nature structural & molecular biology

The cryo-EM structure of the UPF–EJC complex shows UPF1 poised toward the RNA 3′ end

Roberto Melero[1,5], Gretel Buchwald[2,5], Raquel Castaño[1], Monika Raabe[3], David Gil[4], Melisa Lázaro[4], Henning Urlaub[3], Elena Conti[2] & Oscar Llorca[1]

Nonsense-mediated mRNA decay (NMD) is a eukaryotic surveillance pathway that degrades aberrant mRNAs containing premature termination codons (PTCs). NMD is triggered upon the assembly of the UPF surveillance complex near a PTC. In humans, UPF assembly is prompted by the exon junction complex (EJC). We investigated the molecular architecture of the human UPF complex bound to the EJC by cryo-EM and using positional restraints from additional EM, MS and biochemical interaction data. The heptameric assembly is built around UPF2, a scaffold protein with a ring structure that closes around the CH domain of UPF1, keeping the helicase region in an accessible and unwinding-competent state. UPF2 also positions UPF3 to interact with the EJC. The geometry is such that this transient complex poises UPF1 to elicit helicase activity toward the 3′ end of the mRNP.

The expression of eukaryotic genes is regulated at multiple levels to control the production of functional proteins at the appropriate amount, location and time in different cell types. The modulation of mRNA levels by targeted degradation is a widespread mechanism to downregulate gene expression post-transcriptionally. Several pathways mediate the depletion of the translatable pool of physiological and nonphysiological transcripts (reviewed in refs. 1,2). NMD was originally discovered as the surveillance pathway that detects and degrades mRNAs with PTCs (reviewed in refs. 3,4). These aberrant mRNAs arise frequently because of germline mutations in inherited genetic disorders, pre-mRNA processing errors and nonproductive rearrangements at the DNA or RNA level (reviewed in refs. 5,6). NMD also modulates the steady-state level of physiological mRNAs, amounting to ~10% of the transcriptome (reviewed in ref. 7).

The NMD pathway is evolutionary conserved in eukaryotes and essential in humans (reviewed in ref. 8). Work in different model organisms has shown that NMD requires translating ribosomes and a combination of *cis*-acting elements and *trans*-acting factors to signal whether the context of translation termination is physiological or aberrant[9,10]. *Cis*-acting elements can originate from the 3′ untranslated region (UTR), whose length and features influence the process of translation termination (the 'faux 3′ UTR' model)[9,11,12]. In addition, a major determinant that promotes NMD in human cells derives from splice junctions[3,4]. Here, four proteins assemble onto mRNA upon splicing to form the EJC, a stable constituent of the spliced mRNP[13]. In humans, NMD is elicited when a stop codon is present at least 50–54 nucleotides (nt) upstream of a splice junction. This observation, made decades ago, is now interpreted in molecular terms as the requirement of a minimal distance for a ribosome stalled at a PTC to establish the appropriate network of interactions with a downstream EJC. Although this mechanism was originally thought of as a specialization to increase the efficiency of NMD in vertebrates, EJC-mediated NMD has recently also been discovered in flies[14].

The cross-talk between the terminating ribosome and the EJC is mediated by the *trans*-acting factors UPF1 (also known as SMG2 and RENT1), UPF2 (also known as SMG3) and UPF3 (also known as SMG4). UPF1 is an RNA helicase that associates with ribosomes via interaction with release factors[15–17]. The catalytic activity of UPF1 is essential for NMD and is triggered upon formation of the surveillance complex[18–20]. The switch from 'off' to 'on' state is mediated by a conformational change that occurs upon binding of UPF2, a multidomain protein that concomitantly binds UPF3 (refs. 20,21). Although these features are conserved from yeast to humans, the choreography of interactions centered at the UPF complex is markedly more complex in metazoans. In humans, UPF3 contains a C-terminal low-complexity region that binds to the EJC[22,23]. UPF1 also contains N-terminal and C-terminal low-complexity regions that are phosphorylated upon formation of the surveillance complex by the SMG1 kinase in conjunction with its regulators SMG8 and SMG9 (refs. 24–27). UPF1 phosphorylation is thought to induce translational repression[28], cause dissociation from the ribosome[25] and recruit downstream *trans*-acting factors (SMG6 and SMG5–SMG7)[29], eventually leading to mRNA degradation. Additional layers of complexity have also been reported, including interaction of UPF1 with the mRNA 5′ end[30] and with the 3′ UTR[31], as well as alternative pathways[32].

According to the prevalent model, formation of the UPF surveillance complex in human cells occurs when UPF2–UPF3 bound to an EJC comes close to UPF1 bound to a terminating ribosome.

[1]Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas (Spanish National Research Council), Madrid, Spain. [2]Max Planck Institute of Biochemistry, Department of Structural Cell Biology, Martinsried, Germany. [3]Max Planck Institute of Biophysical Chemistry, Göttingen, Germany. [4]Structural Biology Unit, CIC bioGUNE, Parque Tecnológico de Bizkaia, Bizkaia, Spain. [5]These authors contributed equally to this work. Correspondence should be addressed to O.L. (ollorca@cib.csic.es) or E.C. (conti@biochem.mpg.de).

Received 18 November 2011; accepted 22 March 2012; published online 22 April 2012; doi:10.1038/nsmb.2287

After dissociation from the ribosome, the UPF1 helicase is thought to move along the nucleic acid in the 5′→3′ direction, stripping the proteins bound to it and disassembling the downstream mRNP[33]. How is the UPF complex formed on the EJC such that it can elicit helicase activity toward the 3′ end of the mRNP? Partial structural information from X-ray crystallography of the EJC[23,34,35] and of UPF proteins and subcomplexes[21,36–38] is available, but how these components are pieced together in the surveillance complex is unknown. This is partly because it is difficult to obtain structural information about intermediates in transient macromolecular complexes formed by structured and unstructured regions. In this study, we produced and stabilized the heptameric UPF–EJC complex biochemically; we report its cryo-EM structure at a resolution of 16 Å. We have interpreted the cryo-EM reconstruction on the basis of available atomic models and additional structural, MS and biochemical data.

RESULTS

Purification of chemically homogeneous UPF–EJC complexes

The domain structure of individual UPF proteins is known from earlier structural studies and from computational predictions (**Fig. 1a**). Human UPF1 has a folded core (residues 115–914) that contains a regulatory zinc-knuckle domain (the CH domain) and the catalytic helicase region (formed by domains 1B, 1C, RecA1 and RecA2). The low-complexity N- and C-terminal regions are not involved in binding UPF or EJC proteins[20] and were not included in this study. The CH domain of UPF1 binds UPF2, a 147-kDa protein with three predicted middle domain of eukaryotic initiation factor 4G (MIF4G) domains as well as regions with limited secondary structure (particularly at the N and C termini; **Fig. 1a**). The C-terminal region of UPF2 contains the main binding determinant for UPF1, as has been shown both biochemically[20,39] and with structural studies[38]. The third MIF4G domain ($MIF4G_3$) binds UPF3. The interacting region of UPF3 is a domain with the fold of an RNA-recognition motif (RRM)[36]. Human UPF3 also contains a C-terminal low-complexity region that is not required to form the UPF complex[20], but contains the EJC-binding motif (EBM)[22,23]. The EJC core is formed by MAGO, the RRM domain of Y14, eIF4AIII, a low-complexity region of Barentsz (Btz) (the SELOR domain, residues 137–286), RNA and ATP[34,35,40].

We purified the human EJC core, $UPF1_{115–914}$ (referred to below as UPF1), full-length UPF2 and one of the variants of UPF3 (UPF3b, also known as UPF3X) as described[21,23]. The individual recombinant components were combined to reconstitute and purify the UPF1–UPF2–UPF3–EJC complex by size-exclusion chromatography (**Supplementary Fig. 1a**). To facilitate the interpretation of the EM maps, we also purified the UPF2–UPF3–EJC and UPF1–UPF2–$UPF3_{RRM}$ complexes. Despite the presence of low-complexity regions, we obtained samples of these multisubunit complexes with relatively high chemical homogeneity and in amounts suitable for structural studies (**Fig. 1b**).

Stabilization of structurally homogeneous UPF–EJC

The UPF1–UPF2–UPF3–EJC complex was stable at protein concentrations used for biochemical reconstitutions (~25 μM) but dissociated at the low concentrations required in EM analysis to prevent contacts between individual molecules (about 0.05–0.2 μM; data not shown). We proceeded by stabilizing UPF1–UPF2–UPF3–EJC using the GraFix method, a sample preparation procedure for EM single-particle analysis that involves mild and gradual chemical fixation during sucrose gradient centrifugation[41]. We evaluated the cross-linked particles that sedimented in each fraction of the gradient according to their size and content. Two fractions showed a sharp, well-defined band in SDS-PAGE (fractions f7 and f8, **Supplementary Fig. 1b,c**).

