Functional and Structural Characterization of the DEAH-box Helicase Prp43

Dissertation

for the award of the degree

"Doctor rerurm naturalium"

of the Georg-August-Universität Göttingen

within the doctoral program

"Biomolecules: Structure-Function-Dynamics"

of the Georg-August University School of Science (GAUSS)

Submitted by

Marieke Enders

from Wittenberg

Göttingen, 2022

Members of the thesis advisory committee

Prof. Dr. Ralf Ficner (Reviewer 1)
Georg-August-University Göttingen
Department of Molecular Structural Biology

Dr. Sarah Adio (Reviewer 2)
Georg-August-University Göttingen
Department of Molecular Structural Biology

Prof. Dr. Reinhard Lührmann
Max-Planck-Institute for Multidisciplinary Sciences
Department of Cellular Biochemistry

Members of the examination board

Prof. Dr. Markus Bohnsack
University Medical Center Göttingen
Department of Molecular Biology

Prof. Dr. Kai Tittmann
Georg-August-University Göttingen
Department of Molecular Enzymology

Prof. Dr. Henning Urlaub
Max-Planck-Institute for Multidisciplinary Sciences
Department of Bioanalytical Mass Spectrometry

Date of oral examination: 12th July, 2022

Table of contents

Chapter 1: Introduction 1
1.1. RNA helicase families 1
1.2. RNA helicases in splicing 3
1.3. RNA helicases in ribosome biogenesis 6
1.4. Translocation and unwinding mechanisms of DEAD- and DEAH-box helicases 8
DEAD-box helicases 8
DEAH-box helicases 10
1.5. Regulation of RNA helicases 13
Cellular functions of Prp43 and their regulation by G-patch proteins 14
Binding mode of the G-patch motif 16
A model for G-patch induced stimulation 17
1.6. Observing helicase domain motility in real-time 18
1.7. Scope of this thesis 21
Chapter 2: Regulation of Prp43 by the G-patch factor Pfa1 22
2.1. Abstract 23
2.2. Significance 23
2.3. Introduction 24
2.4. Results 26
Pfa1(gp) stabilizes the open conformation of Prp43 RecA domains 26
Pfa1(gp) promotes ADP release by Prp43 27
Pfa1(gp) enhances RNA binding of Prp43 while ATP binding generates force and induces partial unwinding 28
Pfa1(gp) accelerates transitions between closed and open RecA domain conformation during ATP turnover, facilitating processive translocation 30
Motility cycle of the Prp43-Pfa1(gp) complex 31
2.5. Discussion 32
2.6. Material and methods 34
2.7. Acknowledgements and funding sources 34
2.8. References 35
2.9. Figures and legends 40
2.10. Supplementary information 49
Chapter 3: Conformational dynamics of the RNA binding channel 67
3.1. Abstract 68
3.2. Introduction 68
3.3. Results 70
smFRET label positions on Prp43 to monitor the structural dynamics of the RNA binding channel 70
ATPase activity, RNA binding and RNA unwinding by $Prp43_{Cys}$ 70
Experiment scheme to monitor structural changes of the RNA binding channel in Prp43 72
The Prp43 RNA binding channel shows a dynamic equilibrium between open and closed conformations 72

The interplay of ATP and Pfa1(gp) leads to efficient RNA loading and stable binding ... 74
3.4. Discussion ... 75
RNA loading mechanism of Prp43 ... 75
Comparison to other DEAH-box helicases ... 76
Role of the G-patch ... 76
Comparison to DEAD-box helicases ... 77
Comparison to other helicase families ... 77
3.5. Material and methods ... 77
Protein expression and purification ... 77
Fluorescence-labeling of ctPrp43 ... 77
Determination of dye to protein ratio ... 78
Sample preparation for TIRF microscopy ... 78
TIRF microscopy ... 78
Data analysis ... 79
ATPase activity assay ... 79
RNA binding assay ... 79
RNA unwinding assay ... 79
3.6. Acknowledgements and funding sources ... 80
3.7. References ... 80
3.8. Supplementary information ... 83
Chapter 4: Discussion and future perspectives ... 89
4.1. A model for Prp43 stimulation by G-patch proteins ... 89
4.2. Helicase processivity and implications for Prp43 processivity ... 90
4.3. Double strand separation by Prp43 ... 94
4.4. Conformational regulation of helicases ... 97
4.5. Combining structural data and smFRET to characterize helicase domain motility ... 99
Chapter 5: Abstract ... 101
Bibliography ... 102
Abbreviations ... 117
Acknowledgements ... 120

Chapter 1: Introduction

Nucleic acids are key macromolecules in all domains of life. While deoxyribonucleic acid (DNA) carries the genetic instructions that are necessary to build a living organism, ribonucleic acid (RNA) is essential to convert this information into the amino acid sequence of proteins. However, the role of RNAs is not limited to linking the genetic information to the final protein product as messenger (mRNA) and transfer (tRNA) RNAs. While some RNAs are essential regulators of gene expression, others have been found to act like enzymes and display catalytic activity (Moazed, 2009; Smith *et al*, 2010; Arnold *et al*, 2020). Either alone as so-called ribozymes or supported by proteins in ribonucleoprotein particles (RNPs) they catalyze important reactions in various cellular processes. To function correctly, it is crucial for RNA molecules to fold into certain structures and many catalytic RNAs undergo substantial conformational rearrangements during their functional cycles (Gilman *et al*, 2017). The required structural transitions are often guided by RNA helicases that are consequently involved in all aspects of RNA metabolism (Jarmoskaite & Russell, 2014). This diverse class of enzymes is thought to make up 1% of all genes in an eukaryotic genome and was originally characterized as NTP-dependent unwindases, disrupting the base-pairing of RNA duplexes (Fairman-Williams *et al*, 2010; Wu, 2012). However, the functions of RNA helicases are not limited to classical unwinding. They are also known to displace proteins from RNA strands, mediate protein-protein interactions, stabilize folding intermediates or even promote annealing of RNA strands (Pyle, 2011; Bleichert & Baserga, 2007).

1.1 RNA helicase families

Generally, helicases can be classified into six different **s**uper**f**amilies (SF1-6) based on conserved sequence motifs and structural as well as mechanistic characteristics (Singleton *et al*, 2007; Leitão *et al*, 2015). The majority of SF1 and SF2 helicases is monomeric and their helicase cores are composed of two RecA-like domains, named after the homologous *E. coli* DNA repair protein RecA (Singleton *et al*, 2007). In contrast to this, SF3 to SF6 helicases assemble as toroidal hexamers built from six RecA-like or AAA^+ domains (ATPases associated with diverse cellular activities) (Iyer *et al*, 2004; Singleton *et al*, 2007). Similar to the ubiquitous functions of RNA helicases, many aspects of DNA metabolism rely on DNA helicases which are distributed among all six superfamilies. In contrast, most RNA helicases belong to the SF2 superfamily, where they split into five different sub-families: DEAD-box, DEAH/RHA, RIG-I-like, Ski2-like and the exclusively viral NS3/NPH-II helicases (Leitão *et al*, 2015). SF2 helicases share eight to nine conserved sequence motifs that are located in the helicase core and are thought

to fulfill similar functions in different proteins (Figure 1.1a). Motifs I, II, V and VI show the highest level of conservation and are necessary for nucleotide binding and hydrolysis (Walker *et al*, 1982; Pause *et al*, 1993; Fairman-Williams *et al*, 2010). The eponymous amino acid sequence for DEAH (Asp-Glu-Ala-His) and DEAD-box (Asp-Glu-Ala-Asp) helicases is contained in motif II (Pyle, 2008). While motifs Ia, Ib and IV are involved in the interaction with the RNA substrate (Caruthers & McKay, 2002; Cordin *et al*, 2006; Fairman-Williams *et al*, 2010), motif III is responsible for coupling ATP hydrolysis and unwinding (Pause & Sonenberg, 1992; Schwer & Meszaros, 2000). The Q motif, which coordinates the adenine base and thus mediates specificity for ATP, is less conserved as it is missing in DEAH/RHA and NS3 helicases (Tanner, 2003). Although these conserved motifs are spread over both RecA-like domains, they come into close proximity to form the binding interfaces for NTP and RNA which are located in between and on top of the domains, respectively (Figure 1.1b).

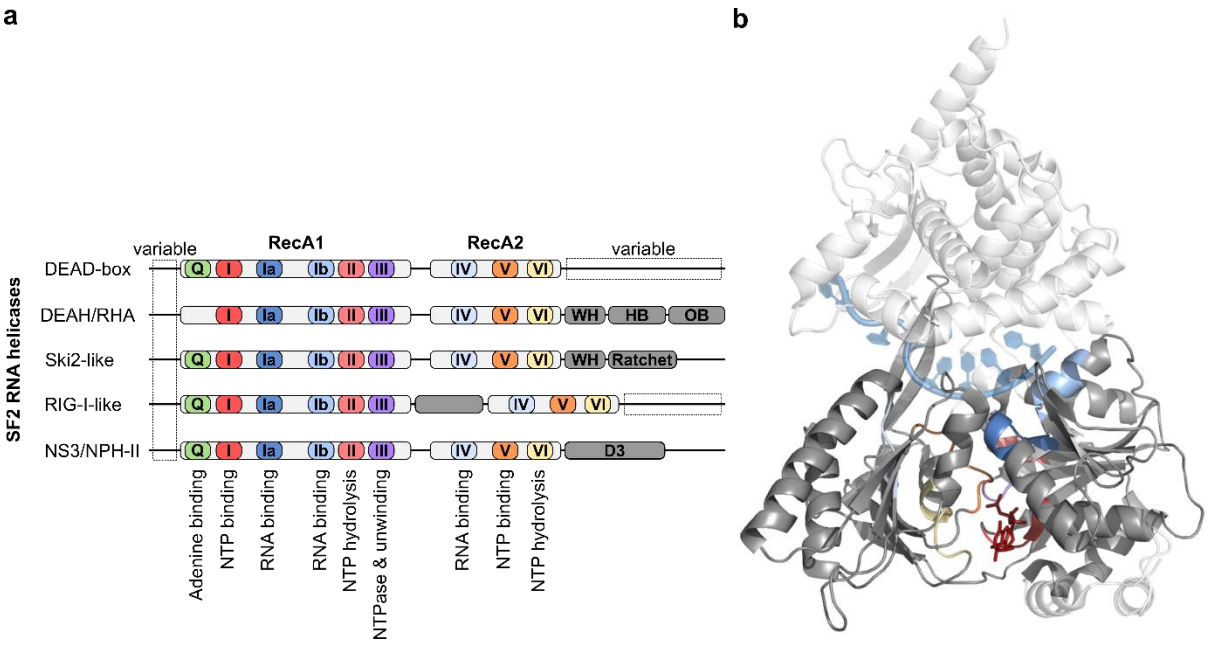


Figure 1.1: Conserved sequence motifs and domains of SF2 RNA helicases. (a) RNA helicases belonging to the SF2 share eight to nine conserved sequence motifs spread over both RecA-like domains which are involved in ATP binding and hydrolysis, RNA binding and coupling between ATP hydrolysis and RNA unwinding. Some subfamilies share conserved C-terminal domains. **(b)** Crystal structure of the DEAH-box helicase Prp43 in complex with U_7-RNA and ADP-BeF_3^- (PDB: 5lta) in cartoon representation. The helicase core is shown in grey with the conserved motifs colored as in (a) and the auxiliary domains are depicted in white. The motifs come together to form binding sites for the nucleotide and ssRNA depicted as red sticks and blue cartoon, respectively.

Flanking the conserved core, most SF2 helicases exhibit auxiliary N- and C-terminal domains which often adopt defined folds with specific functions and are responsible for physiological specificity of the enzymes (Fairman-Williams *et al*, 2010). Consistent with this function, the accessory domains are usually not conserved within a helicase family. However, structural conservation of the C-terminal domains has been observed for the Ski2-like and DEAH/RHA families (He *et al*, 2010; Büttner *et al*, 2007; Fairman-Williams *et al*, 2010). DEAH/RHA helicases share a C-terminal architecture composed of a winged-helix (WH), helix-bundle (HB) and oligonucleotide-binding (OB) domain. Ski2-like helicases also exhibit a conserved WH domain which is always succeeded by a ratchet domain, while other C-terminal domains such as helix-loop-helix or arch domains are only shared by some (Cordin *et al*, 2012; Johnson & Jackson, 2013). RIG-I-like helicases share a prominent insertion between the conserved RecA domains, while an α-helical C-terminal domain is common in NS3/NPH-II.

While RNA helicases fulfill essential functions in all aspects of RNA metabolism, the majority are involved in ribosome biogenesis or pre-mRNA splicing (Bleichert & Baserga, 2007). Both processes require a plethora of protein and RNA components that work in tandem to stepwise assemble huge dynamic RNP machineries on primary transcripts. In both processes, RNA helicases dynamically remodel RNA secondary structures and facilitate binding as well as dissociation of RNPs. In the following sections the impact of helicase action on both major pathways will be described in greater detail.

1.2 RNA helicases in splicing

The spliceosome is a multimegadalton molecular machine responsible for the removal of most non-coding intron sequences from eukaryotic pre-mRNAs (Will & Lührmann, 2011; Hoskins & Moore, 2012). It is composed of five U-snRNPs (uridine-rich small nuclear ribonucleic particles) and numerous non-snRNP proteins (Schmidt *et al*, 2014). Introns are recognizable by short sequences at the 5’ splice site (5’SS), the branch site (BS) and the 3’ splice site (3’SS), which are almost invariant in budding yeast but more varied in higher metazoans (Cordin & Beggs, 2013). The spliceosomal excision of an intron involves two subsequent transesterification reactions. In the first reaction which is also called branching, the 2’hydroxyl group of a conserved adenosine at the BS is aided by the surrounding snRNPs to nucleophilically attack the 5’SS, leading to cleavage of its phosphodiester bond. This reaction yields two intermediates, the 5’-exon and the intron-3’-exon whose 5’-end is attached to the BS adenosine, forming a lariat structure. In the second step the 3’ SS is attacked by the 3’ hydroxyl group of the 5’-exon, leading to joined exons and an excised intron in lariat form. The spliceosome is assembled in a step-wise manner and needs to undergo further conformational

changes to reach a state that is capable of catalyzing the two individual transesterification reactions (Figure 1.2). The same active site is used for both reactions, therefore reorganizations that reposition the products of the first reaction as substrates for the second one need to take place in between. Acting as a single-turnover enzyme, the spliceosome is disassembled after each exon ligation and the components can be used in the next cycle (Will & Lührmann, 2011; Hoskins & Moore, 2012).

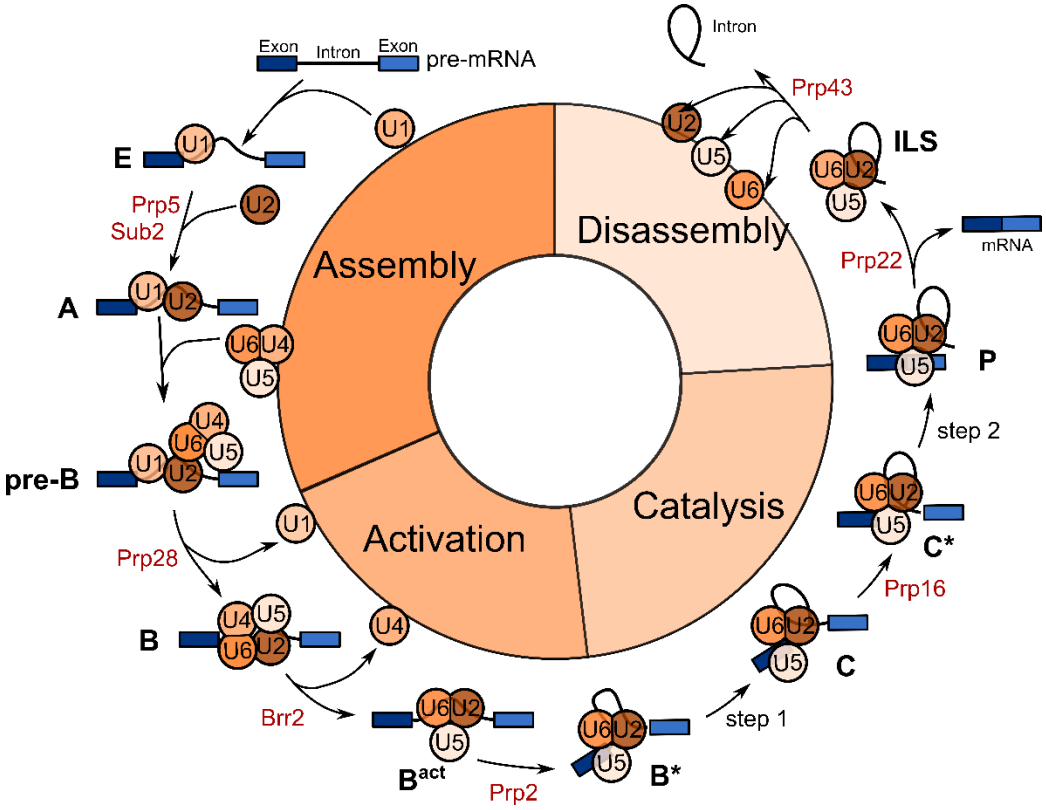


Figure 1.2: Schematic representation of the splicing cycle. Characterized spliceosomal complexes during sequential assembly, activation, the two transesterification steps and final disassembly are depicted. The U snRNPs are shown as circles and the involved RNA helicases are indicated in red. For clarity, no other non-snRNP components are shown. This figure was adopted from Absmeier *et. al.* (2016).

In *S. cerevisiae*, the action of eight different SF2 RNA helicases is required for splicing. The DEAD-box helicases Sub2, Prp5 (**p**re-m**R**NA-**p**rocessing factor 5) and Prp28 are involved in spliceosome assembly and remodeling necessary for activation. Spliceosome assembly on the pre-mRNA begins with the binding of the U1 snRNP to the 5'SS and the recognition of the BS by two non-snRNP (Seraphin *et al*, 1988; Siliciano & Guthrie, 1988; Abovich & Rosbash, 1997; Berglund & Chua, 1997). Thereby, the complex E is formed in an ATP independent manner. By the subsequent association of the U2 snRNP with the BS, the pre-spliceosome or complex A is

formed. In this step, the activity of both Sub2 and Prp5 is required. While Prp5 has been suggested to remodel the U2 snRNA to facilitate binding to the BS, Sub2 makes the BS accessible for base-pairing by displacing bound proteins (Kistler & Guthrie, 2001; Xu & Query, 2007). Addition of the pre-assembled U4/U6/U5 tri-snRNP leads to the formation of the pre-B-complex (Boesler *et al*, 2016; Bai *et al*, 2018) that contains almost all components known to be necessary for the splicing reactions. However, major rearrangements are still required to form the active site. In a first step towards activation, Prp28 destabilizes the interaction between the U1 snRNP and the 5'SS. The U1 snRNP is released from the complex, allowing the U6 snRNP to occupy its former position and thereby generate complex B (Chen *et al*, 2001; Staley & Guthrie, 1999). Then, the Ski-like helicase Brr2 mediates the unwinding of the U4/U6 duplex, triggering the release of the U4 snRNP as well as massive reorganizations in the snRNA interaction network, forming the B^{act} complex (Laggerbauer *et al*, 1998; Raghunathan & Guthrie, 1998). While this leads to the displacement of numerous splicing factors, several others are recruited (Fabrizio *et al*, 2009). Contrary to its name, the B^{act} complex is not yet catalytically active. Final activation yielding the B* complex is achieved by Prp2, which belongs, as all helicases involved in the subsequent steps, to the DEAH/RHA family (Kim & Lin, 1996). The action of Prp2, which is strictly dependent on its cofactor Spp2 (Suppressor of PRP protein 2), makes the BS adenosine accessible for the first nucleophilic attack by displacing several splicing factors and creates binding sites for new ones (Lardelli *et al*, 2010; Warkocki *et al*, 2009; Roy *et al*, 1995; Warkocki *et al*, 2015; Liu *et al*, 2007; Chiu *et al*, 2009). The B* complex is capable of catalyzing the first transesterification reaction between the BS and the 5'SS. The resulting complex, containing the freed 5'-exon and the lariat-intron-3'-exon, is denoted as C complex and requires further remodeling before the second transesterification reaction can take place (Krishnan *et al*, 2013; Galej *et al*, 2016; Ohrt *et al*, 2013; Konarska *et al*, 2006; Fica *et al*, 2017). Here, Prp16 employs a kinetic proof-reading mechanism to ensure fidelity of the BS recognition and enables exon ligation by facilitating the release of splicing factors (Burgess & Guthrie, 1993; Semlow *et al*, 2016; Tseng *et al*, 2011). The resulting complex C* is able to perform the second transesterification reaction, yielding complex P containing ligated exons and the excised intron in lariat (Bai *et al*, 2017; Liu *et al*, 2017; Wilkinson *et al*, 2017). In a further quality control step, Prp22 can remodel the spliceosome into a conformation incapable of exon ligation upon encountering a suboptimal 3'SS (Mayas *et al*, 2006). Subsequent to the splicing reaction, the post-catalytic spliceosome has to be disassembled. In a first step, Prp22 facilitates the release of the mature mRNA by disrupting its base-pairings with the U5 snRNA, obtaining the intron-lariat spliceosome (ILS) complex (Company *et al*, 1991; Mayas *et al*, 2006; Schwer, 2008; Semlow *et al*, 2016). The ILS is further disassembled by Prp43, which disrupts interactions of the U6 snRNA with the U2 snRNA and the intron, leading to dissociation of the single components (Arenas & Abelson, 1997; Martin *et al*, 2002;

Toroney *et al*, 2019). For this function, Prp43 depends on regulation by its cofactor Ntr1 (**N**ine**t**een complex-**r**elated protein 1 (Fourmann *et al*, 2016, 2017; Tsai *et al*, 2007; Tanaka *et al*, 2007)). Additionally, Prp43 is crucial for the disassembly of spliceosomes stalled at earlier stages (Pandit *et al*, 2006; Koodathingal *et al*, 2010; Fourmann *et al*, 2017). After spliceosome disassembly, most of the involved proteins and snRNPs are recycled to be used in subsequent cycles of splicing (Raghunathan & Guthrie, 1998).

Generally, the role of RNA helicases can be described as essential regulators of splicing as they promote conformational rearrangements of other spliceosome components that drive the transitions from state to state and are also involved in quality control by only allowing appropriate substrates to proceed through the cycle.

1.3 RNA helicases in ribosome biogenesis

Ribosomes are highly complex molecular machines responsible for the translation of the nucleotide sequence of mRNAs into the amino acid sequence of proteins. Eukaryotic ribosomes sediment at 80S and consist of a large 60S and a small 40S subunit, formed by ribosomal RNAs (rRNAs) and a multitude of ribosomal proteins (Trapman *et al*, 1975; Martin *et al*, 2013; Baßler & Hurt, 2019). While the small subunit contains only the 18S rRNA, three rRNAs (25S, 5.8S and 5S) are present in the large subunit (Martin *et al*, 2013; Baßler & Hurt, 2019). The 18S, 5.8S and 25S rRNAs are derived from a common precursor, the 35S pre-rRNA which is transcribed by RNA polymerase II, whereas the 5S rRNA is transcribed separately by RNA polymerase III (Martin *et al*, 2013; Baßler & Hurt, 2019). The initial step of ribosome biogenesis is the transcription of the 35S pre-rRNA in the nucleolus. Various ribosomal proteins and cofactor proteins assemble co-transcriptionally, forming the 90S pre-ribosome (Grandi *et al*, 2002). External as well as internal transcribed spacers are distributed around the rRNA sequences and contain sites for exo- and endonucleases, allowing sequential processing of the initial transcript. By early cleavage of the 35S pre-rRNA, which can also occur already during transcription, the 90S particle is split into the pre-40S and pre-60S pre-ribosomal particles, precursors to the small and large ribosomal subunits (Martin *et al*, 2013; Turowski & Tollervey, 2015). These particles follow independent maturation processes while they move from the nucleolus through the nucleoplasm and are then exported through the nuclear pores into the cytoplasm (Henras *et al*, 2015; Turowski & Tollervey, 2015). The maturation of pre-60S particles in the nucleoplasm is rather extensive and includes multiple processing events as well as several compositional and structural changes. Meanwhile, pre-40S particles are exported rapidly to the cytoplasm where they undergo their final maturation to translation-competent ribosomal subunits (Baßler & Hurt, 2019) (Figure 1.3). Additionally to the

processing, rRNAs undergo a series of base modifications, mostly mediated by small nucleolar RNPs (snoRNPs) and predicted to occur co-transcriptionally (Watkins & Bohnsack, 2012; Sloan *et al*, 2017). The snoRNPs contain snoRNAs which guide them to their target sites by basepairing with the pre-rRNA, where they introduce 2'-hydroxyl methylations, isomerize uridine to pseudouridine or facilitate processing reactions. (Kiss, 2002; Decatur & Fournier, 2003). Besides the snoRNPs and the previously mentioned nucleases, more than 200 non-ribosomal *trans* acting factors are involved in ribosome biogenesis. These include various structural/scaffolding proteins, GTPases, RNA modifying enzymes and RNA helicases (Martin *et al*, 2013; Baßler & Hurt, 2019). In yeast, 19 RNA helicases have been connected to ribosome assembly (Figure 1.3).

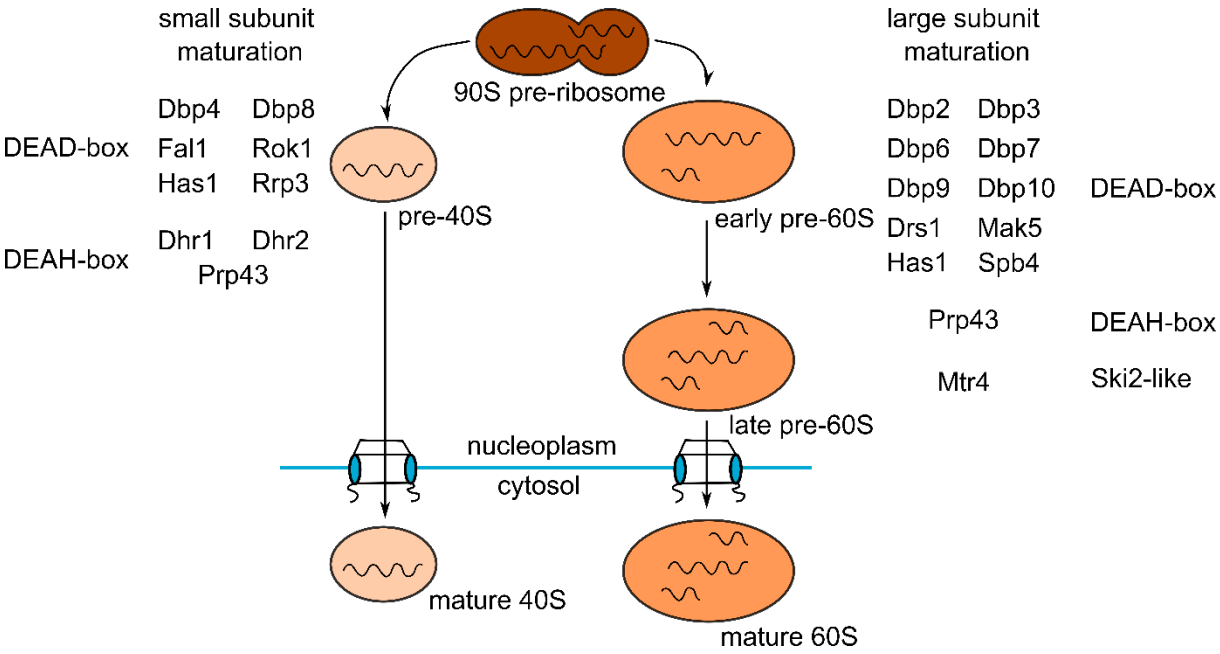


Figure 1.3: Schematic representation of ribosome biogenesis in *S. cerevisiae*. The major pre-ribosomal particles occurring during biogenesis of the small and large ribosomal subunit are shown. They follow independent maturation pathways in the nucleus and are exported to the cytoplasm for final maturation. RNA helicases involved in the synthesis of the small and large subunits are listed on the sides. This figure was inspired by Martin (2014) and Pelletier *et. al.* (2018).

While the majority of these RNA helicases are essential for cell viability, their precise roles remain largely unknown. With the exception of the Ski2-like helicase Mtr4, they belong to the DEAD- or DEAH-box families. Analyzing the accumulation of pre-rRNA intermediates upon depletion of a selected helicase allowed to identify the maturation pathway they are involved

in. Most are exclusively implicated in biogenesis of the large or small subunit, while two, Prp43 and Has1, could be shown to be required in both pathways (Martin *et al*, 2013).

For several RNA helicases acting early in ribosome assembly, a function in dismantling snoRNAs from the pre-rRNA has been described. The DEAD-box proteins Has1, Dbp4 and Rok1 as well as the DEAH-box protein Dhr1 release snoRNAs during maturation of the small subunit (Koš & Tollervey, 2005; Liang & Fournier, 2006; Sardana *et al*, 2015; Martin *et al*, 2014). During large subunit maturation, the DEAH-box helicase Prp43 is not only required for the release of snoRNAs but has also been suggested to promote association of certain snoRNPs with pre-60S complexes (Bohnsack *et al*, 2009). Another known function of RNA helicases is to remodel the pre-rRNA to promote correct folding and efficient processing. For Mtr4 and the DEAD-box protein Dbp3 roles in the processing of large subunit pre-rRNAs have been described while Prp43 has been shown to remodel cytoplasmic small subunit particles to facilitate the final processing step (Weaver *et al*, 1997; Fromm *et al*, 2017; Pertschy *et al*, 2009).

The limited understanding of the precise functions of RNA helicases in ribosome biogenesis is likely a result of their mostly transient association to pre-ribosomal particles as well as the high complexity and dynamics of the pathway and the relative inaccessibility of the nucleolus (Martin *et al*, 2013; Baßler & Hurt, 2019). Nevertheless, their remodeling activities are indispensable for the assembly of functional ribosomes, making them key effectors of compositional and structural changes of pre-ribosomal complexes.

1.4 Translocation and unwinding mechanisms of DEAD- and DEAH-box helicases

In splicing as well as ribosome biogenesis, mostly RNA helicases belonging to the DEAD- and DEAH-box families are involved. Interestingly, a clear division of labor between the two families can be observed in splicing. The early assembly steps are facilitated by DEAD-box helicases whereas DEAH-box proteins act in activation, catalysis and disassembly. While members of both families use energy provided by ATP hydrolysis to remodel their RNA substrates, they apply fundamentally different mechanisms. It is likely that the requirements on the remodeling helicases change throughout the cycle and are therefore best accommodated by different modes of action.

DEAD-box helicases

As all SF2 family members, DEAD box proteins share the highly conserved helicase core formed by two RecA-like domains. However, the composition and function of the adjacent N- and C-terminal extensions is highly diverse (Cordin & Beggs, 2013). Some of them direct the helicases to their respective site of action while others modulate the activity of the helicase core (Cordin

et al, 2006; Hilbert *et al*, 2009). DEAD-box helicases interact exclusively with RNA, as direct contacts with the 2'OH groups of the ribose moieties allow discrimination against DNA, and show no sequence specificity (Peck & Herschlag, 1999; Sengoku *et al*, 2006). Their substrates are unwound by local strand displacement in a non-processive manner (Yang *et al*, 2007). Single-stranded regions adjacent to the target duplex can facilitate loading of the helicase, however, no preference for 3' or 5' overhangs is discernable (Bizebard *et al*, 2004; Rogers *et al*, 2001; Yang & Jankowsky, 2006).

A multitude of structural and biochemical studies have contributed to understanding their mechanism of action (Figure 1.4): In the apo form, the RecA domains, which are connected by a flexible linker, are widely apart, adopting a 'non-productive' open conformation with no pre-formed active site for ATP hydrolysis. While the RecA domains each conserve their folds, the interdomain orientation differs between individual proteins (Caruthers *et al*, 2000; Cheng *et al*, 2005; Andersen *et al*, 2006; Zhang *et al*, 2013). The joint binding of ATP and an RNA substrate triggers domain closure, creating a functional ATPase active site formed by an intricate network of interactions between the conserved sequence motifs of the RecA domains (Sengoku *et al*, 2006; Andersen *et al*, 2006; Von Moeller *et al*, 2009; Collins *et al*, 2009; Theissen *et al*, 2008). In the closed conformation, a conserved α-helix extruded from the RecA1 domain forces the RNA in a bend conformation that is incompatible with duplex structure, resulting in local strand unwinding (Sengoku *et al*, 2006; Andersen *et al*, 2006; Collins *et al*, 2009; Von Moeller *et al*, 2009; Del Campo & Lambowitz, 2009). The use of single-molecule fluorescence resonance energy transfer (smFRET) techniques allowed the direct observation of these conformational dynamics in real-time for several DEAD-box helicases (Theissen *et al*, 2008; Beier *et al*, 2019; Aregger & Klostermeier, 2009). While ATP hydrolysis is not needed for duplex destabilization, it enables the RecA domains to return to the open conformation and release the remaining RNA strand (Liu *et al*, 2008; Cao *et al*, 2011). Interestingly, DEAD-box helicases can also act as RNA clamps, when they are trapped in the closed RNA-bound conformation by inhibition of ATP hydrolysis (Ballut *et al*, 2005; Alexandrov *et al*, 2012). As result of their mode of substrate unwinding, DEAD-box helicases lack directionality and are limited to short duplex regions (Gilman *et al*, 2017). Nevertheless, even the remodeling of complex tertiary structured RNA can be achieved through local unwinding. In a mechanism termed 'helix capture', binding to transiently exposed helices prevents them from re-forming tertiary contacts and their unwinding further disrupts the local structure. Thereby, even more helices are exposed to be bound and disrupted, eventually yielding a completely remodeled substrate (Pan *et al*, 2014).

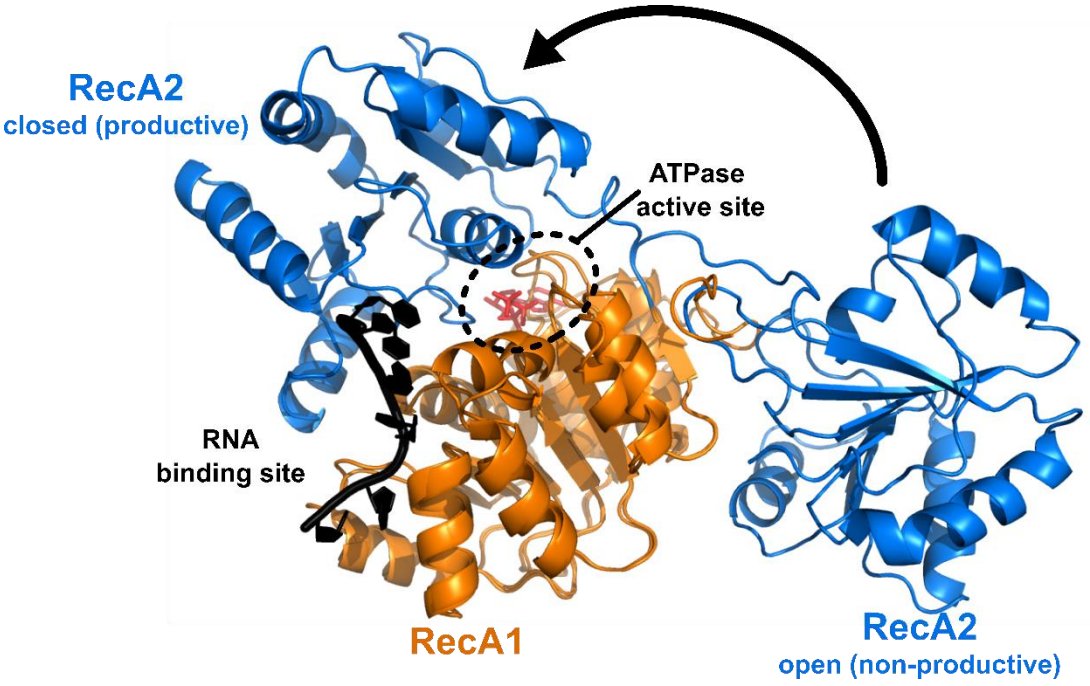


Figure 1.4: Conformational changes of DEAD-box helicases. The crystal structures of the DEAD-box helicase eIF4A-III in the apo form (PDB: 2hxy) and in complex with an ATP analog and an U_6-RNA (PDB: 2hyi) are superposed *via* their RecA1 domains. The RecA domains of the helicase in the apo state, shown in cartoon mode and colored orange and blue, are in an open unproductive conformation. Joint binding of ATP and RNA (depicted as red sticks and black cartoon, respectively) initiates domain closure, creating a functional ATPase active site.

DEAH-box helicases

DEAH-box helicases form a subfamily of the SF2 DEAH/RHA helicases. In contrast to DEAD-box helicases, their C-terminal domain architecture, consisting of a WH, HB and OB domain, is highly conserved (Figure 1.5a) (Walbott *et al*, 2010; He *et al*, 2010). According to the homologous domain of the DNA helicase Hel308, the HB domain was originally named ratchet domain (Büttner *et al*, 2007). Studies on the RHA helicase MLE showed no ratcheting function for this domain, consequently it has been termed ratchet-like or helix-bundle in later publications (Prabu *et al*, 2015). The N-terminal extensions of DEAH-box helicases show only little conservation, varying greatly in length as well as in function, and can be involved in subcellular targeting or recruitment of the helicase to its site of action (Cordin & Beggs, 2013; Fouraux *et al*, 2002; Schneider & Schwert, 2001; Wang & Guthrie, 1998). Due to the absence of the Q motif, any type of NTP can be hydrolyzed. DEAH-box helicases are not specific for RNA, some members of the family also act on DNA substrates (Fairman-Williams *et al*, 2010; Vaughn *et al*, 2005; Pyle, 2008). Double-stranded nucleic acids are preferably unwound with a 3'-5' polarity and a single stranded overhang is required for binding, however some DEAH-

box helicases have been shown to unwind with dual polarity (Fairman-Williams *et al*, 2010). The single stranded RNA is bound in a channel formed by the helicase core and the C-terminal domains, mostly by contacts with the sugar-phosphate backbone, resulting in a sequence-independent interaction (He *et al*, 2017; Prabu *et al*, 2015; Tauchert *et al*, 2017; Hamann *et al*, 2019). Crystal structures of Prp43 in complex with an ATP analog in the absence and presence of U_7 RNA suggested an RNA loading mechanism (Figure 1.5b). Structural rearrangements of the C-terminal domains lead to opening of the binding tunnel to form a cleft that closes again upon RNA binding (Tauchert *et al*, 2017; Roychowdhury *et al*, 2019). The unwinding activity of a Prp43 mutant trapped with a closed binding channel was shown to be severely impaired, underlining the importance of the opening motion (Tauchert *et al*, 2017).

The mechanism employed by DEAH-box helicases to unwind double-stranded RNA has been the subject of longstanding debates. They are generally thought to translocate with a 3'-5' directionality along RNA strands, thereby inducing duplex unwinding in a potentially processive manner. Guided by crystal structures of DEAH-box helicases in different conformational states, a mechanistic model for ssRNA translocation has been developed recently (Figure 1.5c) (Tauchert *et al*, 2017; He *et al*, 2017; Boneberg *et al*, 2019; Hamann *et al*, 2019; Chen *et al*, 2018; Robert-Paganin *et al*, 2017): The RecA domains of the helicase core cycle between open, nucleotide free and closed nucleotide-bound conformations. In the closed conformation, a stack of four RNA nucleotides is formed in the binding channel between a conserved loop in RecA1, and a β-hairpin in RecA2. When transitioning to the open state upon nucleotide release, the RecA2 domain shifts toward the 5' end of the RNA, leading to the accommodation of one additional nucleotide in the channel, resulting in a five-stack. Domain closure shifts the RecA1 domain towards RecA2, omitting the 3' most nucleotide from the channel. Thereby, translocation in 3'-5' direction with a step-size of 1 nt per hydrolyzed ATP is achieved (Hamann *et al*, 2019; Boneberg *et al*, 2019).

While features that are implicated in unwinding have been identified, the mechanistic connection between translocation and duplex unwinding remains unclear. A conserved β-turn in the RecA1 domain, the hook-turn, has been shown to be crucial for Prp43 to unwind double stranded RNA and disassemble ILSs (Tauchert *et al*, 2017). The interaction of a loop, termed C-loop, protruding from the HB domain with the 5' RNA region turned out to be required for the unwinding activity of Prp43. In Prp2, which does not display *in vitro* unwinding activity, this C-loop interacts with a loop in the RecA2 domain instead. As the sequence of the C-loop differs between the four spliceosomal DEAH-box helicases but is conserved in each individual helicase among different organisms, it likely plays a role in regulating their unwinding activity (Hamann *et al*, 2021).

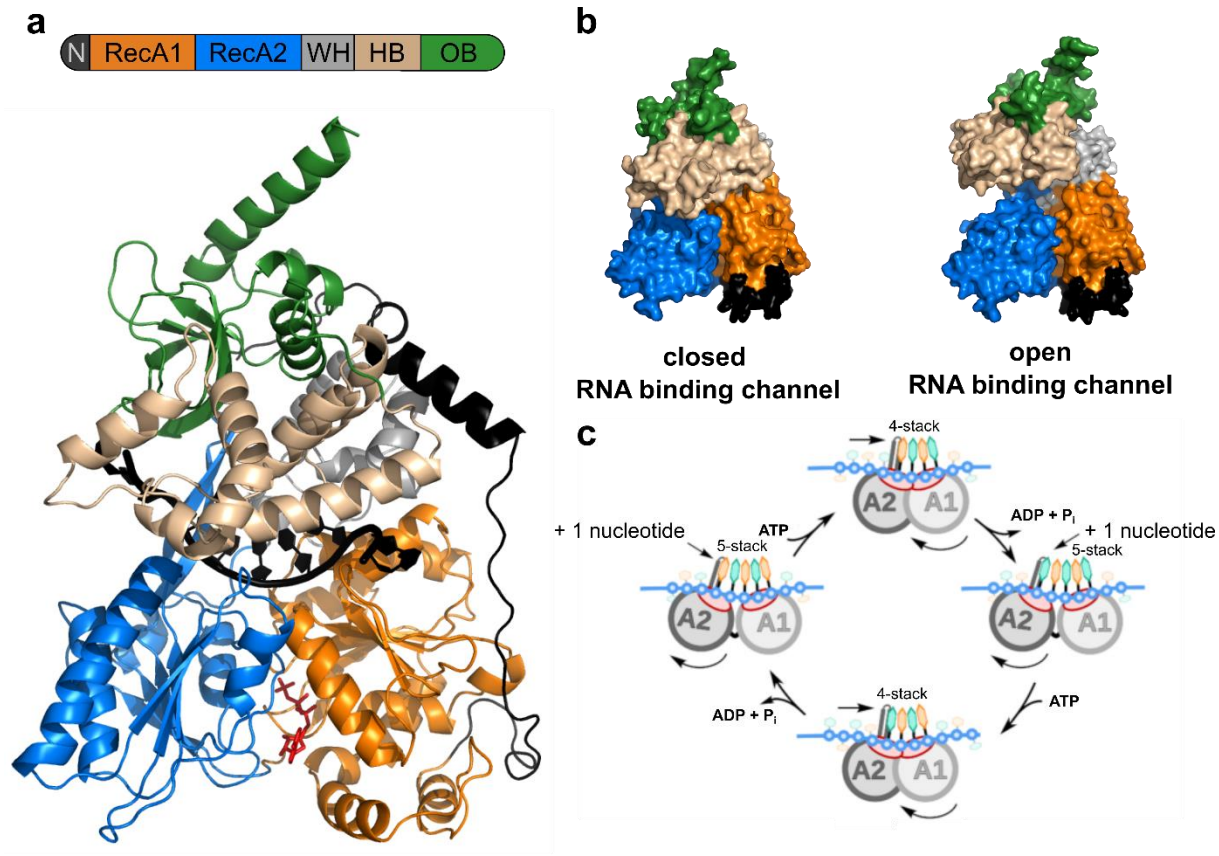


Figure 1.5: Overall structure and RNA interaction of DEAH-box helicases. (a) Crystal structure of the DEAH-box helicase Prp43 in complex with U_{25}-RNA and AMPPNP (PDB: 5i8q) in cartoon representation. The helicase core is formed by two RecA-like domains (RecA1: orange, RecA2: blue), the C-terminal domains comprise a winged-helix (WH: grey), helix-bundle (HB: wheat) and an oligonucleotide-binding (OB: green) domain. The N-terminal extension is shown in black. The NTPase active site is located between the RecA domains with the nucleotide depicted in red. In a channel formed by the helicase core and the C-terminal domains, the ssRNA (black sticks, nine nucleotides are resolved) is bound. **(b)** Comparison of Prp43 in complex with ADP-BeF_3^- in presence and absence of U_7-RNA (PDB: 5lta and 5ltk). The structures are depicted in surface representation and colored as in (a). In the RNA-free structure, the binding channel is widened to form a shallow groove. **(c)** Model for translocation of DEAH-box helicases. Continuous between open and closed conformations of the helicase core during ATP hydrolysis enables DEAH-box helicases to translocate in 3' to 5' direction along a single-stranded RNA. This part of the figure was adopted from Hamann *et. al.* (2019).

