

## Review

# The function and architecture of DEAH/RHA helicases

Yangzi He<sup>1,2</sup>, Gregers R. Andersen<sup>1,2,\*</sup> and Klaus H. Nielsen<sup>1,2,\*</sup>

<sup>1</sup>Department of Molecular Biology, Aarhus University, Gustav Wieds Vej 10C, DK-8000 Aarhus C, Denmark

<sup>2</sup>Center for mRNP Biogenesis and Metabolism, Aarhus University, C.F. Møllers Allé 3, DK-8000 Aarhus C, Denmark

\*Corresponding authors

e-mail: gra@mb.au.dk; khn@mb.au.dk

## Abstract

Helicases are ubiquitous enzymes that participate in every aspect of nucleic acid metabolism. The DEAH/RHA family of helicases are involved in a variety of cellular processes including transcriptional and translational regulation, pre-mRNA splicing, pre-rRNA processing, mRNA export and decay, in addition to the innate immune response. Recently, the first crystal structures of a DEAH/RHA helicase unveiled the unique structural features of this helicase family. These structures furthermore illuminate the molecular mechanism of these proteins and provide a framework for analysis of their interaction with nucleic acids, regulatory proteins and large macromolecular complexes.

**Keywords:** DEAH/RHA helicase; pre-mRNA splicing; ribosome biogenesis; RNA binding site; RNA helicase A (RHA).

## Introduction

Helicases are essential enzymes that use nucleoside triphosphates (NTPs) to bind or remodel nucleic acids and nucleic acid-protein complexes (1–4). They are involved in almost all aspects of nucleic acid metabolism. Based on their sequence, structural and mechanistic features helicases can be classified into six superfamilies (SFs) (3, 4). SFs 3–6 consist of toroidal enzymes, whereas those not forming rings comprise SFs 1 and 2 (3). SF1 and SF2 helicases can be further classified into a number of distinct protein families (5).

The DEAH/RNA Helicase A (RHA) family of helicases are found in all eukaryotes and are members of SF2. The DEAH proteins were first identified as a distinct family in the early 1990s when several pre-mRNA processing (PRP) mutants were characterized in *Saccharomyces cerevisiae*, namely Prp2p (6), Prp16p (7) and Prp22p (8). Shortly after, the human RNA Helicase A (RHA, also known as DHX9) was identified (9) and found to be homologous to the male-

less protein (MLE) in *Drosophila melanogaster* (10). The core structure of DEAH/RHA helicases was determined by recent crystal structures of yeast Prp43p (11, 12), a protein involved in pre-mRNA splicing and ribosome biogenesis (13–16). This review focuses on these two structures and how they help advance our understanding of the mechanisms and biological functions of this family of proteins.

## Biological processes involving DEAH/RHA helicases

In *S. cerevisiae* six helicases of the DEAH/RHA family have been characterized, all of which have orthologs in humans: Prp2p (DHX16 in humans), Prp16p (DHX38), Prp22p (DHX8), Prp43p (DHX15), Dhr1p (DHX37) and Dhr2p (DHX32). Prp2p, Prp16p, Prp22p and Prp43p are all essential factors involved in nuclear pre-mRNA splicing. Spliceosomes catalyze the excision of introns and ligation of exons from pre-mRNAs in a two-step transesterification reaction (17). Prp2p is required for activation of the spliceosome for the first chemical step (known as ‘5′ splice site cleavage’) (18, 19). It has been proposed to catalyze the ATP-dependent displacement of the SF3a and SF3b complexes from the 2′-OH of the branch point adenosine, allowing nucleophilic attack of the 5′ splice site (SS) (19). The Prp16p helicase promotes conformational rearrangements of the spliceosome during the transition between the first and the second (known as ‘exon ligation’) steps (20–22). After the 5′ SS cleavage it mediates the dissociation of Yju2 and Cwc25, two factors required for the first transesterification reaction, from the spliceosome (21). Prp22p promotes exon ligation at 3′ SS distal from the branch point sequence (23) and after the second step it furthermore mediates release of spliced mRNA from the spliceosome (23–25). The helicase was cross linked to the intron region just upstream of the 3′ SS prior to the second step (26). After exon ligation but before mRNA release, cross linking data suggest that Prp22p is located close to a segment of the mRNA downstream of the exon-exon junction. At this location, the helicase is proposed to disrupt the mRNA/U5 small nuclear ribonucleoprotein (snRNP) contacts using its 3′–5′ helicase activity, thereby releasing the mRNA from the spliceosome (25). Prp43p catalyzes spliceosome disassembly by promoting release of the intron lariat (13, 27, 28). Prp43p was significantly cross linked to U6 small nuclear RNA (snRNA) but not U2 and U5 snRNAs of the spliceosome (29). However, its precise site of action for lariat intron release and spliceosome disassembly is yet to be defined.

At least two DEAH helicases are also required for the kinetic proofreading mechanism of the spliceosome in order

to ensure fidelity during splicing (30). Prior to the first step, Prp16p proofreads 5' SS cleavage and discriminates against aberrant substrates with a mutated branch point sequence (21, 30–33). Prp22p proofreads exon ligation prior to the second step and discriminates against aberrant splicing intermediates with mutated 5' SS, branch point sequence, or 3' SS (34). These splicing substrates or intermediates rejected by Prp16p and Prp22p are stalled at the spliceosome until Prp43p catalyzes their release for subsequent degradation (32, 35).

In addition to splicing, DEAH helicases also play crucial roles in ribosome biogenesis, another major RNA processing pathway. Chemical modification and nucleolytic cleavage of the polycistronic 35S pre-rRNA are required to give rise to mature 18S, 5.8S and 25S rRNAs. Concomitantly, pre-rRNAs assemble with numerous ribosomal and non-ribosomal proteins and small nucleolar ribonucleoprotein particles (snoRNPs) (36). Specifically, Prp43p is essential for the processing of pre-rRNA of both the small and large ribosomal subunits (14–16). It promotes cleavage of the 20S pre-rRNA at site D by endonuclease Nob1, giving rise to mature 18S rRNA (37). In agreement with this model, Prp43p has been cross linked to helix 44 at the 3' end of 18S rRNA (29). Furthermore, Prp43p was cross linked to several sites on the 25S rRNA and to a number of Box C/D snoRNAs that guides 2'-O-methylation of specific rRNA nucleotides. Depletion or mutation of Prp43p blocked the release of several of these snoRNAs. These results suggest that Prp43p might facilitate the dissociation of certain snoRNAs from the preribosomes (29). Dhr1p and Dhr2p are involved in earlier steps of 18S rRNA synthesis. Both Dhr1p and Dhr2p are required for cleavage at sites A<sub>1</sub> and A<sub>2</sub>, whereas Dhr2p is also required for cleavage at site A<sub>0</sub> (38). The Box C/D snoRNP U3 is required for cleavage at these three sites (39, 40). Dhr1p stably associates with the U3 snoRNP and was proposed to remodel the interaction between U3 snoRNA and the 18S region of pre-rRNA (38).

The functions of Prp2p and Prp43p are regulated by proteins containing a G-patch domain, which is a conserved glycine-rich sequence approximately 48 amino acids in length, including six highly conserved glycines. This domain is found in many proteins involved in RNA processing (41). Spliceosomal activation by Prp2p requires its interaction with the G-patch protein Spp2p (42). Prp43p interacts with three G-patch proteins: Ntr1p, Pfa1p and Gno1p. Both Ntr1p and Pfa1p have been shown to bind directly to Prp43p and stimulate its NTPase activity (43, 44) and the association of Prp43p with the spliceosome requires Ntr1p (43). Pfa1p and Gno1p have been implicated to interact with Prp43p during ribosome biogenesis (14, 45) and activation of Prp43p by Pfa1p is required for efficient 20S to 18S rRNA processing (37, 44).

In addition to the homologues of the six yeast precursor RNA processing factors, three other DEAH helicases have been characterized in humans. RHA (DHX9) interacts with a number of cellular and viral transcription regulators (46–51), whereas its *D. melanogaster* homologue MLE is a transcriptional regulator required for male X-chromosome

dosage compensation (52, 53). RHA can unwind triple-helical DNA structures *in vitro* (54). It facilitates export and translation of several retroviral mRNAs (55–59) and senses microbial Class B CpG oligodeoxynucleotide in the cytosol of plasmacytoid dendritic cells (pDCs), triggering innate immune responses such as TNF- $\alpha$  and IL-6 production and NF- $\kappa$ B activation (60). DHX36 senses Class A CpG oligodeoxynucleotide in pDCs and triggers IFN- $\alpha$  production and the activation of IRF7 (60). This helicase is also known as RHAU-RNA helicase associated with AU-rich element (ARE), and was initially found to bind to the ARE of urokinase plasminogen activator mRNA (ARE<sup>uPA</sup>) and facilitate its deadenylation and decay (61). DHX36 has also been implicated in transcriptional regulation (62). This helicase recognizes DNA and RNA G-quadruplexes (G4) via its N-terminal RSM (RHAU-specific motif) domain and resolves the G4 structures (63, 64). The RSM domain is also required for the recruitment of RHAU to stress granules (65). Finally, the DHX29 helicase is required for cap-dependent translation initiation of mammalian mRNAs with structured 5' UTRs, ensuring base-by-base inspection of 5' UTRs for the AUG start codon by scanning 43S pre-initiation complexes (66–68). DHX29 has been shown to bind specifically to the 40S ribosomal subunit, likely near the mRNA entry channel (67). The known DEAH/RHA helicases and their respective biological functions are summarized (Table 1).

