Crystal structure of the UPF2-interacting domain of nonsense-mediated mRNA decay factor UPF1

JAN KADLEC, DELPHINE GUILLIGAY, RAIMOND B. RAVELLI, and STEPHEN CUSACK

European Molecular Biology Laboratory, Grenoble Outstation, BP 181, 38042 Grenoble Cedex 9, France

ABSTRACT

UPF1 is an essential eukaryotic RNA helicase that plays a key role in various mRNA degradation pathways, notably nonsense-mediated mRNA decay (NMD). In combination with UPF2 and UPF3, it forms part of the surveillance complex that detects mRNAs containing premature stop codons and triggers their degradation in all organisms studied from yeast to human. We describe the 3 Å resolution crystal structure of the highly conserved cysteine–histidine-rich domain of human UPF1 and show that it is a unique combination of three zinc-binding motifs arranged into two tandem modules related to the RING-box and U-box domains of ubiquitin ligases. This UPF1 domain interacts with UPF2, and we identified by mutational analysis residues in two distinct conserved surface regions of UPF1 that mediate this interaction. UPF1 residues we identify as important for the interaction with UPF2 are not conserved in UPF1 homologs from certain unicellular parasites that also appear to lack UPF2 in their genomes.

Keywords: nonsense-mediated mRNA decay; NMD; surveillance complex; UPF1; X-ray crystallography

INTRODUCTION

Nonsense-mediated mRNA decay (NMD) is an mRNA degradation pathway that detects and eliminates aberrant coding transcripts containing premature termination codons (PTC) originating from nonsense or frameshift mutations. The PTC-containing transcripts would otherwise be translated into truncated proteins that might have a deleterious effect on the cell (Maquat 2004).

According to the current consensus model of NMD in mammalian cells, the recognition of a PTC requires a splicing-dependent deposition of a multiprotein complex, the exon junction complex (EJC), 20–24 nucleotides (nt) upstream of a splice junction (Le Hir et al. 2001). During mRNA export to the cytoplasm, a perinuclear protein, UPF2, is recruited to the EJC via UPF3 (Lykke-Andersen et al. 2000). In mammals nonsense-mediated mRNA decay has been proposed to occur during the first “pioneer” round of translation (Ishigaki et al. 2001). With the first passage of a ribosome, the EJCs would normally be stripped

from the mRNA. If, however, translation terminates at a PTC upstream of an EJC, UPF2 associated with a downstream EJC can be bound by UPF1 that is recruited to the terminating ribosome within the so-called SURF complex, which also includes the translation release factors eRF1 and eRF3 and the UPF1 kinase, Smg1 (Kashima et al. 2006). The association of SURF and UPF2-EJC (the DECID complex) is followed by phosphorylation of UPF1 by Smg1 (Conti and Izaurralde 2005; Kashima et al. 2006). The phosphorylated form of UPF1 is then recognized by Smg7, which targets the aberrant transcript for decay in mRNA degradation foci (Unterholzner and Izaurralde 2004; Fukuhara et al. 2005).

Human UPF1 (hUPF1, also known as RENT1) consists of 1118 amino acid residues. In its N-terminal region, UPF1 has a conserved cysteine–histidine-rich (CH-rich) region (residues 123–213), while centrally it possesses the seven conserved motifs characteristic of eukaryotic group I RNA helicases (Applequist et al. 1997). UPF1 displays nucleic acid-dependent ATPase activity and ATP-dependent 5′–3′ helicase activity, which is required for its NMD function (Bhattacharya et al. 2000).

UPF1 has been assigned an increasing number of functions and consequently interacts with numerous proteins. Yeast and/or human UPF1 interacts with several NMD factors including UPF2, Smg1, Smg5, Smg6, Smg7, and HrpI (Gonzalez et al. 2000; Conti and Izaurralde 2005).

Reprint requests to: Stephen Cusack, European Molecular Biology Laboratory, Grenoble Outstation, 6 rue Jules Horowitz, BP 181, 38042 Grenoble Cedex 9, France; e-mail: cusack@embl-grenoble.fr; fax: (33) 476207199.

Article published online ahead of print. Article and publication date are at <http://www.rnajournal.org/cgi/doi/10.1261/rna.177606>.

et al. 2002) and in *Caenorhabditis elegans*, together with Smg5 and Smg7, is important for persistence of mRNA silencing by RNA interference (Domeier et al. 2000). Recently, UPF1 was shown to interact with StauI, an interaction essential for a newly identified mRNA degradation pathway called StauI-mediated decay (Kim et al. 2005) and with SLBP for the specific degradation of histone mRNA (Kaygun and Marzluff 2005). Probably as a result of its multiple important functions, UPF1 is essential for mammalian embryonic viability (Medghalchi et al. 2001).

