Structure, Volume *22*

Supplemental Information

Translation Initiation Factor eIF3b Contains a Nine-Bladed β-Propeller and Interacts with the 40S Ribosomal Subunit

Yi Liu, Piotr Neumann, Bernhard Kuhle, Thomas Monecke, Stephanie Schell, Ashwin Chari, and Ralf Ficner

Supplemental Information

Translation initiation factor eIF3b contains a nine-bladed β-propeller and interacts with the 40S ribosomal subunit

Yi Liu, Piotr Neumann, Bernhard Kuhle, Thomas Monecke, Stephanie Schell, Ashwin Chari, Ralf Ficner

Inventory of Supplemental Information

Supplemental Data

Figure S1, related to Figure 1, showing the SDS-PAGE of full-length *ct*eIF3b, *ct*eIF3b WD40 domain and dissolved crystals, reflecting that the crystallized fragment only contains the WD40 domain of eIF3b.

Figure S2, related to Figure 2, showing the multiple sequence alignment of eIF3b orthologs.

Figure S3, related to Figure 2, displaying the electrostatic potential of conserved surface areas of cteIF3b-WD40.

Figure S4, related to Figure 3, showing the interaction studies between the eIF3b,i,g subcomplex and eIF3c using analytical SEC. No eIF3b,i,g,c complex was observed.

Figure S5, related to Figure 4, comparing the best fit of eIF3b-WD40 (blue) and eIF3i (yellow, PDB code: 3ZWL)

Figure S6, related to Figure 4, showing an eIF3b truncation (eIF3b$^{167\text{-}670}$) is able to bind to the 40S, but eIF3b$^{167\text{-}670}$, eIF3b$^{167\text{-}704}$-3i complex, eIF3i, eIF3g cannot interact with rpS9e.

Supplemental Experimental Procedures

Supplemental References

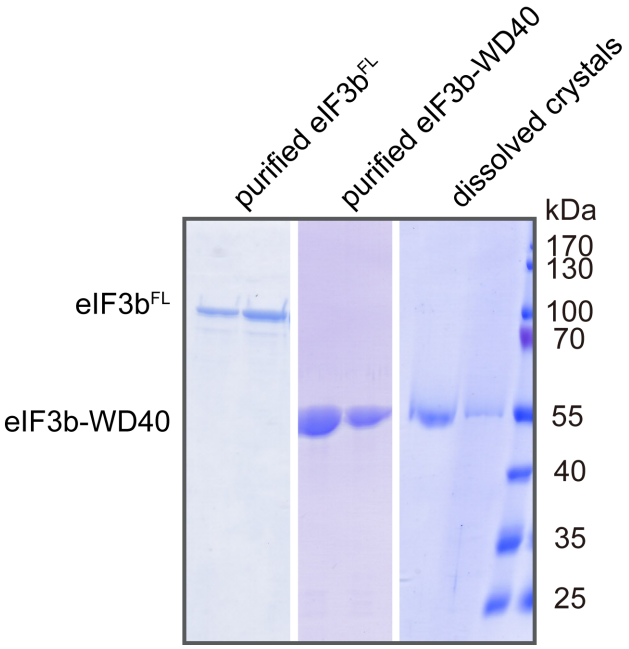

Figure S1, related to Figure 1. SDS-PAGE of full-length *cte*IF3b, *cte*IF3b WD40 domain and dissolved crystals.

The molecular weight of full-length *cte*IF3b and its WD40 domain are ~90 kDa and ~55 kDa, respectively. The size of the crystallized fragment is the same as *cte*IF3b WD40 domain.

