2648–2659 *Nucleic Acids Research, 2018, Vol. 46, No. 5* *Published online 25 January 2018*
doi: 10.1093/nar/gky040

A conserved structural element in the RNA helicase UPF1 regulates its catalytic activity in an isoform-specific manner

Manjeera Gowravaram[1], Fabien Bonneau[2], Joanne Kanaan[3], Vincent D. Maciej[1], Francesca Fiorini[3], Saurabh Raj[3,4], Vincent Croquette[3,4], Hervé Le Hir[3] and Sutapa Chakrabarti[1,*]

[1]Institute of Chemistry and Biochemistry, Freie Universität Berlin, Takustr. 6, D-14195 Berlin, Germany, [2]Max Planck Institute of Biochemistry, Structural Cell Biology Department, Am Klopferspitz 18, D-82152 Martinsried, Germany, [3]IBENS, Département de Biologie, Ecole Normale Supérieure, CNRS, Inserm, PSL Research University, F-75005 Paris, France and [4]LPS, Département de physique de l'ENS, École normale supérieure, PSL Research University, Université Paris Diderot, Sorbonne Paris Cité, Sorbonne Universités, UPMC Univ. Paris 06, CNRS, F-75005 Paris, France

Received September 19, 2017; Revised January 12, 2018; Editorial Decision January 15, 2018; Accepted January 17, 2018

ABSTRACT

The RNA helicase UPF1 is a key component of the nonsense mediated mRNA decay (NMD) pathway. Previous X-ray crystal structures of UPF1 elucidated the molecular mechanisms of its catalytic activity and regulation. In this study, we examine features of the UPF1 core and identify a structural element that adopts different conformations in the various nucleotide- and RNA-bound states of UPF1. We demonstrate, using biochemical and single molecule assays, that this structural element modulates UPF1 catalytic activity and thereby refer to it as the regulatory loop. Interestingly, there are two alternatively spliced isoforms of UPF1 in mammals which differ only in the lengths of their regulatory loops. The loop in isoform 1 ($UPF1_1$) is 11 residues longer than that of isoform 2. We find that this small insertion in $UPF1_1$ leads to a two-fold increase in its translocation and ATPase activities. To determine the mechanistic basis of this differential catalytic activity, we have determined the X-ray crystal structure of the helicase core of $UPF1_1$ in its apo-state. Our results point toward a novel mechanism of regulation of RNA helicases, wherein alternative splicing leads to subtle structural rearrangements within the protein that are critical to modulate enzyme movements and catalytic activity.

INTRODUCTION

Gene expression in eukaryotes is a complex, multi-step process that is subject to stringent regulation at every stage. This regulation can be mediated at the level of DNA (differential transcription) or protein (translation and selective degradation) or at an intermediate step, at the level of mRNA. Post-transcriptional gene regulation occurs at several stages in the lifetime of an mRNA, such as processing, export, translation and degradation [reviewed in (1) and (2)]. Each of these processes is a complex multi-step event involving the dynamic assembly, remodeling and disassembly of messenger ribonucleoprotein particles (mRNPs) (3). Such events are often facilitated by a group of ATP-dependent enzymes called RNA helicases, which utilize the energy derived from ATP binding or hydrolysis to alter the conformation of RNA, and thereby, unwind RNA duplexes or remodel RNPs (4). RNA helicases are ubiquitously present in eukaryotes and are involved in every step of mRNA metabolism, from transcription to degradation (5). Due to their pervasiveness, in part, and functional importance, many RNA helicases are essential for cell viability and are stringently regulated by intra- and inter-molecular mechanisms (6). While many RNA helicases actively unwind RNA duplexes or remodel RNPs, some function as place-holders, to stabilize the interaction of a protein or protein complex with RNA [(7–11), reviewed in (12)]. As such, all RNA helicases can be defined as RNA-dependent ATPases, referring to their ability to hydrolyze ATP in the presence of RNA.

*To whom correspondence should be addressed. Tel: +49 30 83875094; Email: chakraba@zedat.fu-berlin.de
Present addresses:
Francesca Fiorini, Institut de Biologie et Chimie des Protéines, BMSSI-IBCP, UMR 5086 CNRS Université Lyon 1, 7, Passage du Vercors, 69367 Lyon Cedex 07, France.
Saurabh Raj, Molecular Biophysics Group, Institute for Experimental Physics 1, Universität Leipzig, Linnéstr. 5, D-04103 Leipzig, Germany.

Although there exist six superfamilies of nucleic acid-dependent ATPases (SF1 to SF6), all eukaryotic RNA helicases are members of the SF1 or SF2 superfamilies (13,14). They share a conserved core architecture, consisting of two RecA-like domains, which form a nucleotide-binding pocket and a composite RNA-binding surface. Additionally, many RNA helicases have auxiliary domains that are folded independently of the RecA core and exert a regulatory effect on the catalytic activity of the helicase [reviewed in (6)]. The mechanisms by which RNA helicases mediate unwinding of duplexes or remodeling of RNPs vary depending on their sub-family—DEAD box helicases of the SF2 superfamily act by local strand-separation of RNA duplexes, whereas DExH helicases of the SF2 superfamily as well as the SF1 superfamily of helicases are thought to mediate their effect by translocation on the nucleic acid [reviewed in (15) and (4)]. Interestingly, translocating helicases have a preferred directionality; while all DExH helicases translocate in the 3′-5′ direction, RNA helicases of the SF1 superfamily translocate in the 5′-3′ direction (13).

RNA helicases of the SF1 superfamily are also referred to as UPF1-like helicases due to their similarity with the prototype member, UPF1. UPF1 is a central component of the nonsense mediated mRNA decay (NMD) pathway, which involves the step-wise assembly and disassembly of several protein factors to mediate target mRNA degradation (16). In the NMD pathway, UPF1 employs its catalytic activity to remodel mRNPs containing premature termination codons (9,17–20). In addition to the two conserved RecA domains, UPF1 contains a number of additional domains located at different positions within its primary structure (21). The helicase core is flanked at the N-terminus by a cysteine-histidine rich (CH) domain and at the C-terminus by a stretch of unstructured amino acids rich in serine-glutamine (SQ) motifs (Figure 1A). Additionally, UPF1 contains two subdomains 1B and 1C that are inserted within the sequence of the helicase core (Figure 1A). Subdomain 1B adopts a β-barrel fold and is connected to the RecA1 domain by two long 'stalk' helices, while the all α-helical 1C subdomain packs against RecA1 (Figure 1B) (21–23). Previous biochemical and structural data indicate that each of these additional domains plays a role in regulating the catalytic activity of UPF1. The CH domain was found to repress catalytic activity of UPF1 (24). It interacts with the RecA2 domain of UPF1 and indirectly positions domain 1B to clamp down on the 3′ end of the bound RNA and inhibit translocation of the RNA relative to the protein (22). This inhibition is relieved upon binding of the conserved NMD factor UPF2 to the CH domain, which elicits a large conformational change and switches the helicase from a 'clamping'-mode to an 'unwinding'-mode (23). In addition to the CH domain, the SQ motifs at the C-terminus of UPF1 interact with the helicase core and inhibit its catalytic activity, possibly by retaining the helicase in a 'weak' RNA-binding state or by constraining the small conformational changes that are associated with ATP binding, hydrolysis and exchange (25). Lastly, deletion of the subdomain 1C was found to abolish RNA-binding and ATPase activity of UPF1, establishing that this domain is integral to the RNA-dependent ATPase activity of UPF1 (21). The position of 1C at the very 3′ end of the RNA suggests that it functions as a ratchet, conferring directionality to the RNA threading process.

In this study, we examine known structures of the UPF1 helicase core and identify a structural element within domain 1B that undergoes considerable conformational changes upon nucleotide and RNA binding. Using a combination of biochemical assays, we demonstrate that this element, referred to as the regulatory loop, modulates the catalytic activity of UPF1, possibly by interfering with the translocation of the helicase on RNA. Interestingly, UPF1 exists as two alternatively-spliced isoforms in mammals, which differ only in the length of this regulatory loop. Isoform 1 ($UPF1_1$) contains an insertion in domain 1B which extends its regulatory loop to 22 residues, while isoform 2 [$UPF1_2$, the more abundant species, (26)] has a 11-residue long loop, similar to those found in lower eukaryotes (Figure 1A and Supplementary Figure S1A, residue numbers are that of $UPF1_2$). Although the $UPF1_2$ isoform has been extensively characterized using structural and biochemical tools, very little is known about the catalytic activity and regulation of the isoform $UPF1_1$. We find that insertion of 11 amino acids within the regulatory loop considerably alters the RNA-binding and catalytic activity of UPF1 *in vitro*. In order to elucidate the basis of this differential catalytic activity, we have determined the X-ray crystal structure of the $UPF1_1$ helicase core in its apo-state. Our structural and biochemical studies suggest an elegant mechanism for the regulation of catalytic activity by introducing subtle structural changes through alternative splicing of mammalian UPF1.