Here we present the crystal structure of the CH-rich domain of human UPF1 at a resolution of 3 Å. Overall the domain has a unique fold coordinating three zinc atoms.

However, two substructures coordinating one and two zinc atoms, respectively, have remarkable and unexpected structural similarities to the RING domain commonly found in ubiquitin ligases. Structure-based mutagenesis of conserved residues revealed two distinct UPF1 surface regions likely to be involved in the interaction with UPF2.

RESULTS AND DISCUSSION

The crystal structure of CH-rich domain of UPF1

Expression of the putative UPF2-interacting region of hUPF1 (residues 115–245) in *Escherichia coli* yielded insoluble protein. However, a C-terminally extended fragment, spanning residues 115–272 (Fig. 1), could be expressed in a soluble form, and crystals grown. Previous measurements using the proton-induced X-ray emission (microPIXE) (Garman and Grime 2005) technique had shown that the domain contains three zinc atoms (data not shown). This enabled the atomic structure to be solved at 3 Å resolution using the anomalous scattering of the native zinc atoms (Table 1).

Residues 118–272 of hUPF1 are visible in the electron density maps, and these form a single domain with a compact fold containing three structural zinc atoms

(Fig. 2). The main feature of the structure is a central, pseudo-twofold symmetric, four-stranded antiparallel β -sheet (strands β 1– β 4) with symmetric flanking loops L1 and L6 (Fig. 2A). While L1 is involved in coordinating Zn1, L6 is stabilized by Zn3. The β -sheet packs against helix α 1 that follows β 2. L6 connects this β 1 β 2 α 1 β 3 β 4 arrangement to a three-stranded antiparallel β -sheet (strands β 5– β 7) that is in a perpendicular orientation to the first β -sheet. β 6 is followed by L8, which participates in coordination of Zn3 and a one-turn helix, α 2. The C terminus of the domain is formed by long meandering loops L10 and L11 and two helices, α 3 and α 4, that wrap around the N-terminal part of the structure (Fig. 2B). L10 packs against strands β 3 and β 4, helix α 1, and loops L5 and L3. Helices α 3 and α 4 and the intervening linker (L11) bind between loop L3, helix α 1, and loop L1 (Fig. 2). Using the DALI server (Holm and Sander 1993), no similar structures were found, indicating that overall this domain represents a previously unseen fold.

Zinc coordination within the UPF1 CH-rich domain

The three zinc atoms occur in different variants of zinc fingers. Zn1 is coordinated by three cysteines and one histidine within a treble clef zinc-finger-like motif (Fig. 2D; Krishna et al. 2003). This motif consists of a zinc knuckle (a unique turn with the consensus sequence CPXCG) providing two zinc ligands, the two others coming from a β -hairpin and C-terminal part of an α -helix. For Zn1, ligands Cys123 and Cys126 are provided by a Zn-knuckle within loop L1, Cys145 comes from the C terminus of β 2 of the following β -hairpin (β 1– β 2), and His155 from the N terminus of downstream helix α 1. Zn2, with two ligands (Cys137 and Ser140) being provided by a Zn-knuckle in a β -hairpin (loop L2), and His159 and Cys165 coming, respectively, from helix α 1, which packs against the β -hairpin and the following loop L4 (Fig. 2D). This motif is, in fact, a classical two-cysteine, two-histidine-like (C2H2) zinc finger and as shown in Figure 1B, most UPF1 homologs, indeed, have the conventional C2H2 ligands. Only in vertebrates has C2H2 evolved to become CSHC, the structure suggesting that Ser140 is actually a zinc ligand, although this cannot be certain at the current resolution. Finally, Zn3 is coordinated within a second treble-clef zinc-finger motif, with Cys183 and Cys186 provided by a Zn-knuckle (L6), Cys209 coming from the C terminus of β 6 of a following β -hairpin (β 5– β 6) and Cys213 coming from L8 (Fig. 2E). A feature of this zinc site is that it includes a *cis*-proline, Pro212. Both treble-clef motifs of UPF1 can be very well superposed on L24E, a ribosomal protein interacting with 23S RNA via the corresponding helix (Klein et al. 2004). All the zinc coordinating residues are absolutely conserved among known UPF1 homologs (Fig. 1B), suggesting that the zinc atoms are crucial for UPF1 structure and hence function.