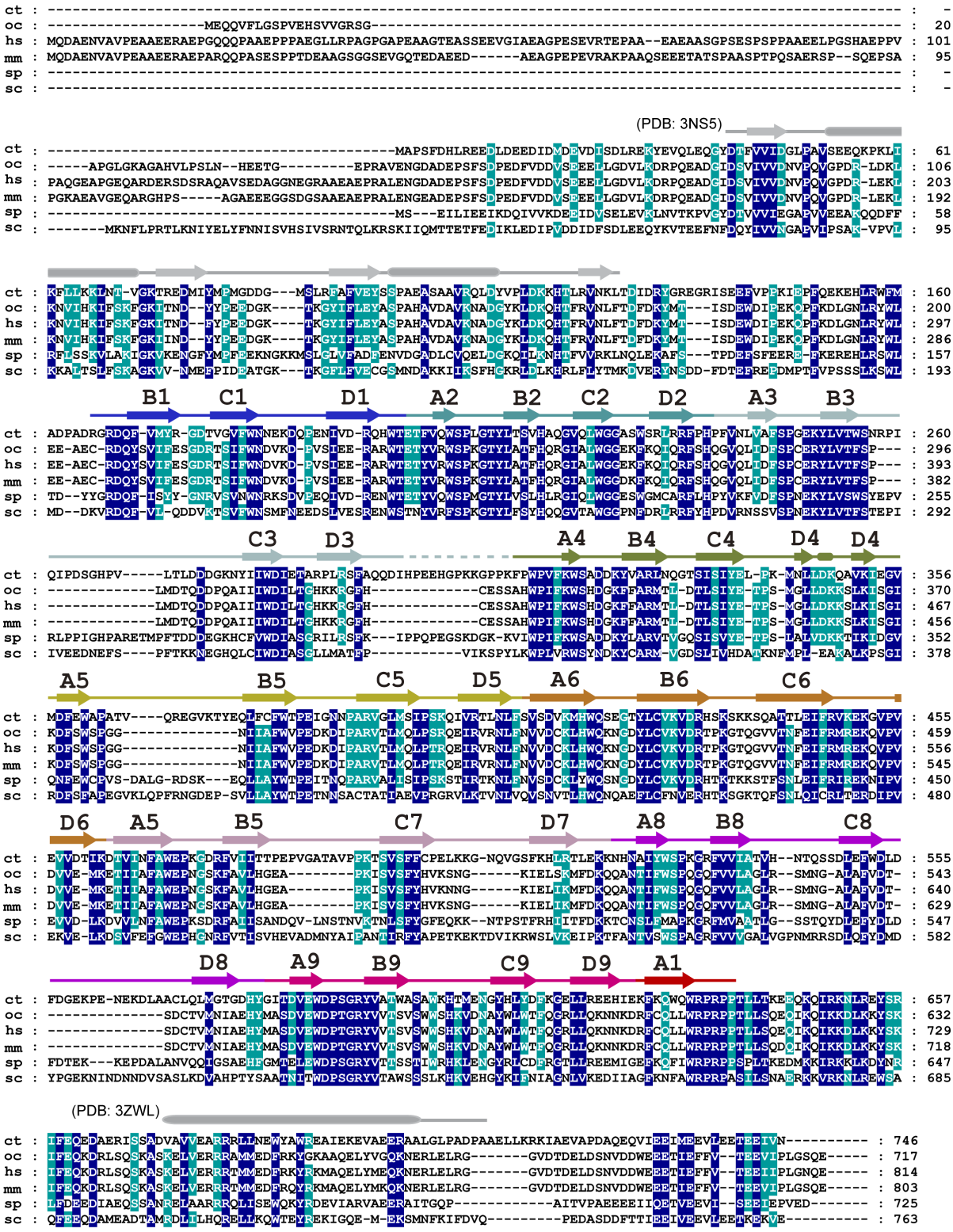

Figure S2, related to Figure 2. Multiple sequence alignment of eIF3b orthologs.

Amino acid sequences of eIF3b from *Chaetomium thermophilum* (ct), *Oryctolagus cuniculus*

(oc), *Homo sapiens* (hs), *Mus musculus* (mm), *Schizosaccharomyces pombe* (sp), and *Saccharomyces cerevisiae* (sc) are aligned using ClustalW2 (www.ebi.ac.uk/Tools/msa/clustalw2/, (Chenna et al., 2003)). Identical residues are highlighted with *blue* background and conserved residues with *cyan* background. Secondary structural elements (β-strands: arrows; α-helices: rounded rectangles; loops: lines) are indicated above the alignment (residues that were not built in the structure are represented as dashed lines). The WD40 domain of eIF3b is colored as in Fig 1B, while the N-terminal RRM domain and the C-terminal α-helix, which are not present in our structure, are colored light gray.