Based on structural comparisons with the Ski2-like helicase Hel308, the conserved β-hairpin in RecA2 has been suggested to be involved in double-strand separation (Büttner *et al*, 2007; He *et al*, 2010; Walbott *et al*, 2010). Contradictory, a crystal structure of Prp22 in complex with ssRNA shows the 5' end of the RNA locked in a conformation with the bases pointing away from the β-hairpin, positioning a potential double strand on the opposite side of this structural feature (Hamann *et al*, 2019).

Interestingly, direct interaction with double-stranded RNA regions might not be necessary for all DEAH-box helicase to fulfill their functions. Cryo-EM structures of various spliceosomal complexes revealed that the targets of DEAH-box helicases in this pathway are buried within large RNPs and thus inaccessible by direct translocation (Liu *et al*, 2017; Wan *et al*, 2017; Fica *et al*, 2017; Yan *et al*, 2016). This is in line with the fact that Prp16 and Prp22 have been shown to remodel their target RNAs inside the spliceosome by translocating towards but not through the interactions they disrupt. From this observation a mechanism described as winching was derived, where the force generated by translocation is transmitted to the site of remodeling at the interior of the RNP, disrupting RNA structures distant from the helicase binding site (Semlow *et al*, 2016). Comparing this mechanism to the local strand displacement employed by DEAD-box helicases could provide an explanation for the division of helicase induced remodeling of the spliceosome. While DEAD-box proteins act efficiently on accessible duplexes, DEAH-box helicases are required when more spliceosomal components are present in the complex and reduce access to the target sites (Gilman *et al*, 2017).

1.5 Regulation of RNA helicases

The careful regulation of RNA helicases is essential for normal cellular function, as they play crucial roles in various aspects of RNA metabolism and their dysregulation is often connected to disease (Steimer & Klostermeier, 2012). Most RNA helicases show intrinsically low specificity for their RNA substrates because they primarily interact with the sugar-phosphate backbone (Ozgur *et al*, 2015; Sloan & Bohnsack, 2018). Therefore, mechanisms that reduce indiscriminate interactions and ensure target specificity are needed. Additionally, many helicases show inherently low enzymatic activity, the modulation of which provides an additional layer of regulative potential (Sloan & Bohnsack, 2018).

Regulation of RNA helicases can be achieved by various means. Some RNA helicases contain additional domains that are able to recognize specific features of their targets, such as secondary structure elements or certain RNA modifications (Kretschmer *et al*, 2018; Luo *et al*, 2011; Wojtas *et al*, 2017). Others are kept in a low activity native state by autoinhibitory domains until structural rearrangements are prompted by substrate binding (Absmeier *et al*, 2020; Chakrabarti *et al*, 2011; Gowravaram *et al*, 2018). For some RNA helicases, regulation through post-translational modifications that allow specific activation upon interaction with a target site has been observed (Jacobs *et al*, 2007; Mathew *et al*, 2008; Song *et al*, 2017). The most common mechanism of regulation, however, is through the interaction with cofactor proteins that bind to the helicase to modulate their activity and in many cases also recruit them to their site of action (Sloan & Bohnsack, 2018; Bohnsack *et al*, 2021). Although many

helicases interact with a single specialized cofactor, some cofactor classes contain common domains that allow the interaction with a specific type of helicase (Chakrabarti *et al*, 2011; Erkizan *et al*, 2015; Granneman *et al*, 2006; Young *et al*, 2013; Schütz *et al*, 2008; Sloan & Bohnsack, 2018; Bohnsack *et al*, 2021). This concept is especially relevant for multifunctional helicases such as Prp43 that act on various targets at different cellular locations. The following section will explain in detail how the activity of Prp43 is modulated in its functions in splicing and ribosome biogenesis by cofactors belonging to the family of G-patch proteins.

Cellular functions of Prp43 and their regulation by G-patch proteins

The G-patch protein family comprises a variety of highly diverse proteins that only share the eponymous 40-50 amino-acid-long glycine-rich consensus motif and a general implication in RNA metabolism (Aravind & Koonin, 1999; Bohnsack *et al*, 2021). All G-patch proteins contain only a single consensus motif that is always located in an intrinsically disordered region. Besides seven highly conserved glycine residues, the G-patch motif contains three hydrophobic patches and an invariable aromatic amino acid following the second glycine (Figure 1.6). The occurrence of the G-patch motif is restricted to eukaryotes and some retroviruses (Aravind & Koonin, 1999). While five different G-patch factors have been identified in yeast, the increased regulative complexity of higher eukaryotes is reflected by the repertoire of more than 20 human G-patch proteins (Bohnsack *et al*, 2021). The majority of these proteins has been shown to be involved in the regulation of Prp43 in yeast or its human homologue DHX15. The only other interactions characterized so far involve a single G-patch partner for Prp2 and its human homologue DHX16 (Roy *et al*, 1995; Hegele *et al*, 2012). For several human G-patch proteins, no target helicase has been identified so far (Bohnsack *et al*, 2021).

Out of the four G-patch partners of Prp43 in yeast, only Ntr1 acts in the functional context of the spliceosome. It is part of the NTR complex (**n**ine**t**een-**r**elated complex) and recruits Prp43 to the spliceosome, where the helicase mediates the disassembly of the ILS in the late states of splicing (Wan *et al*, 2017; Tanaka *et al*, 2007; Fourmann *et al*, 2016) and where it disrupts stalled spliceosome complexes associated with suboptimal or mutated pre-mRNA substrates (Mayas *et al*, 2010; Fourmann *et al*, 2017). In this context, the C-terminal domain of Ntr1 hinders a productive interaction of its G-patch motif with Prp43 in functional spliceosomes before the ILS stage, preventing their premature disassembly (Fourmann *et al*, 2017). In human cells, DHX15 interacts with the Ntr1 homolog TFIP11 during ILS disassembly, implying a functional conservation (Yoshimoto *et al*, 2009). Additionally, several other human G-patch proteins have been reported to interact with DHX15 in the spliceosomal context, supposedly during maturation of the spliceosome (Niu *et al*, 2012; Chen *et al*, 2017).

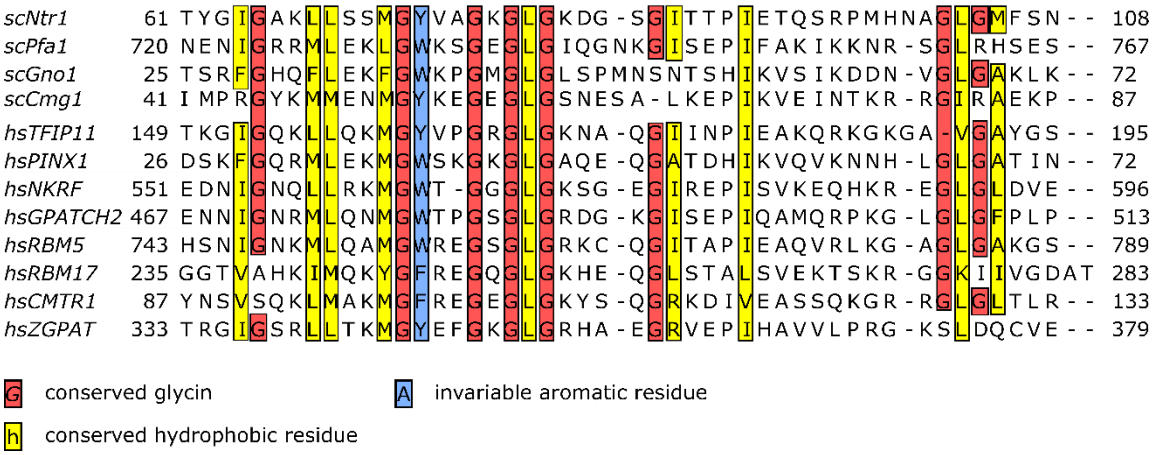


Figure 1.6: The G-patch motif. Sequence alignment of the G-patch motifs that interact with Prp43 in *S. cerevisiae* or its homolog DHX15 in *H. sapiens*. Conserved hydrophobic residues are highlighted in yellow, glycine residues in red and the invariable aromatic residue in blue.

In yeast ribosome biogenesis, Prp43 interacts with the G-patch proteins Gno1 and Pfa1 (**G**-patch **n**ucle**o**lar protein 1 and **P**rp **f**orty-three **a**ssociated 1). During early pre-rRNA processing steps Gno1 links Prp43 to precursors of the 60S ribosome, promoting release of snoRNAs from their pre-mRNA binding sites located in the 25S rRNA (Robert-Paganin *et al*, 2017; Chen *et al*, 2014; Bohnsack *et al*, 2009). The human homolog PINX1 was shown to stimulate yeast Prp43 *in vitro,* but the cellular role of the DHX15 PINX1 interaction remains uncharacterized so far (Chen *et al*, 2014). During the maturation of 40S ribosome subunits, the stimulation of Prp43 by Pfa1 facilitates 20S to 18S rRNA processing (Lebaron *et al*, 2009; Pertschy *et al*, 2009). Prp43 structurally rearranges the 3'-end of the 18S rRNA sequence, allowing the endonuclease Nob1 to gain access to its cleavage site (Pertschy *et al*, 2009). Although Pfa1 has a human homolog, GPATCH2, that stimulates DHX15, the helicase does not associate with late pre-40S particles, indicating a different function of this complex than in yeast (Lin *et al*, 2009; Sloan *et al*, 2019). The human G-patch factor NKRF interacts with DHX15 to facilitate an early endonucleatic cleavage event of the initial pre-rRNA transcript (Memet *et al*, 2017).

In yeast, a fourth G-patch protein interacts with Prp43. Cmg1 (**c**ytoplasmic and **m**itochondrial **G**-patch protein 1) localizes to the cytoplasm, where it stimulates Prp43 activity in a so far unidentified process (Heininger *et al*, 2016). Besides splicing and ribosome biogenesis the human DHX15 acts in pre-mRNA capping and transcription regulation where it associates with the G-patch protein CMTR1 (Toczydlowska-Socha *et al*, 2018; Inesta-Vaquera *et al*, 2018).

All G-patch partners of Prp43 characterized thus far stimulate the enzymatic activity of the helicase and enhance its affinity for ssRNA (Chen *et al*, 2014; Christian *et al*, 2014; Fourmann *et al*, 2017; He *et al*, 2017; Heininger *et al*, 2016; Lebaron *et al*, 2009; Tauchert *et al*, 2017). Additionally, their isolated G-patch motifs can bind to the helicase and are sufficient for efficient stimulation (Tauchert *et al*, 2017; Robert-Paganin *et al*, 2017; Lebaron *et al*, 2009; Chen *et al*, 2014; Christian *et al*, 2014). Interestingly, the identity of the G-patch motif is not connected to the cellular context as the G-patch motif of Pfa1 (Pfa1(gp)) can replace Ntr1(gp) during ILS disassembly (Fourmann *et al*, 2017). Therefore, the G-patch motif seems to primarily stimulate Prp43 in its functional context, whereas guidance to the target sites involves the complete G-patch partners.

Binding mode of the G-patch motif

While the G-patch motif was easily identified as minimal interacting region of G-patch proteins, the corresponding interaction sites on the helicase were only recently unambiguously characterized. Crystal structures of two helicase-G-patch complexes show the almost identical interaction mode of human DHX15-NKRF(gp) and Prp2-Spp2(gp) from *Chaetomium thermophilum* (ct) (Figure 1.7) (Studer *et al*, 2020; Hamann *et al*, 2020). In both complexes the mostly unstructured gp motif reaches across the backside of the RNA binding channel, connecting the WH and RecA2 domains of the helicase. Only the N-terminal part of the peptide forms a short α-helix that binds to a hydrophobic patch on the WH domain. A loop at the C-terminus of the peptide constitutes the second major interaction surface and packs into a hydrophobic cleft on the RecA2 domain. The connecting linker region barely interacts with the helicase, especially the conserved eponymous glycine residues make only few contacts. However, they are mainly responsible for the flexibility of this region, allowing conserved hydrophobic sidechains to contact hydrophobic cavities on the helicase surface. The N-terminal helix contributes most to the overall binding affinity and is strictly required for complex formation while the C-terminal loop provides a weaker interaction site. Both major interaction sites show strong sequence conservation among both G-patch proteins and gp-interacting helicases. Comparison to the cryo-EM density of the scPrp43-Ntr1 complex in the ILS shows a similar binding mode of the gp motif, suggesting a common interaction strategy (Hamann *et al*, 2020; Studer *et al*, 2020; Wan *et al*, 2017).

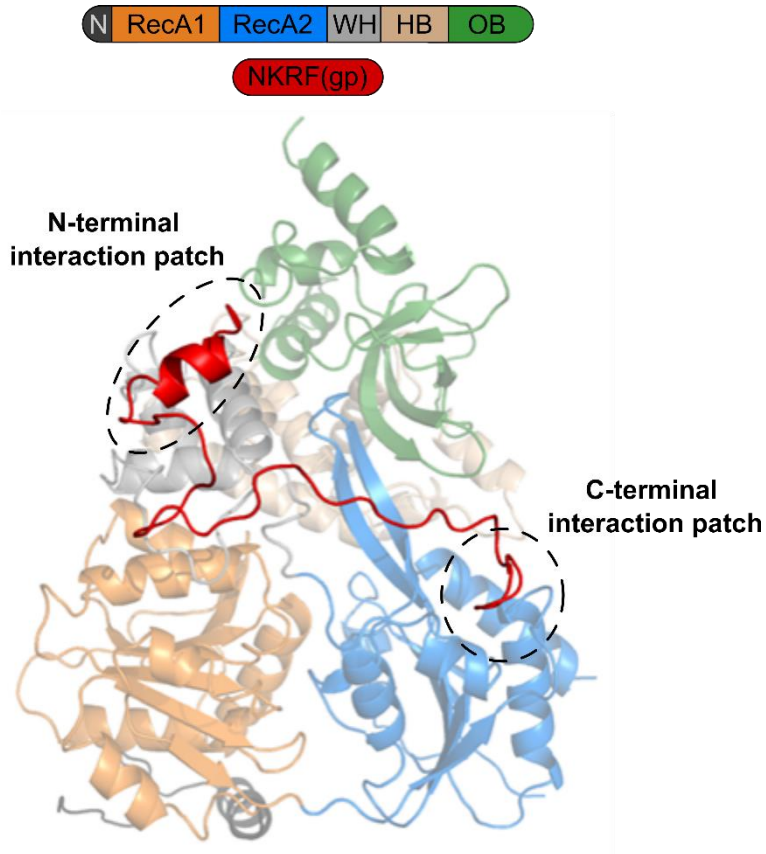


Figure 1.7: Binding mode of the G-patch motif. Cartoon representation of a DHX15-NKRF(gp) complex crystal structure (PDB: 6sh7). The domains of the helicase are shown in opaque and colored as in figure 1.5, the G-patch motif of NKRF is depicted in red. The N- and C-terminal interaction patches at the WH and RecA2 domains are connected by a flexible linker that stretches across the backside of the RNA binding channel.

A model for G-patch induced stimulation

As described earlier, DEAH-box helicases exhibit a high conformational plasticity. On the one hand the C-terminal domains can shift and rotate with respect to the helicase core, disrupting the RNA binding channel. On the other hand, the RecA domains adopt various conformations during the stages of ATP hydrolysis. However, the overall helicase conformation in the DHX15-NKRF(gp) complexes strongly resembles the RNA-bound structure of Prp43 in complex with ADP-BeF_3^- with closed RecA domains and intact RNA binding channel (Studer *et al*, 2020; Tauchert *et al*, 2017). Based on these findings a model for G-patch induced stimulation was proposed (Studer *et al*, 2020). The helicase is thought to be in a highly flexible autoinhibitory state in solution, where only few molecules assume a conformation compatible with efficient RNA binding and ATP hydrolysis. Upon G-patch binding, the conformational flexibility is restricted to stabilize a state with high RNA affinity. Due to the flexibility of the linker region,

domain movements required for ATP hydrolysis and translocation are still possible, allowing productive substrate interaction (Studer *et al*, 2020). According to this model, the observed stimulation of ATPase activity in presence of G-patch partners is mainly caused by constricting the conformational freedom of the RecA2 domain. Despite the high degree of flexibility of this domain observed in various crystal structures of DEAH-box helicases, none shows it in a position that is not in line with the range of motions executed during translocation (Hamann *et al*, 2019; Boneberg *et al*, 2019; Roychowdhury *et al*, 2019). It is therefore unclear if the restriction of flexibility alone is responsible for ATPase stimulation or if other contributions have to be considered. Additionally, a strong reinforcement of a conformation with closed RNA binding channel would come at the cost of less efficient RNA loading, requiring the single strand to enter by threading through a narrow opening. In the cellular context this might be compensated by the G-patch holding Prp43 in close proximity to its target RNA. However, the interaction with a G-patch partner increases the affinity for ssRNA from the µM to the nM range, indicating a reasonably efficient mechanism for RNA loading (Studer *et al*, 2020). Although the proposed mechanism is clearly reasonable on the basis of the underlying complex structures, some uncertainties remain. This is mainly due to the fact that it describes the regulation of a highly dynamic protein known to undergo several conformational changes during catalysis based on static observations provided by crystal structures. To meaningfully characterize the impact of G-patch interaction on the domain motility of Prp43, the repertoire of employed methods needs to be expanded to include real-time observations of dynamic motions.

1.6 Observing helicase domain motility in real-time

The wealth of structural information on DEAH-box helicases obtained by X-ray crystallography and cryo-EM studies offers detailed insights on the molecular interactions involved in substrate binding, various conformations the helicases can occupy and their positions in the context of huge cellular machineries such as the spliceosome. Biochemical studies deepen this knowledge by characterizing general enzymatic features such as their affinity and specificity for certain substrates or their kinetic properties during catalysis. However, methods that couple conformational rearrangements to easily observable changes of fluorescence signals are well suited to characterize the dynamic domain motions occurring during helicase activity. Utilizing the principle of Förster Resonance Energy Transfer (FRET) is a convenient way to visualize distance changes between mobile parts of a protein.

In FRET, energy is non-radiatively transferred from an initially exited donor fluorophore to an acceptor fluorophore whose excitation spectrum needs to overlap with the donor emission spectrum. The efficiency of this highly distance dependent process is described in the Förster equation,

$$E = \frac{1}{1 + (\frac{r}{R_0})^6}$$

where r is the donor acceptor distance and R_0 is the Förster radius, i.e. the intramolecular distance corresponding to an energy transfer of 50 % that is determined by properties of both fluorophores (Förster, 1946; Masters, 2014). Due to the inverse sixth power dependence on the distance, FRET can be used as a molecular ruler for distances relatively close to R_0. Depending on the choice of fluorophores, distance changes in a range of approximately 1 to 10 nm are easily observable using common donor acceptor pairs (Andreou & Klostermeier, 2012). While ensemble FRET measurements can be used to monitor structural dynamics of helicases, they describe the average population, limiting their ability to detect intermediate states or distinguish between heterogeneously behaving molecules. In contrast to this, single-molecule methods allow the observation of domain movements of individual molecules in a time resolved manner. For example, fluorescently labeled molecules can be immobilized on a surface while the individual FRET pairs are monitored by Total Internal Reflection Fluorescence (TIRF) microscopy (Figure 1.8a) (Roy *et al*, 2008). This method utilizes the evanescent field generated upon total internal reflection of light on an interface between media with increasing refractive indices (e.g. a glass cover slip and an aqueous solution) to exite the fluorophores (Figure 1.8b) (Poulter *et al*, 2014). Due to the amplitude of the evanescent field decaying exponentially, only fluorophores attached to surface immobilized molecules are exited, allowing the detection of single molecules. Emission of donor and acceptor are separated using dichroic mirrors and recorded as simultaneous time courses from which FRET can be calculated. Each trajectory represents the continuous observation of one individual molecule during the time frame of the experiment, until the fluorophores photobleach (Figure 1.8c) (Roy *et al*, 2008). Alternatively, confocal microscopy can be applied to observe freely diffusing labeled molecules. Here, a laser is focused to a diffraction-limited spot, creating a suitably small observation volume. Only molecules diffusing through this confocal volume are exited, generating bursts of fluorescence on the millisecond timescale. Thereby, snapshots of several individual single-molecules are collected during the course of the experiment (Andreou & Klostermeier, 2012).

Several studies used single molecule approaches to characterize the conformational changes associated with opening and closing of the DEAD-box helicase core, contributing substantially to the understanding of their mechanism of action and the impact of interaction with cofactor proteins (Aregger & Klostermeier, 2009; Andreou & Klostermeier, 2012; Hilbert *et al*, 2011; Theissen *et al*, 2008; Beier *et al*, 2019). Although comparable studies targeting DEAH-box helicases are not available so far, smFRET is well suited to study the impact of G-patch interaction on the domain motility of Prp43.

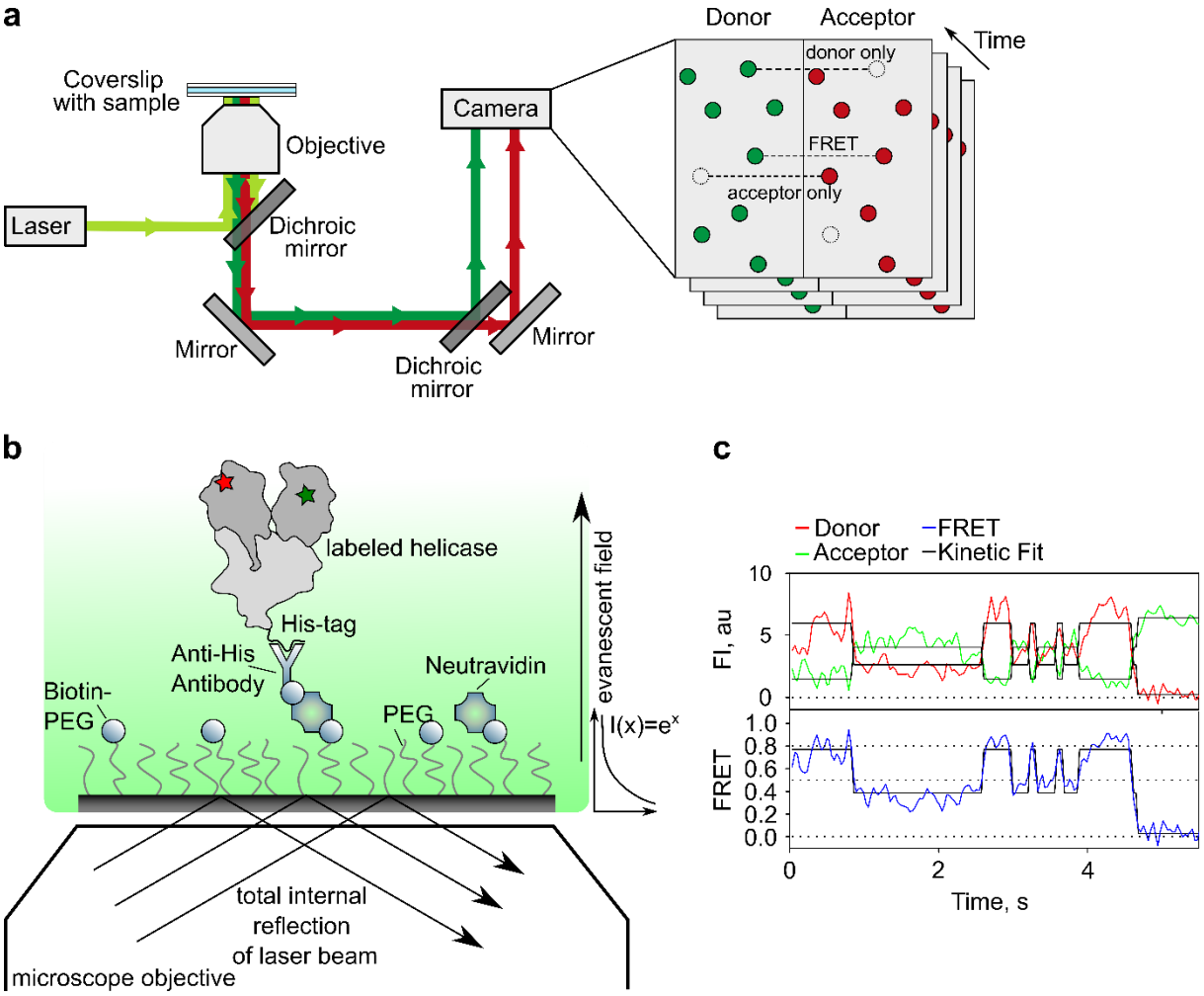


Figure 1.8: Single molecule detection by TIRF microscopy. (a) Schematic representation of an objective-based TIRF microscopy setup. The excitation laser beam (light green) is reflected at the interface between coverslip and sample. Dichroic mirrors split the emission fluorescence of donor (green) and acceptor (red) that are then guided towards the camera chip, where they are displayed in separate channels. The figure was inspired by Lerner *et. al.* (2021). **(b)** Immobilization of labelled helicase molecules *via* biotin-neutravidin interactions of the functionalized coverslip and an antibody against the C-terminal His_6-tag **(c)** Representative smFRET time trace showing donor Cy3 (green) and acceptor Cy5 (red) fluorescence intensities, calculated FRET (blue) and kinetic fitting of the data (black).

1.7 Scope of this thesis

This thesis focusses on the elucidation of the mechanistic principles behind the activation of the DEAH-box helicase Prp43 by G-patch proteins. The results of this work are summarized in one manuscript that is currently under peer-review (chapter 2) and one manuscript in preparation (chapter 3).

At the beginning of this thesis, structural data on Prp43 in various nucleotide- and ssRNA-bound complexes were available, but the interaction with G-patch proteins had only been characterized biochemically. Shortly after this thesis was started, the structural repertoire describing the catalytic cycle of DEAH-box helicases was extended further to also include adenosine nucleotide-free states in presence and absence of ssRNA and a model for RNA translocation was postulated. However, the impact of G-patch proteins on this mechanism was not addressed. Recent structural studies were able to characterize the overall binding mode of the G-patch motif and resulted in the model for G-patch induced stimulation described earlier. In the present work, a single-molecule approach was chosen to accommodate the known structural plasticity of DEAH-box helicases and complement the available structure based models with time resolved information on conformational dynamics. Using smFRET combined with TIRF microscopy, the effect of G-patch interaction on the domain motility and RNA interaction of Prp43 was studied.

In a first project, a smFRET system reporting on the RecA domain movement of Prp43 was developed and analyzed, giving insights in conformational dynamics during ATP hydrolysis and their modulation by interaction with Pfa1(gp). To complement this, a second reporter used smFRET labels on the RNA substrate to follow RNA binding and unwinding by the Prp43-G-patch complex. Combining observations from both systems revealed the regulation mechanism of Prp43 by G-patch partners and led to a mechanistic model of its motility cycle.

A second project was focused on the initial step of Prp43 activation: the loading of the helicase onto ssRNA. A smFRET system reporting on the conformation of the RNA binding channel of Prp43 elucidated the dynamic equilibrium of the channel and its modulation through G-patch interaction. The results illustrate the role of the channel conformation for productive RNA interaction of Prp43 during loading and processive translocation.

Chapter 2: Regulation of Prp43 by the G-patch factor Pfa1

This manuscript was submitted to PNAS and is currently in review. Reviewer comments were received 13th April 2022 and resubmission is scheduled for July 2022.

Regulation of the DEAH/RHA helicase Prp43 by the G-patch factor Pfa1

Marieke Enders[1], *Ralf Ficner*[1], *Sarah Adio*[1*]

Department of Molecular Structural Biology, Institute of Microbiology and Genetics, Georg-August-University Göttingen, Justus-von-Liebig-Weg 11, D-37077 Göttingen, Germany

*Sarah Adio

Email: sarah.adio@uni-goettingen.de

Author Contributions: M.E. and S.A. performed experiments and data analysis. M.E., R.F. and S.A. designed experiments and interpreted the data. M.E. and S.A. and wrote the paper with contributions of R.F.

Competing Interest Statement: The authors do not have any competing interests.

Classification: Biological Sciences, biochemistry

Keywords: DEAH/RHA helicases, G-patch proteins, ribosome biogenesis, splicing, molecular machines

Abstract

The DEAH/RHA helicase Prp43 remodels protein-RNA complexes during pre-mRNA splicing and ribosome biogenesis. The Helicase activity and ATP turnover are intrinsically low and activated by G-patch (gp) factors in the specific cellular context. The gp motif connects the helicase core to the flexible C-terminal domains, but it is unclear how this affects RecA domain movement during catalysis and the unwinding of RNA substrates.

We developed single-molecule (sm) FRET reporters to study RecA domain movements within Prp43 in real-time. In the apo and nucleotide-bound state, the domains move together adopting a closed conformation. Addition of Pfa1(gp) induces an open state that is incompatible with nucleotide-binding and shows rapid ADP release rates accelerating the transition from the weak (ADP) to the strong (apo) RNA binding state.

Using smFRET labels on the RNA to probe substrate binding and unwinding, we demonstrate that Pfa1(gp) enables Prp43(ADP) to switch between bound and unbound states instead of dissociating from the RNA. ATP binding to the apo-enzyme induces translocation along the RNA, generating the unwinding force required to melt proximal RNA structures.

During turnover, Pfa1(gp) stimulates ATP hydrolysis, promoting cycling of the RecA domains between open and closed states. Consequently, translocation becomes faster than drop-off from the substrate in the ADP-state, allowing processive movement along the RNA.

We provide a mechanistic model of DEAH/RHA helicase motility and reveal the principles of regulation by their G-patch partners.

Significance

DEAH/RHA helicases remodel RNP complexes in central processes of RNA metabolism such as transcription, splicing or translation at the cost of ATP hydrolysis. These enzymes require activation to function efficiently in their specific cellular context. We show that G-patch factors do not only recruit the DEAH/RHA helicase Prp43 to its cellular localization, but also interfere with its mechanism of motion along RNA substrates. Instead of dropping off the substrate after a single round of catalysis, Prp43 bound to G-patch translocates continuously along the RNA, generating force required to resolve stable RNA structures. DEAH/RHA helicases are implicated in severe human diseases, i.e. cancer and viral infections. Understanding the principles underlying their regulation might contribute to the development of new therapeutic strategies.

Introduction

RNA helicases of the DEAH/RHA family are involved in key steps of RNA metabolism by reorganizing structured RNAs and RNA-protein complexes (RNPs) through ATP-dependent motility in 3' to 5' direction along single-strands (1). They share conserved sequence motifs (I, Ia, Ib, II, III, IV, V and VI) with the helicase superfamily 2 (SF2), which spread over two RecA domains that together form the helicase core (2). These essential motifs are responsible for ATP and RNA binding, ATP hydrolysis and the coupling between hydrolysis and unwinding (2-4). Further domains include a common set of C-terminal domains composed of a winged helix, a helix bundle (previously called ratchet-like domain), and an OB-fold domain as well as varying N-terminal extensions (5, 6). RNA is bound in a sequence independent manner in a channel between the RecA and the C-terminal domains (7-12). A mechanistic model for ssRNA translocation has been developed from crystal structures of DEAH/RHA helicases in different conformational states (7–10, 13, 14): Cycling between open, nucleotide-free and closed, nucleotide-bound conformations of the helicase core leads to alternate accommodation of either four or five RNA nucleotides stacked in the binding channel with a conserved loop in RecA1 and a β-hairpin in RecA2 serving as 'bookends' preventing slippage. Thereby, translocation in 3'-5' direction with a step-size of 1 nt per hydrolyzed ATP is achieved. *In vivo* targets of DEAH/RHA helicases often lie buried within large RNPs and are not accessible by direct translocation (15-18). In a mechanism described as winching, it was proposed that the force generated by translocation is transmitted to the interior of the RNP disrupting RNA structures distant from the helicase binding site (19). Nevertheless, the precise mechanism linking translocation and the unwinding of duplex RNA structures is still unknown.

Some specialized DEAH-box helicases target their substrates with high sequence specificity, such as DHX36, which binds RNA- and DNA G4-quadruplex structures with high affinity (13, 20). However, the vast majority of DEAH-box helicases bind RNA with low intrinsic specificity and employ auxiliary protein factors for the recruitment to their cellular place of action (21-23). Prp43 and its human homologue DHX15 are such ubiquitous DEAH-box helicases which are activated in different functional contexts by so-called G-patch (gp) factors, named after a 40-50 amino-acid-long glycine-rich consensus motif (24, 25). In yeast, Prp43 interacts with four different G-patch partners. Ntr1 recruits Prp43 to the spliceosome where it mediates the disassembly of the intron lariat complex in late states of splicing (16, 26, 27) and where it disrupts stalled spliceosome complexes associated with suboptimal or mutated pre-mRNA substrates (28, 29). Together with Pfa1 Prp43 promotes 20S to 18S rRNA processing during the maturation of 40S ribosome subunits (30, 31), while the G-patch protein Gno1 links Prp43 to precursors of the 60S ribosome promoting release of small nucleolar RNAs from their pre-mRNA binding sites (14, 32, 33). In the cytoplasm, Prp43 interacts with the G-patch protein

Cmg1, but the function and target of this complex remain to be identified (34). Functions of the human DHX15 protein additionally include RNA unwinding during mRNA capping, and transcription regulation (35, 36).

So far, eight of the more than 20 human G-patch factors are known to regulate DHX15 activity, however the interaction partners of most remaining gp factors are still undiscovered (25). Besides the G-patch motif, these factors share little similarity regarding size, structure and domain organization except for the high prevalence of RNA binding domains linking them to the various pathways of RNA metabolism (25). However, there is mounting evidence that they employ a common mechanism regulating the activity of Prp43/DHX15.

G-patch factors enhance the ATPase and helicase activities of Prp43/DHX15; interestingly, the gp motif alone is sufficient for efficient stimulation (7, 14, 31, 32, 36–39). In the spliceosome, Ntr1(gp) and even Pfa1(gp) can fully replace full-length Ntr1 during Intron Lariat Spliceosome (ILS) disassembly (27, 29). Together, these findings demonstrate that the gp motif primarily stimulates Prp43 activity in a specific cellular context rather than guiding Prp43 to its binding site. The gp motif is intrinsically disordered and only forms secondary structure elements upon contact with a helicase (37, 40). This allows sampling of the whole helicase surface, making recruitment to the target RNP more efficient compared to a structured domain (41). In the crystal structure of the DHX15-NKRF(gp) complex, the gp-motif forms a flexible brace between the RecA2 and WH domain in DHX15 reaching across the back side of the RNA binding channel (42). The overall conformation strongly resembled the RNA-bound structure of Prp43 in complex with ADP-BeF_3^{2-} with closed RecA domains (7). Based on these findings the authors proposed that gp binding stabilizes a compact state of the helicase competent of ATP turnover and with a tight grip on the RNA substrate (42). Comparison of the DHX15—NKRF(gp) complex with the scPrp43—Ntr1 complex in the ILS (16) shows a similar binding mode of the gp motif to the helicase suggesting that activation occurs by comparable mechanisms. It is important to notice that Wan et al. report the opening of RecA domains in Prp43 bound to NTR1(gp) within the ILS complex which differs from the observations by Studer et al. (16, 42).

Overall, the molecular mechanism of helicase stimulation by gp factors remains unclear, which is mainly due to the lack of studies investigating the structural dynamics of Prp43 during catalysis. Open questions concern how gp factors alter structure and function of Prp43 during ATP turnover and how gp-dependent changes affect Prp43 motility on the RNA. Here, we use single molecule Förster Resonance Energy Transfer (smFRET) to monitor in real-time how the G-patch motif affects the conformational cycling of Prp43 RecA domains. Additionally, using smFRET labels on the RNA substrate, we follow RNA binding and unwinding by the Prp43-G-patch complex. The method allows us to detect transient conformational intermediates of

Prp43 and the RNA during turnover. We show that the G-patch motif induces the open state of RecA domains which is incompatible with nucleotide- binding and facilitates nucleotide release. We confirm that the G-patch motif enhances RNA binding and demonstrate that ATP binding generates sufficient force to melt short RNA duplexes. During turnover, the G-patch accelerates transitions of the RecA domains between the open and closed state which couples ATP hydrolysis to processive motion along the RNA. Our results reveal the principal regulation mechanism of DEAH-box helicases by gp partners and lead to a mechanistic model of their motility cycle. As dysregulation of DEAH-box helicases is often involved in the development of diseases, our model could provide valuable insights towards understanding the underlying principles.

Results

Pfa1(gp) stabilizes the open conformation of Prp43 RecA domains

To analyze domain movements in Prp43, we generated fluorescence labeling sites at position C170 and C303 in RecA1 and RecA2, respectively, of *Chaetomium thermophilum* Prp43 (ctPrp43) (Figure 1A, Methods). ctPrp43 shows high sequence and structural similarity with the homolog in *Saccharomyces cerevisiae* (scPrp43) and can fully replace the function in yeast spliceosome disassembly (43) but turned out to be more stable in biochemical experiments and was better compatible with the labeling process.

We validated that rates of ATP turnover by labeled ctPrp43 in the absence and presence of saturating concentrations of ctPfa1(gp) and ssRNA are similar to wild type ctPrp43(wt) and that the extent of stimulation is in the same range as for scPrp43 (Supplementary Figure S1A and S1B). This indicates that the catalytic activity is not impaired by the fluorescence tags and that ctPrp43 is a suitable substitute for the yeast protein. We immobilized Prp43 on streptavidin-functionalized glass cover slips (44) using a biotin-tagged antibody against the C-terminal His-tag and followed FRET on individual Prp43 molecules at steady state conditions by Total Internal Reflection of Fluorescence (TIRF) microscopy (Figure 1B). In the absence of ligands, Prp43 shows a single population of signals with E_{FRET}= 0.77 ± 0.03 (Figure 1C, Supplementary Figure S2A, Supplementary Table 1), corresponding to an estimated donor-acceptor distance of 4.3 nm. This value is comparable to the distance between the C_α atoms of the labeled residues in the crystal structure of ctPrp43—ADP where the RecA domains are in the closed (C) conformation (4.2nm, PDB: 5d0u (43), Supplementary Table 2). Next, we followed smFRET of the Prp43-Pfa1(gp) complex in order to test if the G-patch alters the conformation of RecA domains in Prp43 (Figure 1D). In addition to the C state, we observe a

second population of signals with lower FRET efficiency (E_{FRET}= 0.57 ± 0.04, Supplementary Figure S2B, Supplementary Table 1) indicating the opening (O) of the RecA domains in Prp43 in response to Pfa1(gp) binding. O and C states are nearly evenly distributed (Supplementary Table 1) and a small fraction of smFRET traces (7%) shows transitions between the states suggesting slow exchange between the populations (Supplementary Figure S2C and S2D). In the presence of ADP, Prp43 and the Prp43-Pfa1(gp) complex show similar FRET states compared to the apo-enzyme (Figures 1E and 1F, Supplementary Table 1). However, for the gp complex, the O state is less prevalent in the presence of ADP, which agrees with the notion that contacts between RecA domains and the nucleotide stabilize the C conformation (42). Interestingly, when bound to ssRNA Prp43 can access the O state in the absence of Pfa1(gp), while the conformation of the Prp43-Pfa1(gp) complex is almost entirely shifted towards the O state (Figure 1G and 1H). Hence, opening of the Prp43 RecA domains is principally possible through binding of the RNA substrate, similar as for other DEAH-box helicases such as DHX36 or Prp22 (10, 13). Binding of Pfa1(gp) stabilizes the O conformation such that the occupancy of the C state becomes less prominent in the population distribution of Prp43-Pfa1(gp) complexes (Figure 1H).

Pfa1(gp) promotes ADP release by Prp43

To explore the implications of the open RecA domains on Prp43 function, we compared the position of the C_α atoms of the labeled residues C170 (K in the wt protein) and C303 in the structures of ctPrp43—ADP with closed Rec A domains (PDB: 5d0u (43)) with the position of the corresponding residues in the scPrp43—Ntr1 complex in the spliceosomal context (PDB: 5y88(16)) and in the highly homologous DEAH-box helicase ctPrp22 in complex with a ssRNA (PDB: 6i3p (10)) where the RecA domains are open (Supplementary Table 2). In the open state structures, the RecA domains were about 0.8 nm further apart from each other than in the closed state structures, which agrees with the distance change of approximately 1 nm derived from the smFRET data (Supplementary Table 2). Further analysis of the nucleotide binding site showed that the open RecA domain conformation is incompatible with ADP binding by Prp43 (Figure 2A). While ADP is bound through interactions with both RecA domains in the closed state structure, only the interactions to one of the domains could be maintained when modeling ADP in the open state structures, suggesting that ADP is released when Prp43 enters the O state. To test this idea, we measured the dissociation rate of *N*-methylanthraniloyl-ADP (mant-ADP) from Prp43 by stopped-flow (Figure 2B). In the absence of Pfa1(gp) and ssRNA, mant-ADP release followed a single exponential time course with a slow rate of k_{off}=0.002 s^{-1}. Mant-ADP release was accelerated by 3-fold in the presence of ssRNA and by 100-fold in the presence of Pfa1(gp) (Figure 2B) indicating that the opening of RecA domains observed by smFRET correlates with fast ADP release rates (Figure 1D and 1G). Previous studies have

shown that Prp43 and Prp43-gp complexes in the nucleotide-free state target their RNA substrates with higher affinity than in the ADP-state (10, 42). The low affinity state allows Prp43 to move along the RNA, presumably because of reduced contacts between RecA2 and the RNA (10). Hence, the fast ADP release induced by Pfa1(gp) accelerates the transition back to the tight RNA binding state and might prevent Prp43 drop-off during translocation.

Pfa1(gp) enhances RNA binding of Prp43 while ATP binding generates force and induces partial unwinding

To further investigate how Pfa1(gp) modulates Prp43 interaction with the RNA substrate, we determined the binding affinity in our system (Figure 3A). Consistent with previous results (10, 34, 42) the affinity of Prp43 for ssRNA in the absence of nucleotide was in the µM range without Pfa1(gp) and in the nM range in the presence of Pfa1(gp) (K_D =7 µM and K_D =0.03 µM, respectively, Figure 3A). In the presence of AMPPNP, the affinity of Prp43 was slightly increased, while the Prp43-Pfa(gp) complex bound with similar affinity compared to the nucleotide-free states (K_D=0.4µM and K_D=0.08µM, respectively). In the presence of ADP however, the affinity decreases severely without Pfa1(gp), in fact the K_D was so low that a full titration was not feasible (K_D>100µM). The Prp43-Pfa1(gp) complex showed moderate affinity in the presence of ADP (K_D=0.5 µM), confirming that the G-patch promotes binding of the helicase to the substrate and that the choice of the nucleotide (ADP, AMPPNP or nt-free) regulates the switch between weak or strong binding states.

Next, we aimed to visualize RNA binding in a time-resolved manner. To this end, we designed an RNA probe with a single-strand 3'-overhang allowing for Prp43 binding, followed by a stem loop, which can be unwound during Prp43 translocation (Figure 3B). The 3'-end of the substrate was biotinylated to immobilize the probe on the cover slip, the Cy3-dye was placed in the single-strand region (position 76) and the Cy5-dye was placed at the 5'-end (position 1) such that the distance between the dyes is 15 nucleotides. In the absence of the helicase the probe shows stable FRET signals with mean values of E_{FRET}=0.8± 0.01 (Supplementary Figure S3A and S3B). To test if Prp43 binding changes the FRET state of the RNA probe, we incubated the RNA with Prp43 in the absence of nucleotide and observed two types of signals (Figure 3C, Supplementary Figure S3A, Supplementary Table 3). The majority of RNA molecules (76 %) showed stable signals with E_{FRET}=0.79 ± 0.01, similar to the free (F) RNA probe. A smaller fraction (24 %) displayed signals with E_{FRET}=0.51 ± 0.02, corresponding to RNA bound (B) by the helicase. Addition of ADP shifted the distribution of FRET values entirely to the F state, which agrees with the very low binding affinities observed by fluorescence polarization (Figure 3A, Supplementary Figure S3C). Binding of the Prp43-Pfa1(gp) complex shifted the population towards the B state (88 % of molecules are in B), while high concentrations of Pfa1(gp) without

Prp43 resulted only in the F state, confirming that the G-patch enhances Prp43 binding to the RNA (Figure 3E, Supplementary Figure S3D, Supplementary Table 3). Upon addition of ADP, the probe interconverted between F and B states indicating binding and dissociation of the Prp43-Pfa1(gp)-complex (Figure 3F, Supplementary Table 3). Traces with B to F transitions were used to estimate the dissociation rate of the complex in the low affinity ADP state (k_{off}=0.45 s^{-1}, Figure 3F, Supplementary Table 4). This rate is one order of magnitude lower compared to the rate of ATP turnover (k_{cat}=5 s^{-1}, Supplementary Figure S1A). Hence, binding, hydrolysis and release of ATP are much faster than drop-off of the complex from the RNA in the weak-affinity state which suggests that Prp43-Pfa1(gp) can perform multiple rounds of ATP turnover while staying bound to the substrate.

