Structural insights into nonsense-mediated mRNA decay

(NMD) by electron microscopy

Oscar Llorca

1 Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas (CSIC), Ramiro de Maetzu 9, 28040 Madrid, Spain

Running Title: EM of NMD complexes

* Corresponding author: Tel: +34 91 837 31 12 ext 4446; Fax: +34 91 536 04 32; E-mail: ollorca@cib.csic.es

Key words: NMD, mRNA decay, SMG1, PIKK, UPF1, UPF2, UPF3, EJC, cryoEM
ABSTRACT

Nonsense-mediated mRNA decay (NMD) is a pathway that detects and degrades mRNAs containing premature translation termination codons (PTCs). In humans, recognition of these aberrant mRNAs requires an exon-junction-complex (EJC) placed downstream of a PTC and the dynamic interaction of several UPF/SMG proteins, the ribosome and the EJC. These interactions promote UPF1 phosphorylation by SMG1 kinase, triggering mRNA degradation. Recent progress in our understanding of NMD has been achieved by combining cryo-electron microscopy, high-resolution structures, interaction data and functional assays. These studies uncovered a mechanism regulating SMG1 kinase as well as the architecture and functional implications of a complex containing UPF1, UPF2, UPF3 and EJC. Collectively these findings reveal the role of protein scaffolds and conformational changes in NMD regulation.

Large macromolecular complexes and transient interactions regulate NMD

Nonsense-mediated mRNA decay (NMD) is an evolutionary conserved eukaryotic post-transcriptional surveillance pathway that discriminates normal versus aberrant translation termination events in mRNAs containing premature translation termination codons (PTCs) [1]. PTCs result from germ-line mutations in a large number of inherited genetic disorders as well as errors during transcription and splicing. NMD detects PTC-containing mRNAs during the pioneer round of translation, and these are targeted for
degradation to prevent the synthesis of truncated proteins. NMD also regulates the
steady-state level of a wide range of physiological substrates [2].

Work over the years has provided significant insights into the molecular mechanisms of
aberrant mRNA detection in the NMD pathway [1,3-7]. NMD requires translating
ribosomes and a signal downstream of the stop codon to define the context of
translation termination as physiological or aberrant. The mechanistic details and
specificities in each organism are a matter of debate and intense research [8-16]. In
humans, the exon-junction-complex (EJC), four proteins assembled onto mRNA at a
few nucleotides upstream of the exon–exon boundary upon splicing, is a major
downstream element promoting NMD (Figure 1a). NMD is elicited when termination
of translation occurs prematurely at least 50–55 nucleotides upstream of a 3’ exon–exon
junction [1,3]. Despite intensive research, we do not fully understanding the molecular
basis of this 50-55 nucleotide rule. According to current models, a large complex
bridging the ribosome and an EJC, and containing the three UP-Framesshift proteins,
UPF1, UPF2 and UPF3, assembles only when a downstream EJC comes into proximity
of a ribosome that has terminated prematurely (Figure 1a) [1,3-7,17]. UPF1 is an RNA
helicase that associates with ribosomes by interacting with eukaryotic release factors
(eRFs) eRF1 and eRF3. UPF1 can interact also with UPF2. UPF2 can bind UPF3,
which is anchored to the EJC. UPF3 is found as two homologous proteins UPF3a and
UPF3b in human cells. Upon formation of this large complex, SMG1 (Suppressor with
Morphogenetic effect on Genitalia 1), a ~ 430 kDa serine-threonin kinase belonging to
the phosphatidylinerinositol 3-kinase-related protein kinase (PIKK) family [18],
phosphorylates UPF1 [19]. Phospho-UPF1 determines the fate of the mRNA towards
degradation [20].
A molecular view of NMD regulation provided by EM

Despite recent advances in the field, our understanding of the molecular basis of NMD is incomplete by the lack of a detailed structural view of the multiple interactions regulating this pathway. Most of our structural knowledge has been obtained by X-ray crystallography of UPF proteins and partial complexes (Figure 1) [21, 22••, 23, 24, 27]. Elena Conti and colleagues (Max Planck Institute, Martinsried) recently solved several crystal structures of UPF1 in a closed conformation [22••] (Figure 1b), which were compared with the structure UPF1 bound to the C-terminal domain of UPF2 solved by Cusack and colleagues (Max Planck Institute, Grenoble) [23] (Figure 1c). These structures revealed that a large conformational change relocates the cysteine-histidine-rich (CH) domain of UPF1, and the helicase activity is activated. UPF2 interacts with an N-terminal region of UPF3 and the structure of a complex containing the interacting domains of each protein has also been solved [27] (Figure 1d). A C-terminal peptide of UPF3 is sufficient to provide the link between UPF3 and the EJC as revealed by the structure of a UPF3 fragment bound to the EJC [24] (Figure 1e).