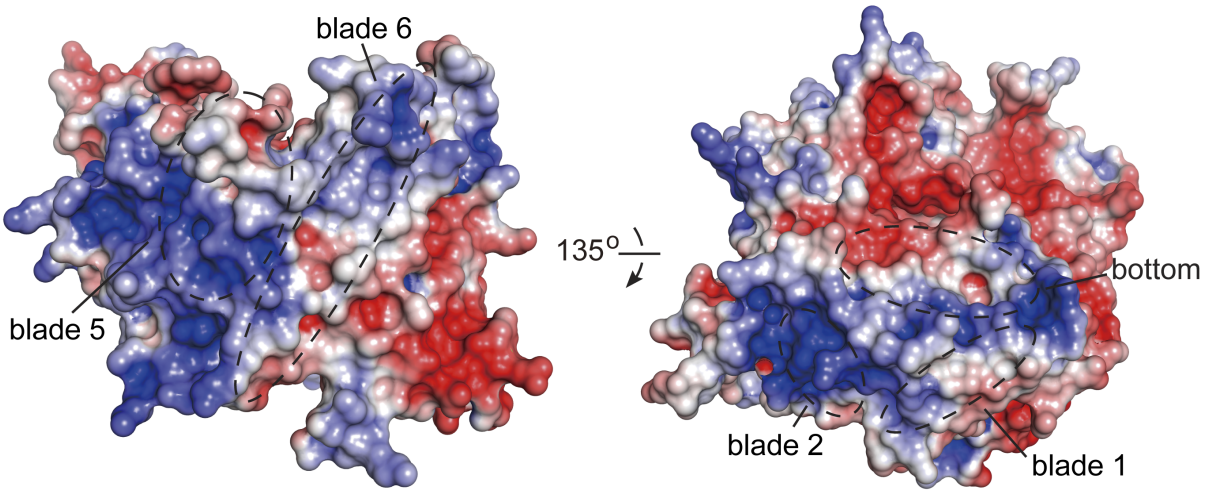

Figure S3, related to Figure 3. Electrostatic potential of conserved surface areas of *cte*IF3b-WD40.

The electrostatic surface potential of *cte*IF3b-WD40 colored from red (-5kT/e) through white (0kT/e) to blue (+5kT/e). Areas containing conserved surface residues are encircled and labeled based on blade numbers or propeller side. The orientations are similar to Fig 2C.

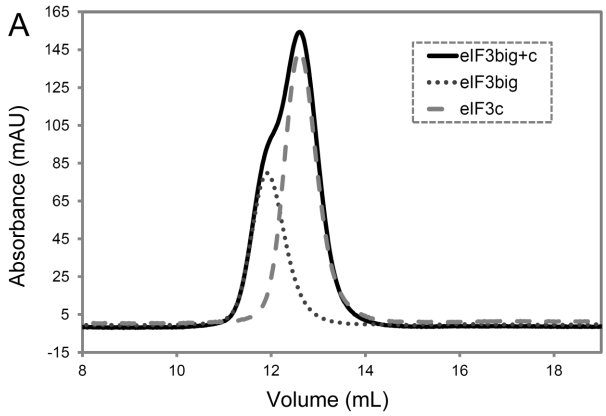

Figure S4, related to Figure 3. Interaction studies between the *ct*eIF3b,i,g subcomplex and *ct*eIF3c using analytical SEC.

Overlay of the analytical SEC profiles of the *ct*eIF3b-i-g subcomplex (dotted *black)*, *ct*eIF3c (dashed *gray*) and the *ct*eIF3b-i-g subcomplex together with subunit c (solid *dark*). In case of *ct*eIF3b,i,g + *ct*eIF3c, two overlapping peaks are observed; the main peak at a higher elution volume corresponds to eIF3c alone, whereas the smaller peak at a lower elution volume (forming the shoulder of the eIF3c peak) corresponds to the *ct*eIF3b,i,g complex. A shift to a lower elution volume does not occur, indicating that the *ct*eIF3b,i,g,c complex is not formed.

