The Ighmbp2 helicase structure reveals the molecular basis for disease-causing mutations in DMSA1

Siew Choo Lim[1,2], Matthew W. Bowler[3], Ting Feng Lai[1,2] and Haiwei Song[1,4,5,*]

[1]Institute of Molecular and Cell Biology, Proteos, Singapore 138673, [2]School of Biological Sciences, Nanyang Technological University, Singapore 637551, Singapore, [3]Structural Biology Group, European Synchrotron Radiation Facility, F-38043 Grenoble, France, [4]Life Sciences Institute, Zhejiang University, Hangzhou, China and [5]Department of Biochemistry, National University of Singapore, Singapore 117543, Singapore

ABSTRACT

Mutations in immunoglobulin μ-binding protein 2 (Ighmbp2) cause distal spinal muscular atrophy type 1 (DSMA1), an autosomal recessive disease that is clinically characterized by distal limb weakness and respiratory distress. However, despite extensive studies, the mechanism of disease-causing mutations remains elusive. Here we report the crystal structures of the Ighmbp2 helicase core with and without bound RNA. The structures show that the overall fold of Ighmbp2 is very similar to that of Upf1, a key helicase involved in nonsense-mediated mRNA decay. Similar to Upf1, domains 1B and 1C of Ighmbp2 undergo large conformational changes in response to RNA binding, rotating 30° and 10°, respectively. The RNA binding and ATPase activities of Ighmbp2 are further enhanced by the R3H domain, located just downstream of the helicase core. Mapping of the pathogenic mutations of DSMA1 onto the helicase core structure provides a molecular basis for understanding the disease-causing consequences of Ighmbp2 mutations.

INTRODUCTION

Distal spinal muscular atrophy type 1 (DSMA1) is an autosomal recessive disorder resulting from the degeneration of α-motoneurons in the spinal cord. It manifests itself in early childhood and is characterized by early respiratory failure due to diaphragmatic paralysis (1). Although clinical and pathological features are similar to spinal muscular atrophy (SMA), DSMA1 and SMA are genetically distinct. SMA is caused by mutations of the survival of motor neurons (SMN) gene coding for SMN protein, a member of the SMN complex (2–5). DSMA1 is caused by genetic mutations in the *Ighmbp2* gene, which encodes immunoglobulin μ-binding protein 2 (Ighmbp2) (6,7). The precise cellular function of Ighmbp2 remains elusive, although it has been implicated in transcription and pre-mRNA processing (8,9). More recently, Ighmbp2 has been reported to associate with ribosomes and tRNAs, suggesting that it is functionally linked to translation (10,11).

Ighmbp2 is a multidomain protein composed of a DNA/RNA helicase domain and an R3H domain and a zinc-finger domain (Figure 1A). Based on sequence homology, Ighmbp2 has been classified as a member of the Upf1-like group within the helicase superfamily 1 (SF1) (12). The other two members of the Upf1-like family of SF1 helicases are Upf1, a key protein involved in nonsense-mediated mRNA decay (13–15), and Senataxin, implicated in transcription regulation (16–18). Biochemical characterization has shown that Ighmbp2 is an ATP-dependent $5' \rightarrow 3'$ DNA/RNA helicase (10). Most of the pathogenic mutations in DSMA1 patients occurred in the helicase domain, suggesting that this domain plays a major role in the DSMA1 disease-causing mechanism. Consistent with this notion, the pathogenic mutants of Ighmbp2 exhibited loss of either ATPase or helicase activity or both activities (10), arguing that the biochemical defects of Ighmbp2 may be the primary causes underlying DSMA1.

SF1 helicases can be divided into two classes, SF1A and SF1B, on the basis of the direction of translocation, with SF1B further branched to Pif1-like (Pif1, RecD2) and Upf1-like subfamily (Upf1, Ighmbp2, Senataxin) (19). The Upf1-like family differs from the Pif1-like family as it can unwind both DNA and RNA duplexes (20). A great deal is known about the mechanisms of SF1A enzymes
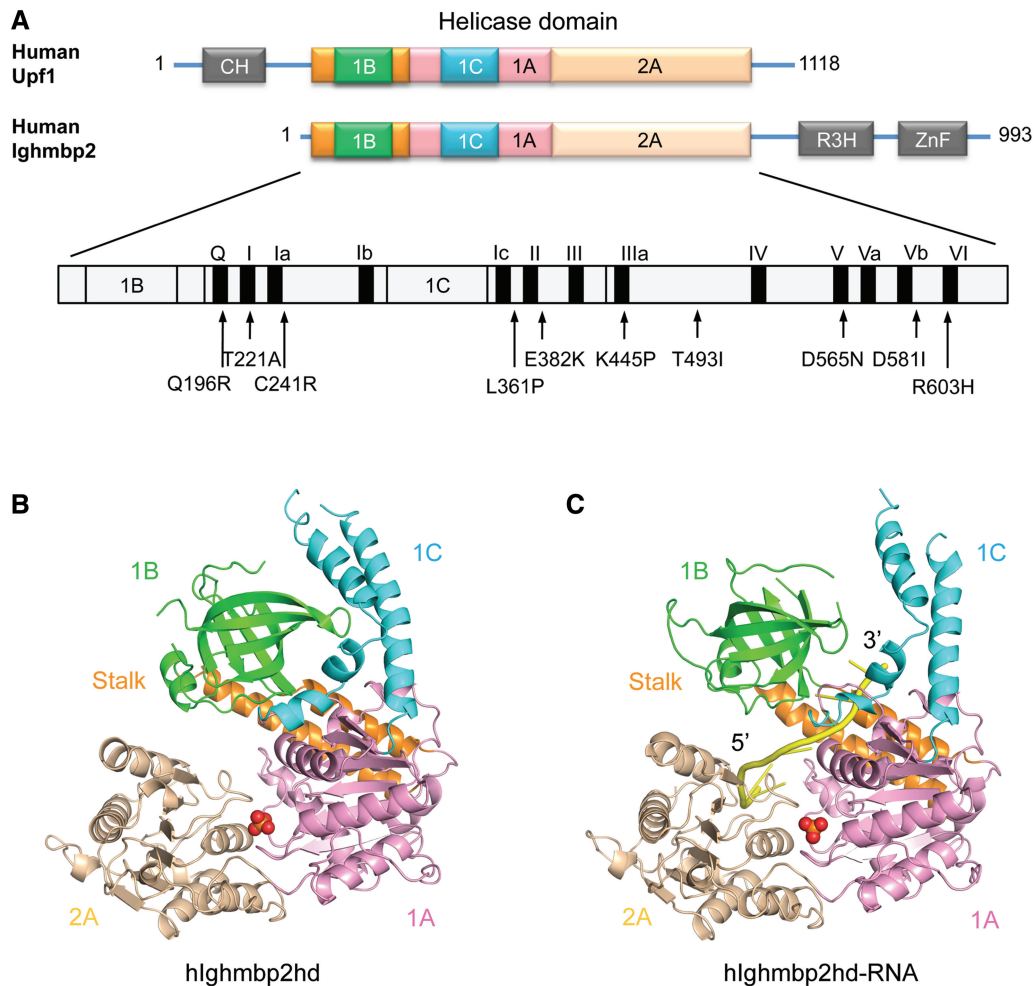
Figure 1. Structural overview of hIghmbp2hd and hIghmbp2hd–RNA. (**A**) Schematic representation of the domain arrangement in Upf1 and Ighmbp2. The helicase core region contains two RecA-like domains: domain 1A (pink) and domain 2A (wheat) and additional regulatory domains: domain 1B (green), domain 1C (cyan) and the 'stalk' (orange). Upf1 contains another regulatory domain, the CH domain (gray) located upstream of the helicase region while Ighmbp2 features two more another regulatory domains: a R3H domain (gray) and an AN-1 type zinc-finger domain (gray) downstream of the helicase region. Missense mutations and canonical motifs in Ighmbp2 are labeled. (**B**) Crystal structures of hIghmbp2hd. (**C**) Crystal structure of hIghmbp2hd–RNA. The coloring schemes for domains are as in (A). The bound phosphate ion and ssRNA are shown as red sphere and yellow cartoon tube, respectively.

(from PcrA, Rep and UvrD) (21–26); however, only two SF1B helicase structures (Upf1 and RecD2) are available. The structures of RecD2 have been determined in three states: the free form, in complex with single-stranded DNA (ssDNA) and in the presence and absence of an ATP analog (27,28). The mechanistic basis for $5' \rightarrow 3'$ translocation by the Pif1-like family has been revealed from the structural characterization of the RecD2 enzyme (28). For the understanding of the mechanism by which Upf1-like helicase functions, several crystal structures of Upf1 in different functional states have been determined. These include Upf1 in complex with either phosphate or ADP or the non-hydrolyzable ATP analog, AMPPNP (29), as well as in transition-state with bound single-stranded RNA (ssRNA) (30), but no structure of the RNA-bound Upf1 in the absence of a nucleotide has been reported.

Here we have determined the crystal structure of the helicase core domain of human Ighmbp2 with and without bound ssRNA in the absence of nucleotide.

These structures allow a molecular interpretation of a number of pathogenic mutations that cause DSMA1. Moreover, structural comparison combined with biochemical characterization reveal the critical role of the R3H domain in modulation of enzymatic and RNA-binding activities of Ighmbp2.