A full mechanistic understanding of the NMD pathway will require defining the structures of large and transient macromolecular complexes by integrating different methodologies. X-ray crystallography provided the essential high-resolution information of those most stable arrangements and sub-complexes. 3D-electron microscopy (3D-EM), particularly EM performed on frozen hydrated samples at liquid nitrogen temperatures (cryo-electron microscopy, cryoEM), is a powerful method to solve larger complexes containing several NMD factors as well as their complexes with nucleic acids, especially when these complexes are labile and/or showing flexible
conformations. Interpretation of the cryoEM densities could be typically performed by fitting atomic structures using positional restraints provided by complementary biochemical and biophysical methods [28●].

**EM combined with functional assays reveals SMG1 kinase regulation by SMG8**

SMG1 triggers NMD by phosphorylating the UPF1 helicase, but little is known about the mechanisms by which this kinase activity is confined to the context of an aberrant mRNA [29]. In 2009, Shigeo Ohno, Akio Yamashita and colleagues at Yokohama University described two proteins, SMG8 and SMG9, that co-purified with SMG1 and which were essential for NMD [30]. In collaboration with Ohno and Yamashita, we have described the overall architecture of SMG1 and the modulation of its kinase activity by the interaction with SMG8 and SMG9 [31,32●●]. SMG1 comprises several conserved domains at the C-terminus, including the catalytic PI3K kinase domain, and a long stretch of multiple helical repeats, mostly HEAT (huntingtin, elongation factor 3, A subunit of PP2A and TOR1) repeats, organized as super-helical structures at the N-terminus [19]. SMG1 and the SMG1-SMG9 complex appeared in EM as “question mark” shaped-molecules with a large “head” region and a protruding “arm” (Figure 2a). SMG1 C-terminal end was located at the “head” region, which showed similarities with the C-terminus of DNA-PKcs (DNA-dependent Protein Kinase catalytic subunit), a member of the PIKK family that has been extensively characterized [33]. SMG1 N-terminal end was arranged as an “arm” extending from the catalytic C-terminus. A combination of EM, mapping experiments and functional assays revealed that this helical N-terminal domain interacted with SMG9, which in turn recruited SMG8 [32●●].
SMG8 was identified by comparing images of SMG1-SMG8-SMG9 and SMG1-SMG9 (Figure 2b) [32••]. Therefore, SMG1 helical region functions as a scaffold for the assembly of a larger complex, a property described for other PIKKs [33] as well as other proteins containing HEAT repeats [34,35]. Interestingly, SMG8 binding down-regulated SMG1 despite interacting far from the kinase domain. Large-scale conformational changes were observed in the N- and C-terminal regions of SMG1-SMG9 after SMG8 recruitment, which we propose down-regulate the kinase (Figure 2c). This model is consistent with the complex regulation of PI3-kinases [18] and the large conformational changes described for HEAT repeats proteins interacting with other proteins [34,35].

These results provide some insights into SMG1 regulation but new questions arise. What steps remove SMG8 inhibition to facilitate UPF1 phosphorylation? SMG1 activation could be achieved by disrupting the SMG1-SMG8-SMG9 complex and/or additional conformational changes induced by other factors or proteins binding to SMG1. The answer maybe found by elucidating how this inhibitory mechanism is coordinated with the rest of events that take place during recognition of an aberrant mRNA. UPF2 binds a C-terminal fragment of SMG1, adding a potential new level of regulation [36].