Apart from these eukaryotic helicases, there are two known prokaryotic DEAH/RHA helicases, HrpA and HrpB. Their exact function is not known but HrpA has been implicated in bacterial mRNA processing (69).

## The architecture of a DEAH/RHA helicase

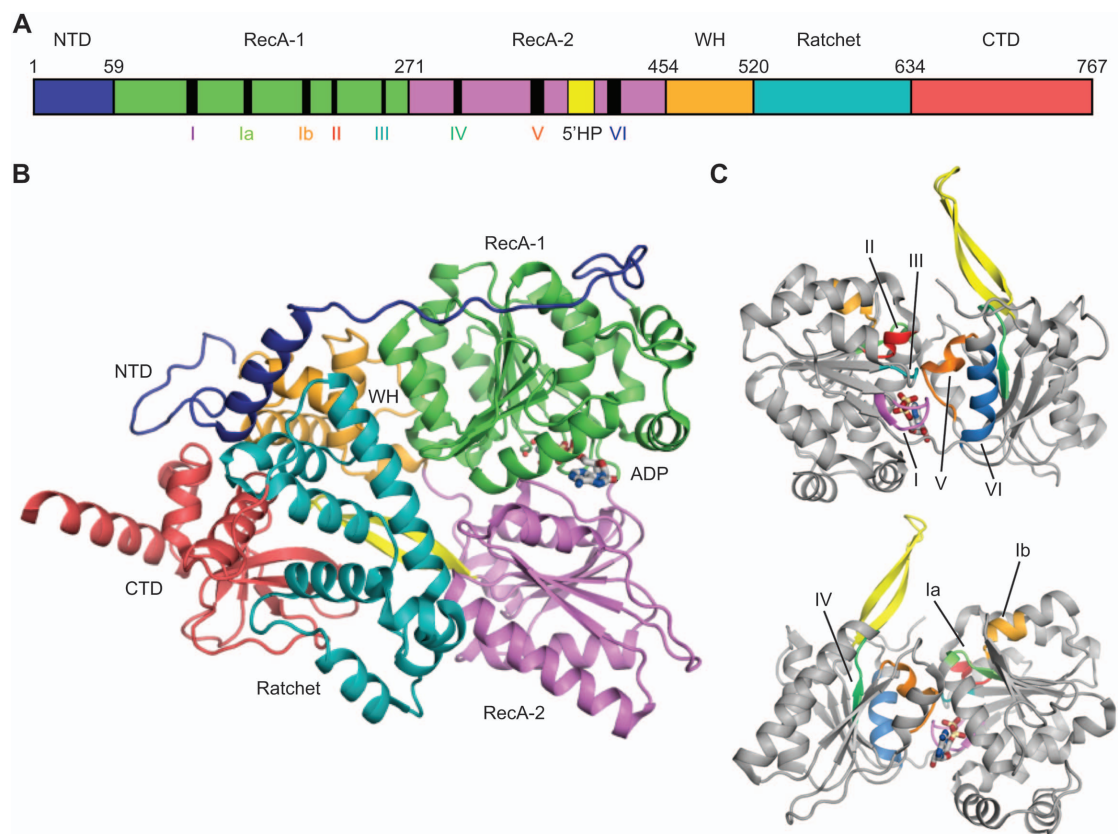
The common structural features of DEAH/RHA proteins were revealed by two recently published crystal structures of yeast Prp43p bound to ADP (11, 12) (Figures 1 and 2). These two structures (RCSB entries 3KX2 and 2XAU) are nearly identical and superimpose with a root mean square deviation of less than 0.3 Å. Residues 90–726 of Prp43p define the structural core of this helicase family, which comprises five domains: two conserved RecA-like helicase domains (RecA-1 and RecA-2) in the N-terminus, followed by a winged helix (WH) domain, a seven-helical bundle named the ratchet domain and a C-terminal domain (CTD) including a five-stranded  $\beta$ -barrel arranged in an oligonucleotide/oligosaccharide-binding (OB) fold. Prp43p-specific regions include the N-terminal domain (NTD, residues 1–58), the first two  $\alpha$ -helices of RecA-1 (residues 59–89) and the last two helices of the CTD (residues 727–755).

RecA-1 and RecA-2 contain eight conserved motifs found in all SF1 and SF2 helicases: Motifs I, Ia, Ib, II and III in RecA-1 and motifs IV, V and VI in RecA-2 (Figure 1A and C). Motif I is also referred to as the Walker A motif or the P-loop. Motif II, after which the DEAH helicase family is named, is also known as the Walker B motif (70) and both motifs I and II are found in many families of NTPases (71). In the structure of Prp43p, the ADP nucleotide is sandwiched

**Table 1** Biological functions of DEAH/RHA helicases.

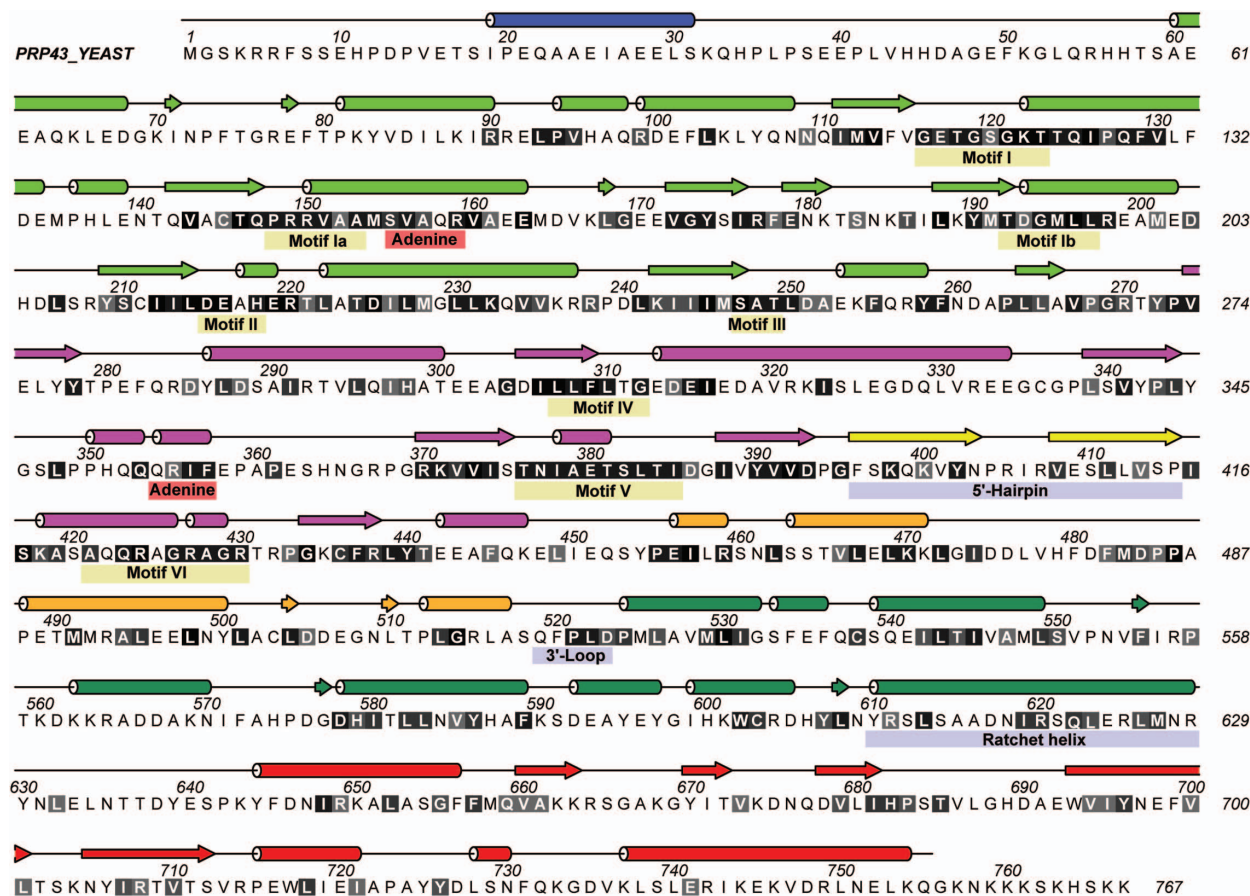
Protein	Cellular processes involved in	Reference
<i>S. cerevisiae</i>		
Prp2p	Pre-mRNA splicing	(18, 19, 41)
Prp16p	Pre-mRNA splicing	(20–22, 29–32)
Prp22p	Pre-mRNA splicing	(23–26)
Prp43p	Pre-mRNA splicing	(13, 27, 32, 34, 42)
	Ribosome biogenesis	(14–16, 28, 36)
Dhr1p	Ribosome biogenesis	(37)
Dhr2p	Ribosome biogenesis	(37)
<i>H. sapiens</i>		
DHX9 (RHA)	Transcription regulation	(45–50)
	Export and translation of retroviral mRNAs	(54–58)
	Innate immunity	(59)
DHX29	Translation initiation	(65–67)
DHX36 (RHAU)	Transcription regulation	(61)
	mRNA deadenylation and decay	(60)
	Innate immunity	(59)

The human orthologs of the yeast factors are not listed. Please refer to the main text.