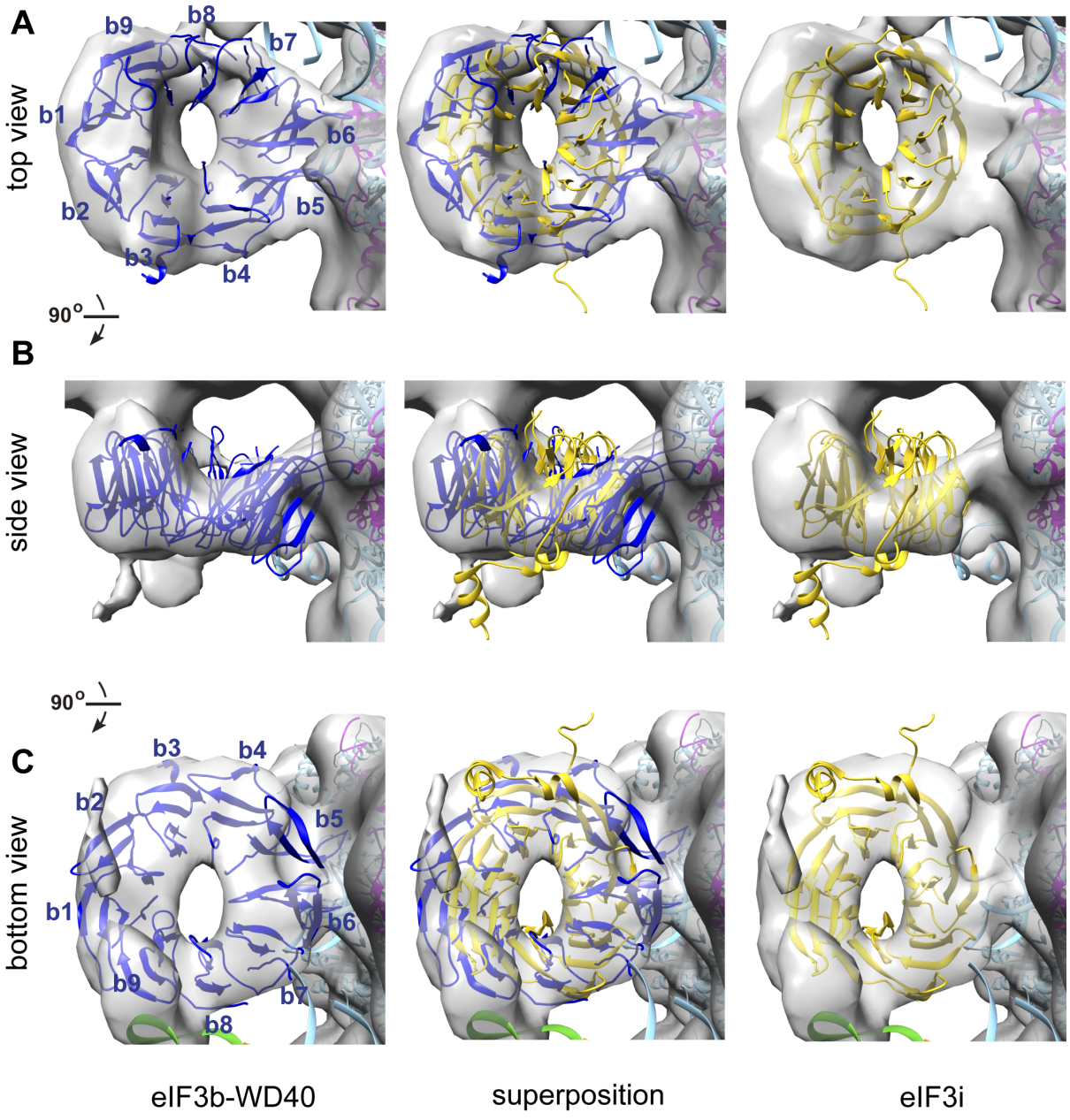

Figure S5, related to Figure 4. Comparison of the best fit of eIF3b-WD40 (blue) and eIF3i (yellow, PDB code: 3ZWL) into the density map of rabbit 43S initiation complex (grey, EMDB ID: 5658). In figure 5A the density assigned to DHX29 located above the doughnut-like density was removed for clarity.

