



# Crystal structure of Prp16 in complex with ADP

Tim Benedict Garbers, Marieke Enders, Piotr Neumann and Ralf Ficner\*

Department of Molecular Structural Biology, Institute of Microbiology and Genetics, GZMB, Georg-August-University Göttingen, Göttingen, Germany. \*Correspondence e-mail: rficner@uni-goettingen.de

Received 24 March 2023

Accepted 29 June 2023

Edited by K. K. Kim, Sungkyunkwan University School of Medicine, Republic of Korea

**Keywords:** DEAH-box helicases; Prp16; *Chaetomium thermophilum*; spliceosome.

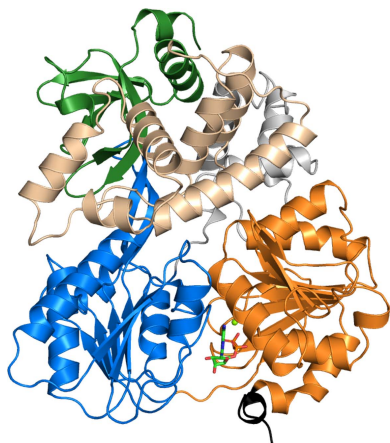
**PDB reference:** Prp16, complex with ADP, 8cnt

**Supporting information:** this article has supporting information at journals.iucr.org/f

DEAH-box helicases play a crucial role in pre-mRNA splicing as they are responsible for major rearrangements of the spliceosome and are involved in various quality-ensuring steps. Prp16 is the driving force during spliceosomal catalysis, remodeling the C state into the C\* state. Here, the first crystal structure of Prp16 from *Chaetomium thermophilum* in complex with ADP is reported at 1.9 Å resolution. Comparison with the other spliceosomal DEAH-box helicases Prp2, Prp22 and Prp43 reveals an overall identical domain architecture. The  $\beta$ -hairpin, which is a structural element of the RecA2 domain, exhibits a unique position, punctuating its flexibility. Analysis of cryo-EM models of spliceosomal complexes containing Prp16 reveals that these models show Prp16 in its nucleotide-free state, rendering the model presented here the first structure of Prp16 in complex with a nucleotide.

## 1. Introduction

The spliceosome is large molecular machinery that is responsible for the removal of noncoding introns from eukaryotic precursor messenger RNAs (pre-mRNAs). For each intron removal, the spliceosome is assembled *de novo* and performs two subsequent transesterification reactions. In the course of each cycle, the components assemble in a step-wise manner and a number of conformational changes take place (Will & Lührmann, 2011). The spliceosome consists of five small nuclear ribonucleoprotein complexes (snRNPs) and several other non-snRNP proteins (Wahl *et al.*, 2009). One cycle can be divided into assembly, activation, catalysis and disassembly stages. The first event in each cycle is the binding of the U1 snRNP to the 5'-splice site of the pre-mRNA, forming the E complex. Binding of the U2 snRNP to the branch-point sequence and recruitment of the pre-assembled tri-snRNP U4/U6/U5 conclude the assembly step, forming the pre-B complex. For the activation of the spliceosome, the U1 and the U4 snRNPs are displaced while the NineTeen complex (NTC) and the NTC-related complex (NTR) bind, forming the catalytically competent B\* spliceosome. Following the first transesterification reaction, the resulting C complex is further remodeled into the C\* complex, which can facilitate the second transesterification reaction. Finally, the spliceosome is disassembled, releasing the remaining snRNPs, NTC and NTR, the severed intron lariat and the spliced mRNA (Ficner *et al.*, 2017). The structural rearrangements of the spliceosome include RNA–RNA, protein–RNA and protein–protein remodeling, which need to be tightly orchestrated and to undergo rigorous quality-control steps, as misspliced mRNA can be linked to diseases such as Duchenne muscular dystrophy (Takeshima *et al.*, 2010) and spinal muscular atrophy (Lorson *et al.*, 1999). All of these remodeling steps are driven and controlled by the action of RNA helicases



OPEN ACCESS

Published under a CC BY 4.0 licence

belonging to helicase superfamily 2 (SF2; Cordin & Beggs, 2013). The SF2 helicases involved in splicing can be further divided into three subfamilies: DEAD-box, DEAH-box and Ski2-like helicases (Jankowsky & Fairman, 2007).