**Figure 1** Structure of yeast Prp43p.

(A) Schematic of the six structural domains of Prp43p. Domain names and boundaries are indicated on top. Canonical helicase motifs in the RecA domains are underneath, annotated in the same color code as in (C). (B) Cartoon representation of Prp43p bound to ADP (RCSB entry 3KX2). (C) Close-up views of the two RecA domains containing eight conserved helicase motifs. The lower panel is rotated by 180° relative to the upper panel. Motif I is colored violet; Ia, lime; Ib, bright orange; II, red; III, teal; IV, lime green; V, orange; VI, sky blue; 5' HP, yellow.





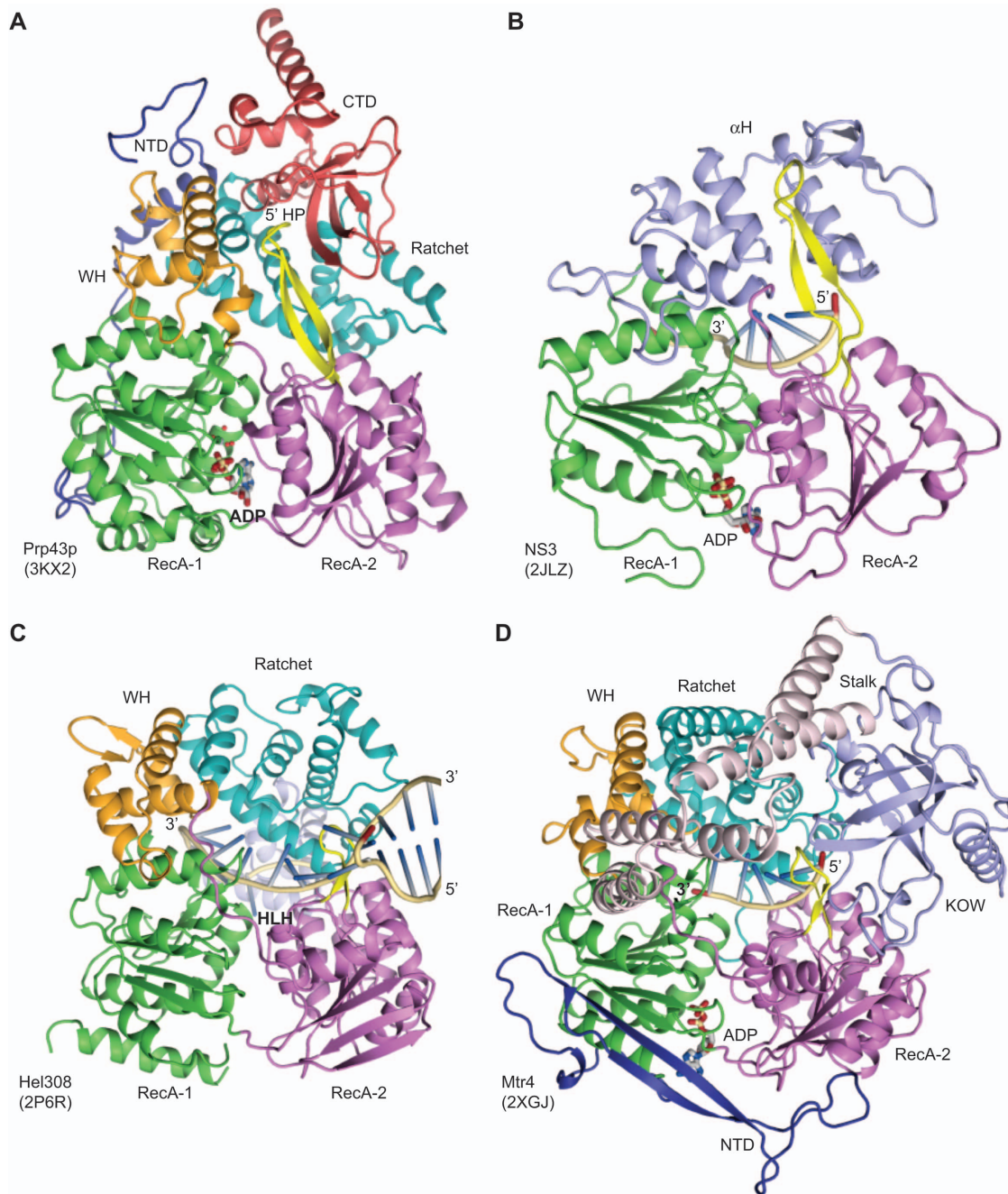
**Figure 2** Structural features of DEAH/RHA helicases mapped on yeast Prp43p with the sequence shaded according to conservation (black, most conserved; white, not conserved) after alignment of known human, *Drosophila*, *Arabidopsis* and yeast DEAH/RHA sequences (see below).

The canonical helicase motifs are indicated with a light brown bar below, residues forming the adenine binding pocket with a red bar and residues lining the putative single stranded RNA binding pocket with a gray-blue bar. The secondary structure is that observed in yeast Prp43. The Figure was prepared with ALINE. For alignment the following Swissprot entries were used: DHX8\_HUMAN, DHX9\_HUMAN, DHX15\_HUMAN, DHX16\_HUMAN, DHX29\_HUMAN, DHX30\_HUMAN, DHX32\_HUMAN, DHX33\_HUMAN, DHX34\_HUMAN, DHX35\_HUMAN, DHX36\_HUMAN, DHX38\_HUMAN, DHX40\_HUMAN, DHX57\_HUMAN, DQX1\_HUMAN, YTDC2\_HUMAN, A1Z9L3\_DROME, KZ\_DROME, MLE\_DROME, Q7K3M5\_DROME, Q8SWT2\_DROME, SPNE\_DROME, Q9VIZ3\_DROME, Q9VL25\_DROME, Q9VR29\_DROME, Q9VWI5\_DROME, Q9VX63\_DROME, Q9VY54\_DROME, Q9VZ55\_DROME, DHX15\_ARATH, O23511\_ARATH, Y4102\_ARATH, Q0WVI8\_ARATH, Q38800\_ARATH, DHX8\_ARATH, Q8VY00\_ARATH, Q93Y16\_ARATH, Q9C6G0\_ARATH, Q9C734\_ARATH, Q9C873\_ARATH, Q9C6F9\_ARATH, Q9FF84\_ARATH, Q9FZC3\_ARATH, Q9LXT8\_ARATH, Q9LZQ9\_ARATH, Q9SHK6\_ARATH, Q9SJ58\_ARATH, Q9SMG9\_ARATH, Q9ZU53\_ARATH, DHR1\_YEAST, DHR2\_YEAST, PRP2\_YEAST, PRP16\_YEAST, PRP22\_YEAST, PRP43\_YEAST, YL419\_YEAST.

between the two RecA domains. A  $\beta$ -hairpin named 5' HP (residues 396–415) is located between motifs V and VI of RecA-2. A similar hairpin, although of variable length, is present at the equivalent location in the flavivirus NS3 helicase (72) of the NS3/NPH-II family, in addition to archeal Hel308 (73) and yeast Mtr4 (74, 75), both members of the Ski-2 like helicase family (Figure 3). The similarity between the DEAH/RHA and the Ski-2 helicases extends even further as both the WH and ratchet domains are also present in Hel308 and Mtr4 (Figure 3). In Prp43p, the WH domain interacts extensively with RecA-I and the CTD is packed against the ratchet domain. In the ADP-bound state, the 5' HP is inserted into a cleft between the WH domain and CTD.