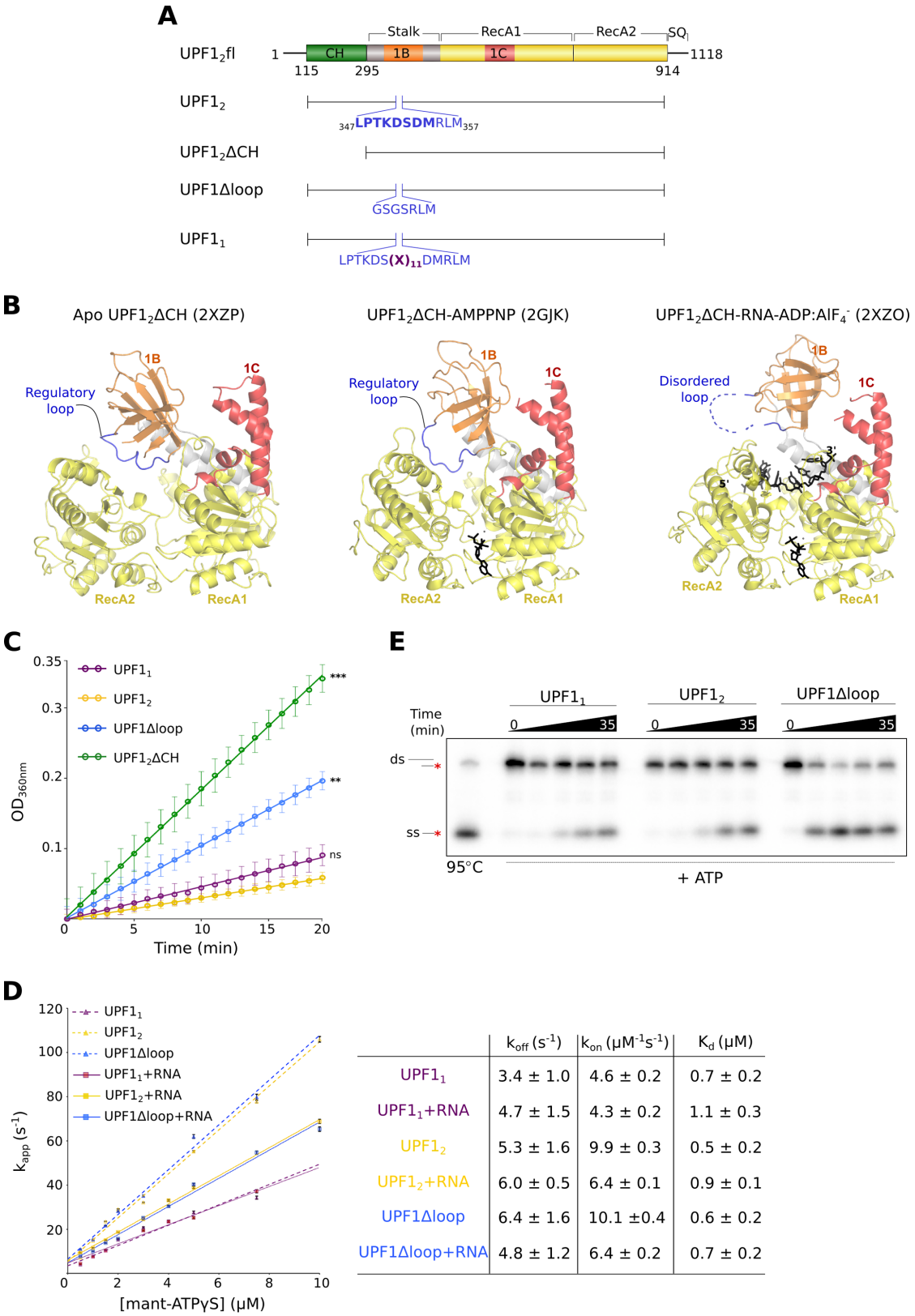
MATERIALS AND METHODS

Cloning, expression and purification

Human $UPF1_2$ and $UPF1_2\Delta CH$ were purified as previously described (22). The UPF1Δloop and UPF1ΔCHΔloop mutants were generated by one-step PCR amplification, using pET28a-$UPF1_2$ and pET15b-$UPF1_2\Delta CH$ as templates and the following primers: 5′-GAATCGCCTACTTCACTGGTTCTGGAAGTCG GCTCATGCAGGG-3′ and 5′-CCCTGCATGAGCCGA CTTCCAGAACCAGTGAAGTAGGCGATTC-3′.

The $UPF1_1$ and $UPF1_1\Delta CH$ plasmids were also generated as above, using pET28a-$UPF1_2$ and pET15b-$UPF1_2\Delta CH$ as a template and the following primers: 5′-GATCTGGTGATTATCTGGCTGCGTGACATGC GGCTCATG-3′ and 5′-GATAATCACCAGATCTTCG TTGCCAGAGTCAGTCTTGGG-3′.

Both $UPF1_1$ and UPF1Δloop were further sub-cloned from the pET28a plasmid into a modified pET vector containing an N-terminal 6X-His-Thioredoxin fusion tag, followed by a 3C protease cleavage site. The resultant fusion proteins were expressed and purified by Ni^{2+}-affinity and ion-exchange (IEX) chromatography in a manner identical to $UPF1_2$, except for an additional step of tag-cleavage by 3C protease (1:100 w/w ratio of protease to target protein). The proteins were concentrated and flash-frozen in liquid nitrogen directly after IEX chromatography. Proper folding of the proteins was confirmed by analytical size-exclusion chromatography. The UPF1Δloop protein used for stopped-flow analysis was additionally subjected to



	k_{off} (s^{-1})	k_{on} ($\mu M^{-1}s^{-1}$)	K_d (μM)
$UPF1_1$	3.4 ± 1.0	4.6 ± 0.2	0.7 ± 0.2
$UPF1_1$+RNA	4.7 ± 1.5	4.3 ± 0.2	1.1 ± 0.3
$UPF1_2$	5.3 ± 1.6	9.9 ± 0.3	0.5 ± 0.2
$UPF1_2$+RNA	6.0 ± 0.5	6.4 ± 0.1	0.9 ± 0.1
UPF1Δloop	6.4 ± 1.6	10.1 ±0.4	0.6 ± 0.2
UPF1Δloop+RNA	4.8 ± 1.2	6.4 ± 0.2	0.7 ± 0.2

Figure 1. A conserved loop in domain 1B of UPF1 modulates its catalytic activity. (**A**) Schematic representation of the domain organization of human $UPF1_2$ full length (fl) and constructs used for the biochemical assays shown below. The helicase core comprises two RecA domains (in yellow) and the subdomains 1B (in orange) and 1C (in red). Additional regulatory domains, the CH domain (in green) and the SQ domain are present N- and C-terminal to the helicase core, respectively. The residue numbers of all UPF1 constructs used in this study correspond to that of the more abundant, short isoform ($UPF1_2$). The position and the residues of the UPF1 regulatory loop are indicated. Amino acids in bold denote those replaced by the GSGS linker in

preparative size-exclusion chromatography, as described in (22). The $UPF1_1\Delta CH$ and $UPF1\Delta CH\Delta loop$ proteins were purified in a manner identical to $UPF1_2\Delta CH$.

A fragment of UPF2 containing the third MIF4G domain in addition to the UPF1-binding site (UPF2 761–1227) was expressed and purified as described in (22)

Crystallization and structure determination

Crystals of $UPF1_1\Delta CH$ were grown by the sitting-drop vapor diffusion method, by mixing together equal volumes of protein and reservoir solution (1.4 M sodium malonate, 50 mM sodium citrate pH 6.2). Crystals were cryoprotected by step-wise soaking in increasing concentrations of sodium malonate, up to a final concentration of 3 M and flash frozen in liquid N_2 prior to data collection.

Diffraction data were collected at 100 K on beamline 14.1 of the *BESSY II* storage ring (Berlin, Germany) and were processed by the program XDS (27). The structure was determined by molecular replacement with Phaser (28) using the RecA domains and domain 1C of $UPF1_2\Delta CH$ as a search model, while domain 1B was built *de novo*. Iterative model building and refinement were performed using Coot and Phenix, respectively (29,30). The final model was validated with the program Molprobity (31).

ATPase assay

6 pmol of the UPF1 proteins were pre-incubated with 2 μg poly-U RNA, 40 nmol MESG (2-amino-6-mercapto-7-methylpurine ribonucleoside) and 0.5 U purine-nucleoside phosphorylase (EnzChek Phosphate Assay kit, Invitrogen) in 1× ATPase buffer (50 mM MES pH 6.5, 50 mM potassium acetate, 5 mM magnesium acetate, 2 mM DTT) at 30°C for 20 min. The reaction was initiated by the addition of ATP to a final concentration of 1 mM. Generation of 2-amino-6-mercapto-7-methylpurine from MESG and phosphate (released from ATP hydrolysis) was detected by measuring absorbance at 360 nm on an Infinite M1000 Pro (Tecan). The reaction was allowed to proceed for 20 min and monitored during this period by measurement of A_{360} at 60-s intervals. Wherever indicated, 7.5 pmol of UPF2 (1.25× of UPF1) were added to the reaction mixture. To determine the significance of the differences observed in ATPase activity, unpaired t-tests were performed and used to determine the two-tailed *P* values. The significance threshold was set at 0.05.