The protein studied in this work is pre-mRNA-splicing factor 16 (Prp16) and is required for the transition of the spliceosome from the C to the C\* state (Wilkinson *et al.*, 2021). Therefore, Prp16 binds the 3'-end sequence of the intron RNA and translocates towards the spliceosome, leading to dissociation of the branching factors Yju2 and Cwc25 and thereby enabling exon ligation (Tseng *et al.*, 2011). The translocation also plays a role in splicing-fidelity control in a kind of kinetic proofreading mechanism where Prp16 antagonizes suboptimal substrates and promotes optimal substrates for 5'-splice site cleavage (Burgess & Guthrie, 1993; Koodathingal *et al.*, 2010; Koodathingal & Staley, 2013; Semlow *et al.*, 2016). Prp16 belongs to the DEAH-box proteins, which share a highly conserved helicase core formed by two RecA-like domains. These two domains provide the necessary architecture that enables DEAH-box helicases to be functional NTPases (Schwer & Guthrie, 1992). The nucleotide binds at the domain interface, where the eight conserved sequence motifs I, Ia, Ib, II, III, IV, V and VI are located. Motifs I, II and VI are involved in nucleic acid binding of the helicase and motif II contains the eponymous sequence DEAH. While motifs Ia, Ib and IV are necessary to bind and hydrolyze the nucleoside triphosphate, motif III is required to couple NTP hydrolysis to the process of unwinding (Campodonico & Schwer, 2002; Schneider *et al.*, 2004). All DEAH-box helicases have a common C-terminal domain architecture comprising winged-helix (WH), helix-bundle (HB) and oligosaccharide-binding fold (OB-fold) domains. These domains have a regulatory effect on the hydrolyzation rate of the helicase (Kudlinzki *et al.*, 2012). It has also been shown that this part can serve as a platform for the binding of interaction partners (Cordin & Beggs, 2013). The N-terminal regions of the spliceosomal DEAH-box helicases greatly vary in their length and are the least conserved region (Cordin & Beggs, 2013).

Recent cryo-EM structures (three-dimensional reconstructions) of spliceosomal complexes in the C or C\* state from human or yeast contain map areas which can be related to parts of or even complete Prp16 molecules (Wilkinson *et al.*, 2021; Galej *et al.*, 2016; Yan *et al.*, 2017; Bertram *et al.*, 2020; Zhan *et al.*, 2018). All atomic models of the Prp16 structures were modeled based on the structurally related yeast Prp43 in complex with ADP (PDB entries 3kx2 and 2xau; He *et al.*, 2010; Walbott *et al.*, 2010). As the local resolution of the cryo-EM maps at the position of Prp16 ranges between 8 and 15 Å, the derived atomic models of Prp16 provide only limited information about the structure of this helicase.

Here, we report the first crystal structure of Prp16 from the thermophilic eukaryotic ascomycete *Cheatomium thermophilum* in its ADP-bound state at 1.9 Å resolution. Analysis of the structure shows the same domain architecture as observed for other spliceosomal DEAH-box helicases. Moreover, the interaction pattern of Prp16 with ADP also seems to be conserved within this helicase family. In addition, a new

**Table 1**  
Macromolecule-production information.

Source organism	<i>C. thermophilum</i> var. <i>thermophilum</i> DSM 1495
DNA source	Codon-optimized synthetic DNA
Forward primer	CGGGATCCCTGAAAGAACAGCGTGAATTC TGCC
Reverse primer	GGAATTCCTATACGGCGGAGGGTCCAGTT CTGCCAGCCAATGCG
Cloning vector	pETM-13
Expression vector	pGEX-6P-3
Expression host	<i>Escherichia coli</i> Rosetta 2 (DE3)
Complete amino-acid sequence of the construct produced	<b>GPLGSPNTLKEQREFLPAFAVRELLLRVIR</b> DNQVVIVIGETGSGKTTQLTQFLYEDGY GKTGMI GCTQPRRVAAMSVAKRVAEEME VKLGLTVGYAIRFEDCTSKETVIKYMTD GVLLRESLNEPDLDRYSCIIMDEAHERA LNTDVLMLGFKKILQRRDLKLIITSAT MNSKRFSDFFGGAPEFTIPGRTFPVLL PHRSVVEDYVDQAVQVLAIHVSKPAGD ILVFM TGQEDIEVTCELIQRERLALNDP PKLSVLP IYSQMPADLQAKIFDRAPPGV RKCI VATNIAETSLTVDGIMYVVDGYS KLVVYNPRMGMDTLQITPISQANAAQRA GRAGRTGPGQAYRLYTEKQFRDEMYMQT IPEIQRTNLSNTVLLKSLGVKDLLDFD FMDPPQDTITTSLYDLWALGALDNLGE LTELGRKMNAFPMDPPLAKLLIMSEEYG CSEEMVTIVSMLSVPNVFYRPKERQES DAAREKFFVPESDHLTYLHVYQWKANG YSDAWCARHFLHSKSLRRAREVRDQLLD IMKMQHMRMVSCGTDWDIIRKICISGYI HQAARKVKGIGEYINLRSTVTVQLHPTSA LYGLGFLPDYVVYHELIIITSKEYMSTVT AVDPHWAEL <b>EPSAV</b>

position of the  $\beta$ -hairpin could be observed, punctuating its proposed flexibility.