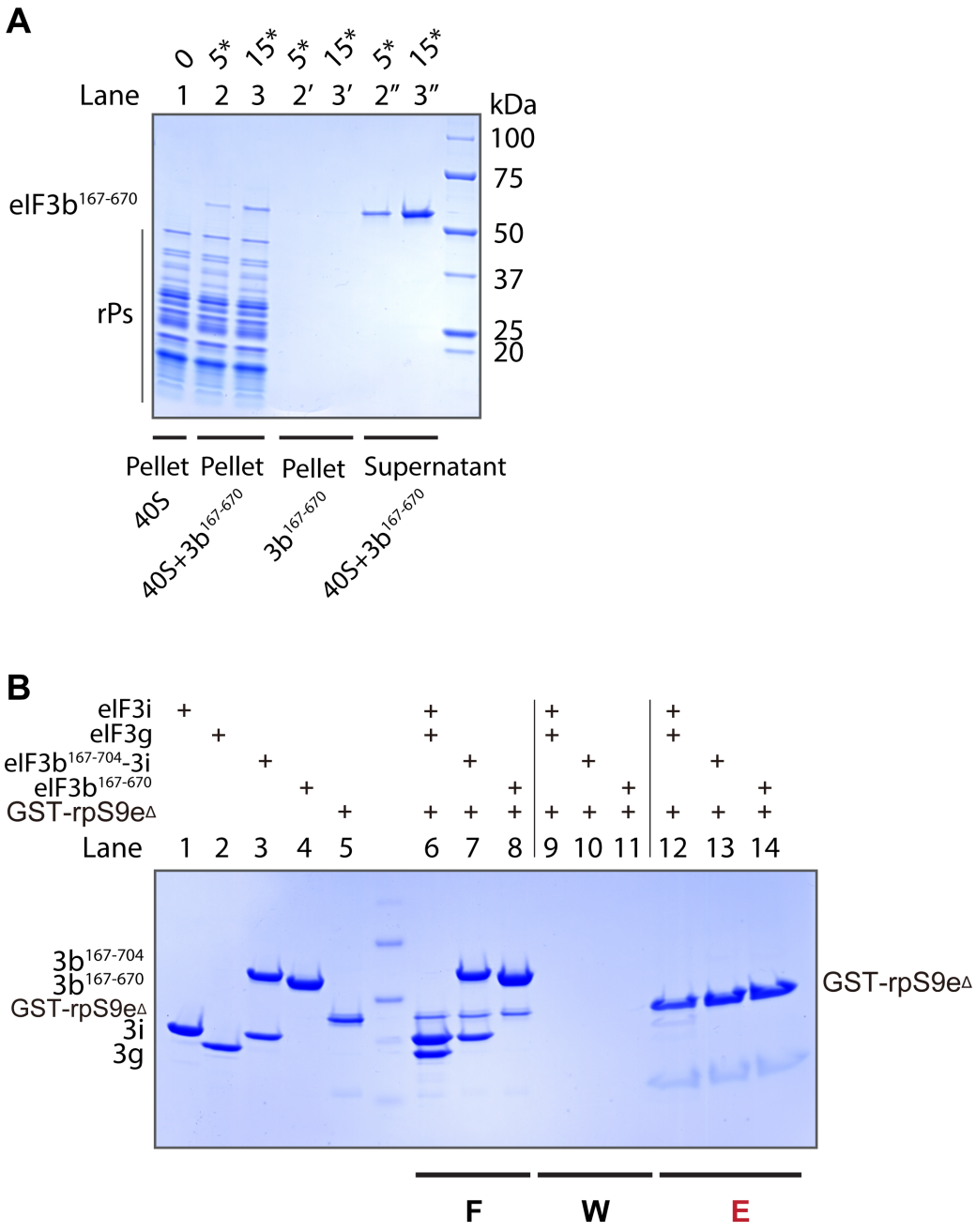

Figure S6, related to Figure 4. Interaction assays between eIF3b truncations and the 40S as well as rpS9e