Fluorescence anisotropy assay

Fluorescence anisotropy measurements were performed with a 5′-6-carboxy-fluorescein (6-FAM)-labeled 26-mer RNA at 25°C in 40 μl-reactions in a Victor plate reader (PerkinElmer). The RNA was dissolved to a concentration of 5 nM and incubated with the UPF1 proteins at different concentrations in binding buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 1 mM $MgCl_2$, 5% glycerol and 2 mM DTT). Where indicated, ATP was added to a final concentration of 10 μM. The excitation and emission wavelengths were 485 and 535 nm, respectively and each titration point was measured three times. The data were analyzed by non-linear regression fitting as described in (32) using GraphPad Prism (GraphPad Software, Inc.). Unpaired *t*-test derived *P* values were calculated to determine the significance of the differences in K_d between $UPF1_1$ and $UPF1_2$ or between $UPF1_1\Delta CH$ and $UPF1_2\Delta CH$, in the absence and presence of ATP. The significance threshold was set at 0.05.

Stopped flow assay

The kinetics of interaction of $UPF1_1$, $UPF1_2$ and $UPF1\Delta loop$ with ATP were investigated by fluorescence stopped-flow, performed in a stopped-flow apparatus (SX-20MV) using a fluorescent non-hydrolyzable ATP analogue, mant-ATPγS. The fluorescence of mant-ATPγS was excited at 290 nm by FRET from aromatic residues in the proximity of the ATP binding pocket of UPF1 and measured at 90° after passing through a 409 nm cut-off filter (KV 408). Increasing amounts of mant-ATPγS (0.5–10 μM, final concentrations) were successively titrated into a fixed amount of UPF1 (0.1 μM final concentration). Experiments were performed by rapidly mixing equal volumes (70 μl) of the protein and labeled nucleotide in buffer containing 20 mM Tris pH 7.5, 150 mM NaCl, 1 mM $MgCl_2$ and 1 μM $ZnCl_2$ and monitoring the change in fluorescence over a period of 20 s. For measurements in the presence of RNA, a 1.5-fold molar excess of U_{15} RNA was added to the protein solution and incubated at room temperature for at least 30 min. All other steps were carried out in a manner identical to the measurements in the absence of RNA. At least six replicates were performed in every case. Data were visualized using the software Pro-Data Viewer. Data analysis is described in detail in the Supplementary Information (Supplementary Figure S1).

←

UPF1Δloop. The insertion of 11 residues within the regulatory loop of $UPF1_1$ is indicated by $(X)_{11}$. (**B**) X-ray crystal structures of apo- (left panel) and AMPPNP-bound $UPF1_2\Delta CH$ (middle panel) and of $UPF1_2\Delta CH$ in its transition state, bound to RNA and ADP:AlF_4^- (right panel). The domains are colored according to the schematic in (A). RNA and nucleotides are shown as black sticks. The 11-amino acid regulatory loop in domain 1B (colored blue) occupies part of the RNA-binding surface in the apo- and AMPPNP-bound structures, but is disordered in the transition-state structure. This and all other structure figures have been generated using PyMOL (http://www.pymol.org). (**C**) RNA-dependent ATPase assay of $UPF1_1$, $UPF1_2$, $UPF1_2\Delta CH$ and $UPF1\Delta loop$ performed using an enzyme-coupled phosphate-detection assay. The data points and error bars represent mean values and standard deviations from three experiments, respectively. The symbols ns, ** and *** denote unpaired *t*-test derived *P* values of >0.05, 0.001 and <0.001 respectively. Enhancement of ATPase activity upon deletion of the CH domain is shown for comparison. (**D**) Stopped-flow analyses of binding of mant-ATPγS to $UPF1_1$, $UPF1_2$ and $UPF1\Delta loop$, in the presence and absence of RNA. The plot shows the concentration dependence of k_{app} values for UPF1:mant-ATP association in each case. The data points represent values derived from fitting an average of at least six time-course replicates while the error bars denote the standard error of mean (SEM). The table presents the rate constants of association and dissociation (k_{on} and k_{off}), along with their SEM, and the resultant equilibrium dissociation constant (K_d). Unpaired *t*-test derived *P* values suggest that the differences in K_d between $UPF1_1$ and $UPF1_2$ and between $UPF1_2$ and $UPF1\Delta loop$, both in the presence and absence of RNA, are not significant ($P > 0.05$ in every case). (**E**) ATP-dependent unwinding assays with UPF1 proteins, using a RNA:DNA hybrid substrate with a 5′-RNA overhang. Positions of the duplex, the single-stranded nucleic acid and of the ^{32}P label (*) are indicated.

Nucleic acid unwinding assay

The substrate used in the unwinding assay was identical to that described in ref (22). Briefly, a 75-mer RNA strand (5′-GGGACCGGAUGAGCGGUAUUGAGUUUGA AUUUAUCGAUGGUAUCAGAUCUGGAUCCU CGAGAAGCUGCGGGUACC-3′) was hybridized to a 21-mer DNA (5′-GGAGCTCTTCGACGCCCATGG-3′) to yield a RNA:DNA hybrid with a 54-nucleotide overhang. The duplex was labeled with ^{32}P at the 5′-end of the DNA strand. Reactions were carried out at 30°C under conditions described in (33), with 22.5 fmoles of the labeled duplex and 4.5 pmole of each UPF1 protein per reaction. Reactions were resolved by 12% native PAGE and visualized by phosphorimaging.

RT-PCR

RNA was extracted from ∼1 × 10^6 cells using RNATri (Bio&Sell) and 0.5 μg of total RNA was used for reverse transcription with the UPF1 reverse primer (sequence given below). To detect the mRNA corresponding to the two isoforms of human UPF1, we designed primers specific to sequences in exons 7 and 8, which flank the alternative 5′-splice site and, upon amplification, would generate products of 166 and 133 bp for human $UPF1_1$ and $UPF1_2$ isoforms, respectively. The primer sequences are as follows: forward 5′-GGGACCTGGGCCTTAACAAGAAGAGA-3′ and reverse 5′-ATCCCTTTCCACAGGGGCGCAAGGT-3′. Plasmids expressing the human $UPF1_1\Delta$CH and $UPF1_2\Delta$CH proteins were used as positive controls. For quantification of mRNA levels of the two isoforms low-cycle PCRs were performed with the ^{32}P-labeled forward primer (sequence given above), products were resolved by denaturing-PAGE and analyzed using a Phosphoimager and the ImageQuantTL software (GE Life Sciences). The mRNA preparations of different mouse tissues were analyzed as described above, using the human-specific forward primer (sequence given above, identical to the mouse sequence in this region) and the following mouse-specific reverse primer: 5′-ATCCCCTTCCACAGGGGCGCCAGAT-3′.

Single molecule experiments

Recombinant proteins used for single molecule experiments were cloned in a pET28a (Novagen) derivative plasmid and expressed in *Escherichia coli* BL21 *(DE3)* cells grown in LB medium and induced overnight at 16°C. The proteins were CBP-tagged at the N-terminus and His-tagged at the C-terminus. After cell harvest and lysis, proteins were purified on Nickel columns (Ni-NTA, Qiagen) and further purified on a calmodulin affinity column (Agilent technologies) before dialysis and storage at –80°C. Buffer composition for purifications can be further found in (24) and (34).

The DNA substrate used in the single-molecule studies corresponds to a 1239 bp hairpin with a 4-nt loop, a 76-nt 5′-biotinylated ssDNA tail and a 146 bp 3′-digoxigenin-labelled dsDNA tail previously described (34,35). We used a PicoTwist magnetic tweezers instrument (www.picotwist.com) to manipulate individual DNA hairpin molecules. The DNA hairpins were attached by the 5′-biotinylated extremity to streptavidin-coated magnetic beads (Dynabeads MyOne streptavidin T1, Thermofisher Scientific) and by a 3′-digoxigenin modified extremity to an anti-Dig-coated glass surface. The glass coverslip was treated with anti-digoxigenin antibody (Roche) and passivated with 1× Passivation Buffer [1× PBS pH 7.5, 0.2% pluronic surfactant, 5 mM EDTA, 10 mM sodium azide, and 0.2% of BSA (Sigma-Aldrich)]. The beads were trapped in the magnetic field generated by a pair of magnets located above the reaction chamber. Experiments were conducted at 37°C. The helicase buffer was 20 mM Tris–HCl pH 7.5, 75 mM potassium acetate, 3 mM magnesium chloride, 2% BSA, 0.5 mM DTT and 2 mM ATP. The enzyme concentration was the lowest possible to observe helicase activity in single-molecule conditions. During the helicase assays, DNA hairpins were maintained at a constant force of 10 pN.

In order to calculate processivity, we exclusively selected the bursts starting in the first 50 bp of the hairpin. $UPF1_1\Delta$CH reached the end of the 1.2 kb DNA hairpin in 27 unwinding bursts out of 29, and only two bursts were aborted. We calculated a processivity factor, fp, as the ratio of the number of enzymes reaching the hairpin apex over the total number of bursts. Assuming that the helicase has a constant probability of detaching in time and sequence, we extrapolated an average processivity of 17 ± 8 kb (34).

To calculate the velocity of unwinding and translocation, we analyzed the instantaneous rates of each enzyme at a 10 pN force, and generated histograms of distances travelled at each rate. In (34), the unwinding and translocation velocities corresponded to the histograms maximum values. However in this work, we selected the local rate derivative as an average velocity measure for both isoforms during unwinding and translocation, to avoid the bias generated by UPF1 low velocity and high number of pauses during unwinding (peak close to 0 in the histogram of unwinding rates).