Negative-stain EM of the cross-linked samples that migrated with higher apparent molecular mass in SDS-PAGE showed the presence of large aggregates (f5 and f6, data not shown). Samples of lower apparent molecular mass (f8 and f9) contained heterogeneous particles of small size, indicating complex dissociation. The sample selected for subsequent EM analysis (f7) showed monodisperse particles that were homogeneous in size and shape (**Supplementary Fig. 2a,b**). We used a similar procedure to assess the content and quality

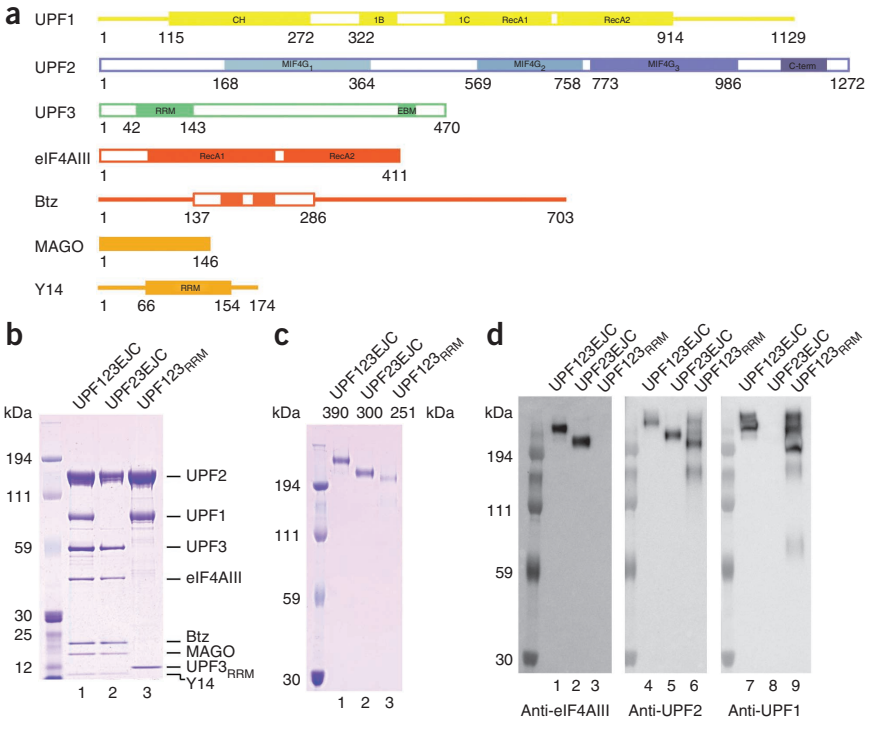

Figure 1 Reconstitution of biochemically pure and stable UPF–EJC complexes. (**a**) Domain arrangement of human proteins used in this study. Color-filled rectangles, regions known from crystal structures or homology models (in the case of $MIF4G_1$ and $MIF4G_2$ of UPF2). Individual domains of UPF1 (yellow), UPF2 (light blue, homology models; dark blue, X-ray structures), UPF3 (green, UPF3b residue numbering), eIF4AIII and Btz (red), MAGO and Y14 (orange) are indicated. RecA, recombinase A; CH, cysteine-histidine-rich; MIF4G, middle domain of eukaryotic initiation factor 4. (**b**) Samples from peak fractions of final size-exclusion chromatography purification of UPF1–UPF2–UPF3–EJC (UPF123EJC), UPF2–UPF3–EJC (UPF23EJC) and UPF1–UPF2–$UPF3_{RRM}$ ($UPF123_{RRM}$) complexes, separated by SDS-PAGE together with protein marker (left lane). (**c**) Samples of UPF123EJC, UPF23EJC and $UPF123_{RRM}$ complexes after cross-linking with glutaraldehyde, separated by SDS-PAGE together with protein marker (left lane). (**d**) Western blot analysis of the protein complexes using antibodies to eIF4AIII (a subunit EJC) and to UPF1 and UPF2 (Online Methods).

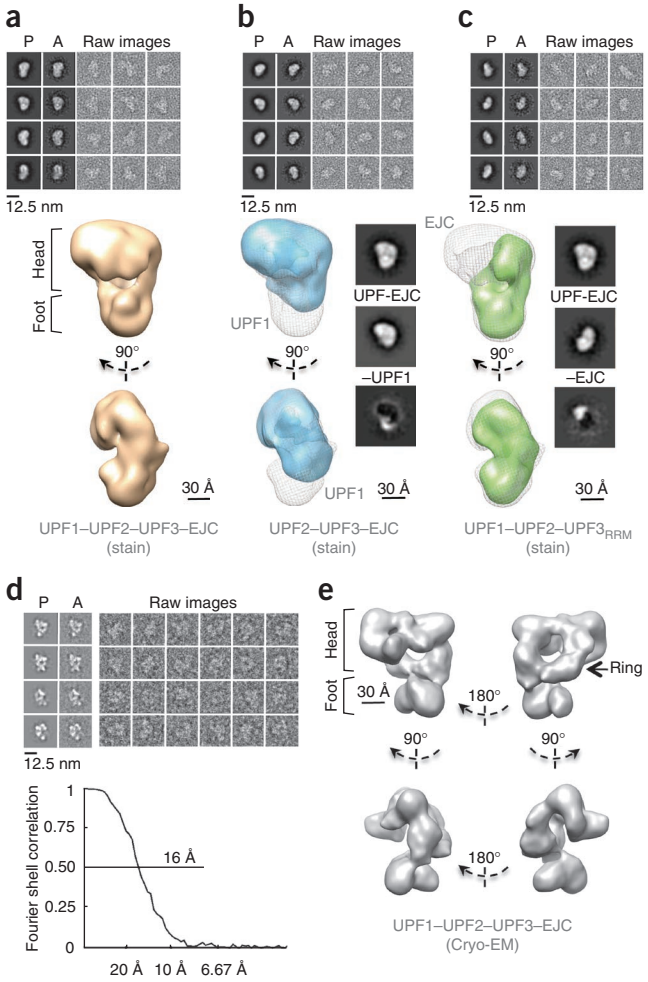

Figure 2 EM analysis of UPF–EJC complexes. (**a**) Negative-stain structure of UPF1–UPF2–UPF3–EJC. Top, projections (P), 2D averages (A) and selected raw images. Bottom, reconstruction in two views related by 90° rotation. (**b**) Negative-stain structure of UPF2–UPF3–EJC. Top, EM analysis as in **a**. Bottom, reconstruction of UPF2–UPF3–EJC (blue) fitted within negative-stain structure of UPF1–UPF2–UPF3–EJC (gray mesh) in same orientations as in **a**. Difference density identifies location of UPF1 in complex. Inset, difference between a reference-free average for UPF1–UPF2–UPF3–EJC and UPF2–UPF3–EJC. (**c**) Negative-stain structure of UPF1–UPF2–UPF3$_{RRM}$. Top, EM analysis as in **a**. Bottom, reconstruction of UPF1–UPF2–UPF3$_{RRM}$ (green) is fitted within negative-stain structure of UPF1–UPF2–UPF3–EJC (gray mesh) in same orientations as in **a**. Difference density identifies location of the EJC in complex. Inset, difference between reference-free averages of UPF1–UPF2–UPF3–EJC and UPF1–UPF2–UPF3$_{RRM}$. (**d**) Cryo-EM analysis of unstained UPF1–UPF2–UPF3–EJC in vitrified ice. Top, as in **a**. Bottom, Fourier shell correlation obtained by angular refinement methods. (**e**) Cryo-EM reconstruction of UPF1–UPF2–UPF3–EJC. Structure is shown in four different orientations, as indicated. Orientations on top and bottom left correspond to views of negative-stain structure in **a**.

of the UPF2–UPF3–EJC and UPF1–UPF2–UPF3$_{RRM}$ samples (**Supplementary Figs. 1d**,**e** and **2c**,**d**). The three complexes used in the structural analysis migrated by SDS-PAGE as we expected on the basis of their relative molecular mass (**Fig. 1c**). Finally, we probed the three cross-linked complexes by MS (discussed below) and by western blotting using antibodies to the EJC subunit eIF4AIII and UPF subunits. The results confirmed the protein composition we expected: we detected eIF4AIII only in the UPF1–UPF2–UPF3–EJC and UPF2–UPF3–EJC samples (**Fig. 1d**, left), UPF1 only in the UPF1–UPF2–UPF3–EJC and UPF1–UPF2–UPF3$_{RRM}$ samples (**Fig. 1d**, right) and UPF2 in all three samples (**Fig. 1d**, middle).

Cryo-EM structure of the UPF1–UPF2–UPF3–EJC complex

To generate a starting model suitable for cryo-EM reconstructions, we first analyzed the UPF1–UPF2–UPF3–EJC complex in negative stain. We obtained an *ab initio* model with the random conical tilt (RCT) method[42] and used it as reference for the refinement of >10,000 individual images (**Fig. 2a** and **Supplementary Figs. 2b** and **3**). The resulting structure at a resolution of 32 Å showed an elongated particle with a bulky density region (referred to as 'head') connected to a thinner region (the 'foot'; **Fig. 2a**). We next examined the location of UPF1 and of the EJC in the complex by determining and comparing the negative-stain structures of complexes missing either of these components. In the case of UPF2–UPF3–EJC, the negative-stain images showed a shorter particle as compared with the intact complex (**Fig. 2b**, top). The 3D structure of UPF2–UPF3–EJC showed that the foot region was missing (**Fig. 2b**, bottom), as we also observed by calculating the difference between 2D averages of the full complex and UPF2–UPF3–EJC (**Fig. 2b**, insert). In the case of the UPF1–UPF2–UPF3$_{RRM}$ complex, the difference between 2D averages as well as the 3D structure showed that a portion of the head was missing as compared with the intact complex (**Fig. 2c**). Thus, the difference maps located UPF1 at the foot and the EJC within the head, at opposite ends of the complex. With the exception of the missing portions, the EM reconstructions of three complexes had markedly similar features (**Fig. 2a–c**).

