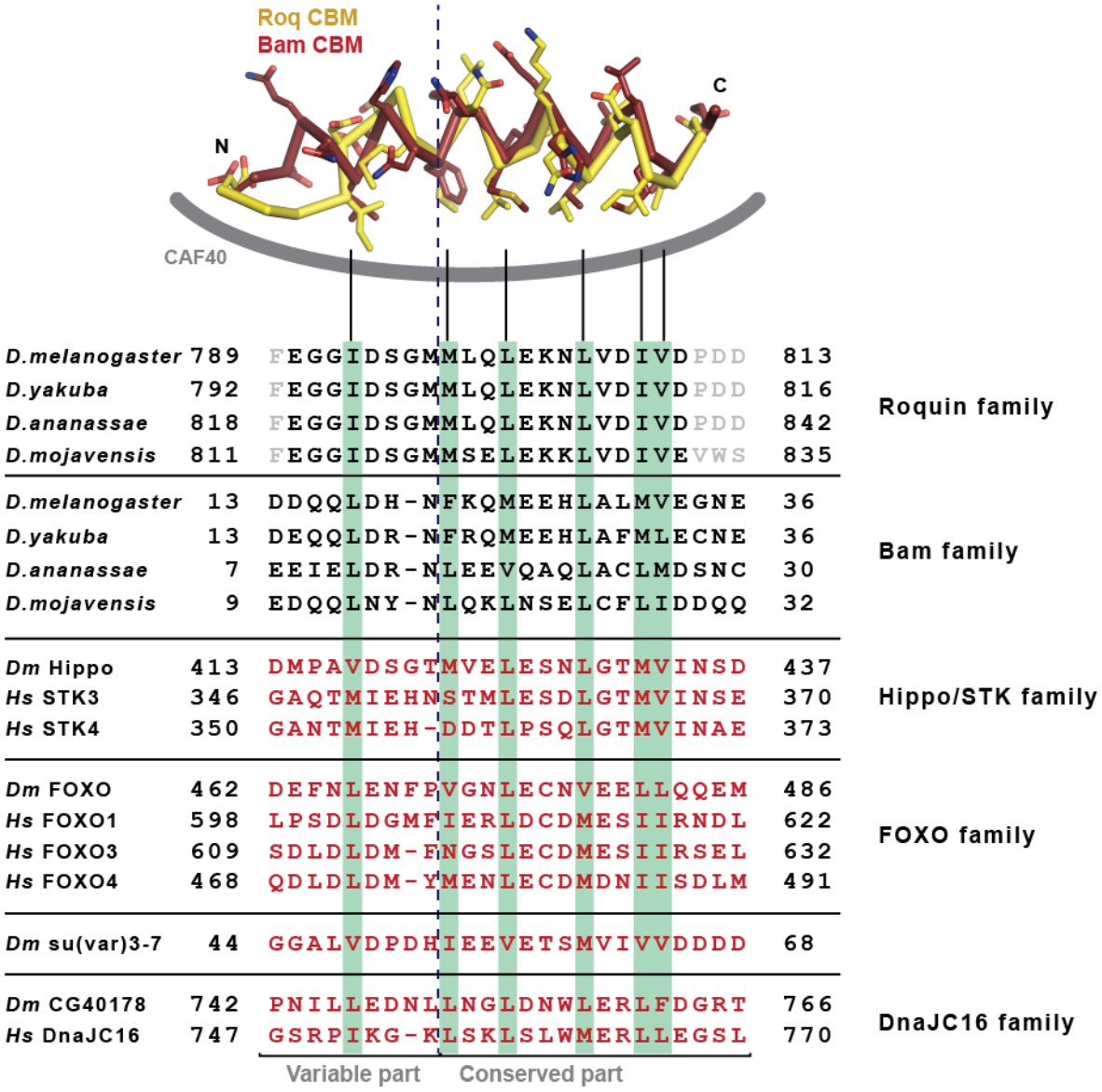
Supplemental Figure S7



Supplemental Figure S7. The CBM is required for Bam activity. (*A*) Representative isothermal titration calorimetry thermogram showing the interaction of the MBP-tagged Bam CBM with the NOT1 CN9BD-CAF40 complex. The upper panel shows raw data, and the lower panel shows the integration of heat changes associated with each injection. Data were fitted using a one-site binding model. (*B*) Coimmunoprecipitation assay showing the interaction of HA-tagged Bam with GFP-tagged CAF40 in S2 cell lysates treated with RNase A. GFP-F-Luc served as a negative control. Inputs (1% for the HA-tagged proteins and 3% for the GFP-tagged proteins) and immunoprecipitates (30% for the HA-tagged proteins and 10% for the GFP-tagged proteins) were analyzed by western blotting. (*C,D*) Tethering assay using the F-Luc reporter lacking BoxB sites and λN-HA-tagged Bam (wild-type or the indicated mutants) in S2 cells. The samples were analyzed as described in Figure 1B–D. The corresponding experiment with the F-Luc-5BoxB reporter is shown in Figure 6A and B. (*E*) Tethering assays using the β-globin-6xMS2bs reporter and MS2-HA-tagged Bam (full-length, MBP-Bam CBM or the 4xMut) in human HEK293T cells. A plasmid expressing a β-globin mRNA reporter lacking MS2-binding sites (Control) served as a transfection control. The β-globin-6xMS2bs mRNA level was normalized to that of the control mRNA and set to 100% in cells expressing MS2-HA. The mean values ± s.d. from three independent experiments are shown in (E). (*F*) Representative northern blot of samples shown in (E). (**G**) Western blot showing the equivalent expression of the MS2-HA-tagged Bam constructs used in (E) and (F). (*H*) Western blot showing the efficiency of the CAF40 depletion in HEK293T cells corresponding to the experiment shown in Figure 6D and E. Dilutions of control cell lysates were loaded in lanes (1–4) to estimate the efficacy of the depletion. Tubulin served as a loading control. KO: knockout. Protein size markers are shown on the right in each panel. (*I,J*) Coimmunoprecipitation assays showing the interaction of HA-tagged Bam with GFP-