RESULTS

A flexible loop in domain 1B affects the catalytic activity of UPF1

To date, several X-ray crystal structures of different states of the $UPF1_2$ isoform have been determined. These include structures of a human $UPF1_2$ construct encompassing the CH and helicase domains (referred to as $UPF1_2$ in this study, Figure 1A) bound to its activator, UPF2, (23) as well as that of yeast UPF1 in its transition state, bound to RNA and ADP-aluminum tetrafluoride (ADP:AlF_4^-) (22). Additionally, structures of the human $UPF1_2$ helicase core (referred to as $UPF1_2\Delta$CH) in its apo-, transition- and nucleotide-bound states are available (21,22). Together, these structures provide insight into the conformational changes associated with RNA- and nucleotide binding of UPF1 and its activation by UPF2.

We analyzed the X-ray crystal structures of the different states of $UPF1_2\Delta$CH and observed that, in addition to lobe closure between the two RecA domains, domain 1B undergoes considerable movement upon nucleotide and RNA binding while domain 1C is relatively rigid. In particular, an 11-residue conserved loop (amino acids 347–357, shown in

blue) within 1B adopts strikingly different conformations in the different structures of the helicase (Figure 1B). In apo-$UPF1_2\Delta CH$, part of this loop folds into an over-wound 3_{10} helix (36), while in the ATP- (AMPPNP-) bound state, it adopts an extended conformation (Figure 1B, left and middle panels). In both these conformations, the loop occludes part of the RNA-binding surface of UPF1 and appears to impede RNA binding (Supplementary Figure S1B). Interestingly, this loop is disordered in the transition state of $UPF1_2\Delta CH$ and no longer occupies the RNA-binding surface (Figure 1B, right panel). Due to a possible role of the loop in modulating RNA binding in different states of the helicase, we refer to this loop as the 'regulatory' loop. The conformational flexibility of the regulatory loop is probably due to large movements in domain 1B which either position the loop in close proximity to the helicase core, where it is well-ordered or render it solvent accessible and as a result, disordered.

Since the conformation of the regulatory loop appears to be coupled to the nucleotide- and RNA-bound state of UPF1, we wished to determine if it impacts the catalytic activity of UPF1. To this end, we generated a mutant UPF1 protein, where residues 347–354 in domain 1B have been replaced by a 4-residue GSGS linker. We reasoned that the GSGS linker is long enough to connect the flanking β-strands without compromising the overall fold of domain 1B, but not sufficiently long to exhibit the conformational flexibility observed with the native regulatory loop. This mutant protein, referred to as UPF1Δloop (Figure 1A), was found to be identical to $UPF1_2$ in terms of its oligomeric state and thermal stability, suggesting that shortening the loop in domain 1B does not perturb the overall fold of the protein (Supplementary Figure S1C). In order to discern the effect of the regulatory loop on the catalytic activity of UPF1, we performed RNA-dependent ATPase assays with $UPF1_2$, $UPF1_2\Delta CH$ and UPF1Δloop proteins in the presence of poly-(U) RNA. As reported earlier, $UPF1_2$ has a low basal ATPase activity, which is greatly stimulated upon deletion of the CH domain (Figure 1C, compare yellow and green traces) (22,24). We observed that deletion of the regulatory loop also led to an increase in ATPase activity compared to $UPF1_2$, albeit to a lower extent than that for $UPF1_2\Delta CH$ (Figure 1C, compare yellow and blue traces).

Previous studies show that the affinity of $UPF1_2$ for RNA is reduced in the presence of ATP and that deletion of residues 351–355 or 352–354 within the regulatory loop abolish this effect (21). This raises the question of whether the regulatory loop might have a role in coupling the binding of ATP and RNA to the helicase core. To investigate this, we carried out stopped-flow binding kinetic experiments of the UPF1 proteins with a fluorescently labeled non-hydrolyzable ATP analogue, mant-ATPγS, in the presence and absence of U_{15} RNA. A time-course for the binding of mant-ATPγS to $UPF1_2$, as determined by monitoring mant fluorescence, is shown in Supplementary Figure S1D. The data were fitted to a 1:1 binding model in every case to obtain an apparent rate constant, k_{app} (described in Supplementary Figure S1). A plot of k_{app} as a function of mant-ATP concentration yielded a linear dependence and was used to determine the rate constants of association (k_{on}) and dissociation (k_{off}), and the equilibrium dissociation constant, K_d, in each case (Figure 1D). We observed that binding of RNA to $UPF1_2$ did not significantly alter its affinity for ATP (K_d of 0.9 μM and 0.5 μM in the presence and absence of RNA, respectively). The marginal difference in K_d can be attributed to the higher k_{on} of ATP in the absence of RNA. A similar trend was observed with the UPF1Δloop protein, where despite a small decrease of k_{off} and k_{on}, the overall K_d for ATP binding remained unchanged upon addition of RNA (Figure 1D, right panel). Interestingly, we observed no differences in the ATP-binding affinities of the $UPF1_2$ and UPF1Δloop proteins, suggesting that although the regulatory loop participates in ATP-driven conformational changes, it does not affect binding of ATP to the helicase core.

Auxiliary domains of helicases occasionally act as a 'pin' to separate strands of a nucleic acid duplex or as a 'ratchet' to impose directionality of nucleic acid unwinding. The pin or ratchet domains do not affect the rate of ATP hydrolysis by the helicase *per se*, but influence the mechanism of helicase activity. In the SF1B DNA helicase, RecD2, deletion of the pin domain uncouples ATPase and helicase activities (37). The mutant protein is capable of hydrolyzing ATP but is no longer capable of unwinding a DNA duplex. Likewise, deletion of the ratchet domain in the helicase Hel308 (DExH DNA helicase) leads to a loss in DNA-unwinding activity without a concomitant loss of ATPase activity (38). To determine if the regulatory loop observed in the $UPF1_2\Delta CH$ structures mimics a pin or a ratchet domain, we carried out ATP-dependent helicase assays using a RNA:DNA hybrid substrate. Since UPF1 binds single-stranded (ss) RNA and unwinds nucleic acids in the 5′-3′ direction *in vitro*, the hybrid was designed to have a 5′ ssRNA overhang of 54 nucleotides [Figure 1E, described in (22)]. The ATP-dependent unwinding activity of UPF1Δloop on this substrate was found to be significantly higher than that of $UPF1_2$, as shown by the accumulation of labeled ssDNA (Figure 1E, compare right and middle panels). Remarkably, a substrate with a 3′ ssRNA overhang was not unwound by either of the two proteins, indicating that deletion of the regulatory loop does not affect the directionality of unwinding by UPF1 (Supplementary Figure S1E). In light of our observations, we deduce that the regulatory loop does not act as a classical pin or ratchet domain but rather plays a role in modulating UPF1 catalytic activity.

The two isoforms of UPF1 differ in their RNA-binding and catalytic activities

As mentioned earlier, mammalian UPF1 exists as two isoforms that are generated by the presence of an alternative 5′ splice site in intron 7 of the UPF1 gene. Other metazoans and lower eukaryotes such as flies, worms and yeast contain only one UPF1 isoform, possibly due to the absence of an alternative 5′ splice site at an equivalent position in the gene (Supplementary Figure S2A). Although the shorter isoform ($UPF1_2$) is the more abundant one (26), the relative amounts of the longer isoform ($UPF1_1$) appeared to vary considerably among different human cell types and across different mouse tissues (Figure 2A). The difference between the two isoforms lies in the length of their regulatory loops; the regulatory loop of $UPF1_2$ con-