We used the negative-stain structure of UPF1–UPF2–UPF3–EJC as a reference template to start the angular refinement of images obtained from unstained frozen samples (**Supplementary Fig. 3**). We processed >85,000 unstained particles to obtain a cryo-EM structure of UPF1–UPF2–UPF3–EJC, estimated at a resolution of 16 Å (using a Fourier shell correlation criterion of 0.5; **Fig. 2d**). The cryo-EM map showed the overall features of the corresponding negative-stain map but with higher resolution and greater details (**Fig. 2e**). The cryo-EM map showed that the foot has a bilobal shape and the head contains a distorted ring-like structure at the front connected to a compact density at the back end (**Fig. 2e**). The ring-like structure is also connected to the foot. We obtained a pseudoatomic model for the three-dimensional architecture of the UPF1–UPF2–UPF3–EJC surveillance complex by fitting known atomic structures or homology models according to the features of the map and to the topological information obtained from additional negative-stain EM reconstructions, as well as from MS and biochemical data (discussed below).

The helicase region is exposed in the UPF–EJC complex

We first docked UPF1 to the foot (**Fig. 3**), the region where we had localized it by comparing the UPF1–UPF2–UPF3–EJC and the UPF2–UPF3–EJC negative-stain reconstructions (**Fig. 2b**). In the cryo-EM map, the foot comprises a larger lobe (~45 Å in diameter) and a smaller lobe (~30 Å in diameter). In the crystal structure of UPF1 bound to the C-terminal domain of UPF2 (UPF1–UPF2$_{Cterm}$), the helicase region is organized in two modules, a smaller module encompassing the RecA2 domain and a larger one encompassing the RecA1, 1B and 1C domains[38]. The two modules are in an open conformation compared with the closed conformation that is induced upon RNA binding[21]. We docked the atomic coordinates of UPF1–UPF2$_{Cterm}$ such that the RecA1-1B-1C and RecA2 modules fit into the large and small lobes of density of the cryo-EM map, respectively

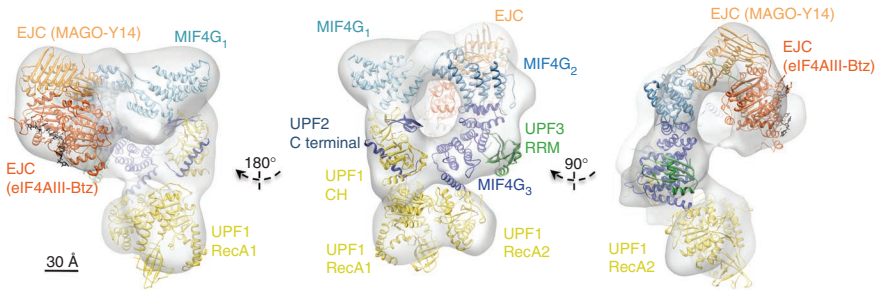

Figure 3 Molecular architecture of UPF1–UPF2–UPF3–EJC complex. Three views of cryo-EM structure of UPF1–UPF2–UPF3–EJC fitted with atomic structures of EJC bound to a small fragment of UPF3 and containing a short RNA fragment (black; PDB 2XB2)[23], of UPF1 bound to the C-terminal domain of UPF2 (PDB 2WJV)[38], of the C-terminal MIF4G bound to an N-terminal fragment of UPF3 (PDB 1UW4)[36] and of two homology models for $MIF4G_1$ and $MIF4G_2$. Colors are as in **Figure 1a**.

(cross-correlation = 0.71). In this model, the CH domain and $UPF2_{Cterm}$ fit into the density of the ring-like structure, but parts of domains RecA1 and 1C in the helicase region were outside the density of the EM map (**Supplementary Fig. 4a**). We independently validated this model by labeling the UPF–EJC complex with an antibody recognizing residues 250–300 of UPF1 (in the CH domain). The antibody labeled a region of the ring compatible with the model (**Fig. 4a**).

The fit of UPF1 within the foot could be improved by a small rigid-body movement of the entire helicase region such that only a portion of domain 1B would be out of the density (a domain known to assume different conformations in different crystal structures of UPF1; refs. 21,37,38; cross-correlation = 0.73) and concomitantly by a rotation of the CH domain relative to the helicase region (cross-correlation = 0.88; **Fig. 3** and **Supplementary Fig. 4b**). Movement of the CH domain relative to the helicase region would be feasible in the context of earlier structural information[21,38]. Independent of the details of the exact orientation of the CH domain relative to the helicase region, cryo-EM analysis shows that the CH domain is linked to the large lobe (that is, to RecA1, corresponding to an unwinding-competent state). It also shows that the helicase region sticks out from the surveillance complex. The RecA2 domain in particular is not restrained by contacts with other subunits, rationalizing how RNA could access the binding site on UPF1 when the helicase is within the surveillance complex.

UPF2 is a structural ring-like scaffold

We next interpreted the ring-like density that is connected both to the foot (UPF1) and to the back of the head (where we localized the EJC) by fitting the missing folded regions of the UPF complex: the CH domain of UPF1 (~17 kDa), the three MIF4G domains of UPF2 (~90 kDa) and the RRM domain of UPF3 (19 kDa; **Figs. 2e** and **3**). As UPF2 is the largest component of the three, we hypothesized that it accounts for most of the ring-like density. To validate this, we investigated the non-cross-linked negative-stain structure of UPF2. Single molecules of UPF2 appeared in negative-stain EM as distorted circles, with apparent conformational heterogeneity (**Fig. 4b**). We used RCT reconstruction to obtain an *ab initio* model of UPF2 from homogenous subclasses of images. The resulting structure showed an open

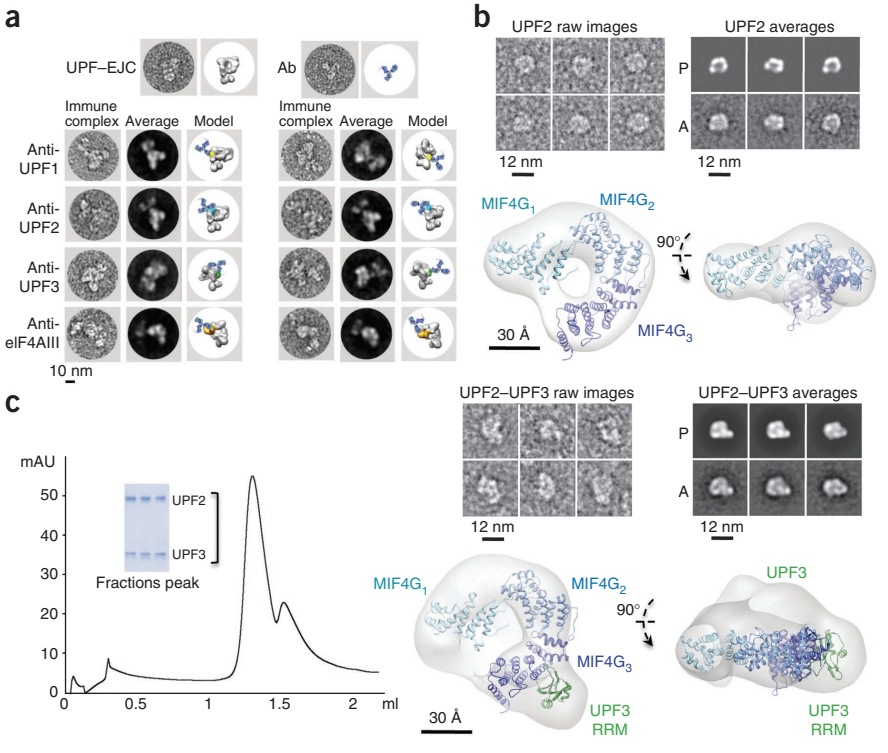

Figure 4 Validation of structural model of UPF–EJC complex. (**a**) Antibody labeling of UPF–EJC complex using polyclonal antibodies to residues 250–300 of UPF1, residues 100–200 of UPF2, residues 1–50 of UPF3 and eIF4AIII (also used for **Fig. 1d**). Two typical labeled complexes for each case together with 2D averages for several images of antibody–UPF–EJC complexes in similar orientations. Using these images and after comparison with the projections from UPF–EJC we inferred a model for each of the immune complexes. Putative location of epitopes recognized by each antibody according to pseudoatomic model is colored in model (UPF1, yellow; UPF2, blue; UPF3, green; eIF4AIII, orange). Images of UPF–EJC and an isolated antibody (Ab) at the same magnification are reference for comparison. (**b**) EM and structure of UPF2. Top, representative images of particles, projections from structure (P) and 2D averages (A). Bottom, two views of structure of UPF2 in which pseudoatomic model of UPF2 proposed within UPF–EJC complex (**Fig. 3**) was fitted. (**c**) Purification, EM and structure of UPF2–UPF3 complex. Peak fraction of size-exclusion chromatography containing UPF2 and UPF3 (inset, SDS-PAGE of these fractions) was analyzed by EM. Top, representative images of raw particles, projections from structure (P) and 2D averages (A). Bottom, two views of structure of UPF2–UPF3. Structure of UPF2-UPF3 was fitted with atomic model proposed for these proteins in the context of the intact complex (**Fig. 3**).