tagged CCR4 (I) and NOT2 (J) in S2 cell lysates treated with RNaseA. GFP-F-Luc served as a negative control. Inputs (1% for the HA-tagged proteins and 3% for the GFP-tagged proteins) and immunoprecipitates (30% for the HA-tagged proteins and 10% for the GFP-tagged proteins) were analyzed by western blotting. Protein size markers (kDa) are shown on the right in each panel.

Supplemental Figure S8



Supplemental Figure S8. Profile-based sequence alignment. Profile-based sequence alignment of CBMs from Roquin and Bam of the indicated *Drosophila* species, as well as putative CBMs of proteins from *Hs* and *Dm* shown in red. Residues known or expected to interact with CAF40 are highlighted by a light green background. Gray letters indicate residues that were not included in the crystallization setup. Numbers on both sides of the alignment indicate the residue numbers of the respective fragment boundaries.

SUPPLEMENTAL MATERIAL AND METHODS

DNA constructs

The plasmids used for the expression of subunits of the human and *Dm* CCR4-NOT complex and *Dm* Roq in cells have been previously described (Brau et al. 2011; Bawankar et al. 2013; Sgromo et al. 2017). The plasmids for the expression of *Hs* NOT2-C, NOT3-C, CAF40 ARM domain and the NOT1 MIF4G, CN9BD, CD and SHD domains in *Escherichia coli* have been previously described (Petit et al. 2012; Boland et al. 2013; Chen et al. 2014a; Sgromo et al. 2017). The plasmids for expression of the β-globin-6xMS2bs and the control β-globin-GAP mRNA in human cells were kindly provided by Dr. Lykke-Andersen and have been previously described (Lykke-Andersen et al. 2000). The plasmids for tethering assays in S2 cells (F-Luc-5BoxB, F-Luc-V5, F-Luc-5BoxB-$A_{95}C_7$-HhR, and R-Luc) have been previously described (Behm-Ansmant et al. 2006; Zekri et al. 2013).