**Architecture of the UPF-EJC complex resolved by cryoEM**

A recent study has provided a pseudo-atomic model of a complex containing UPF1, UPF2, UPF3 and EJC (UPF-EJC hereafter) by combining a cryoEM structure of the complex with partial crystal structures, and integrating positional restraints obtained by
labeling experiments of the components [37••] (Figure 3a). UPF-EJC was reconstituted from the individual components by Elena Conti and colleagues (Max Planck Institute, Martinsried) and analyzed by our group using EM. Interestingly, UPF-EJC was only stable at high concentrations, whereas it disassembled at the concentrations required for EM analysis. Therefore, the complex was stabilized with a mild fixation coupled to purification in a sucrose gradient using the GraFix method [38]. We determined the 3D structure of this 390 kDa complex at ~16 Å resolution using cryoEM. The positions of the individual components were mapped by antibody labeling, the structural characterization of several sub-complexes (UPF1-UPF2-UPF3-EJC, UPF2-UPF3-EJC, UPF1-UPF2-UPF3, UPF2-UPF3 and UPF2) by EM and the fitting of available crystal structures and homology models. The pseudo-atomic model suggested that the requirement for a minimum of 50-55 nucleotides between a stop codon and a downstream splice junction might be interpreted in molecular terms as physical and geometrical constraints. A minimal distance of several tens of nucleotides between the ribosome and the exon-exon junction could be required to accommodate UPF-EJC, considering the distance of an RNA-bound EJC to the exon-exon junction, the RNA covered by the EJC itself and the distance between the EJC and the RNA-binding region of UPF1 in the context of the UPF-EJC [37••].

UPF1 is part of SURF (SMG1-Upf1-Release-Factors), a probably transient NMD complex assembled at the ribosome that contains SMG1, UPF1 and release factors eRF1 and eRF3 [36]. UPF1 would be expected to locate at the 5’end of the EJC bound RNA as part of SURF (Figure 1a). We assessed whether UPF1 was placed towards the 5’ or the 3’ end of the EJC by reconstituting UPF-EJC with an EJC bound to RNAs labeled with streptavidin at each end. Mechanistically, this allowed addressing that
UPF1 was positioned towards the substrate of its helicase activity (the RNA 3’ end) in the context of UPF-EJC rather than towards the ribosome (Figure 3b). Remarkably, recent findings have revealed that UPF1 ATPase activity in the 5’ to 3’ direction is essential to strip proteins and dismantle the downstream mRNP during NMD [39]. This suggested that cryoEM had trapped a transient NMD complex at a stage when UPF1 is ready to function as a helicase towards the 3’ end of the mRNP. How is UPF1 that is initially in complex with SMG1 and the ribosome as part of SURF relocated to the RNA 3’ end at a later stage of the NMD reaction? This transition might be part of the dynamics connecting early and late events during recognition of aberrant mRNAs with short-lived transient complexes regulating SMG1 activation and UPF1 phosphorylation.

EM and mass spectrometry show that UPF2 is an open ring-like scaffold coordinating the assembly of the surveillance complex

The UPF-EJC structure revealed that the crosstalk between UPF1, UPF2 UPF3 and the EJC involves a set of interactions with UPF2 as a core scaffold as well as multiple conformational changes [37••]. UPF2 appears as a distorted open-ring and its two ends, corresponding to the N- or C-terminal regions of UPF2, grab and reposition the CH domain of UPF1, inducing an open unwinding-competent conformation. UPF1 helicase regions stick out in UPF-EJC, being accessible to RNA binding (Figure 3b). Other regions of the UPF2 ring interact with UPF3. Interestingly, EM images of UPF2 alone showed open rings with certain flexibility, suggesting conformational changes in the ring as it closes around UPF1. These were confirmed by a mass spectrometry (MS) analysis performed by Henning Urlaub and colleagues (Max Planck Institute, Göttingen) [37••,40]. UPF1-UPF2-UPF3-EJC, UPF2-UPF3-EJC and UPF1-UPF2-
UPF3 complexes, after stabilization by glutaraldehyde cross-linking, were treated with trypsin. Detected peptides were those that had not been modified by the crosslinker either because they were the core of the proteins or because they corresponded to inaccessible regions resulting from protein-protein interactions within UPF-EJC. The comparison of peptides obtained for UPF-EJC with those from complexes lacking either UPF1 or EJC confirmed the EM model by revealing that the N- and C-terminal regions of UPF2 interact with UPF1. Interestingly, these experiments also revealed conformational changes in UPF2 when the ring closes around UPF1. Some regions in UPF2 were accessible to solvent or protected from crosslinking by the EJC when either bound or not to UPF1 respectively.