TABLE 1. Data collection and refinement statistics

	UPF1
Data collection	
Space group	$P2_12_12_1$
Cell dimensions	
a, b, c (Å)	64.6, 73.0, 73.2
α, β, γ (°)	90.0, 90.0, 90.0
Resolution (Å)	20–2.95 (3.0–2.95) ^a
R_{sym} or R_{merge}	0.124 (0.424)
$I/\sigma I$	10.2 (2.9)
Completeness (%)	90.9 (96.5)
Redundancy	4.9 (3.4)
Refinement	
Resolution (Å)	20–2.95
No. reflections	6685 (work), 315 (test)
$R_{\text{work}}/R_{\text{free}}$	0.221/0.259
No. atoms ^b	
Protein	2458
Ligands/ions	6 zinc
Average B -factor (Å ²)	
Protein	34.5
RMS deviations	
Bond lengths (Å)	0.009
Bond angles (°)	1.23
Ramachandran plot	
Favorable regions	89.1%
Additionally allowed	10.9%

^aThe highest resolution shell is shown in parentheses.

^bTotal atom count for two molecules in the asymmetric unit.

FIGURE 2. Crystal structure of the CH-rich domain of human UPF1. (A) Ribbon diagram of the structure of hUPF1(115–272). The polypeptide chain is colored from the N terminus (blue) to the C terminus (red). The coordinated zinc atoms are in brown. This and other ribbon diagrams were generated with Pymol (<http://www.pymol.org>). (B) The same ribbon model as in A rotated 180° around vertical axes. (C) Schematic diagram of the UPF1 CH-rich domain topology. The figure was generated with TopDraw (Ohi et al. 2003). (D) A view of the N-terminal part of the CH-rich domain (residues 118–172) coordinating Zn1 and Zn2 within a treble-clef and C2H2 zinc-finger motifs, respectively. Ser140 is shown in a position where other UPF1 homologs have a fourth cysteine. (E) Coordination of Zn3 within the second treble clef zinc-finger motif (residues 181–216) of the UPF1 CH-rich domain.

Several reports aiming to characterize UPF1 function relied on mutagenesis of cysteine and histidine residues within the CH-rich region (Weng et al. 1996; de Pinto et al. 2004; Kashima et al. 2006). Most of the results can be rationalized on the basis of whether or not the mutations were of zinc ligands and hence whether they disrupted the structural integrity of the domain. For example, in yeast, individual mutations C72S, H110R (Cys133, His171 in hUPF1) had no effect on UPF1 function (Weng et al. 1996), consistent with these residues not being involved in zinc binding, whereas mutations H94R, H98R, C122S, and C125S (His155, His159, Cys183, and Cys186 in hUPF1), which are involved, respectively, in Zn1, Zn2, Zn3, and Zn3 binding, result in complete loss of UPF1 activities in NMD as well as in preventing nonsense codon read-through (Weng et al. 1996). Our structural results enable design of more rational mutations that selectively perturb putative interaction surfaces without affecting the overall domain structure and stability (see below).

The UPF1 CH-rich domain contains two modules similar to ubiquitin ligase RING or U-box domains

A search of the PDB with the protein structure comparison service SSM at the European Bioinformatics Institute

(<http://www.ebi.ac.uk/msd-srv/ssm>) revealed that the CH-rich domain of UPF1 contains two similar structural modules that both resemble the RING-box or U-box found most often in E3 ubiquitin ligases (Fig. 3; Ohi et al. 2003; Petroski and Deshaies 2005). Neither of these modules could be identified by sequence analysis, and their juxtaposition in UPF1 is unique. The RING-like module 1 of hUPF1 spans residues 121–172 (loops L1–L4, strands β 1– β 3, and helix α 1), while the RING-like module 2 includes residues 180–233 (loops L6–L9, strands β 5– β 7, and helix α 2) (Fig. 3A,B). RING-box domains usually coordinate two structural zinc atoms, while U-box domains are zinc free, the fold being instead maintained by a hydrogen-bond network (Fig. 3C,D; Zheng et al. 2000; Ohi et al. 2003; Andersen et al. 2004). UPF1 module 1 has the two zinc atoms, but compared to canonical RING-boxes, contains a lengthened loop L3 between β 2 and α 1, and a shortened loop L4 connecting α 1 and β 3 (Fig. 3E, cf. also A with C and D). Unusually, UPF1 module 2 coordinates only a single zinc atom and also has a minimal α -helix.

Additionally, both RING/U-box domains, as well as both UPF1 RING-like modules, are similar in structure to the ZZ domain of CREB-binding protein (CBP), another two-zinc binding module found in diverse proteins (Fig. 3F; Legge et al. 2004). The ZZ domain has been proposed to function in protein–protein interactions but is not well characterized.