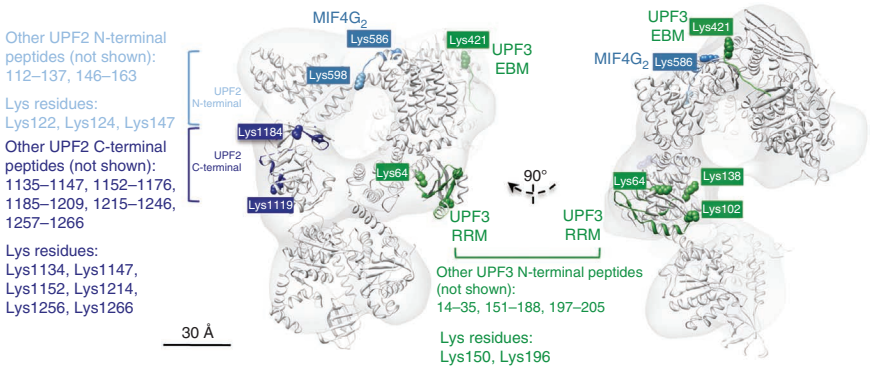

Figure 5 Regions of protein-protein interactions detected by MS. Peptides detected by ESI-MS analysis that show areas of contact between components of the complex, colored according to color codes in **Figure 1** (UPF2: 587–598, 1120–1128, 1185–1209; UPF3: 57–64, 103–138 and 422–434). Lysine residues buried specifically because of protein-protein interactions within UPF–EJC are shown as atoms (UPF2: Lys586, Lys598, Lys1119, Lys1184; UPF3: Lys64, Lys102, Lys138 and Lys421). Peptides and lysine residues outside known atomic structures, which cannot be represented, are indicated next to model.

ring-like molecule with similar dimensions to those in the cryo-EM map, corroborating the assignment of the ring to UPF2 (**Fig. 4b**).

To fit the available atomic models or homology models into the ring-like density (**Fig. 3**), we obtained additional restraints by MS analysis. We treated the UPF1–UPF2–UPF3$_{RRM}$ cross-linked sample with trypsin and sequenced the peptides resulting from the proteolytic cleavage by ESI-MS analysis (**Supplementary Fig. 5**)[43]. With this approach, only regions of the proteins in which lysine residues have not been modified by glutaraldehyde can be detected; this suggests that they are protected and not accessible to solvent (either because they are in the core of molecule or because they are protected by interactions in the UPF–EJC complex). The comparison of peptides detected for UPF1–UPF2–UPF3$_{RRM}$, UPF2–UPF3–EJC and UPF1–UPF2–UPF3–EJC shows exposed regions specifically protected by protein-protein interactions within UPF–EJC. The analysis showed the presence of peptides that we expected from the interaction interfaces in the UPF2 MIF4G$_3$–UPF3$_{RRM}$ crystal structure (UPF3 residues 57–64 and 103–138; **Fig. 5** and **Supplementary Fig. 5**; ref. 36) and the UPF1–UPF2$_{Cterm}$ crystal structures (UPF2 residues 1120–1128, 1152–1176 and 1185–1209; **Fig. 5** and **Supplementary Fig. 5**; ref. 38). Through MS analysis, we also detected other peptides in the C and N termini of UPF2 (residues 1215–1246 and 1257–1266; residues 112–137 and 146–163; **Supplementary Fig. 5**). These results indicate that the N-terminal region and almost the entire C-terminal region of UPF2 are buried in the complex, supporting a ring-like architecture of UPF2.

In our model (**Fig. 3**), we arranged the three MIF4G domains of UPF2 consecutively in a circular fashion, with the N and C termini approaching each other and the CH domain of UPF1 (cross-correlation MIF4G$_1$ = 0.88; cross-correlation MIF4G$_2$ = 0.79; cross-correlation MIF4G$_3$ = 0.89). This spatial arrangement of the MIF4G domains is consistent with topological features of the proteins and of the cryo-EM density. First, MIF4G$_2$ and MIF4G$_3$ are close together as in the primary structure of UPF2. Second, there is sufficient density to accommodate the 200-residue connecting region between MIF4G$_1$ and MIF4G$_2$ (**Fig. 1a**) and to accommodate the bound UPF3$_{RRM}$ to MIF4G$_3$. This configuration also places the C-terminal end of MIF4G$_3$ near the N terminus of the UPF2 C-terminal domain (which binds the CH domain in UPF1; **Fig. 3**). Notably, the structure of non-cross-linked UPF2 fitted this pseudoatomic model (**Fig. 4b**). Although the docking is probably not accurate in detail, the ring-like architecture of UPF2 closed around the CH domain of UPF1 is in agreement with results from earlier coimmunoprecipitation studies showing that deletions at the N and C termini of UPF2 (Δ94–133 and Δ1094–1272, respectively) both diminished the interaction with UPF1, albeit to a different extent[39].

To corroborate this model we used a double strategy. First, residues 100–200 of UPF2 and 1–50 of UPF3 were localized in the images of UPF–EJC by labeling with polyclonal antibodies that were found to bind regions compatible with the model (**Fig. 4a**). Second, we analyzed the non-cross-linked negative-stain structure of the UPF2–UPF3 complex (**Fig. 4c**). The images of UPF2–UPF3 were slightly larger and less rounded than those of UPF2. The structure of the complex showed a ring with extra densities that fitted the UPF2–UPF3 model proposed from the cryo-EM structure. In the UPF2–UPF3 structure, density for full-length UPF3 outside the RRM domain extends toward the EJC-interacting region.

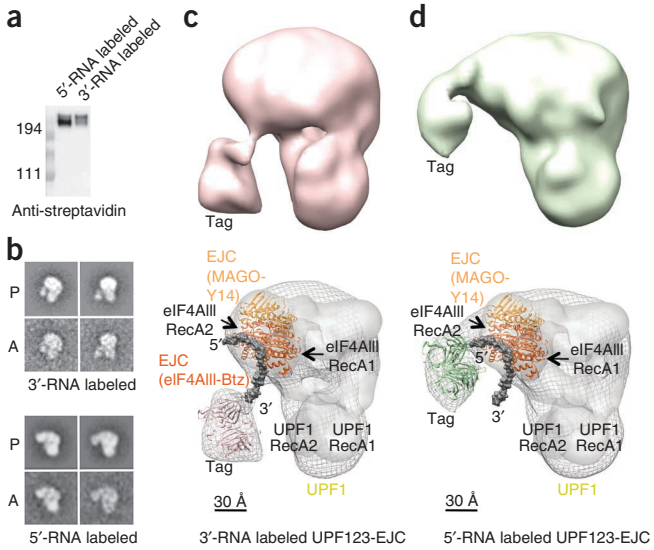

Figure 6 Labeling of RNA 3′ and 5′ ends within UPF–EJC complex. (**a**) Western blot analysis of UPF1–UPF2–UPF3–EJC complexes, reconstituted with a 15-mer poly(U) ssRNA containing a biotin moiety at the 5′ or 3′ end and bound to streptavidin. (**b**) Projections (P) and averages (A) for each of labeled complexes. Label is a dense spot in the vicinity of the UPF–EJC complex. (**c**) Negative-stain structure of UPF1–UPF2–UPF3–EJC (pink) with RNA 3′ end labeled with a biotin moiety bound to streptavidin. Bottom, fitting of cryo-EM structure of UPF1–UPF2–UPF3–EJC (white transparent density) within 3′-labeled complex to facilitate mapping of RNA end. Atomic structure of EJC complex fitted within the cryo-EM reconstruction has same color codes as in **Figure 3**. RNA in complex (black) was modeled by extending RNA crystallized with the EJC (PDB 2XB2)[23]. The structure of a streptavidin tetramer (PDB 1MEP)[50] is fitted within the density of the tag (pink). (**d**) Negative-stain structure of UPF1–UPF2–UPF3–EJC complex with RNA 5′ end with a biotin moiety bound to streptavidin (in green). Bottom, fitting of cryo-EM structure of UPF1–UPF2–UPF3–EJC with a similar representation as for the 3′-labeled complex. Tag is in green at bottom.

