The EMBO Journal (2010) 29, 2194–2204 | © 2010 European Molecular Biology Organization | All Rights Reserved 0261-4189/10
www.embojournal.org

THE EMBO JOURNAL

Prp43p contains a processive helicase structural architecture with a specific regulatory domain

Hélène Walbott[1,5], Saïda Mouffok[2,3,5], Régine Capeyrou[2,3,*], Simon Lebaron[2,3,6], Odile Humbert[2,3], Herman van Tilbeurgh[1], Yves Henry[2,3,*] and Nicolas Leulliot[1,4,*]

[1]Institut de Biochimie et de Biophysique Moléculaire et Cellulaire, Université de Paris-Sud, CNRS-UMR8619, IFR115, Orsay Cedex, France, [2]Centre National de la Recherche Scientifique, Laboratoire de Biologie Moléculaire Eucaryote, Toulouse, France, [3]Université de Toulouse, UPS, Toulouse, France and [4]Laboratoire de Cristallographie et RMN Biologiques, Université Paris Descartes, CNRS-UMR8015, Paris Cedex, France

The DEAH/RNA helicase A (RHA) helicase family comprises proteins involved in splicing, ribosome biogenesis and transcription regulation. We report the structure of yeast Prp43p, a DEAH/RHA helicase remarkable in that it functions in both splicing and ribosome biogenesis. Prp43p displays a novel structural architecture with an unforeseen homology with the Ski2-like Hel308 DNA helicase. Together with the presence of a β-hairpin in the second RecA-like domain, Prp43p contains all the structural elements of a processive helicase. Moreover, our structure reveals that the C-terminal domain contains an oligonucleotide/oligosaccharide-binding (OB)-fold placed at the entrance of the putative nucleic acid cavity. Deletion or mutations of this domain decrease the affinity of Prp43p for RNA and severely reduce Prp43p ATPase activity in the presence of RNA. We also show that this domain constitutes the binding site for the G-patch-containing domain of Pfa1p. We propose that the C-terminal domain, specific to DEAH/RHA helicases, is a central player in the regulation of helicase activity by binding both RNA and G-patch domain proteins.

The EMBO Journal (2010) **29,** 2194–2204. doi:10.1038/emboj.2010.102; Published online 28 May 2010
Subject Categories: RNA; structural biology
Keywords: G-patch domain; Prp43p; ribosome biogenesis; RNA helicase structure; splicing

*Corresponding authors. R Capeyrou, Laboratoire de Biologie Moléculaire Eucaryote, 118 route de Narbonne, 31062 Toulouse cedex 09, France. Tel.: 33 5 61 33 58 20; Fax: 33 5 61 33 58 86; E-mail: capeyrou@ibcg.biotoul.fr or Y Henry, Laboratoire de Biologie Moléculaire Eucaryote, 118 route de Narbonne, 31062 Toulouse cedex 09, France. Tel.: 33 5 61 33 59 53; Fax: 33 5 61 33 58 86; E-mail: henry@ibcg.biotoul.fr or N Leulliot, Laboratoire de Cristallographie et RMN Biologiques, Université Paris Descartes, 4 av de l'Observatoire, 75270 cedex 06, France. Tel: 33 1 53 73 15 70; Fax: 33 1 53 73 99 25; E-mail: nicolas.leulliot@parisdescartes.fr
[5]These authors contributed equally to this work
[6]Present address: Wellcome Trust Centre for Cell Biology, Institute for Cell and Molecular Biology, King's Buildings, University of Edinburgh, EH9 3JR, UK

Received: 27 November 2009; accepted: 3 May 2010; published online: 28 May 2010

Introduction

RNA helicases intervene in all aspects of RNA metabolism, including transcription, pre-mRNA splicing, mRNA export, ribosome biogenesis, translation and RNA degradation (Tanner and Linder, 2001; Jankowsky and Fairman, 2007). These enzymes catalyse RNA conformational rearrangements, such as strand separation, secondary-structure melting and RNA–protein dissociation, using ATP hydrolysis (Bleichert and Baserga, 2007). They are found in all organisms, from bacteria and viruses to humans. The superfamily 2 (SF2) DExD/H-box proteins share a helicase core consisting of two tandemly repeated RecA-like domains containing eight conserved motifs (Jankowsky and Fairman, 2007; Pyle, 2008) and are often associated with a wide variety of domains (Caruthers and McKay, 2002). The structures of multidomain DExD/H-box proteins show that the accessory domains act in synergy with the two RecA-like domains, for example by participating in RNA binding or protein–protein interactions, to modulate the function of the helicase. Helicases have been classified according to the identity of the catalytic residues in motif II (DEAD, DExH and DEAH-box helicases) and to their domain architecture (Jankowsky and Fairman, 2007).

The DEAH/RNA helicase A (RHA) is a large RNA helicase family (Sanjuan and Marin, 2001) characterized by two conserved C-terminal domains following the helicase core, identified by sequence homology: a 'helicase-associated domain' HA2, belonging to the Pfam PF04408 superfamily and a domain of unknown function (DUF1605) identified as belonging to the Pfam PF07717 superfamily. The Pfam PF07717 domain is almost always found at the C-terminus of DEAH/RHA proteins and apparently always in association with the HA2 domain. Although these domains are essential *in vivo* (Edwalds-Gilbert *et al*, 2004; Koo *et al*, 2004), neither of them has yet been functionally and structurally characterized. This domain architecture is unique to the DEAH/RHA helicases and distinct from the viral DEAH helicases, which form a separate family.

In yeast *Saccharomyces cerevisiae*, the DEAH/RHA family comprises six essential putative RNA-dependent ATPases with clear homologues in humans: the pre-mRNA splicing factors Prp2p, Prp16p and Prp22p (DHX16, DHX38 and DHX8 in humans), the ribosome biogenesis factors Dhr1p and Dhr2p (DHX37 and DHX32) and the remarkable bifunctional factor Prp43p (DHX15) involved in both those processes. A seventh nonessential yeast DEAH-box protein YLR419w (Shiratori *et al*, 1999; Sanjuan and Marin, 2001) is homologous to human helicases belonging to the RHA subfamily, comprising several DEAH-box proteins (DHX9, DHX29, DHX57, DHX36 and DHX30) associated with transcription regulation, cell-cycle regulation and cancer (Erkizan *et al*, 2009). The bacterial HrpA and HrpB helicases found in many Proteobacteria, which are thought to be involved in mRNA processing (Koo *et al*, 2004), are the only prokaryotic DEAH/RHA helicases.

The precise function of DEAH/RHA helicases in these processes remains mostly elusive. In splicing and ribosome biogenesis, these helicases act in the context of large ribonucleoprotein complexes. DEAH/RHA helicases are thought to function in remodelling the structure and/or composition of these complexes by locally melting RNA double helical strands (Schwer, 2001). For example, Prp16p induces a conformational rearrangement in the spliceosome after the first transesterification step (Mefford and Staley, 2009), Prp22p is proposed to remove mRNA from the spliceosome by translocating along the mRNA and disrupting the U5 snRNP-mRNA contact (Schwer, 2008), and Prp43p is necessary for removal of the lariat intron from the spliceosome (Arenas and Abelson, 1997; Martin *et al*, 2002; Tsai *et al*, 2005) and production of both large and small ribosomal subunits (Lebaron *et al*, 2005; Combs *et al*, 2006; Leeds *et al*, 2006). The function of Prp43p and Prp2p is modulated by proteins containing a G-patch domain, a domain mostly identified in RNA-associated proteins and called after its characteristic glycine-rich sequence signature (Aravind and Koonin, 1999). Prp2p interaction with G-patch protein Spp2p is required for association with the spliceosome (Silverman *et al*, 2004). Three G-patch proteins, Ntr1p, Gno1p and Pfa1p, interact with Prp43p. Ntr1p is involved in splicing, where it is required for Prp43p association with the spliceosome and release of the lariat intron *in vitro* (Tsai *et al*, 2005, 2007; Boon *et al*, 2006; Tanaka *et al*, 2007), whereas Gno1p and Pfa1p are implicated in ribosome biogenesis (Guglielmi and Werner, 2002; Lebaron *et al*, 2005). Pfa1p directly binds to Prp43p and activates its ATPase and helicase activities *in vitro*. Activation of Prp43p by Pfa1p is probably required for efficient 20S pre-rRNA to 18S rRNA conversion (Lebaron *et al*, 2005, 2009; Pertschy *et al*, 2009).

Here, we report the first crystal structure of a DEAH/RHA helicase: the full-length yeast Prp43p in complex with ADP/Mg^{2+} at 1.9 Å resolution. The structure of Prp43p reveals an unexpected structural homology with the Ski2-like DNA helicase Hel308 (Buttner *et al*, 2007). Moreover, we uncover a DEAH/RHA-specific C-terminal domain containing an oligonucleotide/oligosaccharide-binding (OB)-fold and a Prp43p-specific domain at the N-terminus. The ADP is in a binding conformation characteristic of DEAH-box helicases (Luo *et al*, 2008), and very different from what was usually observed for the DEAD-box proteins (Cordin *et al*, 2006; Sengoku *et al*, 2006). We further show that the C-terminal domain of Prp43p containing the OB-fold is required for full ATPase activity and RNA binding of Prp43p. Finally, we show that this domain of Prp43p is also involved in binding to the G-patch domain of Pfa1p and mediates Prp43p activation by its protein partner. Therefore, altogether these results provide a structural model for all the DEAH/RHA helicases and significant advances in understanding their mode of interaction with both RNA and G-patch proteins.

