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

Manjeera Gowravaram1, Fabien Bonneau2, Joanne Kanaan3, Vincent D. Maciej1, Francesca Fiorini3, Saurabh Raj3,4, Vincent Croquette3,4, Hervé Le Hir3 and Sutapa Chakrabarti1,*

1Institute of Chemistry and Biochemistry, Freie Universität Berlin, Takustr. 6, D-14195 Berlin, Germany, 2Max Planck Institute of Biochemistry, Structural Cell Biology Department, Am Klopferspitz 18, D-82152 Martinsried, Germany, 3IBENS, Département de Biologie, École Normale Supérieure, CNRS, Inserm, PSL Research University, F-75005 Paris, France and 4LPS, 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 (UPF11) is 11 residues longer than that of isoform 2. We find that this small insertion in UPF11 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 UPF11 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: chakrabar@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.

© The Author(s) 2018. Published by Oxford University Press on behalf of Nucleic Acids Research. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com
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 RNA-RNA hybrids vary depending on their sub-family—DEAD box helicases of the SF2 superfamily act by local strand-separation of RNA duplexes, whereas DEAH 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 DEAH helicases translocate in the 5′→3′ direction, RNA helicases of the SF1 superfamily translocate in the 3′→5′ 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 its 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 (UPF1Δ1) contains an insertion in domain 1B which extends its regulatory loop to 22 residues, while isoform 2 (UPF1Δ2, 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 UPF1Δ2). Although the UPF1Δ2 isoform has been extensively characterized using structural and biochemical tools, very little is known about the catalytic activity and regulation of the isoform UPF1Δ1. 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 UPF1Δ1 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 UPF1Δ1 and UPF1Δ2ΔCH were purified as previously described (22). The UPF1Δ1loop and UPF1Δ1ΔCH Δloop mutants were generated by one-step PCR amplification, using pET28a-UPF1Δ1 and pET15b-UPF1Δ2ΔCH as templates and the following primers: 5’-GAATTCGCTAATACGTGTTCTGGGAAGTCGGC-3’ and 5’-CCCTGCATGAGCCGACTCTGAGGG-3’ and 5’-CCCTGCATGAGCCGACTCTGAGGG-3’ and 5’-CCCTGCATGAGCCGACTCTGAGGG-3’ and 5’-CCCTGCATGAGCCGACTCTGAGGG-3’ and 5’-CCCTGCATGAGCCGACTCTGAGGG-3’. The UPF1Δ1 and UPF1Δ1ΔCH plasmids were also generated as above, using pET28a-UPF1Δ1 and pET15b-UPF1Δ2ΔCH as a template and the following primers: 5’-GAATTCGCTAATACGTGTTCTGGGAAGTCGGC-3’ and 5’-CCCTGCATGAGCCGACTCTGAGGG-3’. Both UPF1Δ1 and UPF1Δ1ΔCH 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 Ni2+-affinity and ion-exchange (IEX) chromatography in a manner identical to UPF1Δ1, 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 UPF1Δ1loop protein used for stopped-flow analysis was additionally subjected to

Downloaded from https://academic.oup.com/nar/article-abstract/46/5/2648/4824839 by MPI Biochemistry user on 13 November 2018
Figure 1. A conserved loop in domain 1B of UPF1 modulates its catalytic activity. (A) Schematic representation of the domain organization of human UPF1 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 (UPF12). The position and the residues of the UPF1 regulatory loop are indicated. Amino acids in bold denote those replaced by the GSGS linker in
The UPF1ΔCH and UPF1ΔCHΔloop proteins were purified in a manner identical to UPF1ΔCH.

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 UPF1ΔCH 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 N2 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 UPF1ΔCH 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-mercaptop-7-methylpurine ribonucleoside) and 0.5 U purine-nucleoside phosphatase (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-mercaptop-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 ΔA360 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′-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 MgCl2, 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 KD between UPF11 and UPF12 or between UPF1ΔCH and UPF1ΔCH, in the absence and presence of ATP. The significance threshold was set at 0.05.

Stopped flow assay

The kinetics of interaction of UPF11, UPF12 and UPF1Δloop 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 (Kx 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 MgCl2 and 1 μM ZnCl2 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 U15 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).