Next, we investigated the interaction between the RNA and Prp43-Pfa1(gp) at turnover conditions. At physiological ATP concentrations (2 mM) FRET signals vanish immediately because rapid unwinding of the stem loop increases the Cy3-Cy5-distance to 75 nt, which is far beyond the FRET range (data not shown). At very low ATP concentration (2 µM), where the binding rate of the nucleotide limits the turnover of the helicase, the majority of molecules is in the B state. However, a fraction of the population shows a new state with E_{FRET}=0.35 ± 0.05 (Figure 4A). We assume that Prp43 binding straightens the RNA probe leading to the decrease in FRET in the B state. Hence, further reduction of FRET values most likely corresponds to the partial unwinding (PU) of the stem loop. The PU state is transiently sampled from the B state and we observed it in 20% of all molecules (Figure 4B). Analysis of the transition state frequency between B, F and PU states revealed that 33% of observed transitions occur between B and PU, indicating that a significant portion of RNA molecules undergo partial unwinding (Figure 4C). A very small number of molecules showed direct transitions between F and PU, presumably the B state was too short-lived to be captured in our experiment. In an attempt to distinguish events resulting from ATP binding or the hydrolysis of ATP we performed the experiment again using AMPPNP, which is hydrolyzed at a very low rate by Prp43 (Supplementary Fig S4A). However, the distribution of FRET states of the RNA probe, in particular the sampling of the PU state, was similar as in the experiments with ATP (Supplementary Fig S4B). The PU state was also sampled when we incubated the probe with a Prp43 variant ($Prp43_{E216A}$) that hydrolyses ATP at very low rate, comparable to the AMPPNP hydrolysis rate by Prp43(wt) (Supplementary Fig. S4A and S4C) (8, 45). Although we cannot distinguish between binding and hydrolysis, in the time-course of our experiment (33 sec) not more than 1 ATP is hydrolyzed (k_{cat}=0.04 s^{-1}). Therefore, partial unwinding occurs in a single round of ATP turnover. Previous models of DEAH-box helicase translocation along the RNA suggested that ATP binding, rather than hydrolysis provides the power stroke for translocation (9). This explains why our RNA probe is unwound in an ATP-dependent manner but also in the

absence of actual ATP hydrolysis. To estimate how many RNA nucleotides are unwound during the transition from the B to the PU state, we used dsDNA oligos as distance ruler, where Cy3- and Cy5-fluorophores were placed on opposing strands at defined distances between 7 and 19 nucleotides (nt) (Figure 4D, Supplementary Table 5). The 0.5 FRET efficiency of the B state corresponds to a 14 nt distance between Cy3 -and Cy5 in the ruler which agrees with the 15 nt distance in the RNA probe. The 0.3 FRET efficiency of the PU state corresponds to a fluorophore distance of 19 nt in the ruler, indicating the unwinding of approximately 5 base pairs of the stem loop in our RNA probe (Figure 4D). Models derived from structural studies propose that the force resulting from conformational changes between the RecA domains is transmitted to the contact point with the double strand during unwinding (9). Upon ATP binding, RecA2 is shifted by ~0.6 nm towards the RecA1 domain, clamping the nucleotide between the two domains. In this step, the RNA is repositioned by 1 nt in 3'-direction through the binding tunnel resulting in the displacement of Prp43 along the RNA. This could lead to the separation of proximal RNA duplexes by a winching-like mechanism (Boneberg *et al*, 2019; Hamann *et al*, 2019; Semlow *et al*, 2016). Also the hydrolysis of ATP induces a conformational change, the rotation of RecA2 with respect to RecA1 by ~20 degrees (He *et al*, 2010; Tauchert *et al*, 2017; Hamann *et al*, 2019). Hence, our smFRET data show that binding of a single ATP and potentially its hydrolysis is sufficient for the melting of about 5 bp by Prp43, but further unwinding requires continuous ATP hydrolysis at physiological rate.

Pfa1(gp) accelerates transitions between closed and open RecA domain conformation during ATP turnover, facilitating processive translocation

Knowing that continuous turnover is necessary for Prp43 to unwind more than the initial base pairs of an RNA duplex, we investigated the conformational dynamics of Cy3-/Cy5-labeled Prp43 at high ATP concentrations (Figure 5). In absence of Pfa1(gp) or ssRNA, we mainly observed stable signals corresponding to the C state, which were earlier observed for Prp43(apo) and Prp43—ADP (Figure 1C and 1E, Figure 5A, Supplementary Table 1). However, a small part of the population (13%) showed the O state, which was so far only observed in the presence of Pfa1(gp) or ssRNA (Figure 1D and F-H). Few traces (4%) showed transitions between C and O state. Considering that domain opening occurs concomitantly to ADP release (Figure 2), and domain closures results from nucleotide binding, slow dynamics is expected given that the intrinsic ATP hydrolysis rate of the helicase is very low (k_{cat}=0.01 s^{-1}, Figure 5C). To examine the influence of Pfa1(gp) and ssRNA on the motility of the RecA domains, we first monitored smFRET of the Prp43-Pfa1(gp) complex in absence of ssRNA. The complex showed increased levels of the O state (35%) and 27% of molecules toggled between C and O, which is a significant increase compared to unstimulated Prp43 and agrees with elevated rates of ATP turnover (k_{cat}=0.4 s^{-1}, Figure 5C, Supplementary Figure S5A and S5B, Supplementary Table

3). Incubation of the Prp43-Pfa1(gp) complex with ssRNA further stimulated the rate of ATP hydrolysis (k_{cat}=9.6s^{-1}, Figure 5C). The distribution of states did not change, but the activation of Prp43 increased further, which becomes apparent by higher exchange rates between C and O states ($k_{C\rightarrow O}$=1.44 s^{-1} ,$k_{O\rightarrow C}$=1.92 s^{-1}, Figure 5B, Supplementary Table 4). The rate of ATP turnover determined in steady-state kinetic experiments is somewhat faster than the rate of conformational transitions observed by smFRET, presumably due to reduction of activity caused by the immobilization process. Still, $k_{C\rightarrow O}$ is faster than the previously determined drop-off rate from the RNA (k_{off}=0.45 s^{-1}, Figure 3E and 3F, Supplementary Table 4), indicating that the helicases converts fast enough from the weak (ADP) to the strong (nt-free) RNA binding state to complete several conformational cycles before detachment.

Motility cycle of the Prp43-Pfa1(gp) complex

Our data show how Pfa1(gp) promotes transitions between opening and closing of RecA domains of Prp43 during ATP turnover resulting in processive translocation and unwinding of RNA (Figure 6). The Prp43-Pfa1(gp) complex strongly binds ssRNA, ATP binding leads to closure of the RecA domains. During the transition from the open to the closed state Prp43 and other DEAH-box helicases translocate with a step size of one RNA nucleotide per hydrolyzed ATP in 3’ direction along ssRNA (10). In proximity of duplex RNA, the force exerted during domain closure and concomitant translocation can separate double strands of approximately 5 bp. The hydrolysis of ATP induces the opening of the RecA domains, which precedes the release of the bound ADP. While ADP is still bound, the affinity to the RNA is low, because RecA2 has little contact with the substrate (10). In this state, the helicase is susceptible to drop-off. Pfa1(gp) facilitates domain opening, thus accelerating ADP release and the transition back to tight RNA binding. Assuming that the conformational transitions are strictly coupled to the hydrolysis of ATP, the ATP turnover rate (k_{cat}(scPrp43)=5s^{-1}, Supplementary Figure 6) corresponds to the rate of translocation (5 nt/s). Dissociation of the Prp43-Pfa1(gp) complex is much slower (k_{off}(scPrp43)=0.45s^{-1}), indicating that the helicase undergoes successive conformational cycles before it releases from the substrate and is capable to unwind long RNA duplexes in a processive manner. In the absence of Pfa1(gp), ATP hydrolysis is orders of magnitude slower and the affinity to ssRNA, especially in the ADP bound state, is extremely weak such that dissociation is more probable than translocation.

Discussion

RNA unwinding by Prp43 relies on the transmission of force generated during RecA domain closure to proximal RNA duplexes. The underlying principle resembles the winching mechanism applied by the spliceosomal helicases Prp16, which remodels the spliceosomal C complex after the first splicing reaction to enable exon ligation, and Prp22, which subsequently releases the mature mRNA from the complex (46–49). Both helicases change the structure of their target RNAs inside the spliceosome without direct contact, supposedly by transmitting the translocation force to the interior of the RNP (19). As part of the ILS, Prp43 binds to the 3' end of U6 snRNA disrupting its interactions with U2 snRNA and the intron, which triggers spliceosomal disassembly (50). The location of the helicase binding site is distant from the U6 snRNA-intron duplex, indicating that Prp43 can transmit force over long distances similar to Prp16 or Prp22. Together, these observations suggest that DEAH/RHA helicases destabilize RNA structures when direct access to the target site is restricted (51, 52). In this respect, RNA unwinding is clearly distinct from the mechanism employed by the structurally related DEAD-box helicases, which unwind structured RNAs by local strand displacement in a non-processive manner (53). smFRET studies on the translation initiation factor eIF4A, the spliceosomal helicase Prp5 and the ribosomal helicase YxiN as well as crystallographic studies on several other DEAD-box helicases shed light on their conformational cycling during unwinding: Without ligands, RecA domains are open and close upon the joint binding of ATP and RNA forcing it in a bent conformation incompatible with duplex structure (54–63). Upon ATP hydrolysis and presumably after phosphate release the RecA domains return to the open conformation, releasing the remaining RNA strand and allowing the enzyme to move on to the next substrate (58, 64). Analogous to our observations on Prp43, ATP binding initiates the power stroke required for strand displacement. However, DEAD-box helicases do not translocate along the RNA, therefore only local structures can be unwound. Some DEAD-box helicases perform processive unwinding of the same duplex in complex with auxiliary factors. An example is eIF4A, which unwinds the structured 5' untranslated region (5'UTRs) of mRNAs synergistically with components of the m7-G-cap-binding complexes eIF4F, in particular eIF4G, eIF4B, and eIF4H (65–68). eIF4G stabilizes the half-open RecA domain conformation in eIF4A, while eIF4B and eIF4H improve the overall affinity towards RNA, preventing together the dissociation from the substrate (62, 69). Conformational guidance by co-factors stabilizing structures with high substrate affinity seems to be a common theme modulating the processivity of both types of helicases.

Co-factors play a central role in the regulation of DEAH/RHA helicases, which have low intrinsic substrate specificity and therefore require activation in a specific cellular environment avoiding futile ATP hydrolysis and undesired RNA unwinding. Prp2/DHX16 is a further

DEAH/RHA family member known to interact with a G-patch partner (70–73). In yeast spliceosomes, Spp2 tethers Prp2 to adjacent spliceosomal proteins where it facilitates rearrangements making the branch point accessible for the splicing reaction (70, 71, 74–76). Although the Spp2 gp motif braces the RecA2 and WH domain of Prp2 in a similar manner as observed in the DHX15-NKRF(gp) complex, Spp2 does not alter the conformation of Prp2 and is not capable to stimulate the ATPase activity of Prp2 independently (40, 42, 74, 77). Activation occurs exclusively in the presence of RNA, suggesting that the regulation of Prp2 follows different principles than Prp43. Presumably tethering of Prp2 to the single-stranded intron region of the pre-mRNA is more relevant than the stabilization of an open RecA domain conformation. Comparison of other DEAH/RHA helicases with Prp43/DHX15 and Prp2/DHX16 reveals that they are most likely not regulated by G-patch factors because of potentially disruptive substitutions in the binding regions (42). It remains to be studied, whether the many G-patch factors identified in the human proteome (25) regulate other types of helicases or whether they have different targets.

Regulation of DEAH/RHA helicases by a non-gp cofactor is known for DHX37 (Dhr1 in yeast), which is involved in the maturation of 18S rRNA during biogenesis of the small ribosome subunit (22, 78–80). Similar to Pfa1(gp), UTP14A enhances ATP hydrolysis, RNA binding and unwinding by DHX37 (9, 22) , however, the impact of UTP14A on the structure and motility of DHX37 is unknown. Hence, it remains unclear whether helicase activation follows a similar mechanism as described for Prp43 (Figure 6). Several DEAH/RHA helicases employ specialized domains outside the conserved core to mediate substrate specificity. A prominent example is the N-terminal DSM domain in DHX36, which is indispensable for the recognition and subsequent destabilization of G-quadruplex structures and makes interaction with additional co-factors unnecessary to achieve full helicase activity (13, 81, 82) .

We demonstrate for the first time how conformational cycling of a DEAH/RHA helicase correlates with processive motility along the RNA substrate. Remaining open questions concern how translocation is regulated the cellular context, i.e. as part of spliceosome or in nascent ribosome complexes. *In vivo*, different G-patch factors compete for the interaction with Prp43 (34), and might affect the speed of translocation or the degree of processivity to different extents. Knowledge on the conformational dynamics of other DEAH/RHA helicases will provide important insights into the regulation of fundamental cellular processes. They are often involved in the development of severe human diseases. Examples are DHX29 which promotes cell proliferation and tumorigenesis, DHX36 which functions as a tumor suppressor or the closely related NS3 helicases which are indispensable for viral replication (83–85). Mechanistic knowledge on their regulation will advance the development of novel therapeutic strategies.

Material and Methods

Detailed descriptions of the experimental procedures are provided in *SI Appendix, Materials and Methods.*

All Prp43 and Pfa1 constructs were expressed in *E. coli* Rosetta II (DE3) and first purified by glutathione sepharose affinity chromatography using the N-terminal GST-tag. Prp43 constructs were then subjected to affinity chromatography using their C-terminal His_6- or Strep-tag. Finally, all Prp43 and Pfa1 constructs were purified to homogeneity by size exclusion chromatography.

Prp43 was fluorescently labeled with a mixture of Cy3-maleimeide and Cy5-maleimide using reactive cysteines at position C303 and K170C. The labeled protein was separated from excess dye by Ni-sepharose affinity chromatography.

smFRET experiments were performed using a commercial TIRF microscopy set up (Olympus, Japan) described in (44). PEG-Biotin functionalization of cover slips was performed as described in (44). The concentration of Cy3/Cy5-labeled ctPrp was 1 nM. ctPrp43-Pfa1 complex were formed prior to immobilization incubating 1 µM labeled ctPrp43 with 5 µM ctPfa1(gp). The concentration of the labeled RNA construct was 1 nM. TIRF microscopy buffers, single molecule detection and data analysis were adapted from protocols described in (44).

ATP turnover by Prp43 was monitored using a coupled enzymatic assay following nicotinamide adenine dinucleotide (NADH) absorption. Binding of Pfa1(gp) to Prp43 was monitored by ITC with a MicroCal VP-ITC (Malvern Panalytical) and the RNA binding of Prp43 was measured by fluorescence polarization spectroscopy using a VICTOR Nivo Multimode Microplate Reader (PerkinElmer). ADP release from Prp43 was followed through FRET between Prp43's tryptophan residues and mant-labeled ADP using a stopped-flow apparatus (Applied Photophysics).

Acknowledgments and funding sources

We thank Prof. M. Rodnina for access to the TIRF microscopy facility in her department. We thank Prof. K. Tittman for access to the stopped-flow set-up in his department and Dr. F. Rabe von Pappenheim for instruction to the instrument. We thank N. Nitschke for pilot studies on the project. We thank Dr. P. Neumann and Dr. A. Dickmanns for critically reading of the manuscript. The work was supported by the German research council (Deutsche Forschungsgemeinschaft) with project grants to R.F. (SFB860, project A1) and S.A. (SFB860, project A15).

References

1. A. M. Pyle, Translocation and unwinding mechanisms of RNA and DNA helicases. *Annu. Rev. Biophys.* **37**, 317–336 (2008).
2. M. E. Fairman-Williams, U. P. Guenther, E. Jankowsky, SF1 and SF2 helicases: Family matters. *Curr. Opin. Struct. Biol.* **20**, 313–324 (2010).
3. J. M. Caruthers, D. B. McKay, Helicase structure and mechanism. *Curr. Opin. Struct. Biol.* **12**, 123–133 (2002).
4. O. Cordin, J. Banroques, N. K. Tanner, P. Linder, The DEAD-box protein family of RNA helicases. *Gene* **367**, 17–37 (2006).
5. Y. He, G. R. Andersen, K. H. Nielsen, Structural basis for the function of DEAH helicases. *EMBO Rep.* **11**, 180–186 (2010).
6. H. Walbott, *et al.*, Prp43p contains a processive helicase structural architecture with a specific regulatory domain. *EMBO J.* **29**, 2194–2204 (2010).
7. M. J. Tauchert, J. B. Fourmann, R. Lührmann, R. Ficner, Structural insights into the mechanism of the DEAH-box RNA helicase Prp43. *Elife* **6**, 1–25 (2017).
8. Y. He, J. P. Staley, G. R. Andersen, K. H. Nielsen, Structure of the DEAH/RHA ATPase Prp43p bound to RNA implicates a pair of hairpins and motif Va in translocation along RNA. *Rna* **23**, 1110–1124 (2017).
9. F. M. Boneberg, *et al.*, Molecular mechanism of the RNA helicase DHX37 and its activation by UTP14A in ribosome biogenesis. *RNA* **25**, 685–701 (2019).
10. F. Hamann, M. Enders, R. Ficner, Structural basis for RNA translocation by DEAH-box ATPases. *Nucleic Acids Res.* **47**, 4349–4362 (2019).
11. F. Hamann, *et al.*, The structure of Prp2 bound to RNA and ADP-BeF3-reveals structural features important for RNA unwinding by DEAH-box ATPases. *Acta Crystallogr. Sect. D Struct. Biol.* **77**, 496–509 (2021).
12. L. M. Grass, *et al.*, Large-scale ratcheting in a bacterial DEAH/RHA-type RNA helicase that modulates antibiotics susceptibility. *Proc Natl Acad Sci U S A* **118**, e2100370118 (2021).
13. M. C. Chen, *et al.*, Structural basis of G-quadruplex unfolding by the DEAH/RHA helicase DHX36. *Nature* **558**, 465–483 (2018).
14. J. Robert-Paganin, *et al.*, Functional link between DEAH/RHA helicase Prp43 activation and ATP base binding. *Nucleic Acids Res.* **45**, 1539–1552 (2017).
15. S. Liu, *et al.*, Structure of the yeast spliceosomal postcatalytic P complex. *Science (80-.).* **358**, 1278–1283 (2017).
16. R. Wan, C. Yan, R. Bai, J. Lei, Y. Shi, Structure of an Intron Lariat Spliceosome from Saccharomyces cerevisiae. *Cell* **171**, 120-132.e12 (2017).
17. S. M. Fica, *et al.*, Structure of a spliceosome remodelled for exon ligation. *Nature* **542**, 377–380 (2017).
18. C. Yan, R. Wan, R. Bai, G. Huang, Y. Shi, Structure of a yeast activated spliceosome at 3.4 Å resolution. *Science (80-.).* **353**, 904–911 (2016).
19. D. R. Semlow, M. R. Blanco, N. G. Walter, J. P. Staley, Spliceosomal DEAH-Box ATPases Remodel Pre-mRNA to Activate Alternative Splice Sites. *Cell* **164**, 985–998 (2016).

20. Y. J. Tseng, *et al.*, The RNA helicase DHX36–G4R1 modulates C9orf72 GGGGCC hexanucleotide repeat–associated translation. *J. Biol. Chem.* **297**, 1–15 (2021).
21. K. E. Sloan, M. T. Bohnsack, Unravelling the Mechanisms of RNA Helicase Regulation. *Trends Biochem. Sci.* **43**, 237–250 (2018).
22. J. Zhu, X. Liu, M. Anjos, C. C. Correll, A. W. Johnson, Utp14 Recruits and Activates the RNA Helicase Dhr1 To Undock U3 snoRNA from the Preribosome. *Mol. Cell. Biol.* **36**, 965–978 (2016).
23. R.-T. Tsai, *et al.*, 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 (2007).
24. L. Aravind, E. V Koonin, G-patch: a new conserved domain in eukaryotic RNA-processing proteins and type D retroviral polyproteins. *Trends Biochem. Sci.* **24**, 342–344 (1999).
25. K. E. Bohnsack, R. Ficner, M. T. Bohnsack, S. Jonas, Regulation of DEAH-box RNA helicases by G-patch proteins. *Biol. Chem.* **402**, 561–579 (2021).
26. N. Tanaka, A. Aronova, B. Schwer, Ntr1 activates the Prp43 helicase to trigger release of lariat-intron from the spliceosome. *Genes Dev.* **21**, 2312–2325 (2007).
27. J. B. Fourmann, *et al.*, The target of the DEAH-box NTP triphosphatase Prp43 in saccharomyces cerevisiae spliceosomes is the U2 snRNP-intron interaction. *Elife* **5**, e15564 (2016).
28. R. M. Mayas, H. Maita, D. R. Semlow, J. P. Staley, Spliceosome discards intermediates via the DEAH box ATPase Prp43p. *Proc. Natl. Acad. Sci.* **107**, 10020–10025 (2010).
29. J.-B. B. Fourmann, M. J. Tauchert, R. Ficner, P. Fabrizio, R. Lührmann, Regulation of Prp43-mediated disassembly of spliceosomes by its cofactors Ntr1 and Ntr2. *Nucleic Acids Res.* **45**, 4068–4080 (2017).
30. B. Pertschy, *et al.*, RNA helicase Prp43 and its co-factor Pfa1 promote 20 to 18 S rRNA processing catalyzed by the endonuclease Nob1. *J. Biol. Chem.* **284**, 35079–91 (2009).
31. S. Lebaron, *et al.*, The ATPase and helicase activities of Prp43p are stimulated by the G-patch protein Pfa1p during yeast ribosome biogenesis. *EMBO J.* **28**, 3808–19 (2009).
32. Y.-L. L. Chen, *et al.*, The telomerase inhibitor Gno1p/PINX1 activates the helicase Prp43p during ribosome biogenesis. *Nucleic Acids Res.* **42**, 7330–7345 (2014).
33. M. T. Bohnsack, *et al.*, Prp43 Bound at Different Sites on the Pre-rRNA Performs Distinct Functions in Ribosome Synthesis. *Mol. Cell* **36**, 583–592 (2009).
34. A. U. Heininger, *et al.*, Protein cofactor competition regulates the action of a multifunctional RNA helicase in different pathways. *RNA Biol.* **13**, 320–330 (2016).
35. D. Toczydlowska-Socha, *et al.*, Human RNA cap1 methyltransferase CMTr1 cooperates with RNA helicase DHX15 to modify RNAs with highly structured 50 termini. *Philos. Trans. R. Soc. B Biol. Sci.* **373**, 1–10 (2018).
36. F. Inesta-Vaquera, *et al.*, DHX15 regulates CMTR1-dependent gene expression and cell proliferation. *Life Sci. Alliance* **1**, e201800092 (2018).
37. H. Christian, R. V. Hofele, H. Urlaub, R. Ficner, Insights into the activation of the helicase Prp43 by biochemical studies and structural mass spectrometry. *Nucleic Acids Res.* **42**, 1162–1179 (2014).

38. I. Memet, C. Doebele, K. E. Sloan, M. T. Bohnsack, The G-patch protein NF-κB-repressing factor mediates the recruitment of the exonuclease XRN2 and activation of the RNA helicase DHX15 in human ribosome biogenesis. *Nucleic Acids Res.* **45**, 5359–5374 (2017).
39. Z. Niu, W. Jin, L. Zhang, X. Li, Tumor suppressor RBM5 directly interacts with the DExD/H-box protein DHX15 and stimulates its helicase activity. *FEBS Lett.* **586**, 977–983 (2012).
40. F. Hamann, *et al.*, Structural analysis of the intrinsically disordered splicing factor Spp2 and its binding to the DEAH-box ATPase Prp2. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 2948–2956 (2020).
41. B. A. Shoemaker, J. J. Portman, P. G. Wolynes, Speeding molecular recognition by using the folding funnel: The fly-casting mechanism. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 8868–8873 (2000).
42. M. K. Studer, L. Ivanovic, M. E. Weber, S. Marti, S. Jonas, Structural basis for DEAH-helicase activation by G-patch proteins. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 7159–7170 (2020).
43. M. J. Tauchert, J. B. Fourmann, H. Christian, R. Lührmann, R. Ficner, Structural and functional analysis of the RNA helicase Prp43 from the thermophilic eukaryote Chaetomium thermophilum. *Acta Crystallogr. Sect. Struct. Biol. Commun.* **72**, 112–120 (2016).
44. S. Adio, *et al.*, Fluctuations between multiple EF-G-induced chimeric tRNA states during translocation on the ribosome. *Nat. Commun.* **6**, 1–11 (2015).
45. A. Martin, S. Schneider, B. Schwer, Prp43 is an essential RNA-dependent ATPase required for release of lariat-intron from the spliceosome. *J. Biol. Chem.* **277**, 17743–50 (2002).
46. C. K. Tseng, H. L. Liu, S. C. Cheng, DEAH-box ATPase Prp16 has dual roles in remodeling of the spliceosome in catalytic steps. *RNA* **17**, 145–154 (2011).
47. T. Ohrt, *et al.*, Molecular dissection of step 2 catalysis of yeast pre-mRNA splicing investigated in a purified system. *RNA* **19**, 902–915 (2013).
48. M. Company, J. Arenas, J. Abelson, Requirement of the RNA helicase-like protein PRP22 for release of messenger RNA from spliceosomes. *Nature* **349**, 487–493 (1991).
49. B. Schwer, C. H. Gross, Prp22, a DExH-box RNA helicase, plays two distinct roles in yeast pre-mRNA splicing. *EMBO J.* **17**, 2086–2094 (1998).
50. R. Toroney, K. H. Nielsen, J. P. Staley, Termination of pre-mRNA splicing requires that the ATPase and RNA unwindase Prp43p acts on the catalytic snRNA U6. *Genes Dev.* **33**, 1–20 (2019).
51. P. Donsbach, D. Klostermeier, Regulation of RNA helicase activity: Principles and examples. *Biol. Chem.* **402**, 529–559 (2021).
52. F. De Bortoli, S. Espinosa, R. Zhao, DEAH-Box RNA Helicases in Pre-mRNA Splicing. *Trends Biochem. Sci.* **46**, 225–238 (2021).
53. Q. Yang, M. Del Campo, A. M. Lambowitz, E. Jankowsky, DEAD-Box Proteins Unwind Duplexes by Local Strand Separation. *Mol. Cell* **28**, 253–263 (2007).
54. J. M. Caruthers, E. R. Johnson, D. B. Mckay, Crystal structure of yeast initiation factor 4A, a DEAD-box RNA helicase. *PNAS* **97**, 13080–13085 (2000).

55. Z. Cheng, J. Coller, R. Parker, H. Song, Crystal structure and functional analysis of DEAD-box protein Dhh1p. *RNA* **11**, 1258–1270 (2005).
56. C. B. F. Andersen, *et al.*, Structure of the Exon Junction Core Complex with a Trapped DEAD-Box ATPase Bound to RNA. *Science (80-.).* **313**, 1968–1972 (2006).
57. B. Theissen, A. R. Karow, J. Rgen Kö Hler ‡, A. Gubaev, D. Klostermeier, Cooperative binding of ATP and RNA induces a closed conformation in a DEAD box RNA helicase. *PNAS* **105**, 548–553 (2008).
58. F. Liu, A. Putnam, E. Jankowsky, ATP hydrolysis is required for DEAD-box protein recycling but not for duplex unwinding. *PNAS* **51**, 20209–20214 (2008).
59. T. Sengoku, O. Nureki, A. Nakamura, S. Kobayashi, S. Yokoyama, Structural Basis for RNA Unwinding by the DEAD-Box Protein Drosophila Vasa. *Cell* **125**, 287–300 (2006).
60. R. Collins, *et al.*, The DEXD/H-box RNA helicase DDX19 is regulated by an α-helical switch. *J. Biol. Chem.* **284**, 10296–10300 (2009).
61. H. Von Moeller, C. Basquin, E. Conti, The mRNA export protein DBP5 binds RNA and the cytoplasmic nucleoporin NUP214 in a mutually exclusive manner. *Nat. Struct. Mol. Biol.* **16**, 247–254 (2009).
62. M. Hilbert, F. Kebbel, A. Gubaev, D. Klostermeier, eIF4G stimulates the activity of the DEAD box protein eIF4A by a conformational guidance mechanism. *Nucleic Acids Res.* **39**, 2260–2270 (2011).
63. D. H. Beier, *et al.*, Dynamics of the DEAD-box ATPase Prp5 RecA-like domains provide a conformational switch during spliceosome assembly. *Nucleic Acids Res.* **47**, 10842–10851 (2019).
64. W. Cao, *et al.*, Mechanism of Mss116 ATPase reveals functional diversity of DEAD-Box proteins. *J. Mol. Biol.* **409**, 399–414 (2011).
65. U. Harms, A. Z. Andreou, A. Gubaev, D. Klostermeier, eIF4B, eIF4G and RNA regulate eIF4A activity in translation initiation by modulating the eIF4A conformational cycle. *Nucleic Acids Res.* **42**, 7911–7922 (2014).
66. A. R. Özeş, K. Feoktistova, B. C. Avanzino, C. S. Fraser, Duplex unwinding and ATPase activities of the DEAD-box helicase eIF4A are coupled by eIF4G and eIF4B. *J. Mol. Biol.* **412**, 674–687 (2011).
67. G. W. Rogers, N. J. Richter, W. F. Lima, W. C. Merrick, Modulation of the Helicase Activity of eIF4A by eIF4B, eIF4H, and eIF4F. *J. Biol. Chem.* **276**, 30914–30922 (2001).
68. J. Brito Querido, *et al.*, Structure of a human 48S translational initiation complex. *Science (80-.).* **369**, 1220–1227 (2020).
69. C. García-García, K. L. Frieda, K. Feoktistova, C. S. Fraser, S. M. Block, Factor-dependent processivity in human eIF4A DEAD-box helicase. *Science (80-.).* **348**, 1486–1488 (2015).
70. J. Roy, K. Kim, J. R. Maddock, J. G. Anthony, J. L. Woolford, The final stages of spliceosome maturation require Spp2p that can interact with the DEAH box protein Prp2p and promote step 1 of splicing. *RNA* **1**, 375–390 (1995).
71. E. J. Silverman, *et al.*, Interaction between a G-patch protein and a spliceosomal DEXD/H-box ATPase that is critical for splicing. *Mol. Cell. Biol.* **24**, 10101–10 (2004).
72. A. Hegele, *et al.*, Dynamic Protein-Protein Interaction Wiring of the Human Spliceosome. *Mol. Cell* **45**, 567–580 (2012).

73. S. Zang, *et al.*, GPKOW is essential for pre-mRNA splicing in vitro and suppresses splicing defect caused by dominant-negative DHX16 mutation in vivo. *Biosci. Rep.* **34**, 841–850 (2014).
74. R. Bai, *et al.*, Mechanism of spliceosome remodeling by the ATPase/helicase Prp2 and its coactivator Spp2. *Science (80-.).* **371**, eabe8863 (2021).
75. Z. Warkocki, *et al.*, Reconstitution of both steps of Saccharomyces cerevisiae splicing with purified spliceosomal components. *Nat. Struct. Mol. Biol.* **16**, 1237–1243 (2009).
76. P. Bao, C. Höbartner, K. Hartmuth, R. Lührmann, Yeast Prp2 liberates the 5 ′ splice site and the branch site adenosine for catalysis of pre-mRNA splicing. *RNA* **23**, 1770–1779 (2017).
77. Z. Warkocki, *et al.*, The G-patch protein Spp2 couples the spliceosome-stimulated ATpase activity of the deah-box protein Prp2 to catalytic activation of the spliceosome. *Genes Dev.* **29**, 94–107 (2015).
78. A. Colley, J. D. Beggs, D. Tollervey, D. L. J. Lafontaine, Dhr1p, a Putative DEAH-Box RNA Helicase, Is Associated with the Box CD snoRNP U3. *Mol. Cell. Biol.* **20**, 7238–7246 (2000).
79. R. Sardana, *et al.*, The DEAH-box Helicase Dhr1 Dissociates U3 from the Pre-rRNA to Promote Formation of the Central Pseudoknot. *PLoS Biol.* **13**, e1002083 (2015).
80. P. Choudhury, P. Hackert, I. Memet, K. E. Sloan, M. T. Bohnsack, The human RNA helicase DHX37 is required for release of the U3 snoRNP from pre-ribosomal particles. *RNA Biol.* **16**, 54–68 (2019).
81. S. D. Creacy, *et al.*, 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.* **283**, 34626–34634 (2008).
82. R. Tippana, M. C. Chen, N. A. Demeshkina, A. R. Ferré-D'Amaré, S. Myong, RNA G-quadruplex is resolved by repetitive and ATP-dependent mechanism of DHX36. *Nat. Commun.* **10**, 1–10 (2019).
83. A. Parsyan, *et al.*, The helicase protein DHX29 promotes translation initiation, cell proliferation, and tumorigenesis. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 22217–22222 (2009).
84. Y. Zeng, *et al.*, Identification of DHX36 as a tumour suppressor through modulating the activities of the stress-associated proteins and cyclin-dependent kinases in breast cancer. *Am J Cancer Res* **10**, 4211–4233 (2020).
85. A. M. I. Lam, D. N. Frick, Hepatitis C Virus Subgenomic Replicon Requires an Active NS3 RNA Helicase. *J. Virol.* **80**, 404–411 (2006).

Figures and Legends

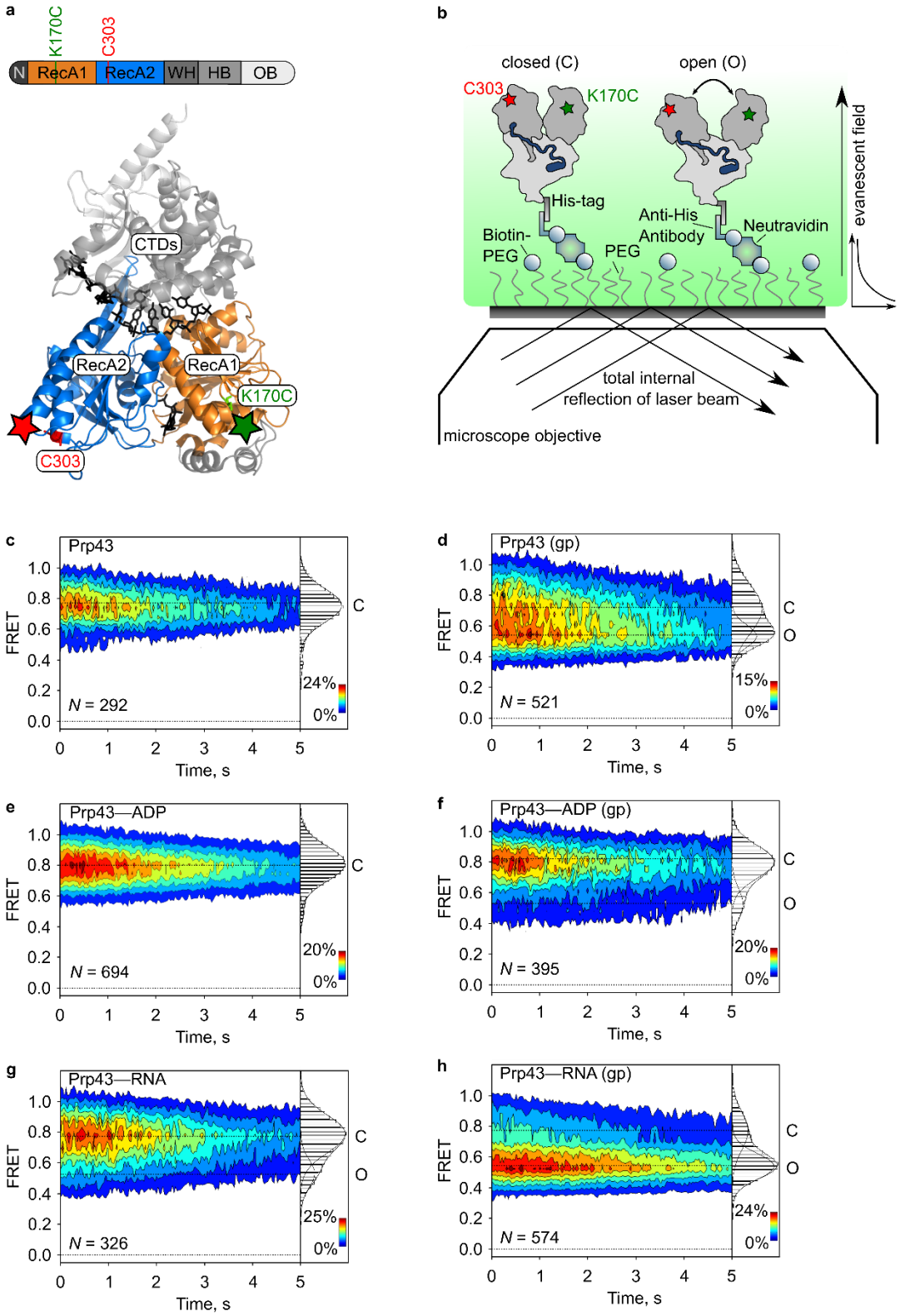
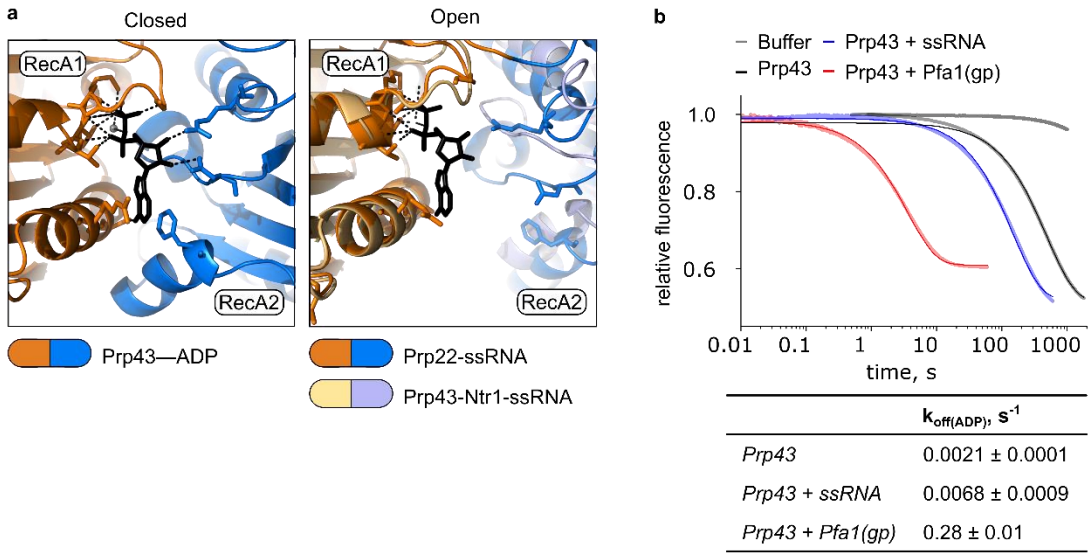
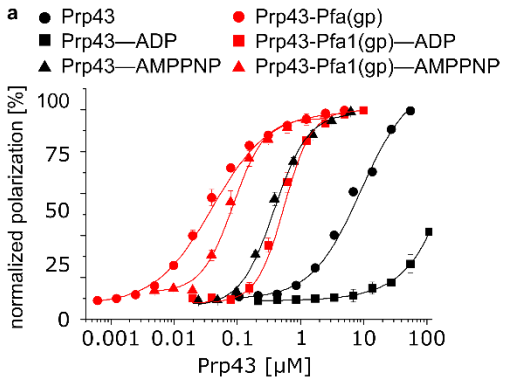


Figure 1: Pfa1(gp) promotes opening of RecA domains in Prp43.
a Domain organization and structure of ctPrp43 (residues 61-764) with U_7-RNA and ADP-BeF_3^- (PDB 5lta). Remainders of the N-terminal domain (residues 61-96) and the C-terminal domains (winged-helix: residues 459-526, helix-bundle: residues 527-640, and OB-fold: residues 641-764) are shown in grey shades, RecA1 (residues 97-273) and RecA2 (residues 274-458) domains are shown in orange and blue, respectively. RNA and ADP-BeF_3^- are shown in black. Cy3- and Cy5-label positions at K170C and C303 are indicated as red and green stars.
b TIRF microscopy experiment scheme monitoring RecA domain movement in Prp43. The C-terminal His-tag of Prp43 interact with a biotinylated antibody for attachment on neutravidin-functionalized coated coverslips. Distance changes between Cy3 and Cy5 report on the conformation (closed or open) of the helicase core.
c-h Contour plot and 2D histogram showing the distribution of FRET values (mean±sd) of (**c**) Prp43 in the apo state (0.77±0.03), (**d**) Prp43-Pfa1(gp) complex (0.75±0.05 and 0.56±0.04), (**e**) Prp43—ADP (0.81±0.02), (**f**) Prp43—ADP-Pfa1(gp) complex (0.82±0.03 and 0.54±0.03), (**g**) Prp43—RNA (0.79±0.02 and 0.53±0.04) and (**h**) Prp43—RNA-Pfa1(gp) complex (0.77±0.02 and 0.54±0.01) corresponding to closed (C) or open (O) conformations. Normalization was performed here and in all further FRET distributions by the number of FRET counts. *N* is here and all further plots the number of individual traces. Data are from N=3 independent experiments.

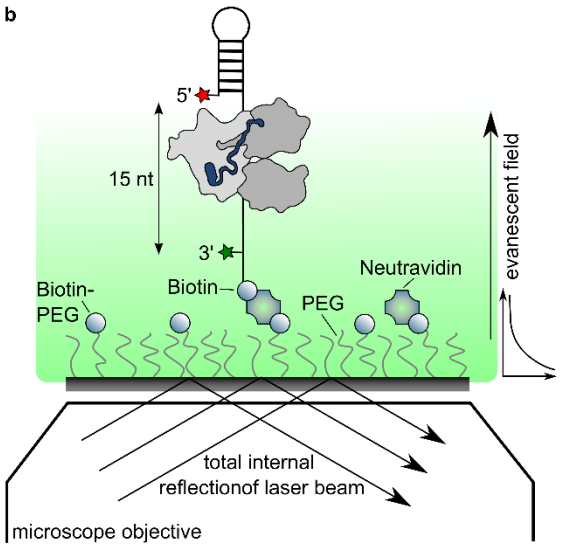


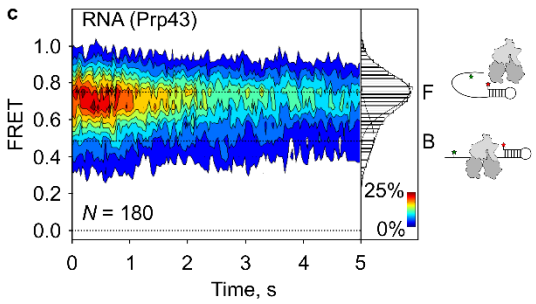
	$k_{off(ADP)}$, s^{-1}
Prp43	0.0021 ± 0.0001
Prp43 + ssRNA	0.0068 ± 0.0009
Prp43 + Pfa1(gp)	0.28 ± 0.01

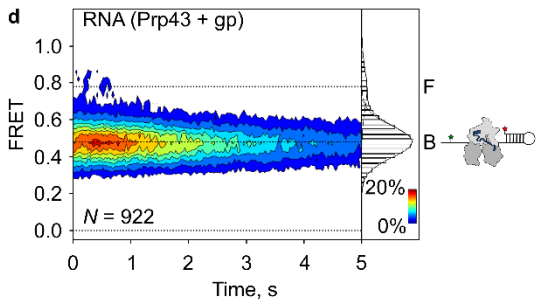
Figure 2: The opening of RecA domains facilitates ADP release.
a Modeling of ADP binding by Prp43 with closed and open RecA domains. Left panel: Nucleotide binding site of ctPrp43—ADP in the closed state ((43); PDB 5d0u), where ADP is sandwiched between RecA1 (orange) and RecA2 (blue). The adenine moiety (black) is bound by π- π electron stacking with F360 in the RecA2 and R162 in the RecA1 domain. The ribose is bound by hydrogen bonding with D391 and R435 in Rec2 and phosphates hydrogen bond with G122, G124, K125, T126 and T127 in RecA1. Right panel: Overlay of the nucleotide binding site of scPrp43 as part of the intron-lariat spliceosome ((16); PDB: 5y88, pale orange and light blue) and the closely related ctPrp22 bound to U_{12}-RNA ((10); PDB: 6i3p, orange and blue) with open RecA domains. The position of ADP was derived by alignment of RecA1 domains in open and closed state structures, assuming that contacts with RecA1 are maintained during domain opening. RecA2 is too far distant to maintain the interactions with ADP. Due to limited resolution, the scPrp43 structure does not contain side chains. Therefore, the residues that interact with ADP in the closed state structure are only shown for ctPrp22.
b Release of mant-ADP from Prp43 monitored by stopped-flow. Averages of fluorescence time traces, normalized by fluorescence intensity, are shown for Prp43, Prp43—ssRNA and Prp43-Pfa1(gp). Table indicates mant-ADP dissociation rates ($k_{off(ADP)}$±s.d.) derived from at least N=3 independent measurements.