Results

Prp43p is homologous to Ski2-like DNA helicase Hel308

The yeast Prp43p protein was purified as described in the Materials and methods section and crystallized in the presence of ADP and magnesium. The structure of Prp43p was solved at 1.9 Å resolution by single anomalous diffraction (SAD) on selenomethionine-substituted protein (see Table I for data collection and refinement statistics). The final model contains residues 3 to 749, one ADP molecule and one magnesium atom (Figure 1B). The two molecules per asymmetric unit have essentially the same conformation, with a root mean square deviation (r.m.s.d.) of 0.5 Å for all corresponding Cα atoms.

Table I Data collection and refinement statistics

	Native	SeMet
Data collection		
Space group	$P3_221$	$P3_221$
Cell dimensions		
a, *b*, *c* (Å)	117.6 117.6 254.6	117.7 117.7 253.3
α, β, γ (deg)	90 90 120	90 90 120
Wavelength (Å)	0.9793	0.9789
Resolution (Å)	34.0–1.9	35.0–2.8
Outer resolution shell (Å)	2.0–1.9	2.97–2.8
Number of observed reflections/unique	843 635/160 290	554 789/96 552
Completeness (%) (outer shell)	99.8 (98.6)	99.6 (98.2)
Multiplicity (outer shell)	5.3 (4.1)	5.7 (5.7)
I/σ (I) (outer shell)	13.4 (2.2)	13.4 (4.0)
R_{merge} (%) (outer shell)	8.7 (59.5)	11.1 (37.7)
Refinement		
Resolution (Å)	34.3–1.9	
Reflections (working/test)	160 181/8044	
R/R_{free} (%)	18.4/22.4	
R.m.s.d.		
Bond lengths (Å)	0.017	
Bond angles (deg)	1.557	
B-factors (Å^2)	32.1	
Protein	31.4	
ADP/Mg^{2+}	27.4	
Solvent	38.6	
Ramachandran statistics (%)		
Most favoured	97.1	
Allowed	2.9	

The protein can be divided into six domains, which make a claw-like structure. The N-terminal domain (green in Figure 1) is a Prp43p-specific domain, whereas the C-terminal domain is common to all DEAH/RHA helicases (red in Figure 1). The N-terminal extension of Prp43p (green in Figure 1) contains three α helices: α helix αN1 packs between the domains 4 and 5, whereas α helices αN2, αN3 interact with the domain 2. These helices are separated by an ordered 20-residue linker, which braces the helicase ring structure. The helices therefore interact with distinct functional domains on opposite sides of the helicase. Domains 2 and 3 correspond to the two RecA-like domains ubiquitous to helicase structures, involved in ATP binding, hydrolysis and nucleic acid binding (hereafter referred to as RecA1 and RecA2, respectively, cyan and dark blue in Figure 1). The structure of these domains is structurally most homologous to the viral NS3 DEAH helicases (r.m.s.d. 3.3 Å, PDB 2JLW; Luo *et al*, 2008). Notably, the RecA2 domain also contains an antiparallel β-hairpin (light blue in Figure 1B) protruding from the domain between motifs V and VI. This β-hairpin element, which is not found in DEAD-box proteins, seems therefore to be a conserved feature of the DExH-box helicases.

Strikingly, domains 4 and 5 are homologous to the corresponding domains of Hel308 Ski2-like family helicase involved in DNA repair (r.m.s.d. 3.1 Å, PDB 2P6R; Buttner *et al*, 2007; Supplementary Figure S1). Domain 4 is a winged helix (WH)-fold (orange in Figure 1) that packs against the RecA1 domain. Domain 5, a seven-helix bundle that binds

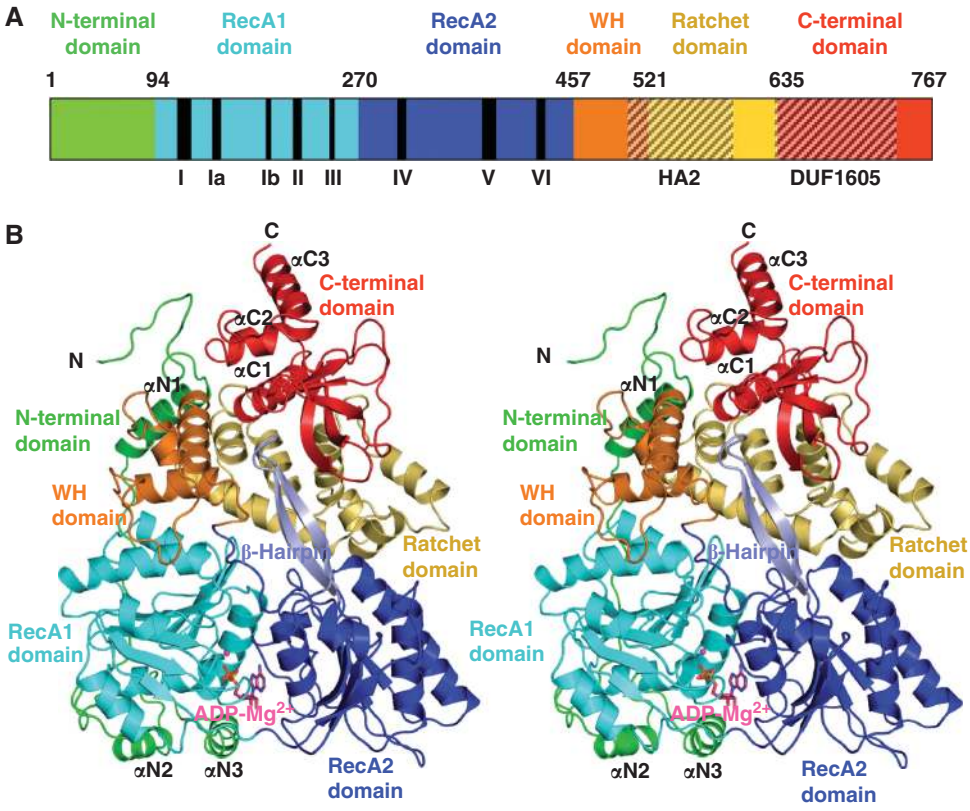

Figure 1 Structure of yeast Prp43p. (**A**) Schematic representation of the domain organization of Prp43p. Structural domain names and boundaries are indicated on top, conserved sequence motifs (roman numerals) and domains (hatched areas) beneath. (**B**) Stereo view of the overall structure of Prp43p in complex with ADP-Mg^{2+}. The six structural domains of Prp43p are coloured as in (**A**). The β-hairpin element (light blue) protrudes from the RecA2 domain (dark blue) between motifs V and VI. The ADP molecule is shown as magenta sticks, and the magnesium ion is represented as a magenta sphere.

across the two RecA-like domains, has been described as the 'ratchet' domain (yellow in Figure 1; Buttner *et al*, 2007). The WH and ratchet domains of Prp43p are annotated in domain databases as the 'helicase-associated domain' HA2, which is present in all DEAH/RHA protein sequences (Figure 1A). The structure of Prp43p therefore reveals an unforeseen structural homology of the HA2 domain with the characteristic WH and ratchet domains of the Ski2-like family helicases. This result confirms that Prp43p, and therefore all the DEAH/RHA helicases, have a distinct C-terminal domain structure compared with viral DEAH helicases.

The Prp43p helicase core binds ATP in a DEAH-specific mode

The ADP molecule bound to Prp43p is sandwiched between the RecA1 and RecA2 domains, interacting with the conserved motifs I, II, V and VI already known to be required for ATP binding and hydrolysis (Cordin *et al*, 2006; Pyle, 2008), but also with unexpected additional residues between motifs Ia and Ib and between motifs IV and V (Figures 1B and 2A). As in all helicase structures solved to date, the P-loop (motif I) of Prp43p binds the phosphate moiety of the ADP molecule and the α and β phosphates are therefore superposable with other bound nucleotides in helicase structures (Figure 2A).

The binding mode of the base/ribose moieties of the ADP strongly differs from that reported so far in DEAD-box and Ski2-like helicase structures but is more similar to NS3 viral helicase bound nucleotides (Figure 2A and B). In comparison to the Prp43p conformation, the ADP α angle (O3α-Pα-O5′-C5′) is rotated by about 150° in the DEAD-box protein Vasa (Sengoku *et al*, 2006) and in the Hel308 homologue Hjm (PDB 2ZJ5; Oyama *et al*, 2009; Figure 2B). As a result, in Prp43p, the adenine ring is sandwiched between the two RecA-like domains instead of being buried inside the RecA1 domain. The adenine base is stacked between the F357 phenyl ring (RecA2) and the R159 side chain (RecA1) (Figure 2A). In the viral NS3 DEAH helicases, the adenine base of the bound ADP is totally exposed to solvent without any stacking interactions (Wu *et al*, 2005; Luo *et al*, 2008). Moreover, in the Prp43p conformation, the 2′ and 3′ hydroxyl groups of the ribose moiety make H-bonds with the carboxyl group of D386 (motif V) and the amino group of R430 (motif VI), respectively (Figure 2A). In the viral NS3 DEAH helicases, the 3′OH group of the ribose interacts with the main chain of the arginine of motif VI (Luo *et al*, 2008), whereas in the DEAD-box helicases, the 3′OH group of the ribose is hydrogen bonded with the aspartate of motif V and residues from motif VI only interact with the phosphate moiety (Sengoku *et al*, 2006). It should be noted that the interaction of the nucleotides with the different motifs are generally influenced by binding with nucleic acids, as in the RNA bound and RNA-free structures of the related viral NS3 DEAH helicases (Luo *et al*, 2008), and that the observed interactions in Prp43p could therefore change upon RNA binding.