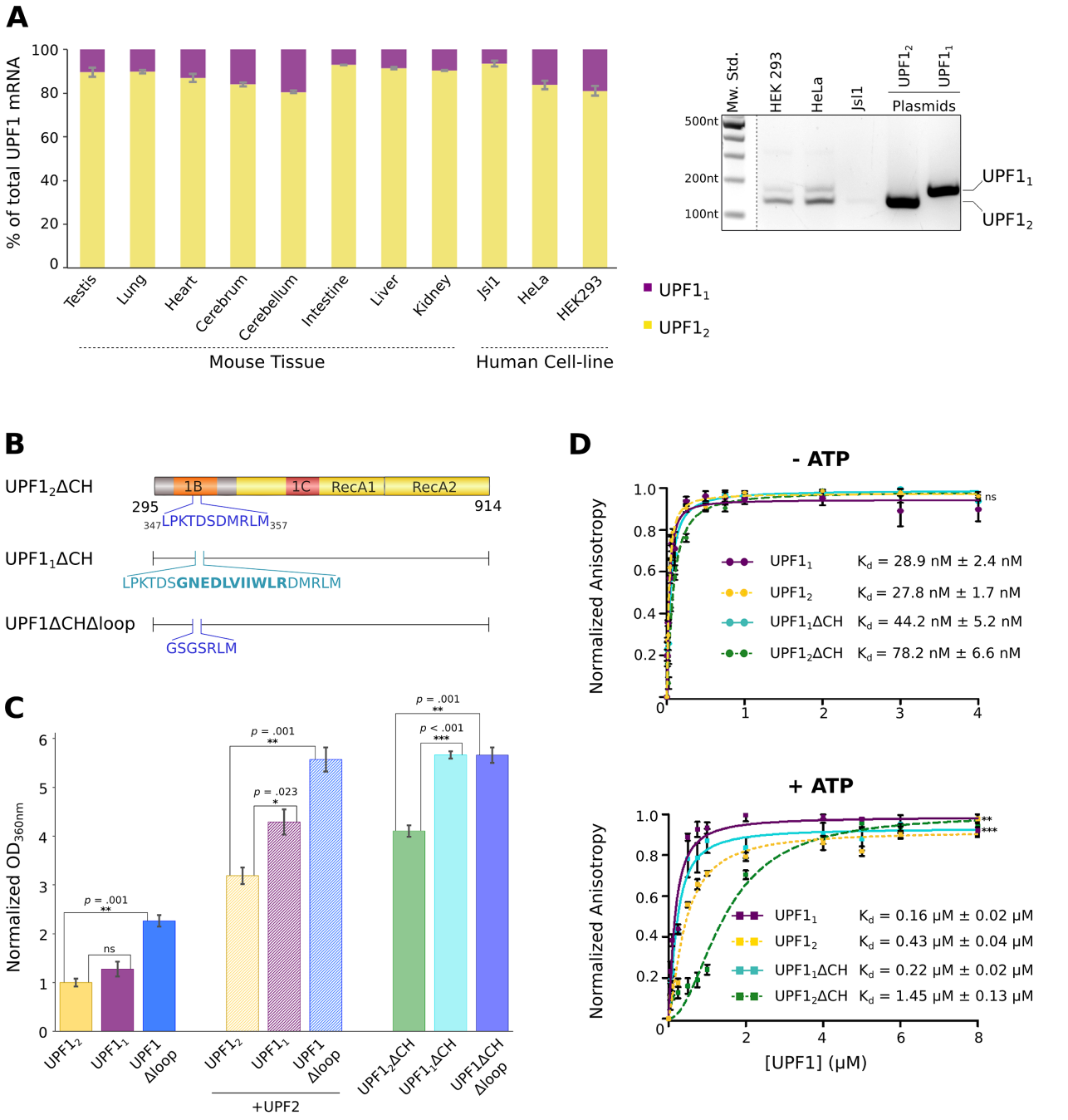


Figure 2. Mammalian UPF1 exists as two splice isoforms that differ in catalytic and RNA binding activity. (**A**) Quantitative RT-PCR analysis of mRNA levels corresponding to the two UPF1 isoforms in different mouse tissues and human cell-lines. The right panel is an ethidium bromide-stained agarose gel of the products obtained after RT-PCR analysis of the human cell-line samples. The left panel shows a quantification of the mRNA levels of the UPF1 isoforms (derived from analysis of radioactive PCR products). Biological duplicates of mRNA samples from mouse tissue were used for this analysis. The data represent the mean of at least five experiments and error bars (in grey) denote the standard deviation from the mean. (**B**) Schematic representation of the UPF1ΔCH constructs used in biochemical assays shown below. The regulatory loop of $UPF1_1$ is indicated and the 11-amino acid insertion specific to isoform 1 is denoted in bold. The construct UPF1ΔCHΔloop contains the same deletion as UPF1Δloop in Figure 1. (**C**) RNA-dependent ATPase assay of the UPF1 isoforms and UPF1Δloop proteins carried out using an enzyme-coupled phosphate-detection assay. Each column represents the ATPase activity of the indicated protein (or protein mixture) at the end point of the reaction (20 minutes), normalized against the activity of $UPF1_2$. The data represent mean values and standard deviations from three experiments. A time-course of all the reactions is shown in Supplementary Figure S2C. The ATPase activities of $UPF1_1$ and $UPF1_2$ have been compared in the absence and presence of UPF2 and to the activities of their corresponding ΔCH constructs (left, middle and right stacks). Unpaired t-test derived *P* values are indicated. (**D**) Quantitative measurements of RNA-binding affinities of UPF1 isoforms, with and without the CH domain, in the absence and presence of ATP (top and bottom panels, respectively) by fluorescence anisotropy assays. The K_ds and their corresponding errors are the mean and SEM of at least 3 independent experiments. Error bars on the traces denote standard deviation from the mean. Unpaired *t*-test derived *P* values suggest that the difference in K_d between $UPF1_1$ and $UPF1_2$ in the absence of ATP is not significant (ns, $P > 0.05$) while the difference in the presence of ATP is very significant ($^{**}P = 0.004$). Similarly, the difference in K_d between $UPF1_1$ΔCH and $UPF1_2$ΔCH in the presence of ATP is extremely significant ($^{***}P < 0.001$).

Table 1. X-ray crystallographic data collection and refinement statistics

Data collection	
Beamline	*BESSY II* BL 14.1
Wavelength (Å)	0.918
Space group	$P2_13$
Unit cell parameters (Å)	$a = 146.9$
Resolution range (Å)	19.62–3.34 (3.61–3.34)
Unique reflections	15 480 (1512)
Multiplicity	6.7 (6.9)
Completeness	99.2 (98.9)
Mean I/sigma(I)	13.9 (2.3)
R-merge	0.099 (0.802)
$CC_{1/2}$	0.999 (0.757)
Refinement	
Reflections used for R_{free}	777
R_{free} (%)	31.3
R_{work} (%)	29.2
RMS bonds (Å)	0.003
RMS angles (°)	0.76
Average *B* factor (Å^2)	111.92
Ramachandran values	
Favored	96.0
Allowed	4.0
Outliers	0

*Values within parentheses correspond to the highest resolution shell.

tains 11 amino acids, whereas that of $UPF1_1$ contains 22 amino acids due to insertion of an additional 11 amino acids upon usage of the alternative 5′ splice site (Figure 2B). Although the catalytic activity of $UPF1_2$ has been extensively studied by structural and biochemical means, not much is known about $UPF1_1$. Since deletion of the regulatory loop enhanced the catalytic activity of UPF1, we wondered if, conversely, the presence of a longer regulatory loop would have a strong inhibitory effect on its activity. To test this hypothesis, we generated UPF1 constructs containing the 22-residue long regulatory loop, with and without the CH domain (referred to as $UPF1_1$ and $UPF1_1\Delta$CH, Figures 1A and 2B, respectively). Both proteins were similar to their corresponding shorter isoforms and loop-deletion mutants in terms of their oligomeric state and thermal stability (Supplementary Figures S1C and S2B). We compared the ATPase activities of $UPF1_1$ and $UPF1_2$ and found that a longer regulatory loop did not inhibit the catalytic activity of the helicase (Figure 1C, compare purple and yellow traces). In fact, $UPF1_1$ showed a marginally higher ATPase activity than $UPF1_2$. ATP-dependent helicase assays with $UPF1_1$ and $UPF1_2$ also showed that the unwinding activity of UPF1 was not inhibited by a longer regulatory loop (Figure 1E, compare left and middle panels). As observed with $UPF1_2$, the affinity of $UPF1_1$ for ATP is not significantly modulated in the presence of RNA (Figure 1D, purple traces), suggesting that the longer regulatory loop also does not influence binding of ATP to the helicase core.

We next proceeded to determine if the catalytic activity of $UPF1_1$ is regulated by intra- and inter-molecular interactions, in a manner similar to $UPF1_2$. To this end, we compared the ATPase activities of $UPF1_1$ and $UPF1_2$ with the corresponding constructs lacking the CH domain ($UPF1_1\Delta$CH and $UPF1_2\Delta$CH, respectively). We found that, consistent with a role for auto-inhibition by the CH domain, $UPF1_1\Delta$CH shows about 4-fold higher activity than $UPF1_1$ (Figure 2C, compare left and right stacks). As in the case of $UPF1_2$, repression of $UPF1_1$ by its CH domain is relieved upon addition of UPF2 (Figure 2C, compare left and middle stacks). Our observations suggest that regulation of UPF1 catalytic activity by the CH domain and activation by UPF2 is not perturbed by the longer loop of $UPF1_1$.

Contrary to our expectations of stronger inhibition of catalytic activity by the long regulatory loop, we find that activated $UPF1_1$ shows significantly higher ATPase activity than activated $UPF1_2$ (Figure 2C, compare first two columns in the middle and right stacks). Furthermore, the ATPase activity of $UPF1_1\Delta$CH is identical to that of the UPF1ΔCHΔloop mutant (Figure 2C, compare second and third column in the right stack), indicating that deletion or elongation of the regulatory loop mediates similar effects on the activity of UPF1.

A previous study by Cheng *et al*. reported that the RNA-binding affinity of UPF1 is reduced in the presence of ATP and that deletion of four residues (351–354) within the regulatory loop abolishes this effect (21). We sought to determine whether presence of a longer regulatory loop in UPF1 would also mediate a similar effect on its RNA-binding affinity. To address this, we performed fluorescence anisotropy assays with $UPF1_1\Delta$CH and $UPF1_2\Delta$CH using a fluorescein-labeled 26-mer RNA, in the presence and absence of ATP. As previously reported, $UPF1_2\Delta$CH shows a significant (18-fold) reduction in RNA-binding affinity in the presence of ATP while $UPF1_1\Delta$CH shows only a modest 5-fold decrease (Figure 2D, compare cyan and green traces) (21,22). A similar trend was observed when comparing constructs encompassing the CH domain; $UPF1_2$ shows a 16-fold reduction in RNA-binding affinity in the presence of ATP in comparison to the 5-fold decrease exhibited by $UPF1_1$ (Figure 2D, compare purple and yellow traces). However, perturbing the length of the regulatory loop did not affect the length of RNA bound by UPF1 in the presence or absence of nucleotides, as observed in RNase protection assays (Supplementary Figure S2D and data not shown). Our results suggest that deletion or elongation of the regulatory loop modulates the RNA-binding affinity as well as the catalytic activity of UPF1 in a similar manner.