### Insights into DEAH/RHA helicase-nucleic acid interactions

The primary nucleic acid-binding sites of SF2 proteins are known to be located on the surface of the two RecA domains, opposite the nucleotide binding site (2, 4). Based on homology to structures of other SF2 proteins bound to nucleic acid substrates, the putative nucleic acid binding pocket of Prp43p has been proposed (11, 12). The core structure of DEAH/RHA helicases is highly homologous to that of Ski-2 like helicases and in Prp43p-ADP the relative orientation of the two RecA domains is similar to a semi-closed conformation of these domains observed in the structures of



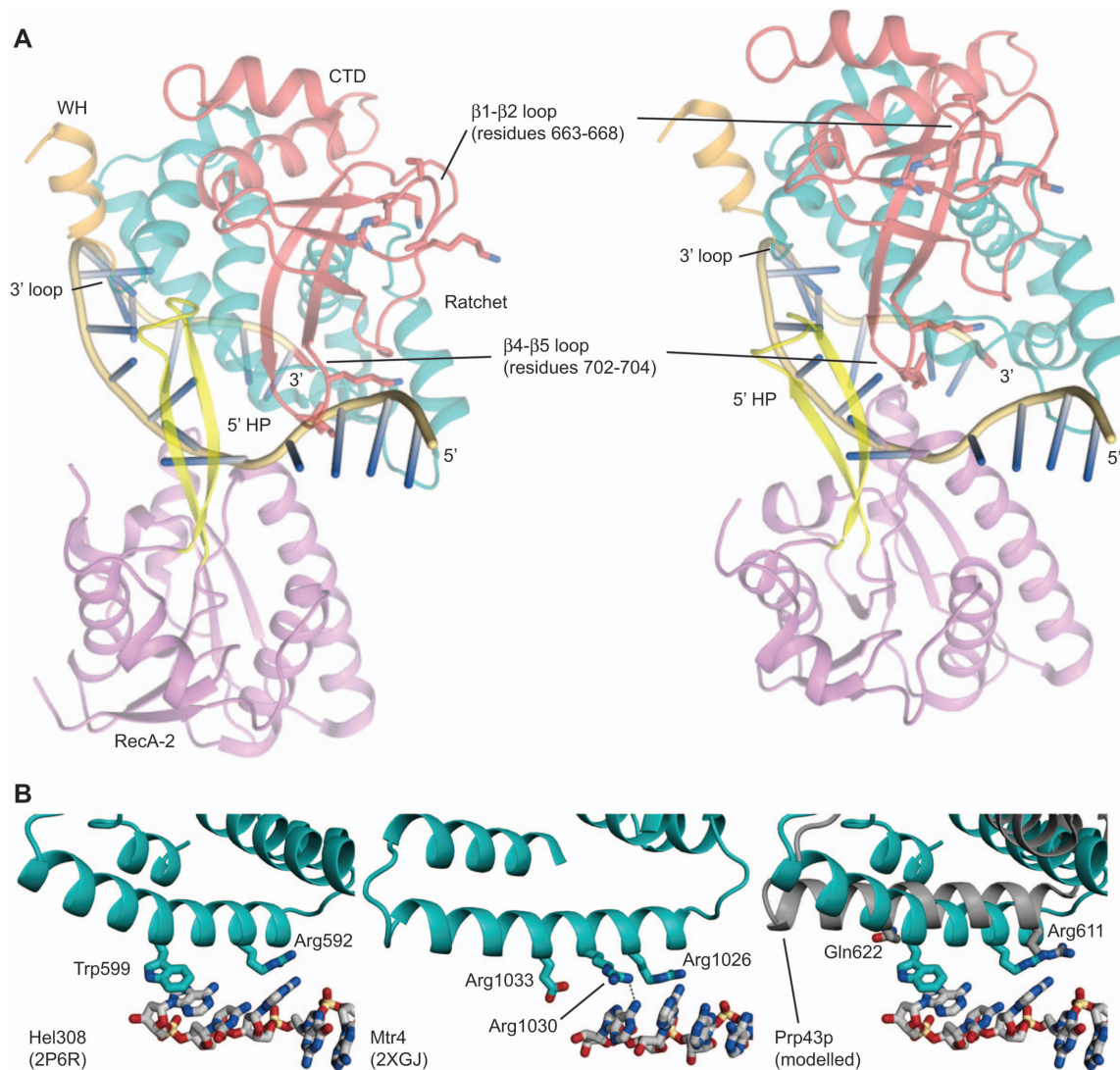
**Figure 3** Structural homology between DEAH/RHA helicases and other SF2 families.

(A) Structure of Prp43p-ADP (RCSB entry 3KX2). (B) Flavivirus NS3 helicase bound to ADP and RNA (RCSB entry 2JLZ). (C) Archeal Hel308 bound to DNA (RCSB entry 2P6R). (D) Yeast Mtr4 bound to ADP and RNA (RCSB entry 2XGJ). (B, C, D) are shown in the same orientation and color code for the common structural domains as in (A). The helical domain of NS3 ( $\alpha$ H in B), the helix-loop-helix domain of Hel308 (HLH in C) and the Kyprides-Ouzounis-Woese domain of Mtr4 (KOW in D) are shown in light blue and the stalk domain of Mtr4 in light pink.

Hel308 (in complex with a partially unwound DNA substrate) (73) and Mtr4 (in complex with ADP and a single-stranded RNA substrate) (75). We can therefore superimpose the RecA domains of Prp43p (RCSB entry 3KX2) and Hel308 (RCSB entry 2P6R), and use the longer of the two DNA strands in the Hel308 structure to identify the putative nucleic acid binding pocket of Prp43p. In the ADP-bound

state of Prp43p, the putative RNA binding site is occluded, with the 3' end of this binding pocket sealed by the loop connecting the WH and ratchet domains (which we term the '3' loop', Figure 2), and the 5' end blocked by the 5'HP and the  $\beta$ 4– $\beta$ 5 loop in the OB-fold of CTD (Figure 4A). This was experimentally verified by pull-down experiments demonstrating that ADP prevented Prp43p interaction with a





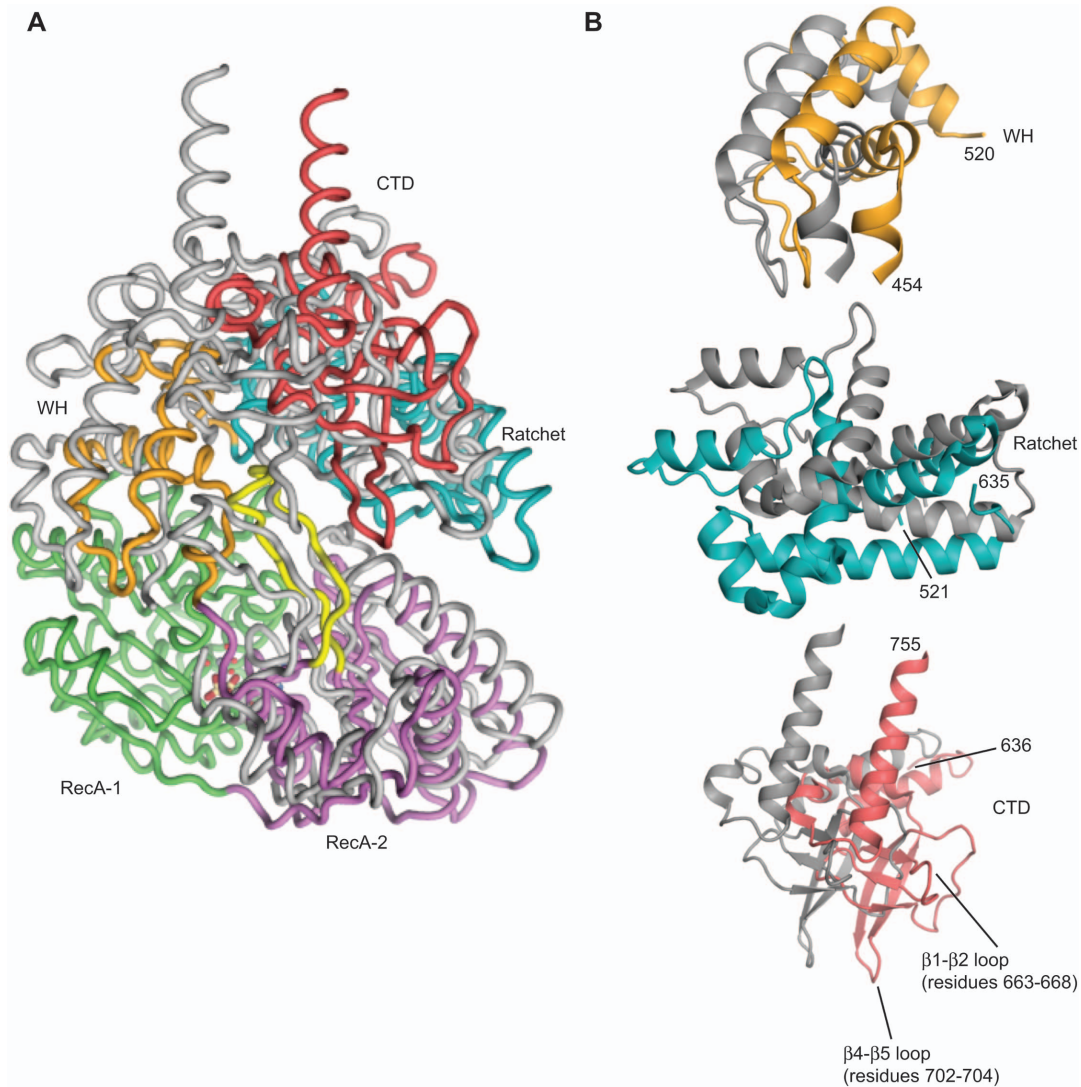
**Figure 4** The putative nucleic acid-binding pocket of Prp43p.