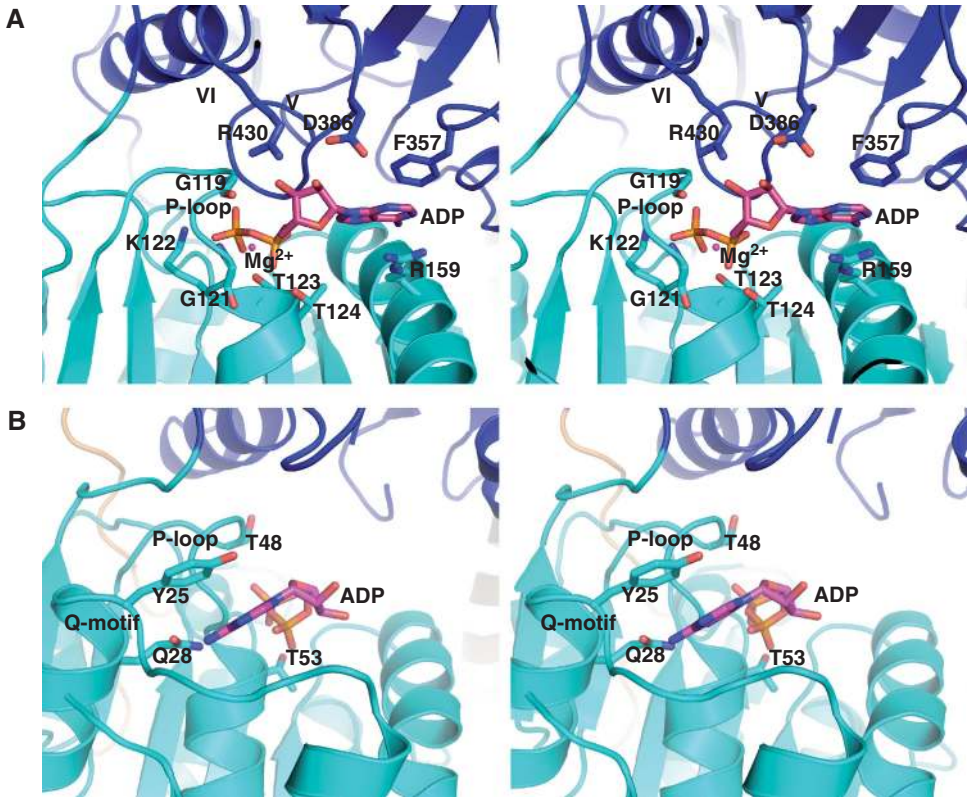

Figure 2 Detailed comparison of the ATPase sites of Prp43p and the Hel308 homologue Hjm. The RecA1 and RecA2 domains are coloured as in Figure 1. Residues that directly interact with the ADP molecule (magenta sticks) are shown as sticks and labelled. The conserved sequence motifs are also indicated. (**A**) Close-up stereo view of the Prp43p ADP-binding site. The magnesium ion is pictured as a magenta sphere. (**B**) Close-up stereo view of the Hel308 homologue Hjm ADP-binding site (PDB 2ZJ5; Oyama *et al*, 2009) in the same orientation as in (**A**).

An additional difference between DEAD and DEAH-box helicases is the absence of the DEAD-box-specific Q-motif upstream of motif I in DEAH-box helicases (Tanner *et al*, 2003). This motif is characterized by a conserved glutamine, which hydrogen bonds to the N6 and N7 positions of the adenine base (Sengoku *et al*, 2006; Oyama *et al*, 2009; Figure 2B), thereby providing ATP-binding specificity. The nonspecific recognition of the base could explain why Prp43p and viral NS3 DEAH helicases are NTPases (Warrener *et al*, 1993; Tanaka and Schwer, 2006) rather than strict ATPases-like DEAD-box proteins (Cordin *et al*, 2006). Therefore, this novel ATP-binding mode observed in the structure of the Prp43p/ADP complex seems to be a specific feature of the viral DEAH and DEAH/RHA helicase families.

Prp43p contains all the elements of a processive helicase

The structure of Hel308 in complex with DNA provided a structural model for the processive helicase activity of Hel308 in which the WH and ratchet domains had a central function (Buttner *et al*, 2007). The structural homology of Prp43p with Hel308 provides a template to model the interaction of Prp43p with nucleic acids and determine whether both proteins function by the same mechanism. Figure 3B presents a model of a Prp43p/DNA complex modelled against the Hel308/DNA complex (PDB 2P6R; Buttner *et al*, 2007; Figure 3A) by simple superposition of the proteins. This model shows that Prp43p contains a similar single-stranded nucleic acid-binding cavity, formed by the 2 RecA-like, the ratchet and the WH domains. The 2 RecA-like domains bind this nucleic acid strand through the phosphate backbone. The RecA1 domain binds to the 3′ end of the modelled ssDNA and RecA2 interacts with the 5′ end.

In an indirect validation of this model, an acetate molecule from the crystallization buffer is found bound to the RecA2 domain in a position occupied by one of the phosphates of the modelled nucleic acid (not shown). The T376V mutation in motif V, which is in close proximity to this acetate molecule and conserved in Hel308 structure where it binds DNA, is lethal *in vivo* and displays low helicase activity *in vitro* (Tanaka and Schwer, 2006). R150 from motif Ia of the RecA1 domain also probably contributes to nucleic acid recognition based on this model. Interestingly, the R150A mutant displays a cold-sensitive phenotype *in vivo* (Tanaka and Schwer, 2006).

A unique structural trait of Hel308 is the long 'ratchet' helix, which binds across the entire length of the single strand in the nucleic acid-binding cavity. The stacking of R592 and W599 from this helix on the DNA bases is proposed to act as hooks on the nucleic acid strand, which would allow the Hel308 enzyme to processively translocate along the single strand to unwind the DNA duplex (Buttner *et al*, 2007). Prp43p contains a similar ratchet helix with R611 and R625 side chains protruding from the underside of this helix (not shown). Although the positions of these residues along the ratchet helix are different compared with Hel308, they probably make stacking interactions with the base moieties of the

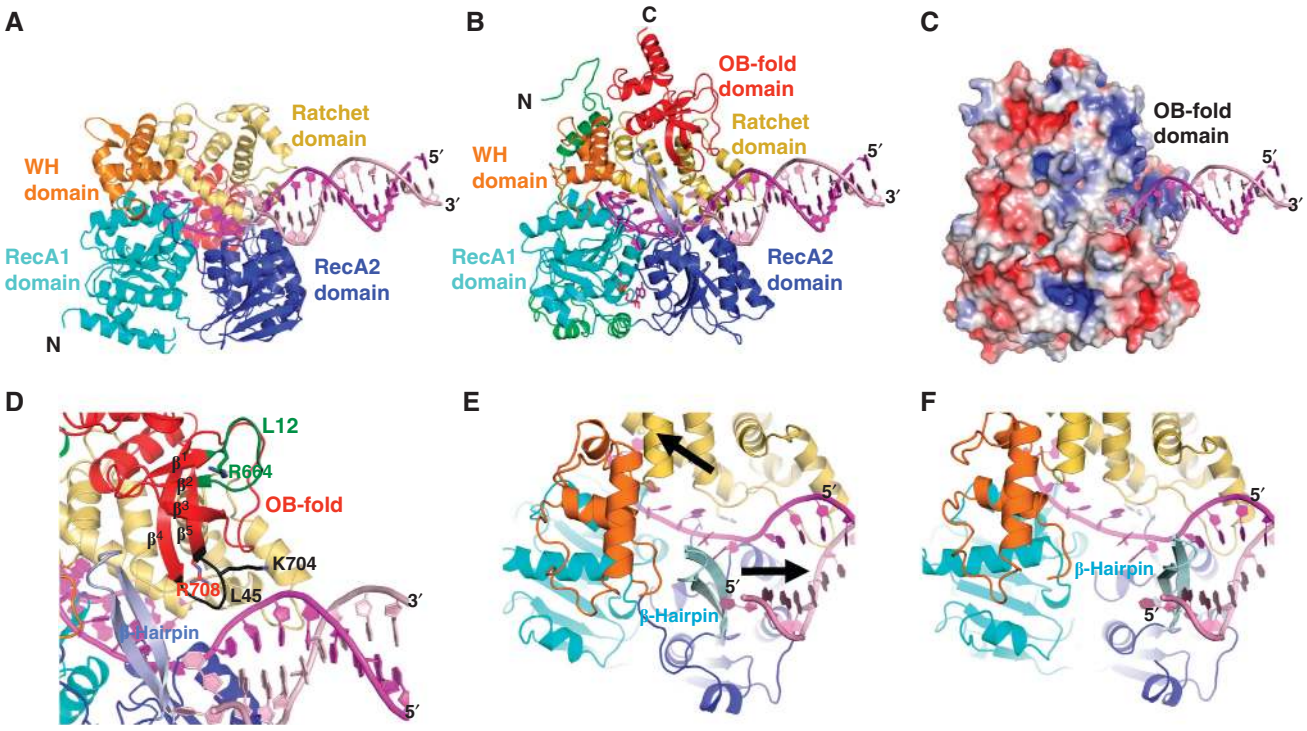

Figure 3 RNA unwinding by the Prp43p helicase. (**A**) Overview of the Hel308/DNA complex (PDB 2P6R; Buttner *et al*, 2007) in the same orientation and colour code as in Figure 1B. The DNA strands are shown in magenta (threaded strand) and pink cartoons. (**B**) Model of the Prp43p/DNA complex modelled against the Hel308/DNA complex by simple superposition of the proteins (same orientation and colour code as in (**A**)). (**C**) Surface representation of Prp43p in complex with modelled DNA (magenta and pink), coloured according to its electrostatic potential calculated with APBS (Baker *et al*, 2001) and presented in the same orientation as in (**B**). (**D**) Close-up view of the nucleic acid-binding cavity entrance of the Prp43p/DNA complex. The β-hairpin element and the OB-fold domain are coloured light blue and red, respectively. The loops L12 (dark green) and L45 (black) and specific positively charged residues (pictured as sticks) of the OB-fold, which are mutated in the probed Prp43p variants, are also indicated. (**E**, **F**) Close-up views of the conformational rearrangements of each of the Prp43p domains (**E**, initial conformation) to adopt the Hel308 conformation (**F**, final conformation). The Prp43p domains and modelled DNA are pictured as in (**B**). The two black arrows in (**E**) show the concerted movement of the RecA-like and ratchet domains, which provides a possible model for RNA translocation and unwinding mechanism of Prp43p helicase: the ratchet domain seems to pull the threaded single strand inside the cavity (left arrow), whereas the β-hairpin slices through the nucleic acid duplex (right arrow).

bound nucleic acid. Therefore, the corresponding ratchet helix might be used in an analogous manner in Prp43p helicase activity.