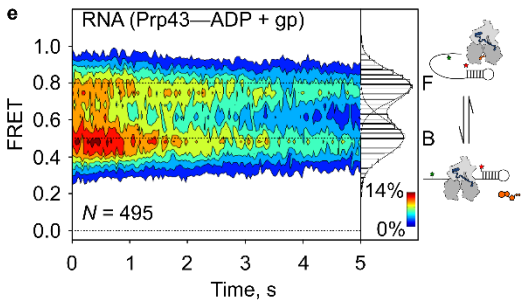


	K_D(ssRNA), μM (±sd)
Prp43	7±0.3
Prp43—ADP	> 100
Prp43—AMPPNP	0.4±0.02
Prp43+Pfa1(gp)	0.03±0.003
Prp43—ADP+Pfa1(gp)	0.5±0.04
Prp43—AMPPNP+Pfa1(gp)	0.08±0.005









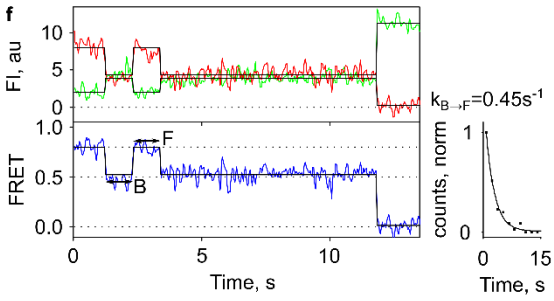


Figure 3: Pfa1(gp) enhances RNA binding of Prp43.
a scPrp43 binding affinity to ssRNA determined by fluorescence polarization spectroscopy in the absence (black) or presence (red) of Pfa1(gp) in the nt-free (circles), ADP- (squares) or AMPPNP- (triangles) state. Shown are mean values; error bars correspond to the sd derived from N=3 independent measurements. Normalization was performed by polarization values. The table indicates affinity constants ($K_{D(ssRNA)}$±sd). For Prp43—ADP less than 50% polarization was reached, indicated K_D is an estimate of the lower limit.
b TIRF microscopy experiment scheme to monitor single-strand RNA binding and duplex unwinding by Prp43. The 3'-end of the RNA probe is biotinylated for attachment on neutravidin-functionalized coated coverslips. Distance changes between Cy3 (position 76, red star) and Cy5 (position 1, green star) report on conformational changes of the probe related to the binding of Prp43 and the unwinding of the stem loop.
c-e Contour plots and 2D histograms showing the distribution of FRET values (mean±sd, derived from N=3 independent data sets) of the RNA probe in the presence of (**c**) Prp43(apo) (0.79 ± 0.01 and 0.51 ± 0.02), (**d**) Prp43-Pfa1(gp) (0.78 ± 0.03 and 0.48 ± 0.02), or (**e)** Prp43—ADP-Pfa1(gp) (0.80 ± 0.02 and 0.51 ± 0.02). Cartoons indicate the state of the RNA (free (F) or bound (B)) by Prp43).
f Representative time traces of Cy3- (green) and Cy5- (red) fluorescence intensity (FI) and FRET (blue, bottom plots) corresponding to transitions between the B and the F state. Solid lines represent the Hidden-Markov fit of the traces. The distribution of dwell times of the B state, normalized by the number of FRET counts, was fitted by an exponential function to determine the drop-off rate of Prp43 from the RNA ($k_{B\rightarrow F}$=0.45s^{-1}). n=76 transitions were included in the analysis.

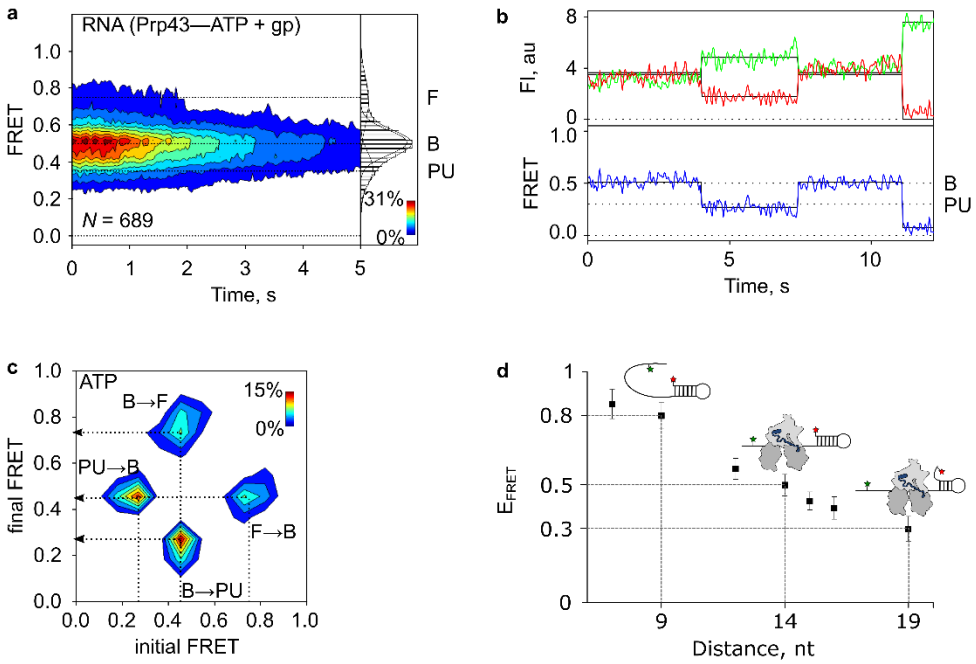
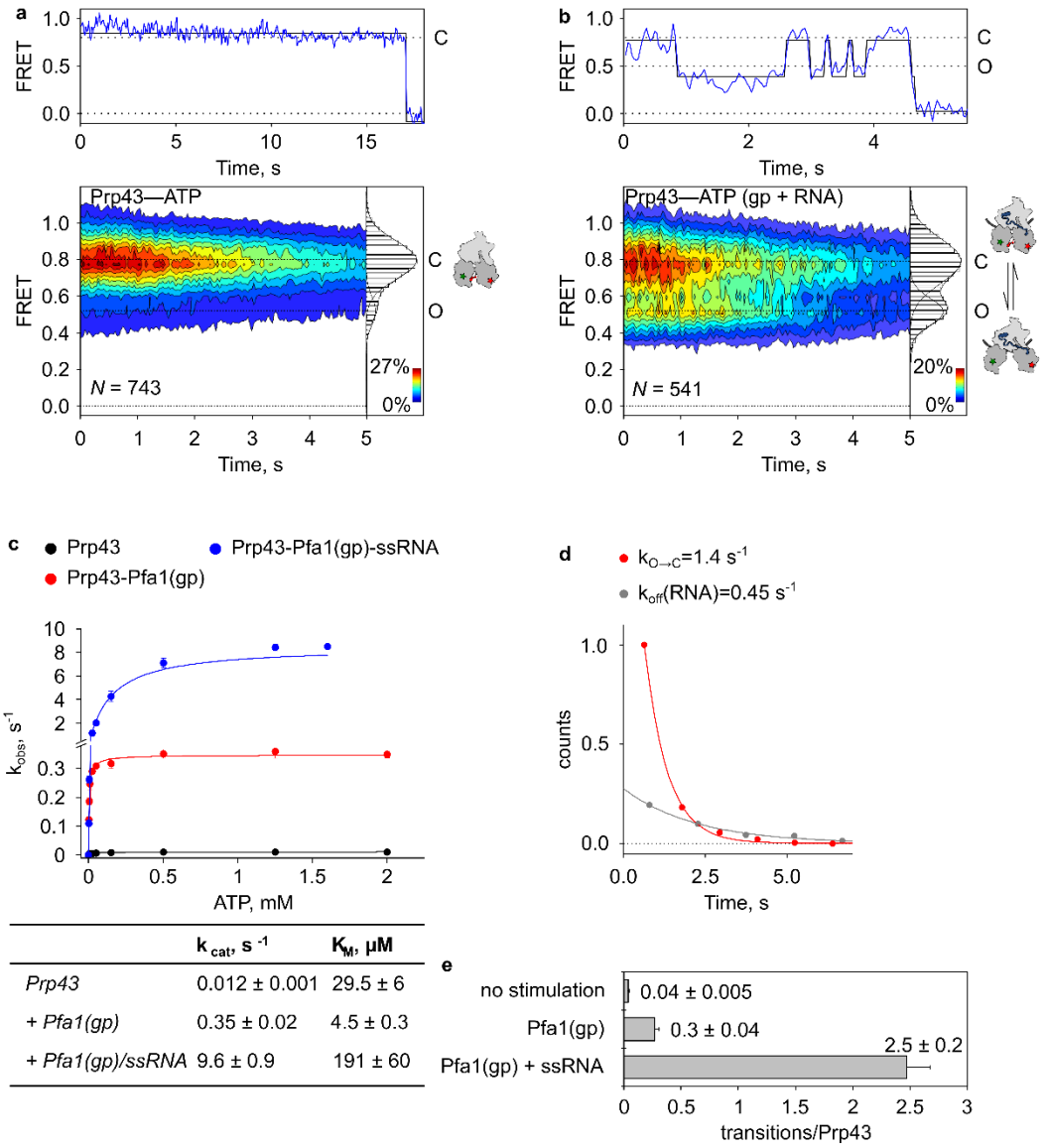


Figure 4: Duplex RNA unwinding by Prp43-Pfa1(gp) complex upon ATP binding.
a Contour plot and 2D histogram showing the distribution of FRET values (mean±sd, derived from N=3 independent data sets) of the RNA probe in the presence of the Prp43-Pfa1(gp) complex at 2 µM ATP (0.76 ± 0.02 (F), 0.51 ± 0.02 (B) and 0.35 ± 0.05 (partially unwound, PU). **b** Representative time trace of Cy3 (green) and Cy5 (red) FI (top), and FRET (blue, bottom) of the RNA probe in the presence of the Prp43-Pfa1(gp) complex and 2 µM ATP. Solid lines represent the Hidden-Markov fit.
c Transition density plot visualizing the frequency of transitions between B and F (B→F, F→B) or B and PU (B→PU, PU→B) states.
d FRET states (mean mean±sd, derived from N=3 independent data sets) of dsDNA oligos with Cy3- and Cy5-labels at defined nucleotide distances on opposing strands. Cartoons show conformations of the RNA probe fitting best within the distance range.



	k_{cat}, s^{-1}	K_M, μM
Prp43	0.012 ± 0.001	29.5 ± 6
+ Pfa1(gp)	0.35 ± 0.02	4.5 ± 0.3
+ Pfa1(gp)/ssRNA	9.6 ± 0.9	191 ± 60

Figure 5: Pfa1(gp) and ssRNA accelerate conformational cycling of RecA domains during ATP turnover.

a-b Representative FRET trace (above), contour plots and 2D histograms (below) showing the distribution of FRET values (mean±sd, derived from N=3 independent data sets) of (**a**) Cy3-/Cy5-labeld Prp43—ATP (0.80±0.02 and 0.53±0.02) or (**b**) Cy3-/Cy5-labeld Prp43—ATP in complex with Pfa1(gp) and in presence of ssRNA (0.80±0.01 and 0.52±0.01).

c Michealis-Menten titration and table with ATP turnover numbers (k_{cat}) and dissociation constants (K_M) showing the stimulation of Prp43 ATPase activity by Pfa1(gp) and ssRNA

d Rates of RecA domain closure ($k_{O\rightarrow C}$, red) and the RNA drop-off rate (k_{off}(RNA), grey) of the Prp43-Pfa1(gp) complex on ssRNA.

e Transitions per molecule Prp43 determined by dividing the number of transitions between C and O state by the number of individual traces (*N*). In the case of stimulation by Pfa1(gp) and ssRNA only detectably activated molecules, i.e. traces that showed at least one transition, were included.

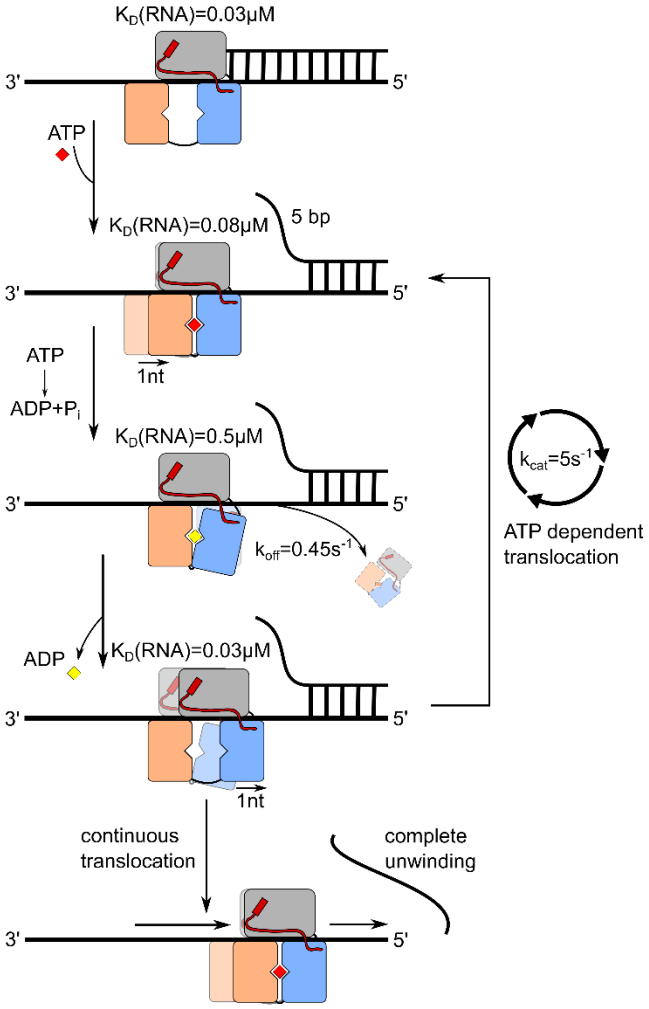


Figure 6. Model of the motility cycle of the Prp43-Pfa1(gp) complex.
Pfa1(gp) is depicted in red, the C-terminal domains of Prp43 are colored in grey, and RecA1 and RecA2 are shown in orange and blue, respectively. The model describes how the ATP dependent conformational cycling of the Prp43-Pfa1(gp) complex induces weak and strong RNA binding states leading to processive translocation and unwinding of RNA. Prp43 affinity constants for ssRNA are derived from fluorescence polarization experiments (Figure 2A), the Prp43 drop-off rate corresponds to $k_{B \rightarrow F}$ of the RNA probe (Figure 3E, F) and the ATP turnover rate was determined by Michaelis-Menten titration (Supplementary Figure 6).

Supplementary Information for

Regulation of the DEAH/RHA helicase Prp43 by the G-patch factor Pfa1

Marieke Enders[1], Ralf Ficner[1], Sarah Adio[1]*

Department of Molecular Structural Biology, Institute of Microbiology and Genetics, Georg-August-University Göttingen, Justus-von-Liebig-Weg 11, D-37077 Göttingen, Germany

*Sarah Adio

Email: sarah.adio@uni-goettingen.de

This PDF file includes:

Supplementary Information Text

Figures S1 to S6

Tables S1 to S5

SI References

Supplementary Information text

Materials and Methods

Protein expression and purification

The homologues of Prp43 and Pfa1 from *C. thermophilum (ctPrp43, ctPfa1) are annotated as* 'hypothetical protein CTHT_0005780' and 'hypothetical protein CTHT_0048220'. Full-length ctPrp43 and the gp motif of ctPfa1 (residues 662-742) were cloned from genomic DNA of *C. thermophilum var. thermophilum* DSM 1495 into pGEX-6P-1 (Tauchert *et al*, 2016). For ctPrp43, a C-terminal His_6-tag was added to the GST-fusion protein by site-directed mutagenesis. *S. cerevisiae* Pfa1(gp) (residues 701-767)) was cloned into pGEX-6P-1 (Fourmann *et al*, 2017). Prp43 from *S. cerevisiae* was amplified from genomic DNA of *S. cerevisiae S288c* using primers CCGTCGACATGGGTTCCAAAAGAAGATTC and GGGCGGCCGCCTA TTTCTTGGAGTGCTTAC and cloned into pGEX-6P-1 using the SalI and NotI restriction sites. All dispensable bases between the PreScission Protease cleavage site and the starting methionine of scPrp43 were deleted and a Strep-tag was added at the C-terminus of the GST-fusion protein *by* site-directed mutagenesis.

scPrp43, ctPrp43 and ctPfa(gp) constructs were recombinantly expressed in Rosetta 2 (DE3) cells using an autoinduction protocol adapted from Studier (Studier, 2005). Expression cultures were grown in ZY medium, supplemented with 25 mM K_2HPO_4, 25 mM NaH_2PO_4, 50 mM NH_4Cl, 5 mM Na_2SO_4, 0.5 % (v/v) glycerol, 0.05 % (w/v) glucose, 0.2 % (w/v) α-lactose, 2 mM $MgSO_4$ and a mixture of trace metals (Studier, 2005) at 37°C at 220 rpm until an OD600 of 0.4 was reached. Then, the temperature was reduced to 16°C and the cultures were incubated at 220 rpm for ~60h.

scPfa1(gp) was recombinantly expressed in BL21 (DE3) cells. Expression cultures were grown in 2YT medium at 37°C at 220 rpm until an OD600 of 0.8 was reached. The expression was induced by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM, and the cultures were incubated at 16°C and 220 rpm overnight.

Harvested cells were disrupted using a microfluidizer (Microfluidics) in 50 mM Tris/HCl (pH 7.5), 500 mM NaCl, 5% (v/v) glycerol and 10 mM ethylenediaminetetraacetic acid (EDTA), soluble proteins were isolated by ultracentrifugation at 35000x*g* for 30min. The clarified lysate was loaded onto Glutathione Sepharose 4B (Cytiva) at 20°C. Bound nucleic acids were removed by intensive washing with lysate buffer supplemented with 2 M LiCl and the target protein was eluted with 30mM reduced glutathione. The GST-tag was proteolytically cleaved

off with 1:100 (w/w) PreScission protease overnight at 4°C. Strep-tagged Prp43 was loaded onto StrepTactin HP Sepharose (GE Healthcare) in 50 mM Tris/HCl (pH 7.5), 400 mM NaCl, 5% (v/v) glycerol, 2 mM $MgCl_2$, and ctPrp43-His was loaded onto Ni Sepharose HP (GE Healthcare) in 50 mM Tris/HCl (pH 7.5), 400 mM NaCl, 5% (v/v) glycerol, 2 mM $MgCl_2$ and 10 mM imidazole. Elution of target protein was realized with 3 mM D-desthiobiotin (scPrp43-Strep) or 500 mM imidazole (ctPrp43-His). All proteins were purified to homogeneity by size exclusion chromatography (Superdex 200 or Superdex 75, Cytiva for Prp43 and Pfa1 respectively) in 20 mM Tris/HCl (pH 7.5), 200 mM NaCl, 5% (v/v) glycerol, 2 mM $MgCl_2$. For the gp constructs, the buffer contained 400 mM NaCl. The proteins were concentrated to a final concentration of 70-300 µM (Prp43; Amicon Ultra 50K, Millipore) or 150-700 µM (Pfa1(gp); Amicon Ultra 3K, Millipore). scPrp43-E216A was expressed and purified using the same protocol as described for scPrp43.

Fluorescence-Labeling of ctPrp43

To generate two maleimide-reactive fluorescence-labeling sites in ctPrp43, eight native cysteine residues had to be considered. C148, C214, and C377 are buried inside the protein and inaccessible to the coupling group. C303, C323, C441, C508, and C543 are surface exposed and therefore accessible to the fluorescence dye. C303 is located in the RecA2 domain and remained unchanged while the other exposed cysteines were replaced by site directed mutagenesis (C323V, C441A, C508A and C543S). The second labeling site was generated by introducing a cysteine residue at position K170 in the RecA1. The mutant protein was expressed and purified as described above with the addition of 5 mM DTT to the lysis buffer and 1 mM TCEP to the His-trap loading and elution buffers. The purified protein was mixed with Cy3-maleimeide and Cy5-maleimide (Cytiva) dissolved in dimethylsulfoxide at a molar ratio of 1:2:3 (protein:Cy3:Cy5) and incubated for 10 min at 20°C. Excess dye was removed by Ni-sepharose affinity chromatography, labeled protein was eluted in 50 mM Tris/HCl (pH 7.5), 400 mM NaCl, 5% (v/v) glycerol, 2 mM $MgCl_2$, 250 mM imidazole. The labeled protein was dialyzed twice against 50 mM Tris-HCl, pH 7.5, 300 mM KCl, 3 mM $MgCl_2$ using Slide-A Lyzer Dialsis Casette G2 3.5K (Thermo) for 1 h at 4°C. The labeled protein was concentrated to final concentrations between 40 and 70 µM (Amicon Ultra 50K, Millipore).

Sample preparation for TIRF microscopy

Cover slips and objective slides were cleaned by bath sonication in 1 M KOH and exposure to plasma (FEMTO plasma cleaner, Diener Electronic GmbH, Germany). Surfaces were then silanized by sonication in 3.9 mM N1-[3-(trimethoxysilyl) propyl] diethylenetriamine (Sigma-Aldrich) and 1.7 mM acetic acid, and baked for 20 min at 110 °C. PEG/PEG-Biotin functionalization of silanized surfaces was carried out by incubation with 20 mM PEG-NHS

(MeO-PEG-NHS, IRIS Biotech GmbH, PEG1165), 0.2 mM Biotin-PEG-NHS (IRIS Biotech, PEG1057) and 20 mM KOH in 100 mM H_3BO_3 solution for 1 h at room temperature. Excess PEG was removed by washing with H_2O. Cover slips were dried at 60 °C and stored under vacuum. For TIRF experiments, flow chambers were generated by combining objective slides and cover slips with double-sided sticky tape.

The ctPrp43-ctPfa1(gp) complex was formed by incubating 1 µM labeled ctPrp43 with 5 µM ctPfa1(gp) in TIRF buffer A (50 mM Tris-HCl, pH 7.5, 300 mM KCl, 3 mM $MgCl_2$) for 10 min at room temperature. Prior to the experiment, the complex was diluted to 1 nM ctPrp43 with TIRF buffer supplemented with 5 µM ctPfa1(gp). For experiments without ctPfa1(gp), labeled ctPrp43 was diluted to 1 nM with TIRF buffer A. Biotin/PEG-functionalized cover slips were incubated for 5 min at room temperature with TIRF buffer A containing additionally 10 mg ml^{-1} BSA and 1 µM neutravidin (Thermo Scientific). Excess neutravidin was removed by washing the cover slip with the same buffer containing 1 mg ml^{-1} BSA. A biotinylated anti-His antibody (Rabbit monochlonal, Sigma) was applied at 1.25 µg/ml to the cover slip. Excess antibody was removed by washing with TIRF buffer A containing 1 mg ml^{-1} BSA. Labeled ctPrp43 or ctPrp43-ctPfa1(gp) complex was applied to the surface and incubated for 1 min at room temperature. Images were recorded after washing with TIRF buffer C (TIRF buffer A with 2.5 mM protocatechuic acid, 50 nM protocatechuate-3,4-dioxygenase (from *Pseudomonas*), 1 mM trolox (6-hydroxy-2,5,7,8-tetramethylchromane-2-carboxylic acid) and 1 mM methylviologen). To study the influence of ssRNA and ADP, the imaging buffer was supplemented with 50 µg/ml PolyU RNA (Sigma Aldrich) or 100 µM ADP. To observe domain movement during ATP turnover, the imaging buffer was supplemented with 2 mM ATP, 0.1 mg/ml pyruvate kinase and 3 mM phosphoenolpyruvate.

The labeled RNA (G[Cy5]CGCCUACGCCACCAGCUCCGUAGGCGCAGGAGCGCCUACGGAGCU GGUGGCGUAGGCGCAAAAAAAAAAAAAAAU[Cy3]AAAAAAAAAAAAAAAAAAAA-Biotin) was diluted to 1 nM with TIRF buffer B (50 mM Tris-HCl, pH 7.5, 150 mM KCl, 3 mM $MgCl_2$). Biotin-PEG-functionalized cover slips were incubated for 5 min at room temperature with TIRF buffer B containing additionally 10 mg ml^{-1} BSA and 1 µM neutravidin (Thermo Scientific). Excess neutravidin was removed by washing with the same buffer containing 1 mg ml^{-1} BSA. The labeled RNA was applied to the surface and images were recorded after the addition of TIRF buffer C. To observe binding of Prp43 to the RNA, the buffer was supplemented with 5 µM scPrp43 or 0.5 µM scPrp43 and 2.5 µM scPfa1(gp). The influence of nucleotides was studied by adding 2 mM ADP, AMPPNP or ATP with an energy recycling system (see above) to TIRF buffer C.

To relate measured FRET efficiencies to distances on the RNA probe, we used DNA oligos as distance ruler:

Cy5-strand ([CY5]GGACTGCCGCCTGGGGAGCCGCACGACGACACGACAAAG[Biotin]),

Cy3-7 (CGTGTCGTCGTGCGGCTCCCCAGGCGG[Cy3]CAGTCC),

Cy3-9 (CGTGTCGTCGTGCGGCTCCCCAGGC[Cy3]GGCAGTCC),

Cy3-12 (CGTGTCGTCGTGCGGCTCCCCA[Cy3]GGCGGCAGTCC),

Cy3-14 (CGTGTCGTCGTGCGGCTCCC[Cy3]CAGGCGGCAGTCC),

Cy3-15 (CGTGTCGTCGTGCGGCTCC[Cy3]CCAGGCGGCAGTCC),

Cy3-16 (CGTGTCGTCGTGCGGCTC[Cy3]CCCAGGCGGCAGTCC),

Cy3-19 (CGTGTCGTCGTGCGG[Cy3]CTCCCCAGGCGGCAGTCC).

The Cy5-labeled DNA oligo was mixed with each of the Cy3-labeled DNA oligos at a ratio of 1:1.2 (Cy3:Cy5) and annealed by incubating at 95 °C for 5 min and subsequent cooling to room temperature. The annealed oligos were diluted to 500 pM with TIRF buffer B, immobilization and image recording was performed analogue to the labeled RNA probe.

TIRF microscopy

TIRF imaging was performed on an IX 81 inverted microscope using a PLAPON 60 × 1.45 numerical aperture objective (Olympus, Japan). Fluorescence was excited by a 561 nm solid-state laser operated at a power of 25 mW. Images were recorded with an electron multiplying CCD (charge-coupled device) camera (CCD-C9100-13, Hamamatsu, Japan). In FRET experiments, color channels were separated by projecting donor and acceptor emission on different parts of the CCD chip using an image splitter (dual view micro imager DV2, Photometrics, USA), filter specifications HQ 605/40, HQ 680/30 (Chroma Technology). Movies were recorded at a rate of 30 frames per second. The experiments were carried out at 22 °C.

Data analysis

Fluorescence time courses for donor (Cy3) and acceptor (Cy5) were extracted using custom-made Matlab (MathWorks) software as described (Adio *et al*, 2015; Roy *et al*, 2008). A semi-automated algorithm (Matlab) was used to select anti-correlated fluorescence traces (correlation coefficient <0.1) exhibiting characteristic single fluorophore intensities. The bleed-through of Cy3 signal into the Cy5 channel was corrected using an experimentally determined coefficient (∼0.13 in our setup; (Adio *et al*, 2015)). All trajectories were smoothed over three data points and truncated to remove photobleaching and photoblinking events.

Traces with lifetimes of Cy3 or Cy5 less than 20 frames (0.66 s) or with multiple photobleaching steps were excluded from the analysis. The FRET efficiency was defined as the ratio of the measured emission intensities, Cy5/(Cy3+Cy5) (Roy *et al*, 2008). FRET time courses were fitted by Hidden Markov modeling using the vbFRET software package (http://vbfret.sourceforge.net/) (Bronson *et al*, 2009). Models with different number of states were considered for each data set. FRET changes of <0.1 in idealized trajectories were not considered as transitions. Transitions lasting for only one frame were not included in the analysis as well. About 5% of all traces were poorly idealized by Hidden Markov modelling and eliminated from subsequent analysis. Two-dimensional contour plots were generated from time-resolved FRET trajectories. The set of all FRET traces for a given condition was compiled in a histogram, which was fitted to a sum of Gaussian functions using Matlab code (Adio *et al*, 2015). Mean FRET values (mean±sd) and population distribution (p=area under the curve±sd) were calculated from three independent datasets and are summarized in Table 1 and Table 3. Dwell times of different FRET states of fluctuating traces were extracted from idealized trajectories. The dwell time histogram for each transition was fitted to an exponential function, $y=y_0+Ae^{-t/\tau}$. Rates (k) were calculated by taking the inverse of dwell times (τ).

ATPase activity assay

ATP turnover by Prp43 was monitored using a coupled enzymatic assay following nicotinamide adenine dinucleotide (NADH) absorption at 340 nm over time in a VICTOR Nivo Multimode Microplate Reader (PerkinElmer) (Agarwal *et al*, 1978). Triplicate measurements were performed at room temperature in 20 mM Tris/HCl (pH 7.5), 150 mM KCl and 3 mM $MgCl_2$,250 nM NADH, 500 nM phosphoenolpyruvate, 6–8.3 U/ml pyruvate kinase and 9–14 U/ml lactic dehydrogenase. The ATP concentration ranged between 0 to 2 mM. To obtain suitable reaction velocities, Prp43 was used at a concentration of 2 µM (no stimulation), 0.5 µM (stimulation with Pfa1(gp)) or 0.2 µM (stimulation with Pfa1(gp) and RNA). For measurements in the presence of Pfa1(gp) a 5-fold molar excess over Prp43 was used. Measurements in the presence of RNA were conducted with a 5-fold molar excess of an A_{20}-ssRNA (AXOlabs) over Prp43.

The ATP consumption per minute (k_{obs}) was calculated using

$$k_{obs} = \frac{[\frac{\Delta A_{340}}{\Delta t}]}{\varepsilon_{340} \cdot d \cdot c}$$

where $\Delta A_{340}/\Delta t$ is the slope of the NADH decrease, ε_{340} is the extinction coefficient of NADH, d is the optical pathlength and c is the protein concentration. K_M and k_{cat} ($\frac{V_{max}}{[E]}$) were calculated by fitting the experimental data with the Michaelis-Menten equation

$$v = \frac{V_{max} \cdot [S]}{K_M + [S]}$$

where v is the reaction velocity, V_{max} is the maximal velocity of the system [S] is the substrate concentration and [E] is the total Prp43 concentration, using OriginPro 9.1.

Isothermal titration calorimetry (ITC)

The binding affinity of ctPfa1(gp) to ctPrp43 was measured by ITC with a MicroCal VP-ITC (Malvern Panalytical) using a concentration of 4 µM Prp43 in the cell and 53 µM Pfa1(gp) in the syringe. The reaction buffer contained 20 mM Tris/HCl (pH 7.5), 200 mM NaCl, 5% (v/v) glycerol and 2 mM $MgCl_2$. The initial injection of 5 µl was followed by 15 µl injections performed at a speed of 0.5 μls^{-1} with intervals of 360 s between injections. The reference energy was set to 10 $\mu Cals^{-1}$ and the binding was monitored at 25 °C. Stoichiometry of binding and dissociation constant were determined in three independent experiments, data analysis was carried out using MicroCal VP-ITC Analysis software (Malvern Panalytical).

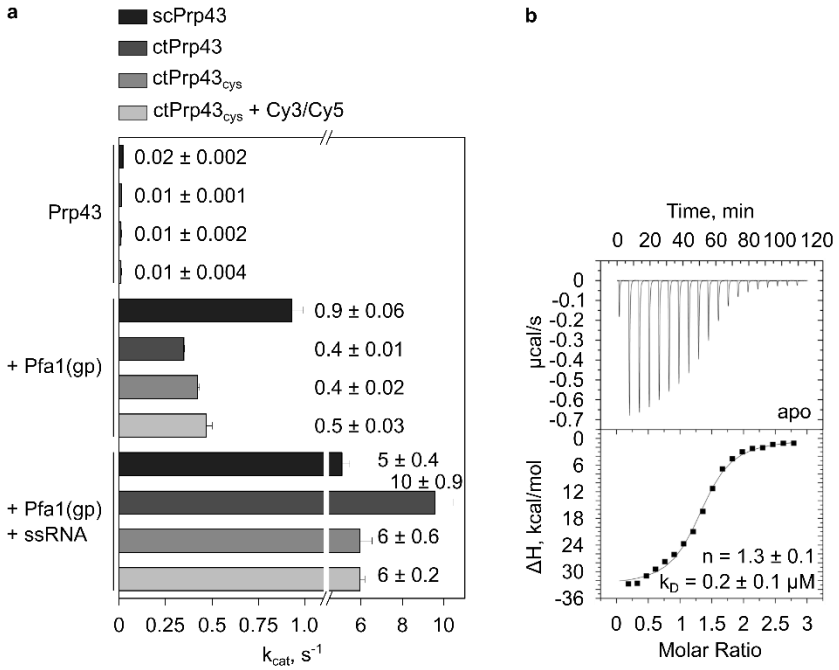
RNA binding assay

The RNA binding of Prp43 was measured by fluorescence polarization spectroscopy using a VICTOR Nivo Multimode Microplate Reader (PerkinElmer). The binding of 6 nM 3' 6-carboxyfluorescein-labeled A_{20}-RNA to up to 100 µM scPrp43 was monitored as triplicates at room temperature in 20 mM Tris/HCl (pH 7.5), 200 mM NaCl, 5% glycerol and 3 mM $MgCl_2$. For measurements in presence of Pfa(gp), the complex was formed by adding a 5-fold molar excess over Prp43 and incubating for 10 min at room temperature. Experiments in presence of ADP or AMPPNP were performed at a constant concentration of 3.5 mM throughout all measurements. The excitation wavelength was 480 nm and the emission was detected at 530 nm for 500 ms. The data were normalized by the maximum of measured polarization and fitted by nonlinear regression using the analysis software OriginPro 9.1.

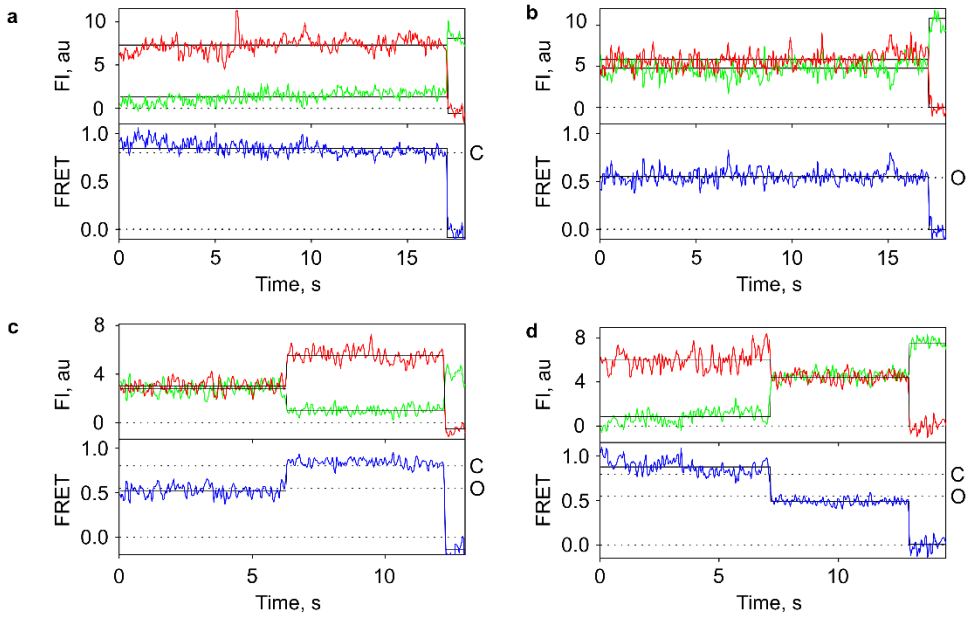
ADP release assay

ADP release from ctPrp43 was followed through FRET, with Prp43's tryptophan residues acting as the FRET donor, and the mant-label on the nucleotide as the FRET acceptor. For the determination of dissociation rates (k_{off}), 10 µM of labelled nucleotide was preincubated for 10min at room temperature with 1 µM ctPrp43 in 20 mM Tris/HCl (pH 7.5), 200 mM NaCl, 5% glycerol and 3 mM $MgCl_2$, in presence or absence of a 5-fold molar excess of ctPfa1(gp) or A_{20} RNA. The sample was rapidly mixed with 1 mM unlabeled "dark" ADP using a Stopped Flow Spectrometer SX20 (Applied Photophysics). The excitation wavelength was set to 280 nm and the emission of the mant-label was detected at 450 nm using a 395 nm filter to eliminate stray light. The reaction was followed for 60 to 1800 s and the time curves were fitted to a one phase exponential function using the analysis software OriginPro 9.1.

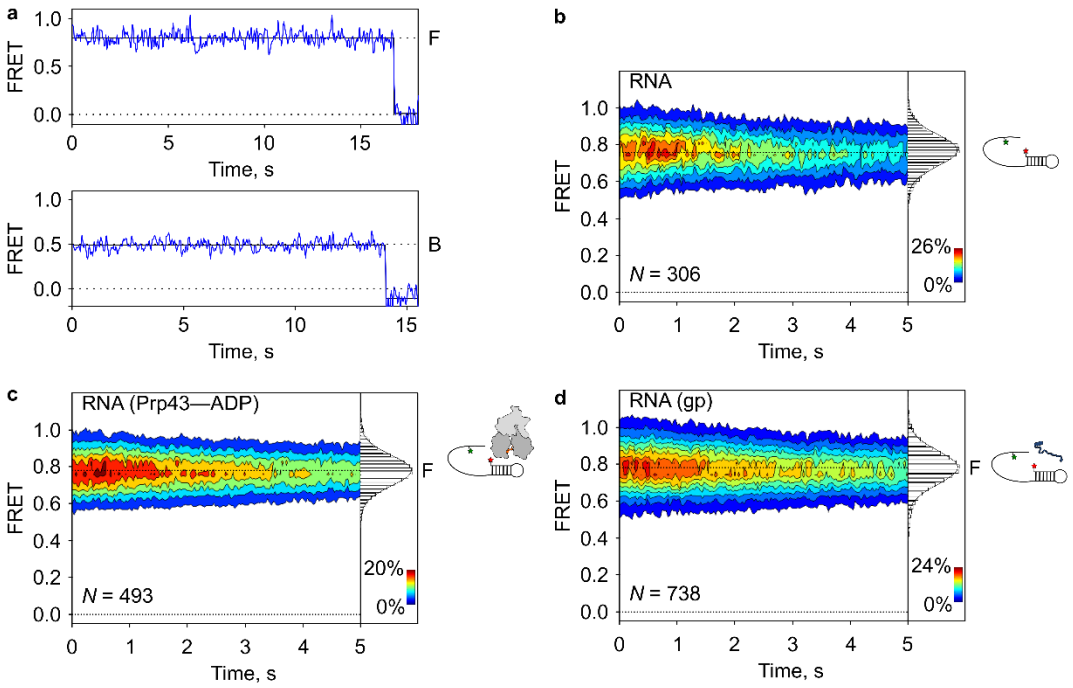
Supplementary Figures



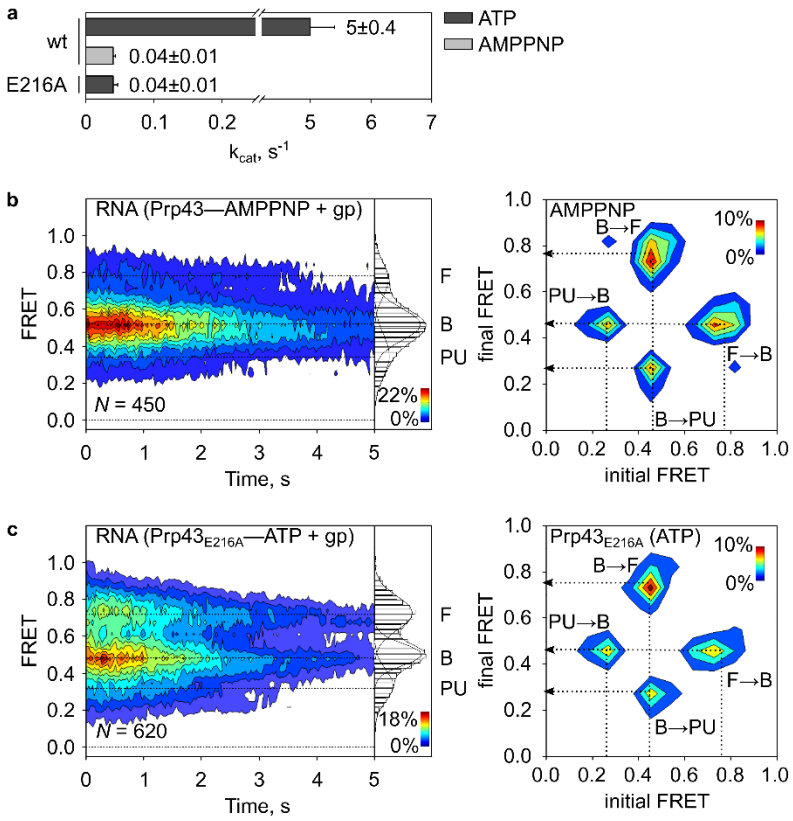
Supplementary Figure 1: ATPase activity and binding affinity of the Prp43-Pfa1(gp) complex.
a Steady-state ATP hydrolysis rates of scPrp43, ctPrp43, mutant ctPrp43 ($Prp43_{Cys}$) with reactive cysteines in RecA1 and RecA2, and Cy3/Cy5 labeled $Prp43_{Cys}$ at 2 mM ATP. Pfa1(gp) and A_{20}-ssRNA were added at saturation. k_{cat}±sd were obtained in N=3 independent measurements.
b Representative binding titration of Prp43 with Pfa1(gp) monitored by isothermal titration calorimetry (ITC). Cell and syringe contained 4 µM Prp43 and 53 µM Pfa1(gp). Stoichiometry of binding (n=1.3±0.1) and dissociation constant (K_D=0.19±0.02 µM) were determined in N=3 independent experiments.