(A) Comparison of the binding pocket between the ADP state (left panel) and the modeled nucleic acid-bound state (right panel). Residues of the  $\beta 1$ – $\beta 2$  and  $\beta 4$ – $\beta 5$  loops in the CTD proposed to bind RNA are shown as sticks. The longer DNA strand from the Hel308 structure (Figure 3C) was docked by superimposition of the Prp43p RecA domains with the equivalent Hel308 domains. (B) The long helix in the ratchet domains of Hel308 (left panel) and Mtr4 (middle panel) interact with bound DNA and RNA. Conserved residues pointing towards the binding pocket are shown as sticks. The corresponding helix in Prp43p in the modeled nucleic acid bound state (gray) is shown in comparison with the Hel308 helix interacting with DNA (right panel). Conserved residues Arg611 and Gln622 of Prp43p are shown as sticks.

single-stranded 30 mer uracil oligo, whereas ADPNP or the absence of nucleotide allowed RNA-Prp43p interaction (11). With respect to functional states and their nucleotide dependence, Prp43p probably cycles through ATP-ADP-*apo*-ATP bound forms, where the ATP state contains RNA whereas both the ADP and the *apo* states might or might not be bound to RNA. An RNA-bound ADP state originating from a Prp43p-ATP-RNA state might exist transiently but after dissociation RNA most likely cannot rebind until ADP dissociates.

To further explore the structural changes necessary in order to accommodate a nucleic acid substrate in a DEAH/RHA helicase, the following modeling steps were carried

out. First, we superimposed the RecA-1 domains of Prp43p and eIF4AIII-RNA-ADPNP (RCSB entry 2HYI), and then matched RecA-2 of Prp43p to its counterpart in eIF4AIII-RNA-ADPNP. We then superimposed the RecA-1 domains of Prp43p and Hel308-DNA (RCSB entry 2P6R) and subsequently matched the WH and ratchet domains of Prp43p to their counterparts in Hel308-DNA. The CTD of Prp43p was allowed to move along with the WH and ratchet domains. In the resulting model, RecA-2 was rotated by approximately  $19^\circ$  compared to the Prp43p-ADP structure, whereas the WH, ratchet and CTD domains were rotated by approximately  $17^\circ$  (Figure 5). The tip of the 5'HP moves 9 Å and the tip of the  $\beta 4$ – $\beta 5$  loop in CTD moves 11 Å (Figures



**Figure 5** The conformational changes from the crystallized ADP state to the modeled nucleic acid-bound state. (A) Overview of domain movements between the ADP state (color) and the modeled nucleic acid-bound state (gray). The position of the RecA-1 is kept constant. The NTD is not shown. (B) Movements of the WH (top), Ratchet (middle) and CTD domains (bottom) between the ADP state (color) and the modeled nucleic acid-bound state (gray). The first and last residues in the domains are labeled.

4A and 5B). These modeled domain rearrangements would open up the 5' end of the binding pocket for a single-stranded nucleic acid. Likewise, the 3' loop connecting the WH and ratchet domains moves 5 Å, opening up the 3' end of the nucleic acid binding pocket (Figure 4A).

In the resulting model, the last helix in the ratchet domain (residues 610–630) would be aligned along the nucleic acid strand opposite the two RecA domains (Figure 4B, right panel). The equivalent helix in Hel308 was proposed to serve as a ratchet for unidirectional movement of ssDNA through the binding pocket (73). The conserved residues Arg592 and Trp599 in Hel308 stack with the third and the fifth unpaired bases (positions +3 and +5) of the partially unwound DNA duplex, respectively (Figure 4B, left panel). A similar situation is found in Mtr4, where residues Arg1026, Arg1030 and Glu1033, which are conserved in Ski2 and Mtr4 helicases in all eukaryotic species, also point towards the RNA

binding pocket (Figure 4B, middle panel). The side-chain of Mtr4 Arg1030 forms a hydrogen bond with the most 3' base of the 5 mer adenine RNA (equivalent to position +4 in Hel308). Clearly, our modeling attempt involves only inter-domain movements and cannot account for intra-domain conformational changes, which are likely to occur upon ATP and RNA binding. In particular, comparison of the resulting model with the structures of DNA-bound Hel308 and RNA-bound Mtr4 suggests that the equivalent long ratchet helix in Prp43p is likely to align more in parallel to the bound nucleic acid strand than what our present model predicts. This would orient the side-chains of Arg611 (either arginine or lysine in most DEAH/RHA helicases) and Gln622 (highly conserved) towards the RNA binding pocket (Figure 4B). These side chains might interact with bases via stacking or hydrogen bonding. Mutation in either the  $\beta$ 1- $\beta$ 2 or  $\beta$ 4- $\beta$ 5 loop of the OB-fold in CTD (Figures 4A and 5B) leads to

significant decrease of RNA-stimulated ATPase activity of Prp43p (12). It is therefore likely that regions within the CTD facilitate the entry of one nucleic acid strand to the binding pocket. The 5' HP is probably responsible for duplex unwinding, as is observed with the Hel308 hairpin (73) or nucleic acid-protein complex remodeling.

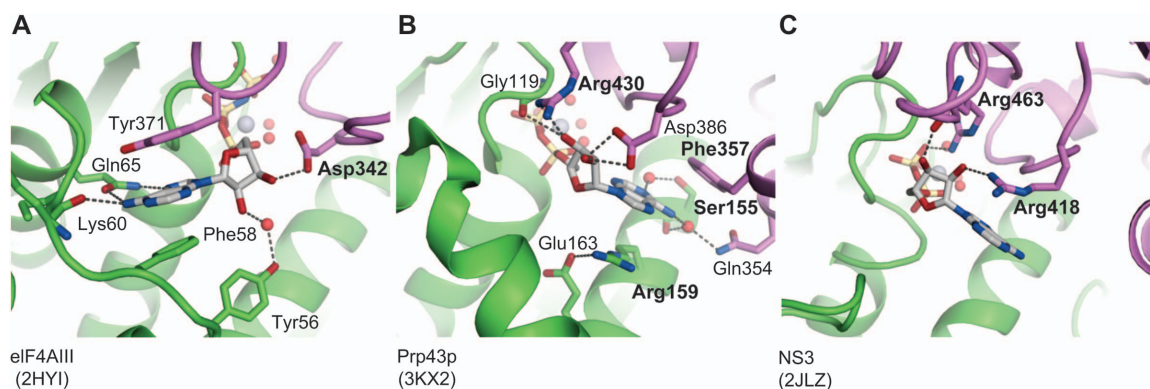
### Nucleotide binding and NTPase activity within DEAH/RHA helicases

Yeast Prp2p (76), Prp16p (20), Prp22p (77), Prp43p (78), mammalian RHA (9) and DHX29 (67) have all been shown to be able to hydrolyze various NTPs and dNTPs in a nucleic acid-stimulated fashion. NS3/NPH-II is the only other family in SF2 that is not ATP-specific. All other families of helicases in SF2 specifically recognize the adenine ring of ATP/ADP through the Q motif in the N-terminal region of RecA-1 (79), which is absent in DEAH/RHA and NS3/NPH-II proteins. As expected, the ADP phosphates are coordinated by interactions primarily with residues from motif I of Prp43p, similar to the structures of other nucleotide-bound SF2 proteins. The orientation of the ribose and the adenine base is, however, drastically different (Figure 6). In the ATP-specific DEAD-box protein eIF4AIII (Figure 6A), the 3' OH of the ribose forms a hydrogen bond with the side-chain carboxyl of Asp342 and the interaction between 2' OH and Tyr56 is mediated by a water molecule. The base is stacked between Phe58 and Tyr371. Recognition of the adenine ring involves hydrogen bonding of N-6 with both the backbone carbonyl of Lys60 and the side-chain amide of Gln65 from the Q motif. Another hydrogen bond is formed between N-7 of the adenine base and the Gln65 amide (80, 81). In Prp43p (Figure 6B), the 3' OH of the ribose forms hydrogen bonds with the backbone carbonyl of Gly119 from motif I and the side-chain of Arg430 from motif VI, whereas the 2' OH engages in a hydrogen bond with the side-chain carboxyl of Asp386. The base is stacked between Arg159 and Phe357, both high-

ly conserved in the DEAH/RHA family (Figure 2). Ser155 and Gln354 coordinate two water molecules that in turn form hydrogen bonds with the adenine but the base is not specifically recognized via any direct interactions with Prp43p. In the flavivirus NS3 helicase (Figure 6C), the side-chain of Arg418 and the backbone carbonyl of Arg463 form hydrogen bonds with 2' and 3' OH of the ribose respectively, whereas the adenine base is completely exposed to solvent (72). Therefore, DEAH/RHA helicases possess a novel nucleotide binding site different from all the other SF2 proteins.