The Prp43p RecA2 domain possesses an antiparallel β-hairpin (light blue in Figures 1B, 3B and D) protruding at the entrance of the nucleic acid-binding cavity, the second hallmark of processive helicases, as revealed by Hel308 (Buttner *et al*, 2007; light blue in Figure 3A). This β-hairpin element, also found in viral NS3 DEAH helicases, is proposed to act as a physical barrier promoting nucleic acid strand separation in unwinding activity, although this function has not yet been addressed experimentally (Kim *et al*, 1998; Buttner *et al*, 2007; Luo *et al*, 2008). The Prp43p β-hairpin is notably longer than the corresponding loop in Hel308, but most similar to the DENV NS3h β-hairpin.

The relative orientation of the 2 RecA-like, WH and ratchet domains is significantly different in Prp43p and Hel308 (Figure 3A and B). As major conformational changes are thought to occur during the helicase ATP cycle (Luo *et al*, 2008), it is tempting to speculate that the Hel308 and Prp43p structures selected different representative conformations of the helicase during its catalytic cycle. Observing different conformational states of a helicase is inherently difficult because of the effect of crystal packing, which selects certain conformations compatible with crystal growth, providing only a snapshot of the conformation space sampled by the helicase that might not be biologically relevant. Moreover, in the absence of the nucleic acid substrate, helicases are thought to be more flexible and therefore more prone to be influenced by crystal lattice interactions. However, as all Hel308 homologues in the presence or absence of nucleic acid or different nucleotides crystallize in the same conformation (Buttner *et al*, 2007; Oyama *et al*, 2009), no information on the other conformational states sampled by this helicase is available. The conformation of Prp43p provides us with a possible model, which could give insight into the dynamic changes during the catalytic cycle of this helicase superfamily. Indeed, Prp43p would have to undergo substantial conformational rearrangements of each of the domains to adopt the Hel308 conformation (Figure 3E and F). Although no experimental evidence exists to support such a conformational change, the concerted movement of the 2 RecA-like and ratchet domains resembles that of a 'can opener' and provides an appealing model for strand separation activity: the ratchet domain seems to pull the threaded RNA strand inside the cavity, whereas the β-hairpin slices through the RNA double helix (Figure 3E and F). This model will need to be addressed experimentally by solving structures in different conformational states or other methods probing protein dynamics.

The C-terminal OB-fold domain of Prp43p defines a new helicase subfamily

In Hel308, a α-helical C-terminal domain following the ratchet domain binds the single-stranded DNA at the exit of the helicase tunnel (Buttner *et al*, 2007). Prp43p and all the putative eukaryotic DEAH/RHA helicases also contain a characteristic C-terminal domain of unknown function (DUF1605/Pfam PF07717; Figure 1A). Our structure reveals that part of this domain contains a five-stranded β-barrel topology with strand order β1β2β3β5β4 (red in Figures 1B, 3B and D). Structural alignments performed with EBI SSM (Krissinel and Henrick, 2004) clearly classify it as an OB-fold (Murzin, 1993), with no structural homology to the Hel308 C-terminal domain.

The Prp43p OB-fold domain (red in Figure 1B) is linked to the ratchet domain by an α helix (αC1), which extensively packs with the ratchet domain. Two helices (αC2 and αC3) follow the OB-fold, αC2 stacks on the αC1 helix, whereas αC3 stacks on both the αC2 helix and on the β1 and β2 strands of the OB-fold. The OB-fold is fixed on one side by stacking of the β4 and β5 strands and the β3–β4 loop on helix αC1, on the ratchet domain and to a lesser extent on helix αC2. On the other side, the strand β1, the C-terminus of strand β4 and the β1–β2, β2–β3 and β4–β5 loops stack with the helix αC3, the β-hairpin of the RecA2 domain and to a lesser extent with the WH domain. This places the OB-fold domain on the opposite side of the helicase compared with the Hel308 C-terminal domain (Figure 3A and B).

The DUF1605/Pfam PF07717 domain (Figure 1A) is therefore structurally assignable to an OB-fold domain preceded and followed by one and two α helices, respectively. Sequence alignments suggest that all DEAH/RHA helicases contain the same structural elements in this domain, with the exception of helix αC3, which is less conserved. OB-folds are found in proteins involved in a wide variety of functions, generally associated with nucleic acid binding, although it has also been found to be involved in protein–protein interactions (Arcus, 2002). As all the DEAH/RHA helicases share the same domain organization, where this domain is always associated with the HA2 domain (WH and ratchet), the structure of Prp43p can therefore be used as a model to study the function of this domain in DEAH/RHA helicase function.

The C-terminal domain of Prp43p contributes to RNA-stimulated ATPase activity and increases its affinity for RNA

Whereas the C-terminal domain of Hel308 acts as a molecular 'brake' for helicase activity (Richards *et al*, 2008), the OB-fold-containing domain of Prp43p has been shown to be essential for activity *in vivo*. Several amino-acid substitutions and truncations within Prp43p lead to phenotypes ranging from temperature sensitivity to lethality (Martin *et al*, 2002; Tanaka and Schwer, 2006), showing a prominent function for the OB-fold-containing domain in Prp43p function. The effect on the biochemical activity of Prp43p has not been precisely investigated.

In the Prp43p/DNA model, the OB-fold is located at the entrance of the nucleic acid-binding tunnel and interacts with the β-hairpin (Figure 3B and D). Although OB-folds from different proteins share little sequence homology, they are generally associated with nucleic acid-binding activity and use a common theme for nucleic acid recognition. The nucleic acid-binding surface is generally located on the groove formed by the β2–β3 strands, and the β1–β2, β3–β4 and β4–β5 loops (Arcus, 2002; Theobald *et al*, 2003). In Prp43p, this groove is solvent exposed and conveniently placed at the entrance of the nucleic acid cleft, and the electrostatic potential mapped on the surface of the protein (Figure 3C) shows that this surface is positively charged. The OB-folds also generally bind nucleic acids in a defined polarity. In the context of a nucleic acid that binds Prp43p in the same manner as the DENV NS3h and Hel308 proteins, the Prp43p OB-fold would be in a perfect orientation to interact with the RNA substrate entering the cavity.

To directly test the importance of the C-terminal domain of Prp43p for binding RNA, we expressed and purified a truncated version of Prp43p, Prp43p(1–657), lacking the C-terminus up to the αC1 helix (Figure 4A), and used it in band retardation assays (Figure 4B). We also performed ATPase assays in the absence or presence of RNA with this truncated protein (Figure 5A and B). ATP hydrolysis time courses reveal that the basal ATPase activity of Prp43p(1–657) in the absence of RNA is identical to that of wild-type Prp43p (Figure 5A). In contrast, the initial ATP hydrolysis rate of Prp43p(1–657) in the presence of total yeast RNA is about 20-fold lower than that of wild-type Prp43p (Figure 5B). This decrease in RNA-stimulated ATP hydrolysis rate is consistent with a 200-fold decrease in the affinity for the 5′ half of snR5 snoRNA (Figure 4B). The decrease in RNA-stimulated ATPase activity and affinity for RNA is unlikely the result of an overall unfolding of the truncated protein for the following reasons: (1) Prp43p(1–657) is efficiently produced in a soluble form in *Escherichia coli*, (2) the basal unstimulated ATPase activity of Prp43p(1–657) is similar to that of the wild-type protein and (3) it retains the ability to bind RNA, albeit much more weakly than wild type. The residual binding of Prp43p(1–657) to RNA could be anticipated, as the RecA helicase core generally represents the minimum RNA-binding module for helicases. These data show that the C-terminal domain of Prp43p is not absolutely required for its interaction with RNA, but strongly increases its affinity for RNA and consistently, its RNA-stimulated ATPase activity.