MATERIALS AND METHODS

Protein expression and purification

The cDNA encoding the helicase domain of human Ighmbp2 (hIghmbp2hd, residues 1–652) was cloned into the pGEX6p-1 vector (GE Healthcare) with a GST tag at the N-terminus. The hIghmbp2hd protein was expressed in *Escherichia coli* BL21 Rosetta (DE3) pLysS strain (Novagen). Cells were grown in Luria–Bertani medium supplemented with 100 µg/ml ampicillin at 37°C until the optical density at 600 nm (OD600) reached ~0.6. Protein expression was induced by the addition of 0.4 mM

isopropyl-ß-D-1-thiogalactopyranoside and further cell growth was performed at 18°C for 16 h. Cells were harvested by centrifugation at 4000 rpm and resuspended in lysis buffer (25 mM HEPES, 500 mM NaCl, 3 mM DTT, 3 mM MgCl$_2$, pH 7.3) supplemented with 1 mg/ml lysozyme from chicken egg-white (Sigma). Cells were disrupted by sonication on ice and cell debris was cleared by centrifugation at 18 000 rpm. The clarified supernatant was loaded into a column self-packed with Glutathione Sepharose 4B beads (GE Healthcare), pre-equilibrated with the lysis buffer. The column was washed with the lysis buffer and the protein was eluted with elution buffer (25 mM HEPES, 500 mM NaCl, 3 mM DTT, 3 mM MgCl$_2$, 1% w/v reduced glutathione, pH 7.3). The GST tag is then cleaved by incubating with PreScission protease overnight. The protein sample was then further purified by a second glutathione sepharose column and gel filtration using HiPrep Superdex-75 column on an AKTA Xpress system (GE Healthcare) in lysis buffer. hIghmbp2hd was pooled and concentrated to 12 mg/ml. We failed to obtain SeMet-substituted protein as it became insoluble after PreScission protease cleavage (data not shown).

mIghmbp2hd-R3H, a construct containing both helicase and R3H domains (residues 1–786) of mouse Ighmbp2, was cloned and purified in the same way except a Superdex-200 gel filtration column was used in gel filtration.

Crystallization

In the initial crystallization screening, 5 mM potassium phosphate was added to the protein. Plate-like crystals were obtained at 15°C in a reservoir solution containing 0.2 M sodium malonate pH 7.0 and 20% (w/v) PEG 3350. Mercury-derivative crystals were obtained by soaking native crystals in the reservoir solution supplemented with 0.1 mM EMTS (ethyl mercury thiosalicylate) for 6 h. Iodide-derivative crystals were obtained by quick soaking the crystals in the reservoir solution containing 500 mM potassium iodide for 5 s.

To crystallize the hIghmbp2hd–RNA complex, a single-stranded A$_{10}$ RNA was mixed with hIghmbp2hd at a molar ratio of 2:1. Crystals were obtained by hanging drop method with equal volumes of hIghmbp2hd–RNA solution and reservoir solution containing 0.1 M imidazole pH 8.0, 1.0 M sodium citrate at 15°C. All crystals were cryoprotected with 25% ethylene glycol before being flash-cooled in liquid nitrogen.

Data collection and structure determination

Native data sets of the hIghmbp2hd and hIghmbp2hd–RNA complexes were collected at 100K at beamlines ID14-4 and ID23-1 at the European Synchrotron Facility (ESRF, Grenoble, France). The best diffraction volumes of the crystals were defined using diffraction cartography (31). For the hIghmbp2hd–RNA complex crystals the *c* cell edge was aligned parallel to the spindle using a mini-kappa goniometer (32). This allowed better separation of Bragg spots due to the length of this axis. Mercury and iodide derivative data sets were collected at

100K at beamline 13B1 at the National Synchrotron Radiation Research Centre (NSRRC, Taiwan, R.O.C).

The structure of hIghmbp2hd was determined by the multiple isomorphous replacement with anomalous scattering (MIRAS) method in the space group *C*2. The data sets were integrated using MOSFLM (33). Scaling of intensities was carried out using SCALA from the CCP4i package (34). Heavy atom search was carried out by Autosol in the PHENIX software suite (35). Three mercury and 17 iodide sites, corresponding to one molecule of hIghmbp2hd in the asymmetric unit, were found. Automatic partial model building was carried out by AutoBuild in the PHENIX software suite (35). The model was built manually using COOT (36). Crystallographic refinement was performed using the PHENIX software suite (35) to a final R_{free} of 25.23%. TLS refinement was introduced and a free phosphate group was included at the final stages of refinement.

The crystals of hIghmbp2hd–RNA belong to space group $P3_121$, with two molecules in the asymmetric unit. The data for hIghmbp2hd–RNA were integrated with XDS (37,38), then scaled and reduced using SCALA in CCP4i. The structure was determined by molecular replacement with PHASER (39), using the hIghmbp2hd structure as the search model. A good match for domains 1A and 2A was found. Domain 1B was manually built with the aid of $2F_o$–F_c and F_o–F_c maps using COOT. Model refinement was carried out with the PHENIX software suite to a final R_{free} of 29.28%. RNA molecules were included in the final stages of refinement. Difference Fourier maps clearly showed the electron density for nine and eight ribonucleotides in molecules A and B, respectively (Supplementary Figure S1). As there is no substantial difference between the two molecules in the ASU (r.m.s.d. of 0.85 Å for all equivalent Cα atoms of hIghmbp2hd), all subsequent analyses are based on molecule A as it is more complete. The stereochemical geometry of the structures was validated using PROCHECK (40). The *Ramachandran* plot showed that one outlier (residue 580) is in hIghmbp2hd, and three outliers of molecule A (residues 82, 270 and 580) and one outlier of molecule B (residue 580) are in hIghmbp2hd–RNA. All these residues have poorly-defined electron density except for residue 270 which has relatively well defined electron density and is located at a loop region. The statistics for data collection and refinement and the quality of the final models are summarized in Table 1.

Fluorescence anisotropy

Fluorescence anisotropy measurements were performed with single-stranded U$_{15}$ RNA labeled at the 5′-end with 6-carboxy-fluorescein (6-FAM) (Metabion International AG). 100 µl reactions were carried out in a 96-well plate at 25°C and measured using Safire II fluorescent plate reader (Tecan). 100 nM of the labeled oligonucleotide was incubated with increasing protein concentration at ambient temperature for 30 min prior to measurements, in a buffer containing 20 mM HEPES, 2 mM MgCl$_2$, 2 mM DTT, 100 mM NaCl, 10% glycerol and pH 7.4.

Table 1. Data collection and refinement statistics

Data collection	hIghmbp2hd	hIghmbp2hd–Hg	hIghmbp2hd–Iodide	hIghmbp2hd–RNA
Derivative	–	Mercury	Iodide	–
Wavelength (Å)	0.9795	0.99315	1.54	0.9760
Resolution limit (Å)	2.5	3.0	3.0	2.85
Space group	C2	C2	C2	$P3_121$
Cell parameters				
a/b/c (Å)	116.57/76.72 88.54	115.83/76.65/87.29	117.67/76.67/87.97	87.29/87.29/372.69
$\alpha/\beta/\gamma$ (°)	90/107.32/90	90/107.40/90	90/107.72/90	90/90/120
Unique reflections (N)	25 840	14 640	15006	39 419
Total reflections (N)	114 847	95 240	95 220	222 197
I/σ	16.0 (1.4)	13.0 (2.2)	8.4 (1.8)	21.6 (1.5)
Mean (I/σ)	18.9 (2.3)	24.4 (3.6)	26.0 (3.5)	12.5 (1.9)
Completeness (%)	91.8 (99.7)	99.9 (99.7)	99.8 (99.7)	90.2 (99.5)
R_{merge} [a]	0.04 (0.512)	0.044 (0.341)	0.065 (0.418)	0.026 (0.551)
Number of derivative sites	–	3	17	–
Refinement statistics				
Data range (Å)	63.16–2.50			19.97–2.85
Used reflections (N)	23 987			38 694
Protein residues	625			1222
Protein atoms	4853			18 460
Solvent molecules	99			140
Phosphate molecules	1			2
RNA bases	–			17
R_{work} [b] (%)	19.02			20.58
R_{free} [c] (%)	25.23			29.28
R.m.s.d.				
Bond length (Å)	0.008			0.009
Bond angles (°)	1.181			1.338
Ramchandran plot (% residues)				
Allowed	98.5			97.2
Generously allowed	1.2			2.3
Disallowed [d]	0.2			0.5

Values in parentheses indicate the specific values in the highest resolution shell.
[a]$R_{merge} = \sum |I_j - <I>|/\sum I_j$, where I_j is the intensity of an individual reflection and $<I>$ is the average intensity of that reflection.
[b]$R_{work} = \sum ||F_o| - |F_c||/\sum |F_c|$, where F_o denotes the observed structure factor amplitude and F_c denotes the structure factor amplitude calculated from the model.
[c]R_{free} is as for R_{work} but calculated with 5.0% of randomly chosen reflections omitted from the refinement.
[d]Disallowed: Val580 in hIghmbp2hd and Ser82 (A), Arg270 (A), Val580 (A) and Val580 (B) in hIghmbp2hd–RNA.