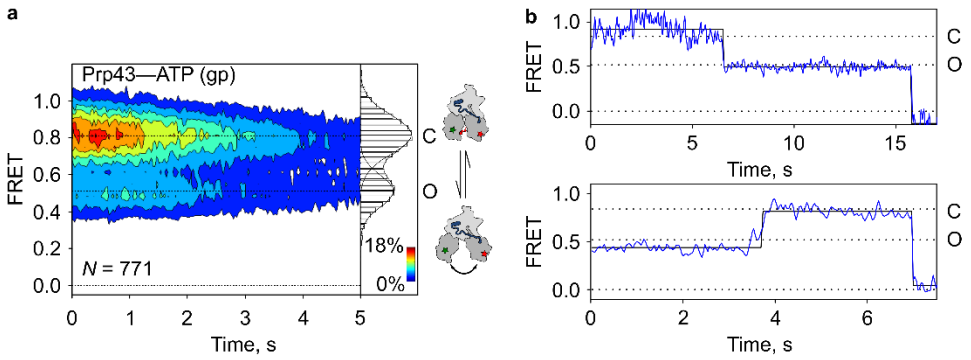
Supplementary Figure 2: smFRET signals of Cy3/Cy5 labeled Prp43$_{Cys}$.
a-d Representative time traces of Cy3- (green) and Cy5- (red) fluorescence intensity (FI) and FRET (blue, bottom plots) corresponding to (**a**) closed (C) RecA domains with E_{FRET}=0.8 or (**b**) open (O) RecA domains with E_{FRET}=0.55 and to transitions from (**c**) the O to C or (**d**) from the C to O state. Solid lines represent the Hidden-Markov fit of the traces.



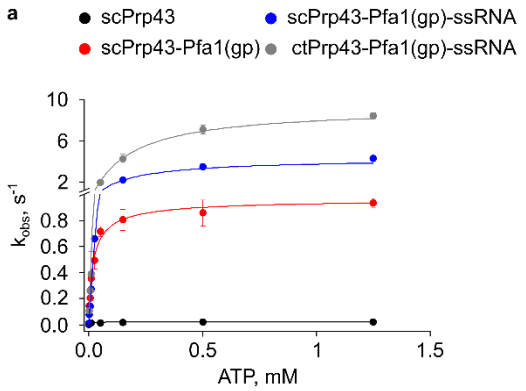
Supplementary Figure 3: Assignment of conformational states of the RNA probe.
a Representative smFRET time traces corresponding to the F (top plot) and B (bottom plot) RNA states. Solid lines represent the Hidden-Markov fit.
b-d Contour plots and 2D histograms of FRET values (mean±sd, derived from N=3 independent data sets) of the RNA probe (**b**) in isolation (0.79±0.01), (**c**) in the presence of 5 µM Prp43 and 2 mM ADP (0.78±0.02) and (**d**) in the presence of 2.5 µM Pfa1(gp) (0.80±0.01).



Supplementary Figure 4: Unwinding of dsRNA in the absence of continuous ATP hydrolysis.
a Steady-state ATP hydrolysis rates of scPrp43$_{wt}$ at 2mM ATP or 2 mM AMPPNP and scPrp43$_{E216A}$ at 2 mM ATP. Pfa1(gp) and A_{20}-ssRNA were added at saturation. k_{cat}±sd were obtained in N=3 independent measurements.
b Contour plot and 2D histogram showing the distribution of FRET values (mean±sd, derived from N=3 independent data sets) of the RNA probe in the presence of the Prp43-Pfa1(gp) complex at 2mM AMPPNP (0.77±0.02 (F), 0.51±0.02 (B), and 0.28±0.05 (PU) and corresponding transition density plot visualizing the frequency of transitions between B and F or B and PU states.
c Contour plot and 2D histogram showing the distribution of FRET values (mean±sd, derived from N=3 independent data sets) of the RNA probe in the presence of the Prp43$_{E216A}$-Pfa1(gp) complex at 2mM ATP (0.74±0.02 (F), 0.50±0.01 (B) and 0.34±0.04 (PU) and corresponding transition density plot visualizing the frequency of transitions between B and F or B and PU states.



Supplementary Figure 5: Acceleration of conformational cycling by Pfa1(gp).
a Contour plot and 2D histogram showing the distribution of FRET values (mean±sd, derived from N=3 independent data sets) of Cy3-/Cy5-labeld Prp43—ATP in complex with Pfa1(gp) (0.83±0.02 and 0.51±0.01).
b Representative FRET traces of Cy3-/Cy5-labeld Prp43—ATP in complex with Pfa1(gp) showing transitions from C to O (above) and from O to C (below) states.



	k_{cat}, s^{-1}	K_M, µM
scPrp43	0.02 ± 0.02	11.5 ± 6
scPrp43 + scPfa(gp)	0.9 ± 0.06	18.0 ± 1
scPrp43 + scPfa(gp)/ssRNA	5.1 ± 0.4	210 ± 42
ctPrp43 + ctPfa(gp)/ssRNA	9.6 ± 0.9	191 ± 60

Supplementary Figure 6: Comparison of ATPase activity between scPrp43 and ctPrp43.
Michealis-Menten titration and table with ATP turnover numbers (k_{cat}) and dissociation constants (K_M) showing the stimulation of scPrp43 ATPase activity by Pfa1(gp) and ssRNA compared to ctPrp43 stimulated by Pfa1(gp) and ssRNA.

Supplementary Tables

Supplementary Table 1: Analysis of RecA domain conformation.

	N	$FRET_C$	$FRET_O$	P_C/P_O	Dyn
Steady-state					
Prp43	292	0.77±0.03	0.42±0.08	0.94±0.01/0.06±0.01	
Prp43 + Pfa1(gp)	521	0.75±0.05	0.56±0.03	0.66±0.08/0.34±0.08	7%
Prp43—ADP	694	0.81±0.02	0.49±0.03	0.97±0.02/0.03±0.02	
Prp43—ADP + Pfa1(gp)	395	0.82±0.03	0.54±0.03	0.79±0.04/0.21±0.04	
Prp43—RNA	326	0.79±0.02	0.53±0.04	0.83±0.02/0.17±0.02	
Prp43—RNA + Pfa1(gp)	574	0.77±0.02	0.54±0.01	0.35±0.06/0.65±0.06	6%
ATP hydrolysis					
Prp43—ATP	743	0.80±0.02	0.53±0.02	0.87±0.02/0.13±0.02	4%
Prp43—ATP + Pfa1(gp)	689	0.83±0.02	0.51±0.01	0.65±0.05/0.35±0.05	27%
Prp43—RNA—ATP + Pfa1(gp)	541	0.80±0.01	0.52±0.01	0.70±0.02/0.30±0.02	37%

N – number of traces obtained in three independent measurements

FRET – mean of FRET distribution ± sd

P_C – fraction of population in the C state ± sd

P_O – fraction of population in the O state ± sd

Dyn – percentage of traces with C→O and O→C transitions

Supplementary Table 2: RecA Inter domain distances of Prp43 in C and O conformations.

	R_{D-A} (Förster Equation)	**$R_{C\alpha}$ (Structural data)**
Closed RecA domains	4.3 nm	4.2 nm (Prp43-ADP, PDB: 5D0U)
Open RecA domains	5.2 nm	5.0 nm (Prp22-RNA, PDB: 6I3P) 5.1 nm (Prp43, PDB: 5y88)
Difference	1 nm	0.7-0.8 nm

R_{D-A} – distance between Cy3 and Cy5 calculated using the Förster equation with R_0 = 5.4 nm (Son *et al*, 2020)

$R_{C\alpha}$ – distance between the C_α atoms of K170C and C303 in ctPrp43 or their equivalents in scPrp43 (K167 and T300) and ctPrp22 (Q629 and T762)

Supplementary Table 3: FRET state analysis of the Cy3- and Cy5-labeled RNA substrate.

Protein factors added	**N**	**$FRET_F$**	**$FRET_B$**	**$FRET_{PU}$**	**$P_F/P_B/P_{PU}$**	**Dyn**
No protein	306	0.79±0.01				
Pfa1(gp)	406	0.80±0.01				
Prp43	180	0.77±0.04	0.49±0.02		0.79±0.05/ 0.21±0.05	4%
Prp43 + Pfa1(gp)	922	0.78±0.03	0.48 ±0.02		0.12±0.01/ 0.88±0.01	5%
Prp43—ADP	493	0.78±0.02				
Prp43—ADP + Pfa1(gp)	481	0.80±0.02	0.51±0.02		0.48±0.09/ 0.52±0.09	16%
Prp43—ATP						
Prp43—AMPPNP	340	0.79±0.01	0.52±0.02		0.63±0.04/ 0.37±0.04	12%
Prp43—ATP + Pfa1(gp)	576	0.76±0.03	0.51±0.01	0.35±0.08	0.24±0.12/ 0.55±0.07/ 0.2±0.08	20%
Prp43—AMPPNP + Pfa1(gp)	450	0.77±0.02	0.51±0.02	0.28±0.05	0.21±0.02/ 0.68±0.08/ 0.11±0.09	30%
$Prp43_{E216A}$—ATP + Pfa1(gp)	620	0.74±0.02	0.50±0.01	0.34±0.04	0.44±0.03/ 0.34±0.04/ 0.22±0.04	34%

N – number of traces obtained in three independent measurements

FRET – mean of FRET distribution ± sd

P_F – fraction of population in the free (F) state ± sd

P_B – fraction of population in the bound (B) state ± sd

P_{PU} – fraction of population in the partially unwound (PU) state ± sd

Dyn – percentage of traces with transitions between F, B and PU states

Supplementary Table 4: Transition frequency analysis.

RNA binding	**$k_{B\rightarrow F}$ (n), s^{-1}**	
Prp43—ADP + Pfa1(gp)	0.45 (76)	
RecA domain conformation	**$k_{O\rightarrow C}$ (n), s^{-1}**	**$K_{C\rightarrow O}$ (n), s^{-1}**
Prp43—ATP + Pfa1(gp)	0.47 (56)	0.90 (104)
Prp43—RNA—ATP + Pfa1(gp)	1.92 (223)	1.44 (228)

n – number of transitions

$k_{B\rightarrow F}$ – transition rate from the bound (B) to the free (F) state

$k_{C\rightarrow O}$/ $k_{O\rightarrow C}$ – transition rate from the closed (C) to the open (O) state and *vice versa*

Supplementary Table 5: FRET state analysis of DNA oligos with defined Cy3 and Cy5 distance.

$D_{Cy3\text{-}Cy5}$	**N**	**FRET**
7	195	0.86±0.01
9	206	0.81±0.01
12	177	0.58±0.01
14	302	0.51±0.01
15	123	0.44±0.01
16	65	0.41±0.01
19	64	0.32±0.01

$D_{Cy3\text{-}Cy5}$ – number of nucleotides between Cy3 and Cy5 located on opposite DNA strands

N – number of traces obtained in three independent measurements

FRET – mean of FRET distribution ± sd

SI References

1. M. J. Tauchert, J. B. Fourmann, H. Christian, R. Lührmann, R. Ficner, Structural and functional analysis of the RNA helicase Prp43 from the thermophilic eukaryote Chaetomium thermophilum. *Acta Crystallogr. Sect. Struct. Biol. Commun.* **72**, 112–120 (2016).
2. J.-B. B. Fourmann, M. J. Tauchert, R. Ficner, P. Fabrizio, R. Lührmann, Regulation of Prp43-mediated disassembly of spliceosomes by its cofactors Ntr1 and Ntr2. *Nucleic Acids Res.* **45**, 4068–4080 (2017).
3. F. W. Studier, Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* **41**, 207–234 (2005).
4. S. Adio, *et al.*, Fluctuations between multiple EF-G-induced chimeric tRNA states during translocation on the ribosome. *Nat. Commun.* **6**, 1–11 (2015).
5. R. Rahul, S. Hohng, T. Ha, A Practical Guide to Single Molecule FRET. *Nat. Methods* **5**, 507–516 (2008).
6. J. E. Bronson, J. Fei, J. M. Hofman, R. L. Gonzalez, C. H. Wiggins, Learning rates and states from biophysical time series: A Bayesian approach to model selection and single-molecule FRET data. *Biophys. J.* **97**, 3196–3205 (2009).
7. K. C. Agarwal, R. P. Miech, R. E. Parks, Guanylate Kinases from Human Erythrocytes, Hog Brain, and Rat Liver. *Methods Enzymol.* **51**, 483–490 (1978).
8. H. Son, W. Mo, J. Park, J. W. Lee, S. Lee, Single-Molecule FRET Detection of Sub-Nanometer Distance Changes in the Range below a 3-Nanometer Scale. *Biosensors* **10**, bios10110168 (2020).

Chapter 3: Conformational dynamics of the RNA binding channel

This manuscript is currently in preparation.

Conformational dynamics of the RNA binding channel regulate loading and translocation of the DEAH-box helicase Prp43

Marieke Enders[1], *Ralf Ficner*[1], *Sarah Adio*[1*]

Department of Molecular Structural Biology, Institute of Microbiology and Genetics, Georg-August-University Göttingen, Justus-von-Liebig-Weg 11, D-37077 Göttingen, Germany

*Sarah Adio

Email: sarah.adio@uni-goettingen.de

Author Contributions: M.E. and S.A. performed experiments and data analysis. M.E., R.F. and S.A. designed experiments and interpreted the data. M.E. and S.A. and wrote the paper with contributions of R.F.

Conformational dynamics of the RNA binding channel regulate loading and translocation of the DEAH-box helicase Prp43

Marieke Enders[a], Ralf Ficner[a], Sarah Adio[a]

[a]Department of Molecular Structural Biology, Institute of Microbiology and Genetics, Georg-August-University Göttingen, Justus-von-Liebig-Weg 11, D-37077 Göttingen, Germany

The DEAH-box helicase Prp43 has essential functions in pre-mRNA splicing and ribosome biogenesis, where it remodels structured RNAs and RNA-protein complexes. G-patch (gp) factors activate Prp43 in its cellular context enhancing the intrinsically low ATPase and RNA unwinding activity. To initiate unwinding, Prp43 must first accommodate a single stranded RNA segment preceding the structured RNA element into its RNA binding channel. This allows translocation of helicase on the RNA. To date, it is unclear how the RNA loading process is accomplished by Prp43 and how it is regulated by its substrates, ATP and RNA, and the gp binding partners. In this study, we developed single-molecule (sm) FRET reporters on Prp43 to monitor the conformational dynamics of the RNA binding channel in Prp43 in real-time. We show that the channel can alternate between two conformations, the open and closed state, in the absence of binding partners. Binding of Pfa1(gp) and ATP shifts the distribution of states towards the open state, facilitating the accommodation of RNA. After the completion of the loading process, the channel remains firmly closed during successive cycles of ATP hydrolysis, ensuring stable interaction with the RNA. In the absence of Pfa1(gp), it remains predominantly closed preventing efficient RNA loading. Our data reveal how the ligands of Prp43 regulate the structural dynamics of the RNA binding channel thereby controlling the initial binding to the RNA.

The DEAH-box helicase Prp43 remodels structured RNAs and RNA-protein complexes (RNPs) on ribosome precursor complexes and the spliceosome (1-4). Members of this helicase family reorganize their substrates through ATP-dependent motility in 3' to 5' direction along single-strands (5, 6). During maturation of 60S ribosome subunits, Prp43 releases small nucleolar RNAs from their pre-mRNA binding sites (4, 7, 8), and it promotes the processing of 20S to 18S rRNA in precursors of the 40S ribosome (2, 3). During the late states of splicing, Prp43 disassembles the intron lariat complex and disrupts stalled spliceosomes bound to pre-messenger RNA (mRNA) with suboptimal or erroneous splice sites, thus playing a key role in mRNA quality control (9-13). In the absence of binding partners provided by ribosome or spliceosome complexes, the affinity of Prp43 for RNA is low, and the ability to unwind structured RNA is negligible (10, 14-17). So called G-patch proteins recruit the helicase to its cellular target sites, where they locally stimulate helicase and ATPase activity (2-4, 7-9, 15, 18). Interestingly, the conserved glycine-rich consensus sequence motif shared by all the G-patch (gp) proteins is sufficient to achieve efficient stimulation (3, 7, 8, 15, 18, 19). We have recently demonstrated that the gp motif of Pfa1 (Pfa1(gp)) enables Prp43 to translocate processively along the RNA controlling the movement of its RecA domains such that dissociation from the substrate is prevented during turnover (17). Still, it remains unresolved how Prp43 gains access to its specific RNA substrates which often lie buried in the interior of complex RNA structures or RNPs. Efficient unwinding of double stranded RNA (dsRNA) by DEAH-box helicases requires a loading step onto a single stranded RNA (ssRNA) overhang preceding the duplex (20-22). The ssRNA is accommodated in a binding channel located between the two RecA domains comprising the helicase core and the C-terminal domains (Figure 1), mostly by contacts with the sugar-phosphate backbone, resulting in sequence-independent interaction (14-16). How

the ssRNA strands enter the channel when Prp43 is part of cellular complexes is unclear. It can either insert through the narrow opening formed by the RecA2 and the C-terminal helix bundle (HB) and oligosaccharide binding (OB) domains, which is unlikely, or the entire channel widens up into a groove that can then fold around the RNA. Support for the latter idea comes from crystal structures of Prp43 bound to the ATP analog $ADPBeF_3^-$, which show the RNA binding channel in an open conformation (Figure 1, (15). In these structures the center of mass of the C-terminal HB and OB domains is shifted by approximately 15 Å with respect to the helicase core while the winged helix (WH) domain acts as a hinge between the RecA2 and the HB domain. The opening of the RNA binding channel is essential for Prp43 function as mutations trapping the channel in the closed conformation abolish the unwinding activity (15). Together, these findings suggest a nucleotide-dependent RNA loading mechanism where ATP binding induces the opening of the channel such that ssRNA can enter in a single step. The subsequent closure of the channel around the RNA as seen in the Prp43—$ADPBeFe_3^-$—RNA complex would prime the helicase for translocation on the substrate (15). Although this mechanism seems plausible on first sight, it is not fully consistent with structural studies on other DEAH-box helicases. So far, the only other DEAH-box helicase observed with an open RNA binding channel is Dhr1 (DHX37 in humans) in the absence of ligands and co-factors (23). Nucleotide-dependent RNA loading is only compatible with this structure assuming a dynamic equilibrium of open and closed states which is shifted in favor of the open state by the binding of ATP. The channel is closed in structures of DHX15, (the human homolog of Prp43) in complex with NKRF(gp) with and without ADP (24). The G-patch connects the RecA2 and WH domains of the helicase across the back side of the RNA binding channel and induces no apparent changes in this region, suggesting it does not promote the opening of the channel. On the other hand, G-patch binding substantially strengthens the affinity of Prp43 for RNA from the µM to the nM range, which indicates that it has a massive impact on the RNA binding properties of the enzyme (17, 24, 25). Whether this also effects the initial RNA binding step was not addressed so far and structures of helicase-G-patch-complexes with an open channel are not available. We have shown that when Prp43 undergoes successive cycles of ATP hydrolysis to processively translocate on the RNA, the likelihood for drop-off is highest in the ADP-state while it is much lower in the nucleotide-free and ATP-bound states (17). Tight RNA binding of Prp43—ATP disagrees with ATP induced channel opening unless further contacts govern the movement of the channel. Evidence proving the physiological relevance of channel opening is lacking, due to the absence of information on the structural dynamics of DEAH-box helicases during interaction with their binding partners, including the natural ATP substrate. This limits the understanding how their activation is primarily achieved and how regulation of helicases with similar architecture of the RNA binding, for example the viral NS3 helicases, is accomplished (26). Here, we use single molecule Förster Resonance Energy Transfer (smFRET) to monitor the conformational dynamics of the RNA binding channel in Prp43 in real-time using total internal reflection of fluorescence (TIRF) microscopy. We show that in the absence of binding partners the channel can alternate between open and closed conformations. While the channel is mainly closed in the apo state, allosteric binding of ATP and Pfa1(gp), the cellular activators of the helicase, shifts the equilibrium towards the open state, allowing the accommodation of ssRNA. When the loading step is accomplished, the channel remains exclusively in the closed state as required for processive translocation. In the absence of Pfa1(gp), efficient loading does not take place. Our data explain the role of channel opening for Prp43 function and allow the integration of different channel conformations into a mechanistic model of the RNA loading process by DEAH-box helicases.

Results

smFRET label positions on Prp43 to monitor the structural dynamics of the RNA binding channel. To monitor the structural dynamics of the RNA binding channel by smFRET, we placed Cy3 and Cy5 fluorescence labels in the RecA1 domain and the winged-helix (WH) domain located in the C-terminal region of Prp43 from *Chaetomium thermophilum* (ctPrp43) (Figure 1).

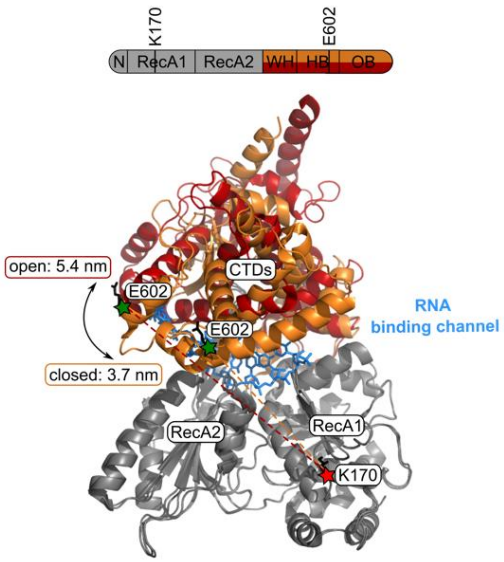
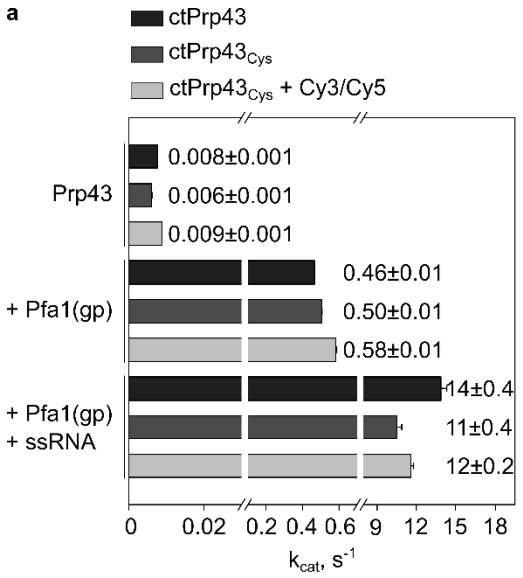
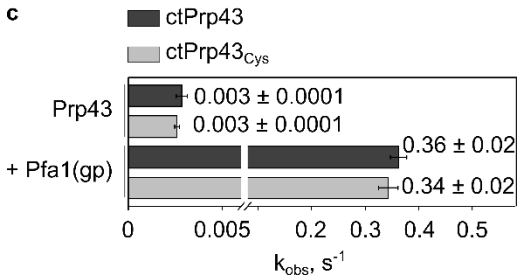


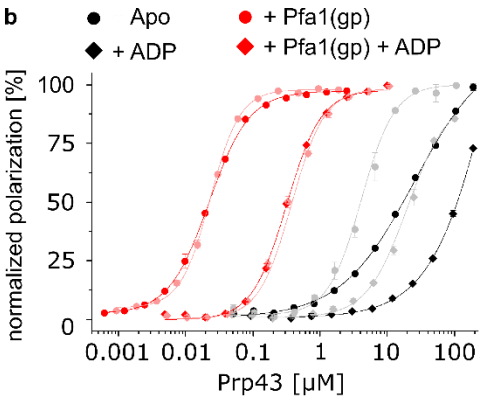
Figure 1. Open to closed transition of the RNA binding channel in Prp43. Domain organization and structures of ctPrp43 (residues 61-764) in complex with ADP-BeF_3^- in presence and absence of U_7-RNA (PDB: 5lta, 5ltk) are shown. The truncated N-terminal domain (residues 61-96) and the RecA domains are depicted in grey shades, (RecA1: residues 97-273; RecA2: residues 274-458). The C-terminal domains (winged-helix (WH): residues 459-526, helix-bundle (HB): residues 527-640, and OB-fold: residues 641-764) are shown in red for the RNA-bound structure and in orange for the RNA-free structure. Cy3- and Cy5-label positions at K170 and E602 are indicated as red and green stars.

We generated two maleimide-reactive fluorescence-labeling sites by introducing cysteine residues at positions K170 and E602. To achieve selective labeling, we had to consider the eight native cysteine residues within ctPrp43. C148, C214, and C377 are buried inside the protein, inaccessible to the maleimide coupling group and remained in their native state. C303, C323, C441, C508, and C543 are exposed on the protein surface and were replaced by site directed mutagenesis (C303T, C323V, C441A, C508A and C543S). Furthermore, we attached a non-cleavable histidine-tag (His_6) to the C-terminus of the protein, which allows immobilization of Prp43 on glass coverslips. The resulting Prp43 variant ($Prp43_{Cys}$) was labeled with a mixture of Cy3- and Cy5-maleimide, yielding dual-labeled $Prp43_{Cys}$ with labeling efficiencies of 81% Cy3 and 98% Cy5 as determined by photometrical analysis. In a previous study we have shown that the protein is stable during the labeling process and in smFRET experiments (17). The chosen label positions undergo a large distance change during opening of the channel (Figure 1). Comparison of crystal structures of ctPrp43 in complex with ADP-BeF_3^- in the presence (channel closed) and absence (channel open) of ssRNA revealed that the distance between the C_α atoms of K170 in RecA1 and E602 in WH increases from 3.7 nm in the closed channel structure to 5.4 nm in the open channel structure (closed: PDB: 5lta; open: PDB:5ltk; (15)). The Cy3/Cy5 FRET dye pair has a Förster radius (R_0) of approximately 5.4 nm (27) and is therefore highly suitable to analyze this range of the movement.

ATPase activity, RNA binding and RNA unwinding by $Prp43_{Cys}$. To ensure that the functionality of $Prp43_{Cys}$ is comparable to Prp43(wt), we measured the rates of ATP turnover by $Prp43_{Cys}$ and Cy3/Cy5-labeled $Prp43_{Cys}$ and addressed in how far ATP hydrolysis can be stimulated by the presence of the G-patch factor ctPfa1(gp) and the RNA substrate (Figure 2A). The low basic ATPase activity of Prp43Cys increases by about 50-fold in the presence of Pfa1(gp) and by about 1000-fold in the presence of Pfa1(gp) and ssRNA, which is similar to wt Prp43 and indicates that the catalytic activity is neither impaired by the mutations nor by the fluorescence tags (Figure 2A). The binding affinity of $Prp43_{Cys}$ for ssRNA was in the µM range without Pfa1(gp) and increased strongly by addition of Pfa1(gp) (k_D =25 µM and k_D =0.02 µM,







	K_D(ssRNA), μM (±sd)
Prp43	4.1±0.24
Prp43$_{Cys}$	25.1±2.5
Prp43 —ADP	20.2±2.2
Prp43$_{Cys}$—ADP	> 100
Prp43+Pfa1(gp)	0.02±0.002
Prp43$_{Cys}$+Pfa1(gp)	0.02±0.002
Prp43—ADP+Pfa1(gp)	0.38±0.03
Prp43$_{Cys}$—ADP+Pfa1(gp)	0.32±0.01

Figure 2. Prp43$_{Cys}$ shows wt-like ATPase activity, RNA binding and unwinding. **(a)** Steady-state ATP hydrolysis rates of ctPrp43, Prp43Cys, and Cy3/Cy5-labeled Prp43Cys at 2 mM ATP, determined using a coupled enzymatic assay. Pfa1(gp) and poly-U ssRNA were added at concentrations of 5 μM and 50 μg/ml respectively. kcat±sd were obtained in N=3 independent measurements. **(b)** Prp43$_{Cys}$ (dark shade) and Prp43 wt (light shade) binding affinity to poly-U ssRNA determined by fluorescence polarization spectroscopy in the absence (black) or presence (red) of Pfa1(gp) in the nt-free (circles) or ADP-bound (squares) state. Mean values are derived from N=3 independent measurements are shown; error bars correspond to the sd. The measured polarization was normalized by the maximum value. Affinity constants (k_D(ssRNA)±sd) are indicated in the table. **(c)** Unwinding of a 28 nt RNA duplex with 20 nt 3'-ss overhang by ctPrp43 and Prp43$_{Cys}$ monitored in a fluorescence based assay in absence and presence of 5 μM Pfa1(gp). k_{obs}±sd were obtained in N=3 independent measurements.

respectively, Figure 2B). Binding of ADP generally reduces the affinity of Prp43Cys for RNA, leading to an extremely low dissociation constant (k_D>100μM) without the G-patch and a moderate dissociation constant of the Prp43Cys-Pfa1(gp) complex (k_D=0.4 μM). k_D values of Prp43$_{Cys}$ overall agree with values of the Prp43(wt) (Figure 2B) and with previous measurements on ctPrp43 and the human homologue DHX15 with its G-patch factor NKRF (Hamann *et al*, 2019; Studer *et al*, 2020). Hence, the general RNA binding properties of Prp43$_{Cys}$ are not substantially changed by the mutations, albeit they seem to some extent decrease the already weak RNA affinity in absence of Pfa1(gp) (Figure 2B). Finally, we monitored the unwinding of a dsRNA substrate with a 3' ss overhang in a fluorescence-based assay. One of the RNA strands carries a Cy5-fluorophore on the 5'- and a BBQ-quencher on the 3'-end. Upon unwinding of the dsRNA, the labeled strand forms an internal hairpin, bringing fluorophore and quencher in close proximity

resulting in a decrease of fluorescence. In absence of Pfa1(gp), both Prp43(wt) and Prp43$_{Cys}$ unwind the substrate extremely slowly (k_{obs}=0.003 s-1, Figure 2C), while the addition of Pfa1(gp) leads to comparably faster unwinding (k_{obs}=0.36 s^{-} and kobs=0.34 s^{-}). This indicates a strong stimulation of Prp43 helicase activity by Pfa1(gp) and is in line with previous measurements on ctPrp43 (Tauchert *et al*, 2017; Hamann *et al*, 2021). In summary, the central determinants of Prp43 function, i.e. the ability to bind and hydrolyze ATP, to bind and unwind RNA and the stimulation of these functions by the G-patch factor Pfa1(gp) are very similar in Prp43$_{Cys}$ and the wild-type counterpart. This allows us to interpret the smFRET data obtained in this study in the context of the functional wild-type protein.

Experiment scheme to monitor structural changes of the RNA binding channel in Prp43. To monitor smFRET signals reporting on the structural dynamics of the RNA binding channel in Prp43 by Total Internal Reflection of Fluorescence (TIRF) microscopy, we immobilized Cy3/Cy5-labeled Prp43$_{Cys}$ on neutravidin-functionalized glass cover slips (29).To this end, we attached a biotin-conjugated antibody directed against the C-terminal His$_6$-tag of Prp43$_{Cys}$ to the neutravidin-coated surface (Figure 3). In a previous study we have shown that the His$_6$-tag does not interfere with Prp43 function (17). The fluorescence labels reported on the conformation of the RNA binding channel by generating a high FRET signal in the closed state that would decrease with increasing fluorophore distance upon channel opening. To elucidate under which conditions channel opening occurs and how this movement is regulated, we monitored smFRET of the labeled molecules in presence and absence of Pfa1(gp), ssRNA and different nucleotides.

The Prp43 RNA binding channel shows a dynamic equilibrium between open and closed conformations. First, we analyzed the conformation of the RNA binding channel in Prp43 in absence of ligands. We observed two populations of smFRET signals with E_{FRET}=0.89 and E_{FRET}=0.57 (Figure 4A, Supplementary Table 1, Supplementary Figure 1A and 1B), indicating the presence of two structurally distinct states.

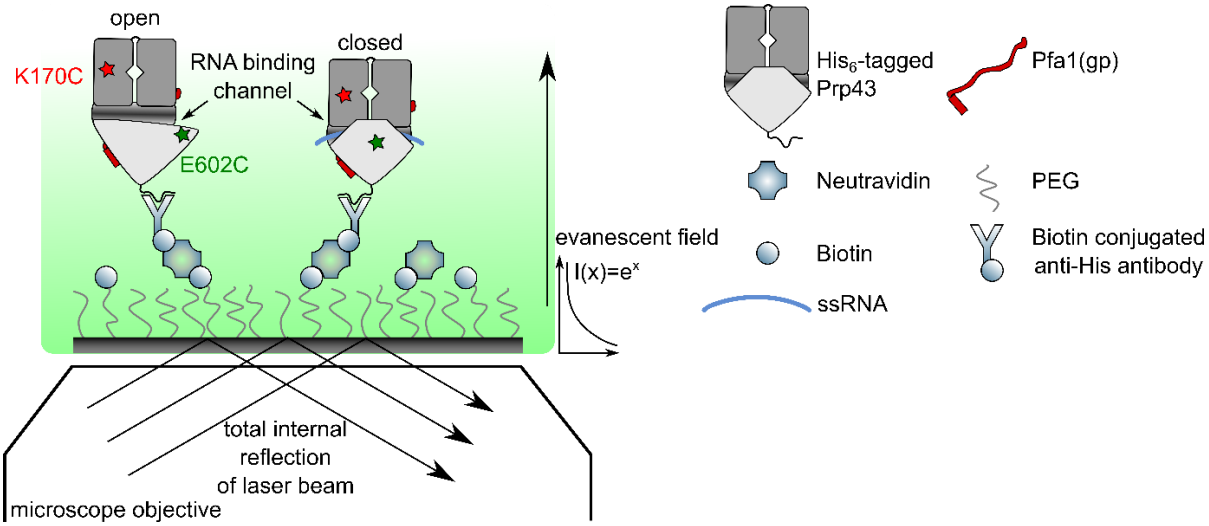


Figure 3. Experimental setup to monitor structural changes of the RNA binding channel in Prp43. A TIRF microscopy experiment scheme monitoring conformational changes of the RNA binding channel in Prp43 by smFRET is shown. Labeled Prp43$_{Cys}$ is immobilized on PEG/PEG-Biotin functionalization glass cover slips, using neutravidin to connect a biotinylated antibody directed against the C-terminal His$_6$-tag of Prp43. A high FRET signal reports on a low Cy3-Cy5 distance corresponding to a closed conformation of the binding channel. Opening of the channel increases the fluorophore distance, leading to decrease of the FRET signal.

Based on these values we estimated the Cy3-Cy5 distance in the two states using the Förster equation (30). The high FRET state corresponds to a distance of 3.8 nm, comparable to the distance between the C_α atoms of the labeled residues in the closed state structure of ctPrp43—ADP ($D_{C\alpha}$=3.2 nm, PDB: 5d0u ((15), Supplementary Table 2). The low FRET state corresponds to a distance of 5.2 nm, which is in good agreement with the open state structures of ctPrp43—ADPBeF$_3^-$ ($D_{C\alpha}$=5.1 nm, PDB: 5ltj, and $D_{C\alpha}$=5.4 nm, PDB: 5ltk ((15), Supplementary Table 2). Although the majority of the helicase molecules (87%, Supplementary Table 1) shows a closed binding tunnel, in a small but significant fraction of molecules the tunnel is open. So far, the open conformation of the RNA binding channel was only observed in the structures of ctPrp43 bound to the ATP hydrolysis transition state analogue ADPBeF$_3^-$ and of the DEAH-box helicase Dhr1 in the apo state (15, 23). The fact that we can monitor the open conformation by smFRET shows that this state occurs indeed in physiological solution. A small fraction of smFRET traces (5%) shows transitions between open and closed states (Supplementary Figure 1C and 1D), suggesting the presence of an equilibrium with slow exchange between the populations. While the closed channel conformation seems to be favored under these conditions, the open state is also transiently occupied. This is in line with the existence of apo state DEAH-box- helicase structures with both closed and open RNA binding channel (16, 23). The dynamic of the equilibrium might be influenced by the binding partners of the helicase, i.e. G-patch, nucleotide

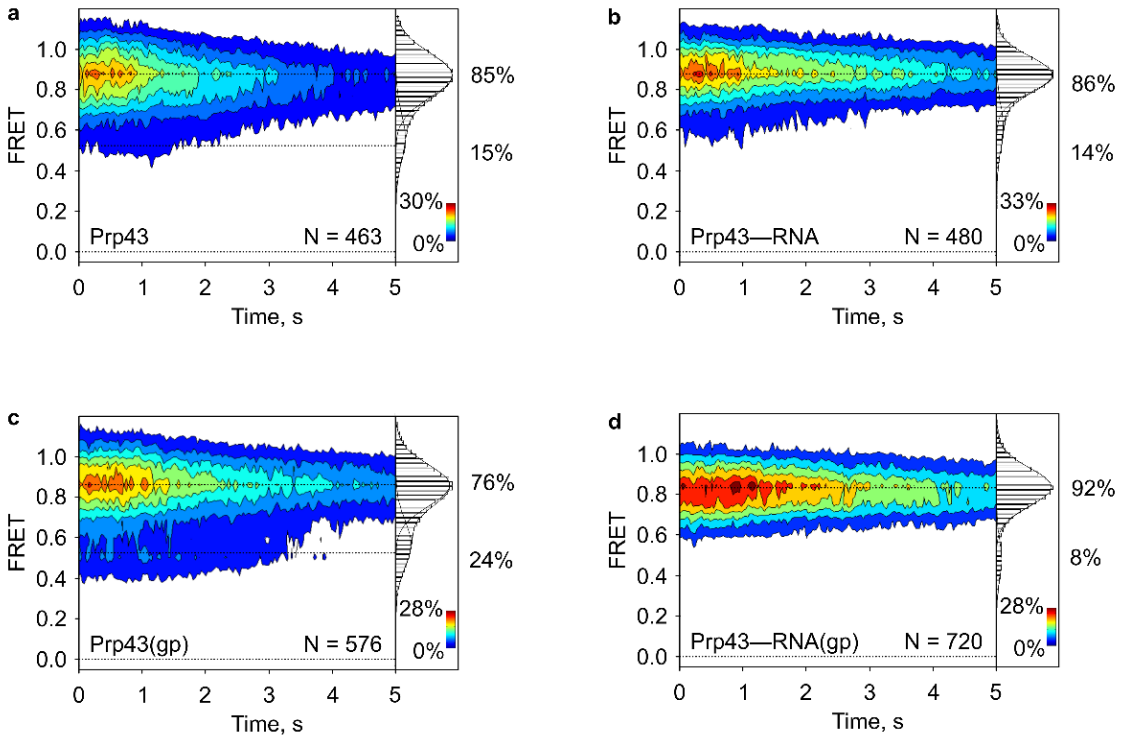


Figure 4. Pfa1(gp) and RNA modulate the conformation of the RNA binding channel. Contour plots and 2D histograms showing the distribution of FRET values (mean±sd) of (**a**) Prp43 in the apo state (0.89±0.02 and 0.57±0.02), (**b**) Prp43-Pfa1(gp) complex (0.89±0.01 and 0.57±0.02), (**c**) Prp43—RNA (0.89±0.01 and 0.62±0.01) and (**d**) Prp43—RNA-Pfa1(gp) complex (0.85±0.01 and 0.48±0.02). Normalization was performed here and in all further FRET distributions by the number of FRET counts. N here and in all further plots indicates the number of individual traces. Data are from n=3 independent experiments.

or ssRNA, providing a basis for regulation. To test if the interaction with a G-patch partner alters the observed equilibrium, we next monitored smFRET of the Prp43-Pfa1(gp) complex. The fraction of the helicase molecules in the open state increased notably from 15% to 24% compared to the apo protein, and more molecules (9%) showed transitions from the closed towards the open state (Figure 4C, Supplementary Figure 2A, Supplementary Table 1). Interestingly, when we added ssRNA to the complex, we observed the reverse trend and the fraction of Prp43 molecules with an open channel decreased to 8% (Figure 4C, Supplementary Table 1). Additionally, the FRET values of the closed state show a small shift from E_{FRET}=0.89 observed in the previous conditions to E_{FRET}=0.85 (Supplementary Figure 3, Supplementary Table 1). This value corresponds to an increase of the Cy3-Cy5 distance by 0.2 nm to 4.0 nm (Supplementary Table 2) suggesting that the fluorophores move further apart when the binding channel closes around the RNA. Indeed, this interpretation is supported by comparison of the Prp43—ADP and Prp43—ADPBeF$_3^-$—ssRNA structures, where the distance between the C_α atoms of the labeled residues increases by 0.4 nm (Supplementary Table 2) (15, 31). In the absence of Pfa1(gp) the distribution of FRET values of the Prp43-RNA complex is similar to the apo state with E_{FRET}=0.89 for the closed and E_{FRET}=0.62 for the open conformation. Given the low affinity of $Prp43_{Cys}$ for ssRNA in absence of Pfa1(gp) (Figure 2B) this expectably indicates that only a minor amount of the present RNA is stably bound under the observed conditions. Together, our data support the idea that the RNA binding channel in Prp43 can alternate between open and closed conformations. The co-factor Pfa1(gp) modulates the equilibrium of these structural states promoting the channel opening. Based on our distance estimations (E_{FRET}=0.57; D_{D-A}≈5.2 nm, Supplementary Table 1, Supplementary Table 2), the channel opens wide enough for ssRNA to enter in one step. This argues against a mechanism where RNA binding occurs by threading through the opening of a closed channel.

The interplay of ATP and Pfa1gp leads to efficient RNA loading and stable binding. Next, we addressed how nucleotide binding influences the conformation of the RNA binding channel (Figure 5). In the ADP state, FRET values and the population distribution of open and closed states strongly resemble the values observed for Prp43 in the apo state (Figure 4A and 5A). 21% of the helicase molecules show an open RNA binding channel and open-closed transitions are visible in 4% of the total population (Supplementary Figure 4, Supplementary Table 1). In presence of ATP however, the equilibrium is clearly shifted towards the open state (Figure 5A, Supplementary Table 1). Here, 32% of the population adopted the open conformation and an increased fraction of molecules showed transitions between conformations, most of them corresponding to opening of the channel (9%, Supplementary Table 1, Supplementary Figure 2B). Interestingly, a very small fraction of the helicase population shows a state with a very low E_{FRET} around 0.25 (Figure 5A). Visual inspection of traces showed that this state is only transiently sampled, which indicates that in presence of ATP the binding channel can open even further than in the previously studied conditions or the crystal structures, leading to estimated Cy3-Cy5 distances of up to 6.5 nm. The addition of Pfa1(gp) to the Prp43—ATP complex did not significantly change the FRET efficiencies and the distribution of states (Figure 5B, Supplementary Table1). However, less of the extreme wide opening events were observed than in the absence of Pfa1(gp) (Supplementary Figure 2C). It is therefore conceivable that the G-patch restricts the opening range of the channel. This could increase the productivity of the RNA binding as the distance between the RNA backbone interacting with the RecA domains and the HB domain forming the upper side of the binding channel would be restrained. In fact, when we added ssRNA to the Prp43—ATP-Pfa1(gp) complex, nearly the entire Prp43 population (97%) shows the closed channel conformation previously observed upon stable RNA binding by the Prp43-Pfa1(gp) complex (E_{FRET}=0.85, Figure 4D) while almost no

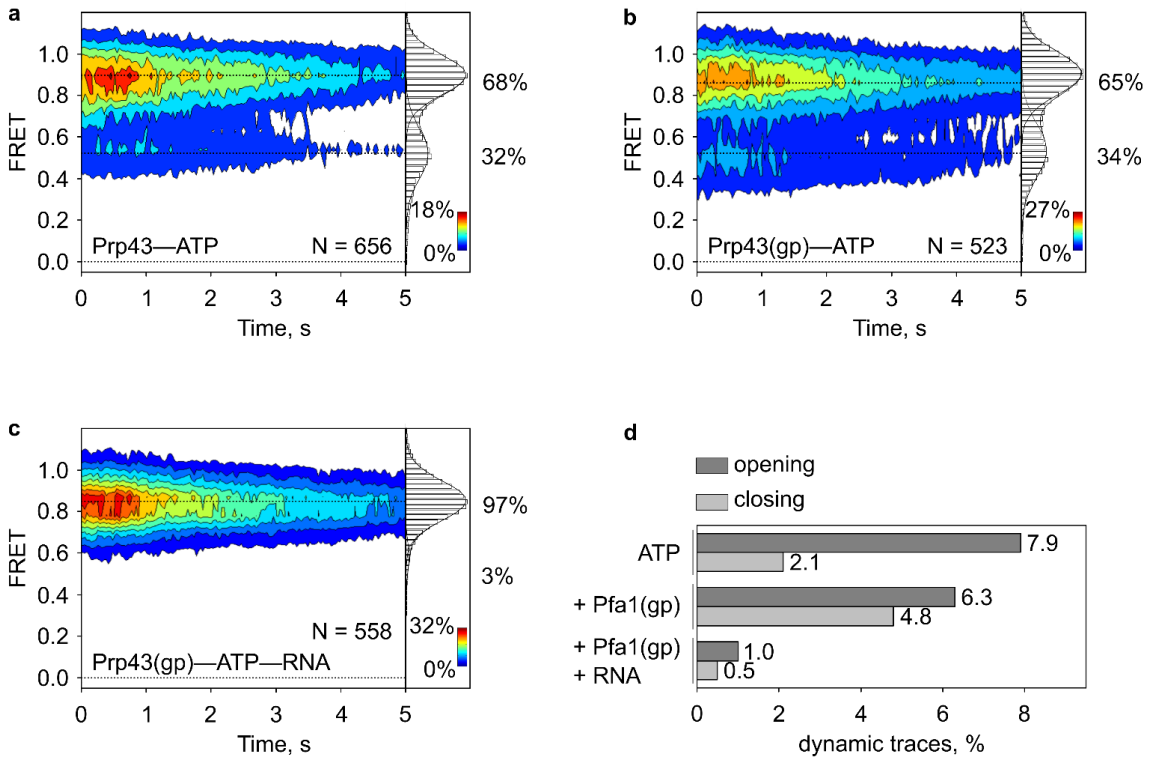


Figure 5. Influence of ATP on the conformation of the RNA binding channel. **(a-c)** Contour plots and 2D histograms showing the distribution of FRET values (mean±sd) of (**a**) Prp43—ATP (0.90±0.01 and 0.52±0.01), (**b**) Prp43—ATP-Pfa1(gp) complex (0.90±0.01 and 0.54±0.04) and (**c**) Prp43—ATP—RNA-Pfa1(gp) complex (0.85±0.01 and 0.57±0.03). **(d)** Transition frequency, determined by dividing the total number of transitions between closed and open conformation observed in one condition by the total number of traces (*N*) of that condition.

transitions occur (Figure 5C and 5D, Supplementary Table 1). The interplay of Pfa1(gp) and RNA seems to lock the channel in a closed conformation, keeping it stably bound around the RNA substrate. In summary, our data show that in the absence of the RNA substrate ATP binding strongly shifts the dynamic equilibrium of channel conformations towards the open state, resulting in an increased part of the population being primed for RNA loading. While the G-patch slightly reinforces this effect, it likely also restricts the opening range of the channel to facilitate efficient RNA loading. Upon binding of the RNA substrate, the G-patch supports the closed channel conformation to allow stable binding which is an essential prerequisite for processive motion.