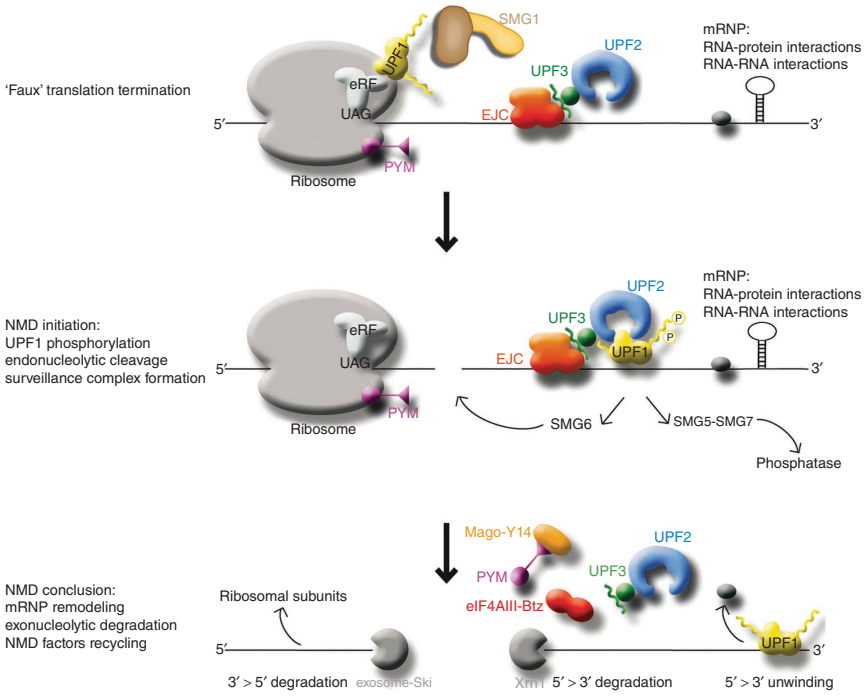

Figure 7 Model of surveillance complex during NMD. Scheme recapitulates some of the most important steps in mammalian NMD, incorporating information from the UPF–EJC cryo-EM structure in a simplified version. Premature translation termination leads to recruitment of UPF1. If the recruitment occurs >30 nt from a downstream UPF2–UPF3–EJC complex, NMD proceeds through recruitment of the SMG1 kinase, UPF1 phosphorylation and binding to UPF2–UPF3–EJC, recruitment of SMG5-SMG7 and SMG6 factors and endonucleolytic cleavage of the mRNA by SMG6. The final stages of NMD involve mRNP remodeling and degradation. The architecture of the surveillance complex places the UPF1 helicase at the appropriate position to remodel the 3′ end of the mRNP (RNA-RNA and RNA-protein interactions, gray sphere). Remodeling of the EJC (at the 5′ end of UPF1) requires ribosome-binding protein PYM. mRNP remodeling allows RNA decay exonucleases (Xrn1 and exosome-Ski) to degrade the body of the nucleic acid.

Connections between UPF3 and the EJC

The last element of the UPF1–UPF2–UPF3–EJC cryo-EM map that we interpreted was the compact density at the back end of the head where we had located the EJC from the analysis of the difference density between the UPF1–UPF2–UPF3–EJC and UPF1–UPF2–UPF3$_{RRM}$ negative-stain maps (**Figs. 2** and **3**). This location was also supported by labeling experiments using antibodies to the EJC subunit eIF4AIII (**Fig. 4a**). When rendered at the resolution of the cryo-EM map, the EJC has an overall elliptical shape. Half of the ellipse (including the RecA1 domain of eIF4AIII and the C terminus of the Btz SELOR domain) is slightly smaller than the other half (including MAGO, Y14, the RecA2 domain of eIF4AIII and the N terminus of Btz SELOR). The larger lobe is also where a segment from the C terminus of UPF3 binds[23]. These features pointed to the orientation of the complex in the density, where it could be docked computationally with a cross-correlation coefficient of 0.84 (**Supplementary Fig. 6**). An earlier negative-stain structure of an EJC reconstituted with full-length Btz[44] fits with our cryo-EM map, with an additional density that probably corresponds to regions of Btz outside the SELOR domain (not shown).

The cryo-EM map showed a connection in density between the back of the head (the EJC) and the ring-like structure at the front (UPF2–UPF3), indicating a point of contact. To map the interacting proteins and surfaces, we used the MS approach described above to assess the protection pattern of UPF samples with and without the EJC. The comparative analysis showed no marked differences in UPF2. The only exception was a peptide at the N-terminal side of MIF4G$_2$ (residues 587–598) that we detected in the UPF1–UPF2–UPF3$_{RRM}$ and UPF1–UPF2–UPF3–EJC samples, but not in UPF2–UPF3–EJC (**Fig. 5**). This indicates that the N-terminal side of MIF4G$_2$ might change orientation when the UPF2 ring closes upon binding UPF1. The major differences were instead located in UPF3. The EJC-bound samples showed the presence of the EJC-binding motif of UPF3 (residues 422–434, **Fig. 5**) that is known from the EJC–UPF3$_{Cterm}$ crystal structure to interact with MAGO-Y14 and with the RecA2 domain of eIF4AIII[23]. The analysis of EJC-bound samples also showed additional peptides within the regions of UPF3 preceding and following the RRM domain (residues 14–35, 151–188 and 197–205), which are therefore buried in the presence of the EJC[23].

The EJC points the 3′ end of bound RNA toward UPF1

We next investigated how the surveillance complex is oriented onto RNA. In the EJC–UPF3$_{Cterm}$ crystal structure, the 3′ and 5′ ends of an 8-nt single-stranded RNA (ssRNA) are at the antipodes of the elliptical assembly, protruding from the RecA1 and RecA2 domains, respectively[23]. Thus, the positioning of the RNA ends could be extrapolated from the fitting of the crystal structure in the cryo-EM map of UPF–EJC (**Fig. 3**). For an independent validation of the pseudoatomic model, we reconstituted two variants of the EJC using a 15-mer ssRNA labeled with a biotin moiety at either the 3′ or the 5′ end. We assembled the corresponding UPF1–UPF2–UPF3–EJC complexes, coupled them to streptavidin and cross-linked them with the same GraFix procedure used for the untagged complex (**Supplementary Fig. 2**). We verified the presence of the streptavidin tag in both cross-linked samples by western blotting (**Fig. 6a**).

The negative-stain EM structures of the streptavidin-labeled complexes showed a similar architecture to that of UPF1–UPF2–UPF3–EJC, but with additional density connected to the head (**Fig. 6b–d**). We overlaid the negative-stain maps of the streptavidin-labeled complexes onto the cryo-EM map of the unlabeled complex. The fitting was done computationally and showed the additional density (the putative streptavidin tag) in the vicinity of the RecA1 domain of eIF4AIII in the case of the 3′-labeled sample (**Fig. 6c**, bottom) and near the position corresponding to the RecA2 domain of eIF4AIII in the case of the 5′-labeled sample (**Fig. 6d**, bottom). The size of the extra density was what we expected for a streptavidin tetramer (60 kDa) and its distance from the EJC was compatible with the attachment of streptavidin to the 3′ and 5′ ends of the 15-mer biotinylated RNAs used in the biochemical reconstitutions (modeled in **Fig. 6c**,**d**, black). In the cryo-EM map, the 3′ end of the RNA embedded in the EJC points toward the small lobe of UPF1 (**Fig. 6c**, bottom), where we expected the 5′ end of RNA to bind in the course of the unwinding reaction[21] (**Fig. 7**). Modeling of an RNA bound to UPF1 according to known crystal structures[21] within the cryo-EM map of UPF–EJC indicates that the RNA 3′ end in the EJC would be ~80 Å from the RNA 5′ end in UPF1, a distance that would be spanned by a minimum

of 13 nt. With the opposite geometry (that is, if in the surveillance complex UPF1 were positioned at the 5′ end of the EJC), the RNA 3′ end in UPF1 would be ~120 Å from the RNA 5′ end in the EJC. Such distance would be spanned by a minimum of 19 nt and would be incompatible with the 50-nt rule of EJC-dependent NMD, as we discuss below. In the context of the surveillance complex, the UPF1 helicase is thus positioned downstream of the EJC.

DISCUSSION

One of the central steps in the nonsense-mediated mRNA decay pathway is the activation of the UPF1 helicase to remodel the 3′ end of mRNPs undergoing NMD[33]. UPF1 is activated in the context of the surveillance complex with UPF2 and UPF3, whose recruitment in human cells generally depends on the EJC. In this work, we trapped a transient assembly of the NMD pathway, namely the surveillance complex, after its recruitment onto the EJC and before it unlashes its helicase activity toward the 3′ end of the mRNP. It is challenging to obtain structural information on the UPF–EJC complex because of its transient nature. We circumvented this problem by stabilizing the samples with gentle cross-linking, a method successfully used before to overcome the instability of complexes at low concentrations[41,45]. Several observations suggest that the cryo-EM structure we obtained faithfully represents the overall architecture of the assembly. The cryo-EM density has features similar to negative-stained EM maps from both cross-linked and non-cross-linked samples and is consistent with known crystal structures. The MS data also show a pattern we expected on the basis of earlier structural and biochemical data and new interaction sites.