The excitation and emission wavelengths were 470 and 535 nm, respectively. All anisotropy measurements were performed in triplicate. Each titration point was measured using three reads with an integration time of 40 µs. Dissociation constants (K_d) values were determined by fitting the experimental data to a binding equation describing a single-site binding model (shown below) using Prism software 4.0 (GraphPad).

$$A = A_0 + (A_f - A_0)$$
$$\times \frac{([L]_t + [P]_t + K_d) - \sqrt{([L]_t + [P]_t + K_d)^2 - 4[L]_t[[P]_t}}{2[L]_t}$$

where A is the anisotropy measured, A_0 is the initial anisotropy in the absence of protein, A_f is the final anisotropy, $[P]$ is the protein concentration, $[L]$ is the labeled RNA concentration

In vitro ATPase assay

The ATPase assay was performed in 50 µl reaction volume containing 40 pmol of purified Ighmbp2 proteins (hIghmbp2hd or mIghmbp2hd-R3H) in 50 mM MES (pH 6.0), 50 mM KAc, 2.5 mM Mg(Ac)$_2$, 5 µg of 15-mer poly(A), 30 µM cold ATP and 0.3 µCi [γ^{32}P]ATP. Upon incubation at 37°C for 1 h, 5 µl of 500 mM EDTA was added to terminate the reaction. In order to identify the ATPase hydrolysis products, a standard reaction with the same substrates and conditions was performed using 0.1 U of Calf Intestine Phosphatase (CIP) at 37°C for 6 min. Aliquots (10 µl) of the reaction products were analyzed by thin-layer chromatography on a polyethyleneimine-cellulose plate (Sigma) and developing it in 0.3 M K$_2$HPO$_4$ (pH 7.6).

RESULTS

Overall structure

The crystal structure of the helicase domain of human Ighmbp2 (residues 1–652, designated as hIghmbp2hd) was solved by the MIRAS method using mercury and iodide derivatives to a resolution of 2.5 Å. Residues 73–75, 120–126, 310–315, 503–507 and 649–652 were not visible in the electron density map and assumed to be disordered in the structure. The helicase core of Ighmbp2 has essentially the same fold as that of the Upf1

protein (29,30), which contains four domains: two RecA-like domains (domains 1A and 2A) with two subdomains 1B and 1C inserted into domain 1A (Figure 1B). Domains 1A (residues 159–270 and 347–440) and 2A (residues 441–648) share the α–β fold with a central seven-stranded parallel β-sheet surrounded by α helices, resembling the RecA-like fold seen in all SF1 and SF2 helicases (20,41). Domain 1B (residues 34–141) consisting of a short α-helix and a β-barrel, is connected to domain 1A by two α-helices (residues 3–33 and 142–158) also referred to as the 'stalk' (30). Domain 1C (residues 271–346) is inserted between the Walker A and Walker B motifs and forms four helices situated above domain 1A (Figure 1B). The canonical motifs I, II, III, V and VI known from other SF1 helicases to be involved in ATP binding [reviewed in (20)] and motifs III and Va known to coordinate between NTP and nucleic acid-binding sites (19), are present in the equivalent positions in hIghmbp2hd (Figure 1).

To gain insights into how the helicase interacts with nucleic acid, we solved the crystal structure of hIghmbp2hd in complex with a ssRNA (thereafter referred to as hIghmbp2hd–RNA). Residues 32–36, 94–97, 121–124, 308–320, 502–506 and 649–652 were disordered in the structure. Although a 10-nt ssRNA (A_{10}) was used in crystallization, we only observed nine adenine bases in the hIghmbp2hd–RNA complex. For simplicity, we have numbered nucleotides A_1 to A_9 in a $5' \rightarrow 3'$ direction. The ssRNA binds in an extended conformation with the 5'-end at domain 2A and the 3'-end pointing towards the channel between domains 1B and 1C (Figure 1C). Both the structures of hIghmbp2hd and hIghmbp2hd–RNA contain a phosphate situated in the nucleotide-binding cleft between domains 1A and 2A.

RNA recognition

The ssRNA binds to a composite RNA-binding channel formed by domains 1A, 2A, 1B and 1C (Figure 1C). The backbone phosphates of ssRNA mainly contact domains 1A and 2A with the adenine bases pointing either to the solvent region or towards domain 1B. Domains 1B and 1C directly interact with the 3'-end of RNA, similar to that observed in the Upf1-RNA-bound structure with the presence of the CH domain (30). Protein–RNA interactions were mediated mainly by a combination of hydrogen bonds and electrostatic interactions as depicted in Figure 2. Briefly, A_1 is well ordered by stacking its adenine base with the side chain of His411. The adenine bases A_1 to A_3 are in parallel orientation to each other at an average distance of $\sim$3.7 Å, a common feature of base stacking in ssRNA (42). As the adenine base of A_3 is perpendicular to that of A_4, it protrudes outwards causing a bend in the RNA backbone between A_3 and A_4. The side chain of residue Gln507 forms hydrogen bonding with the O2' atom of the ribose sugar of base A_3 while the O2' atom of the ribose sugar of A_4 interacts with the side chains of Asp565 and Arg596. The hydroxyl group of residue Thr407 forms a hydrogen bond with the N3 atom of the nitrogenous base of A_5 which in turn forms base stacking interactions with A_6 to A_8. Motif

Ia interacts with the phosphate backbone of A_7 via the side chain of Asn245. Residues Thr351 and Thr353 from motif Ic interact with the phosphate backbone of A_7 and A_8, respectively. A second bend of the RNA backbone is observed whereby A_9, located $\sim$10 Å from A_8, flips away from the base stacking configuration. Arg270 makes two important contacts with the RNA at this position. First, its NH_2 group interacts with the N7 atom of the adenine base of A_8. Second, its NH_1 group forms electrostatic interactions with the O1P atom of the phosphodiester backbone of A_9. Since only residues Gln507, Asp565 and Arg596 interact with the ribose 2'-OH groups of A3 and A4 (Figure 2) through polar interactions, the ability of Ighmbp2 to unwind both DNA and RNA duplexes (10) could be explained by the lack of specific recognition of 2'-OH group of the sugar ribose ring that renders Ighmbp2 unspecific to DNA or RNA.

RNA binding induced conformational changes

Superposition of domains 1A and 2A of hIghmbp2hd and hIghmbp2hd–RNA showed little change in the positions of Cα atoms (r.m.s.d. of 0.87 Å, Figure 3A), indicating that the relative orientations of domains 1A and 2A are largely conserved in the presence and absence of RNA. In contrast to this observation, large conformational changes are observed in domains 1B and 1C. When RNA is bound, domain 1B rotates $\sim$30° away from domain 1C such that it bends towards domain 2A to form a channel (diameter of $\sim$15 Å and length $\sim$50 Å) that could accommodate a single stranded but not a duplex nucleic acid (Figure 3A). This marked conformational change widens the channel between domains 1B and 1C where the 3'-end of ssRNA is situated. The opening of the channel between 1B and 1C also requires the outwards tilting of domain 1C by 10°. In addition to these inter-domain movements, a prominent intra-domain conformational change occurs, which involves the reorganization of a loop (residues 264–273 in domain 1A, located just before domain 1C). This results in reorientation of the side chain of Arg270 from the nucleotide-binding pocket to the RNA-binding pocket (between bases A_8 and A_9) upon ssRNA binding (Figure 3B). These conformational changes demonstrate the conformational flexibility of Ighmbp2, similar to that observed in Upf1 (30).

Structural comparison of Ighmbp2 with Upf1

Superposition of the structure of hIghmbp2hd (Figure 4A) with that of hUpf1ΔCH-PO$_4^-$ (Figure 4B) showed that these two structures are very similar. The best fitting lies in domains 1A and 2A and the largest change in orientation occurs in domain 1B. Domain 1B of hIghmbp2hd interacts with domain 1C but has no contacts with domain 2A. In contrast, domain 1B in hUpf1ΔCH-PO$_4^-$ has no contacts with domain 1C but interacts with domain 2A.