The OB-fold of Prp43p C-terminal domain is a RNA-binding module important for RNA-stimulated ATPase activity

As the C-terminal domain contains an OB-fold module often associated with nucleic acid binding, we constructed more subtle mutations and truncations within the OB-fold to determine whether the generic nucleic acid-binding surface, located at the entrance of the helicase cavity in the Prp43p structure, is implicated in RNA binding. We chose to mutate the β1–β2 and β4–β5 loops in the OB-fold, as these loops generally constitute the major determinants in nucleic acid recognition and contain positively charged residues exposed to solvent in the Prp43p structure (Arcus, 2002; Theobald *et al*, 2003). Two mutants of the β1–β2 loop were chosen as either a single mutation of R664 into A (R664A) or a more drastic combination of K663 to A, R664 to G substitutions and deletion of amino acids 665–668 (Prp43pΔL12). Two similar mutants of the β4–β5 loop, the single mutation of K704 to A (K704A) and the combined T702 to G, S703 to G substitutions and deletion of K704 (Prp43pΔL45) were constructed

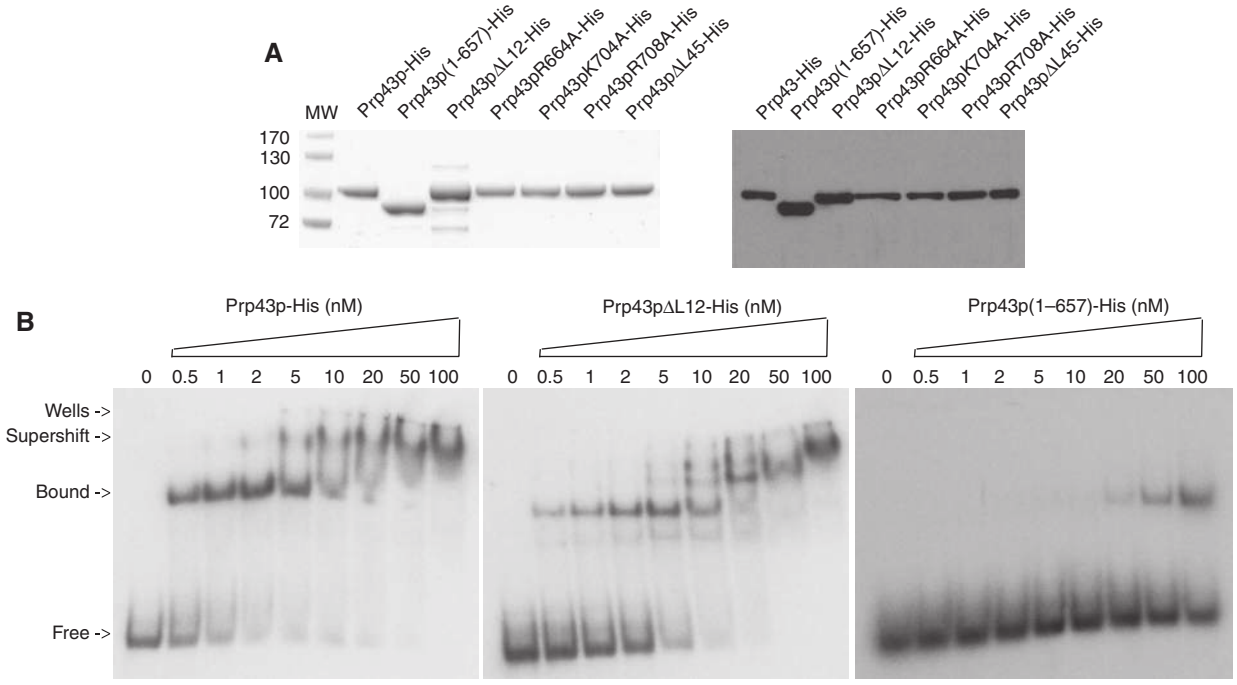

Figure 4 Analysis of the production and RNA binding of Prp43p variants. (**A**) Coomassie blue staining (left) and western blot analysis (right) using anti-histidine antibodies of purified histidine-tagged wild-type Prp43p and the indicated histidine-tagged Prp43p variants. (**B**) Band retardation assays performed with the 5′ half of snR5 snoRNA at 1 nM hybridized to a 21 nucleotide-long radiolabelled ssDNA (Lebaron *et al*, 2009) and wild-type Prp43p (left), Prp43pΔL12 (middle) or Prp43p(1–657) (right) at the indicated concentrations.

(Figure 3D). The R708 to A mutation (R708A) was also considered because this residue (on β5) is not at the surface of the protein but instead buried inside the putative Prp43p nucleic acid-binding cavity. In our model, this residue contacts the single-stranded RNA inside the cavity (Figure 3D). All mutants were purified from *E. coli* (Figure 4A) and tested in band retardation (Figure 4B) and ATPase assays (Figure 5B). Circular dichroism analyses show that all these Prp43p variants are properly folded (data not shown). Moreover, in the absence of RNA, all these variants exhibited a basal ATPase activity similar to that of wild-type Prp43p (data not shown).

To assess the effect of the mutations on RNA binding, band retardation analyses were carried out with all the mutants using an snR5 RNA template (Figure 4B and data not shown). As could be expected, mutation of R708 that is not surface exposed, as well as single mutations in the β1–β2 and β4–β5 loops did not display any significant decrease in RNA-binding affinity (data not shown). In our experimental conditions, only Prp43pΔL12 displayed a modest but reproducible 2.5-fold decrease in RNA-binding affinity (Figure 4B). More interestingly, in the presence of RNA, all the Prp43p variants displayed reduced ATPase activity compared with wild type (Figure 5B). The most severely affected variants were the Prp43pΔL12 and Prp43pΔL45 mutants in which the β1–β2 and β4–β5 loops are removed, and single mutations in these loops (R664A and K704A) had a smaller but measurable effect on activity. Prp43pΔL45 displayed an ATPase activity close to that of Prp43p(1–657).

Interestingly, the R708A mutation, which is not surface exposed, was also severely impaired in RNA-stimulated ATPase activity. Therefore, the effect on ATPase activity arises not only from the putative recruitment of the RNA to the helicase through the generic OB-fold-binding surface, but probably also from an active participation of the OB-fold module in the ATPase cycle of the helicase. However, although in our model, R708A and the β4–β5 loop evidently contact the RNA threaded in the helicase cavity, the β1–β2 loop is located at 15 Å away from the nucleic acid backbone (Figure 3D). These data show that RNA binding by a loop of the OB-fold distant from the helicase RNA-binding cavity is important for the RNA-stimulated ATPase activity. It is therefore possible that interaction with the OB-fold provides selectivity towards certain RNA structural features, which are not visible using our model substrate and that the Prp43p variants display a more pronounced reduction of binding to potential high affinity/specificity RNA targets present in the total yeast RNA sample used in ATPase assays. The precise *in vivo* RNA substrates of Prp43p will need to be identified to further investigate the interplay between RNA binding and ATPase activity.

The C-terminal domain of Prp43p is involved in binding the G-patch protein Pfa1p

In yeast Prp2p, deletion of the C-terminal 42 amino acids, which removes the αC2 and αC3 C-terminal helices and two strands of the OB-fold domain, results in a protein that can no longer interact with the spliceosome *in vitro* or function *in vivo* (Edwalds-Gilbert *et al*, 2004). Moreover, the altered proteins Prp2p-D845N/C846Y (OB-fold β5) or Prp2p-W854A/L855A (αC2 helix) are no longer able to bind to the G-patch protein GST-Spp2p, indicating that the C-terminal domain is implicated in binding (Silverman *et al*, 2004).

As the C-terminal domain containing the OB-fold is specific to DEAH/RHA helicases and given that interaction with G-patch proteins seems to be a conserved feature of these

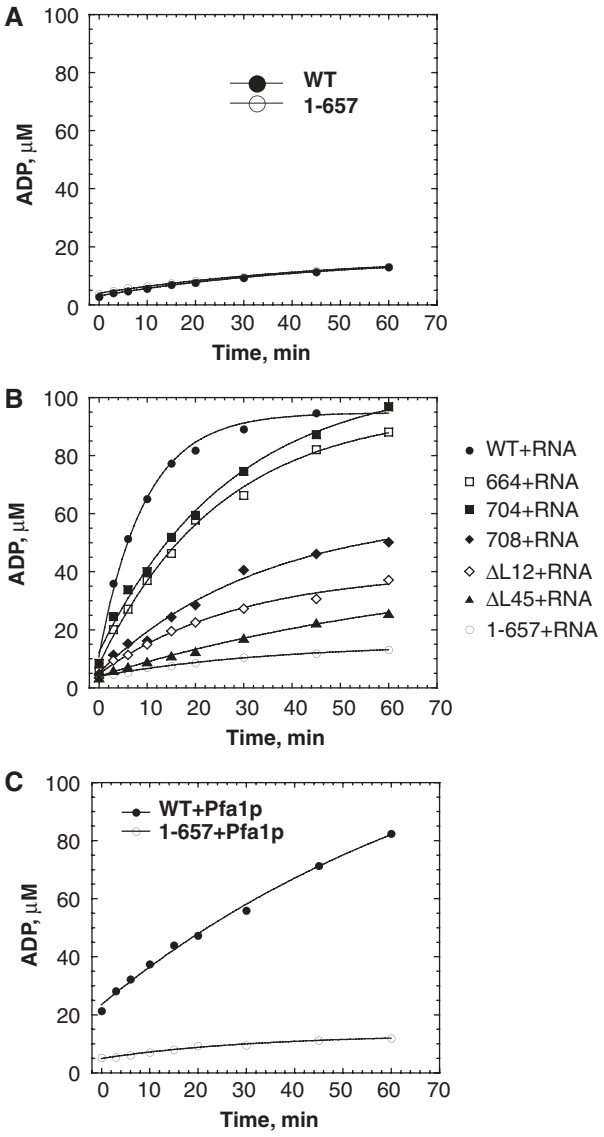

Figure 5 Analysis of the ATPase activity of Prp43p variants. ATP hydrolysis time courses were performed with the indicated proteins either in the absence of RNA (**A**), in the presence of RNA (**B**) or in the absence of RNA and presence of Pfa1p (**C**). Wild-type or Prp43p variants at 100 nM were incubated for the indicated times with cold ATP at 100 μM, mixed with α-^{32}P-ATP tracer. In addition, total yeast RNA was added at 150 μM in (**B**) and Pfa1p at 500 nM in (**C**). Aliquots of reaction mixes were then subjected to thin layer chromatography. ATP and ADP were quantified by phosphorimager to derive ADP concentrations. The abbreviations 664, 704, 708, ΔL12, ΔL45 and 1–657 refer to Prp43pR664A-His, Prp43pK704A-His, Prp43pR708A-His, Prp43pΔL12-His, Prp43pΔL45-His and Prp43p(1–657)-His, respectively.