(**A**) Co-sedimentation of eIF3b$^{167-670}$ and 40S ribosomes. eIF3b$^{167-670}$ co-sediments with 40S (Lanes 2-3) and excess eIF3b$^{167-670}$ remains in the supernatant (Lanes 2"-3"). Lane 1 and Lanes 2'-3' show controls of 40S alone and protein, respectively. (**B**) Interaction study between eIF3b truncations, eIF3i, eIF3g and rpS9e by GST pull-down assays. Lanes 1-5 show

the single proteins as reference. Lanes 12-14 show the elution fractions of the GST pull-down after the removal of unbound proteins (Lanes 6-8) and extensive washing (Lanes 9-11). rpS9e$^{\Delta}$ corresponds to rpS9e$^{1\text{-}183}$. No binding of eIF3b$^{167\text{-}670}$ or eIF3b$^{167\text{-}704}$-eIF3i complex to rpS9e can be detected. A wear binding between eIF3i and eIF3g complex and rpS9e could be detected (lane 12).

Supplemental Experimental Procedures

Cloning, protein expression and purification

Genes encoding eIF3b, a, c, i, and g were amplified by PCR from *Chaetomium thermophilum* genomic DNA, and the *ct*rpS9e gene from cDNA. Each gene was cloned into the pGEX-6P-1 vector (GE Healthcare), with the exception of eIF3a, which was cloned into the pET-28b vector (Merck), following the In-Fusion HD cloning kit user manual (Clontech). The ORFs of eIF3b$^{167\text{-}670}$ and eIF3b$^{167\text{-}704}$ were subcloned into the pGEX-6P-1 vector (GE Healthcare), while eIF3b$^{167\text{-}C}$ into a modified pETDuet-1 vector. The GST fusion proteins (containing a Prescission protease cleavage site), eIF3a with C-terminal His_6-tag and eIF3b$^{167\text{-}C}$ with N-terminal His_6-tag were individually expressed in *Escherichia coli* strain BL21 (DE3) grown in 2×YT medium. Target protein expression was induced by addition of isopropyl β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.25 mM at an OD_{600} of ~1.0. Cells were harvested after incubation overnight at 16 °C by centrifugation (4,800 xg, 20 min, 4 °C; Beckman) and resuspended in lysis buffer (25 mM HEPES/NaOH, pH 7.5, 500 mM NaCl, 5% glycerol and 5 mM β-mercaptoethanol (β-ME)). A microfluidizer (Microfluidics, Newton, US) was used to rupture the bacterial cells and the cell debris was separated via centrifugation (30,000 xg, 35 min, 4 °C). For the GST fusion proteins the supernatant was applied to a GSTrap column (GE Healthcare) equilibrated in lysis buffer. After sample loading and extensive washing with lysis buffer, the column was equilibrated in a low salt buffer (25 mM HEPES/NaOH, pH 7.5, 100 mM NaCl, 5% glycerol and 5 mM β-ME). The GST fusion proteins were eluted with the low salt buffer containing additional 30

mM reduced glutathione. The GST-tag was cleaved by Prescission protease at 4 °C overnight at a 1:100 mass ratio of protease to fusion protein.

For further purification of full-length eIF3b and eIF3g, respectively, the protein was loaded onto an ion exchange column (GE healthcare) (Source 30Q for eIF3b; SP sepharose for eIF3g) equilibrated in buffer A (25 mM HEPES/NaOH, pH 7.5, 100 mM NaCl, 5% glycerol and 5 mM β-ME) and the target protein was eluted applying a linear gradient to buffer B (25 mM HEPES/NaOH, pH 7.5, 1 M NaCl, 5% glycerol and 5 mM β-ME).