Comparison of hIghmbp2hd–RNA (Figure 4C) with the RNA-bound Upf1 structures (30) showed that the conformation of domain 1B in Ighmbp2 is drastically different from those in Upf1 in response to RNA binding. In Ighmbp2, upon RNA binding, domain 1B rotates $\sim$30° towards domain 2A, therefore forming a composite
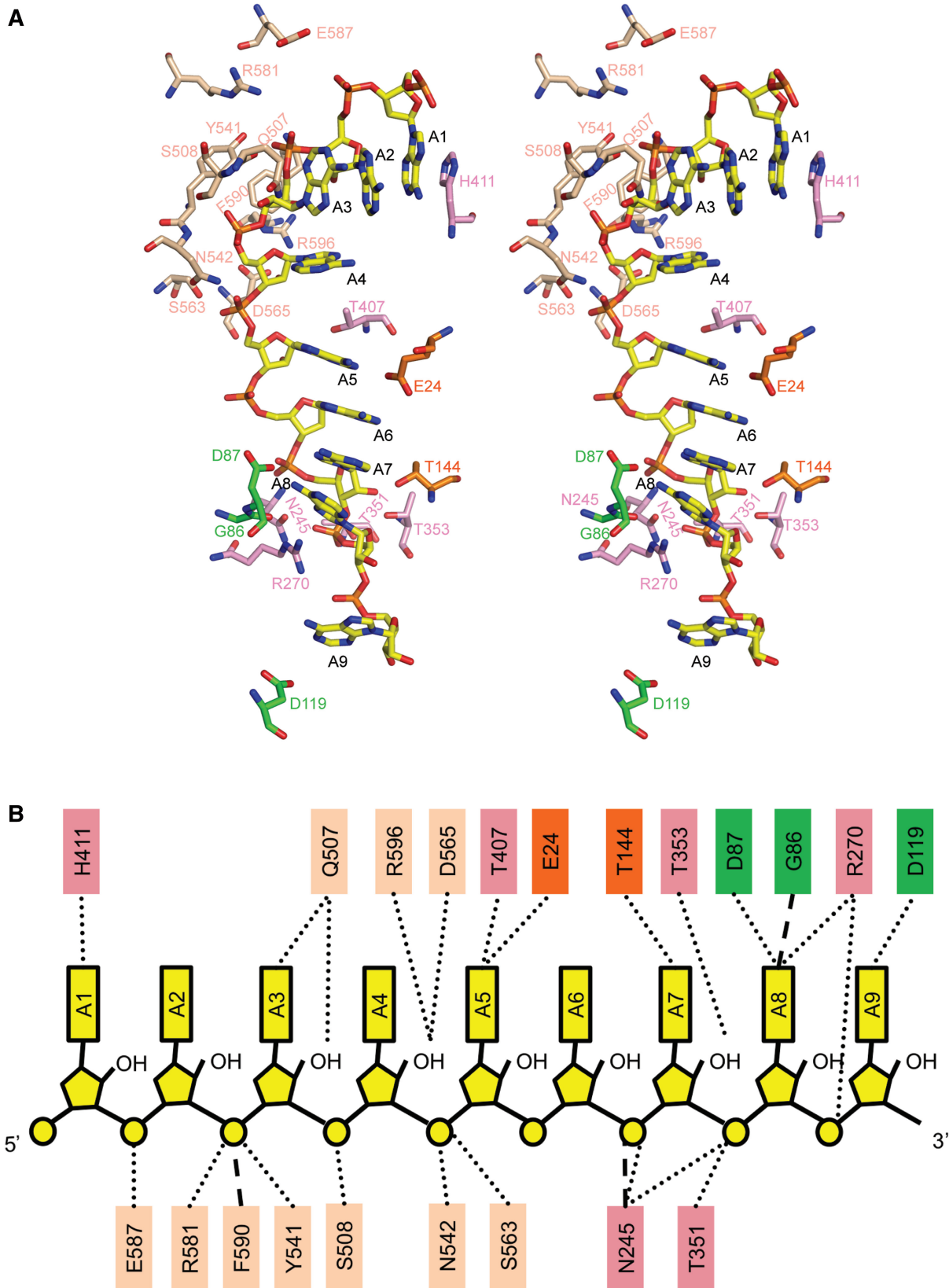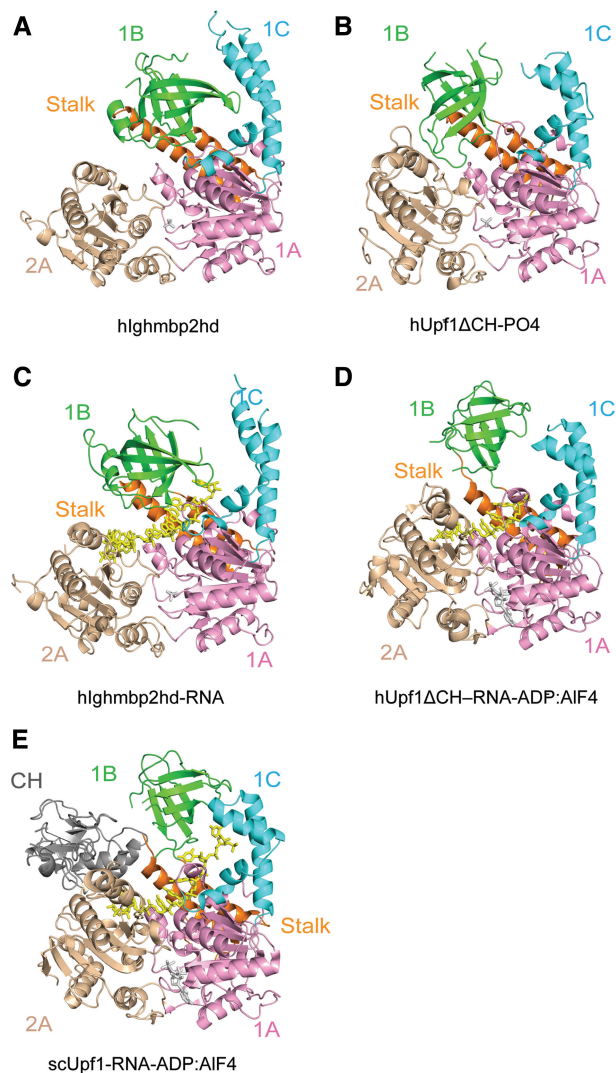
Figure 2. Protein–RNA interactions. (**A**) Stereo view showing the ssRNA-binding site in hIghmbp2hd–RNA. The bound ssRNA and the residues involved protein–RNA interactions are shown in ball-stick models. The color scheme is as in Figure 1. (**B**) Schematic representation of the contacts between the hIghmbp2hd protein and RNA.
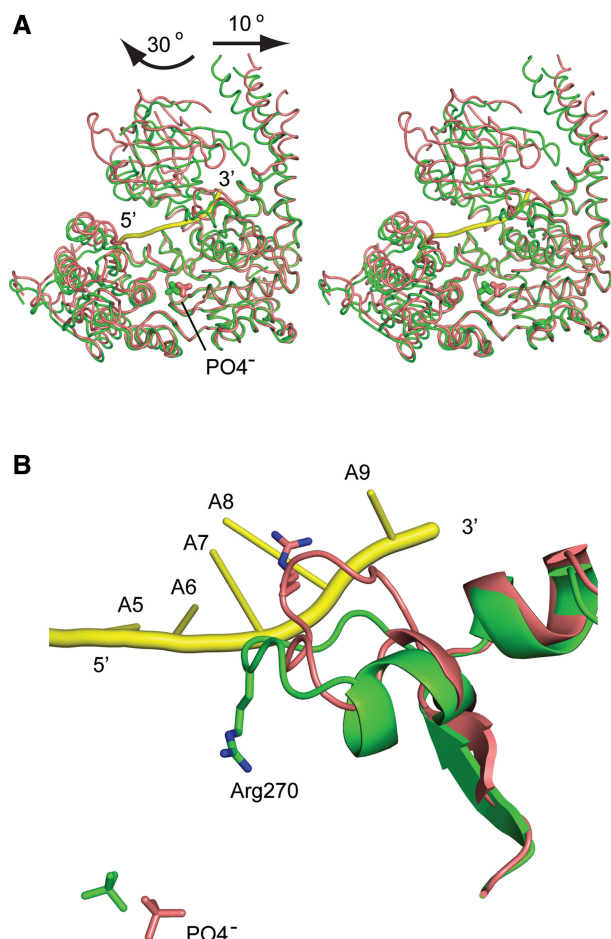
Figure 3. Conformational changes of hIghmbp2hd upon RNA binding. (**A**) Stereo view showing the superposition of hIghmbp2hd–RNA (pink) with hIghmbp2hd (green). Phosphate ions are shown as sticks and the bound ssRNA is shown in yellow cartoon. The arrows indicate the movement of domain 1B and 1C in the RNA-bound state compared to the RNA-free state. (**B**) Conformational change of a loop region (residues 264–273) and the reorientation of residue Arg270 in response to ssRNA binding. The position of the phosphate ions indicates the location of the nucleotide-binding site.

Figure 4. Structural comparison of Ighmbp2 with Upf1. The ribbon diagrams are drawn with domains 1A (pink) and 2A (wheat) in the same orientation. The bound phosphate ion and nucleotide are shown in gray sticks. ssRNA is show in yellow sticks. The coloring scheme for domains are as in Figure 1. (**A**) hIghmbp2hd. (**B**) hUpf1ΔCH-PO$_4^-$ (PDB code: 2GK7). (**C**) hIghmbp2hd–RNA. (**D**) hUpf1ΔCH-RNA-ADP:AlF$_4^-$ (PDB code 2XZO). (**E**) Yeast Upf1-RNA-ADP:AlF$_4^-$ (PDB code: 2XZL).

RNA-binding channel together with domains 1A, 2A and 1C (Figure 4C versus A). In contrast, upon RNA binding, domain 1B in the structure of human Upf1ΔCH-RNA-ADP:AlF$_4^-$ rotates $\sim$35° and moves away from the helicase core, thereby relieving its steric hindrance imposed on domain 2A but making no contacts with RNA (30) (Figures 4D versus B). In addition to this conformational change, domain 1B in the structure of yeast Upf1-RNA-ADP:AlF$_4^-$ shifts by $\sim$20° toward domain 1C, thereby narrowing the channel between them (Figure 4E versus D). In contrast to the large conformational change of domain 1B, domain 1C in hIghmbp2hd–RNA tilts only 10° away from domain 1B with respect to its position in the RNA-free structure, whereas no conformational change of domain 1C has been observed in the Upf1 structures with and without RNA (Figures 3A and 4).