For expression of Bam (full-length and fragments) in *Dm* S2 cells, the corresponding cDNA was amplified from total *Dm* oocyte cDNA and cloned between the XhoI and ApaI restriction sites of the pAc5.1-λN-HA and pAc5.1-GFP vectors (Rehwinkel et al. 2005; Tritschler et al. 2007). For expression in HEK293T cells, the cDNA encoding Bam was inserted between the BglII and BamHI restriction sites of the pT7-V5-SBP-C1 and pT7-MS2-HA vectors (Jonas et al. 2013). The plasmids for expression of Bam (full-length, Bam CBM and Bam fragments) in *Escherichia coli* were obtained by inserting the corresponding Bam cDNA fragments between the XhoI and AvrII restriction of the pnYC-vM plasmid (Diebold et al. 2011), thus yielding fusion proteins carrying N-terminal MBP tags cleavable by the TEV protease. For expression of the *Hs* NOT1-10-11 complex, two plasmids were generated. A cDNA fragment encoding the *Hs* NOT1 N-terminus (residues M1–D1000) was inserted into the AvrII restriction site of the pnYC vector,

which does not encode a solubility tag. cDNA fragments encoding *Hs* NOT10 (residues D25–Q707) and *Hs* NOT11 (residues D257–D498) were cloned in a bicistronic plasmid based on the pnEA backbone, thus resulting in the expression of untagged NOT10 and NOT11 with a C-terminal, TEV-cleavable His_6 tag. For expression of the human catalytic module, the His_6-tagged human NOT1 MIF4G domain (residues E1093-S1317) was coexpressed with a bicistronic plasmid expressing untagged CAF1 and CCR4a with an N-terminal MBP-tag cleavable by the HRV3C protease. *Hs* NOT1-CD cDNA was cloned in the pnYC-pM plasmid (Diebold et al. 2011), thereby generating a fusion protein containing an N-terminal MBP tag that is cleavable by the HRV3C protease. The cDNA encoding the NOT3-N fragment (residues A2–D212) was inserted between the XhoI and BamH1 restriction sites of the pnEA-pM vector, thus resulting in an N-terminally MBP-tagged protein.

Coimmunoprecipitation and SBP-pull-down assays

For coimmunoprecipitation assays in S2 cells, 2.5×10^6 cells were seeded per well in 6-well plates and transfected using Effectene transfection reagent (Qiagen). The transfection mixtures contained plasmids expressing GFP-tagged CCR4-NOT of subunits (2 μg) or HA-tagged Bam (1 μg). Cells were harvested 3 days after transfection, and coimmunoprecipitation assays were performed using RIPA buffer [20 mM HEPES (pH 7.6), 150 mM NaCl, 2.5 mM $MgCl_2$, 1% NP-40, 1% sodium deoxycholate supplemented with protease inhibitors (Complete protease inhibitor mix, Roche)] as previously described (Tritschler et al. 2008). For SBP pull-down assays in human cells, HEK293T cells (ATCC, wild-type or CAF40-null cells) were grown in 10-cm dishes (4×10^6 / 10-cm dish) and transfected using TurboFect transfection reagent (Thermo Fisher Scientific). The transfection mixtures contained 20 μg, 5 μg and 25 μg of plasmids expressing

Bam, MBP-CBM and Bam 4xMut, respectively. For the pull-down assays in Figure 6E and G, cells were also co-transfected with 8 μg of a plasmid expressing HA-tagged CCR4. The cells were harvested 2 days after transfection, and pull-down assays were performed as previously described (Bhandari et al. 2014).

Generation of the CAF40-null cell line

An sgRNA (sequence: 5' CCCATGCTGTGGCATTCATT 3') targeting the second exon of the *Hs* CAF40 gene was designed using CHOPCHOP (http://chopchop.cbu.uib.no) and inserted 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 pSp-CAF40-sgRNA-Cas9(BB)-2A-Puro plasmid and selected with puromycin (3 μg/ml) to obtain stable CAF40 knockout cells. To obtain clonal cell lines, single cells were distributed in 96-well plates using serial dilutions. Expansion of single-cell clones was performed under non-selective conditions. CAF40-null clones were identified by western blotting using anti-CAF40 antibodies (Supplemental Table S2). Genomic DNA from single clones was isolated using a Wizard SV Genomic DNA Purification System (Promega) and the targeted CAF40 locus was amplified by PCR and sequenced to confirm gene editing. We observed a deletion of 22nt in one allele and an insertion of one nucleotide in the second exon of CAF40 in the other allele, both of which cause a frameshift.