For the eIF3a-His_6 protein the supernatant after cell rupture and centrifugation was applied to an IMAC column (His60 Ni superflow, Clontech). After sample loading and extensive washing with lysis buffer, the column was equilibrated in a low salt buffer (25 mM HEPES/NaOH, pH 7.5, 200 mM NaCl, 5% glycerol and 5 mM β-ME). eIF3a-His_6 was eluted with the low salt buffer containing additional 400 mM imidazole. Subsequently eIF3a-His_6 was loaded onto a HiTrap Heparin HP column (GE healthcare) equilibrated in buffer A (20 mM HEPES/NaOH, pH 7.3, 200 mM NaCl, 5% glycerol and 5 mM β-ME) and was eluted with a linear gradient to buffer B (20 mM HEPES/NaOH, pH 7.5, 1 M NaCl, 10% glycerol and 5 mM β-ME).

Final purification of the proteins (eIF3b, eIF3b$^{167\text{-}670}$, eIF3b$^{167\text{-}C}$, c, i, g, rpS9e and eIF3b$^{167\text{-}704}$-eIF3i complex as well as eIF3a-b-c complex) was achieved by a Superdex 200 (26/60) gel filtration column (GE Healthcare) in 10 mM HEPES/NaOH, pH 7.5, 150 mM NaCl, 2% glycerol and 2 mM DTT. Protein-containing fractions were pooled, concentrated, flash-frozen in liquid nitrogen and stored at -80 °C.

The selenomethionine (SeMet) substituted *ct*eIF3b was expressed in the methionine auxotroph *E. coli* strain B834 (DE3). The cells were initially grown in M9 mineral medium supplied with 50 $mg \cdot L^{-1}$ methionine until an OD_{600} of ~0.6 was reached. The medium was then removed by centrifugation (2,500 xg, 15 min, 4 °C) and the cells were resuspended in new M9 medium, which was supplemented with 50 $mg \cdot L^{-1}$ selenomethionine after residual methionine was depleted. Target protein expression was induced by addition of IPTG to a final concentration of 0.25 mM. The purification of selenomethionine labeled *ct*eIF3b was analogous to the wild type protein purification.

Crystallization and X-ray data collection

Initial crystallization screening for full-length *ct*eIF3b (12 $mg \cdot ml^{-1}$) was performed using the sitting drop vapor diffusion method with droplet compositions of 0.25µl : 0.25µl or 0.25µl : 0.125µl for protein and reservoir solutions, respectively. In total, 1056 commercially available conditions at 4 °C and 20 °C have been tested. Bipyramidal crystals were obtained at 20 °C from both droplets from JBScreen Nuc-Pro HTS (Jena Bioscience) with 15% (w/v) PEG 20k, 0.08 M $MnCl_2$, 0.1 M MES, pH 6.5 after approximately 30 days. Crystals of SeMet derivative protein were grown under the same conditions. Prior to X-ray data collection, crystals were transferred into the cryoprotectant buffer (crystallization condition supplemented with 23% (v/v) of ethylene glycol) for 5 seconds and then plunged into liquid nitrogen. A native diffraction dataset was collected from a single crystal at 100K at beamline P13 at PETRA III, DESY (Hamburg, Germany), while the MAD datasets of Se-Met substituted protein crystals were collected at wavelengths of 0.9798 Å (peak), 0.9799 Å

(inflection) and 0.9775 Å (remote) at beamline 14.1, BESSY II (Berlin, Germany, (Mueller et al., 2012)).