Although there is no evidence that ubiquitination plays a role in NMD or any other function of UPF1, structural homology with RING/U-box domains can perhaps be used to define potential protein–protein interacting surfaces. The structure of the RING domain from c-Cbl complexed with the E2 ubiquitin conjugating enzyme UbcH7 serves a model for interactions between RING and U-box E3s with E2 enzymes (Zheng et al. 2000). In this structure, Ile383, Trp408, Pro417, and Phe418 of c-Cbl make a largely hydrophobic interacting surface contacting loops 1 and 2 of UbcH7 (Fig. 3C). Similar residues are exposed in the U-box domain of splicing factor Prp19 (Fig. 3D; Ohi et al. 2003). In UPF1 RING-like module 1, a corresponding surface formed by Tyr125, His159, Lys142, and Phe144 is partially occluded by interaction with residues from the C-terminal helix α 4 (Fig. 3A). On the other hand, in UPF1 module 2, it is formed by conserved Tyr184, Trp224, Val204, Val206, and

FIGURE 3. UPF1(115–272) RING-like modules are similar to RING and U-box domains of E3 ubiquitin ligases. (A) RING-like module 1 (121–172) of the UPF1 CH-rich domain. Conserved Tyr125 and His159 together with conserved Lys142 and Phe144 are shown that interact with residues of $\alpha 4$ and form a putative protein-binding surface. Tyr125 and His159 are in positions corresponding to Ile383 and Trp408 of c-Cbl (C) and Ile5 and Tyr31 of Prp19 in D. (B) RING-like module 2 (180–233) of the UPF1 CH-rich domain. A hydrophobic surface is shown formed by conserved residues Tyr184, Trp224, Val204, and Leu219. Superposing the two modules of UPF1 gives a root mean square (RMS) deviation of 1.57 Å for 38 C_{α} atoms (DALI Z-score of 4.1). (C) The RING domain (residues 381–431) of c-Cbl showing residues involved in the binding of E2 enzyme UbcH7 (Zheng et al. 2000) (PDB entry 1FBV). The figure is based on the superposition with the UPF1 RING-like module 1 with an RMS deviation of 1.95 Å for 32 C_{α} atoms (DALI Z-score of 1.5). (D) U-box domain (residues 1–56) of splicing factor Prp19 (PDB entry 1N87) and the residues of a putative E2 enzyme interaction interface (Ohi et al. 2003). The UPF1 RING-like module 1 and the Prp19 domain superpose with an RMS deviation of 1.89 Å for 32 C_{α} atoms (DALI Z-score of 3.4). (E) A superposition of C_{α} trace of UPF1 RING-like module 1 (115–172) in blue, UPF1 module 2 (180–233) in green, and the RING-box domain of c-Cbl in red, showing also the zinc atoms (the first zinc atom of c-Cbl overlaps with Zn1 of UPF1). (F) ZZ domain of CREB-binding protein (Legge et al. 2004) (PDB entry 1TOT). The UPF1 RING-like module 1 and the ZZ domain superpose with an RMS deviation of 2.03 Å for 32 C_{α} atoms (DALI Z-score of 3.6).

Leu219 and being solvent-exposed, could be available for protein–protein interactions (Fig. 3B).

Mutational analysis of UPF1–UPF2 interaction

Recently, it has been confirmed that the interaction between hUPF1 and hUPF2 is, indeed, essential for NMD. It promotes the contact between the SURF complex (Smg1, UPF1, eRF3, and eRF1) at the premature termination codon and a downstream hUPF2–EJC complex, thus triggering UPF1 phosphorylation (Kashima et al. 2006). We tested whether the hUPF1 CH-rich domain can bind hUPF2 in vitro. Indeed, when His-tagged hUPF1(115–272) is coexpressed with hUPF2(761–1207) in *E. coli*, a binary complex can be purified using a Ni^{2+} resin (Fig. 4E, lane 1). The complex is stable enough to be purified by gel filtration. Complex formation did not require the presence of CBP80, which has been recently described

as enhancing the interaction in mammalian cells (Hosoda et al. 2005). After adding to this binary complex purified hUPF3b(42–143), which binds tightly to the third MIF4G domain of UPF2 (residues 761–1054) (Kadlec et al. 2004), a stable ternary UPF complex could be formed (data not shown).

Mapping of phylogenetically conserved residues to the molecular surface of the CH-rich domain of UPF1 revealed two major concentrations, both of which are at the interface of the two β -sheets (Fig. 4). The first one overlaps with the hydrophobic surface of RING-box-like module 2 (see above) and is formed by residues of $\beta 4$, $\beta 6$, and loops L6, L9, and L10 (Fig. 4A,C). The other conserved surface is formed by residues from loops L6, L9, L10, and helix $\alpha 1$ (Fig. 4B,D).