enzymes, we reasoned that the C-terminal domain might contain the binding site for the G-patch. To test this hypothesis, we performed protein pull-down assays using wild type or truncated versions of Prp43p and Pfa1p (Figure 6). We have shown earlier that the N-terminal (Pfa1p(1–202)) and the G-patch-containing C-terminal (Pfa1p(574–767)) domains of Pfa1p can independently directly bind to Prp43p (Lebaron *et al*, 2009). Here, we show that, although the N-terminal domain of Pfa1p is still able to interact with Prp43p(1–657) (Figure 6, lane 11), the G-patch-containing C-terminal domain of Pfa1p is unable to do so (Figure 6, lanes 13 and 14). These results further confirm that Prp43p(1–657) is properly folded, as it can interact efficiently with Pfa1p(1–202). More importantly, they show that the G-patch-containing C-terminal domain of Pfa1p binds to the C-terminal domain of Prp43p.

We have shown earlier that Pfa1p activates the ATPase and helicase activities of Prp43p both in the absence and presence of RNA through its C-terminal G-patch-containing domain (Lebaron *et al*, 2009). From the above results, we would predict that the basal ATPase activity of Prp43p(1–657) observed in the absence of RNA should not be stimulated by Pfa1p, as the G-patch-containing domain of Pfa1p cannot interact with Prp43p(1–657). Indeed, although the ATPase activity of full-length Prp43p in the absence of RNA is stimulated ~10-fold by Pfa1p, the ATPase activity of Prp43p(1–657) is not stimulated (Figure 5C). Altogether, these results show that removal of the C-terminal domain of Prp43p renders it unresponsive to stimulation by Pfa1p, because this truncated Prp43p can no longer interact with the G-patch-containing domain of Pfa1p.

Discussion

The structure of Prp43p reveals a novel domain architecture that makes it the founding structure of the DEAH/RHA helicase family. These proteins group three modules that have never been associated before: a DEAH/RHA helicase core similar to viral DEAH helicases, a WH and ratchet domain with similarity to Ski2-like Hel308 helicase and an OB-fold domain. The structure of Prp43p also reveals the folding of the Prp43p-specific N-terminal domain before the helicase core. DEAH/RHA helicases differ in their N-terminal extensions, ranging from 100 residues (Prp43p/Prp2p) to >500 residues (Prp16p/Prp22p). Sequence alignments (not shown) suggest that the Prp43p-specific N-terminal extension including the αN1 and αN2 helices are poorly conserved in yeast genomes and are probably altogether absent in higher eukaryotes, whereas αN3 is probably present in Prp43p orthologues of other organisms. This is corroborated by the fact that the deletion of the first 90 residues of Prp43p does not impair cell growth, indicating that the αN1, αN2 and αN3 helices, which are removed in this mutant, are dispensable for optimal *in vivo* activity (Martin *et al*, 2002). Further deletion to residue 104 leads to a lethal phenotype (Martin *et al*, 2002). Our structure reveals that this truncation removes half of the first helix of the RecA1 domain, and therefore probably destabilizes the protein.

The finding of the Hel308 scaffold within the structure of Prp43p has important implications for the mechanism of action of all the DEAH/RHA helicases. The presence of the single-stranded nucleic acid-binding cavity, a ratchet helix in the ratchet domain and a β-hairpin in RecA2 strongly suggest that DEAH/RHA helicases function in the same manner as Hel308. This mechanism involves threading a single strand inside the cavity, translocating along this strand in a 3′ to 5′ direction and unwinding the double helix by separating the two strands on the β-hairpin physical barrier. This is consistent with the function of Prp22p, which is proposed to remove the mRNA from the spliceosome by translocating along the mRNA and therefore disrupting mRNA-U5 snRNP interaction (Schwer, 2008). Prp43p might function in an analogous manner by translocation along the lariat intron to strip-associated proteins and/or RNA. In ribosome

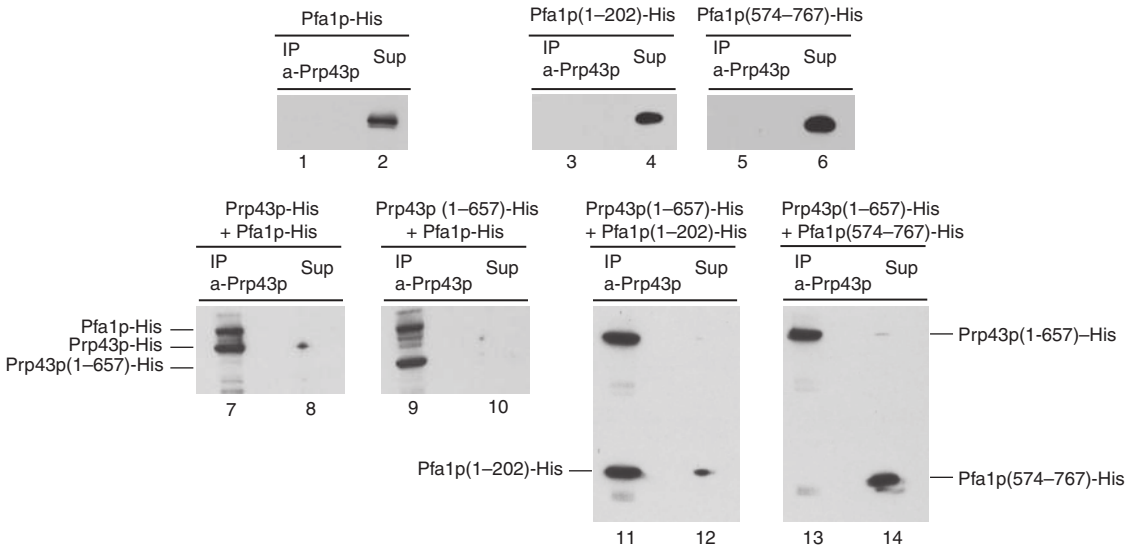

Figure 6 Prp43p(1–657) fails to interact with the G-patch-containing domain of Pfa1p. The indicated combinations of His-tagged wild-type or truncated Prp43p and Pfa1p proteins were incubated at 33 nM for the wild type and 130 nM for the truncated proteins and precipitated with anti-Prp43p antibodies. Proteins in the pellet (IP) and supernatant (Sup) fractions were detected by western blot analysis using anti-histidine antibodies.

biogenesis, Prp43p is suggested to restructure the rRNA around helix 44 and/or 45 to allow Nob1p access to the D cleavage site (Pertschy *et al*, 2009). Surprisingly, Prp43p has a rather poor helicase activity and preferentially unwinds 5′–3′ *in vitro* with the tested model substrates (Tanaka and Schwer, 2006). However, addition of either Pfa1p (Lebaron *et al*, 2009) or Ntr1p (Tanaka *et al*, 2007) boosts Prp43p activity and allows it to unwind both 5′–3′ and 3′–5′ model substrates. How Prp43p protein cofactors influence Prp43p helicase activity and processivity remains to be uncovered and will provide additional clues to the regulation and function of these proteins in splicing and ribosome biogenesis.

Recent reports have revealed the presence of the WH and ratchet domains in the Sec63 domains of Brr2, a Ski2-like helicase required for spliceosome activation and disassembly (Pena *et al*, 2009; Zhang *et al*, 2009). The Sec63 domains had a poor sequence similarity with the corresponding domains of Hel308 and Prp43p. It therefore seems that the Hel308 domain organization is a common architecture used by both Ski2-like and DEAH/RHA families. The common 'Hel308 scaffold' is supplemented by a C-terminal domain, which has a different structure and position in Hel308, Brr2 and Prp43p, and could contribute to modulate the helicase activity in the different families. Prp43p is the first example of a SF2 helicase containing an OB-fold domain, but OB-fold domains have been found associated with helicase domains from other superfamilies (Enemark and Joshua-Tor, 2008). The N-terminal domain of two archaeal orthologues of the MCM protein, a SF3 helicase, contains an OB-fold domain, which is crucial for hexamerization of the protein and is proposed to have a function in interacting with extruded single-stranded DNA (Fletcher *et al*, 2003; Liu *et al*, 2008). The hexameric RuvBL1, a SF6 helicase, also possesses an OB-fold at its N-terminus, which may have a similar function in binding ssDNA, but it is not implicated in oligomerization (Matias *et al*, 2006). The structural context of the OB-fold in the monomeric Prp43p protein is therefore unique to the DEAH/RHA family.