Structure determination and refinement

Data were processed and scaled using the programs XDS and XSCALE (Kabsch, 2010a, b) to a final resolution of 2.72 Å and 3.30 Å for native and Se-Met derivative *ct*eIF3b, respectively. Using the Se-MAD peak, inflection and remote diffraction datasets, six (out of seven expected) selenium atoms were found using the program SHELXD (Sheldrick, 2008) and initial phases were calculated at 4 Å resolution. Se positions were further refined using SHARP (Vonrhein et al., 2007) followed by density modification using the program Solomon (Abrahams and Leslie, 1996). An initial poly-alanine model was built by ARP/wARP (Langer et al., 2008) and refined against Se-Met data with Refmac5 (Morris et al., 2004; Perrakis et al., 1999), which was used for manual rebuilding and sequence assignment in Coot (Emsley et al., 2010). Due to the lack of isomorphism between Se-Met and native crystals (the c axis differs in length by 14.7 Å, see Supplementary Table S1), this initial model of the *ct*eIF3b WD40 domain was positioned by molecular replacement into the unit cell of the native crystal. Further model optimization and completion has been performed against the native data. Prior to structural refinement, randomly selected 5% test set of the reflections were set aside for the calculation of R_{free} as a quality monitor (Brunger, 1992, 1993). Refinement was performed with the PHENIX package (Adams et al., 2010). The electrostatic surface potential was calculated with PDB2PQR (Dolinsky et al., 2007) and displayed using the APBS plugin of the PyMol software.

Analytical size exclusion chromatography

An analytical Superdex 200 (10/300) column (GE healthcare) was used for the *in vitro* reconsititution of eIF3 subcomplexes. The buffer contained 25 mM HEPES/NaOH, pH 7.5, 150 mM NaCl, 5% glycerol and 2.5 mM β-ME. In each case, ~50 µg protein in a volume of 400 µl was injected on the column at a flow rate of 0.5 ml·min^{-1}.

GST pull-down assays

50 µg GST fusion protein was mixed with two-fold molar excess of non-tagged protein in a buffer containing 25 mM HEPES/NaOH, pH 7.5, 100 mM NaCl, 5% glycerol and 2.5 mM β-ME and incubated for 30 min with 100 µl glutathione beads. After washing four times with 1 mL buffer, bound protein was eluted with the same buffer containing additional 30 mM reduced glutathione.

Supplemental References

Abrahams, J.P., and Leslie, A.G. (1996). Methods used in the structure determination of bovine mitochondrial F1 ATPase. Acta Crystallogr. D *52*, 30-42.

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.

Brunger, A.T. (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature *355*, 472-475.

Brunger, A.T. (1993). Assessment of phase accuracy by cross validation: the free R value. Methods and applications. Acta Crystallogr D Biol Crystallogr *49*, 24-36.

Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T.J., Higgins, D.G., and Thompson, J.D. (2003). Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Res *31*, 3497-3500.

Dolinsky, T.J., Czodrowski, P., Li, H., Nielsen, J.E., Jensen, J.H., Klebe, G., and Baker, N.A. (2007). PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. Nucleic Acids Res. *35*, W522-525.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr D Biol Crystallogr *66*, 486-501.

Kabsch, W. (2010a). Integration, scaling, space-group assignment and post-refinement. Acta Crystallogr D Biol Crystallogr *66*, 133-144.

Kabsch, W. (2010b). XDS. Acta Crystallogr D Biol Crystallogr *66*, 125-132.

Langer, G., Cohen, S.X., Lamzin, V.S., and Perrakis, A. (2008). Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. Nat Protoc *3*, 1171-1179.

Morris, R.J., Zwart, P.H., Cohen, S., Fernandez, F.J., Kakaris, M., Kirillova, O., Vonrhein, C., Perrakis, A., and Lamzin, V.S. (2004). Breaking good resolutions with ARP/wARP. J Synchrotron Radiat *11*, 56-59.

Mueller, U., Darowski, N., Fuchs, M.R., Forster, R., Hellmig, M., Paithankar, K.S., Puhringer, S., Steffien, M., Zocher, G., and Weiss, M.S. (2012). Facilities for macromolecular crystallography at the Helmholtz-Zentrum Berlin. J Synchrotron Radiat *19*, 442-449.

Perrakis, A., Morris, R., and Lamzin, V.S. (1999). Automated protein model building combined with iterative structure refinement. Nat Struct Biol *6*, 458-463.

Sheldrick, G.M. (2008). A short history of SHELX. Acta Crystallogr A *64*, 112-122.

Vonrhein, C., Blanc, E., Roversi, P., and Bricogne, G. (2007). Automated structure solution with autoSHARP. Methods Mol. Biol. *364*, 215-230.