Another notable difference lies in the RNA-binding sites. In the structure of hUpf1ΔCH-RNA-ADP:AlF$_4^-$

(Figure 4D), domains 1A and 2A form an extended surface to bind RNA while domain 1B changes from an inhibitory conformation seen in hUpf1ΔCH-PO$_4^-$ (Figure 4B) to a conformation that neither hinders nor promotes RNA binding. Consequently, Upf1 lacking the CH domain (human Upf1ΔCH-RNA-ADP:AlF$_4^-$) could only bind to a ssRNA of 6 nt and domains 1B and 1C have no contribution to RNA binding. In the presence of CH domain (yeast Upf1-RNA-ADP:AlF$_4^-$), domain 1B is pushed towards domain 1C and the RNA-binding channel of Upf1 is extended to accommodate eight bases with domains 1B and 1C directly interacting with RNA (30) (Figure 4E). In this conformation, Upf1 binds more tightly to the 3′-end of the RNA, 'clamping' it and reduces its helicase activity (30). Ighmbp2 lacks the CH domain

but has a R3H domain and a zinc-finger domain located downstream of the helicase core domain. Without these two accessory domains, a 9-nt ssRNA is bound to Ighmbp2hd in the absence of a nucleotide (Figure 4C) because all four domains contribute to the formation of the composite RNA-binding channel. One simple explanation for the ability of Ighmbp2hd to bind longer RNA without the contribution of the accessory domain is that domain 1B in Ighmbp2hd is not in an inhibitory conformation (no contact with domain 2A but interacts with domain 1C), and a simple 30° rotation coupled with the outward movement of domain 1C is sufficient to allow RNA binding. Given that helicases make use of accessory domains for their specific and regulatory functions and that the R3H domain has been proposed to bind nucleic acids (9), it is envisaged that the R3H domain might have a role in regulating Ighmbp2's RNA-binding and catalytic activities.

R3H domain has a regulatory role on Ighmbp2

The R3H domain is a conserved sequence motif characterized by an invariant arginine residue and a highly conserved histidine residue that are separated by three residues (43). The solution structure of the R3H domain from human Ighmbp2 (PDB:1MSZ) showed that the R3H domain consists of three-stranded anti-parallel β-sheet and two α-helices and has a positively charged surface (44). To examine the functional role of R3H domain in Ighmbp2 protein, we expressed and purified the protein fragment composed of the helicase core and the R3H domain from mouse mIghmbp2 (mIghmbp2hd-R3H) since the human protein containing the corresponding regions degraded severely during purification (data not shown). To investigate the contribution of R3H domain to RNA binding, we carried out fluorescence anisotropy experiments to determine the binding affinities of the proteins (hIghmbp2hd and mIghmbp2hd-R3H) to RNA. In these experiments, the tightest RNA affinity was observed with the mIghmbp2hd-R3H in the absence of any nucleotides with a K_d value of 107 ± 29 nM (Figure 5A). In comparison, hIghmbp2hd interacted ∼10-fold less tightly with a K_d value of 1077 ± 130 nM (Figure 5A). These strikingly different K_d values strongly suggest that R3H domains have a role in promoting nucleic acid binding, consistent with the results from Fukita and coworkers (9). Similar to Upf1, both hIghmbp2hd and mIghmbp2hd-R3H showed higher RNA-binding affinity in the absence of a nucleotide (Figure 5A) (30,45,46). Although the presence of nucleotides did lower RNA-binding affinity, it was not significant (Figure 5A). The difference in RNA binding could be due to conformational changes associated with nucleotide binding. Hence, the variation in the RNA-binding affinities of Ighmbp2 is mainly contributed by the R3H domain.

We hypothesize that tighter RNA binding will lead to enhanced ATPase activity as Ighmbp2 is a nucleic acid-dependent ATPase (10). ATPase assays in the absence and presence of nucleic acids as shown in Figure 5B (lanes 3–6) indicated that ATP hydrolysis is stimulated by the

presence of ssRNA. Importantly, mIghmbp2hd-R3H exhibits 4-fold stronger ATPase activity compared to hIghmbp2hd in the presence of ssRNA (Figure 5B, lanes 4 and 6), suggesting R3H domain stimulates ATPase activity through enhancing RNA binding. Altogether, our results reveal the synergy between the R3H domain and the helicase domain as demonstrated by enhanced RNA binding which in turn potentiates its ATPase activity.

Structural and functional placement of DSMA1 missense mutations on Ighmbp2

Of all the disease-causing missense-mutations in *IGHMBP2*, which are located within or adjacent to the hecliase domain of Ighmbp2 (7,47–52), only nine representative mutants have been biochemically characterized (10). We mapped these nine missense mutants onto the hIghmbp2hd–RNA structure (Figure 6A) and classified them according to their locations: residing in or close to the nucleotide-binding pocket (class I), in RNA-binding channel (class II) and in neither ligand-binding sites (class III).

Class I mutants comprise Q96R, T221A, H445P and R603H. Gln196 is the invariant residue in the Q motif, a motif previously proposed to be unique to the DEAD box family of helicases to control ATP binding and hydrolysis (53,54). When AMPPNP from the Upf1 structure is superimposed onto Ighmbp2 nucleotide-binding pocket, residue Gln196 is shown to make a Van der Waals (VDW) contact with the adenine base of AMPPNP with the shortest distance of 3.6 Å (Figure 6B). The mutation of Gln196 to an arginine would result in a longer side chain protruding into the nucleotide-binding pocket. This would prevent ATP binding, thereby leading to loss of the ATPase activity.

Residue Thr221 is part of the Walker A sequence in motif I and could form hydrogen bonds with the α- and β-phosphate groups of modeled AMPPNP (Figure 6B) and also coordinate the catalytic magnesium ion. Motif I is well characterized to be important for nucleotide binding and hydrolysis (55,56). Hence mutation of Thr221 to alanine will inevitably impact the ATPase activity.

Residue His445 is well conserved in mouse, human and hamster Ighmbp2. It is located at one end of motif IIIa which is the loop linking domains 1A and 2A (Figure 6A and B). Since the movement of motor domains 1A and 2A relative to each other in response to ATP binding/hydrolysis is critical for the SF1 helicases [reviewed in (20)], this loop might act as a hinge to regulate the conformational change between domains 1A and 2A. Mutation of His445 to a proline would reduce the flexibility of this loop, thus restraining the movement of motor domains 1A and 2A and affect the ATPase activity.

Arg603 lies in motif VI (Figure 5), which has been shown to function in ATP hydrolysis and RNA binding (55,57). From the hIghmbp2hd–RNA structure, Arg603 forms electrostatic interactions with the γ-phosphate of AMPPNP. The equivalent residue in PcrA has been proposed as an 'arginine finger' to stabilize the transition

A

B
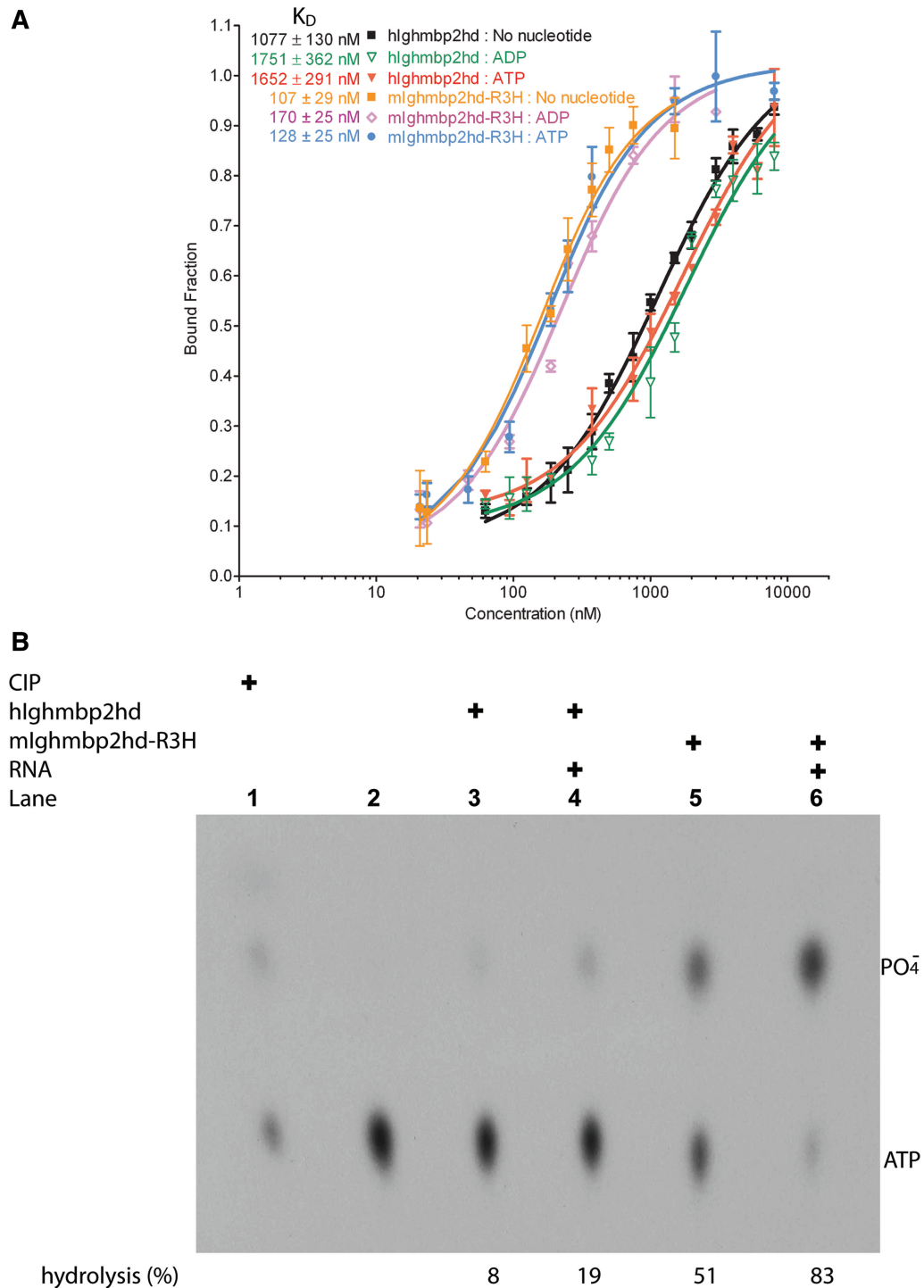
Figure 5. RNA-binding and ATPase activities of Ighmbp2. (**A**) Quantitative measurements of RNA-binding affinities of hIghmbp2hd and mIghmbp2hd-R3H in solution by fluorescence anisotropy. Dissociation constants (K_D) values were determined by fitting the experimental data to a binding equation describing a single-site binding model. The K_D values and their corresponding errors are the mean and standard deviation of three independent experiments. (**B**) ATPase activities of hIghmbp2hd and hIghmbp2hd-R3H in the absence or presence of ssRNA. The position of phosphate was determined by treating the substrate with CIP.