We prepared several point mutants of solvent-exposed residues in these two conserved surfaces as well as in the putative interaction surfaces identified by similarity to c-Cbl and tested them for binding to UPF2(761–1207) in His-tag pull-down assays (Fig. 4E–G). The interaction with UPF2 was substantially reduced when Val204 was mutated to aspartate (V204D). Similar effects were observed for the single mutant V206E and double mutant E182R + Y184D (Fig. 4F, lanes 3–5). All these residues are in one conserved surface (Fig.

4A,C). Interestingly, the UPF2 binding was also affected by mutations in the other conserved patch (e.g., F192E and V161E + R162E) (Fig. 4E, lanes 4,5). Additionally, a triple mutation, encompassing both regions, V206E + V161E + R162E essentially abolished the interaction with UPF2 (Fig. 4G, lane 4). These data suggest that the full UPF2-binding activity of UPF1 involves both of these two regions, and this correlates with the previous two-hybrid analysis of the UPF1-binding site in UPF2 that was mapped to two adjacent regions in the UPF2 C terminus, separated by a flexible linker (He et al. 1996). The mutations in the putative interaction surfaces identified by similarity to c-Cbl did not affect UPF2 binding (Fig. 4E). The correct folding of the UPF1 mutants that affected UPF2 binding (V204D, V206E, E182R + Y184D, F192E, V161E + R162E, V206E + V161E + R162E, Y125E) was confirmed by size-exclusion chromatography, which showed them to behave identically to the native protein (data not shown). Moreover these experiments

FIGURE 4. Identification of UPF2 binding surfaces. (A,B) Surface representation of the domain, highlighting areas of conserved surface residues. The conservation of the surface residues, based on the first 12 sequences in the alignment shown in Figure 1B, is represented from gray to green (green is 100% conserved) according to the color scale bar. (C,D) Ribbon representation of the UPF1 CH-rich domain in the same orientation as in A and B, showing the conserved surface residues of UPF1 labeled in A and B. (E–G) His-tagged UPF1 mutants indicated above the lanes were coexpressed with UPF2 (761–1207) and purified using Ni^{2+} resin. The resin was extensively washed with 20 mM and 100 mM imidazole. Bound proteins were eluted with 250 mM imidazole and analyzed by 15% SDS-PAGE.

showed that whereas by pull-down assays these UPF1 mutants show some residual binding to UPF2, the complex is not stable to gel filtration, unlike the wild-type complex.

In the vast majority of eukaryotic genomes, including metazoans, yeasts, and the parasites *Trypanosoma brucei*, *Trypanosoma cruzi* and *Plasmodium falciparum*, the residues E182, Y184, and 204-VVV-206 are absolutely conserved in the UPF1 homologs (Fig. 1), and furthermore, a corresponding putative UPF2 homolog can be identified in each case (data not shown). The importance of these residues for interacting with UPF2 is also supported by the following phylogenetic data. In the intracellular parasite *Encephalitozoon cuniculi*, there are small deletions and substitutions in the CH-rich domain of the putative UPF1 (accession code Q8SR02) that would completely alter the putative UPF2-interacting surface. In particular,

a part of loop L7 and $\beta 6$ containing 204-VVV-206 is missing, and E182 and Y184 are mutated to, respectively, lysine and serine (Fig. 1). Moreover, the minimal genome of this parasite does not apparently code for UPF2. The absence of UPF2 and conserved UPF2-interacting residues in *E. cuniculi* UPF1 suggests first that if NMD occurs in this organism, it involves a different, UPF2-independent, mechanism than that normally invoked. Second, the conservation of the CH-rich domain of UPF1, despite the absence of UPF2, suggests that this domain has additional roles, perhaps related to conserved non-NMD functions of UPF1. A similar situation occurs in the putative UPF1 homologs of *Entamoeba histolytica* (XP_657569 and EAL50744) and *Giardia lamblia* (Q7QU62), where, again, UPF2-interacting residues E182, Y184, and 204-VVV are mutated (Fig. 1), correlating with the apparent absence of UPF2 in the available genome sequence. On the other hand, UPF2 homologs are apparent in the *T. brucei*, *T. cruzi*, and *P. falciparum* genomes.

In conclusion, the structure of the CH-rich domain of UPF1 is the first structural information on this key NMD factor and provides a starting point for further functional studies to help elucidate its exact functions in gene expression including NMD.