In this study, we found that the UPF–EJC complex is organized into three distinct structural modules formed by a combination of domains of individual proteins: UPF2, the CH domain of UPF1 and the RRM domain of UPF3 form the central module, which is flanked on one side by the helicase region of UPF1 (the 3′ module) and on the other side by the EJC and the C-terminal domain of UPF3 (the 5′ module; **Fig. 7**, middle). The central module is a ring-like structure built around UPF2, the largest subunit of the surveillance complex. The three MIF4G domains of UPF2 probably form the circular scaffold of the ring, which closes upon intermolecular and/or intramolecular interactions of the N- and C-terminal regions of UPF2 with the CH domain of UPF1. This interaction connects the central module to the helicase region of UPF1 and keeps it in an active form, as it sequesters the CH domain in a position where it cannot allosterically inhibit the unwinding activity of UPF1. In the complex, the helicase region of UPF1 is exposed and thus accessible to RNA binding. The central ring also accommodates the RRM domain of UPF3, which is recruited to the third MIF4G domain of UPF2. UPF3 connects the central module to the EJC. The major interaction occurs at the EJC-binding motif within the C terminus of UPF3. Additional contacts at the N- and C-terminal low-complexity regions of UPF3 contribute to the interaction with the EJC and possibly stabilize the orientation of the EJC relative to central module and hence to the helicase region of UPF1.

Current models of NMD envision that, in the surveillance complex, the UPF proteins are positioned at the 5′ end of the EJC, pointing toward the ribosome stalled at a premature stop codon (**Fig. 7**). A minimal distance of 50 nt between the stop codon and the downstream exon-exon boundary is required to elicit EJC-dependent NMD[3,4,9,11,12]. Considering that the 3′ end of an RNA-bound EJC is 20 nt from the exon-exon boundary[13], that the EJC itself covers 8 nt (ref. 13) and that an mRNA needs ~10 nt at the 3′ of a stop codon to exit from the ribosome[46], it follows that the minimal physical distance between the ribosome and the EJC to elicit NMD is ~12 nt. This distance would be sufficient to account for UPF1 binding in isolation[20,21], but not in the context of the UPF–EJC structure. Instead, the pseudoatomic model of the UPF–EJC complex suggests that the UPF1 helicase is positioned at least 13 nt at the 3′ end of the EJC (**Fig. 7**, middle). With hindsight, these findings rationalize how UPF1 is poised to unwind the 3′ end of the mRNP in the final stages of EJC-dependent NMD[21,33], creating the ssRNA segment required for Xrn1-mediated degradation[47] (**Fig. 7**, bottom). They also rationalize the requirement for an additional factor, the ribosomal-binding protein PYM, to remodel the EJC component of the mRNP at the 5′ end of UPF1 (ref. 48; **Fig. 7**, bottom). These findings also raise questions and predictions. First, UPF1 associates with release factors on ribosomes terminating translation prematurely, at the 5′ end of the EJC[15–17] (**Fig. 7**, top). What steps would lead to the repositioning of UPF1 from the 5′ to the 3′ end of the EJC? Is UPF1 the missing link in the 50-nt rule of NMD and, if so, is UPF1 in a helicase-on conformation when bound to the release factors? When does UPF1 phosphorylation take place and how does it affect RNA binding? What is the involvement of the large SMG1 kinase and its partners SMG-8 and SMG-9 in the initial stages of NMD[26]? Finally, the architecture of the UPF–EJC complex predicts that large stem-loop structures in the vicinity of the EJC 3′ end would be sterically incompatible with the recruitment of the UPF proteins and would thus interfere with mRNA degradation. Notably, such a stem-loop structure has been recently identified in the case of EJC-dependent *oskar* mRNA localization in *Drosophila melanogaster*[49].

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/nsmb/.

Accession codes. Electron Microscopy Data Bank: the cryo-EM map of the UPF1–UPF2–UPF3–EJC complex has been deposited under accession code EMD-2048 (http://www.emdatabank.org/index.html).

Note: Supplementary information is available on the Nature Structural & Molecular Biology website.

ACKNOWLEDGMENTS

We thank E. Arias-Palomo, E. Torreira and U. Jayachandran for help in the initial stages of this project, and C. Boulegue and S. Uebel (at the Max Planck Institute Martinsried Core Facility) and members of our labs for critical reading of the manuscript. This work was funded by the Spanish Government (SAF2008-00451 and SAF2011-22988 to O.L.) and the Red Temática de Investigación Cooperativa en Cáncer from the Instituto de Salud Carlos III (RD06/0020/1001 to O.L. and contract to R.M.). O.L. is additionally supported by the Human Frontiers Science Program (RGP39/2008 to O.L.), the Fundación Ramón Areces and the Government from the Autonomous Region of Madrid (S2010-BMD-2316). This work was also supported by the Max Planck Gesellschaft, the Sonderforschungsbereich SFB646, the Gottfried Wilhelm Leibniz Program of the Deutsche Forschungsgemeinschaft and the Center for Integrated Protein Science Munich (E.C.).

AUTHOR CONTRIBUTIONS

R.M., R.C., D.G. and M.L. carried out EM. R.M. did all the image processing. R.C. purified and analyzed UPF2–UPF3 complexes. M.R. and H.U. did MS experiments. G.B. purified all proteins and complexes used, cross-linked complexes and analyzed data. H.U., E.C. and O.L. designed experiments, analyzed results and wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/nsmb/.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Garneau, N.L., Wilusz, J. & Wilusz, C.J. The highways and byways of mRNA decay. *Nat. Rev. Mol. Cell Biol.* **8**, 113–126 (2007).
2. Houseley, J. & Tollervey, D. The many pathways of RNA degradation. *Cell* **136**, 763–776 (2009).