Discussion

RNA loading mechanism of Prp43. Our data show that in the absence of ligands and co-factors the RNA binding channel in Prp43 can exist in two different conformations, which we denominate the open and closed states. Apo-Prp43 alternates slowly between the two states, but shows a strong preference for the closed state (Figure 6). Binding of the co-factor Pfa1(gp) slightly shifts the dynamic equilibrium of channel conformations towards the open state. Binding of ATP to the Prp43-Pfa1(gp) complex further enhances channel opening. The subsequent accommodation of ssRNA promotes the closure of the channel, and the open state is near absent indicative of the productive loading of the helicase onto ssRNA. The closed channel state allows stable binding of Prp43 during continuous

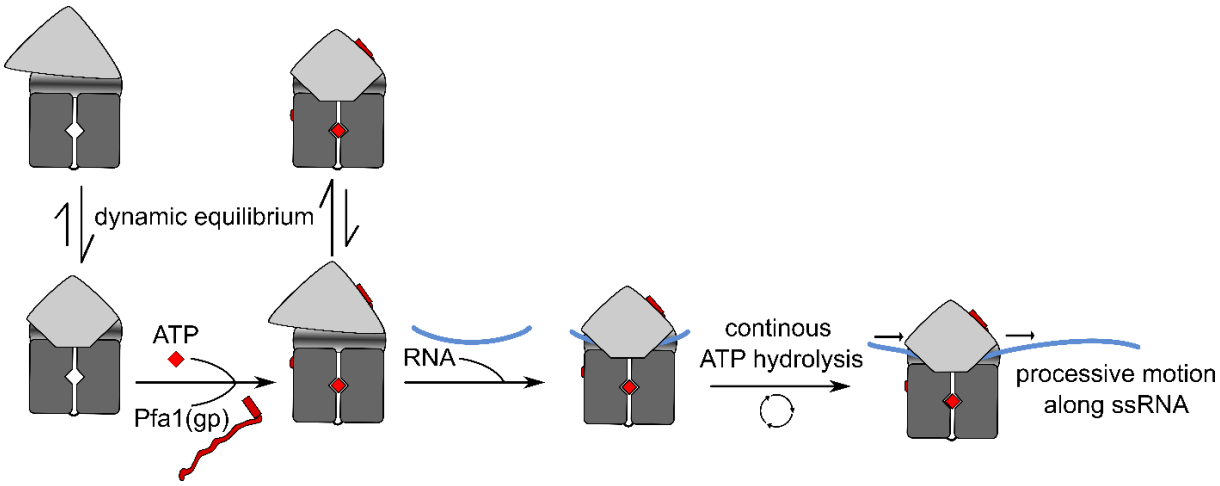


Figure 6. Model of RNA loading by Prp43. Pfa1(gp) is depicted in red, the RecA and C-terminal domains of Prp43 are colored in shades of grey and the RNA is depicted in blue. The model describes how the interaction with Pfa1(gp) and ATP shifts the RNA binding channel towards an open conformation, increasing the efficiency of RNA loading. In the RNA-bound state, the channel remains closed which allows for processive motion without losing contact to the ssRNA.

ATP hydrolysis cycles. In the absence of the G-patch, the closure of the channel around RNA is not observed and Prp43 dissociates rapidly (17). This preventing futile interactions with RNA when unwinding is not required. Hence, the main function of the G-patch in the RNA loading process relates to the opening of the channel prior to RNA binding. Although the channel is closed after the accommodation of RNA, we cannot distinguish whether Pfa1(gp) actively stabilizes the closed state or whether interactions between the RNA and residues of the tunnel are sufficient to maintain the closed state during turnover. In either case, closure of the channel enables Prp43 to translocate along ssRNA without detachment.

Comparison to other DEAH-box helicases. Crystal structures of other DEAH-box helicases support our finding that the RNA binding channel can alternate between open and closed conformations. A closed channel conformation was observed for the spliceosomal helicase Prp22 in the apo state (16), while apo-Dhr1 (DHX37 in humans), a DEAH-box helicase acting in ribosome assembly, showed an open channel (23). As the RNA interaction mode is conserved in the DEAH-box family of helicases (15, 16, 32) a common mechanism of RNA loading is highly probable. Loading of the RNA into an opened channel is advantageous compared to the threading of a single-strand through the entrance of the closed channel because RNA binding can take place in one single event. Additionally, with an open channel Prp43 can bind on single-strand RNA segments of roughly the size of the helicase footprint. The unobstructed 3'-overhangs are often buried inside RNA structures or RNPs. Therefore, this mechanism seems to be well adapted to the cellular context.

Role of the G-patch. Besides shifting the conformation of Prp43 towards channel opening, the G-patch likely is responsible for the channel to stay closed around RNA, even during continuous ATP hydrolysis. Prp2 is a further DEAH-box family member known to interact with a G-patch partner. Despite a similar binding mode of the gp motif, the interaction with the helicase shows striking differences (24, 33). The G-patch protein Spp2 does not alter the conformation of Prp2 and only stimulates its ATPase activity in the presence of RNA (34, 35). Prp2 has higher affinity for ssRNA than Prp43 that is not influenced by the presence of Spp2 (35), indicating the G-patch interaction is not needed for Prp2 to maintain a

stable grip on the RNA. Presumably, tethering Prp2 close to its target site in the spliceosome is more relevant than directly altering its conformational dynamic. DEAH-box helicases that have no G-patch partner either need to maintain the grip during translocation on their own or might be supported by other cofactors. UTP14A is the co-factor of Dhr1/DHX37 and enhances ATP hydrolysis, RNA binding and unwinding and might therefore play a similar role than Pfa1 in RNA loading (32,36). However, the affinity of Dhr1/DHX37 to RNA in the absence of cofactors is considerably higher compared to Prp43, suggesting that it might stay bound on the RNA during ATP turnover without assistance of a co-factor (32).

Comparison to DEAD-box helicases. DEAH-box helicases bind RNA in a different manner than the structurally related DEAD-box helicases, which unwind structured RNAs by local strand displacement by a non-processive mechanism (37). DEAD-box helicases do not require single stranded overhangs as starting point for unwinding, but merely utilize unpaired substrate regions as a loading aid (38). The composition and function of N- and C-terminal extensions flanking the helicase core is highly diverse. In some DEAD-box helicases, these domains mediate the interaction with RNA (39). The RNA binding site localizes between both RecA domains and is perpendicular to the inter-domain cleft (40). While all DEAD-box helicases initially bind ATP with their N-terminal RecA domain (RecA_N), the recognition of dsRNA differs (41, 42). In YxiN, which is involved in ribosome biogenesis in *Bacillus subtilis*, both RecA domains simultaneously interact with the RNA, but only after ATP has been bound by RecA_N (42, 43). Similar to this binding mode, Hera, which acts in ribosome and tRNA metabolism of *Thermus thermophilus*, also binds dsRNA with both RecA domains (42, 44). In contrast to YxiN, RNA binding by Hera does not depend on the presence of ATP (42). Another variety of dsRNA interaction was observed for Mss116 which is involved in group I and II intron splicing (41, 45). Here, the C-terminal RecA domain functions as RNA-duplex recognition domain while RecA_N only contacts the RNA after ATP is bound and the RecA domains close (41). This diversity of initial binding mechanisms implies a huge potential for differential regulation of DEAD-box helicases that might also facilitate their targeting to particular biological processes (42).

Comparison to other helicase families. In contrast to DEAD-box helicases, other processive superfamily 2 helicases bind ssRNA in a similar manner to DEAH-box helicases, i.e. in a cleft formed between the helicase core and their respective C-terminal domains. Examples are the Ski2-like helicases Hel308 and Mtr4, involved in DNA repair and nuclear RNA processing, or the NS3 helicase of hepatitis C virus (26, 46-50). Comparison between NS3 helicases from different hepatitis C virus strains showed that slight changes in the electrostatic properties of the RNA binding channel can have a huge impact on the efficiency of RNA binding and unwinding indicating that helicase activity is regulated by variations in the RNA binding channel (51).

Material and Methods

Protein expression and purification. The homologues of Prp43 and Pfa1 from *C. thermophilum (ctPrp43, ctPfa1) are annotated as* 'hypothetical protein CTHT_0005780' and 'hypothetical protein CTHT_0048220'. Full-length ctPrp43 and the gp motif of ctPfa1 (residues 662-742) were cloned from genomic DNA of *C. thermophilum var. thermophilum* DSM 1495 into pGEX-6P-1 (Tauchert *et al*, 2016). For ctPrp43, a C-terminal non-cleavable His_6-tag was added to the GST-fusion protein by site-directed mutagenesis.
ctPrp43 and ctPfa(gp) constructs were recombinantly expressed in Rosetta 2 (DE3) cells using an autoinduction protocol adapted from Studier *et. al.* (55) and purified by affinity chromatography followed by size exclusion chromatography as described in (17).

Fluorescence-labeling of ctPrp43. To generate two maleimide-reactive fluorescence-labeling sites in ctPrp43, eight native cysteine residues had to be considered. C148, C214, and C377 are buried inside the protein and inaccessible to the coupling group. C303, C323, C441, C508, and C543 are surface exposed

and therefore accessible to the fluorescence dye. They were replaced by site directed mutagenesis (C303T, C323V, C441A, C508A and C543S). The labeling sites were generated by introducing cysteine residues at position K170 in the RecA1 and E602 in the WH. The mutant protein was expressed and purified as described in (17). The purified protein was mixed with Cy3-maleimeide and Cy5-maleimide (Cytiva), dissolved in dimethylsulfoxide at a molar ratio of 1:2:2 (protein:Cy3:Cy5) and incubated for 30 min at 20°C. Excess dye was removed by Ni-sepharose affinity chromatography, labeled protein was eluted in 50 mM Tris/HCl (pH 7.5), 400 mM NaCl, 5% (v/v) glycerol, 2 mM $MgCl_2$, 250 mM imidazole. The labeled protein was dialyzed twice against 50 mM Tris-HCl, pH 7.5, 300 mM KCl, 3 mM $MgCl_2$ using Slide-A Lyzer Dialsis Casette G2 3.5K (Thermo) for 1 h at 4°C. The labeled protein was concentrated to final concentrations between 40 and 70 µM (Amicon Ultra 50K, Millipore).

Determination of dye to protein ratio. To determine the degree of labeling, i.e. the average number of fluorophore molecules per molecule Prp43, the absorption of the labeled protein was measured at 280 nm as well as at the absorption maxima of Cy3 and Cy5, 552 nm and 650 nm. Both Cy3 and Cy5 also show absorption at 280 nm, thereby increasing the A_{280} for the labeled protein. The correction factors (CF) required to eliminate the contribution of the dyes at 280 nm were provided by the manufacturer as CF_{Cy3}=0.08 and CF_{Cy5}=0.05 (Cytiva). The dye to protein ratio was then calculated for both dyes using

$$c(Prp43) = \frac{A_{280} - \left[CF_{Cy3} \cdot A_{552}\right] - \left[CF_{Cy5} \cdot A_{650}\right]}{\varepsilon_{280}},$$

$$D/_P\,(Cy3) = \frac{A_{552}/_{\varepsilon_{552}}}{c(Prp43)} \text{ and}$$

$$D/_P\,(Cy5) = \frac{A_{650}/_{\varepsilon_{650}}}{c(Prp43)}$$

where A_{xxx} is the absorption of labeled Prp43 at the specified wavelength, ε_{280} is the extinction coefficient of Prp43 at 280 nm, and ε_{552}=150000 $M^{-1}cm^{-1}$ and ε_{650}=250000 $M^{-1}cm^{-1}$ are the extinction coefficients of Cy3 and Cy5 at their absorption maxima.

Sample preparation for TIRF microscopy.
Cover slips and objective slides were cleaned by bath sonication in 1 M KOH and exposure to plasma (FEMTO plasma cleaner, Diener Electronic GmbH, Germany). Surfaces were then silanized by sonication in 3.9 mM N1-[3-(trimethoxysilyl) propyl] diethylenetriamine (Sigma-Aldrich) and 1.7 mM acetic acid, and baked for 20 min at 110 °C. PEG/PEG-Biotin functionalization of silanized surfaces was carried out by incubation with 20 mM PEG-NHS (MeO-PEG-NHS, IRIS Biotech GmbH, PEG1165), 0.2 mM Biotin-PEG-NHS (IRIS Biotech, PEG1057) and 20 mM KOH in 100 mM H_3BO_3 solution for 1 h at room temperature. Excess PEG was removed by washing with H_2O. Cover slips were dried at 60 °C and stored under vacuum. For TIRF experiments, flow chambers were generated by combining objective slides and cover slips with double-sided sticky tape.
The ctPrp43-ctPfa1(gp) complex was formed by incubating 1 µM labeled ctPrp43 with 5 µM ctPfa1(gp) in TIRF buffer A (50 mM Tris-HCl, pH 7.5, 300 mM KCl, 3 mM $MgCl_2$) for 10 min at room temperature. Prior to the experiment, the complex was diluted to 1 nM ctPrp43 with TIRF buffer supplemented with 5 µM ctPfa1(gp). For experiments without ctPfa1(gp), labeled ctPrp43 was diluted to 1 nM with TIRF buffer A. Biotin/PEG-functionalized cover slips were incubated for 5 min at room temperature with TIRF buffer A containing additionally 10 mg ml^{-1} BSA and 1 µM neutravidin (Thermo). Excess neutravidin was removed by washing the cover slip with the same buffer containing 1 mg ml^{-1} BSA. A biotinylated anti-His antibody (Rabbit monochlonal, Sigma) was applied at 1.25 µg/ml to the cover slip. Excess antibody was removed by washing with TIRF buffer A containing 1 mg ml^{-1} BSA. Labeled ctPrp43 or ctPrp43-ctPfa1(gp) complex was applied to the surface and incubated for 1 min at room temperature. Images were recorded after washing with TIRF buffer C (TIRF buffer A with 2.5 mM protocatechuic acid, 50 nM protocatechuate-3,4-dioxygenase (from *Pseudomonas*), 1 mM trolox (6-hydroxy-2,5,7,8-tetramethylchromane-2-carboxylic acid) and 1 mM methylviologen). To study the influence of ssRNA, ADP and ATP, the imaging buffer was supplemented with 50 µg/ml PolyU RNA (Sigma Aldrich), 2 mM ADP or 2 mM ATP. To ensure a constant concentration of ATP during turnover, the imaging buffer was additionally supplemented with 0.1 mg/ml pyruvate kinase and 3 mM phosphoenolpyruvate.

TIRF microscopy.
TIRF imaging was performed on an IX 81 inverted microscope using a PLAPON 60 × 1.45 numerical aperture objective (Olympus, Japan). Fluorescence was excited by a 561 nm solid-state laser operated at a power of 25 mW. Images were recorded with an electron multiplying CCD (charge-coupled device) camera (CCD-C9100-13, Hamamatsu, Japan). In FRET

experiments, color channels were separated by projecting donor and acceptor emission on different parts of the CCD chip using an image splitter (dual view micro imager DV2, Photometrics, USA), filter specifications HQ 605/40, HQ 680/30 (Chroma Technology). Movies were recorded at a rate of 30 frames per second. The experiments were carried out at 22 °C.

Data analysis.
Fluorescence time courses for donor (Cy3) and acceptor (Cy5) were extracted using custom-made Matlab (MathWorks) software as described (29, 52). A semi-automated algorithm (Matlab) was used to select anti-correlated fluorescence traces (correlation coefficient <0.1) exhibiting characteristic single fluorophore intensities. The bleed-through of Cy3 signal into the Cy5 channel was corrected using an experimentally determined coefficient (~0.13 in our setup (29)). All trajectories were smoothed over three data points and truncated to remove photobleaching and photoblinking events. Traces with lifetimes of Cy3 or Cy5 less than 20 frames (0.66 s) or with multiple photobleaching steps were excluded from the analysis. The FRET efficiency was defined as the ratio of the measured emission intensities, Cy5/(Cy3+Cy5) (52). FRET time courses were fitted by Hidden Markov modeling using the vbFRET software package (http://vbfret.sourceforge.net/) (53). Models with different number of states were considered for each data set. FRET changes of <0.1 in idealized trajectories were not considered as transitions. Transitions lasting for only one frame were not included in the analysis as well. About 5% of all traces were poorly idealized by Hidden Markov modelling and eliminated from subsequent analysis. Two-dimensional contour plots were generated from time-resolved FRET trajectories. The set of all FRET traces for a given condition was compiled in a histogram, which was fitted to a sum of Gaussian functions using Matlab code (29). Mean FRET values (mean±sd) and population distribution (p=area under the curve±sd) were calculated from three independent datasets and are summarized in Supplementary Table 1.

ATPase activity assay. ATP turnover by ctPrp43 was monitored using a coupled enzymatic assay following nicotinamide adenine dinucleotide (NADH) absorption at 340 nm over time in a VICTOR Nivo Multimode Microplate Reader (PerkinElmer) (54). Triplicate measurements were performed at room temperature in 20 mM Tris/HCl (pH 7.5), 150 mM KCl and 3 mM $MgCl_2$,250 nM NADH, 500 nM phosphoenolpyruvate, 6–8.3 U/ml pyruvate kinase and 9–14 U/ml lactic dehydrogenase at saturating ATP concentration of 2mM. To obtain suitable reaction velocities, Prp43 was used at a concentration of 2 µM (no stimulation), 0.5 µM (stimulation with Pfa1(gp)) or 0.2 µM (stimulation with Pfa1(gp) and RNA). For measurements in the presence of Pfa1(gp) a 5-fold molar excess over Prp43 was used. Measurements in the presence of RNA were conducted at a concentration of 50 µg/ml PolyU RNA (Sigma). The ATP consumption per minute (k_{obs}) was calculated using

$$k_{obs} = \frac{[\frac{\Delta A_{340}}{\Delta t}]}{\varepsilon_{340} \cdot d \cdot c}$$

where $\Delta A_{340}/\Delta t$ is the slope of the NADH decrease, ε_{340} is the extinction coefficient of NADH, d is the optical path length and c is the protein concentration.

RNA binding assay. The RNA binding of Prp43 was measured by fluorescence polarization spectroscopy using a VICTOR Nivo Multimode Microplate Reader (PerkinElmer). The binding of 6 nM 3' 6-carboxyfluorescein-labeled U_{12}-RNA to up to 200 µM ctPrp43 was monitored as triplicates at room temperature in 20 mM Tris/HCl (pH 7.5), 200 mM NaCl, 5% glycerol and 3 mM $MgCl_2$. For measurements in presence of Pfa(gp), the complex was formed by adding a 5-fold molar excess over Prp43 and incubating for 15 min at room temperature. Experiments in presence of ADP were performed at a constant concentration of 3.5 mM throughout all measurements. The excitation wavelength was 480 nm and the emission was detected at 530 nm for 500 ms. The data were normalized by the maximum of measured polarization and fitted by nonlinear regression using the analysis software OriginPro 9.1.

RNA unwinding assay. A fluorescence-based unwinding assay was used to monitor the RNA unwinding activity of ctPrp43 (56). The disruption of a dsRNA substrate with a 3'-ssRNA overhang, consisting of 5'-GCG CCU ACG GAG CUG GUG GCG UAG GCG CAA AAA AAA AAA AAA AAA AAA-3' and 5'-(Cy5)-GCG CCU ACG CCA CCA GCU CCG UAG GCG C-(BBQ)-3', was tracked by recording the decrease in fluorescence due to the quenching of Cy5 by BBQ. The labeled RNA strand forms an internal hairpin upon unwinding. The resulting proximity of BBQ and Cy5 leads to quenching of the fluorophore. Measurements were performed in 25 m*M* Tris–HCl pH 7.5, 150 m*M* KCl, 3 m*M* $MgCl_2$, 2 m*M* ATP at 25°C and were recorded with a VICTOR Nivo Multimode Microplate Reader (Perkin Elmer). Prp43 was used at a concentration of 250 n*M* and the

3'-overhang dsRNA (Axolabs, Germany) at 500 n*M*. For measurements in the presence of Pfa1(gp) a 5-fold molar excess over Prp43 was used. The excitation and emission wavelengths were set to 640 nm and 685 nm respectively. The helicase reaction speed (k_{obs}) was calculated using the initial slope (20-110s) of each reaction as described in (28).

Acknowledgements and funding sources. We thank Prof. M. Rodnina for access to the TIRF microscopy facility in her department and Dr. A. Dickmanns for critical reading of the manuscript. The work was supported by the German research council (Deutsche Forschungs-gemeinschaft) with project grants to R.F. (SFB860, project A1) and S.A. (SFB860, project A15).

1. J.E. Arenas, J. N. Abelson, Prp43: An RNA helicase-like factor involved in spliceosome disassembly. *Proc. Natl. Acad. Sci.* **94**, 11798–11802 (1997).
2. B. Pertschy, *et al.*, RNA helicase Prp43 and its co-factor Pfa1 promote 20 to 18 S rRNA processing catalyzed by the endonuclease Nob1. *J. Biol. Chem.* **284**, 35079–91 (2009).
3. S. Lebaron, *et al.*, The ATPase and helicase activities of Prp43p are stimulated by the G-patch protein Pfa1p during yeast ribosome biogenesis. *EMBO J.* **28**, 3808–19 (2009).
4. M. T. Bohnsack, *et al.*, Prp43 Bound at Different Sites on the Pre-rRNA Performs Distinct Functions in Ribosome Synthesis. *Mol. Cell* **36**, 583–592 (2009).
5. A. M. Pyle, Translocation and unwinding mechanisms of RNA and DNA helicases. *Annu. Rev. Biophys.* **37**, 317–336 (2008).
6. F. De Bortoli, S. Espinosa, R. Zhao, DEAH-Box RNA Helicases in Pre-mRNA Splicing. *Trends Biochem. Sci.* **46**, 225–238 (2021).
7. J. Robert-Paganin, *et al.*, Functional link between DEAH/RHA helicase Prp43 activation and ATP base binding. *Nucleic Acids Res.* **45**, 1539–1552 (2017).
8. Y.-L. L. Chen, *et al.*, The telomerase inhibitor Gno1p/PINX1 activates the helicase Prp43p during ribosome biogenesis. *Nucleic Acids Res.* **42**, 7330–7345 (2014).
9. R. Wan, C. Yan, R. Bai, J. Lei, Y. Shi, Structure of an Intron Lariat Spliceosome from Saccharomyces cerevisiae. *Cell* **171**, 120-132.e12 (2017).
10. N. Tanaka, A. Aronova, B. Schwer, Ntr1 activates the Prp43 helicase to trigger release of lariat-intron from the spliceosome. *Genes Dev.* **21**, 2312–2325 (2007).
11. J. B. Fourmann, *et al.*, The target of the DEAH-box NTP triphosphatase Prp43 in saccharomyces cerevisiae spliceosomes is the U2 snRNP-intron interaction. *Elife* **5**, e15564 (2016).
12. R. M. Mayas, H. Maita, D. R. Semlow, J. P. Staley, Spliceosome discards intermediates via the DEAH box ATPase Prp43p. *Proc. Natl. Acad. Sci.* **107**, 10020–10025 (2010).
13. J.-B. B. Fourmann, M. J. Tauchert, R. Ficner, P. Fabrizio, R. Lührmann, Regulation of Prp43-mediated disassembly of spliceosomes by its cofactors Ntr1 and Ntr2. *Nucleic Acids Res.* **45**, 4068–4080 (2017).
14. Y. He, J. P. Staley, G. R. Andersen, K. H. Nielsen, Structure of the DEAH/RHA ATPase Prp43p bound to RNA implicates a pair of hairpins and motif Va in translocation along RNA. *Rna* **23**, 1110–1124 (2017).
15. M. J. Tauchert, J. B. Fourmann, R. Lührmann, R. Ficner, Structural insights into the mechanism of the DEAH-box RNA helicase Prp43. *Elife* **6**, 1–25 (2017).
16. F. Hamann, M. Enders, R. Ficner, Structural basis for RNA translocation by DEAH-box ATPases. *Nucleic Acids Res.* **47**, 4349–4362 (2019).
17. M. Enders, R. Ficner, S. Adio, Regulation of the DEAH/RHA helicase Prp43 by the G-patch factor Pfa1. *under Rev. PNAS* (2022).
18. H. Christian, R. V. Hofele, H. Urlaub, R. Ficner, Insights into the activation of the helicase Prp43 by biochemical studies and structural mass spectrometry. *Nucleic Acids Res.* **42**, 1162–1179 (2014).
19. L. Aravind, E. V Koonin, G-patch: a new conserved domain in eukaryotic RNA-processing proteins and type D retroviral polyproteins. *Trends Biochem. Sci.* **24**, 342–344 (1999).
20. N. Tanaka, B. Schwer, Characterization of the NTPase, RNA-binding, and RNA helicase activities of the DEAH-box splicing factor Prp22. *Biochemistry* **44**, 9795–9803 (2005).
21. P. J. Smaldino, *et al.*, Mutational dissection of telomeric DNA binding requirements of G4 Resolvase 1 shows that G4-structure and certain 3'-tail sequences are sufficient for tight and complete binding. *PLoS One* **10** (2015).
22. B. Gilman, P. Tijerina, R. Russell, Distinct RNA-unwinding mechanisms of DEAD-box and DEAH-box RNA helicase proteins in remodeling structured RNAs and RNPs. *Biochem. Soc. Trans.* **45**, 1313–1321 (2017).

23. A. Roychowdhury, *et al.*, The DEAH-box RNA helicase Dhr1 contains a remarkable carboxyl terminal domain essential for small ribosomal subunit biogenesis. *Nucleic Acids Res.* **47**, 7548–7563 (2019).
24. M. K. Studer, L. Ivanovic, M. E. Weber, S. Marti, S. Jonas, Structural basis for DEAH-helicase activation by G-patch proteins. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 7159–7170 (2020).
25. A. U. Heininger, *et al.*, Protein cofactor competition regulates the action of a multifunctional RNA helicase in different pathways. *RNA Biol.* **13**, 320–330 (2016).
26. T. C. Appleby, *et al.*, Visualizing ATP-dependent RNA translocation by the NS3 helicase from HCV. *J. Mol. Biol.* **405**, 1139–1153 (2011).
27. H. Son, W. Mo, J. Park, J. W. Lee, S. Lee, Single-Molecule FRET Detection of Sub-Nanometer Distance Changes in the Range below a 3-Nanometer Scale. *Biosensors* **10**, bios10110168 (2020).
28. F. Hamann, *et al.*, The structure of Prp2 bound to RNA and ADP-BeF3-reveals structural features important for RNA unwinding by DEAH-box ATPases. *Acta Crystallogr. Sect. D Struct. Biol.* **77**, 496–509 (2021).
29. S. Adio, *et al.*, Fluctuations between multiple EF-G-induced chimeric tRNA states during translocation on the ribosome. *Nat. Commun.* **6**, 1–11 (2015).
30. T. Förster, Energiewanderung und Fluoreszenz. *Naturwissenschaften* **33**, 166–175 (1946).
31. M. J. Tauchert, J. B. Fourmann, H. Christian, R. Lührmann, R. Ficner, Structural and functional analysis of the RNA helicase Prp43 from the thermophilic eukaryote Chaetomium thermophilum. *Acta Crystallogr. Sect. Struct. Biol. Commun.* **72**, 112–120 (2016).
32. F. M. Boneberg, *et al.*, Molecular mechanism of the RNA helicase DHX37 and its activation by UTP14A in ribosome biogenesis. *RNA* **25**, 685–701 (2019).
33. F. Hamann, *et al.*, Structural analysis of the intrinsically disordered splicing factor Spp2 and its binding to the DEAH-box ATPase Prp2. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 2948–2956 (2020).
34. R. Bai, *et al.*, Mechanism of spliceosome remodeling by the ATPase/helicase Prp2 and its coactivator Spp2. *Science (80-.).* **371**, eabe8863 (2021).
35. Z. Warkocki, *et al.*, The G-patch protein Spp2 couples the spliceosome-stimulated ATpase activity of the deah-box protein Prp2 to catalytic activation of the spliceosome. *Genes Dev.* **29**, 94–107 (2015).
36. J. Zhu, X. Liu, M. Anjos, C. C. Correll, A. W. Johnson, Utp14 Recruits and Activates the RNA Helicase Dhr1 To Undock U3 snoRNA from the Preribosome. *Mol. Cell. Biol.* **36**, 965–978 (2016).
37. Q. Yang, M. Del Campo, A. M. Lambowitz, E. Jankowsky, DEAD-Box Proteins Unwind Duplexes by Local Strand Separation. *Mol. Cell* **28**, 253–263 (2007).
38. Q. Yang, E. Jankowsky, The DEAD-box protein Ded1 unwinds RNA duplexes by a mode distinct from translocating helicases. *Nat. Struct. Mol. Biol.* **13**, 981–986 (2006).
39. M. G. Rudolph, D. Klostermeier, When core competence is not enough: Functional interplay of the DEAD-box helicase core with ancillary domains and auxiliary factors in RNA binding and unwinding. *Biol. Chem.* **396**, 849–865 (2015).
40. T. Sengoku, O. Nureki, A. Nakamura, S. Kobayashi, S. Yokoyama, Structural Basis for RNA Unwinding by the DEAD-Box Protein Drosophila Vasa. *Cell* **125**, 287–300 (2006).
41. A. L. Mallam, M. Del Campo, B. Gilman, D. J. Sidote, A. M. Lambowitz, Structural basis for RNA-duplex recognition and unwinding by the DEAD-box helicase Mss116p. *Nature* **490**, 121–125 (2012).
42. B. Samatanga, D. Klostermeier, DEAD-box RNA helicase domains exhibit a continuum between complete functional independence and high thermodynamic coupling in nucleotide and RNA duplex recognition. *Nucleic Acids Res.* **42**, 10644–10654 (2014).
43. K. Kossen, O. C. Uhlenbeck, "Cloning and biochemical characterization of Bacillus subtilis YxiN, a DEAD protein specifically activated by 23S rRNA: delineation of a novel sub-family of bacterial DEAD proteins" (1999).
44. P. Donsbach, *et al.*, The Thermus thermophilus DEAD-box protein Hera is a general RNA binding protein and plays a key role in tRNA metabolism. *RNA* **26**, 1557–1574 (2020).
45. J. P. Potratz, M. Del Campo, R. Z. Wolf, A. M. Lambowitz, R. Russell, ATP-dependent roles of the DEAD-box protein Mss116p in group II intron splicing in vitro and in vivo. *J. Mol. Biol.* **411**, 661–679 (2011).

46. A. Van Hoof, P. Lennertz, R. Parker, "Yeast Exosome Mutants Accumulate 3-Extended Polyadenylated Forms of U4 Small Nuclear RNA and Small Nucleolar RNAs" (2000).
47. C. P. Guy, E. L. Bolt, Archaeal Hel308 helicase targets replication forks in vivo and in vitro and unwinds lagging strands. *Nucleic Acids Res.* **33**, 3678–3690 (2005).
48. A. M. I. Lam, D. N. Frick, Hepatitis C Virus Subgenomic Replicon Requires an Active NS3 RNA Helicase. *J. Virol.* **80**, 404–411 (2006).
49. K. Büttner, S. Nehring, K. P. Hopfner, Structural basis for DNA duplex separation by a superfamily-2 helicase. *Nat. Struct. Mol. Biol.* **14**, 647–652 (2007).
50. J. R. Weir, F. Bonneau, J. Hentschel, E. Conti, J. Feigon, Structural analysis reveals the characteristic features of Mtr4, a DExH helicase involved in nuclear RNA processing and surveillance. *PNAS* **107**, 12139–12144 (2010).
51. T. Zhou, X. Ren, R. L. Adams, A. M. Pyle, NS3 from Hepatitis C Virus Strain JFH-1 Is an Unusually Robust Helicase That Is Primed To Bind and Unwind Viral RNA. *J. Virol.* **92** (2018).
52. R. Roy, S. Hohng, T. Ha, A Practical Guide to Single Molecule FRET. *Nat. Methods* **5**, 507–516 (2008).
53. J. E. Bronson, J. Fei, J. M. Hofman, R. L. Gonzalez, C. H. Wiggins, Learning rates and states from biophysical time series: A Bayesian approach to model selection and single-molecule FRET data. *Biophys. J.* **97**, 3196–3205 (2009).
54. K. C. Agarwal, R. P. Miech, R. E. Parks, Guanylate Kinases from Human Erythrocytes, Hog Brain, and Rat Liver. *Methods Enzymol.* **51**, 483–490 (1978).
55. F. W. Studier, Protein production by auto-induction in high-density shaking cultures. *Protein expression & purification.* **41**, 207-234 (2005).
56. C. Belon, D. Frick, Monitoring Helicase Activity with Molecular Beacons. *Biotechniques*. **45**, 433-442 (2009).

Supplementary Information for

Conformational dynamics of the RNA binding channel regulate loading and translocation of the DEAH-box helicase Prp43

Marieke Enders[1], Ralf Ficner[1], Sarah Adio[1]*

Department of Molecular Structural Biology, Institute of Microbiology and Genetics, Georg-August-University Göttingen, Justus-von-Liebig-Weg 11, D-37077 Göttingen, Germany

*Sarah Adio

Email: sarah.adio@uni-goettingen.de

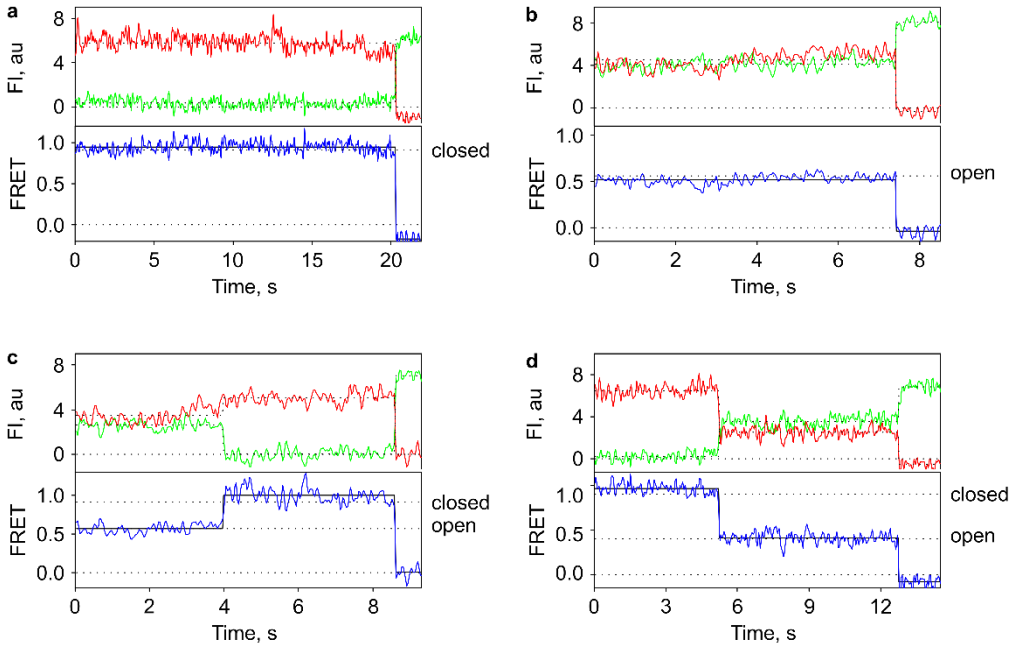
This PDF file includes:

Figures S1 to S6

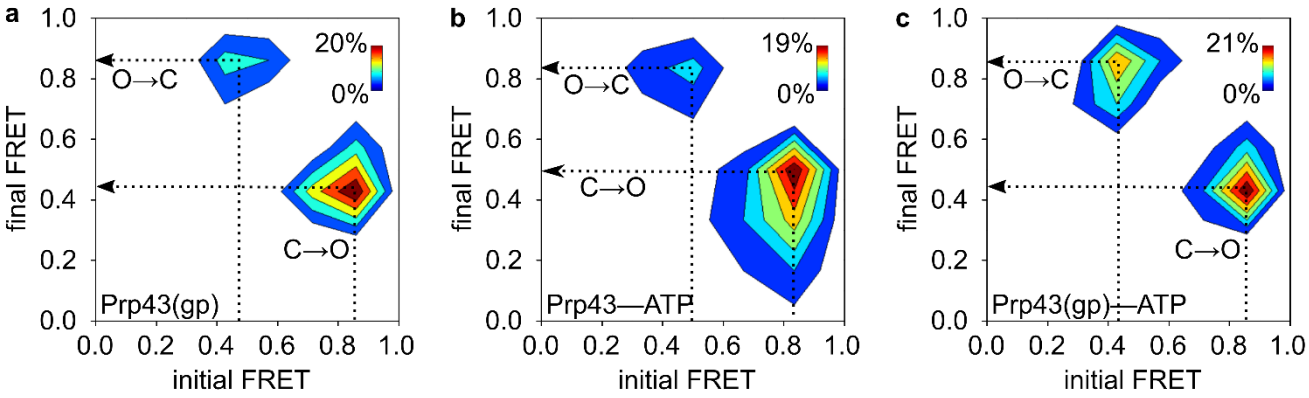
Tables S1 to S5

SI References

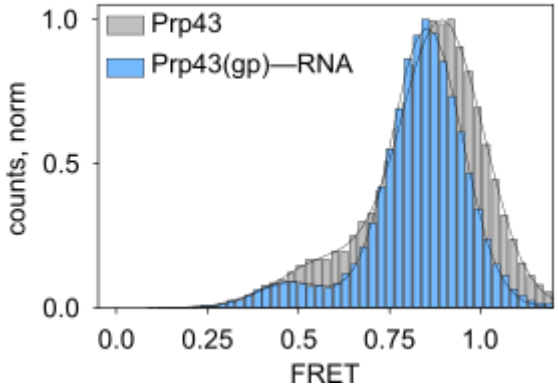
Supplementary Figures



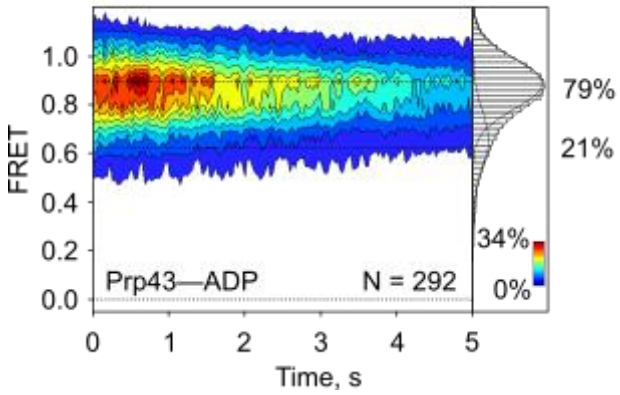
Supplementary Figure 1: smFRET signals of Cy3/Cy5 labeled Prp43$_{Cys}$.
a-d Representative time trajectories of Cy3- (green) and Cy5- (red) fluorescence intensity (FI) and FRET (blue) corresponding to (**a**) closed RNA binding channel or (**b**) open RNA binding channel and to transitions corresponding to (**c**) channel closing or (**d**) opening. Solid lines represent the Hidden-Markov fit of the traces.



Supplementary Figure 2: Pfa1(gp) and ATP influence dynamic of RNA binding channel.
a-c Transition density plots visualizing the frequency of transitions corresponding to opening (C→O) and closing (O→C) of the RNA binding channel in presence of (**a**) Pfa1(gp), (**b**) ATP or (**c**) both.



Supplementary Figure 3: Closed conformation of RNA binding channel slightly shifts upon stable binding of RNA.
Overlay of 2D histograms showing the distribution of FRET values (mean±sd) of Prp43 in absence and presence of Pfa1(gp) and ssRNA. The center of the closed state shifts from E_{FRET}=0.89±01 to E_{FRET}=0.85±01.



Supplementary Figure 4: Influence of ADP on RNA binding channel conformation.
a Contour plot and 2D histogram showing the distribution of FRET values (mean±sd) of Prp43—ADP (0.89±0.01 and 0.65±0.06).

Supplementary Tables

Supplementary Table 1: Analysis of RNA binding channel conformation.

	N	$FRET_C$	$FRET_O$	P_C/P_O	Dyn
Prp43	463	0.89±0.02	0.57±0.02	0.87±0.01/0.13±0.10	5.2
Prp43 + Pfa1(gp)	576	0.89±0.01	0.57±0.02	0.76±0.04/0.24±0.04	8.7
Prp43—RNA	480	0.89±0.01	0.62±0.01	0.86±0.03/0.14±0.03	4.6
Prp43—RNA + Pfa1(gp)	720	0.85±0.01	0.48±0.02	0.92±0.03/0.08±0.03	2.6
Prp43—ADP	292	0.89±0.01	0.65±0.06	0.79±0.11/0.21±0.11	4.1
Prp43—ATP	656	0.90±0.01	0.52±0.01	0.68±0.01/0.32±0.01	9.3
Prp43—ATP + Pfa1(gp)	523	0.90±0.01	0.54±0.04	0.66±0.02/0.34±0.02	9.9
Prp43—ATP—RNA + Pfa1(gp)	558	0.85±0.01	0.57±0.03	0.97±0.01/0.03±0.01	1.4

N – number of traces obtained in three independent measurements

FRET – mean of FRET distribution ± sd

P_C – fraction of population with closed binding channel ± sd

P_O – fraction of population with open binding channel ± sd

Dyn – percentage of traces that showed transitions

Supplementary Table 2: Comparison of channel widths.

	Distance D - A (Förster Equation)	**Distance C_α- C_α (Structural data)**
Open binding channel	E_{FRET} = 0.58 R_{D-A} ≈ 5.1 nm	D = 5.1-5.4 nm (Prp43-ADPBeF$_3^-$, PDB: 5ltj and 5ltk)
Closed binding channel (empty)	E_{FRET} = 0.89 R_{D-A} ≈ 3.6 nm	D = 3.2 nm (Prp43—ADP, PDB: 5d0u)
Closed binding channel (bound RNA)	E_{FRET} = 0.85 R_{D-A} ≈ xx nm	D = 3.6 nm (Prp43—ADPBeF$_3^-$—RNA, PDB: 5lta)

E_{FRET} – average FRET efficiency of C and O states

R_{D-A} – distance between donor and acceptor calculated using the Förster equation ($E_{FRET} = \frac{1}{1+(R/R_0)^6}$ with R_0 = 5.4 nm (Son *et al*, 2020)

D – distance between the C_α atoms of labeled amino acids in ctPrp43 (K170C and E602C)

SI References

1. H. Son, W. Mo, J. Park, J. W. Lee, S. Lee, Single-Molecule FRET Detection of Sub-Nanometer Distance Changes in the Range below a 3-Nanometer Scale. *Biosensors* **10**, bios10110168 (2020).