MATERIALS AND METHODS

Expression, purification, and crystallization of UPF1(115–272)

Human UPF1(115–272) was expressed in *E. coli* BL21Star(DE3) from a pProEXHTb expression vector as a His-tag fusion protein and was purified by affinity chromatography using Ni^{2+} resin. After His-tag cleavage with TEV protease, the protein was further purified by a second Ni^{2+} column and size-exclusion chromatography. Pure UPF1(115–272) was concentrated to ~ 6.5 mg/mL in a buffer containing 20 mM Tris (pH 7.0), 100 mM NaCl, and 10 mM β -mercaptoethanol for crystallization. Diffracting crystals grew within a week at 4°C in a condition containing 30% pentaerythritol ethoxylate (15/4EO/OH) (v/v), 50 mM ammonium sulfate, and 50 mM Bis-Tris (pH 6.5). For data collection at 100 K, crystals were snap-frozen in liquid nitrogen with a solution containing mother liquor and 30% (v/v) glycerol.

Data collection and structure determination

The UPF1(115–272) crystals belong to space group $P2_12_12_1$ with unit cell dimensions $a = 64.6$ Å, $b = 73.0$ Å, $c = 73.2$ Å and a solvent content of 48.2%. The crystals diffract to a resolution of 3.0 Å. Diffraction data were collected at 100 K using a Q4R ADSC CCD detector on beamline ID14-EH4 at the European Synchrotron Radiation Facility (ESRF, Grenoble, France) and processed using XDS (Kabsch 1993). The structure was solved by a zinc single anomalous dispersion (SAD) experiment. Data were collected at the Zn K-absorption edge (X-ray wavelength 1.2821 Å) using the inverse beam data collection method to increase the accuracy of anomalous difference measurements. However, owing to ice rings and the small size of the crystal, data between 4 and 3 Å resolution are of lower quality. Using 4 Å resolution data, SHELXD (Uson and Sheldrick 1999) readily found six sites, corresponding to two molecules of UPF1 per asymmetric unit each containing three zinc atoms. These six sites were refined and used for phasing in SHARP (Bricogne et al. 2003). Phases were then extended to 3 Å with DM (Cowtan 1994). In the initial electron density map calculations, three helices and three β -strands per molecule were identified with the help of FFPEAR. Together with the positions of the zinc sites, this partial model served to establish the NCS operators and to define a mask that was then used for NCS averaging in DM to improve the electron density maps. By iterating this process, a complete model could eventually be built using O (Jones et al. 1991). During refinement, the two UPF1 molecules were constrained by tight noncrystallographic symmetry except for flexible regions 199–204 (L7) and 216–228 ($\alpha 2$ and L9), which differ significantly in conformation. Three TLS groups per subunit (residues 118–183, 184–240, 241–272) were defined with the help of the TLSMD server (<http://skuld.bmsc.washington.edu/~tlsmd/>). TLS parameters accounted for much of the individual B -factor variation and use of this procedure reduced the R_{free} by 2.3%. The final model, obtained with REFMAC5 (Murshudov et al. 1997), has an R_{free} of 25.9% and R -factor of 22.1% with good geometry and all residues in the favored (89.1%) or additionally allowed (10.9%) regions of the Ramachandran plot, as analyzed by PROCHECK (Laskowski et al. 1993). The final structure was also validated by calculating with CNS a systematic simulated annealing omit map covering the whole asymmetric unit. Detailed data and refinement statistics are in Table 1.

His-tag pull-down assays

UPF2(761–1207) and UPF1(115–272) were, respectively, cloned into a pProEXHTb expression vector in order to produce His-tag fusion protein and into pRSFDuet-1 (Novagen) to produce native protein without any tag. Mutations were generated using a QuickChange site-directed mutagenesis kit and confirmed by sequencing. His-tagged UPF1 and UPF2 were coexpressed in *E. coli* BL21Star(DE3), and the resultant complexes were purified using Ni^{2+} resin. The resin was extensively washed with 20 mM and 100 mM imidazole. Bound proteins were eluted with 250 mM imidazole and analyzed by 15% SDS-PAGE.

ACKNOWLEDGMENTS

We thank members of the EMBL-ESRF Joint Structural Biology Group, notably, Andrew McCarthy, for assistance with data

collection on ESRF beamlines. We are grateful to Carlo Petosa, Luc Bousset, and Thibaut Crepin for their help in the crystallographic analysis. We also thank Lynne Maquat (University of Rochester) for providing hUPF1 cDNA. We thank Elisa Izaurralde for critical reading of the manuscript. Coordinates and structure factors have been deposited in the Protein Data Bank with entry code 2IYK.

Received June 6, 2006; accepted July 14, 2006.