Protein expression, purification and competition assays

To purify the *Dm* NOT1 CN9BD-CAF40 complex, MBP-tagged NOT1-CN9BD (residues Y1468-T1719) was co-expressed with His_6-tagged CAF40 (ARM domain, residues E25–G291). The cells were lysed in a buffer containing 50 mM HEPES (pH 7.5), 300 mM NaCl, 20 mM

imidazole and 2 mM β-mercaptoethanol. The complex was purified from cleared cell lysates by Nickel affinity chromatography using a HiTrap IMAC column and eluted by a linear gradient to 500 mM imidazole. The complex was further purified on a HiTrapQ column (GE Healthcare), and this was followed by removal of the His_6 and MBP tags by cleavage with HRV3C protease overnight at 4°C. The complex was separated from the tags by size exclusion chromatography using a Superdex 200 26/600 column in a buffer containing 10 mM HEPES (pH 7.5), 200 mM NaCl and 2 mM DTT.

For competition assays, the CAF40 ARM domain was expressed with an N-terminal GST tag. The cells were lysed in a buffer containing 50 mM HEPES (pH 7.5), 300 mM NaCl and 2 mM DTT. The protein was purified from cleared cell lysates by using Protino glutathione agarose 4B (Macherey-Nagel) followed by a HiTrapQ column and further purified by size exclusion chromatography on a Superdex 200 26/600 column in a buffer containing 10 mM HEPES (pH 7.5), 200 mM NaCl and 2 mM DTT. The Roquin CBM fused to MBP was purified as previously described (Sgromo et al. 2017). Cells expressing either His_6-NusA-tagged Bam CBM or His_6-NusA were lysed in a buffer containing 50 mM potassium phosphate (pH 7.5), 300 mM NaCl, 20 mM imidazole and 2 mM β-mercaptoethanol. The proteins were isolated from the crude cell lysate by Nickel affinity chromatography using a HiTrap IMAC column and eluted by a linear gradient to 500 mM imidazole. The eluted proteins were directly applied to size exclusion chromatography on a Superdex 200 16/600 column in a buffer containing 10 mM HEPES (pH 7.5), 200 mM NaCl and 2 mM DTT.

The assembled *Hs* NOT1-10-11 trimer was obtained by co-expression of C-terminally His_6-tagged NOT11 (residues D257–D498) and untagged NOT1 (residues M1–D1000) and NOT10 (residues D25–Q707). The cells were lysed in 50 mM potassium phosphate (pH 7.6), 300 mM

NaCl, 20 mM imidazole and 2 mM β-mercaptoethanol. The complex was purified from cleared cell lysates by using a HiTrap IMAC column and eluted by a linear gradient to 500 mM imidazole. The complex was dialyzed in a buffer containing 50 mM HEPES (pH 7.5), 200 mM NaCl and 2 mM DTT, and was further purified over a HiTrap Heparin column (GE Healthcare), then subjected to size exclusion chromatography over a Superdex 200 26/600 column in a buffer containing 10 mM HEPES (pH 7.5), 200 mM NaCl and 2 mM DTT.

To purify the assembled catalytic module, His_6-tagged NOT1 MIF4G domain (residues E1093–S1317), untagged CAF1 and MBP-tagged CCR4a were co-expressed. Cells were lysed in a buffer containing 50 mM potassium phosphate (pH 7.5), 300 mM NaCl and 2 mM β-mercaptoethanol. The complex was purified from cleared cell lysates using amylose resin and eluted with in a buffer containing 50 mM potassium phosphate (pH 7.5), 300 mM NaCl, 20 mM imidazole, 25 mM D(+)-maltose and 2 mM β-mercaptoethanol. The complex was further purified using a HiTrap IMAC column (GE Healthcare) and eluted by a linear gradient to 500 mM imidazole. The His_6 and MBP tags were removed by cleavage with the HRV3C protease overnight at 4°C. The catalytic module was further purified over a Superdex200 (26/600 column; GE Healthcare) in a buffer containing 10 mM HEPES (pH 7.5), 200 mM NaCl and 2 mM DTT.

The *Hs* NOT3 N-terminus (residues A2–D212) was expressed with an N-terminal MBP tag. Cells were lysed in a buffer containing 50 mM HEPES (pH 7.5), 300 mM NaCl and 2 mM DTT. The protein was purified from cleared cell lysates with amylose resin, then with a HiTrapQ column. The MBP tag was removed by cleavage using the HRV3C protease. After cleavage of the tag, the protein was further purified on a Superdex 75 26/600 column (GE Healthcare) using a buffer containing 10 mM HEPES (pH 7.5), 200 mM NaCl and 2 mM DTT.