state (21,55). Moreover, it was reported that mutating the corresponding conserved arginine residue in eIF4A would abolish its RNA binding and ATPase and helicase activities (57). In hIghmbp2, mutation of Arg603 to histidine would render this residue unable to interact with the nucleotide, hence affecting the ATPase activity of the enzyme.

The Class II mutants consist of D565N and N583I. Asp565 lies in motif V, which is a flexible loop in domain 2A (Figure 6C). Although the function of motif
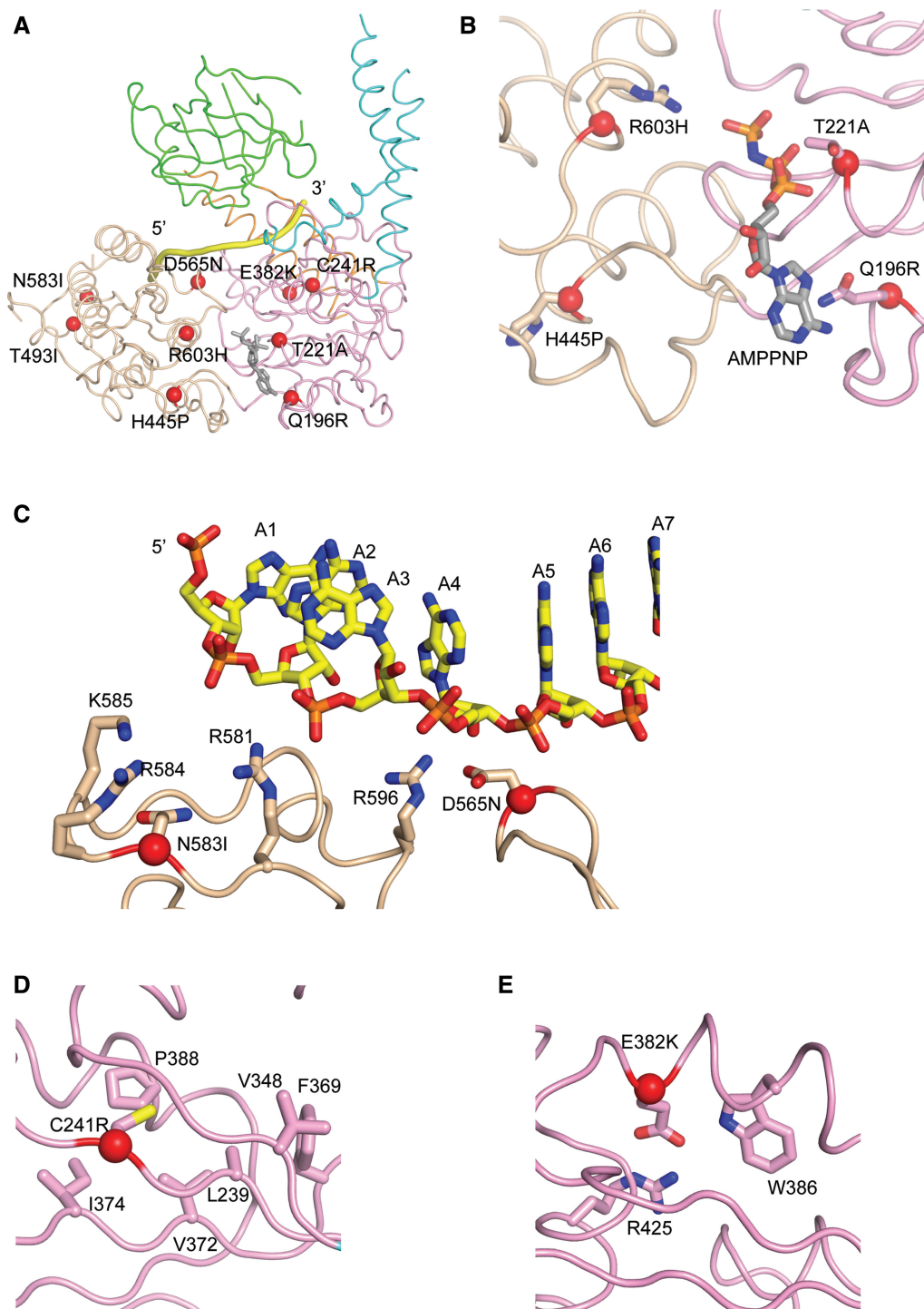
Figure 6. Structural roles of residues associated with disease-causing missense mutations in SMARD1. (**A**) Disease-causing missense mutations in DSMA1 are mapped in the structure of hIghmbp2hd–RNA with the Cα atoms of the mutated residues shown as red spheres. AMPPNP (in stick model) is modeled by superposition of the structure of hIghmbp2hd–RNA with that of human Upf1ΔCH-AMPPNP (PBD code: 2GJK). (**B**) Class I mutations residing in or close to the nucleotide-binding pocket. (**C**) Class II mutations located in the RNA-binding channel. (**D**) and (**E**) Class III mutations located far away from either the nucleotide-binding pocket or the RNA-binding channel.

V is less well studied, there is evidence showing that that it is involved in RNA binding (56) and the translocation activity of hepatitis C virus NS3 helicase (58). The hIghmbp2hd–RNA structure shows that the side-chain of Asp565 contacts the ribose O2′ of A₄ via VDW interactions (Figures 2 and 6C) and forms a hydrogen bond with Arg596, which in turn contacts the ribose group of A₄. Residues Asp565 and Arg596 correspond to Asp827 and Arg858 in human Upf1, which show similar interactions with the RNA backbone (30).

Substitution of Asp565 by Asn would not extensively change the interactions of this residue with the RNA backbone, therefore not affecting the RNA binding to Ighmbp2. Asn583 is located near the 5′-end of the RNA-binding channel but Asn583 does not directly interact with the bound RNA (Figure 6C). We observed that Asn583 is surrounded by several basic residues Arg581, Lys585 and Arg584. Mutation of Asn583 to an isoleucine would place a hydrophobic residue into a positively-charged groove formed by these three residues, thus affecting the local charge distribution and reducing RNA binding.

Class III mutants include C241R, E382K and T493I. Cys241 is located at the central β-sheet of domain 1A and surrounded by several hydrophobic residues (Leu239, Val348, Phe369, Val372, Ile374 and Pro388) (Figure 6D). Therefore, mutation of Cys241 to a positively-charged arginine would most likely disrupt the central β-sheet, hence destabilizing domain 1A.

Residue Glu382 lies on the helix just before motif II. It forms a hydrogen bond with Trp386 in the same helix and also forms a salt-bridge interaction with Arg425 in a neighboring helix, all in domain 1A (Figure 6E). The mutation of Glu382 to a lysine (which has a longer and positively charged side chain) would result in charge-charge repulsion with Arg425 and also could disrupt the hydrogen bond formed between Glu382 and Trp386. Thus it is likely that this mutation would reduce the stability of domain 1A.

Mutant T493I is in a loop region, away from the nucleotide-binding site and is not located in proximity to any helicase motifs (Figure 6A). In addition, the side chain of Thr493 does not make significant contacts with neighboring residues. Based on our current structures, it is difficult to predict how the T493I mutation would affect the conformation and functionality of Ighmbp2.

The mutational effects predicted based on structural analysis are consistent with the previous biochemical data (10). All class I mutants (Q196R, T221A, H445P and R603H) and one of the class II mutants (N583I) and two of the class III mutants (C241R, E382K) exhibited defective ATPase and helicase activities (10). Interestingly, class II mutant D565N retains the ability to bind nucleic acid and has an ATPase activity similar to the wild type but does lose its helicase activity. As unwinding of a nucleic acid duplex requires translocation, the defective helicase activity of this mutant is probably due to its inability to translocate along the RNA substrate. The only mutant with minimal effect on Ighmbp2's catalytic activities is T493I in class III. It has been reported that the T493I mutant has severely reduced steady-state protein level in DSMA1 patient cell lines (50). Therefore, the pathogenicity of this mutant is most likely due to an effect on Ighmbp2 protein levels, rather than its enzymatic activity. Taken together, our structural data rationalized the biochemical defects observed for the disease-causing mutations of DSMA1 and strengthen the notion that the loss of the helicase activity of Ighmbp2 is the primary pathological basis of DSMA1.