Recently, Fiorini *et al.* reported the behavior of $UPF1_2\Delta$CH in single molecule conditions using a magnetic tweezers approach, and observed that this helicase is able to unwind and translocate slowly over long RNA and DNA substrates with a processivity > 10 kb (34). To assess whether the elongation of the regulatory loop has an impact on the properties of single molecules of UPF1, we compared the unwinding and translocation activities of $UPF1_1\Delta$CH and $UPF1_2\Delta$CH onto a long DNA hairpin. Briefly, the magnetic tweezers setup we used consists of a 1239 bp DNA hairpin tethered to a glass surface through its 3′-end and a biotinylated magnetic bead through its 5′-end inside a microfluidic chamber. A constant force was applied to the ends of the hairpin using a pair of magnets placed above the chamber. The extension of single molecules of DNA inside the chamber was tracked over time using video-microscopy (39). Addition of $UPF1_2\Delta$CH to the chamber in saturating concentrations of ATP led to saw-tooth tracks as described in ref (34) (Figure 3, right panel), during which the helicase slowly unwound

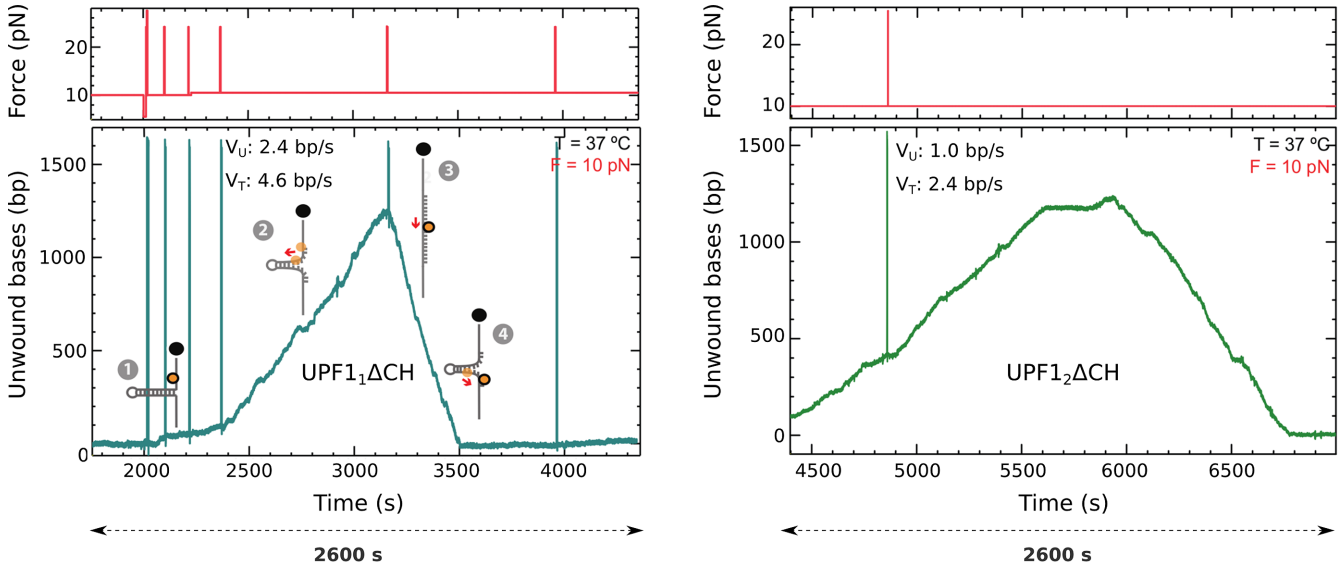


Figure 3. UPF1$_1$ΔCH unwinds nucleic acid with a higher translocation speed than UPF1$_2$ΔCH. Single molecule helicase activity of the two UPF1 isoforms. Experimental traces over 2600 seconds show the activity of UPF1$_1$ΔCH (left) and UPF1$_2$ΔCH (right) in saturating concentration of ATP, at a constant force of 10 pN. Both isoforms unwound the 1239 bp double strand DNA substrate (step 1), reached the apex (2), then translocated over the single strand DNA toward the 3′ end (3) until complete hairpin closure (4). The term V_U refers to the speed of unwinding of the hairpin, while V_T denotes the speed of translocation on single-stranded DNA, after the hairpin has reannealed behind the helicase.

the DNA hairpin (step 1) until it reached the apex (step 2), then translocated toward the 3′-end while the hairpin progressively closed behind it (step 4) (schematic of each step is shown in Figure 3, left panel). These events were interrupted by frequent short pauses, leading to irregular events. In similar conditions, UPF1$_1$ΔCH exhibited a comparable unwinding and translocation activities with a remarkable processivity estimated to more than 10 kb (Figure 3, left panel). Both isoforms also displayed faster speed during the rezipping phase as previously observed for UPF1$_2$ΔCH (34). However, UPF1$_1$ΔCH was two times faster than UPF1$_2$ΔCH with a weighted average rate of 2.4 bp/s during unwinding (1.0 bp/s for UPF1$_2$ΔCH), and 4.6 bp/s during rezipping (2.4 bp/s for UPF1$_2$ΔCH) (Figure 3 and Supplementary Figure S3). This result correlates perfectly with the higher ATP consumption rate of UPF1$_1$ΔCH compared to UPF1$_2$ΔCH (Figure 2C and Supplementary Figure S2C). Therefore, the recording of numerous and independent translocation events of single molecules of UPF1 revealed that its two natural isoforms possess different translocation speeds.

Functional differences between the two UPF1 isoforms can be attributed to the conformation of their regulatory loops

In order to elucidate the mechanistic basis of modulation of catalytic activity by the regulatory loop, we determined the structure of apo-UPF1$_1$ΔCH at 3.3 Å resolution. The final model was refined to an R_{free} of 31.3% and an R_{work} of 29.2% with good stereochemistry (data collection and refinement statistics in Table 1). The domains of UPF1$_1$ΔCH are colored as in Figure 1B, with the exception of the regulatory loop which is in cyan (Figure 4A). Nine of the 22 residues within the regulatory loop are ordered, of which 6 are unique to isoform 1 (Figure 4A, inset). The overall architecture of the helicase core of apo-UPF1$_1$ΔCH is very similar to that of apo-UPF1$_2$ΔCH, with a root mean square deviation (r.m.s.d.) of superposition of 1.1 Å over 96% of the Cα atoms (Figure 4B). There is a slight movement of domain 1B of UPF1$_1$ΔCH by about 10° toward the helicase core. However, despite the overall similarity of the two structures, there are significant differences in their regulatory loops. First, the regulatory loop of UPF1$_2$ΔCH is well ordered and folded into a 3_{10} helix, whereas that of UPF1$_1$ΔCH is partially disordered and in an extended conformation. Additionally, an aspartate residue (Asp_{351}) within the 3_{10} helix interacts with an arginine (Arg_{800}) of the RecA2 domain (Supplementary Figure S4A). As a consequence, the 3_{10} helix of UPF1$_2$ is positioned toward the helicase core. Since the amino acid at an equivalent position in the UPF1$_1$ sequence is a leucine, the interaction of the regulatory loop with the helicase core is lost in this case. Therefore, the extended loop of UPF1$_1$ points away from the helicase core and is oriented toward the solvent (Figure 4B, inset). Our structural observations suggest that the longer regulatory loop of UPF1$_1$ has a higher intrinsic flexibility than that of the regulatory loop of UPF1$_2$.

DISCUSSION

Based on our structural and biochemical observations, we propose a model for the differential catalytic activity of the two isoforms of UPF1. In the more abundant isoform, UPF1$_2$, the short regulatory loop is relatively rigid and collapses onto the RNA binding pocket and partially occludes the 5′ binding site of the RNA. The loop is displaced from this position by a movement of domain 1B and, as a result, is disordered in the transition-state of UPF1. The RNA-binding pocket in apo-UPF1$_2$ΔCH is a relatively open, shallow surface with a distance of approximately 39