The purification procedures for the human NOT1-CD (residues D1607–S1815) and of the NOT and CAF40 modules have been previously described (Chen et al. 2014; Raisch et al. 2016; Sgromo et al. 2017). The NOT module comprises the NOT1-SHD (residues H1833–M2361), NOT2-C (residues M350–F540) and NOT3-C (residues L607–E748). The *Hs* CAF40 module comprises NOT1-CN9BD (residues V1351–L1588) and the CAF40 ARM domain (residues R19-E285). The *Dm* Bam CBM peptide (residues D13–E36) used for crystallization was obtained from EMC microcollections and solubilized in a buffer containing 10 mM HEPES (pH 7.5), 200 mM NaCl and 2 mM DTT.

Crystallization

Crystals of *Hs* CAF40 (ARM domain) bound to Bam CBM peptide (residues D13–E36) were obtained at 22°C using the hanging-drop vapor diffusion method after the protein solution (6 mg/ml CAF40 and 1.1 mg/ml Bam CBM peptide; 200 nl) was mixed with the crystallization reservoir solution (200 nl). Crystals appeared within one day in many conditions. Optimized crystals grew at 18°C in hanging drops consisting of 1 µl protein solution (6 mg/ml CAF40 and 1.1 mg/ml Bam CBM peptide) and 1 µl crystallization reservoir solution containing 100 mM HEPES (pH 7.0), 200 mM $CaCl_2$ and 15% PEG 6,000. Crystals were soaked in reservoir solution supplemented with 15% ethylene glycol for cryoprotection before being flash-frozen in liquid nitrogen.

Crystals of the *Hs* NOT1 CN9BD–CAF40 complex bound to the Bam CBM peptide were obtained at 22°C by using the hanging-drop vapor diffusion method after mixing the protein solution (7.5 mg/ml NOT1 CN9BD–CAF40 and 0.8 mg/ml CBM peptide; 200 nl) with the crystallization reservoir solution (200 nl). Crystals appeared within one day in several conditions.

Optimized crystals grew in drops of 200 nl protein solution (5 mg/ml NOT1 CN9BD–CAF40 complex and 0.5 mg/ml CBM peptide) mixed with 200 nl crystallization reservoir solution comprising 1.5 M ammonium sulfate, 20 mM MES (pH 6.0) and 80 mM MES (pH 6.5). Crystals were soaked in reservoir solution supplemented with 25% glycerol for cryoprotection before flash-freezing in liquid nitrogen.

Data collection and structure determination

X-ray diffraction data for the *Hs* NOT1 CN9BD–CAF40 bound to the Bam CBM were collected at a wavelength of 1.0000 Å on a PILATUS 6M detector (Dectris) at the PXII beamline of the Swiss Light Source (SLS) and processed in space group $P3_221$ by using XDS and XSCALE (Kabsch 2010) to a resolution of 2.7 Å, aiming at a CC(1/2) value (Karplus and Diederichs 2012) of ~70 % as a high resolution cutoff. Initial phases were determined by molecular replacement, with two copies of the NOT1 CN9BD–CAF40 complex (PDB 4CRU) used as a search model in PHASER (McCoy et al. 2007) from the CCP4 package (Winn et al. 2011). The initial model was improved and completed by iterative cycles of building in COOT (Emsley et al. 2010) and refinement in PHENIX (Afonine et al. 2012), also optimizing TLS parameters (one TLS group per macromolecular chain). Finally, two copies of the Bam CBM peptide were manually built into the density (Supplemental Fig. S6A) and improved by further refinement cycles.