The structure of Prp43p shows that the OB-fold domain is positioned at the entrance of the RNA-binding cavity, and we show that the OB-fold-containing C-terminal domain of Prp43p has an important function in RNA binding and consistently, in efficient RNA-stimulated ATPase activity. This is in stark contrast to Hel308, which has a higher helicase activity when the 'autoinhibitory' C-terminal domain is removed (Richards *et al*, 2008). DEAH/RHA helicases often act in the context of large ribonucleoprotein complexes. Prp22p was found to bind a region of the RNA different to the region on which it translocates, suggesting that Prp22p contains two RNA-binding sites (Schwer, 2008). It is therefore possible that the OB-fold-containing domain provides a specific docking module, binding a different RNA to the RNA strand being loaded in the helicase active site. We have shown that the OB-fold-containing domain also constitutes the binding site for the regulatory G-patch-containing protein Pfa1p and is required for Prp43p activation by Pfa1p. It should be noted that Pfa1p (Lebaron *et al*, 2009), and possibly other regulatory proteins, also contain their own RNA-binding domain and might therefore direct the helicases to specific target RNA. DEAH/RHA helicases involved in splicing are usually stably bound to the spliceosome complex until their activity is triggered by an unknown mechanism (Schwer, 2008). We propose that the conserved C-terminal domain of DEAH/RHA helicases serves as a hub providing a link between RNA binding, G-patch protein interaction and helicase activity. The precise mechanisms by which this regulation is achieved will now need to be addressed, as understanding of how these proteins regulate DEAH/RHA helicases is important for human health, especially in light of the finding that the G-patch2 protein, recently shown to regulate human Prp43 function, is overexpressed in breast carcinogenesis (Lin *et al*, 2009).

Materials and methods

Cloning, expression and purification

The transformed *E. coli* strain SE1 with plasmid pSL18 (Lebaron *et al*, 2009) was used to produce recombinant Prp43p-6His. Expression cells were grown at 37°C in 2 × YT medium (MP Biomedicals) supplemented with ampicillin at 50 μg/ml until an OD600 nm of 0.8. Prp43p protein expression was then induced with 0.5 mM isopropyl β-D-thiogalactopyranoside and the cell culture was further incubated at 15°C overnight. Cells were harvested by centrifugation, resuspended in a 20 mM Tris–HCl pH 7.5, 200 mM NaCl, 10 mM β-mercaptoethanol buffer and stored at −20°C. Cell lysis was completed by sonication, and the cell extract was centrifuged at 20 000 *g* for 30 min at 4°C. The Prp43p His-tagged protein from the soluble fraction was purified on a HisTrapTM FF column (Amersham Biosciences) eluted with a linear imidazole gradient, followed by a gel filtration step on a SuperdexTM200 column (Amersham Biosciences). Selenomethionine-substituted protein was produced as described in Leulliot *et al* (2008) and purified as the native protein. Both protein samples were stored in 20 mM Tris–HCl pH 7.5, 200 mM NaCl, 10 mM β-mercaptoethanol at 4°C.

A fragment of the PRP43 open reading frame lacking the sequence encoding amino acids 658–767, followed by the sequence encoding 6-histidines and a stop codon and containing the XbaI and XhoI restriction enzyme sites, was amplified using plasmid pSL18 (Lebaron *et al*, 2009) and oligonucleotides 5′Prp43-DC658 (5′AGCAAAGTTACCCAGAAATTTTACGTTCCAATTTATCCTCCACCG3′) and 3′Prp43-DC658 (5′CCCCCCTCGAGTCAGTGGTGGTGGTGGTGG TGGAAAAACCCAGACGCAAGAGCCTTTCTGATGTTGTCAAAGTATT TAGG3′). This PCR fragment was cut with XbaI and XhoI and used to replace the XbaI-XhoI PRP43 fragment of pSL18, creating plasmid pMFX. This plasmid was transformed into *E. coli* strain SE1, and the resulting transformed strain was used to produce recombinant Prp43p(1–657) as described above for wild-type Prp43p.

Five Prp43p variants, that is three point mutants, namely R664A, K704A and R708A, and two variants combining amino-acid substitutions and deletions, namely Prp43pΔL12 (combined K663A, R664G mutations and deletion of residues 665–668) and Prp43pΔL45 (combined T702G, S703G mutations and K704 deletion) were generated commercially by GenScript Corporation (Piscataway, NJ) in the same vector as wild type (pSL18). All these Prp43p variants were produced as described above for wild-type Prp43p. ATPase assays, band retardation assays and protein–protein interaction assays were performed as described in Lebaron *et al* (2009).

Crystallization and data collection

Both native and mutant proteins (6 mg/ml) crystallized at 18°C by the hanging drop vapour diffusion method from a 1:1 mixture of protein, 1 mM ADP, 1 mM $MgCl_2$ and precipitant containing 8–11% PEG 4000, 100 mM ammonium acetate, 50 mM sodium cacodylate pH 6.5. Crystals were transferred in the mother liquor containing 30% glycerol before flash freezing in liquid nitrogen.

X-ray diffraction data from a native crystal and from a crystal of the mutant SeMet-substituted protein were collected on the Proxima 1 beamline at SOLEIL synchrotron (Gif-sur-Yvette, France) and were processed using MOSFLM (Leslie, 1992) and SCALA (Evans, 2006), and the XDS package (Kabsch, 1993), respectively. Both crystals belong to the P3221 space group with two molecules per asymmetric unit. The cell parameters and data collection statistics are reported in Table I.

Structure solution and refinement

The structure was solved by SAD using 2.8 Å data collected at the Selenium peak wavelength. The Selenium sites were initially found with SHELXD (Schneider and Sheldrick, 2002). The phases were determined in AUTOSHARP (Vonrhein *et al*, 2007) and improved with solvent flattening using SOLOMON (Abrahams and Leslie, 1996). Automatic model building using RESOLVE (Terwilliger, 2004) and ARP/WARP (Perrakis, 1999) allowed to trace 90% of the visible residues. The model was manually completed using O (Jones and Kjeldgaard, 1991) and COOT (Emsley and Cowtan, 2004) and refined to 1.9 Å resolution using REFMAC (Murshudov *et al*, 1997) from the high-resolution native data set.

Refinement statistics are shown in Table I. Electron density was not visible for residues 1–2 and 750–767 and 6-histidine tag, probably due to structural disorder, and are therefore absent from the model. All the residues fall within the allowed regions of the Ramachandran plot as defined by MOLPROBITY (Lovell *et al*, 2003). From the analysis of the electron density maps, the two ADP/Mg^{2+} ligands were unambiguously positioned. Two acetate ions, two nickel ions and two glycerol molecules could also be modelled in the structure. The coordinates have been deposited at the PDB under the accession number 2xau.

Supplementary data

Supplementary data are available at *The EMBO Journal* Online (http://www.embojournal.org).

Acknowledgements

We thank P Legrand and A Thompson (Soleil synchrotron, Gif-sur-Yvette) for their help in data collection and M Grigoriev (LBME) for help with figures, data interpretation and critical reading of the manuscript. Work at Université Paris-Sud was supported by the European 3D-repertoire program (LSHG-CT-2005-512028) and BIORIB grant (BLAN07-1_194553) from the Agence Nationale de la Recherche (ANR Blanche). Work at Université Paul Sabatier/LBME was supported by the CNRS, Université Paul Sabatier and grants from the Agence Nationale de la Recherche (RIBEUC) and the Ligue Contre le Cancer ('équipe labellisée').

Conflict of interest

The authors declare that they have no conflict of interest.

References

Abrahams JP, Leslie AG (1996) Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta Crystallogr D Biol Crystallogr* **52:** 30–42

Aravind L, Koonin EV (1999) G-patch: a new conserved domain in eukaryotic RNA-processing proteins and type D retroviral polyproteins. *Trends Biochem Sci* **24:** 342–344

Arcus V (2002) OB-fold domains: a snapshot of the evolution of sequence, structure and function. *Curr Opin Struct Biol* **12:** 794–801

Arenas JE, Abelson JN (1997) Prp43: An RNA helicase-like factor involved in spliceosome disassembly. *Proc Natl Acad Sci USA* **94:** 11798–11802

Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci USA* **98:** 10037–10041

Bleichert F, Baserga SJ (2007) The long unwinding road of RNA helicases. *Mol Cell* **27:** 339–352

Boon KL, Auchynnikava T, Edwalds-Gilbert G, Barrass JD, Droop AP, Dez C, Beggs JD (2006) Yeast ntr1/spp382 mediates prp43 function in postspliceosomes. *Mol Cell Biol* **26:** 6016–6023

Buttner K, Nehring S, Hopfner KP (2007) Structural basis for DNA duplex separation by a superfamily-2 helicase. *Nat Struct Mol Biol* **14:** 647–652

Caruthers JM, McKay DB (2002) Helicase structure and mechanism. *Curr Opin Struct Biol* **12:** 123–133

Combs DJ, Nagel RJ, Ares Jr M, Stevens SW (2006) Prp43p is a DEAH-box spliceosome disassembly factor essential for ribosome biogenesis. *Mol Cell Biol* **26:** 523–534

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

Edwalds-Gilbert G, Kim DH, Silverman E, Lin RJ (2004) Definition of a spliceosome interaction domain in yeast Prp2 ATPase. *RNA* **10:** 210–220

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

Enemark EJ, Joshua-Tor L (2008) On helicases and other motor proteins. *Curr Opin Struct Biol* **18:** 243–257