Conclusions and outlook

The crosstalk between the stalled ribosome and the EJC involves a number of SMG/UPF proteins that interact, either sequentially or concomitantly, ensuing mRNA degradation. NMD factors interact forming several complexes containing different subunits and conformations. Many of these interactions will be labile and short-lived. Recent EM studies integrating structures resolved by X-ray crystallography and information provided by labeling experiments, functional assays and mass spectrometry have provided important insights into the complexity of NMD [32••, 37••]. UPF1 binds to the ribosome as part of a complex with released factors and SMG1 (SURF) when the ribosome encounters a stop codon (Figure 1a). SMG1 kinase activity is down regulated by SMG8 and SMG9 (Figure 2). Only when translation termination is premature, interactions linking SURF components to a downstream EJC can be established. As part of this large complex, UPF2 forms an open ring that interacts with
UPF1 and induces an active conformation of this RNA helicase, whereas UPF3 brings UPF2 to the EJC (Figure 3). At some stage, SMG1 phosphorylates UPF1, which transits from the SURF complex to the 3’ side of the EJC, poised to strip proteins at the 3’ end of the mRNP (Figure 3). Many questions remain open. The quest for a complete understanding of the molecular basis of the 50-55 nucleotide rule is still open. At what stage does UPF1 phosphorylation take place? Does UPF1 bind mRNA as part of SURF? How is SMG8 inhibition of SMG1 removed? Can UPF1 exist simultaneously bound to components of SURF and UPF2 as part of the large surveillance complex hypothesized? Some evidence suggests that UPF1 probably cannot bind eRFs and UPF2 simultaneously and thus it may transit from one complex to another. Is UPF1 5’ to 3’ helicase activity maintained active in complex with UPF2 or could UPF1 phosphorylation maintain an activated conformation without UPF2?

The answer to these questions will probably benefit from the structural characterization of several NMD complexes, a challenge requiring a combined effort of cryoEM and X-ray crystallography. But a clear mechanistic understanding of NMD will demand in addition defining the correct sequence of interactions, incorporating the time dimension to relate the static structural views provided by cryoEM and X-ray crystallography in the context of the dynamics of NMD.

ACKNOWLEDGEMENTS

This work was funded by the Spanish Government (SAF2011-22988 to O.L.), the “Red Temática de Investigación Cooperativa en Cáncer (RTICC)” (RD06/0020/1001 to O.L.)
and the Human Frontiers Science Program (RGP39/2008 to O.L.). O.L. acknowledges also support by the “Ramón Areces” Foundation. I thank Elena Conti, Akio Yamashita and Shigeo Ohno and their colleagues for their collaboration in the NMD project and helpful discussions. I thank Roberto Melero (CIB, Madrid) and Ernesto Arias-Palomo (UC Berkeley) for his help preparing figures and helpful discussions. I also acknowledge the help of Raquel Castaño (CIB, Madrid) and the contributions and suggestions of Jasminka Boskovic (CNIO, Madrid).

FIGURE LEGENDS

Figure 1. (a) Current models for the NMD pathway propose that the EJC is a major determinant for detection of PTCs and discrimination of aberrant mRNAs in humans [1,3-7]. SMG1, UPF1 and release factors eRF1 and eRF3 interact to form the SURF complex in association with the ribosome [36]. A molecular link between SURF components and a downstream EJC is required to trigger UPF1 phosphorylation by the SMG1 kinase (indicated as an arrow). UPF2 and UPF3 required to bridge UPF1 and the EJC [17]. SMG1 kinase is regulated by SMG8 and SMG9 [30], but interactions of SMG1 and UPF2 have also been described [36]. Structural insights into the mechanism for detecting PTCs in the NMD pathway have been obtained mostly by X-ray crystallography studies. The crystal structures of the UPF1 helicase (yellow color) in closed [22] (b) and open conformations [23] (c) reveal a displacement of the CH domain (indicated with an arrow) upon UPF2 binding (blue color). The CH domain regulates the activation of UPF1 by keeping UPF1 inactive in the closed conformation. The two RecA domains of UPF1 helicase region have been labeled. A complex
comprising interacting fragments of UPF2 (blue color) and UPF3b (yellow color) [27] (d), as well as a complex between a small UPF3b peptide (green color) sufficient to recognize the EJC complex (orange and red colors) [24] (e) have also been solved.