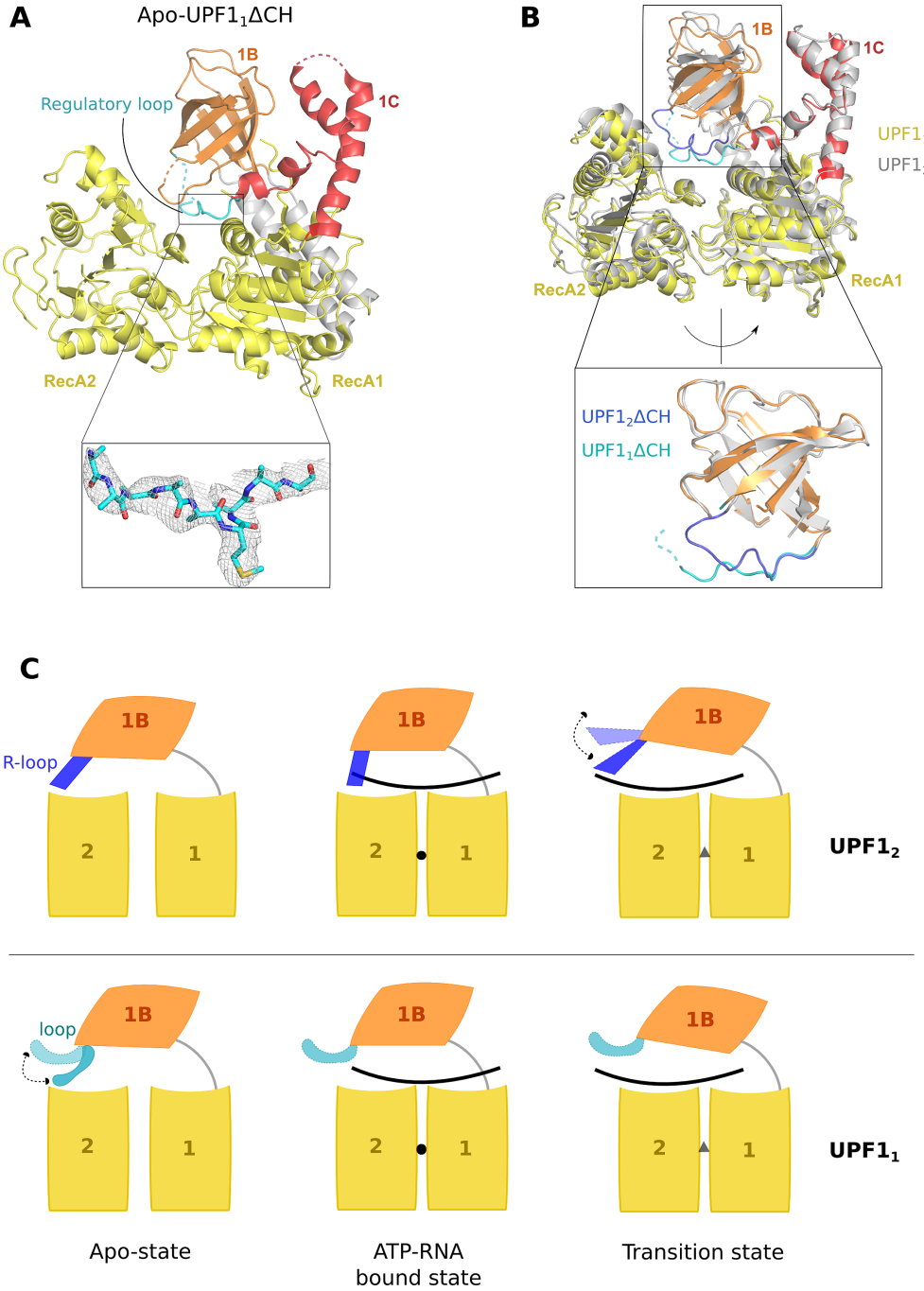


Figure 4. Conformational flexibility of the regulatory loop of UPF1$_1$. (**A**) Crystal structure of human UPF1$_1$ΔCH. The RecA domains, domains 1B and 1C are colored according to the schematic in Figure 1A. The long regulatory loop in cyan lies in an extended conformation close to the RNA binding surface, without significantly blocking RNA binding (Supplementary Figure S4B). The inset shows the $2F_o - F_c$ electron density of the ordered region of the regulatory loop, contoured at 1σ. (**B**) Structural superposition of the RecA1 domains of UPF1$_1$ΔCH (colored yellow, with a cyan regulatory loop) and UPF1$_2$ΔCH (colored gray, with a blue regulatory loop). The loops connecting strands 4 and 5 of domain 1B in both structures have been removed for clarity. The relative orientations of the RecA domains and domain 1C are almost identical, while domain 1B of UPF1$_1$ΔCH undergoes a small movement. Superposition of domains 1B of the two structures shows a considerable divergence in the positions and fold of the regulatory loops (inset). The shorter regulatory loop of UPF1$_2$ΔCH (in blue) is well ordered and folded into a 3_{10} helix which is positioned toward the helicase core while the partially disordered longer loop of UPF1$_1$ΔCH (in cyan) is highly flexible and oriented toward the solvent. (**C**) Model representing the differential effects of the regulatory loop in the two UPF1 isoforms and their impact on catalytic activity of the helicase. The domains are colored as above, while RNA, ATP and the transition-state of the nucleotide are denoted by black lines, black circles and grey triangles, respectively. The short regulatory loop of UPF1$_2$ is more rigid and occupies part of the RNA binding surface in the apo- and AMPPNP-bound states of UPF1. Therefore, it poses as a barrier to translocation of UPF1 upon ATP hydrolysis and has to be displaced from the RNA-binding pocket in each cycle (top panel). In contrast, the longer regulatory loop of UPF1$_1$ is flexible and does not effectively block RNA binding in the presence of ATP or impede translocation (bottom panel). As a consequence, this loop does not have to be removed from the RNA-binding surface each time to allow ATP-dependent translocation of UPF1, resulting in a more active helicase. The CH and 1C domains have been omitted from this model for clarity.

Å between the two edges of the pocket. Upon ATP binding, the pocket is slightly compacted (∼32 Å in width) due to a small movement of the RecA domains toward each other. It should be noted here that, unlike DEAD-box helicases that adopt distinct conformations in their apo- and RNA/ATP bound states, SF1 helicases are maintained in a closed conformation with pre-formed RNA and ATP binding pockets. We reason that the difference in RNA-binding affinity of $UPF1_2$ in the presence and absence of ATP is due to constriction of the RNA-binding pocket, which is further exacerbated by the presence of the regulatory loop (Figure 4C, top left and middle panels). This would explain why deletion of even four of the 11 residues of the inhibitory loop abrogates the difference in RNA-binding affinity in the presence and absence of ATP (21). It is well established that the nucleic acid unwinding and mRNP remodeling activities of UPF1 are a consequence of its ATP-dependent translocation on the nucleic acid (34). Since translocation of UPF1 on RNA in the 5′-3′ direction entails a relative movement of RNA toward the RecA2 domain, it seems likely that the regulatory loop must be forced out of the RNA binding pocket to facilitate this event (Figure 4C, top right panel). We speculate that displacement of the regulatory loop from the RNA-binding pocket might represent the rate-limiting step in the translocation process. As a consequence, permanent removal of the regulatory loop in the UPF1Δloop mutants leads to a stimulation of the catalytic activity of UPF1.

The less abundant isoform, $UPF1_1$, presents a different scenario due to its longer and more flexible regulatory loop. In the apo state, the regulatory loop is mostly positioned away from the RNA binding surface and only marginally blocks RNA binding (Supplementary Figure S4B). We speculate that the regulatory loop also does not inhibit RNA binding in the ATP-bound state, as no drastic reduction in RNA-binding affinity is observed in the $UPF1_1$ isoform in the presence of ATP. We propose that $UPF1_1$ has a higher catalytic activity as the regulatory loop does not impede translocation of the helicase on RNA and does not have to be displaced from the RNA binding pocket in every cycle of ATP hydrolysis (Figure 4C, bottom panels). In this respect, the $UPF1_1$ isoform is similar to the UPF1Δloop protein.

In conclusion, we present here a new mode of regulation of UPF1 activity by a loop in its subdomain 1B. We further demonstrate that the regulatory loop is responsible for the differential catalytic activity of the two UPF1 isoforms. Our study presents a novel mode of regulation of RNA helicases, where catalytic activity is regulated at the post-transcriptional level via subtle structural changes generated by alternative splicing. The existence of two UPF1 splice isoforms with different catalytic activities has interesting implications in the function of this helicase in multiple mRNA degradation pathways. Intriguingly, the more active isoform appears to be less abundant across different mammalian tissues. It would be interesting to determine if the two isoforms target the same set of NMD substrates and are functionally redundant in cells or if each isoform is specific for a certain class of mRNA substrates. Furthermore, it is unclear if a difference in catalytic activity would result in differential NMD efficiency in cells. In a previous study, Nicholson et al. showed that both isoforms of UPF1, when exogenously expressed, are equally competent in rescuing NMD in UPF1-knock out cells (26), but it remains to be seen if this prevails across a wide-range of tissues and cell types. Our studies also suggest that both isoforms of UPF1 are regulated in an identical manner by inter- and intra-molecular interactions with UPF2 and the CH domain, respectively. However, it is not known if additional factors preferentially bind and regulate a specific isoform. The differential regulation of the two isoforms might dictate their usage in different cellular processes. The precise function and regulation of each UPF1 isoform in cells remains a topic for future studies.

DATA AVAILABILITY

The coordinates and structure factors of apo-$UPF1_1$ΔCH have been deposited in the Protein Data Bank (PDB) with the accession code 6EJ5.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR online.

ACKNOWLEDGEMENTS

We thank Elena Conti (MPI of Biochemistry, Martinsried) for support at the initial stages of the project and gratefully acknowledge Florian Heyd and Markus Wahl (Freie Universität Berlin) as well as members of our laboratory for helpful discussions. We are especially grateful to Agnieszka Pietrzyk and Karine Santos for sharing their expertise on stopped-flow experiments and data analysis, Eva Absmeier and Christian Freund for help with fluorescence anisotropy assays, Olga Herdt and Alexander Neumann for analysis of mRNA sequences and RNA-seq data and Bernhard Loll for assistance with X-ray data collection at *BESSY II* (Berlin, Germany). Access to BL14.1 of *BESSY II* was obtained via the Joint Berlin MX laboratory. The mRNA samples from different mouse tissues were a generous gift from Astrid-Solveig Schultz and Florian Heyd. We thank Varun Bhaskar, Atlanta Cook, Felix Halbach and Karine Santos for critically reading the manuscript.