3. Chang, Y.F., Imam, J.S. & Wilkinson, M.F. The nonsense-mediated decay RNA surveillance pathway. *Annu. Rev. Biochem.* **76**, 51–74 (2007).
4. Isken, O. & Maquat, L.E. The multiple lives of NMD factors: balancing roles in gene and genome regulation. *Nat. Rev. Genet.* **9**, 699–712 (2008).
5. Holbrook, J.A., Neu-Yilik, G., Hentze, M.W. & Kulozik, A.E. Nonsense-mediated decay approaches the clinic. *Nat. Genet.* **36**, 801–808 (2004).
6. McGlincy, N.J. & Smith, C.W. Alternative splicing resulting in nonsense-mediated mRNA decay: what is the meaning of nonsense? *Trends Biochem. Sci.* **33**, 385–393 (2008).
7. Rehwinkel, J., Raes, J. & Izaurralde, E. Nonsense-mediated mRNA decay: Target genes and functional diversification of effectors. *Trends Biochem. Sci.* **31**, 639–646 (2006).
8. Culbertson, M.R. & Leeds, P.F. Looking at mRNA decay pathways through the window of molecular evolution. *Curr. Opin. Genet. Dev.* **13**, 207–214 (2003).
9. Singh, G., Rebbapragada, I. & Lykke-Andersen, J. A competition between stimulators and antagonists of Upf complex recruitment governs human nonsense-mediated mRNA decay. *PLoS Biol.* **6**, e111 (2008).
10. Rebbapragada, I. & Lykke-Andersen, J. Execution of nonsense-mediated mRNA decay: what defines a substrate? *Curr. Opin. Cell Biol.* **21**, 394–402 (2009).
11. Amrani, N. *et al.* A faux 3′-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay. *Nature* **432**, 112–118 (2004).
12. Behm-Ansmant, I., Gatfield, D., Rehwinkel, J., Hilgers, V. & Izaurralde, E. A conserved role for cytoplasmic poly(A)-binding protein 1 (PABPC1) in nonsense-mediated mRNA decay. *EMBO J.* **26**, 1591–1601 (2007).
13. Le Hir, H., Gatfield, D., Izaurralde, E. & Moore, M.J. The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. *EMBO J.* **20**, 4987–4997 (2001).
14. Saulière, J. *et al.* The exon junction complex differentially marks spliced junctions. *Nat. Struct. Mol. Biol.* **17**, 1269–1271 (2010).
15. Czaplinski, K. *et al.* The surveillance complex interacts with the translation release factors to enhance termination and degrade aberrant mRNAs. *Genes Dev.* **12**, 1665–1677 (1998).
16. Kashima, I. *et al.* Binding of a novel SMG-1–Upf1–eRF1–eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay. *Genes Dev.* **20**, 355–367 (2006).
17. Ivanov, P.V., Gehring, N.H., Kunz, J.B., Hentze, M.W. & Kulozik, A.E. Interactions between UPF1, eRFs, PABP and the exon junction complex suggest an integrated model for mammalian NMD pathways. *EMBO J.* **27**, 736–747 (2008).
18. Sun, X., Perlick, H.A., Dietz, H.C. & Maquat, L.E. A mutated human homologue to yeast Upf1 protein has a dominant-negative effect on the decay of nonsense-containing mRNAs in mammalian cells. *Proc. Natl. Acad. Sci. USA* **95**, 10009–10014 (1998).
19. Weng, Y., Czaplinski, K. & Peltz, S.W. ATP is a cofactor of the Upf1 protein that modulates its translation termination and RNA binding activities. *RNA* **4**, 205–214 (1998).
20. Chamieh, H., Ballut, L., Bonneau, F. & Le Hir, H. NMD factors UPF2 and UPF3 bridge UPF1 to the exon junction complex and stimulate its RNA helicase activity. *Nat. Struct. Mol. Biol.* **15**, 85–93 (2008).
21. Chakrabarti, S. *et al.* Molecular mechanisms for the RNA-dependent ATPase activity of Upf1 and its regulation by Upf2. *Mol. Cell* **41**, 693–703 (2011).
22. Gehring, N.H., Neu-Yilik, G., Schell, T., Hentze, M.W. & Kulozik, A.E. Y14 and hUpf3b form an NMD-activating complex. *Mol. Cell* **11**, 939–949 (2003).
23. Buchwald, G. *et al.* Insights into the recruitment of the NMD machinery from the crystal structure of a core EJC–UPF3b complex. *Proc. Natl. Acad. Sci. USA* **107**, 10050–10055 (2010).
24. Yamashita, A., Ohnishi, T., Kashima, I., Taya, Y. & Ohno, S. Human SMG-1, a novel phosphatidylinositol 3-kinase-related protein kinase, associates with components of the mRNA surveillance complex and is involved in the regulation of nonsense-mediated mRNA decay. *Genes Dev.* **15**, 2215–2228 (2001).
25. Yamashita, A. *et al.* SMG-8 and SMG-9, two novel subunits of the SMG-1 complex, regulate remodeling of the mRNA surveillance complex during nonsense-mediated mRNA decay. *Genes Dev.* **23**, 1091–1105 (2009).
26. Arias-Palomo, E. *et al.* The nonsense-mediated mRNA decay SMG-1 kinase is regulated by large-scale conformational changes controlled by SMG-8. *Genes Dev.* **25**, 153–164 (2011).
27. Okada-Katsuhata, Y. *et al.* N- and C-terminal Upf1 phosphorylations create binding platforms for SMG-6 and SMG-5:SMG-7 during NMD. *Nucleic Acids Res.* **40**, 1251–1266 (2012).
28. Isken, O. *et al.* Upf1 phosphorylation triggers translational repression during nonsense-mediated mRNA decay. *Cell* **133**, 314–327 (2008).
29. Fukuhara, N. *et al.* SMG7 is a 14–3-3-like adaptor in the nonsense-mediated mRNA decay pathway. *Mol. Cell* **17**, 537–547 (2005).
30. Ishigaki, Y., Li, X., Serin, G. & Maquat, L.E. Evidence for a pioneer round of mRNA translation: mRNAs subject to nonsense-mediated decay in mammalian cells are bound by CBP80 and CBP20. *Cell* **106**, 607–617 (2001).
31. Hogg, J.R. & Goff, S.P. Upf1 senses 3′ UTR length to potentiate mRNA decay. *Cell* **143**, 379–389 (2010).
32. Gehring, N.H. *et al.* Exon-junction complex components specify distinct routes of nonsense-mediated mRNA decay with differential cofactor requirements. *Mol. Cell* **20**, 65–75 (2005).
33. Franks, T.M., Singh, G. & Lykke-Andersen, J. Upf1 ATPase-dependent mRNP disassembly is required for completion of nonsense- mediated mRNA decay. *Cell* **143**, 938–950 (2010).
34. Andersen, C.B. *et al.* Structure of the exon junction core complex with a trapped DEAD-box ATPase bound to RNA. *Science* **313**, 1968–1972 (2006).
35. Bono, F., Ebert, J., Lorentzen, E. & Conti, E. The crystal structure of the exon junction complex reveals how it maintains a stable grip on mRNA. *Cell* **126**, 713–725 (2006).
36. Kadlec, J., Izaurralde, E. & Cusack, S. The structural basis for the interaction between nonsense-mediated mRNA decay factors UPF2 and UPF3. *Nat. Struct. Mol. Biol.* **11**, 330–337 (2004).
37. Cheng, Z., Muhlrad, D., Lim, M.K., Parker, R. & Song, H. Structural and functional insights into the human Upf1 helicase core. *EMBO J.* **26**, 253–264 (2007).
38. Clerici, M. *et al.* Unusual bipartite mode of interaction between the nonsense-mediated decay factors, UPF1 and UPF2. *EMBO J.* **28**, 2293–2306 (2009).
39. Serin, G., Gersappe, A., Black, J.D., Aronoff, R. & Maquat, L.E. Identification and characterization of human orthologues to *Saccharomyces cerevisiae* Upf2 protein and Upf3 protein (*Caenorhabditis elegans* SMG-4). *Mol. Cell. Biol.* **21**, 209–223 (2001).
40. Ballut, L. *et al.* The exon junction core complex is locked onto RNA by inhibition of eIF4AIII ATPase activity. *Nat. Struct. Mol. Biol.* **12**, 861–869 (2005).
41. Stark, H. GraFix: stabilization of fragile macromolecular complexes for single particle cryo-EM. *Methods Enzymol.* **481**, 109–126 (2010).
42. Radermacher, M., Wagenknecht, T., Verschoor, A. & Frank, J. Three-dimensional reconstruction from a single-exposure, random conical tilt series applied to the 50S ribosomal subunit of *Escherichia coli*. *J. Microsc.* **146**, 113–136 (1987).
43. Richter, F.M., Sander, B., Golas, M.M., Stark, H. & Urlaub, H. Merging molecular electron microscopy and mass spectrometry by carbon film-assisted endoproteinase digestion. *Mol. Cell. Proteomics* **9**, 1729–1741 (2010).
44. Stroupe, M.E., Tange, T.O., Thomas, D.R., Moore, M.J. & Grigorieff, N. The three-dimensional architecture of the EJC core. *J. Mol. Biol.* **360**, 743–749 (2006).
45. Wang, H.W. *et al.* Structural insights into RNA processing by the human RISC-loading complex. *Nat. Struct. Mol. Biol.* **16**, 1148–1153 (2009).
46. Yusupova, G.Z., Yusupov, M.M., Cate, J.H. & Noller, H.F. The path of messenger RNA through the ribosome. *Cell* **106**, 233–241 (2001).
47. Jinek, M., Coyle, S.M. & Doudna, J.A. Coupled 5′ nucleotide recognition and processivity in Xrn1-mediated mRNA decay. *Mol. Cell* **41**, 600–608 (2011).
48. Gehring, N.H., Lamprinaki, S., Kulozik, A.E. & Hentze, M.W. Disassembly of exon junction complexes by PYM. *Cell* **137**, 536–548 (2009).
49. Ghosh, S., Marchand, V., Gaspar, I. & Ephrussi, A. Control of RNP motility and localization by a splicing-dependent structure in oskar mRNA. *Nat. Struct. Mol. Biol.* **19**, 441–449 (2012).
50. Hyre, D.E. *et al.* Cooperative hydrogen bond interactions in the streptavidin-biotin system. *Protein Sci.* **15**, 459–467 (2006).

ONLINE METHODS

Purification of proteins and protein complexes. The EJC complex was reconstituted as described[35]. Human UPF3 (full-length UPF3b and residues 42–143), full-length UPF2 (residues 1–1272) and human UPF1 (residues 115–914) were expressed as recombinant His-tagged (tobacco etch virus–cleavable) proteins using *Escherichia coli* BL21-Gold (DE3) pLysS cells (Stratagene), as described[21,23]. Cells grown in TB medium were induced for 12–15 h at 16 °C with 500 μM IPTG. Cells were lysed in 40 mM Tris-HCl, pH 7.5, 200 mM NaCl, 10% (v/v) glycerol, 20 mM imidazole and 5 mM β-mercaptoethanol supplemented with protease inhibitors (Roche). In case of UPF1 cell lysis was carried out in 40 mM Tris-HCl, pH 7.5, 200 mM NaCl, 10% (v/v) glycerol, 20 mM imidazole, 1 mM $MgCl_2$, 1 μM $ZnCl_2$ and 5 mM β-mercaptoethanol. The His-tagged proteins were purified on nickel resin. After cleaving of the His tag with tobacco etch virus protease during dialysis overnight in lysis buffer, the proteins were passed again over a Nickel column to separate the cleaved from uncleaved proteins and from the His-tag. UPF1, UPF2 and $UPF3_{42–143}$ were further purified on a Heparin Sepharose CL-4B column (GE Healthcare) followed by ion-exchange chromatography (Mono S, pH 7.0, in case of UPF1; Mono Q, pH 7.5, in case of UPF2). The proteins were further purified by gel-filtration chromatography (Superdex S200 GE Healthcare, UPF1 and UPF2; Superdex S75 GE Healthcare, $UPF3_{42–143}$).

The UPF1–UPF2–UPF3–EJC complex was formed by incubating purified proteins in 40 mM Tris-HCl, pH 7.5, 300 mM NaCl, 3 mM $MgCl_2$, 1 mM DTT (gel-filtration buffer) using a ~1.3-fold molar excess of UPF3, UPF2 and UPF1. The complex solution was purified by size-exclusion chromatography (Superdex S200, GE Healthcare). The EJC–UPF3–UPF2 complex was purified using the same conditions as described above. The $UPF3_{42–143}$–UPF2–UPF1 complex was formed under similar conditions by incubating purified proteins in 40 mM Tris-HCl, pH 7.5, 150 mM NaCl, 10% (v/v) glycerol, 1 mM $MgCl_2$, 1 μM $ZnCl_2$, 1 mM DTT (gel-filtration buffer) followed by size-exclusion chromatography (Superdex S200, GE Healthcare). In all cases, fractions were analyzed on a 4–12% Bis-Tris precast gradient gel (NuPage, Invitrogen) and were stained with Coomassie blue.