Figure 2. (a) 2D averages of EM images obtained for SMG1 (not shown) and SMG1-SMG9 [32●●] reveal two distinct regions, which were interpreted by comparison with the crystal structure of DNA-PKcs [33] and a simulated density map of DNA-PKcs crystal structure at a resolution comparable to EM. The “head” region comprises the C-terminal domains of SMG1 including the catalytic PI3K domain. The N-terminal helical regions appear as a protruding “arm” in EM images. (b) Positioning SMG8 and SMG9 in SMG1-SMG8-SMG9 was achieved by comparison of EM images for each of the complexes [32●●]. SMG-8 was identified as a spot of density bound to the helical repeat region. (c) Model of SMG1 kinase regulation based on EM, interaction data and functional assays [31-32●●]. N-terminal helical repeats in SMG1 bind SMG9, and SMG9 recruits SMG8 to the complex. SMG8 and SMG9 could bind SMG1 sequentially or as a preformed complex. SMG8 binding induces large-scale conformational changes that correlate with a down-regulation of SMG-1 kinase.

Figure 3. (a) The pseudo-atomic model of the UPF1-UPF2-UPF3-EJC complex is shown in two different views [37●●]. The model is based on the cryoEM structure of the UPF-EJC complex, represented as a transparent density, and the fitting of the atomic structures of EJC-bound to a small fragment of UPF3b and containing a short RNA
fragment (shown in black) (PDB 2XB2) [24], of UPF1 bound to the C-terminal domain of UPF2 (PDB 2WJV) [23], of the C-terminal MIF4G3 domain of UPF2 bound to an N-terminal fragment of UPF3 (PDB 1UW4) [27] and homology models for MIF4G1 and MIF4G2 domains in UPF2 [37••]. Positioning of the atomic structures into the cryoEM map was based on protein and RNA labeling experiments as well as mass spectrometry data [37••]. (b) Cartoon model for the architecture of the UPF-EJC surveillance complex [37••]. UPF2 is a distorted open ring that binds UPF1 and UPF3. UPF3 recruits UPF1-UPF2 to the EJC. UPF2 stabilizes an open “active” conformation of UPF1 by displacing the CH domain (black arrow). The complex places the UPF1 helicase poised to remodel the 3’ end of the mRNPs, rationalizing recent findings showing that UPF1 ATPase activity is required to remodel the 3’ end of the mRNP [39]. Color codes: UPF1 (yellow), UPF2 (blue), UPF3 (green), EJC (orange and red), RNA in EJC (black).

References and recommended reading

Papers of particular interest, published within the period of review have been highlighted as:

• of special interest

•• of outstanding interest


   • This study elucidates the mechanism that regulates UPF1 helicase activity. Several crystal structures of UPF1 bound to RNA and in the presence of ATP and nucleotide analogs were solved and compared with previous work by Cusack and colleagues. UPF1 is converted from an inactive closed conformation to an active open conformation upon UPF2 binding.
   • Understanding the structural bases of complex macromolecular processes requires integration of information provided by complementary methodologies. This paper reviews recent examples of cryoEM studies that combine data from other methods.

This study defines the medium resolution architecture of the SMG1-SMG8-SMG9 complex using EM. This is the first structural study of full length SMG1. Large-scale conformational changes induced after recruitment of SMG8 are described. These regulate SMG1 kinase.


This paper describes the cryoEM structure of an NMD complex containing UPF1, UPF2, UPF3 and the EJC. A pseudo-atomic model was proposed by integrating information from cryoEM, protein and RNA labeling experiments and mass spectrometry. The model provides a 3D view of a surveillance complex with UPF1 helicase poised toward the 3’ end, as well as insights about the role of UPF2 as protein scaffold.