Author contributions: M.G. and S.C. purified the proteins used in this study, performed crystallization trials and determined the structure. M.G., F.B., J.K. and F.F. performed the biochemical experiments. M.G. and V.D.M. carried out the RT-PCR analyses of the UPF1 isoforms. J.K and S.R. performed the single molecule nucleic acid unwinding assays and were supervised by H.L.H. and V.C. S.C. conceptualized the study and wrote the manuscript with input from all the authors.

FUNDING

Deutsche Forschungsgemeinschaft [CH1245/2-1 to S.C.]; Investissements d'Avenir [ANR-10-LABX-54 MEMOLIFE and ANR-10-IDEX-0001–02 PSL* Research University to J.K. and H.L.H.]; Centre National de la Recherche Scientifique, the Ecole Normale Supérieure; Agence Nationale de la Recherche [ANR-13-BSV8-0023 to H.L.H. and ANR-14-CE10-0014 to H.L.H. and V.C.]; Fondation ARC pour la recherche sur le cancer (to J.K.); Max

Planck Gesellschaft (to F.B.). Funding for open access charge: Deutsche Forschungsgemeinschaft [CH1245/2-1]; Freie Universität Berlin.
Conflict of interest statement. None declared.

REFERENCES

1. Lykke-Andersen,J. and Bennett,E.J. (2014) Protecting the proteome: eukaryotic cotranslational quality control pathways. *J. Cell Biol.*, **204**, 467–476.
2. Moore,M.J. (2005) From birth to death: the complex lives of eukaryotic mRNAs. *Science*, **309**, 1514–1518.
3. Stutz,F. and Izaurralde,E. (2003) The interplay of nuclear mRNP assembly, mRNA surveillance and export. *Trends Cell Biol.*, **13**, 319–327.
4. Jankowsky,E. (2011) RNA helicases at work: binding and rearranging. *Trends Biochem. Sci.*, **36**, 19–29.
5. Bleichert,F. and Baserga,S.J. (2007) The long unwinding road of RNA helicases. *Mol. Cell*, **27**, 339–352.
6. Ozgur,S., Buchwald,G., Falk,S., Chakrabarti,S., Prabu,J.R. and Conti,E. (2015) The conformational plasticity of eukaryotic RNA-dependent ATPases. *FEBS J.*, **282**, 850–863.
7. Andersen,C.B., Ballut,L., Johansen,J.S., Chamieh,H., Nielsen,K.H., Oliveira,C.L., Pedersen,J.S., Seraphin,B., Le Hir,H. and Andersen,G.R. (2006) Structure of the exon junction core complex with a trapped DEAD-box ATPase bound to RNA. *Science*, **313**, 1968–1972.
8. Bono,F., Ebert,J., Lorentzen,E. and Conti,E. (2006) The crystal structure of the exon junction complex reveals how it maintains a stable grip on mRNA. *Cell*, **126**, 713–725.
9. Franks,T.M., Singh,G. and Lykke-Andersen,J. (2010) Upf1 ATPase-dependent mRNP disassembly is required for completion of nonsense- mediated mRNA decay. *Cell*, **143**, 938–950.
10. Sengoku,T., Nureki,O., Nakamura,A., Kobayashi,S. and Yokoyama,S. (2006) Structural basis for RNA unwinding by the DEAD-box protein Drosophila Vasa. *Cell*, **125**, 287–300.
11. von Moeller,H., Basquin,C. and Conti,E. (2009) The mRNA export protein DBP5 binds RNA and the cytoplasmic nucleoporin NUP214 in a mutually exclusive manner. *Nat. Struct. Mol. Biol.*, **16**, 247–254.
12. Linder,P. and Jankowsky,E. (2011) From unwinding to clamping - the DEAD box RNA helicase family. *Nat. Rev. Mol. Cell Biol.*, **12**, 505–516.
13. Fairman-Williams,M.E., Guenther,U.P. and Jankowsky,E. (2010) SF1 and SF2 helicases: family matters. *Curr. Opin. Struct. Biol.*, **20**, 313–324.
14. Jankowsky,A., Guenther,U.P. and Jankowsky,E. (2011) The RNA helicase database. *Nucleic Acids Res.*, **39**, D338–D341.
15. Pyle,A.M. (2008) Translocation and unwinding mechanisms of RNA and DNA helicases. *Annu. Rev. Biophys.*, **37**, 317–336.
16. Behm-Ansmant,I., Kashima,I., Rehwinkel,J., Sauliere,J., Wittkopp,N. and Izaurralde,E. (2007) mRNA quality control: an ancient machinery recognizes and degrades mRNAs with nonsense codons. *FEBS Lett.*, **581**, 2845–2853.
17. Lee,S.R., Pratt,G.A., Martinez,F.J., Yeo,G.W. and Lykke-Andersen,J. (2015) Target discrimination in nonsense-mediated mRNA decay requires Upf1 ATPase activity. *Mol. Cell*, **59**, 413–425.
18. 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.
19. Czaplinski,K., Weng,Y., Hagan,K.W. and Peltz,S.W. (1995) Purification and characterization of the Upf1 protein: a factor involved in translation and mRNA degradation. *RNA*, **1**, 610–623.
20. Sun,X., Perlick,H.A., Dietz,H.C. and Maquat,L.E. (1998) 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. U.S.A.*, **95**, 10009–10014.
21. Cheng,Z., Muhlrad,D., Lim,M.K., Parker,R. and Song,H. (2007) Structural and functional insights into the human Upf1 helicase core. *EMBO J.*, **26**, 253–264.
22. Chakrabarti,S., Jayachandran,U., Bonneau,F., Fiorini,F., Basquin,C., Domcke,S., Le Hir,H. and Conti,E. (2011) Molecular mechanisms for the RNA-dependent ATPase activity of Upf1 and its regulation by Upf2. *Mol. Cell*, **41**, 693–703.
23. Clerici,M., Mourao,A., Gutsche,I., Gehring,N.H., Hentze,M.W., Kulozik,A., Kadlec,J., Sattler,M. and Cusack,S. (2009) Unusual bipartite mode of interaction between the nonsense-mediated decay factors, UPF1 and UPF2. *EMBO J.*, **28**, 2293–2306.
24. Chamieh,H., Ballut,L., Bonneau,F. and Le Hir,H. (2008) 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.
25. Fiorini,F., Boudvillain,M. and Le Hir,H. (2013) Tight intramolecular regulation of the human Upf1 helicase by its N- and C-terminal domains. *Nucleic Acids Res.*, **41**, 2404–2415.
26. Nicholson,P., Josi,C., Kurosawa,H., Yamashita,A. and Muhlemann,O. (2014) A novel phosphorylation-independent interaction between SMG6 and UPF1 is essential for human NMD. *Nucleic Acids Res.*, **42**, 9217–9235.
27. Kabsch,W. (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 125–132.
28. McCoy,A.J., Grosse-Kunstleve,R.W., Adams,P.D., Winn,M.D., Storoni,L.C. and Read,R.J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.*, **40**, 658–674.
29. Afonine,P.V., Grosse-Kunstleve,R.W. and Adams,P.D. (2005) The Phenix refinement framework. *CCP4 Newslett.*, **42**, http://www.ccp4.ac.uk/newsletters/newsletter42/content.html.
30. Emsley,P. and Cowtan,K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. Section D, Biol. Crystallogr.*, **60**, 2126–2132.
31. Davis,I.W., Leaver-Fay,A., Chen,V.B., Block,J.N., Kapral,G.J., Wang,X., Murray,L.W., Arendall,W.B. 3rd, Snoeyink,J., Richardson,J.S. *et al.* (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.*, **35**, W375–W383.
32. Ryder,S.P., Recht,M.I. and Williamson,J.R. (2008) Quantitative analysis of protein-RNA interactions by gel mobility shift. *Methods Mol. Biol.*, **488**, 99–115.
33. Fiorini,F., Bonneau,F. and Le Hir,H. (2012) Biochemical characterization of the RNA helicase UPF1 involved in nonsense-mediated mRNA decay. *Methods Enzymol.*, **511**, 255–274.
34. Fiorini,F., Bagchi,D., Le Hir,H. and Croquette,V. (2015) Human Upf1 is a highly processive RNA helicase and translocase with RNP remodelling activities. *Nat. Commun.*, **6**, 7581.
35. Manosas,M., Spiering,M.M., Zhuang,Z., Benkovic,S.J. and Croquette,V. (2009) Coupling DNA unwinding activity with primer synthesis in the bacteriophage T4 primosome. *Nat. Chem. Biol.*, **5**, 904–912.
36. Kabsch,W. and Sander,C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, **22**, 2577–2637.
37. Saikrishnan,K., Griffiths,S.P., Cook,N., Court,R. and Wigley,D.B. (2008) DNA binding to RecD: role of the 1B domain in SF1B helicase activity. *EMBO J.*, **27**, 2222–2229.
38. Buttner,K., Nehring,S. and Hopfner,K.P. (2007) Structural basis for DNA duplex separation by a superfamily-2 helicase. *Nat. Struct. Mol. Biol.*, **14**, 647–652.
39. Hodeib,S., Raj,S., Manosas,M., Zhang,W., Bagchi,D., Ducos,B., Fiorini,F., Kanaan,J., Le Hir,H., Allemand,J.F. *et al.* (2017) A mechanistic study of helicases with magnetic traps. *Protein Sci.*, **26**, 1314–1336.