### The putative interaction surface with G-patch proteins

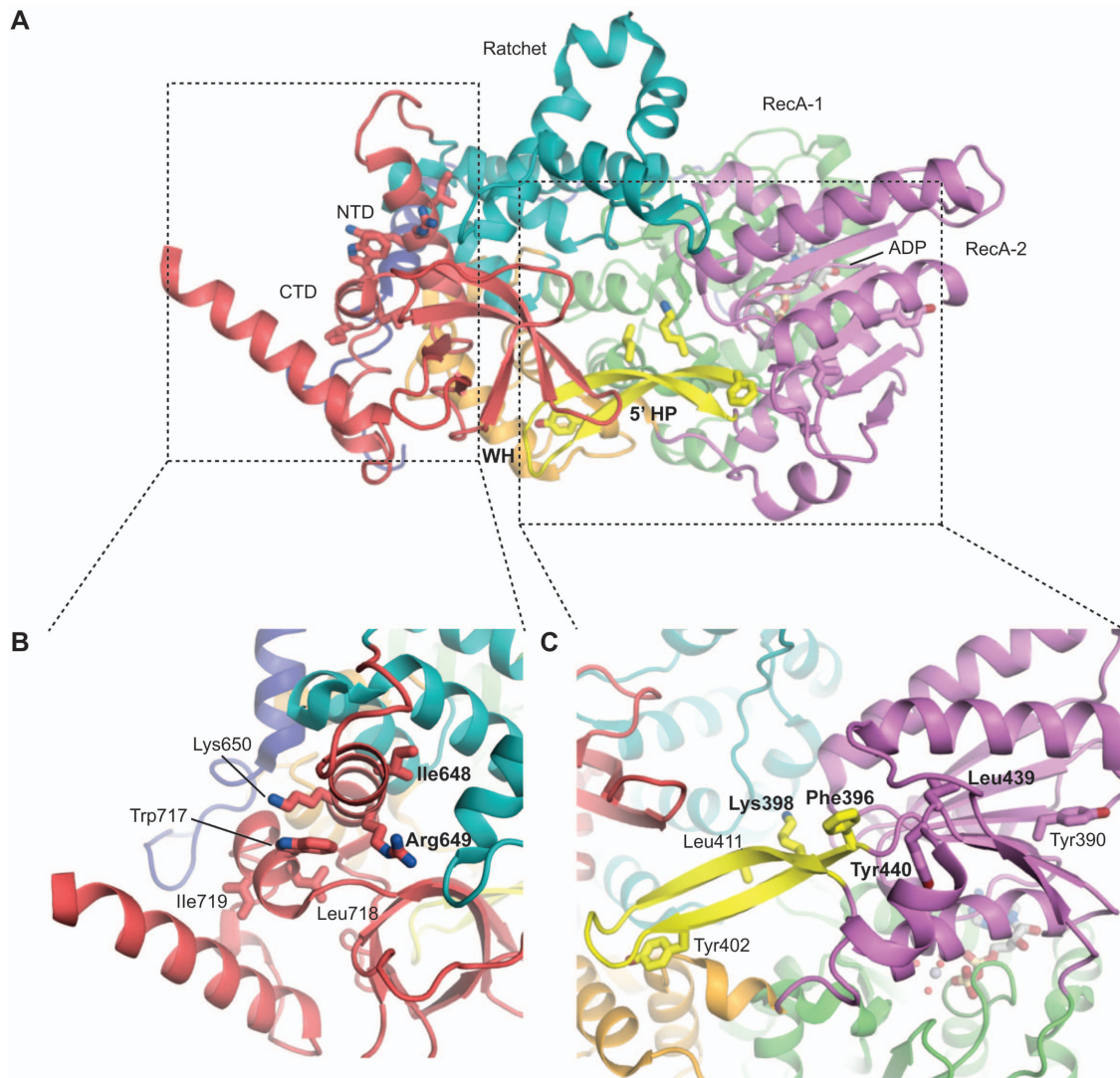
Interaction with various G-patch proteins is essential for the function of Prp2p and Prp43p. Tanaka and colleagues studied genetic interactions between Prp43p and Ntr1p in yeast and identified a number of Prp43p mutants (Y390A, F396A, K398A, Y402A, L411A, L439A-Y440A, I648A-R649A-K650A, W717A-L718A-I719A, and  $\Delta$ C35) that were synthetic lethal or displayed a slow growth phenotype in conjunction with the Ntr1-L68A mutant but showed little or no growth defects in conjunction with wild type Ntr1 at 30°C (43). Leu68 is located within the G-patch motif of Ntr1p. The N-terminal Ntr1p fragment of 120 amino acids suffices for binding to Prp43p and enhances its helicase activity *in vitro* and such interactions are diminished by the Ntr1-L68A mutation. This suggests that the Prp43p mutations mentioned above likely abolished the Prp43p-Ntr1p interaction that was already weakened by the Ntr1-L68A mutation. All of these mutated residues are located in either the RecA-2 domain, the 5' HP or the CTD (Figure 7). The full-length Prp43p can also bind to a C-terminal fragment of Pfa1p (574–767) (44) but this interaction is disrupted by deletion of the 110 residues in the C-terminus of Prp43p, which constitute all but



**Figure 6** The nucleotide binding sites of Superfamily 2 helicases.

(A) ADPNP bound to the DEAD-box protein eIF4AIII (RCSB entry 2HY1). The base is stacked between Phe58 and Tyr371. Lys60 and Gln65 from the Q motif are responsible for direct recognition of adenine. (B) ADP bound to the DEAH/RHA protein Prp43p. The base is stacked between Arg159 and Phe357. Interaction between the adenine ring and the side chains of Ser155 and Gln354 are mediated by two water molecules. (C) ADP bound to NS3/NPH II protein NS3 (RCSB entry 2JLZ). The base is completely exposed to the solvent.





**Figure 7** The putative interaction surface of Prp43p with G-patch proteins.

(A) The RecA-2-5' HP-CTD side of Prp43p might be the primary interaction surface. Mutations of several residues within these domains likely disrupt binding to G-patch proteins. These residues are shown as sticks and highlighted in close-up in (B) and (C).

the first helix of the CTD. Moreover, mutations of Trp854 and Leu855 to alanine in Prp2p, which are equivalent to Prp43 W717A-L718A also abolished the binding between Prp2p and Spp2p *in vitro* (42). These results suggest that the RecA-2-5' HP-CTD side, where the nucleic acid entry path lies, is the primary interaction surface of DEAH/RHA helicases with regulatory G-patch proteins. Phe396, Lys398, Tyr402 and Leu411 are surface exposed residues along the 5' HP that could be directly involved in interaction with G-patch proteins. Tyr390 is a bulky residue in the interior of RecA-2 immediately upstream of the 5' HP, whereas Leu390 and Tyr440 are bulky residues in close vicinity of the 5' HP. Mutation of these three residues might compromise the structural integrity of RecA-2 or the 5' HP. Ile648, Arg649, Lys650, Trp717, Leu718 and Ile719 are bulky residues in the interior of CTD. Mutation at these points could destabilize the structure of CTD.

## Outlook

In the past decade, significant progress has been made in understanding the various functions of DEAH/RHA helicases in several highly complicated biological processes. A structure of a DEAH/RHA helicase in complex with its nucleic acid substrate and a non-hydrolysable ATP analog will be a critical step toward elucidating the molecular mechanism of this helicase family. An equally important endeavour is the structural and biochemical characterization of DEAH/RHA helicases in complex with their well-defined interaction partners, which will help illuminate at the atomic level how their functions are regulated. Most of these helicases are bigger than Prp43p and expected to contain variable N- and/or C-terminal extensions beyond the conserved five-domain structural core. The structure of the RHA N-terminal minimal transactivation domain was recently

revealed (82). This domain mediates the interaction between RHA and RNA polymerase II (48). The functional importance of the N-terminal RSM domain of RHAU has been well documented, whereas Prp16p is recruited to the spliceosome via its non-conserved N-terminal domain. These examples illustrate the importance of extending the structural framework provided by the Prp43p structures to regions outside the conserved regions within the DEAH/RHA family in order to obtain an in-depth understanding of structure-function features of these helicases.

## Acknowledgments

We would like to thank Chrysa Latrick for suggestions and critical reading of the manuscript. This work was supported by the Alfred Benson Foundation (KHN), The Danish Council for Independent Research | Natural Sciences (GRA), a Hallas-Møller stipend from the Novo Nordic foundation (GRA) and the Danish National Research Foundation (GRA).