Purification of streptavidin labeled complexes. The EJC complex was reconstituted in the presence of 2 mM AMPPNP and a 1.5× molar excess of either 5′ or 3′ biotinylated U_{15} RNA (Dharmacon). The complex solution was purified by gel filtration (Superdex S200 GE Healthcare). The UPF1–UPF2–UPF3–EJC–streptavidin complex was formed by incubating purified proteins for 30 min using a 1.8× molar excess of streptavidin (AnaSpec) and ~1.3× molar excess of UPF proteins. The complex solution was cross-linked and purified with the same GraFix procedure used for the untagged complex.

GraFix. For EM the purified complexes were separated and mildly fixed according to the GraFix protocol[41] using a gradient solution of 30% (v/v) sucrose and 0.2% (v/v) glutaraldehyde in buffer containing 50 mM K-phosphate, 150 mM NaCl and 3 mM $MgCl_2$.

Western blots. Fractions were separated on a 4–12% Bis-Tris precast gradient gel (NuPage, Invitrogen), transferred to a nitrocellulose membrane and probed with antibodies to eIF4AIII (abcam), hUPF3 (abcam), hUPF2 (abcam), hUPF1 (abcam) or streptavidin (abcam).

Nanoliter liquid chromatography–mass spectrometry–mass spectrometry. Endoproteolytic and nuclease digestion of complexes were carried out as described[43]. We detected peptides resulting from cleavage of lysine residues that had not been modified by the cross-linker. The comparison of the protection pattern with and without different UPF–EJC components shows areas of contact between components of the complex. Details are in **Supplementary Note**.

Electron microscopy and 3D reconstruction using negative staining. Samples of UPF1–UPF2–UPF3–EJC, UPF2–UPF3–EJC and UPF1–UPF2–$UPF3_{RRM}$ were negatively stained with 2% (m/w) uranyl formate and observed in a JEOL 1230 transmission electron operated at 100 kV and a final magnification of ×68,222.5 (**Supplementary Fig. 2**). We automatically collected 10,458, 9,161, and 8,500 particles for UPF1–UPF2–UPF3–EJC, UPF2–UPF3–EJC and UPF1–UPF2–$UPF3_{RRM}$, respectively, using a 4k × 4k TVIPS CMOS detector under control of EM-TOOLS software for automatic single-particle microscopy (TVIPS). Contrast transfer function for each micrograph was estimated using CTFFIND3 (ref. 51) and corrected using BSOFT[52] before boxing particles. All particle images were initially classified using reference-free methods in EMAN[53] and XMIPP[54] after binning at 4.56 Å per pixel. An *ab initio* structure of the UPF1–UPF2–UPF3–EJC complex was obtained by the RCT method[42] using XMIPP[54], which was then used as template for angular refinement (**Supplementary Fig. 3**). For the RCT reconstruction, 4,483 tilt pairs were collected using 50° tilting of the specimen holder. After classification of the untilted images into highly homogeneous subgroups, the best five classes, containing between 175 and 250 images each, were selected to obtain five RCT structures, all of which were very similar. One of these RCT structures was selected as starting template for refinement (**Supplementary Fig. 3**).

The refined structure of UPF1–UPF2–UPF3–EJC after filtering at 50 Å was used as reference for processing the UPF2–UPF3–EJC and UPF1–UPF2–$UPF3_{RRM}$ data. All angular refinements were done using EMAN[53]. As a control, all data sets were also processed from a featureless Gaussian blob with dimensions estimated from the reference-free averages obtained using makeinitialmodel.py in EMAN[53]. The data processed from these blobs converged to similar structures (**Supplementary Fig. 3**). The resolution was estimated using Fourier shell correlation (FSC) and a correlation coefficient of 0.5, as 32 Å, 30 Å and 35 Å for UPF1–UPF2–UPF3–EJC, UPF2–UPF3–EJC and UPF1–UPF2–$UPF3_{RRM}$, respectively (**Supplementary Fig. 3**).

UPF2 protein freshly collected after gel-filtration chromatography was analyzed in the electron microscope. An *ab initio* structure of UPF2 was reconstructed using a RCT scheme[42] from 2,993 pairs of particles obtained at 0° and 45° using XMIPP[54]. The RCT structure was refined using a data set containing a mixture of 3,052 untilted images and 2,993 images tilted at 45°, using EMAN[53]. The final resolution was estimated as 28.6 Å at FSC cutoff of 0.5.

The UPF2–UPF3 complex was purified by mixing UPF2 with an excess of UPF3 and size-exclusion chromatography in a Superdex S200 (GE Healthcare). A fraction containing both proteins was analyzed as described above. Some 5,649 particles for the UPF2–UPF3 complex were processed using EMAN[53]. The structure of UPF2 and a model obtained from reference-free averages of UPF2–UPF3 and the initial model generator in EMAN2[55] were used as starting references for refinement, and both converged to the same structure. The resolution of the final structure was estimated as 30.3 Å at FSC cutoff of 0.5.

Cryo-electron microscopy and image processing. The UPF1–UPF2–UPF3–EJC complex was vitrified using QUANTIFOIL R 1.2:1.3 holey grids and a Vitrobot (Gatan). The grids were observed in a JEM-2200FS FEG electron microscope at ×69,494 magnification and using an energy filter to increase the signal-to-noise ratio. Data sets were collected in a 4k × 4k Gatan CCD and >85,000 cryo-EM particles were extracted using XMIPP[54] and SPIDER[56] at a final 2.2 Å per pixel. Contrast transfer function correction was done using CTFFIND3 (ref. 51) and BSOFT[52]. Probably owing to the relatively low molecular mass of the UPF1–UPF2–UPF3–EJC complex, only micrographs with defocus values between −2.0 and −5.5 μm showed images of UPF1–UPF2–UPF3–EJC molecules that could be satisfactorily detected over the noisy background. Thus, only particles from these micrographs, showing a good signal-to-noise ratio, were used for subsequent analysis. These images were refined using the structure of UPF1–UPF2–UPF3–EJC from negative stain as starting reference (**Supplementary Fig. 3**). Processing was done in parallel in SPIDER[56], XMIPP[54] and EMAN[53], and we obtained identical results. The hand of the reconstruction was defined using the RCT method and a control experiment with a protein of known handedness defined by X-ray crystallography. The resolution was estimated as 16 Å using FSC (cut-off, correlation coefficient of 0.5). Fitting of atomic structures into the cryo-EM density was carried out using UCSF Chimera[57].

Image processing of streptavidin-labeled UPF1–UPF2–UPF3–EJC complexes. Some 3,291 images of the 3′ streptavidin–labeled and 4,700 of the 5′ streptavidin–labeled UPF1–UPF2–UPF3–EJC complexes were obtained using the methods described above. Reference-free averages were obtained using XMIPP[54], which showed a clear high-density spot near the EU321 complex. These averages were used to obtain an initial structure using common lines and these were refined in EMAN[53] to a resolution of 40 Å and 39 Å for the 3′- and 5′-labeled complexes, respectively.

Antibody labeling of the UPF–EJC complex. UPF–EJC was mixed with a 3× molar excess of each of the antibodies used in **Figure 1d**: antibody to

doi:10.1038/nsmb.2287

hUPF1 (abcam10510), recognizing residues 250–300 of UPF1; antibody to hUPF2 (abcam28712), recognizing residues 100–200 of UPF2; antibody to UPF3 (ab66753), recognizing residues 1–50 of UPF3; and antibody to eIF4AIII (abcam32485), recognizing residues 1–100 of eIF4AIII.

Homology modeling. Two homology models for $MIF4G_1$ and $MIF4G_2$ were obtained using the PHYRE2 server[58].

51. Mindell, J.A. & Grigorieff, N. Accurate determination of local defocus and specimen tilt in electron microscopy. *J. Struct. Biol.* **142**, 334–347 (2003).
52. Heymann, J.B. & Belnap, D.M. Bsoft: image processing and molecular modeling for electron microscopy. *J. Struct. Biol.* **157**, 3–18 (2007).
53. Ludtke, S.J. 3-D structures of macromolecules using single-particle analysis in EMAN. *Methods Mol. Biol.* **673**, 157–173 (2010).
54. Scheres, S.H., Nunez-Ramirez, R., Sorzano, C.O., Carazo, J.M. & Marabini, R. Image processing for electron microscopy single-particle analysis using XMIPP. *Nat. Protoc.* **3**, 977–990 (2008).
55. Tang, G. *et al*. EMAN2: an extensible image processing suite for electron microscopy. *J. Struct. Biol.* **157**, 38–46 (2007).
56. Shaikh, T.R. *et al*. SPIDER image processing for single-particle reconstruction of biological macromolecules from electron micrographs. *Nat. Protoc.* **3**, 1941–1974 (2008).
57. Goddard, T.D., Huang, C.C. & Ferrin, T.E. Visualizing density maps with UCSF Chimera. *J. Struct. Biol.* **157**, 281–287 (2007).
58. Kelley, L.A. & Sternberg, M.J. Protein structure prediction on the Web: a case study using the Phyre server. *Nat. Protoc.* **4**, 363–371 (2009).