Erkizan HV, Kong Y, Merchant M, Schlottmann S, Barber-Rotenberg JS, Yuan L, Abaan OD, Chou TH, Dakshanamurthy S, Brown ML, Uren A, Toretsky JA (2009) A small molecule blocking oncogenic

protein EWS-FLI1 interaction with RNA helicase A inhibits growth of Ewing's sarcoma. *Nat Med* **15:** 750–756
Evans P (2006) Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* **62:** 72–82
Fletcher RJ, Bishop BE, Leon RP, Sclafani RA, Ogata CM, Chen XS (2003) The structure and function of MCM from archaeal *M. Thermoautotrophicum*. *Nat Struct Biol* **10:** 160–167
Guglielmi B, Werner M (2002) The yeast homolog of human PinX1 is involved in rRNA and small nucleolar RNA maturation, not in telomere elongation inhibition. *J Biol Chem* **277:** 35712–35719
Jankowsky E, Fairman ME (2007) RNA helicases—one fold for many functions. *Curr Opin Struct Biol* **17:** 316–324
Jones TA, Kjeldgaard M (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Cryst A* **47:** 110–119
Kabsch W (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J Appl Cryst* **26:** 795–800
Kim JL, Morgenstern KA, Griffith JP, Dwyer MD, Thomson JA, Murcko MA, Lin C, Caron PR (1998) Hepatitis C virus NS3 RNA helicase domain with a bound oligonucleotide: the crystal structure provides insights into the mode of unwinding. *Structure* **6:** 89–100
Koo JT, Choe J, Moseley SL (2004) HrpA, a DEAH-box RNA helicase, is involved in mRNA processing of a fimbrial operon in *Escherichia coli*. *Mol Microbiol* **52:** 1813–1826
Krissinel E, Henrick K (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr* **60:** 2256–2268
Lebaron S, Froment C, Fromont-Racine M, Rain JC, Monsarrat B, Caizergues-Ferrer M, Henry Y (2005) The splicing ATPase prp43p is a component of multiple preribosomal particles. *Mol Cell Biol* **25:** 9269–9282
Lebaron S, Papin C, Capeyrou R, Chen YL, Froment C, Monsarrat B, Caizergues-Ferrer M, Grigoriev M, Henry Y (2009) The ATPase and helicase activities of Prp43p are stimulated by the G-patch protein Pfa1p during yeast ribosome biogenesis. *EMBO J* **28:** 3808–3819
Leeds NB, Small EC, Hiley SL, Hughes TR, Staley JP (2006) The splicing factor Prp43p, a DEAH box ATPase, functions in ribosome biogenesis. *Mol Cell Biol* **26:** 513–522
Leslie AGW (1992) Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography* **26:** 27–33
Leulliot N, Chaillet M, Durand D, Ulryck N, Blondeau K, van Tilbeurgh H (2008) Structure of the yeast tRNA m7G methylation complex. *Structure* **16:** 52–61
Lin ML, Fukukawa C, Park JH, Naito K, Kijima K, Shimo A, Ajiro M, Nishidate T, Nakamura Y, Katagiri T (2009) Involvement of G-patch domain containing 2 overexpression in breast carcinogenesis. *Cancer Sci* **100:** 1443–1450
Liu W, Pucci B, Rossi M, Pisani FM, Ladenstein R (2008) Structural analysis of the *Sulfolobus solfataricus* MCM protein N-terminal domain. *Nucleic Acids Res* **36:** 3235–3243
Lovell SC, Davis IW, Arendall III WB, de Bakker PI, Word JM, Prisant MG, Richardson JS, Richardson DC (2003) Structure validation by Calpha geometry: phi, psi and Cbeta deviation. *Proteins* **50:** 437–450
Luo D, Xu T, Watson RP, Scherer-Becker D, Sampath A, Jahnke W, Yeong SS, Wang CH, Lim SP, Strongin A, Vasudevan SG, Lescar J (2008) Insights into RNA unwinding and ATP hydrolysis by the flavivirus NS3 protein. *EMBO J* **27:** 3209–3219
Martin A, Schneider S, Schwer B (2002) Prp43 is an essential RNA-dependent ATPase required for release of lariat-intron from the spliceosome. *J Biol Chem* **277:** 17743–17750
Matias PM, Gorynia S, Donner P, Carrondo MA (2006) Crystal structure of the human AAA+ protein RuvBL1. *J Biol Chem* **281:** 38918–38929
Mefford MA, Staley JP (2009) Evidence that U2/U6 helix I promotes both catalytic steps of pre-mRNA splicing and rearranges in between these steps. *RNA* **15:** 1386–1397
Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53:** 240–255
Murzin AG (1993) OB (oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO J* **12:** 861–867
Oyama T, Oka H, Mayanagi K, Shirai T, Matoba K, Fujikane R, Ishino Y, Morikawa K (2009) Atomic structures and functional implications of the archaeal RecQ-like helicase Hjm. *BMC Struct Biol* **9:** 2
Pena V, Jovin SM, Fabrizio P, Orlowski J, Bujnicki JM, Luhrmann R, Wahl MC (2009) Common design principles in the spliceosomal RNA helicase Brr2 and in the Hel308 DNA helicase. *Mol Cell* **35:** 454–466
Perrakis A (1999) Automated protein model building combined with iterative structure refinement. *Nat Struct Biol* **6:** 458–463
Pertschy B, Schneider C, Gnadig M, Schafer T, Tollervey D, Hurt E (2009) RNA helicase Prp43 and its co-factor Pfa1 promote 20S to 18S rRNA processing catalyzed by the endonuclease Nob1. *J Biol Chem* **284:** 35079–35091
Pyle AM (2008) Translocation and unwinding mechanisms of RNA and DNA helicases. *Annu Rev Biophys* **37:** 317–336
Richards JD, Johnson KA, Liu H, McRobbie AM, McMahon S, Oke M, Carter L, Naismith JH, White MF (2008) Structure of the DNA repair helicase hel308 reveals DNA binding and autoinhibitory domains. *J Biol Chem* **283:** 5118–5126
Sanjuan R, Marin I (2001) Tracing the origin of the compensasome: evolutionary history of DEAH helicase and MYST acetyltransferase gene families. *Mol Biol Evol* **18:** 330–343
Schneider TR, Sheldrick GM (2002) Substructure solution with SHELXD. *Acta Crystallogr D Biol Crystallogr* **58:** 1772–1779
Schwer B (2001) A new twist on RNA helicases: DExH/D box proteins as RNPases. *Nat Struct Biol* **8:** 113–116
Schwer B (2008) A conformational rearrangement in the spliceosome sets the stage for Prp22-dependent mRNA release. *Mol Cell* **30:** 743–754
Sengoku T, Nureki O, Nakamura A, Kobayashi S, Yokoyama S (2006) Structural basis for RNA unwinding by the DEAD-box protein *Drosophila* Vasa. *Cell* **125:** 287–300
Shiratori A, Shibata T, Arisawa M, Hanaoka F, Murakami Y, Eki T (1999) Systematic identification, classification, and characterization of the open reading frames which encode novel helicase-related proteins in *Saccharomyces cerevisiae* by gene disruption and Northern analysis. *Yeast* **15:** 219–253
Silverman EJ, Maeda A, Wei J, Smith P, Beggs JD, Lin RJ (2004) Interaction between a G-patch protein and a spliceosomal DEXD/H-box ATPase that is critical for splicing. *Mol Cell Biol* **24:** 10101–10110
Tanaka N, Aronova A, Schwer B (2007) Ntr1 activates the Prp43 helicase to trigger release of lariat-intron from the spliceosome. *Genes Dev* **21:** 2312–2325
Tanaka N, Schwer B (2006) Mutations in PRP43 that uncouple RNA-dependent NTPase activity and pre-mRNA splicing function. *Biochemistry* **45:** 6510–6521
Tanner NK, Cordin O, Banroques J, Doere M, 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
Tanner NK, Linder P (2001) DExD/H box RNA helicases: from generic motors to specific dissociation functions. *Mol Cell* **8:** 251–262
Terwilliger T (2004) SOLVE and RESOLVE: automated structure solution, density modification and model building. *J Synchrotron Radiat* **11:** 49–52
Theobald DL, Mitton-Fry RM, Wuttke DS (2003) Nucleic acid recognition by OB-fold proteins. *Annu Rev Biophys Biomol Struct* **32:** 115–133
Tsai RT, Fu RH, Yeh FL, Tseng CK, Lin YC, Huang YH, Cheng SC (2005) Spliceosome disassembly catalyzed by Prp43 and its associated components Ntr1 and Ntr2. *Genes Dev* **19:** 2991–3003
Tsai RT, Tseng CK, Lee PJ, Chen HC, Fu RH, Chang KJ, Yeh FL, Cheng SC (2007) Dynamic interactions of Ntr1-Ntr2 with Prp43 and with U5 govern the recruitment of Prp43 to mediate spliceosome disassembly. *Mol Cell Biol* **27:** 8027–8037
Vonrhein C, Blanc E, Roversi P, Bricogne G (2007) Automated structure solution with autoSHARP. *Methods Mol Biol* **364:** 215–230
Warrener P, Tamura JK, Collett MS (1993) RNA-stimulated NTPase activity associated with yellow fever virus NS3 protein expressed in bacteria. *J Virol* **67:** 989–996
Wu J, Bera AK, Kuhn RJ, Smith JL (2005) Structure of the Flavivirus helicase: implications for catalytic activity, protein interactions, and proteolytic processing. *J Virol* **79:** 10268–10277
Zhang L, Xu T, Maeder C, Bud LO, Shanks J, Nix J, Guthrie C, Pleiss JA, Zhao R (2009) Structural evidence for consecutive Hel308-like modules in the spliceosomal ATPase Brr2. *Nat Struct Mol Biol* **16:** 731–739