REFERENCES

- Andersen, P., Kragelund, B.B., Olsen, A.N., Larsen, F.H., Chua, N.H., Poulsen, F.M., and Skriver, K. 2004. Structure and biochemical function of a prototypical *Arabidopsis* U-box domain. *J. Biol. Chem.* **279**: 40053–40061.
- Applequist, S.E., Selg, M., Raman, C., and Jack, H.M. 1997. Cloning and characterization of HUPF1, a human homolog of the *Saccharomyces cerevisiae* nonsense mRNA-reducing UPF1 protein. *Nucleic Acids Res.* **25**: 814–821.
- Bhattacharya, A., Czaplinski, K., Trifillis, P., He, F., Jacobson, A., and Peltz, S.W. 2000. Characterization of the biochemical properties of the human Upf1 gene product that is involved in nonsense-mediated mRNA decay. *RNA* **6**: 1226–1235.
- Bricogne, G., Vonrhein, C., Flensburg, C., Schiltz, M., and Paciorek, W. 2003. Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0. *Acta Crystallogr. D Biol. Crystallogr.* **59**: 2023–2030.
- Conti, E. and Izaurralde, E. 2005. Nonsense-mediated mRNA decay: Molecular insights and mechanistic variations across species. *Curr. Opin. Cell Biol.* **17**: 316–325.
- Cowtan, K. 1994. Joint CCP4 and ESRF-EACBM. *Newsletter Protein Crystallogr.* **31**: 34–38.
- Czaplinski, K., Ruiz-Echevarria, M.J., Paushkin, S.V., Han, X., Weng, Y., Perlick, H.A., Dietz, H.C., Ter-Avanesyan, M.D., and Peltz, S.W. 1998. The surveillance complex interacts with the translation release factors to enhance termination and degrade aberrant mRNAs. *Genes & Dev.* **12**: 1665–1677.
- de Pinto, B., Lippolis, R., Castaldo, R., and Altamura, N. 2004. Overexpression of Upf1p compensates for mitochondrial splicing deficiency independently of its role in mRNA surveillance. *Mol. Microbiol.* **51**: 1129–1142.
- Domeier, M.E., Morse, D.P., Knight, S.W., Portereiko, M., Bass, B.L., and Mango, S.E. 2000. A link between RNA interference and nonsense-mediated decay in *Caenorhabditis elegans*. *Science* **289**: 1928–1931.
- Fukuhara, N., Ebert, J., Unterholzner, L., Lindner, D., Izaurralde, E., and Conti, E. 2005. SMG7 is a 14–3–3-like adaptor in the nonsense-mediated mRNA decay pathway. *Mol. Cell* **17**: 537–547.
- Garman, E.F. and Grime, G.W. 2005. Elemental analysis of proteins by microPIXE. *Prog. Biophys. Mol. Biol.* **89**: 173–205.
- Gonzalez, C.I., Ruiz-Echevarria, M.J., Vasudevan, S., Henry, M.F., and Peltz, S.W. 2000. The yeast hnRNP-like protein Hrp1/Nab4 marks a transcript for nonsense-mediated mRNA decay. *Mol. Cell* **5**: 489–499.
- He, F., Brown, A.H., and Jacobson, A. 1996. Interaction between Nmd2p and Upf1p is required for activity but not for dominant-negative inhibition of the nonsense-mediated mRNA decay pathway in yeast. *RNA* **2**: 153–170.
- . 1997. Upf1p, Nmd2p, and Upf3p are interacting components of the yeast nonsense-mediated mRNA decay pathway. *Mol. Cell. Biol.* **17**: 1580–1594.
- Holm, L. and Sander, C. 1993. Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233**: 123–138.