Chapter 4: Discussion and future perspectives

The main focus of the presented work was to study the mechanistic principles underlying the activation of the DEAH-box helicase Prp43 by G-patch proteins. The development of different smFRET systems and their observation by TIRF microscopy elucidated several aspects of G-patch-induced stimulation. The impact of G-patch interaction on Prp43 domain motility was studied with two smFRET systems reporting on the RecA domain movements (Chapter 2) and the conformation of the RNA binding channel (Chapter 3). Additionally, a reporter with smFRET labels on the RNA substrate provided information on RNA binding and unwinding by the Prp43-G-patch complex (Chapter 2). The combined observations from these studies reveal the regulatory mechanism of Prp43 by G-patch partners and provide a mechanistic model of its motility during loading onto the RNA substrate and processive translocation, which will be described in detail in the following section. While many aspects concerning the obtained results were already thoroughly discussed in chapters 2 and 3, they will be put into a broader perspective here.

4.1 A model for Prp43 stimulation by G-patch proteins

For a multifunctional helicase such as Prp43 that acts on multiple targets in two major pathways of RNA metabolism, a regulatory mechanism ensuring specific activity at the correct cellular targets is crucial. Without such a mechanism, the helicase might disassemble functional spliceosomal complexes before or during catalysis or strip essential snoRNAs or proteins from rRNAs during ribosome biogenesis. However, the results obtained in this work illustrate how interaction with G-patch proteins regulates Prp43 to only act as a productive helicase in the appropriate context.

The RNA binding mode of Prp43 is not sequence specific, generally enabling it to interact with any ssRNA (Tauchert *et al*, 2017; He *et al*, 2017). However, without interaction partners, Prp43 displays only low affinity for ssRNA (Chapter 2, Figure 3). This can at least be partially attributed to an inefficient RNA loading step as the RNA binding channel is preferably closed in the apo state (Chapter 3, Figure 4). The presence of ATP shifts the equilibrium of the RNA binding channel conformation such that opening becomes more likely, increasing the chances for an unspecific RNA binding event (Chapter 3, Figure 5). However, Prp43 would likely lose contact to the bound RNA during the first cycle of ATP hydrolysis as its affinity for ssRNA is severely decreased in presence of ADP (Chapter 2, Figure 3). Additionally, unstimulated Prp43 hydrolyzes ATP extremely slowly, making it basically impossible to translocate along a single

strand before dropping off (Chapter 2, Figure 5). This also ensures that the helicase does not unnecessarily spend energy when it is not bound to its target complexes.

In presence of a G-patch partner, when Prp43 is part of the spliceosome or a pre-ribosomal complex, the situation changes completely. Both ATP and G-patch now facilitate the opening of the RNA binding channel, allowing Prp43 to efficiently load onto its correct substrate (Chapter 3, Figure 4). When the RNA is bound, the G-patch stabilizes the closed conformation of the channel, reducing the chances for drop-off due to channel opening (Chapter 3, Figure 4). Additionally, the G-patch induces an open conformation of the RecA domains that is incompatible with nucleotide binding and thereby accelerates ADP release (Chapter 2, Figure 1). During ATP turnover, this leads to rapid cycling of the RecA domains between closed and open conformations and consequently shorter periods of time spend in the ADP-bound low RNA affinity state (Chapter 2, Figure 5). Under these conditions, dissociation of the Prp43-G-patch complex from the RNA is much slower than the rate of translocation, enabling the helicase to translocate processively along the single-strand for several conformational cycles. When encountering a RNA duplex, the closure of the RecA domains upon ATP binding likely provides the power stroke for unwinding of the first ~5 nucleotides (Chapter 2, Figure 4).

Overall, the G-patch seems to act as processivity factor for Prp43, allowing continuous productive interaction with its RNA substrates. As processivity is an important aspect of helicase activity, it will be more generally discussed in the next section.

4.2 Helicase processivity and implications for Prp43 activity

Generally, the processivity of a helicase is a measure for the number of base pairs unwound or the number of nucleotides translocated per nucleic acid binding event (Lohman, 1993). Helicases showing high processivity can be found in DNA replication and maintenance. Here, unwinding of several thousand base pairs per helicase binding event at rates as high as ∼1,000 base pairs per second have been determined (Bianco *et al*, 2001; Roman *et al*, 1992; Sikora *et al*, 2006). As the natural substrates of theses helicases are extremely long stretches of DNA, high processivity is needed to efficiently fulfill their functions. For most RNA helicases, especially those involved in the rearrangement of large RNP complexes, the situation is different. Their single stranded binding sites are often surrounded by other components of the complex and the possibilities for long continuous translocation are limited. It is therefore not surprising that Prp43-G-patch complexes exhibit comparably low processivity. Assuming a strict coupling between conformational transitions of the RecA domains and the hydrolysis of ATP, the ATP turnover rate of Prp43-Pfa1(gp) in presence of ssRNA (k_{cat}(scPrp43)=5s^{-1}) corresponds to the rate of translocation (5 nt/s). By comparing this to the dissociation of the

Prp43-Pfa1(gp) complex from ssRNA (k_{off}(scPrp43)=0.45s^{-1}), approximately 10 nucleotides could be translocated per RNA binding event. However, this low processivity is completely sufficient for Prp43 function in spliceosome disassembly, where the helicase binds to the 3' end of U6 snRNA (Toroney *et al*, 2019). Due to the limited length of the 3' overhang, only a few steps of translocation, if any, are required to reach the from this perspective first duplex of the ILS, the U2/U6 helix II (Figure 4.1a). Here, Prp43 is thought to employ a winching mechanism to pull on the 3' overhang to disrupt the interactions between the U6 snRNA and the spliceosome (Toroney *et al*, 2019). It is conceivable from the geometry of the U6 snRNA and its interactions with U2 snRNA and the intron, that pulling by approximately 10 nt would cause major disruptions of the adjacent structures and lead to complete disassembly of the ILS (Figure 4.1b).

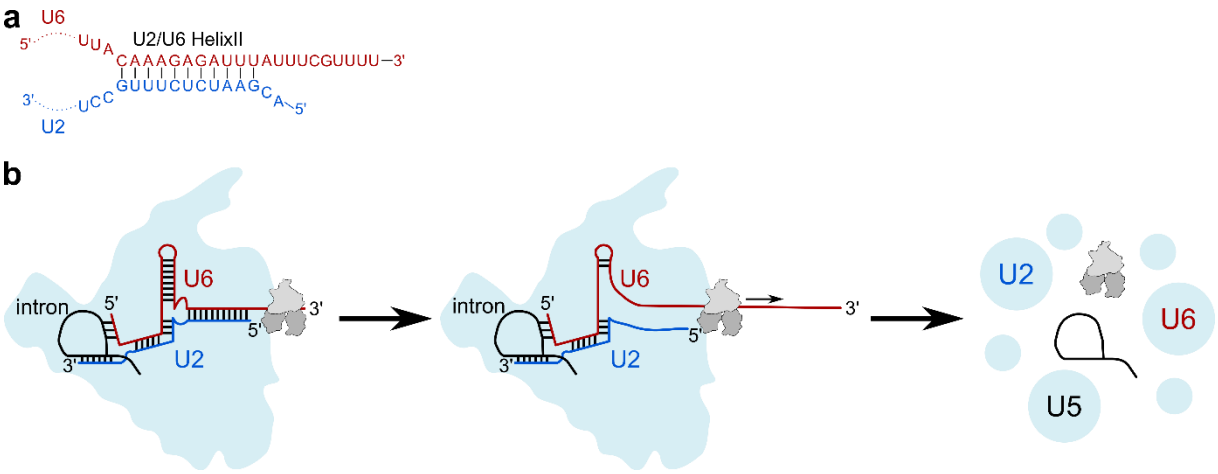


Figure 4.1.: Interaction of Prp43 with its spliceosomal substrate. (a) 3' ss overhang and base-pairing interactions of the U2/U6 Helix II. **(b)** Model for ILS disassembly by Prp43. Secondary structure of interactions between U2 and U6 snRNAs (blue and red) and the intron (black) at the catalytic core of the spliceosome (light blue) are shown. Prp43 binds to the 3' tail of the U6 snRNA and pulls on it through directed translocation. Thereby, the U2/U6 Helix II is disrupted and interactions at the heart of the spliceosome are destabilized, resulting in ILS disassembly. This figure was adapted from Toroney *et. al.*, (2019).

A similar picture is emerging for the other spliceosomal DEAH-box helicases. Due to contacts to other spliceosomal proteins, their range of movement on the ssRNA they bind to is limited and they likely pull on their substrates to induce the major conformational rearrangements of the complex they are associated with (Semlow *et al*, 2016; Bai *et al*, 2021).

To do so, only minor processivity is required whose extend might slightly vary between helicases, depending on the strength of the interactions they are disrupting.

To achieve processivity, a domain architecture of the helicase is required that allows to maintain a strong grip on the nucleic acid during directed motion along the strand. In most cases, a binding channel is formed between the helicase core and auxiliary C-terminal domains. However, many of the highly processive helicases involved in DNA unwinding during replication follow a different design principle. They topological link to their substrate by forming a hexameric ring around the DNA strand. While this efficiently increases the processivity, the ring shape also presents a challenge for substrate loading as usually no free DNA ends are available to thread through the hole. In contrast to monomeric helicases, where a 'simple' conformational change is sufficient to widen the binding channel, hexameric helicases require more substantial rearrangements that typically depend on accessory proteins, (Figure 4.2). On the one hand, so-called 'ring breakers' open a pre-assembled hexameric ring at one subunit interface to allow the DNA to enter. On the other hand, 'ring-makers' act on disassembled helicase subunits, assisting their assembly around the DNA (Davey & O'Donnell, 2003). The non-processive DEAD-box helicases range on the other end of the spectrum. They do not form a nucleic acid binding channel and therefore cannot rely on continuous translocation for strand separation. Instead, they force the bound RNA substrate in a distorted conformation incompatible with duplex structure, leading to local strand displacement (Yang *et al*, 2007).

Strand separation mechanisms of processive helicases can be classified as active or passive. In a passive mechanism, the helicase does not directly interact with the substrate double strand. Instead, it relies on transient base-pair opening, so-called 'breathing', at the junction between single strand and double strand that occurs on a microsecond timescale. By translocating on a recently freed single strand, it is stabilized and re-annealing is prevented. In contrast to this, an actively unwinding helicase directly interacts with the double-stranded region to facilitate duplex destabilization (Pyle, 2008).

Most RNA helicases certainly can be categorized as active unwinders or remodelers. The DEAD-box helicases, which comprise the largest subfamily, are intrinsically incapable of a passive mechanism as they do not translocate. As stated above, spliceosomal DEAH-box helicases are thought to employ a winching mechanism by pulling on single-stranded RNA to disrupt RNA-RNA or RNA-protein interactions that are often located distant from their binding sites (Semlow *et al*, 2016). This is not compatible with a passive mechanism because the path to translocate through potentially passively freed target strands is obstructed by other spliceosomal components (Liu *et al*, 2017; Wan *et al*, 2017; Fica *et al*, 2017; Yan *et al*, 2016).

Although for most RNA helicases involved in ribosome biogenesis the exact binding sites remain uncharacterized, the overall situation is likely similar to the spliceosomal helicases as both pathways involve huge dynamic RNP machineries and little room for directed translocation.

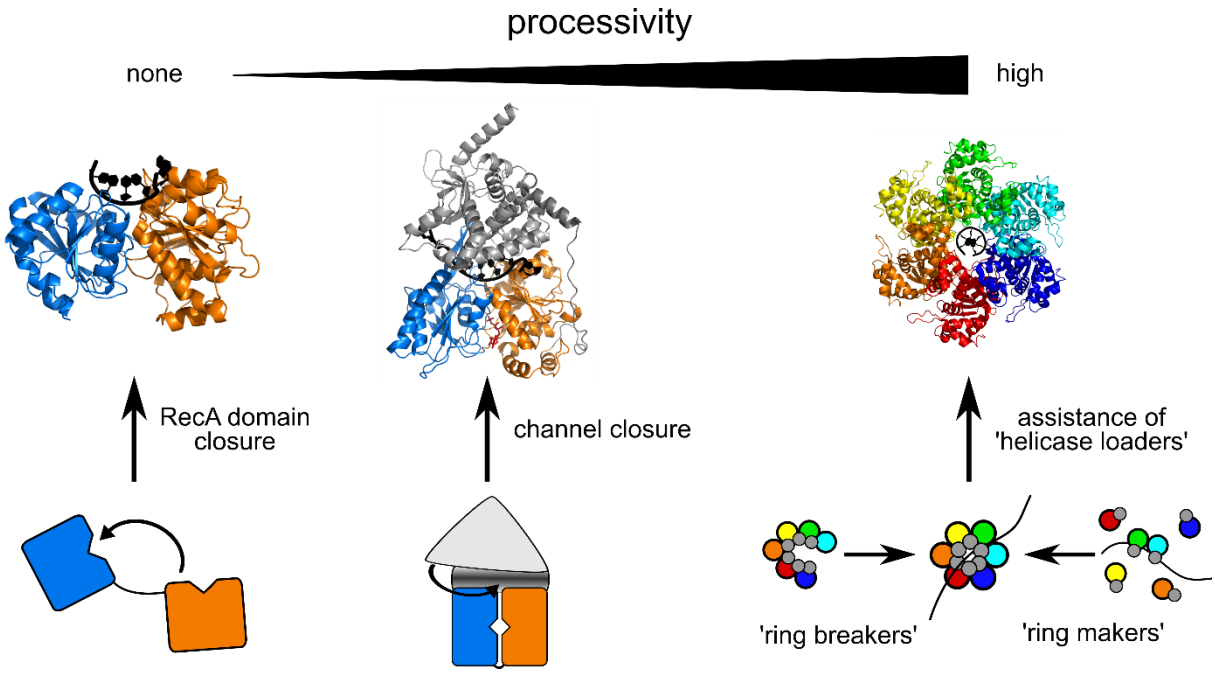


Figure 4.2.: Relationship between helicase architecture and processivity. Crystal structures of the DEAD-box helicase eIF4A-III in complex with U_6-RNA (PDB: 2hyi), the DEAH-box helicase Prp43 in complex with ssRNA and ADP-BeF_3^- (PDB: 5lta) and the hexameric helicase E1 in complex with ssDNA (PDB: 2gxa) are shown in cartoon representation. Non-processive DEAD-box helicases bind RNA upon closure of the RecA domains (orange and blue). Processive monomeric helicases bind their nucleic acid substrate in a channel between the helicase core and auxiliary C-terminal domains (grey), loading requires conformational changes of the binding channel. Hexameric helicases form a ring around their DNA substrate, loading is assisted by cofactor proteins that help to assemble or open the ring.

4.3 Double strand separation by Prp43

To actively separate double strands, different mechanisms are imaginable. The helicase could spread out over the duplex region and destabilize it electro-statically or it could directly form destabilizing interactions with the bases. Alternatively, force or torque might be applied on the double strand to disrupt it (Pyle, 2008). This is most likely the mechanism by which Prp43 achieved partial unwinding of a RNA double strand in the presented work. The melting of the first ~5 base pairs was only observed in the presence of Pfa1(gp) (Chapter 2), implying that the stable ssRNA binding conferred by the G-patch is crucial for Prp43 to productively interact with a double strand. The power stroke for strand separation is likely provided by the domain closure upon ATP binding, where RecA2 shifts towards the RecA1 domain, leading to repositioning of the RNA by 1 nt in 3'-direction through the binding channel. However, the rotation of RecA2 with respect to RecA1 upon ATP hydrolysis might also be involved in the application of force to the adjacent duplex as the experimental setup did not allow to distinguish between the effects of binding ATP and the first cycle of turnover. It is also unclear how strand separation proceeds after the disruption of the first ~5 base pairs observed here. Further unwinding that occurred under continuous ATP hydrolysis at physiological rate led to rapid increase of the donor-acceptor distance beyond the FRET range, impeding a detailed characterization. In principle, two scenarios are imaginable. In both, Prp43 would translocate along the newly freed strand until it encounters the junction between single strand and double strand. Then it could either again disrupt a stretch of ~5 base pairs or it might disrupt the remaining duplex by opening 1 base pair at a time in tandem with its proposed translocation rate of 1 nt per hydrolyzed ATP. The easiest way to approach this question would be by designing a smFRET probe with labels positioned in such a way that the distance increase during unwinding would stay in the detectable range. A similar strategy was used to study the unwinding step-size of the hepatitis C virus helicase NS3. Akin to DEAH-box helicases, NS3 is thought to translocate along the phosphate backbone in single-nucleotide steps. However, a smFRET study revealed that it opens DNA duplexes in three-base-pair increments through a 'spring-loaded' mechanism (Myong *et al*, 2007). While the RecA domains translocate in 1 nt increments, the C-terminal domain remains in fixed position on the DNA, building up tension that is released by a sudden burst of movement after 3-nt translocation (Myong *et al*, 2007). Such mechanisms can enable helicases to overcome large energetic barriers that require the chemical energy from more than one ATP hydrolysis event (Mohapatra *et al*, 2019).

To fully understand how Prp43 acts in strand separation, not only its unwinding step size but also its mode of interaction with the double strand are of interest. Since no structural data on Prp43 or any other DEAH-box helicase in complex with a dsRNA are available so far, putative

mechanisms can only be based on comparisons to more distant relatives and functional studies. In the archaeal SF2 DNA helicase Hel308, which is implicated in DNA repair, a prominent β-hairpin loop in the RecA2 domain was identified as strand separation element (Figure 4.3a) (Büttner *et al*, 2007). However, in recent crystal structures of DEAH-box helicase/G-patch complexes this β-hairpin is buried by the G-patch motif and therefore unlikely to act in strand separation (Studer *et al*, 2020; Hamann *et al*, 2020). In the structurally well characterized SF1 helicases, a domain associated C-terminal to the helicase core contacts the double strand and ATP-dependent unwinding of the duplex is achieved by stepwise rotation and translation (Figure 4.3b). Here, ATP binding induces conformational changes of the interaction surface, destabilizing pairing of the strands (Singleton *et al*, 2007; Soultanas & Wigley, 2001). Auxiliary domains making contacts to structured nucleic acids are also observed within the DEAH-box family. A prominent example is the N-terminal DSM domain in DHX36, which is indispensable for the recognition and subsequent destabilization of G-quadruplex structures (Figure 4.3c) (Chen *et al*, 2018). However, this domain is specific for DHX36 and its function can therefore not be generalized. Based on functional studies and crystal structures of Prp43 in complex with ssRNA, a β-hairpin in the OB-fold was suggested to play a role in strand separation (Walbott *et al*, 2010; Tauchert *et al*, 2017). This structural element is located in close proximity to the bound RNA and is able to undergo pronounced movements depending on the nucleotide state of the helicase (Figure 4.3d) (Tauchert *et al*, 2016, 2017). However, its functional relevance beyond the involvement in the RNA-induced stimulation of ATPase activity was not further investigated in subsequent work (Walbott *et al*, 2010). To unambiguously answer the question which elements of Prp43 are involved in double stand interaction and separation, a crystal structure of the helicase in complex with a dsRNA is strictly necessary.

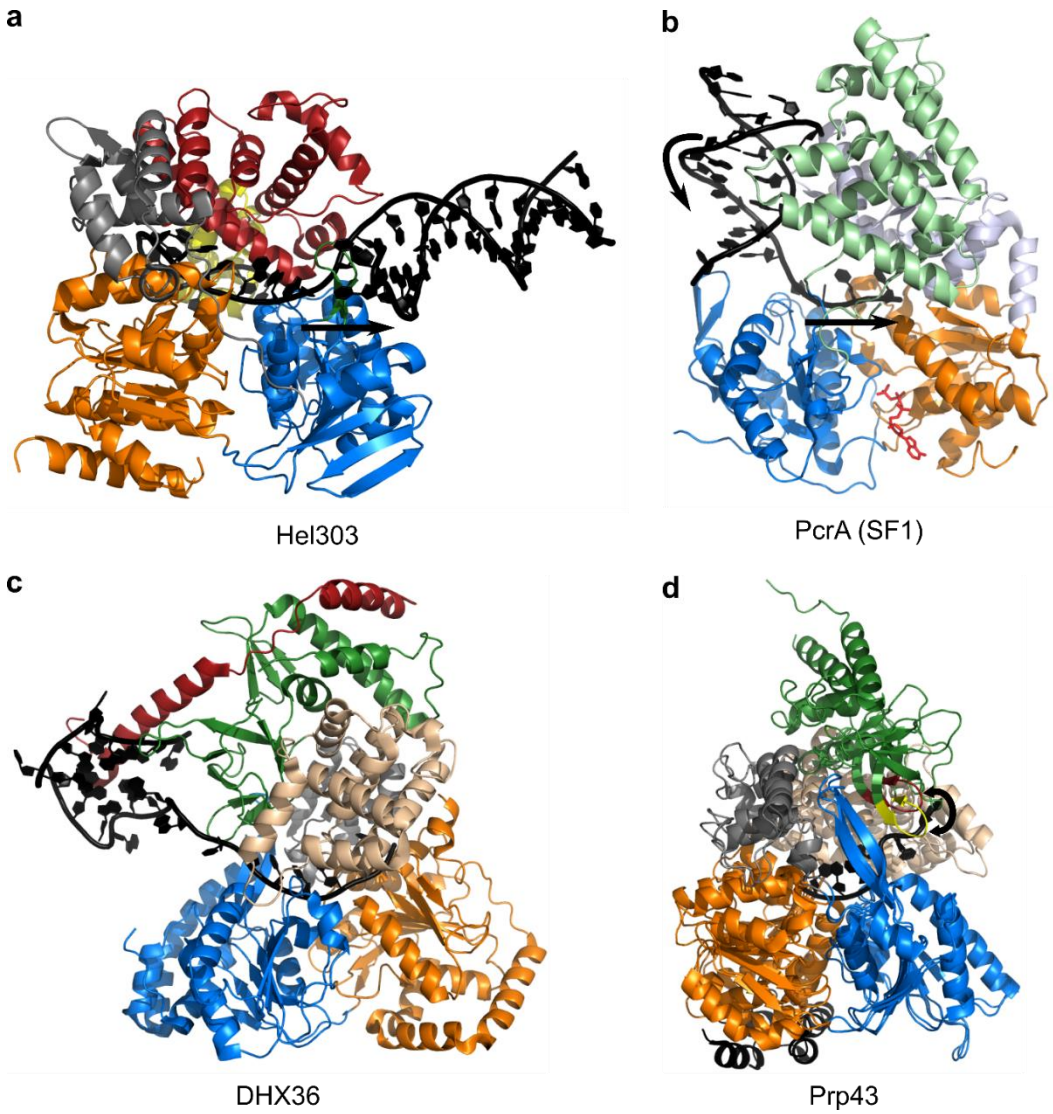


Figure 4.3.: Strand separation mechanisms. Crystal structures of various helicases in complex with nucleic acid substrates are shown in cartoon representation. The RecA domains are colored as before in orange (RecA1) and blue (RecA2). **(a)** Ski2-like helicase Hel308 in complex with a partially unwound DNA-duplex substrate (PDB: 2p6r). The C-terminal WH, ratchet and helix-loop-helix domains are colored in grey, red, and yellow respectively. The RecA2 β-hairpin loop acting as strand separator is highlighted in green. **(b)** SF1 helicase PcrA in complex with a 3'-tailed DNA-duplex (PDB: 3pjr). Helicase core domains 1A and 2A are colored in orange and blue, respectively. The C-terminal domains 1B (light blue) and 2B (light green) interact with the double strand and destabilize it upon translocation of the helicase. **(c)** DEAH/RHA helicase DHX36 in complex with a 3'-tailed G-quadruplex DNA (PDB: 5vhe). The C-terminal WH, ratchet-like and OB domains are colored in grey, wheat and green, respectively. The N-terminal DSM domain (red) is necessary for G-quadruplex destabilization. **(d)** Structures of Prp43 in complex with U_7-RNA and ADP-BeF_3^- (PDB: 5lta) and in complex with ADP (PDB: 5d0u) superposed *via* the RecA2 domain. The C-terminal WH, HB and OB domains are colored in grey, wheat and green, respectively. The nucleotide dependent mobility of the OB β-hairpin that might be involved in strand separation is highlighted in red (ADP-BeF_3^-) and yellow (ADP).

4.4 Conformational regulation of helicases

The regulation of Prp43 by G-patch partners is achieved by the modulation of two major aspects, namely the acceleration of ADP release and the stabilization of RNA binding, which in concert enable Prp43 to act as a processive helicase. This is achieved by both altering the dynamic of the RNA binding channel and controlling the motility of the RecA domains. Generally, this concept can be described as conformational regulation and is a common theme for helicase cofactors. Examples can be found among the regulators of the structurally well characterized DEAD-box helicases which often act by modulating the transition between open and closed conformations of the RecA domains (Ozgur *et al*, 2015). MIF4G (middle domain of eIF4G) domain proteins are known to contact their helicase partners at two interfaces on the helicase core (Schütz *et al*, 2008; Mathys *et al*, 2014) (Figure 4.4a). Similar to G-patch proteins, they combine a high-affinity binding site with a site of lower affinity, allowing one RecA domain to dynamically detach and move towards its counterpart during ATP binding (Hilbert *et al*, 2011). Most commonly, the interaction with a MIF4G domain protein stabilizes a half-open state, pre-aligning the RecA domains for productive ATP and RNA interaction (Hilbert *et al*, 2011; Mathys *et al*, 2014; Alexandrov *et al*, 2011; Sloan & Bohnsack, 2018)). However, they can also act as inhibitors by stabilizing an open unproductive conformation (Buchwald *et al*, 2013) (Figure 4.4b). In an alternative inhibitory mechanism, product release is prevented by locking the RecA domains in the closed state (Ballut *et al*, 2005) (Figure 4.4b). Interestingly, conformational regulation is not limited to modulating the helicase core, as the interaction with cofactor proteins has been shown to induce conformational changes of inhibitory domains resulting in stimulatory effects (Chakrabarti *et al*, 2011; Montpetit *et al*, 2011) (Figure 4.4c). While the stabilization of certain domain conformations can be seen as an indirect way to influence helicase activity, cofactors can also intervene more directly in substrate interaction. They are able to inhibit a helicase by competing with the RNA substrate for its binding site or they can alter the RNA binding properties by contributing either favorable or unfavorable electrostatic interactions (Chang *et al*, 2009; Von Moeller *et al*, 2009; Bono *et al*, 2006; Sharif *et al*, 2013) (Figure 4.4d and 4.4e).

Structural studies have proven extremely valuable in the characterization of the cofactor interactions described above. However, the implications for the motility of the involved domains can only be fully understood by applying methods that allow real-time observations of dynamic motions as has been done in the work presented here. Therefore, the next section will explore how single-molecule methods can deepen our understanding of conformational dynamics and their modulation by cofactors.

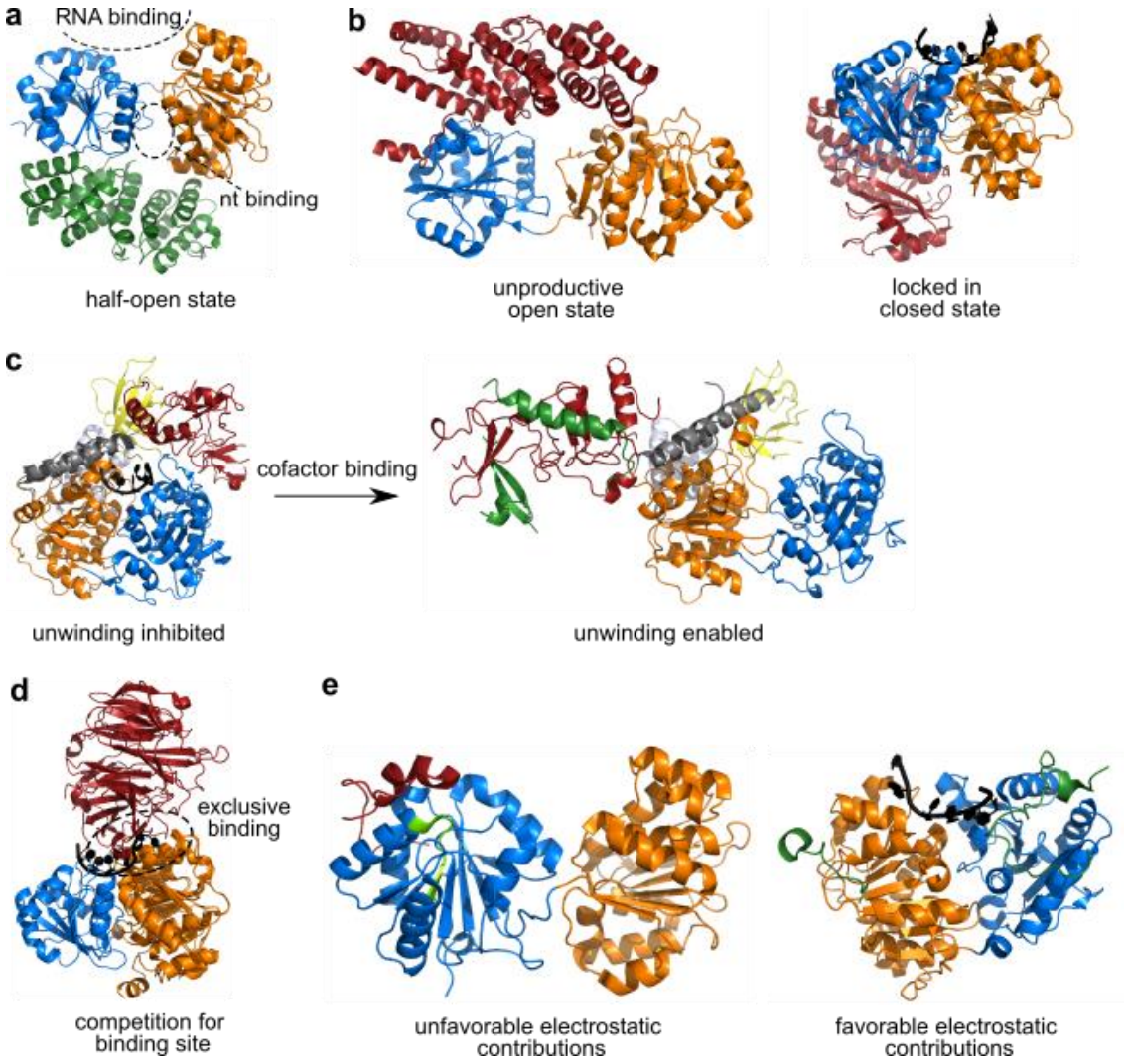


Figure 4.4.: Cofactor regulation of RNA helicases. Crystal structure of various RNA helicases in complex with their cofactors are shown in cartoon representation. The RecA domains are colored as before in orange (RecA1) and blue (RecA2). Activating cofactors are shown in green and inhibiting cofactors in red. **(a)** DDX6 in complex with the MIF4G domain of CNOT1 (PDB: 4ct4). The half-open RecA domain conformation allows efficient binding of ATP and RNA. **(b)** eIF4A-III in complex with the MIF4G domain of CWC22 (PDB: 4c9b, left) and with Mago and Y14 (both red) in the context of the exon junction complex (PDB: 2jos, right). The RecA domains adopt an unproductive open conformation or are locked in the closed state. **(c)** Upf1 in complex with ssRNA (PDB: 2xzl, left) and its cofactor Upf2 (PDB: 2wjv, right). The C-terminal 1B, 1C and stalk domains are colored in yellow, light blue and grey, respectively. In absence of Upf2, unwinding is inhibited by the C-terminal CH domain (red). Binding of Upf2 (green) induces conformational changes that enable unwinding. **(d)** Dbp5 in complex with ssRNA (PDB: 3fht) or Nup214 (PDB: 3fhc) superposed *via* the RecA2 domain. Binding of the cofactor and ssRNA is mutually exclusive as they occupy overlapping binding sites. **(e)** Dhh1 in complex with a fragment of Pat1 (PDB: 4brw, left) and eIF4A-III in complex with Btz in the context of the exon junction complex (PDB: 2jos, right). The Pat1 fragment brings negatively charged residues in proximity to the RNA interacting motif V (light green) in the RecA2 domain and competes with RNA binding, while Btz supports RNA binding by providing positively charged residues.

4.5 Combining structural data and smFRET to characterize helicase domain motility

Structural information on helicases obtained by X-ray crystallography provides insights in the overall domain architecture, the various conformations they can occupy and molecular details of the interactions involved in substrate binding. However, each crystal structure only represents a static snapshot of a conformation adopted under the selected conditions. Mostly, theses snapshots show states that are also relevant under physiological conditions and a good representation of the protein in solution. Nevertheless, certain conformations may be captured with higher probability than others due to the applied crystallization conditions or favorable contacts to other protein molecules within the crystal, while rare but functionally relevant states might be missed. Here lies a major strength of monitoring domain motility by smFRET. This method is not only capable of representing dynamic motions in real-time, but also of visualizing conformations that might be inaccessible otherwise. Combining information gained from a sufficient amount of well distributed FRET pairs and comparing them to available crystal structures even allows modeling of previously unobserved domain conformations (Hilbert *et al*, 2011). Another advantage of observing domain motility by smFRET is that the conformational distribution of the population becomes accessible. The dynamic equilibrium of the RNA binding channel described in chapter 3 is well suited to illustrate this point. The observation of both closed and open channel conformations in crystal structures of DEAH-box helicases in the apo and ATP- analog bound states already suggested the existence of an equilibrium between the conformations (Hamann *et al*, 2019; Roychowdhury *et al*, 2019; Chen *et al*, 2018; Tauchert *et al*, 2017). However, it is impossible to deduce which (if any) conformation is preferentially occupied or if transitions exist between them from crystal structures alone. Here, smFRET allowed to meaningfully characterize this equilibrium and its modulation by Prp43 binding partners. Another example is the G-patch induced opening of the RecA domains described in chapter 2. All crystal structures of DEAH-box helicases in complex with their G-patch partners show a nucleotide free conformation of the RecA domains that is only slightly more open than in the ADP-bound state (Studer *et al*, 2020; Hamann *et al*, 2020). The binding mode of the G-patch suggested its ability to accommodate the conformational range needed for translocation, but it was inconceivable that it would actually promote domain opening. Interestingly, the closed state observed in the smFRET experiment under these conditions is in line with the crystal structures as the donor acceptor distance is slightly increased in comparison to the ADP bound state (Figure 4.5a).

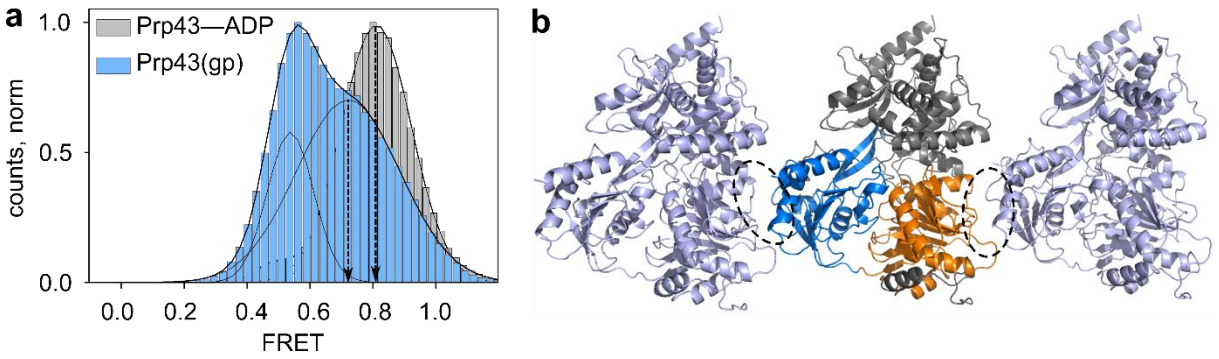


Figure 4.5.: Closed RecA domain conformation in the G-patch complex. (a) Overlay of 2D histograms showing the distribution of FRET values of Prp43—ADP and the Prp43-Pfa1(gp) complex. The center of the closed state shifts from E_{FRET}=0.81 (ADP) to E_{FRET}=0.75 (Pfa1(gp)), indicating a slight opening of the RecA domains. **(b)** Crystal structure of DHX15 in complex with NKRF(gp) (PDB: 6sh7) and two selected symmetry mates (light blue) in cartoon representation. Crystal contacts between the RecA domains of neighboring DHX15 molecules are highlighted.

Nevertheless, so far only the single molecule approach applied in this work was able to detect an open RecA domain conformation of Prp43 in absence of RNA. The preferable crystallization of the closed over the open state might be explained by crystal contacts to neighboring helicase molecules that stabilize the observed conformation (Figure 4.5b).

While smFRET studies utilizing labeled protein domains have been widely applied to understand the mechanism of action of DEAD-box helicases, comparable data on other RNA helicases are sparse. Most available single molecule studies rather focus on observations of labeled nucleic acid substrates. Here, knowledge on properties such as unwinding step size or extend of substrate rearrangement could be gained (Rodgers *et al*, 2016; Semlow *et al*, 2016; Myong *et al*, 2007). Nevertheless, the presented work illustrates that the potential of smFRET experiments on labeled helicase molecules is huge, as they can provide otherwise inaccessible insights into mechanistic properties and their regulation. Comprehensive structural characterization, which is an essential prerequisite to design meaningful smFRET experiments, is available for various RNA helicases. Thereby, a basis is formed to deepen our understanding of the function and regulation of this omnipresent class of enzymes.

Chapter 5: Abstract

RNA helicases of the DEAH-box family fulfill essential functions in various aspects of RNA metabolism by reorganizing structured RNAs and RNPs through ATP-dependent motility in 3' to 5' direction along single-strands. Prp43 acts on several targets in pre-mRNA splicing and ribosome biogenesis. In the specific cellular context Prp43 is activated by G-patch factors which enhance its intrinsically low ATPase and RNA unwinding activity. To initiate unwinding, a single stranded RNA segment preceding the helicase target needs to be accommodated in a binding channel formed between the helicase core and the C-terminal domains. When Prp43 loads onto its substrate, the channel is thought to widen into a groove and fold around the RNA, but how the loading process is regulated by the substrates, ATP and RNA, and the gp binding partners remains unknown. While structural studies revealed that the gp motif connects the flexible C-terminal domains to the helicase core, the impact on RecA domain movement during catalysis, conformation of the binding channel and the unwinding of RNA substrates is unclear. In this thesis, a single-molecule approach was used to gain time resolved information on the effects of G-patch binding on the conformational dynamics of Prp43 and the interaction with its RNA substrate.

In a first project, a smFRET system reporting on the RecA domain movement of Prp43 gave insights in conformational dynamics during ATP hydrolysis and their modulation by interaction with Pfa1(gp). While the RecA domains adopt a closed conformation in the apo and nucleotide-bound states, addition of Pfa1(gp) induces an open state that is incompatible with nucleotide-binding and facilitates ADP release. Thereby, transitions from the weak (ADP) to the strong (apo) RNA binding state are accelerated. In a second smFRET system, labels were placed on the RNA to probe substrate binding and unwinding. In complex with Pfa1(gp), Prp43(ADP) switches between bound and unbound states instead of dissociating from the RNA, while the closure of RecA domains upon ATP binding generates sufficient force to separate proximal RNA structures. During ATP turnover, Pfa1(gp) accelerates the conformational cycling of the RecA domains. Thereby, processive movement along the substrate is enabled, as translocation becomes faster than drop-off from the RNA.

In a second project, a different smFRET system was used to study the conformational dynamics of the RNA binding channel. In the absence of binding partners, the channel alternates between open and closed conformations, but is preferentially closed. Binding of ATP and Pfa1(gp) shifts the equilibrium towards the open state, facilitating the accommodation of RNA. After the initial loading step, the channel remains firmly closed during continuous cycles of ATP hydrolysis, enabling processive translocation. In contrast, no efficient RNA loading takes place in absence of Pfa1(gp).

Combining observations from all systems revealed how G-patch partners regulate the interaction of Prp43 with its substrates ATP and RNA by modulating the motility of its RecA domains and RNA binding channel. This led to a mechanistic model of its motility cycle.

Bibliography

Abovich N & Rosbash M (1997) Cross-Intron Bridging Interactions in the Yeast Commitment Complex Are Conserved in Mammals. *Cell* 89: 403–412

Absmeier E, Santos KF & Wahl MC (2016) Functions and regulation of the Brr2 RNA helicase during splicing. *Cell Cycle* 15: 3362-3377

Absmeier E, Santos KF & Wahl MC (2020) Molecular Mechanism Underlying Inhibition of Intrinsic ATPase Activity in a Ski2-like RNA Helicase. *Structure* 28: 236-243.e3

Adio S, Senyushkina T, Peske F, Fischer N, Wintermeyer W & Rodnina MV (2015) Fluctuations between multiple EF-G-induced chimeric tRNA states during translocation on the ribosome. *Nat Commun* 6: 1–11

Agarwal KC, Miech RP & Parks RE (1978) Guanylate Kinases from Human Erythrocytes, Hog Brain, and Rat Liver. *Methods Enzymol* 51: 483–490

Alexandrov A, Colognori D, Shu M Di & Steitz JA (2012) Human spliceosomal protein CWC22 plays a role in coupling splicing to exon junction complex deposition and nonsense-mediated decay. *Proc Natl Acad Sci USA* 109: 21313–21318

Alexandrov A, Colognori D & Steitz JA (2011) Human eIF4AIII interacts with an eiF4G-like partner, NOM1, revealing an evolutionarily conserved function outside the exon junction complex. *Genes Dev* 25: 1078–1090

Andersen CBF, Ballut L, Johansen JS, Chamieh H, Nielsen KH, Oliveira CLP, Pedersen JS, Séraphin B, Le Hir H, Gregers †, *et al* (2006) Structure of the Exon Junction Core Complex with a Trapped DEAD-Box ATPase Bound to RNA. *Science (80-)* 313: 1968–1972

Andreou AZ & Klostermeier D (2012) Conformational changes of DEAD-Box helicases monitored by single molecule fluorescence resonance energy transfer 1st ed. Elsevier Inc.