The best crystal of the CAF40 (ARM domain) bound to the Bam CBM peptide was recorded at a wavelength of 1.0396 Å on a PILATUS 6M fast detector (DECTRIS) at the DESY beamline P11. The dataset was processed in XDS and XSCALE in space group $P2_12_12$ to a resolution of 3.0 Å, aiming at a CC(1/2) value (Karplus and Diederichs 2012) of ~70 % as a high resolution cutoff. Four copies of the CAF40 ARM domain (PDB 2FV2, chain A) were found in the

asymmetric unit by molecular replacement using PHASER from the CCP4 package. This initial model was improved and completed by iterative cycles of building in COOT (Emsley et al. 2010) and refinement using PHENIX (Afonine et al. 2012) and BUSTER (Bricogne et al. 2011) using NCS restraints and TLS parameters (one TLS group per macromolecular chain). Finally, four copies of the Bam CBM peptide were manually built into the density (Supplemental Fig. S6B) and improved through further refinement cycles.

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

Isothermal titration calorimetry (ITC)

The ITC experiments were performed on a VP-ITC microcalorimeter (Microcal) at 20°C. A solution containing the *Dm* NOT1 CN9BD bound to CAF40 (ARM domain) (6.0 μM in the experiments with the Bam CBM and up to 10 μM in the experiments with the Roq CBM) in a calorimetric cell was titrated with a solution of MBP-tagged Bam CBM (60 μM) or MBP-tagged Roquin CBM (up to 100 μM). All proteins were dissolved in a buffer containing 10 mM HEPES (pH 7.5), 200 mM NaCl and 0.5 mM TCEP. The titration experiments consisted of an initial injection of 2 μl followed by 28 injections of 10 μl at 240 s intervals. The binding experiment was repeated three times. The thermodynamic parameters were calculated using a one-site binding model (ORIGIN version 7.0; Microcal). Correction for dilution heating and mixing was achieved by subtracting the final baseline, which consisted of small peaks of similar size. The first injection was removed from the analysis (Mizoue and Tellinghuisen 2004).

Bioinformatic analysis

To identify proteins featuring potential CBMs in *Dm* and *Hs*, we followed a two-step approach. In the first step, we searched for homologs of *Dm* Bam and Roquin in the nonredundant (nr) protein sequence database using PSI-BLAST (Boratyn et al. 2013), as implemented in the MPI Bioinformatics toolkit (Alva et al. 2016), and extracted the CBMs from the obtained homologs originating from different *Drosophila* species. These motifs were then aligned, and a consensus pattern was derived by manual inspection (x-x-x-[LI]-[DENQ]-x(2,3)-[FLM]-x-x-[ILM]-x-x-x-[IL]-x-x-[ILM]-[LIV]-x-x-x-x). In the second step, the aforementioned consensus pattern was submitted to the PatternSearch tool of the MPI Bioinformatics Toolkit to identify proteins in *Dm* and *Hs* with potential CBMs. This search yielded a total of 1,200 candidate proteins. We next analyzed this set further to discard all proteins in which the detected motifs showed no helical propensity or were embedded within a domain (as opposed to being embedded in an intrinsically disordered region). We also excluded all proteins with obvious functional irrelevance (e.g. membrane proteins) from further consideration. Finally, we chose the *Hs* and *Dm* homologs of four protein families, on the basis of the presence of known or predicted RNA-binding domains in the proteins and on the percentage similarity of the putative CBMs to the CBMs of Bam and Roquin. These candidate CBMs were then expressed as MBP fusions and tested for CAF40 binding in *in vitro* MBP pull-down assays.

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.

Alva V, Nam SZ, Soding J, Lupas AN. 2016. The MPI bioinformatics Toolkit as an integrative platform for advanced protein sequence and structure analysis. *Nucleic Acids Res* **44:** W410-415.

Bawankar P, Loh B, Wohlbold L, Schmidt S, Izaurralde E. 2013. NOT10 and C2orf29/NOT11 form a conserved module of the CCR4-NOT complex that docks onto the NOT1 N-terminal domain. *RNA Biol* **10:** 228-244.

Behm-Ansmant I, Rehwinkel J, Doerks T, Stark A, Bork P, Izaurralde E. 2006. mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. *Genes Dev* **20:** 1885-1898.

Bhandari D, Raisch T, Weichenrieder O, Jonas S, Izaurralde E. 2014. Structural basis for the Nanos-mediated recruitment of the CCR4-NOT complex and translational repression. *Genes Dev* **28:** 888-901.

Boland A, Chen Y, Raisch T, Jonas S, Kuzuoğlu-Öztürk D, Wohlbold L, Weichenrieder O, Izaurralde E. 2013. Structure and assembly of the NOT module of the human CCR4-NOT complex. *Nat Struct Mol Biol* **20:** 1289-1297.