## References

- Lohman TM, Tomko EJ, Wu CG. Non-hexameric DNA helicases and translocases: mechanisms and regulation. *Nat Rev Mol Cell Biol* 2008; 9: 391–401.
- Pyle AM. Translocation and unwinding mechanisms of RNA and DNA helicases. *Annu Rev Biophys* 2008; 37: 317–36.
- Singleton MR, Dillingham MS, Wigley DB. Structure and mechanism of helicases and nucleic acid translocases. *Annu Rev Biochem* 2007; 76: 23–50.
- Jankowsky E, Fairman ME. RNA helicases – one fold for many functions. *Curr Opin Struct Biol* 2007; 17: 316–24.
- Fairman-Williams ME, Guenther UP, Jankowsky E. SF1 and SF2 helicases: family matters. *Curr Opin Struct Biol* 2010; 20: 313–24.
- Chen JH, Lin RJ. The yeast PRP2 protein, a putative RNA-dependent ATPase, shares extensive sequence homology with two other pre-mRNA splicing factors. *Nucleic Acids Res* 1990; 18: 6447.
- Schwer B, Guthrie C. PRP16 is an RNA-dependent ATPase that interacts transiently with the spliceosome. *Nature* 1991; 349: 494–9.
- Company M, Arenas J, Abelson J. Requirement of the RNA helicase-like protein PRP22 for release of messenger RNA from spliceosomes. *Nature* 1991; 349: 487–93.
- Lee CG, Hurwitz J. A new RNA helicase isolated from HeLa cells that catalytically translocates in the 3' to 5' direction. *J Biol Chem* 1992; 267: 4398–407.
- Lee CG, Hurwitz J. Human RNA helicase A is homologous to the maleless protein of *Drosophila*. *J Biol Chem* 1993; 268: 16822–30.
- He Y, Andersen GR, Nielsen KH. Structural basis for the function of DEAH helicases. *EMBO Rep* 2010; 11: 180–6.
- Wallbott H, Mouffok S, Capeyrou R, Lebaron S, Humbert O, van Tilbeurgh H, Henry Y, Leulliot N. Prp43p contains a processive helicase structural architecture with a specific regulatory domain. *EMBO J* 2010; 29: 2194–204.
- Arenas JE, Abelson JN. Prp43: an RNA helicase-like factor involved in spliceosome disassembly. *Proc Natl Acad Sci USA* 1997; 94: 11798–802.
- Lebaron S, Froment C, Fromont-Racine M, Rain JC, Monsarrat B, Caizergues-Ferrer M, Henry Y. The splicing ATPase prp43p is a component of multiple preribosomal particles. *Mol Cell Biol* 2005; 25: 9269–82.
- Combs DJ, Nagel RJ, Ares M Jr, Stevens SW. Prp43p is a DEAH-box spliceosome disassembly factor essential for ribosome biogenesis. *Mol Cell Biol* 2006; 26: 523–34.
- Leeds NB, Small EC, Hiley SL, Hughes TR, Staley JP. The splicing factor Prp43p, a DEAH box ATPase, functions in ribosome biogenesis. *Mol Cell Biol* 2006; 26: 513–22.
- Wahl MC, Will CL, Luhrmann R. The spliceosome: design principles of a dynamic RNP machine. *Cell* 2009; 136: 701–18.
- Kim SH, Lin RJ. Spliceosome activation by PRP2 ATPase prior to the first transesterification reaction of pre-mRNA splicing. *Mol Cell Biol* 1996; 16: 6810–9.
- Lardelli RM, Thompson JX, Yates JR 3rd, Stevens SW. Release of SF3 from the intron branchpoint activates the first step of pre-mRNA splicing. *RNA* 2010; 16: 516–28.
- Schwer B, Guthrie C. A conformational rearrangement in the spliceosome is dependent on PRP16 and ATP hydrolysis. *EMBO J* 1992; 11: 5033–9.
- Tseng CK, Liu HL, Cheng SC. DEAH-box ATPase Prp16 has dual roles in remodeling of the spliceosome in catalytic steps. *RNA* 2011; 17: 145–54.
- Mefford MA, Staley JP. Evidence that U2/U6 helix I promotes both catalytic steps of pre-mRNA splicing and rearranges in between these steps. *RNA* 2009; 15: 1386–97.
- Schwer B, Gross CH. Prp22, a DEXH-box RNA helicase, plays two distinct roles in yeast pre-mRNA splicing. *EMBO J* 1998; 17: 2086–94.
- Wagner JD, Jankowsky E, Company M, Pyle AM, Abelson JN. The DEAH-box protein PRP22 is an ATPase that mediates ATP-dependent mRNA release from the spliceosome and unwinds RNA duplexes. *EMBO J* 1998; 17: 2926–37.
- Schwer B. A conformational rearrangement in the spliceosome sets the stage for Prp22-dependent mRNA release. *Mol Cell* 2008; 30: 743–54.
- McPheeters DS, Muhlenkamp P. Spatial organization of protein-RNA interactions in the branch site-3' splice site region during pre-mRNA splicing in yeast. *Mol Cell Biol* 2003; 23: 4174–86.
- Martin A, Schneider S, Schwer B. Prp43 is an essential RNA-dependent ATPase required for release of lariat-intron from the spliceosome. *J Biol Chem* 2002; 277: 17743–50.
- Pandit S, Lynn B, Rymond BC. Inhibition of a spliceosome turnover pathway suppresses splicing defects. *Proc Natl Acad Sci USA* 2006; 103: 13700–5.
- Bohnsack MT, Martin R, Granneman S, Ruprecht M, Schleiff E, Tollervey D. Prp43 bound at different sites on the pre-rRNA performs distinct functions in ribosome synthesis. *Mol Cell* 2009; 36: 583–92.
- Burgess SM, Guthrie C. A mechanism to enhance mRNA splicing fidelity: the RNA-dependent ATPase Prp16 governs usage of a discard pathway for aberrant lariat intermediates. *Cell* 1993; 73: 1377–91.
- Villa T, Guthrie C. The Isy1p component of the NineTeen complex interacts with the ATPase Prp16p to regulate the fidelity of pre-mRNA splicing. *Genes Dev* 2005; 19: 1894–904.
- Koodathingal P, Novak T, Piccirilli JA, Staley JP. The DEAH box ATPases Prp16 and Prp43 cooperate to proofread 5' splice site cleavage during pre-mRNA splicing. *Mol Cell* 2010; 39: 385–95.
- Query CC, Konarska MM. Suppression of multiple substrate mutations by spliceosomal prp8 alleles suggests functional cor-