DISCUSSION

Previous studies have shown that the Ighmbp2 protein belongs to the SF1B helicase family based on its sequence alignment and the ability to unwind both RNA and DNA duplex in the 5′→3′ direction (10). Indeed, our crystal structure of hIghmbp2hd reveals that the helicase domain of Ighmbp2 is structurally similar to that of Upf1 and both have similar nucleotide and nucleic acid-binding sites. Comparison of the RNA-free and the RNA-bound states of hIghmbp2hd demonstrates that ssRNA binding induces conformational changes in domains 1B and 1C. The importance of these domains was also demonstrated by the Upf1 helicase, whereby conformation changes are observed in domain 1B upon nucleic acid binding (30) and deletion of domain 1C abolished nucleic acid binding (29). These observations indicate that members of the Upf1-like family have a similar domain fold and their mechanistic actions could be very similar.

Many helicases have N-terminal or C-terminal accessory domains. These domains usually have defined folds and can influence the specific function of the helicase. For instance, the N-terminal cysteine–histidine-rich domain (CH-domain, residues 115–275) of Upf1 have been shown to play a regulatory role on the helicase activity of Upf1 by modulating RNA-binding kinetics and the length of the RNA-binding channel through the interaction with Upf2 (30,59). In comparison, Ighmbp2 has an accessory C-terminal R3H domain and an AN1-type zinc-finger motif. Previous biochemical results have argued that the R3H domain of Ighmbp2 is crucial to RNA binding (9). It is clear from our data that although the helicase domain alone was able to bind RNA and has basal ATPase activity, these activities are augmented by the R3H domain. Given that the CH domain of Upf1 extends the RNA-binding channel and enables it to bind longer RNA, the R3H might act in analogy to the CH domain of Upf1 to extend the RNA-binding channel in Ighmbp2, therefore increasing RNA binding. Consistent with this notion, it has been shown that the R3H domain alone is not sufficient for high-affinity RNA binding and that the presence of both the helicase and R3H domains would enhance RNA binding (9). This would suggest that the helicase and R3H domains work in a cooperative manner to enhance RNA binding.

The structure of the Ighmbp2 helicase provides an opportunity to address how genetic defects in patients with DSMA1 impact the molecular function of the protein. Available information indicates specific mutations do not correlate with the severity of the clinical features (47,60). Our structural information, together with the previous biochemical data (10), indicate that seven out of the nine representative missense mutations affect ATPase activity, either through disrupting ATP binding/hydrolysis or through reducing the structural stability of the helicase motors. The two exceptional cases are the mutations T493I and D565N, which do not affect the ATPase activity of the enzyme. Instead, the mutation T493I causes neuronal degeneration through reducing the intracellular Ighmbp2 protein levels while the D565N mutation probably affects the ability of Ighmbp2

translocation along the bound RNA, therefore leading to its enzymatic defect. Our structure-based functional annotation of disease mutations in Ighmbp2 can facilitate further experimental studies to explore the pathology of SMARD1.

There has been an increasing interest in the study of the senataxin protein encoded by the human *SETX* gene, known to be involved in the neurodegenerative diseases ALS4 and AOA2 (16,17,61,62). Senataxin has a C-terminal helicase domain homologous to the helicase domains of Ighmbp2 and Upf1. Mutations in the helicase domain of the *SETX* gene show defects in transcription, RNA processing and DNA damage repair (17,62,63). Recently, Senataxin has been elucidated to function as an helicase that resolves RNA/DNA hybrids so that Xrn2-dependent termination can occur (18). Although senataxin and Ighmbp2 share a homologous helicase domain, the mutations in the helicase domain of senataxin and the disease phenotypes generated by these mutations are different from those of Ighmbp2, suggesting that these two proteins have different and specific cellular functions, presumably determined by their accessory domains outside the helicase domains. Clearly, more studies are required to reveal the exact molecular function of Ighmbp2. These studies, in conjunction with those on senataxin, could help to understand the patho-mechanisms of DSMA1, ALS4 and AOA2.

ACCESSION NUMBERS

The coordinates and structure-factor amplitudes of hIghmbp2hd and hIghmbp2hd–RNA have been deposited in the Protein Data Bank with accession codes 4B3F and 4B3G, respectively.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Figures 1–2.

REFERENCES

1. Mellins,R.B. (1974) Pulmonary physiotherapy in the pediatric age group. *Am. Rev. Respir. Dis.*, **110**, 137–142.
2. Lefebvre,S., Burglen,L., Reboullet,S., Clermont,O., Burlet,P., Viollet,L., Benichou,B., Cruaud,C., Millasseau,P., Zeviani,M. *et al.* (1995) Identification and characterization of a spinal muscular atrophy-determining gene. *Cell*, **80**, 155–165.
3. Gubitz,A.K., Feng,W. and Dreyfuss,G. (2004) The SMN complex. *Exp. Cell Res.*, **296**, 51–56.
4. Paushkin,S., Gubitz,A.K., Massenet,S. and Dreyfuss,G. (2002) The SMN complex, an assemblyosome of ribonucleoproteins. *Curr. Opin. Cell Biol.*, **14**, 305–312.
5. Meister,G., Eggert,C. and Fischer,U. (2002) SMN-mediated assembly of RNPs: a complex story. *Trends Cell. Biol.*, **12**, 472–478.
6. Kaindl,A.M., Guenther,U.P., Rudnik-Schoneborn,S., Varon,R., Zerres,K., Schuelke,M., Hubner,C. and von Au,K. (2008) Spinal muscular atrophy with respiratory distress type 1 (SMARD1). *J. Child Neurol.*, **23**, 199–204.
7. Grohmann,K., Schuelke,M., Diers,A., Hoffmann,K., Lucke,B., Adams,C., Bertini,E., Leonhardt-Horti,H., Muntoni,F., Ouvrier,R. *et al.* (2001) Mutations in the gene encoding immunoglobulin mu-binding protein 2 cause spinal muscular atrophy with respiratory distress type 1. *Nat. Genet.*, **29**, 75–77.
8. Molnar,G.M., Crozat,A., Kraeft,S.K., Dou,Q.P., Chen,L.B. and Pardee,A.B. (1997) Association of the mammalian helicase MAH with the pre-mRNA splicing complex. *Proc. Natl Acad. Sci. USA*, **94**, 7831–7836.
9. Fukita,Y., Mizuta,T.R., Shirozu,M., Ozawa,K., Shimizu,A. and Honjo,T. (1993) The human S mu bp-2, a DNA-binding protein specific to the single-stranded guanine-rich sequence related to the immunoglobulin mu chain switch region. *J. Biol. Chem.*, **268**, 17463–17470.
10. Guenther,U.P., Handoko,L., Laggerbauer,B., Jablonka,S., Chari,A., Alzheimer,M., Ohmer,J., Plottner,O., Gehring,N., Sickmann,A. *et al.* (2009) IGHMBP2 is a ribosome-associated helicase inactive in the neuromuscular disorder distal SMA type 1 (DSMA1). *Hum. Mol. Genet.*, **18**, 1288–1300.
11. de Planell-Saguer,M., Schroeder,D.G., Rodicio,M.C., Cox,G.A. and Mourelatos,Z. (2009) Biochemical and genetic evidence for a role of IGHMBP2 in the translational machinery. *Hum. Mol. Genet.*, **18**, 2115–2126.
12. Jankowsky,E. (2011) RNA helicases at work: binding and rearranging. *Trends Biochem. Sci.*, **36**, 19–29.
13. 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.
14. 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.
15. 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. USA*, **95**, 10009–10014.
16. Chen,Y.Z., Hashemi,S.H., Anderson,S.K., Huang,Y., Moreira,M.C., Lynch,D.R., Glass,I.A., Chance,P.F. and Bennett,C.L. (2006) Senataxin, the yeast Sen1p orthologue: characterization of a unique protein in which recessive mutations cause ataxia and dominant mutations cause motor neuron disease. *Neurobiol. Dis.*, **23**, 97–108.
17. Suraweera,A., Lim,Y., Woods,R., Birrell,G.W., Nasim,T., Becherel,O.J. and Lavin,M.F. (2009) Functional role for senataxin, defective in ataxia oculomotor apraxia type 2, in transcriptional regulation. *Hum. Mol. Genet.*, **18**, 3384–3396.
18. Skourti-Stathaki,K., Proudfoot,N.J. and Gromak,N. (2011) Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. *Mol. Cell*, **42**, 794–805.
19. Fairman-Williams,M.E., Guenther,U.P. and Jankowsky,E. (2010) SF1 and SF2 helicases: family matters. *Curr. Opin. Struct. Biol.*, **20**, 313–324.
20. Singleton,M.R., Dillingham,M.S. and Wigley,D.B. (2007) Structure and mechanism of helicases and nucleic acid translocases. *Annu. Rev. Biochem.*, **76**, 23–50.