Supplementary Information (Supplementary Figure S1).
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′-GGGACCAGAUGCGCGUAUUGAGUUUGA AUUUUCAUGAUGGUACGAUCAGUCAACGGU ACC CGAGAAGCUGCGGGUACC-3′) was hybridized to a 21-mer DNA (5′-GGAGCTTTCCACGCCCATGGAUUAUAUUAUCGAUGGUAUCAGAUCUGGAUCCU-3′) to yield a RNA:DNA hybrid with a 54-nucleotide overhang. The duplex was labeled with 32P at the 5′-end of the DNA strand. Reactions were carried out at 30°C under conditions described in (33), with 22.5 fmol of the labeled duplex and 4.5 pmol 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 and UPF12 isoforms, respectively. The primer sequences are as follows: forward 5′-GGGACCTGGGCTTAAACAAGAAGAGA-3′ and reverse 5′-ATCCCCCTTCACAGGGCGCAAGGT-3′. Plasmids expressing the human UPF1ΔCH and UPF1ΔCH proteins were used as positive controls. For quantification of mRNA levels of the two isoforms low-cycle PCRs were performed with the 32P-labeled forward primer (sequence given above), products were resolved by denaturing-PAGE and analyzed using a Phosphorimager 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′-ATCCCCCTTCACAGGGCGCACGAT-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. UPF1ΔCH 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 k (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 UPF12 isoform have been determined. These include structures of a human UPF12 construct encompassing the CH and helicase domains (referred to as UPF1 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:AlF4−) (22). Additionally, structures of the human UPF1 helicase core (referred to as UPF1ΔCH) 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 UPF1ΔCH and observed that, in addition to lobe closure between the two RecA domains, domain IB undergoes considerable movement upon nucleotide and RNA binding while domain IC 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-UPF1₂ΔCH, part of this loop folds into an over-wound S₁₀ 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 UPF1₂ΔCH 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 UPF1₂ 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 UPF1₂, UPF1₂ΔCH and UPF1Δloop proteins in the presence of poly-(U) RNA. As reported earlier, UPF1₂ 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 UPF1₂, albeit to a lower extent than that for UPF1₂ΔCH (Figure 1C, compare yellow and blue traces).

Previous studies show that the affinity of UPF1₂ for RNA is reduced in the presence of ATP and that deletion of residues 351–355 or 352–354 within the regulatory loop abolishes 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 U₁₅ RNA. A time-course for the binding of mant-ATPγS to UPF1₂, 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 UPF1₂ 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 UPF1₂ and UPF1Δloop proteins, suggesting that although the regulatory loop participates in ATP-dependent 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 UPF1₂ΔCH 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 UPF1₂, 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 (UPF1₂) is the more abundant one (26), the relative amounts of the longer isoform (UPF1₁) 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 UPF1₂ 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 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 UPF1ΔCHΔloop 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. 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 and UPF1ΔCHΔloop have been compared in the absence and presence of ATP 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 $K_d$s 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 and UPF1ΔCH 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 UPF1ΔCH and UPF1ΔCHΔloop in the presence of ATP is extremely significant (**$**P < 0.001$).
Table 1. X-ray crystallographic data collection and refinement statistics

<table>
<thead>
<tr>
<th>Data collection</th>
<th>Beamsline</th>
<th>BESSEY II BL 14.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelength (Å)</td>
<td>0.918</td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>P2₁ 3</td>
<td></td>
</tr>
<tr>
<td>Unit cell parameters (Å)</td>
<td>a = 146.9</td>
<td></td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>19.62–3.34</td>
<td></td>
</tr>
<tr>
<td>Unique reflections</td>
<td>15 480 (1512)</td>
<td></td>
</tr>
<tr>
<td>Multiplicity</td>
<td>67.7 (6.9)</td>
<td></td>
</tr>
<tr>
<td>Completeness</td>
<td>99.2 (98.9)</td>
<td></td>
</tr>
<tr>
<td>Mean I/σ(I)</td>
<td>13.9 (2.3)</td>
<td></td>
</tr>
<tr>
<td>R-merge</td>
<td>0.099 (0.802)</td>
<td></td>
</tr>
<tr>
<td>CC₁/₂</td>
<td>0.999 (0.757)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Refinement</th>
<th>Reflections used for R_free</th>
<th>777</th>
</tr>
</thead>
<tbody>
<tr>
<td>R_free (%)</td>
<td>31.3</td>
<td></td>
</tr>
<tr>
<td>R_work (%)</td>
<td>29.2</td>
<td></td>
</tr>
<tr>
<td>RMS bonds (Å)</td>
<td>0.003</td>
<td></td>
</tr>
<tr>
<td>RMS angles (°)</td>
<td>0.76</td>
<td></td>
</tr>
<tr>
<td>Average B factor (Å²)</td>
<td>111.92</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ramachandran values</th>
<th>Favoured</th>
<th>96.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allowed</td>
<td>4.0</td>
<td></td>
</tr>
<tr>
<td>Outliers</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

*Values within parentheses correspond to the highest resolution shell.

in the case of UPF1₂, repression of UPF1₁ 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 UPF1₁.