- relations with ribosomal ambiguity mutants. *Mol Cell* 2004; 14: 343–54.
34. Mayas RM, Maita H, Staley JP. Exon ligation is proofread by the DEXD/H-box ATPase Prp22p. *Nat Struct Mol Biol* 2006; 13: 482–90.
  35. Mayas RM, Maita H, Semlow DR, Staley JP. Spliceosome discards intermediates via the DEAH box ATPase Prp43p. *Proc Natl Acad Sci USA* 2010; 107: 10020–5.
  36. Henras AK, Soudet J, Gerus M, Lebaron S, Caizergues-Ferrer M, Mouglin A, Henry Y. The post-transcriptional steps of eukaryotic ribosome biogenesis. *Cell Mol Life Sci* 2008; 65: 2334–59.
  37. Pertschy B, Schneider C, Gnadig M, Schafer T, Tollervey D, Hurt E. RNA helicase Prp43 and its co-factor Pfa1 promote 20 to 18 S rRNA processing catalyzed by the endonuclease Nob1. *J Biol Chem* 2009; 284: 35079–91.
  38. Colley A, Beggs JD, Tollervey D, Lafontaine DL. Dhr1p, a putative DEAH-box RNA helicase, is associated with the box C+D snoRNP U3. *Mol Cell Biol* 2000; 20: 7238–46.
  39. Beltrame M, Tollervey D. Base pairing between U3 and the pre-ribosomal RNA is required for 18S rRNA synthesis. *EMBO J* 1995; 14: 4350–6.
  40. Sharma K, Tollervey D. Base pairing between U3 small nucleolar RNA and the 5' end of 18S rRNA is required for pre-rRNA processing. *Mol Cell Biol* 1999; 19: 6012–9.
  41. Aravind L, Koonin EV. G-patch: a new conserved domain in eukaryotic RNA-processing proteins and type D retroviral polyproteins. *Trends Biochem Sci* 1999; 24: 342–4.
  42. Silverman EJ, Maeda A, Wei J, Smith P, Beggs JD, Lin RJ. Interaction between a G-patch protein and a spliceosomal DEXD/H-box ATPase that is critical for splicing. *Mol Cell Biol* 2004; 24: 10101–10.
  43. Tanaka N, Aronova A, Schwer B. Ntr1 activates the Prp43 helicase to trigger release of lariat-intron from the spliceosome. *Genes Dev* 2007; 21: 2312–25.
  44. Lebaron S, Papin C, Capeyrou R, Chen YL, Froment C, Monsarrat B, Caizergues-Ferrer M, Grigoriev M, Henry Y. The ATPase and helicase activities of Prp43p are stimulated by the G-patch protein Pfa1p during yeast ribosome biogenesis. *EMBO J* 2009; 28: 3808–19.
  45. Guglielmi B, Werner M. The yeast homolog of human PinX1 is involved in rRNA and small nucleolar RNA maturation, not in telomere elongation inhibition. *J Biol Chem* 2002; 277: 35712–9.
  46. Friedemann J, Grosse F, Zhang S. Nuclear DNA helicase II (RNA helicase A) interacts with Werner syndrome helicase and stimulates its exonuclease activity. *J Biol Chem* 2005; 280: 31303–13.
  47. Anderson SF, Schlegel BP, Nakajima T, Wolpin ES, Parvin JD. BRCA1 protein is linked to the RNA polymerase II holoenzyme complex via RNA helicase A. *Nat Genet* 1998; 19: 254–6.
  48. Nakajima T, Uchida C, Anderson SF, Lee CG, Hurwitz J, Parvin JD, Montminy M. RNA helicase A mediates association of CBP with RNA polymerase II. *Cell* 1997; 90: 1107–12.
  49. Robb GB, Rana TM. RNA helicase A interacts with RISC in human cells and functions in RISC loading. *Mol Cell* 2007; 26: 523–37.
  50. Mischo HE, Hemmerich P, Grosse F, Zhang S. Actinomycin D induces histone  $\gamma$ -H2AX foci and complex formation of  $\gamma$ -H2AX with Ku70 and nuclear DNA helicase II. *J Biol Chem* 2005; 280: 9586–94.
  51. Jong JE, Park J, Kim S, Seo T. Kaposi's sarcoma-associated herpesvirus viral protein kinase interacts with RNA helicase A and regulates host gene expression. *J Microbiol* 2010; 48: 206–12.
  52. Kuroda MI, Kernan MJ, Kreber R, Ganetzky B, Baker BS. The maleless protein associates with the X chromosome to regulate dosage compensation in *Drosophila*. *Cell* 1991; 66: 935–47.
  53. Aratani S, Kageyama Y, Nakamura A, Fujita H, Fujii R, Nishioka K, Nakajima T. MLE activates transcription via the minimal transactivation domain in *Drosophila*. *Int J Mol Med* 2008; 21: 469–76.
  54. Jain A, Bacolla A, Chakraborty P, Grosse F, Vasquez KM. Human DHX9 helicase unwinds triple-helical DNA structures. *Biochemistry* 2010; 49: 6992–9.
  55. Hartman TR, Qian S, Bolinger C, Fernandez S, Schoenberg DR, Boris-Lawrie K. RNA helicase A is necessary for translation of selected messenger RNAs. *Nat Struct Mol Biol* 2006; 13: 509–16.
  56. Tang H, Wong-Staal F. Specific interaction between RNA helicase A and Tap, two cellular proteins that bind to the constitutive transport element of type D retrovirus. *J Biol Chem* 2000; 275: 32694–700.
  57. Reddy TR, Tang H, Xu W, Wong-Staal F. Sam68, RNA helicase A and Tap cooperate in the post-transcriptional regulation of human immunodeficiency virus and type D retroviral mRNA. *Oncogene* 2000; 19: 3570–5.
  58. Westberg C, Yang JP, Tang H, Reddy TR, Wong-Staal F. A novel shuttle protein binds to RNA helicase A and activates the retroviral constitutive transport element. *J Biol Chem* 2000; 275: 21396–401.
  59. Bolinger C, Sharma A, Singh D, Yu L, Boris-Lawrie K. RNA helicase A modulates translation of HIV-1 and infectivity of progeny virions. *Nucleic Acids Res* 2010; 38: 1686–96.
  60. Kim T, Pazhoor S, Bao M, Zhang Z, Hanabuchi S, Facchinetti V, Bover L, Plumas J, Chaperot L, Qin J, Liu YJ. Aspartate-glutamate-alanine-histidine box motif (DEAH)/RNA helicase A helicases sense microbial DNA in human plasmacytoid dendritic cells. *Proc Natl Acad Sci USA* 2010; 107: 15181–6.
  61. Tran H, Schilling M, Wirbelauer C, Hess D, Nagamine Y. Facilitation of mRNA deadenylation and decay by the exosome-bound, DEXH protein RHAU. *Mol Cell* 2004; 13: 101–11.
  62. Iwamoto F, Stadler M, Chalupnikova K, Oakeley E, Nagamine Y. Transcription-dependent nucleolar cap localization and possible nuclear function of DEXH RNA helicase RHAU. *Exp Cell Res* 2008; 314: 1378–91.
  63. Creacy SD, Routh ED, Iwamoto F, Nagamine Y, Akman SA, Vaughn JP. G4 resolvase 1 binds both DNA and RNA tetramolecular quadruplex with high affinity and is the major source of tetramolecular quadruplex G4-DNA and G4-RNA resolving activity in HeLa cell lysates. *J Biol Chem* 2008; 283: 34626–34.
  64. Lattmann S, Giri B, Vaughn JP, Akman SA, Nagamine Y. Role of the amino terminal RHAU-specific motif in the recognition and resolution of guanine quadruplex-RNA by the DEAH-box RNA helicase RHAU. *Nucleic Acids Res* 2010; 38: 6219–33.
  65. Chalupnikova K, Lattmann S, Selak N, Iwamoto F, Fujiki Y, Nagamine Y. Recruitment of the RNA helicase RHAU to stress granules via a unique RNA-binding domain. *J Biol Chem* 2008; 283: 35186–98.
  66. Abaeva IS, Marintchev A, Pisareva VP, Hellen CU, Pestova TV. Bypassing of stems versus linear base-by-base inspection of mammalian mRNAs during ribosomal scanning. *EMBO J* 2011; 30: 115–29.
  67. Pisareva VP, Pisarev AV, Komar AA, Hellen CU, Pestova TV. Translation initiation on mammalian mRNAs with structured



- 5'UTRs requires DEXH-box protein DHX29. *Cell* 2008; 135: 1237–50.
68. Parsyan A, Shahbazian D, Martineau Y, Petroulakis E, Alain T, Larsson O, Mathonnet G, Tettweiler G, Hellen CU, Pestova TV, Svitkin YV, Sonenberg N. The helicase protein DHX29 promotes translation initiation, cell proliferation, and tumorigenesis. *Proc Natl Acad Sci USA* 2009; 106: 22217–22.
69. Koo JT, Choe J, Moseley SL. HrpA, a DEAH-box RNA helicase, is involved in mRNA processing of a fimbrial operon in *Escherichia coli*. *Mol Microbiol* 2004; 52: 1813–26.
70. Walker JE, Saraste M, Runswick MJ, Gay NJ. Distantly related sequences in the  $\alpha$ - and  $\beta$ -subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J* 1982; 1: 945–51.
71. Gorbalenya AE, Koonin EV, Donchenko AP, Blinov VM. A novel superfamily of nucleoside triphosphate-binding motif containing proteins which are probably involved in duplex unwinding in DNA and RNA replication and recombination. *FEBS Lett* 1988; 235: 16–24.
72. 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. Insights into RNA unwinding and ATP hydrolysis by the flavivirus NS3 protein. *EMBO J* 2008; 27: 3209–19.
73. Buttner K, Nehring S, Hopfner KP. Structural basis for DNA duplex separation by a superfamily-2 helicase. *Nat Struct Mol Biol* 2007; 14: 647–52.
74. Jackson RN, Klauer AA, Hintze BJ, Robinson H, van Hoof A, Johnson SJ. The crystal structure of Mtr4 reveals a novel arch domain required for rRNA processing. *EMBO J* 2010; 29: 2205–16.
75. Weir JR, Bonneau F, Hentschel J, Conti E. Structural analysis reveals the characteristic features of Mtr4, a DEXH helicase involved in nuclear RNA processing and surveillance. *Proc Natl Acad Sci USA* 2010; 107: 12139–44.
76. Kim SH, Smith J, Claude A, Lin RJ. The purified yeast pre-mRNA splicing factor PRP2 is an RNA-dependent NTPase. *EMBO J* 1992; 11: 2319–26.
77. Tanaka N, Schwer B. Characterization of the NTPase, RNA-binding, and RNA helicase activities of the DEAH-box splicing factor Prp22. *Biochemistry* 2005; 44: 9795–803.
78. Tanaka N, Schwer B. Mutations in PRP43 that uncouple RNA-dependent NTPase activity and pre-mRNA splicing function. *Biochemistry* 2006; 45: 6510–21.
79. Tanner NK, Cordin O, Banroques J, Doere M, Linder P. The Q motif: a newly identified motif in DEAD box helicases may regulate ATP binding and hydrolysis. *Mol Cell* 2003; 11: 127–38.
80. Andersen CB, Ballut L, Johansen JS, Chamieh H, Nielsen KH, Oliveira CL, Pedersen JS, Seraphin B, Le Hir H, Andersen GR. Structure of the exon junction core complex with a trapped DEAD-box ATPase bound to RNA. *Science* 2006; 313: 1968–72.
81. Bono F, Ebert J, Lorentzen E, Conti E. The crystal structure of the exon junction complex reveals how it maintains a stable grip on mRNA. *Cell* 2006; 126: 713–25.
82. Schutz P, Wahlberg E, Karlberg T, Hammarstrom M, Collins R, Flores A, Schuler H. Crystal structure of human RNA helicase A (DHX9): structural basis for unselective nucleotide base binding in a DEAD-box variant protein. *J Mol Biol* 2010; 400: 768–82.

Received April 7, 2011; accepted May 24, 2011