- Hosoda, N., Kim, Y.K., Lejeune, F., and Maquat, L.E. 2005. CBP80 promotes interaction of Upf1 with Upf2 during nonsense-mediated mRNA decay in mammalian cells. *Nat. Struct. Mol. Biol.* **12**: 893–901.
- Ishigaki, Y., Li, X., Serin, G., and Maquat, L.E. 2001. 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.
- Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. 1991. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**: 110–119.
- Kabsch, W. 1993. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Crystallogr.* **26**: 795–800.
- Kadlec, J., Izaurralde, E., and Cusack, S. 2004. The structural basis for the interaction between nonsense-mediated mRNA decay factors UPF2 and UPF3. *Nat. Struct. Mol. Biol.* **11**: 330–337.
- Kashima, I., Yamashita, A., Izumi, N., Kataoka, N., Morishita, R., Hoshino, S., Ohno, M., Dreyfuss, G., and Ohno, S. 2006. 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.
- Kaygun, H. and Marzluff, W.F. 2005. Regulated degradation of replication-dependent histone mRNAs requires both ATR and Upf1. *Nat. Struct. Mol. Biol.* **12**: 794–800.
- Kim, Y.K., Furic, L., Desgroseillers, L., and Maquat, L.E. 2005. Mammalian Staufen1 recruits Upf1 to specific mRNA 3'UTRs so as to elicit mRNA decay. *Cell* **120**: 195–208.
- Klein, D.J., Moore, P.B., and Steitz, T.A. 2004. The roles of ribosomal proteins in the structure assembly, and evolution of the large ribosomal subunit. *J. Mol. Biol.* **340**: 141–177.
- Krishna, S.S., Majumdar, I., and Grishin, N.V. 2003. Structural classification of zinc fingers: Survey and summary. *Nucleic Acids Res.* **31**: 532–550.
- Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. 1993. PROCHECK: A program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**: 283–291.
- Le Hir, H., Gatfield, D., Izaurralde, E., and Moore, M.J. 2001. 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.
- Legge, G.B., Martinez-Yamout, M.A., Hambly, D.M., Trinh, T., Lee, B.M., Dyson, H.J., and Wright, P.E. 2004. ZZ domain of CBP: an unusual zinc finger fold in a protein interaction module. *J. Mol. Biol.* **343**: 1081–1093.
- Lykke-Andersen, J. 2002. Identification of a human decapping complex associated with hUpf proteins in nonsense-mediated decay. *Mol. Cell. Biol.* **22**: 8114–8121.
- Lykke-Andersen, J., Shu, M.D., and Steitz, J.A. 2000. Human Upf proteins target an mRNA for nonsense-mediated decay when bound downstream of a termination codon. *Cell* **103**: 1121–1131.
- Maquat, L.E. 2004. Nonsense-mediated mRNA decay: Splicing, translation and mRNP dynamics. *Nat. Rev. Mol. Cell Biol.* **5**: 89–99.
- Medghalchi, S.M., Frischmeyer, P.A., Mendell, J.T., Kelly, A.G., Lawler, A.M., and Dietz, H.C. 2001. Rnt1, a *trans*-effector of nonsense-mediated mRNA decay, is essential for mammalian embryonic viability. *Hum. Mol. Genet.* **10**: 99–105.
- Mendell, J.T., ap Rhys, C.M., and Dietz, H.C. 2002. Separable roles for rnt1/hUpf1 in altered splicing and decay of nonsense transcripts. *Science* **298**: 419–422.
- Murshudov, G.N., Vagin, A.A., and Dodson, E.J. 1997. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**: 240–255.
- Nazareus, T., Cedarberg, R., Bell, R., Cheatle, J., Forch, A., Haifley, A., Hou, A., Wanja Kebaara, B., Shields, C., et al. 2005. Upf1p, a highly conserved protein required for nonsense-mediated mRNA decay, interacts with the nuclear pore proteins Nup100p and Nup116p. *Gene* **345**: 199–212.
- Ohi, M.D., Vander Kooi, C.W., Rosenberg, J.A., Chazin, W.J., and Gould, K.L. 2003. Structural insights into the U-box, a domain associated with multi-ubiquitination. *Nat. Struct. Biol.* **10**: 250–255.
- Petroski, M.D. and Deshaies, R.J. 2005. Function and regulation of cullin-RING ubiquitin ligases. *Nat. Rev. Mol. Cell Biol.* **6**: 9–20.
- Takahashi, S., Araki, Y., Sakuno, T., and Katada, T. 2003. Interaction between Ski7p and Upf1p is required for nonsense-mediated 3'-to-5' mRNA decay in yeast. *EMBO J.* **22**: 3951–3959.
- Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., and Higgins, D.G. 1997. The CLUSTAL_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25**: 4876–4882.
- Unterholzner, L. and Izaurralde, E. 2004. SMG7 acts as a molecular link between mRNA surveillance and mRNA decay. *Mol. Cell* **16**: 587–596.
- Uson, I. and Sheldrick, G.M. 1999. Advances in direct methods for protein crystallography. *Curr. Opin. Struct. Biol.* **9**: 643–648.
- Weng, Y., Czaplinski, K., and Peltz, S.W. 1996. Identification and characterization of mutations in the UPF1 gene that affect nonsense suppression and the formation of the Upf protein complex but not mRNA turnover. *Mol. Cell. Biol.* **16**: 5491–5506.
- Zheng, N., Wang, P., Jeffrey, P.D., and Pavletich, N.P. 2000. Structure of a c-Cbl-UbcH7 complex: RING domain function in ubiquitin-protein ligases. *Cell* **102**: 533–539.