Contrary to our expectations of stronger inhibition of catalytic activity by the long regulatory loop, we find that activated UPF1₁ shows significantly higher ATPase activity than activated UPF1₂ (Figure 2C, compare first two columns in the middle and right stacks). Furthermore, the ATPase activity of UPF₁₁ΔCH is identical to that of the UPF₁₁ΔCHΔloop 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 UPF1₁.

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 UPF₁₁ΔCH and UPF₁₂ΔCH using a fluorescein-labeled 26-mer RNA, in the presence and absence of ATP. As previously reported, UPF₁₁ΔCH shows a significant (18-fold) reduction in RNA-binding affinity in the presence of ATP while UPF₁₁ΔCH 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; UPF₁₁ shows a 16-fold reduction in RNA-binding affinity in the presence of ATP in comparison to the 5-fold decrease exhibited by UPF₁₁ (Figure 2D, compare purple and yellow traces). However, perturbing the length of the regulatory loop did not affect the length of RNA bound by UPF₁₁ 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 UPF₁₁ in a similar manner.

Recently, Fiorini et al. reported the behavior of UPF₁₂ΔCH 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 UPF₁₁, we compared the unwinding and translocation activities of UPF₁₁ΔCH and UPF₁₂ΔCH 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 UPF₁₂ΔCH 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. UPF1ΔCH unwinds nucleic acid with a higher translocation speed than UPF12ΔCH. Single molecule helicase activity of the two UPF1 isoforms. Experimental traces over 2600 seconds show the activity of UPF1ΔCH (left) and UPF12ΔCH (right) in saturating concentration of ATP, at a constant force of 10 pN. Both isoforms unwind 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, UPF1ΔCH 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 UPF12ΔCH (34). However, UPF1ΔCH was two times faster than UPF12ΔCH with a weighted average rate of 2.4 bp/s during unwinding (1.0 bp/s for UPF12ΔCH), and 4.6 bp/s during rezipping (2.4 bp/s for UPF12ΔCH) (Figure 3 and Supplementary Figure S3). This result correlates perfectly with the higher ATP consumption rate of UPF1ΔCH compared to UPF12ΔCH (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-UPF1ΔCH 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 UPF1ΔCH 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-UPF1ΔCH is very similar to that of apo-UPF12ΔCH, with a root mean square deviation (r.m.s.d.) of superposition of 1.1 Å over 96% of the Ca atoms (Figure 4B). There is a slight movement of domain 1B of UPF1ΔCH 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 UPF12ΔCH is well ordered and folded into a 310 helix, whereas that of UPF1ΔCH is partially disordered and in an extended conformation. Additionally, an aspartate residue (Asp351) within the 310 helix interacts with an arginine (Arg800) of the RecA2 domain (Supplementary Figure S4A). As a consequence, the 310 helix of UPF12 is positioned toward the helicase core. Since the amino acid at an equivalent position in the UPF11 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 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 UPF11 has a higher intrinsic flexibility than that of the regulatory loop of UPF12.

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, UPF12, 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-UPF12ΔCH is a relatively open, shallow surface with a distance of approximately 39
Figure 4. Conformational flexibility of the regulatory loop of UPF11. (A) Crystal structure of human UPF11ΔCH. 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 2F₀ − Fᶜ electron density of the ordered region of the regulatory loop, contoured at 1. (B) Structural superposition of the RecA1 domains of UPF11ΔCH (colored yellow, with a cyan regulatory loop) and UPF12ΔCH (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 UPF11ΔCH 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 UPF12 (in blue) is well ordered and folded into a 3₁₀ helix which is positioned toward the helicase core while the partially disordered longer loop of UPF11ΔCH (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 UPF12 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 UPF11 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.
A 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 UPF1Δ2 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, UPF11, 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 UPF11 isoform in the presence of ATP. We propose that UPF1 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 UPF11 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-knockout 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-UPF1ΔCH have been deposited in the Protein Data Bank (PDB) with the accession code 6EJS.

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 MEMO-LIFE 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.]; Fondsation 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