21. Velankar,S.S., Soultanas,P., Dillingham,M.S., Subramanya,H.S. and Wigley,D.B. (1999) Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. *Cell*, **97**, 75–84.

22. Korolev,S., Hsieh,J., Gauss,G.H., Lohman,T.M. and Waksman,G. (1997) Major domain swiveling revealed by the crystal structures of complexes of E. coli Rep helicase bound to single-stranded DNA and ADP. *Cell*, **90**, 635–647.

23. Subramanya,H.S., Bird,L.E., Brannigan,J.A. and Wigley,D.B. (1996) Crystal structure of a DExx box DNA helicase. *Nature*, **384**, 379–383.

24. Tomko,E.J., Fischer,C.J., Niedziela-Majka,A. and Lohman,T.M. (2007) A nonuniform stepping mechanism for E. coli UvrD monomer translocation along single-stranded DNA. *Mol. Cell*, **26**, 335–347.

25. Dillingham,M.S., Soultanas,P., Wiley,P., Webb,M.R. and Wigley,D.B. (2001) Defining the roles of individual residues in the single-stranded DNA binding site of PcrA helicase. *Proc. Natl Acad. Sci. USA*, **98**, 8381–8387.

26. Soultanas,P., Dillingham,M.S., Wiley,P., Webb,M.R. and Wigley,D.B. (2000) Uncoupling DNA translocation and helicase activity in PcrA: direct evidence for an active mechanism. *EMBO J.*, **19**, 3799–3810.

27. 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.

28. Saikrishnan,K., Powell,B., Cook,N.J., Webb,M.R. and Wigley,D.B. (2009) Mechanistic basis of 5′-3′ translocation in SF1B helicases. *Cell*, **137**, 849–859.

29. 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.

30. 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.

31. Bowler,M.W., Guijarro,M., Petitdemange,S., Baker,I., Svensson,O., Burghammer,M., Mueller-Dieckmann,C., Gordon,E.J., Flot,D., McSweeney,S.M. *et al.* (2010) Diffraction cartography: applying microbeams to macromolecular crystallography sample evaluation and data collection. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 855–864.

32. Brockhauser,S., White,K.I., McCarthy,A.A. and Ravelli,R.B. (2011) Translation calibration of inverse-kappa goniometers in macromolecular crystallography. *Acta Crystallogr. A*, **67**, 219–228.

33. Leslie,A.G. (2006) The integration of macromolecular diffraction data. Acta Crystallogr. *D Biol. Crystallogr.*, **62**, 48–57.

34. Potterton,E., Briggs,P., Turkenburg,M. and Dodson,E. (2003) A graphical user interface to the CCP4 program suite. Acta Crystallogr. *D Biol. Crystallogr.*, **59**, 1131–1137.

35. Adams,P.D., Afonine,P.V., Bunkoczi,G., Chen,V.B., Davis,I.W., Echols,N., Headd,J.J., Hung,L.W., Kapral,G.J., Grosse-Kunstleve,R.W. *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. *D Biol. Crystallogr.*, **66**, 213–221.

36. Emsley,P. and Cowtan,K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr. *D Biol. Crystallogr.*, **60**, 2126–2132.

37. Kabsch,W. (1988) Automatic indexing of rotation diffraction patterns. *J. Appl. Crystallogr.*, **21**, 67–72.

38. Kabsch,W. (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Crystallogr.*, **26**, 795–800.

39. 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.

40. Laskowski,R.A., MacArthur,M.W., Moss,D.S. and Thornton,J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr*, **26**, 283.

41. Singleton,M.R. and Wigley,D.B. (2002) Modularity and specialization in superfamily 1 and 2 helicases. *J. Bacteriol.*, **184**, 1819–1826.

42. Šponer,J. and Hobza,P. (2003) Molecular interactions of nucleic acid bases. A review of quantum-chemical studies. *Collect. Czech. Chem. Commun.*, **68**, 2231–2282.

43. Grishin,N.V. (1998) The R3H motif: a domain that binds single-stranded nucleic acids. *Trends Biochem. Sci.*, **23**, 329–330.

44. Liepinsh,E., Leonchiks,A., Sharipo,A., Guignard,L. and Otting,G. (2003) Solution structure of the R3H domain from human Smubp-2. *J. Mol. Biol.*, **326**, 217–223.

45. Weng,Y., Czaplinski,K. and Peltz,S.W. (1998) ATP is a cofactor of the Upf1 protein that modulates its translation termination and RNA binding activities. *RNA*, **4**, 205–214.

46. 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.

47. Grohmann,K., Varon,R., Stolz,P., Schuelke,M., Janetzki,C., Bertini,E., Bushby,K., Muntoni,F., Ouvrier,R., Van Maldergem,L. *et al.* (2003) Infantile spinal muscular atrophy with respiratory distress type 1 (SMARD1). *Ann. Neurol.*, **54**, 719–724.

48. Guenther,U.P., Varon,R., Schlicke,M., Dutrannoy,V., Volk,A., Hubner,C., von Au,K. and Schuelke,M. (2007) Clinical and mutational profile in spinal muscular atrophy with respiratory distress (SMARD): defining novel phenotypes through hierarchical cluster analysis. *Hum. Mutat.*, **28**, 808–815.

49. Maystadt,I., Zarhrate,M., Landrieu,P., Boespflug-Tanguy,O., Sukno,S., Collignon,P., Melki,J., Verellen-Dumoulin,C., Munnich,A. and Viollet,L. (2004) Allelic heterogeneity of SMARD1 at the IGHMBP2 locus. *Hum. Mutat.*, **23**, 525–526.

50. Guenther,U.P., Handoko,L., Varon,R., Stephani,U., Tsao,C.Y., Mendell,J.R., Lutzkendorf,S., Hubner,C., von Au,K., Jablonka,S. *et al.* (2008) Clinical variability in distal spinal muscular atrophy type 1 (DSMA1): determination of steady-state IGHMBP2 protein levels in five patients with infantile and juvenile disease. *J. Mol. Med.*, **87**, 31–41.

51. Pierson,T.M., Tart,G., Adams,D., Toro,C., Golas,G., Tifft,C. and Gahl,W. (2011) Infantile-onset spinal muscular atrophy with respiratory distress-1 diagnosed in a 20-year-old man. *Neuromuscul. Disord.*, **21**, 353–355.

52. Giannini,A., Pinto,A.M., Rossetti,G., Prandi,E., Tiziano,D., Brahe,C. and Nardocci,N. (2006) Respiratory failure in infants due to spinal muscular atrophy with respiratory distress type 1. *Intensive Care Med.*, **32**, 1851–1855.

53. Tanner,N.K., Cordin,O., Banroques,J., Doere,M. and Linder,P. (2003) The Q motif: a newly identified motif in DEAD box helicases may regulate ATP binding and hydrolysis. *Mol. Cell.*, **11**, 127–138.

54. Cordin,O., Tanner,N.K., Doere,M., Linder,P. and Banroques,J. (2004) The newly discovered Q motif of DEAD-box RNA helicases regulates RNA-binding and helicase activity. *EMBO J.*, **23**, 2478–2487.

55. Caruthers,J.M. and McKay,D.B. (2002) Helicase structure and mechanism. *Curr. Opin. Struct. Biol.*, **12**, 123–133.

56. Cordin,O., Banroques,J., Tanner,N.K. and Linder,P. (2006) The DEAD-box protein family of RNA helicases. *Gene*, **367**, 17–37.

57. Pause,A., Methot,N. and Sonenberg,N. (1993) The HRIGRXXR region of the DEAD box RNA helicase eukaryotic translation initiation factor 4A is required for RNA binding and ATP hydrolysis. *Mol. Cell. Biol.*, **13**, 6789–6798.

58. Gu,M. and Rice,C.M. (2010) Three conformational snapshots of the hepatitis C virus NS3 helicase reveal a ratchet translocation mechanism. *Proc. Natl Acad. Sci. USA*, **107**, 521–528.

59. 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.

60. Joseph,S., Robb,S.A., Mohammed,S., Lillis,S., Simonds,A., Manzur,A.Y., Walter,S. and Wraige,E. (2009) Interfamilial phenotypic heterogeneity in SMARD1. *Neuromuscul. Disord.*, **19**, 193–195.

61. Chen,Y.Z., Bennett,C.L., Huynh,H.M., Blair,I.P., Puls,I., Irobi,J., Dierick,I., Abel,A., Kennerson,M.L., Rabin,B.A. *et al.* (2004) DNA/RNA helicase gene mutations in a form of juvenile amyotrophic lateral sclerosis (ALS4). *Am. J. Hum. Genet.*, **74**, 1128–1135.

62. Moreira,M.C., Klur,S., Watanabe,M., Nemeth,A.H., Le Ber,I., Moniz,J.C., Tranchant,C., Aubourg,P., Tazir,M., Schols,L. *et al.* (2004) Senataxin, the ortholog of a yeast RNA helicase, is mutant in ataxia-ocular apraxia 2. *Nat. Genet.*, **36**, 225–227.

63. Suraweera,A., Becherel,O.J., Chen,P., Rundle,N., Woods,R., Nakamura,J., Gatei,M., Criscuolo,C., Filla,A., Chessa,L. *et al.* (2007) Senataxin, defective in ataxia oculomotor apraxia type 2, is involved in the defense against oxidative DNA damage. *J. Cell Biol.*, **177**, 969–979.