Appleby TC, Anderson R, Fedorova O, Pyle AM, Wang R, Liu X, Brendza KM & Somoza JR (2011) Visualizing ATP-dependent RNA translocation by the NS3 helicase from HCV. *J Mol Biol* 405: 1139–1153

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

Aregger R & Klostermeier D (2009) The DEAD box helicase YxiN maintains a closed conformation during ATP hydrolysis. *Biochemistry* 48: 10679–10681

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

Arnold PR, Wells AD & Li XC (2020) Diversity and Emerging Roles of Enhancer RNA in Regulation of Gene Expression and Cell Fate. *Front Cell Dev Biol* 7 doi:10.3389/fcell.2019.00377

Bai R, Wan R, Yan C, Jia Q, Lei J & Shi Y (2021) Mechanism of spliceosome remodeling by the ATPase/helicase Prp2 and its coactivator Spp2. *Science (80-)* 371: eabe8863

Bai R, Wan R, Yan C, Lei J & Shi Y (2018) Structures of the fully assembled Saccharomyces cerevisiae spliceosome before activation. *Science (80-)* 360: 1423–1429

Bai R, Yan C, Wan R, Lei J & Shi Y (2017) Structure of the Post-catalytic Spliceosome from Saccharomyces cerevisiae. *Cell* 171: 1589-1598.e8

Ballut L, Marchadier B, Baguet A, Tomasetto C, Séraphin B & Le Hir H (2005) The exon junction core complex is locked onto RNA by inhibition of eIF4AIII ATPase activity. *Nat Struct Mol Biol* 12: 861–869

Baßler J & Hurt E (2019) Annual Review of Biochemistry Eukaryotic Ribosome Assembly. *Annu Rev Biochem* 88: 281–306

Beier DH, Carrocci TJ, van der Feltz C, Tretbar US, Paulson JC, Grabowski N & Hoskins AA (2019) Dynamics of the DEAD-box ATPase Prp5 RecA-like domains provide a conformational switch during spliceosome assembly. *Nucleic Acids Res* 47: 10842–10851

Berglund JA & Chua K (1997) The Splicing Factor BBP Interacts Specifically with the Pre-mRNA Branchpoint Sequence UACUAAC. *Cell* 89: 781-787

Bianco PR, Brewer LR, Corzett M, Balhorn R, Yehl Y, Kowlczykowski SC & Baskin RJ (2001) Processive translocation and DNA unwinding by individual RecBCD enzyme molecules. *Nature* 409: 374–378

Bizebard T, Ferlenghi I, Iost I & Dreyfus M (2004) Studies on three E. coli DEAD-box helicases point to an unwinding mechanism different from that of model DNA helicases. *Biochemistry* 43: 7857–7866

Bleichert F & Baserga SJ (2007) The Long Unwinding Road of RNA Helicases. *Mol Cell* 27: 339–352

Boesler C, Rigo N, Anokhina MM, Tauchert MJ, Agafonov DE, Kastner B, Urlaub H, Ficner R, Will CL & Lührmann R (2016) A spliceosome intermediate with loosely associated tri-snRNP accumulates in the absence of Prp28 ATPase activity. *Nat Commun* 7

Bohnsack KE, Ficner R, Bohnsack MT & Jonas S (2021) Regulation of DEAH-box RNA helicases by G-patch proteins. *Biol Chem* 402: 561–579

Bohnsack MT, Martin R, Granneman S, Ruprecht M, Schleiff E & Tollervey D (2009) Prp43 Bound at Different Sites on the Pre-rRNA Performs Distinct Functions in Ribosome Synthesis. *Mol Cell* 36: 583–592

Boneberg FM, Brandmann T, Kobel L, Van Den Heuvel J, Bargsten K, Bammert L, Kutay U & Jinek M (2019) Molecular mechanism of the RNA helicase DHX37 and its activation by UTP14A in ribosome biogenesis. *RNA* 25: 685–701

Bono F, Ebert J, Lorentzen E & Conti E (2006) The Crystal Structure of the Exon Junction Complex Reveals How It Maintains a Stable Grip on mRNA. *Cell* 126: 713–725

De Bortoli F, Espinosa S & Zhao R (2021) DEAH-Box RNA Helicases in Pre-mRNA Splicing. *Trends Biochem Sci* 46: 225–238

Bronson JE, Fei J, Hofman JM, Gonzalez RL & Wiggins CH (2009) Learning rates and states from biophysical time series: A Bayesian approach to model selection and single-molecule FRET data. *Biophys J* 97: 3196–3205

Buchwald G, Schüssler S, Basquin C, Le Hir H & Conti E (2013) Crystal structure of the human eIF4AIII-CWC22 complex shows how a DEAD-box protein is inhibited by a MIF4G domain. *Proc Natl Acad Sci U S A* 110

Burgess SM & Guthrie C (1993) A mechanism to enhance mRNA splicing fidelity: The RNA-dependent ATPase Prp16 governs usage of a discard pathway for aberrant lariat intermediates. *Cell* 73: 1377–1391

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

Del Campo M & Lambowitz AM (2009) Structure of the Yeast DEAD Box Protein Mss116p Reveals Two Wedges that Crimp RNA. *Mol Cell* 35: 598–609

Cao W, Coman MM, Ding S, Henn A, Middleton ER, Bradley MJ, Rhoades E, Hackney DD, Pyle AM & De La Cruz EM (2011) Mechanism of Mss116 ATPase reveals functional diversity of DEAD-Box proteins. *J Mol Biol* 409: 399–414

Caruthers JM, Johnson ER & Mckay DB (2000) Crystal structure of yeast initiation factor 4A, a DEAD-box RNA helicase. *PNAS* 97: 13080–13085

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

Chakrabarti S, Jayachandran U, Bonneau F, Fiorini F, Basquin C, Domcke S, Le Hir H & Conti E (2011) Molecular Mechanisms for the RNA-Dependent ATPase Activity of Upf1 and Its Regulation by Upf2. *Mol Cell* 41: 693–703

Chang JH, Cho YH, Sohn SY, Choi JM, Kim A, Kim YC, Key Jang S & Cho Y (2009) Crystal structure of the eIF4A-PDCD4 complex. *PNAS* 106: 3148–3153

Chen JY-F, Stands L, Staley JP, Jackups RR, Latus LJ & Chang T-H (2001) Specific Alterations of U1-C Protein or U1 Small Nuclear RNA Can Eliminate the Requirement of Prp28p, an Essential DEAD Box Splicing Factor. *Molecular Cell* 7: 227-232

Chen MC, Tippana R, Demeshkina NA, Murat P, Balasubramanian S, Myong S & Ferré-D'amaré AR (2018) Structural basis of G-quadruplex unfolding by the DEAH/RHA helicase DHX36. *Nature* 558: 465–483

Chen Y-LL, Capeyrou R, Humbert O, Mouffok S, Kadri Y Al, Lebaron S, Henras AK & Henry Y (2014) The telomerase inhibitor Gno1p/PINX1 activates the helicase Prp43p during ribosome biogenesis. *Nucleic Acids Res* 42: 7330–7345

Chen Z, Gui B, Zhang Y, Xie G, Li W, Liu S, Xu B, Wu C, He L, Yang J, *et al* (2017) Identification of a 35S U4/U6.U5 tri-small nuclear ribonucleoprotein (tri-snRNP) complex intermediate in spliceosome assembly. *J Biol Chem* 292: 18113–18123

Cheng Z, Coller J, Parker R & Song H (2005) Crystal structure and functional analysis of DEAD-box protein Dhh1p. *RNA* 11: 1258–1270

Chiu Y-F, Liu Y-C, Chiang T-W, Yeh T-C, Tseng C-K, Wu N-Y & Cheng S-C (2009) Cwc25 Is a Novel Splicing Factor Required after Prp2 and Yju2 To Facilitate the First Catalytic Reaction. *Mol Cell Biol* 29: 5671–5678

Christian H, Hofele RV., Urlaub H & Ficner R (2014) Insights into the activation of the helicase Prp43 by biochemical studies and structural mass spectrometry. *Nucleic Acids Res* 42: 1162–1179

Collins R, Karlberg T, Lehtiö L, Schütz P, van den Berg S, Dahlgren LG, Hammarström M, Weigelt J & Schüler H (2009) The DEXD/H-box RNA helicase DDX19 is regulated by an α-helical switch. *J Biol Chem* 284: 10296–10300

Company M, Arenas J & Abelson J (1991) Requirement of the RNA helicase-like protein PRP22 for release of messenger RNA from spliceosomes. *Nature* 349: 487–493

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

Cordin O & Beggs JD (2013) RNA helicases in splicing. *RNA Biol* 10: 83–95

Cordin O, Hahn D & Beggs JD (2012) Structure, function and regulation of spliceosomal RNA helicases. *Curr Opin Cell Biol* 24: 431–438

Davey MJ & O'Donnell M (2003) Replicative helicase loaders: Ring breakers and ring makers. *Curr Biol* 13: 594-596

Decatur WA & Fournier MJ (2003) RNA-guided nucleotide modification of ribosomal and other RNAs. *J Biol Chem* 278: 695–698

Donsbach P, Yee BA, Sanchez-Hevia D, Berenguer J, Aigner S, Yeo GW & Klostermeier D (2020) The thermus thermophilus DEAD-box protein Hera is a general RNA binding protein and plays a key role in tRNA metabolism. *RNA* 26: 1557–1574

Enders M, Ficner R & Adio S (2022) Regulation of the DEAH/RHA helicase Prp43 by the G-patch factor Pfa1. *under Rev PNAS*

Erkizan HV, Schneider JA, Sajwan K, Graham GT, Griffin B, Chasovskikh S, Youbi SE, Kallarakal A, Chruszcz M, Padmanabhan R, *et al* (2015) RNA helicase A activity is inhibited by oncogenic transcription factor EWS-FLI1. *Nucleic Acids Res* 43: 1069–1080

Fabrizio P, Dannenberg J, Dube P, Kastner B, Stark H, Urlaub H & Lührmann R (2009) The Evolutionarily Conserved Core Design of the Catalytic Activation Step of the Yeast Spliceosome. *Mol Cell* 36: 593–608

Fairman-Williams ME, Guenther UP & Jankowsky E (2010) SF1 and SF2 helicases: Family matters. *Curr Opin Struct Biol* 20: 313–324

Fica SM, Oubridge C, Galej WP, Wilkinson ME, Bai XC, Newman AJ & Nagai K (2017) Structure of a spliceosome remodelled for exon ligation. *Nature* 542: 377–380

Förster T (1946) Energiewanderung und Fluoreszenz. *Naturwissenschaften* 33: 166–175

Fouraux MA, Kolkman MJ, Van Der Heijden A, De Jong AS, Van Venrooij WJ & Pruijn GJ (2002) The human La (SS-B) autoantigen interacts with DDX15/hPrp43, a putative DEAH-box RNA helicase. *RNA* 8: 1428–1443

Fourmann J-BB, Tauchert MJ, Ficner R, Fabrizio P & Lührmann R (2017) Regulation of Prp43-mediated disassembly of spliceosomes by its cofactors Ntr1 and Ntr2. *Nucleic Acids Res* 45: 4068–4080

Fourmann JB, Dybkov O, Agafonov DE, Tauchert MJ, Urlaub H, Ficner R, Fabrizio P & Lührmann R (2016) The target of the DEAH-box NTP triphosphatase Prp43 in saccharomyces cerevisiae spliceosomes is the U2 snRNP-intron interaction. *Elife* 5: e15564

Fromm L, Falk S, Flemming D, Schuller JM, Thoms M, Conti E & Hurt E (2017) Reconstitution of the complete pathway of ITS2 processing at the pre-ribosome. *Nat Commun* 8

Galej WP, Wilkinson ME, Fica SM, Oubridge C, Newman AJ & Nagai K (2016) Cryo-EM structure of the spliceosome immediately after branching. *Nature* 537: 197–201

Gilman B, Tijerina P & Russell R (2017) Distinct RNA-unwinding mechanisms of DEAD-box and DEAH-box RNA helicase proteins in remodeling structured RNAs and RNPs. *Biochem Soc Trans* 45: 1313–1321

Gowravaram M, Bonneau F, Kanaan J, Maciej VD, Fiorini F, Raj S, Croquette V, Le Hir H & Chakrabarti S (2018) A conserved structural element in the RNA helicase UPF1 regulates its catalytic activity in an isoform-specific manner. *Nucleic Acids Res* 46: 2648–2659

Grandi P, Rybin V, Baßler J, Petfalski E, Strauß D, Marzioch M, Schä T, Kuster B, Tschochner H, Tollervey D, *et al* (2002) 90S Pre-Ribosomes Include the 35S Pre-rRNA, the U3 snoRNP, and 40S Subunit Processing Factors but Predominantly Lack 60S Synthesis Factors. *Molecular Cell* 10: 105-115

Granneman S, Lin CY, Champion EA, Nandineni MR, Zorca C & Baserga SJ (2006) The nucleolar protein Esf2 interacts directly with the DExD/H box RNA helicase, Dbp8, to stimulate ATP hydrolysis. *Nucleic Acids Res* 34: 3189–3199

Guy CP & Bolt EL (2005) Archaeal Hel308 helicase targets replication forks in vivo and in vitro and unwinds lagging strands. *Nucleic Acids Res* 33: 3678–3690

Hamann F, Enders M & Ficner R (2019) Structural basis for RNA translocation by DEAH-box ATPases. *Nucleic Acids Res* 47: 4349–4362

Hamann F, Schmitt A, Favretto F, Hofele R, Neumann P, Xiang SQ, Urlaub H, Zweckstetter M & Ficner R (2020) Structural analysis of the intrinsically disordered splicing factor Spp2 and its binding to the DEAH-box ATPase Prp2. *Proc Natl Acad Sci U S A* 117: 2948–2956

Hamann F, Zimmerningkat LC, Becker RA, Garbers TB, Neumann P, Hub JS & Ficner R (2021) The structure of Prp2 bound to RNA and ADP-BeF3-reveals structural features important for RNA unwinding by DEAH-box ATPases. *Acta Crystallogr Sect D Struct Biol* 77: 496–509

He Y, Andersen GR & Nielsen KH (2010) Structural basis for the function of DEAH helicases. *EMBO Rep* 11: 180–186

He Y, Staley JP, Andersen GR & Nielsen KH (2017) Structure of the DEAH/RHA ATPase Prp43p bound to RNA implicates a pair of hairpins and motif Va in translocation along RNA. *Rna* 23: 1110–1124

Hegele A, Kamburov A, Grossmann A, Sourlis C, Wowro S, Weimann M, Will CL, Pena V, Lührmann R & Stelzl U (2012) Dynamic Protein-Protein Interaction Wiring of the Human Spliceosome. *Mol Cell* 45: 567–580

Heininger AU, Hackert P, Andreou AZ, Boon KL, Memet I, Prior M, Clancy A, Schmidt B, Urlaub H, Schleiff E, *et al* (2016) Protein cofactor competition regulates the action of a multifunctional RNA helicase in different pathways. *RNA Biol* 13: 320–330

Henras AK, Plisson-Chastang C, O'Donohue MF, Chakraborty A & Gleizes PE (2015) An overview of pre-ribosomal RNA processing in eukaryotes. *Wiley Interdiscip Rev RNA* 6: 225–242

Hilbert M, Karow AR & Klostermeier D (2009) The mechanism of ATP-dependent RNA unwinding by DEAD box proteins. *Biol Chem* 390: 1237–1250

Hilbert M, Kebbel F, Gubaev A & Klostermeier D (2011) eIF4G stimulates the activity of the DEAD box protein eIF4A by a conformational guidance mechanism. *Nucleic Acids Res* 39: 2260–2270

Van Hoof A, Lennertz P & Parker R (2000) Yeast Exosome Mutants Accumulate 3-Extended Polyadenylated Forms of U4 Small Nuclear RNA and Small Nucleolar RNAs. *Molecular and Cellular Biology* 20: 441-452

Hoskins AA & Moore MJ (2012) The spliceosome: A flexible, reversible macromolecular machine. *Trends Biochem Sci* 37: 179–188

Inesta-Vaquera F, Chaugule VK, Galloway A, Chandler L, Rojas-Fernandez A, Weidlich S, Peggie M & Cowling VH (2018) DHX15 regulates CMTR1-dependent gene expression and cell proliferation. *Life Sci Alliance* 1: e201800092

Iyer LM, Leipe DD, Koonin E V. & Aravind L (2004) Evolutionary history and higher order classification of AAA+ ATPases. In *Journal of Structural Biology* 146: 11-31

Jacobs AMF, Nicol SM, Hislop RG, Jaffray EG, Hay RT & Fuller-Pace F V. (2007) SUMO modification of the DEAD box protein p68 modulates its transcriptional activity and promotes its interaction with HDAC1. *Oncogene* 26: 5866–5876

Jarmoskaite I & Russell R (2014) RNA helicase proteins as chaperones and remodelers. *Annu Rev Biophys* 83: 697–725

Johnson SJ & Jackson RN (2013) Ski2-like RNA helicase structures: Common themes and complex assemblies. *RNA Biol* 10: 33–43 doi:10.4161/rna.22101 [PREPRINT]

Kim S-H & Lin R-J (1996) Spliceosome Activation by PRP2 ATPase prior to the First Transesterification Reaction of Pre-mRNA Splicing. *Mol Cell Biol* 16: 6810–6819

Kiss T (2002) Small Nucleolar RNAs: An Abundant Group of Noncoding RNAs with Diverse Cellular Functions. *Cell* 109: 145–148

Kistler AL & Guthrie C (2001) Deletion of MUD2, the yeast homolog of U2AF65, can bypass the requirement for Sub2, an essential spliceosomal ATPase . *Genes Dev* 15: 42–49

Konarska MM, Vilardell J & Query CC (2006) Repositioning of the reaction intermediate within the catalytic center of the spliceosome. *Mol Cell* 21: 543–553

Koodathingal P, Novak T, Piccirilli JA & Staley JP (2010) The DEAH Box ATPases Prp16 and Prp43 Cooperate to Proofread 5' Splice Site Cleavage during Pre-mRNA Splicing. *Mol Cell* 39: 385–395

Koš M & Tollervey D (2005) The putative RNA helicase Dbp4p is required for release of the U14 snoRNA from preribosomes in Saccharomyces cerevisiae. *Mol Cell* 20: 53–64

Kossen K & Uhlenbeck OC (1999) Cloning and biochemical characterization of Bacillus subtilis YxiN, a DEAD protein specifically activated by 23S rRNA: delineation of a novel sub-family of bacterial DEAD proteins. *NAR* 27: 3811-3820

Kretschmer J, Rao H, Hackert P, Sloan KE, Höbartner C & Bohnsack MT (2018) The m 6 A reader protein YTHDC2 interacts with the small ribosomal subunit and the 5 ′ -3 ′ exoribonuclease XRN1. *RNA* 24: 1339–1350

Krishnan R, Blanco MR, Kahlscheuer ML, Abelson J, Guthrie C & Walter NG (2013) Biased Brownian ratcheting leads to pre-mRNA remodeling and capture prior to first-step splicing. *Nat Struct Mol Biol* 20: 1450–1457

Laggerbauer B, Achsel T & Lührmann RL (1998) The human U5-200kD DEXH-box protein unwinds U4U6 RNA duplices in vitro. *PNAS* 95: 4188–4192

Lam AMI & Frick DN (2006) Hepatitis C Virus Subgenomic Replicon Requires an Active NS3 RNA Helicase. *J Virol* 80: 404–411

Lardelli RM, Thompson JX, Yates JR & Stevens SW (2010) Release of SF3 from the intron branchpoint activates the first step of pre-mRNA splicing. *RNA* 16: 516–528

Lebaron S, Papin C, Capeyrou R, Chen Y-L, 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: 3808–19

Leitão AL, Costa MC & Enguita FJ (2015) Unzippers, resolvers and sensors: A structural and functional biochemistry tale of RNA helicases. *Int J Mol Sci* 16: 2269–2293

Lerner E, Barth A, Hendrix J, Ambrose B, Birkedal V, Blanchard SC, Börner R, Chung HS, Corde T, Craggs TD et al. (2021) FRET-based dynamic structural biology: Challenges, perspectives and an appeal for open-science practices. *eLife* 10:e60416

Liang X & Fournier MJ (2006) The Helicase Has1p Is Required for snoRNA Release from Pre-rRNA. *Mol Cell Biol* 26: 7437–7450

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 F, Putnam A & Jankowsky E (2008) ATP hydrolysis is required for DEAD-box protein recycling but not for duplex unwinding. *PNAS* 51: 20209–20214

Liu S, Li X, Zhang L, Jiang J, Hill RC, Cui Y, Hansen KC, Zhou ZH & Zhao R (2017) Structure of the yeast spliceosomal postcatalytic P complex. *Science (80-)* 358: 1278–1283

Liu Y-C, Chen H-C, Wu N-Y & Cheng S-C (2007) A Novel Splicing Factor, Yju2, Is Associated with NTC and Acts after Prp2 in Promoting the First Catalytic Reaction of Pre-mRNA Splicing. *Mol Cell Biol* 27: 5403–5413

Lohman TM (1993) Helicase-catalyzed DNA Unwinding*. *J Biol Chem* 268: 2269–2272

Luo D, Ding SC, Vela A, Kohlway A, Lindenbach BD & Pyle AM (2011) Structural insights into RNA recognition by RIG-I. *Cell* 147: 409–422

Mallam AL, Del Campo M, Gilman B, Sidote DJ & Lambowitz AM (2012) Structural basis for RNA-duplex recognition and unwinding by the DEAD-box helicase Mss116p. *Nature* 490: 121–125

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: 17743–50

Martin R, (2014) Functional characterization of ribosome biogenesis cofactors on Saccharomyces cerevisiae. PhD thesis, Georg-August-Universität Göttingen.

Martin R, Hackert P, Ruprecht M, Simm S, Brüning L, Mirus O, Sloan KE, Kudla G, Schleiff E & Bohnsack MT (2014) A pre-ribosomal RNA interaction network involving snoRNAs and the Rok1 helicase. *RNA* 20: 1173–1182

Martin R, Straub AU, Doebele C & Bohnsack MT (2013) DExD/H-box RNA helicases in ribosome biogenesis. *RNA Biol* 10: 4–18

Masters BR (2014) Paths to Förster's resonance energy transfer (FRET) theory. *Eur Phys J H* 39: 87–139

Mathew R, Hartmuth K, Möhlmann S, Urlaub H, Ficner R & Lührmann R (2008) Phosphorylation of human PRP28 by SRPK2 is required for integration of the U4/U6-U5 tri-snRNP into the spliceosome. *Nat Struct Mol Biol* 15: 435–443

Mathys H, Basquin JÔ, Ozgur S, Czarnocki-Cieciura M, Bonneau F, Aartse A, Dziembowski A, Nowotny M, Conti E & Filipowicz W (2014) Structural and Biochemical Insights to the Role of the CCR4-NOT Complex and DDX6 ATPase in MicroRNA Repression. *Mol Cell* 54: 751–765

Mayas RM, Maita H, Semlow DR & Staley JP (2010) Spliceosome discards intermediates via the DEAH box ATPase Prp43p. *Proc Natl Acad Sci* 107: 10020–10025

Mayas RM, Maita H & Staley JP (2006) Exon ligation is proofread by the DExD/H-box ATPase Prp22p. *Nat Struct Mol Biol* 13: 482–490

Memet I, Doebele C, Sloan KE & Bohnsack MT (2017) The G-patch protein NF-κB-repressing factor mediates the recruitment of the exonuclease XRN2 and activation of the RNA helicase DHX15 in human ribosome biogenesis. *Nucleic Acids Res* 45: 5359–5374

Moazed D (2009) Small RNAs in transcriptional gene silencing and genome defence. *Nature* 457: 413–420

Von Moeller H, Basquin C & Conti E (2009) The mRNA export protein DBP5 binds RNA and the cytoplasmic nucleoporin NUP214 in a mutually exclusive manner. *Nat Struct Mol Biol* 16: 247–254

Mohapatra S, Lin C-T, Feng XA, Basu A & Ha T (2019) Single-Molecule Analysis and Engineering of DNA Motors. *Chem Rev*: acs.chemrev.9b00361

Montpetit B, Thomsen ND, Helmke KJ, Seeliger MA, Berger JM & Weis K (2011) A conserved mechanism of DEAD-box ATPase activation by nucleoporins and InsP6 in mRNA export. *Nature* 472: 238–244

Myong S, Bruno MM, Pyle AM & Ha T (2007) Spring-Loaded Mechanism of DNA Unwinding by Hepatitis C Virus NS3 Helicase. *Science (80-)* 317: 513–516

Niu Z, Jin W, Zhang L & Li X (2012) Tumor suppressor RBM5 directly interacts with the DExD/H-box protein DHX15 and stimulates its helicase activity. *FEBS Lett* 586: 977–983

Ohrt T, Odenwälder P, Dannenberg J, Prior M, Warkocki Z, Schmitzová J, Karaduman R, Gregor I, Enderlein J, Fabrizio P, *et al* (2013) Molecular dissection of step 2 catalysis of yeast pre-mRNA splicing investigated in a purified system. *RNA* 19: 902–915

Ozgur S, Buchwald G, Falk S, Chakrabarti S, Prabu JR & Conti E (2015) The conformational plasticity of eukaryotic RNA-dependent ATPases. *FEBS J* 282: 850–863

Pan C, Potratz JP, Cannon B, Simpson ZB, Ziehr JL, Tijerina P & Russell R (2014) DEAD-Box Helicase Proteins Disrupt RNA Tertiary Structure Through Helix Capture. *PLoS Biol* 12

Pandit S, Lynn B & Rymond BC (2006) Inhibition of a spliceosome turnover pathway suppresses splicing defects. *Proc Natl Acad Sci U S A* 103: 13700–13705

Pause A, Methot N & Sonenberg N (1993) The HRIGRXXR Region of the DEAD Box RNA Helicase Eukaryotic Translation Initiation Factor 4A Is Required for RNA Binding and ATP Hydrolysis. *Mol. Cell. Biol.* 13: 6789-6798

Pause A & Sonenberg N (1992) Mutational analysis of a DEAD box RNA helicase: the mammalian translation initiation factor eIF-4A. *EMBO J* 1: 2643–2654

Peck ML & Herschlag D (1999) Effects of oligonucleotide length and atomic composition on stimulation of the ATPase activity of translation initiation factor eIF4A. *RNA* 5: 1210–1221

Pelletier J & Volarevic S (2018) Ribosome Biogenesis in Cancer: New Players and Therapeutic Avenues. *Nat Rev Cancer* 18: 51-63

Pertschy B, Schneider C, Gnädig M, Schäfer T, Tollervey D & Hurt E (2009) RNA helicase Prp43 and its co-factor Pfa1 promote 20 to 18 S rRNA processing catalyzed by the endonuclease Nob1. *J Biol Chem* 284: 35079–91

Potratz JP, Del Campo M, Wolf RZ, Lambowitz AM & Russell R (2011) ATP-dependent roles of the DEAD-box protein Mss116p in group II intron splicing in vitro and in vivo. *J Mol Biol* 411: 661–679

Poulter NS, Pitkeathly WTE, Smith PJ & Rappoport JZ (2014) The physical basis of total internal reflection fluorescence (Tirf) microscopy and its cellular applications. *Methods Mol Biol* 1251

Prabu JR, Müller M, Thomae AW, Schüssler S, Bonneau F, Becker PB & Conti E (2015) Structure of the RNA Helicase MLE Reveals the Molecular Mechanisms for Uridine Specificity and RNA-ATP Coupling. *Mol Cell* 60: 487–499

Pyle AM (2008) Translocation and unwinding mechanisms of RNA and DNA helicases. *Annu Rev Biophys* 37: 317–336

Pyle AM (2011) RNA helicases and remodeling proteins. *Curr Opin Chem Biol* 15: 636–642

Raghunathan PL & Guthrie C (1998) A Spliceosomal Recycling Factor That Reanneals U4 and U6 Small Nuclear Ribonucleoprotein Particles. *Science (80-)* 279: 857–860

Robert-Paganin J, Halladjian M, Blaud M, Lebaron S, Delbos L, Chardon F, Capeyrou R, Humbert O, Henry Y, Henras AK, *et al* (2017) Functional link between DEAH/RHA helicase Prp43 activation and ATP base binding. *Nucleic Acids Res* 45: 1539–1552

Rodgers ML, Tretbar US, Dehaven A, Alwan AA, Luo G, Mast HM & Hoskins AA (2016) Conformational dynamics of stem II of the U2 snRNA. *RNA* 22: 225–236

Rogers GW, Richter NJ, Lima WF & Merrick WC (2001) Modulation of the Helicase Activity of eIF4A by eIF4B, eIF4H, and eIF4F. *J Biol Chem* 276: 30914–30922

Roman LJ, Eggleston AK & Kowalczykowski SC (1992) Processivity of the DNA Helicase Activity of Escherichia coli recBCD Enzyme*. *J Biol Chem* 267: 4207–4214

Roy J, Kim K, Maddock JR, Anthony JG & Woolford JL (1995) The final stages of spliceosome maturation require Spp2p that can interact with the DEAH box protein Prp2p and promote step 1 of splicing. *RNA* 1: 375–390

Roy R, Hohng S & Ha T (2008) A Practical Guide to Single Molecule FRET. *Nat Methods* 5: 507–516

Roychowdhury A, Joret C, Bourgeois G, Heurgué-Hamard V, Lafontaine DLJ & Graille M (2019) The DEAH-box RNA helicase Dhr1 contains a remarkable carboxyl terminal domain essential for small ribosomal subunit biogenesis. *Nucleic Acids Res* 47: 7548–7563

Rudolph MG & Klostermeier D (2015) When core competence is not enough: Functional interplay of the DEAD-box helicase core with ancillary domains and auxiliary factors in RNA binding and unwinding. *Biol Chem* 396: 849–865

Samatanga B & Klostermeier D (2014) DEAD-box RNA helicase domains exhibit a continuum between complete functional independence and high thermodynamic coupling in nucleotide and RNA duplex recognition. *Nucleic Acids Res* 42: 10644–10654

Sardana R, Liu X, Granneman S, Zhu J, Gill M, Papoulas O, Marcotte EM, Tollervey D, Correll CC & Johnson AW (2015) The DEAH-box Helicase Dhr1 Dissociates U3 from the Pre-rRNA to Promote Formation of the Central Pseudoknot. *PLoS Biol* 13: e1002083

Schmidt C, Grønborg M, Deckert J, Bessonov S, Conrad T, Lührmann R & Urlaub H (2014) Mass spectrometry-based relative quantification of proteins in precatalytic and catalytically active spliceosomes by metabolic labeling (SILAC), chemical labeling (iTRAQ), and label-free spectral count. *RNA* 20: 406–420

Schneider S & Schwert B (2001) Functional Domains of the Yeast Splicing Factor Prp22p. *J Biol Chem* 276: 21184–21191

Schütz P, Bumann M, Oberholzer AE, Bieniossek C, Trachsel H, Altmann M & Baumann U (2008) Crystal structure of the yeast eIF4A-eIF4G complex: An RNA-helicase controlled by protein-protein interactions. *PNAS* 105: 9564–9569

Schwer B (2008) A conformational rearrangement in the spliceosome sets the stage for Prp22-dependent mRNA release. *Mol Cell* 30: 743–754

Schwer B & Meszaros T (2000) RNA helicase dynamics in pre-mRNA splicing. *EMBO J* 19: 6582–6591

Semlow DR, Blanco MR, Walter NG & Staley JP (2016) Spliceosomal DEAH-Box ATPases Remodel Pre-mRNA to Activate Alternative Splice Sites. *Cell* 164: 985–998

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

Seraphin B, Kretzner L & Rosbash M (1988) A Ul snRNA:pre-mRNA base pairing interaction is required early in yeast spliceosome assembly but does not uniquely define the 5' cleavage site. *EMBO* 7: 2533-2538

Sharif H, Ozgur S, Sharma K, Basquin C, Urlaub H & Conti E (2013) Structural analysis of the yeast Dhh1-Pat1 complex reveals how Dhh1 engages Pat1, Edc3 and RNA in mutually exclusive interactions. *Nucleic Acids Res* 41: 8377–8390

Sikora B, Eoff RL, Matson SW & Raney KD (2006) DNA unwinding by Escherichia coli DNA helicase I (TraI) provides evidence for a processive monomeric molecular motor. *J Biol Chem* 281: 36110–36116

Siliciano PG & Guthrie C (1988) 5' Splice site selection in yeast: genetic alterations in base-pairing with U1 reveal additional requirements. *Genes Dev* 2: 1258–1267

Singleton MR, Dillingham MS & Wigley DB (2007) Structure and mechanism of helicases and nucleic acid translocases. *Annu Rev Biochem* 76: 23–50

Sloan KE & Bohnsack MT (2018) Unravelling the Mechanisms of RNA Helicase Regulation. *Trends Biochem Sci* 43: 237–250

Sloan KE, Knox AA, Wells GR, Schneider C & Watkins NJ (2019) Interactions and activities of factors involved in the late stages of human 18S rRNA maturation. *RNA Biol* 16: 196–210

Sloan KE, Warda AS, Sharma S, Entian KD, Lafontaine DLJ & Bohnsack MT (2017) Tuning the ribosome: The influence of rRNA modification on eukaryotic ribosome biogenesis and function. *RNA Biol* 14: 1138–1152

Smaldino PJ, Routh ED, Kim JH, Giri B, Creacy SD, Hantgan RR, Akman SA & Vaughn JP (2015) Mutational dissection of telomeric DNA binding requirements of G4 Resolvase 1 shows that G4-structure and certain 3'-tail sequences are sufficient for tight and complete binding. *PLoS One* 10

Smith AM, Fuchs RT, Grundy FJ & Henkin TM (2010) Riboswitch RNAs: Regulation of gene expression by direct monitoring of a physiological signal. *RNA Biol* 7: 104-110

Son H, Mo W, Park J, Lee JW & Lee S (2020) Single-Molecule FRET Detection of Sub-Nanometer Distance Changes in the Range below a 3-Nanometer Scale. *Biosensors* 10: bios10110168

Song C, Hotz-Wagenblatt A, Voit R & Grummt I (2017) SIRT7 and the DEAD-box helicase DDX21 cooperate to resolve genomic R loops and safeguard genome stability. *Genes Dev* 31: 1370–1381

Soultanas P & Wigley DB (2001) Unwinding the 'Gordian knot' of helicase action. *TRENDS Biochem Sci* 26: 47–54

Staley JP & Guthrie C (1999) An RNA Switch at the 5' Splice Site Requires ATP and the DEAD Box Protein Prp28p. *Mol Cell* 3: 55–64

Steimer L & Klostermeier D (2012) RNA helicases in infection and disease. *RNA Biol* 9: 751–771

Studer MK, Ivanovic L, Weber ME, Marti S & Jonas S (2020) Structural basis for DEAH-helicase activation by G-patch proteins. *Proc Natl Acad Sci U S A* 117: 7159–7170

Studier FW (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* 41: 207–234

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 (2005) Characterization of the NTPase, RNA-binding, and RNA helicase activities of the DEAH-box splicing factor Prp22. *Biochemistry* 44: 9795–9803

Tanner NK (2003) The newly identified Q motif of DEAD box helicases is involved in adenine recognition. *Cell cycle* 2: 18–19

Tauchert MJ, Fourmann JB, Christian H, Lührmann R & Ficner R (2016) Structural and functional analysis of the RNA helicase Prp43 from the thermophilic eukaryote Chaetomium thermophilum. *Acta Crystallogr Sect Struct Biol Commun* 72: 112–120

Tauchert MJ, Fourmann JB, Lührmann R & Ficner R (2017) Structural insights into the mechanism of the DEAH-box RNA helicase Prp43. *Elife* 6: 1–25

Theissen B, Karow AR, Köhler J, Gubaev A & Klostermeier D (2008) Cooperative binding of ATP and RNA induces a closed conformation in a DEAD box RNA helicase. *PNAS* 105: 548–553

Toczydlowska-Socha D, Zielinska MM, Kurkowska M, Astha, Almeida CF, Stefaniak F, Purta E & Bujnicki JM (2018) Human RNA cap1 methyltransferase CMTr1 cooperates with RNA helicase DHX15 to modify RNAs with highly structured 50 termini. *Philos Trans R Soc B Biol Sci* 373: 1–10

Toroney R, Nielsen KH & Staley JP (2019) Termination of pre-mRNA splicing requires that the ATPase and RNA unwindase Prp43p acts on the catalytic snRNA U6. *Genes Dev* 33: 1–20

Trapman J, Retel J & Planta RJ (1975) Ribosomal Precursor Particles from Yeast. *Exp Cell Res* 90: 95–104

Tsai R-T, Tseng C-K, Lee P-J, Chen H-C, Fu R-H, Chang K, Yeh F-L & Cheng S-C (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

Tseng CK, Liu HL & Cheng SC (2011) DEAH-box ATPase Prp16 has dual roles in remodeling of the spliceosome in catalytic steps. *RNA* 17: 145–154

Turowski TW & Tollervey D (2015) Cotranscriptional events in eukaryotic ribosome synthesis. *Wiley Interdiscip Rev RNA* 6: 129–139 doi:10.1002/wrna.1263

Vaughn JP, Creacy SD, Routh ED, Joyner-Butt C, Jenkins GS, Pauli S, Nagamine Y & Akman SA (2005) The DEXH protein product of the DHX36 gene is the major source of tetramolecular quadruplex G4-DNA resolving activity in HeLa cell lysates. *J Biol Chem* 280: 38117–38120

Walbott H, Mouffok S, Capeyrou R, Lebaron S, Humbert O, Van Tilbeurgh H, Henry Y & Leulliot N (2010) Prp43p contains a processive helicase structural architecture with a specific regulatory domain. *EMBO J* 29: 2194–2204

Walker JE, Saraste M, Runswick MJ & Gay NJ (1982) Distantly related sequences in the a-and f-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J* 8: 945-951

Wan R, Yan C, Bai R, Lei J & Shi Y (2017) Structure of an Intron Lariat Spliceosome from Saccharomyces cerevisiae. *Cell* 171: 120-132.e12

Wang Y & Guthrie C (1998) PRP16, a DEAH-box RNA helicase, is recruited to the spliceosome primarily via its nonconserved N-terminal domain. *RNA* 4: 1216–1229

Warkocki Z, Odenwälder P, Schmitzová J, Platzmann F, Stark H, Urlaub H, Ficner R, Fabrizio P & Lührmann R (2009) Reconstitution of both steps of Saccharomyces cerevisiae splicing with purified spliceosomal components. *Nat Struct Mol Biol* 16: 1237–1243

Warkocki Z, Schneider C, Mozaffari-Jovin S, Schmitzová J, Höbartner C, Fabrizio P & Lührmann R (2015) The G-patch protein Spp2 couples the spliceosome-stimulated ATpase activity of the deah-box protein Prp2 to catalytic activation of the spliceosome. *Genes Dev* 29: 94–107

Watkins NJ & Bohnsack MT (2012) The box C/D and H/ACA snoRNPs: Key players in the modification, processing and the dynamic folding of ribosomal RNA. *Wiley Interdiscip Rev RNA* 3: 397–414

Weaver PL, Sun C & Chang T-H (1997) Dbp3p, a Putative RNA Helicase in Saccharomyces cerevisiae, Is Required for Efficient Pre-rRNA Processing Predominantly at Site A 3. *Mol Cell Biol* 17: 1354–1365

Weir JR, Bonneau F, Hentschel J, Conti E & Feigon J (2010) Structural analysis reveals the characteristic features of Mtr4, a DExH helicase involved in nuclear RNA processing and surveillance. *PNAS* 107: 12139–12144

Wilkinson ME, Fica SM, Galej WP, Norman CM, Newman AJ & Nagai K (2017) Postcatalytic spliceosome structure reveals mechanism of 3'–splice site selection. *Science (80-)* 358: 1283–1288

Will CL & Lührmann R (2011) Spliceosome structure and function. *Cold Spring Harb Perspect Biol* 3: 1–2

Wojtas MN, Pandey RR, Mendel M, Homolka D, Sachidanandam R & Pillai RS (2017) Regulation of m6A Transcripts by the 3′→5′ RNA Helicase YTHDC2 Is Essential for a Successful Meiotic Program in the Mammalian Germline. *Mol Cell* 68: 374-387.e12

Wu Y (2012) Unwinding and rewinding: Double faces of helicase? *J Nucleic Acids* 2012: 1-14

Xu Y-Z & Query CC (2007) Competition between the ATPase Prp5 and branch region-U2 snRNA pairing modulates the fidelity of spliceosome assembly. *Mol Cell* 28: 838–849

Yan C, Wan R, Bai R, Huang G & Shi Y (2016) Structure of a yeast activated spliceosome at 3.4 Å resolution. *Science (80-)* 353: 904–911

Yang Q, Del Campo M, Lambowitz AM & Jankowsky E (2007) DEAD-Box Proteins Unwind Duplexes by Local Strand Separation. *Mol Cell* 28: 253–263

Yang Q & Jankowsky E (2006) The DEAD-box protein Ded1 unwinds RNA duplexes by a mode distinct from translocating helicases. *Nat Struct Mol Biol* 13: 981–986

Yoshimoto R, Kataoka N, Okawa K & Ohno M (2009) Isolation and characterization of post-splicing lariat-intron complexes. *Nucleic Acids Res* 37: 891–902

Young CL, Khoshnevis S & Karbstein K (2013) Cofactor-dependent specificity of a DEAD-box protein. *Proc Natl Acad Sci U S A* 110

Zhang ZM, Yang F, Zhang J, Tang Q, Li J, Gu J, Zhou J & Xu YZ (2013) Crystal Structure Of Prp5P Reveals Interdomain Interactions That Impact Spliceosome Assembly. *Cell Rep* 5: 1269–1278

Zhou T, Ren X, Adams RL & Pyle AM (2018) NS3 from Hepatitis C Virus Strain JFH-1 Is an Unusually Robust Helicase That Is Primed To Bind and Unwind Viral RNA. *J Virol* 92

Zhu J, Liu X, Anjos M, Correll CC & Johnson AW (2016) Utp14 Recruits and Activates the RNA Helicase Dhr1 To Undock U3 snoRNA from the Preribosome. *Mol Cell Biol* 36: 965–978

Abbreviations

°C	Degree Celsius
2' OH	2’ hydroxyl
3' SS	3’ splice site
5' SS	5’ splice site
A	Adenine
ADP	Adenosine diphosphate
ADP	Adenosine diphosphate
AMPPNP	Adenosine-5'-[(β,γ)-imido]triphosphate
ATP	Adenosine triphosphate
BS	Branch site
Cryo-EM	Cryo-electron microscopy
ct	*Chaetomium thermophilum*
C-terminal	Carboxy-terminal
Cys	Cysteine
DEAD	Aspartate-Glutamate-Alanine-Aspartate
DEAH	Aspartate-Glutamate-Alanine-Histidine
DFG	Deutsche Forschungsgemeinschaft
DHX	DEAH-box helicase
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
ds	Double-stranded
DTT	Dithiothreitol
E. coli	*Escherichia coli*
E_{FRET}	FRET efficiency
FI	Fluorescence intensity
FRET	Förster resonance energy transfer
G	Guanosine

gp	G-patch, Glycine-rich patch
GTP	Guanosine triphosphate
H. sapiens	*Homo sapiens*
HB	Helix-bundle
His	Histidine
ILS	Intron-lariat spliceosome
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ITC	Isothermal titration calorimetry
K_{cat}	Turnover numbered
K_D	Dissociation constant
K_M	Michaelis constant
k_{off}	Dissociation rate constant
M	Molar
min	Minute
mRNA	Messenger RNA
NADH	Nicotinamide adenine dinucleotide
Ni	Nickel
nt	Nucleotide
N-terminal	Amino-terminal
NTP	Nucleoside triphosphate
NTR	Nineteen-related
OB	Oligonucleotide-binding
OD	Optical density
PDB	Protein Data Bank
PEG	Polyethylene glycol
P_i	Inorganic phosphate
Prp	Pre-mRNA-processing factor
RNA	Ribonucleic acid
RNPs	Ribonucleoprotein particles

rpm	Revolutions per minute
rRNA	Ribosomal RNA
S	Svedberg unit
S. cerevisiae	*Saccharomyces cerevisiae*
sc	*Saccharomyces cerevisiae*
sd	Standard deviation
SF	Superfamily
SFB	Sonderforschungsbereich
smFRET	Single-molecule fluorescence resonance energy transfer
snoRNA	Small nucleolar RNA
snoRNP	Small nucleolar RNP
snRNP	Small nuclear ribonucleic particle
ss	Single-stranded
TIRF	Total internal reflection fluorescence
Tris	2-Amino-2-(hydroxymethyl)propane-1,3-diol
tRNA	Transfer RNA
Trolox	6-hydroxy-2,5,7,8-tetramethylchromane-2-carboxylic acid
U	Uracil
U-snRNP	Uridine-rich small nuclear ribonucleic particle
V_{max}	Maximum reaction velocity
WH	Winged-helix

Acknowledgements

Foremost, I want to express my gratitude to Prof. Dr. Ralf Ficner for giving me the opportunity to work on this exciting project. I appreciate his competent and kind supervision, his continuous advice and suggestions concerning my project and the freedom he gave me to develop and execute own ideas.

I am deeply thankful to Dr. Sarah Adio for patiently guiding me through the sometimes confusing and frustrating world of smFRET. Her constant support and advice during our many discussions greatly helped me to develop my project, to gain new interesting perspectives and to improve my scientific writing skills.

I would also like to thank Prof. Dr. Reinhard Lührmann for his participation in my Thesis Advisory Committee and providing useful input during the thesis committee meetings.

Furthermore, I am grateful to Prof. Dr. Markus Bohnsack, Prof. Dr. Kai Tittmann and Prof. Dr. Henning Urlaub for accepting to be part of my extended examination board.

Also, I thank Prof. Dr. Marina Rodnina for granting me access to the TIRF microscopy facility at her department.

Special thanks go to all present and former members of the Department of Molecular Structural Biology for creating a great working atmosphere. In particular, I want to thank Dr. Achim Dickmanns for his continuous support, already starting with the supervision of my bachelor thesis, and his helpful feedback on anything from scientific presentations and manuscripts to the design of experiments. I am thankful to Dr. Piotr Neumann for patiently helping with all crystallographic (and Linux related) concerns and his comments on my manuscripts. Also, I am grateful to Marita Kalck and Susanne van Beckum for kindly navigating me through all bureaucratic pitfalls. A huge thank you goes to my fellow colleagues: Tim, Katharina and Patrick. You make our lab a great place to work at! I also would like to thank Natalie and Panos for their practical advice during microscopy sessions and the whole Deparment of Physical Biochemistry at the MPI for the welcoming atmosphere during my time there.

Finally, I want to thank my family and friends for always being supportive of me and my work. This goes especially for Christoph who put up with many science related mood swings and weird working hours, listened to the countless rehearsals of my talks and gave his best to keep me grounded. Thank you for always being there for me!