Boratyn GM, Camacho C, Cooper PS, Coulouris G, Fong A, Ma N, Madden TL, Matten WT, McGinnis SD, Merezhuk Y. et al. 2013. BLAST: a more efficient report with usability improvements. *Nucleic Acids Res* **41:** W29-33.

Braun JE, Huntzinger E, Fauser M, Izaurralde E. 2011. GW182 proteins directly recruit cytoplasmic deadenylase complexes to miRNA targets. *Mol Cell* **44:** 120-33.

Bricogne G, Blanc E, Brandl M, Flensburg C, Keller P, Paciorek W, Roversi P, Sharff A, Smart O, Vonrhein C. et al. 2011. BUSTER. *Cambridge, United Kingdom: Global Phasing Ltd.*

Chen VB, Arendall WB3rd, 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.

Chen Y, Boland A, Kuzuoğlu-Öztürk D, Bawankar P, Loh B, Chang CT, Weichenrieder O, Izaurralde E. 2014a. A DDX6-CNOT1 complex and W-binding pockets in CNOT9 reveal direct links between miRNA target recognition and silencing. *Mol Cell* **54:** 737-750.

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.

Garces RG, Gillon W, Pai EF. 2007. Atomic model of human Rcd-1 reveals an armadillo-like-repeat protein with in vitro nucleic acid binding properties. *Protein Sci* **16:** 176-188.

Jonas S, Weichenrieder O, Izaurralde E. 2013. An unusual arrangement of two 14-3-3-like domains in the SMG5-SMG7 heterodimer is required for efficient nonsense-mediated mRNA decay. *Genes Dev* **27:** 211-225.

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

Karplus PA, Diederichs K. 2012. Linking crystallographic model and data quality. *Science* **336:** 1030-1033.

Lykke-Andersen J, Shu MD, Steitz JA. 2000. Human Upf proteins target an mRNA for nonsense-mediated decay when bound downstream of a termination codon. *Cell* **103:** 1121-1131.

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.

Petit AP, Wohlbold L, Bawankar P, Huntzinger E, Schmidt S, Izaurralde E, Weichenrieder O. 2012. The structural basis for the interaction between the CAF1 nuclease and the NOT1 scaffold of the human CCR4-NOT deadenylase complex. *Nucleic Acids Res* **40:** 11058-11072.

Raisch T, Bhandari D, Sabath K, Helms S, Valkov E, Weichenrieder O, Izaurralde E. 2016. Distinct modes of recruitment of the CCR4-NOT complex by Drosophila and vertebrate Nanos. *EMBO J* **35:** 974-990.

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.

Rehwinkel J, Behm-Ansmant I, Gatfield D, Izaurralde E. 2005. A crucial role for GW182 and the DCP1:DCP2 decapping complex in miRNA-mediated gene silencing. *RNA* **11:** 1640-1647.

Sgromo A, Raisch T, Bawankar P, Bhandari D, Chen Y, Kuzuoğlu-Öztürk D, Weichenrieder O, Izaurralde E. 2017. A CAF40-binding motif facilitates recruitment of the CCR4-NOT complex to mRNAs targeted by Drosophila Roquin. *Nat Commun* **8:** 14307.

Tritschler F, Eulalio A, Truffault V, Hartmann MD, Helms S, Schmidt S, Cole ,M, Izaurralde E, Weichenrieder O. 2007. A divergent Sm fold in EDC3 proteins mediates DCP1 binding and P-body targeting. *Mol Cell Biol* **27:** 8600-8611.

Tritschler F, Eulalio A, Helms S, Schmidt S, Coles M, Weichenrieder O, Izaurralde E, Truffault V. 2008. Similar modes of interaction enable Trailer Hitch and EDC3 to associate with DCP1 and Me31B in distinct protein complexes. *Mol Cell Biol* **28:** 6695-6708.

Valdar WS. 2002. Scoring residue conservation. *Proteins* **48:** 227-241.

Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, McCoy A. et al. 2011. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67:** 235-242.

Zekri L, Kuzuoğlu-Öztürk D, Izaurralde E. 2013. GW182 proteins cause PABP dissociation from silenced miRNA targets in the absence of deadenylation. *EMBO J* **32:** 1052-1065.