## 2. Materials and methods

### 2.1. Macromolecule production

The Prp16-encoding gene of *C. thermophilum* (*ctPrp16*) was identified by the NCBI *BLAST* search tool (Sequence ID EGS23320.1). The codon-optimized sequence encoding *ctPrp16* residues 302–920 was cloned into the pGEX-6P-3 vector utilizing the BamHI and EcoRI restriction sites. The recombinant protein was expressed using *Escherichia coli* Rosetta2 (DE3) cells in 2×YT medium. The cells were induced with 0.5 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside at an optical density (OD<sub>600</sub>) of 0.6 and were incubated at 16°C for 20 h. The cells were harvested and washed with phosphate-buffered saline. Prior to cell disruption via a microfluidizer, the cells were mixed at 4 ml g<sup>-1</sup> with a buffer consisting of 50 mM Tris–HCl pH 7.5, 500 mM NaCl, 10 mM EDTA. Subsequently, the lysate was centrifuged at 20 000g and 4°C for 30 min. The supernatant was applied onto a Glutathione Sepharose 4B column (GE Healthcare). A washing step with additional 2 M LiCl was performed to remove bound nucleotides and nucleic acids. Finally, the protein was eluted using 30 mM reduced glutathione. Cleavage of the GST tag was performed by incubation with PreScission protease [100:1(w:w)]. Further impurities and the GST tag were removed by applying the protein onto a Superdex 200

**Table 2**  
Crystallization.

Method	Vapor diffusion
Plate type	Sitting drop
Temperature (K)	293
Protein concentration	4
Buffer composition of protein solution	10 mM Tris–HCl pH 7.5, 200 mM NaCl, 2 mM MgCl <sub>2</sub>
Composition of reservoir solution	20% (w/v) PEG 4000, 100 mM MES pH 6.5, 5 mM MgCl <sub>2</sub>
Volume and ratio of drop	2 µl, 1:1 ratio
Volume of reservoir (µl)	500

gel-filtration column (GE Healthcare) coupled to a Glutathione Sepharose 4B column (GE Healthcare) in 10 mM Tris–HCl pH 7.5, 200 mM NaCl, 2 mM MgCl<sub>2</sub>. A final check for remaining nucleotides was performed by measuring the  $A_{260}/A_{280}$  absorbance. The protein was concentrated to 20 mg ml<sup>-1</sup> using an Amicon Ultra centrifugal concentrator (Merck). A total of 8 mg *ctPrp16* (302–920) could be obtained from 2 l medium. The protein solution was aliquoted into PCR tubes in 105 µl samples, flash-frozen in liquid nitrogen and stored at –80°C for further usage. Macromolecule-production information is summarized in Table 1.

## 2.2. Crystallization

Crystallization was performed by the sitting-drop vapor-diffusion method at 293 K. *ctPrp16* (302–920) was diluted to 4 mg ml<sup>-1</sup> (55.98 µM) with gel-filtration buffer. To obtain *ctPtp16*–ADP complex crystals, the protein was mixed with a tenfold molar excess of ADP (559.8 µM) and reservoir solution consisting of 20% (v/v) PEG 4000, 100 mM MES pH 6.5, 5 mM MgCl<sub>2</sub>. The total volume of the drop was 2 µl with a 1:1 ratio of protein and reservoir solution. Crystals were observed after one week. Crystallization information is summarized in Table 2.

## 2.3. Data collection and processing

For cryoprotection, the *ctPrp16*–ADP complex crystals were transferred to mother liquor containing an additional 10% (w/v) PEG 4000 and 550.8 µM ADP before being plunged into liquid nitrogen. Oscillation diffraction images were collected on beamline P13 at DESY, Hamburg, Germany. Data were processed using *XDS* (Kabsch, 2010). Data-collection and processing statistics are summarized in Table 3.

## 2.4. Structure solution and refinement

The structure of *ctPrp16* (302–920) was solved by molecular replacement using *Phaser* (McCoy *et al.*, 2007). The structure of *ctPrp2* in complex with ADP (PDB entry 6fac; Schmitt *et al.*, 2018) was used as the search model. Initial rounds of refinement were performed using *Phenix* (Liebschner *et al.*, 2019). Manual model building was performed using *Coot* (Emsley *et al.*, 2010) and the model was subsequently refined using *REFMAC* (Murshudov *et al.*, 2011). No  $I/\sigma(I)$  cutoff was applied during refinement. Final validation of the model was conducted via *MolProbity* (Chen *et al.*, 2010). All figures were

**Table 3**  
Data collection and processing.

Diffraction source	P13, DESY
Wavelength (Å)	0.97625
Temperature (K)	100
Detector	EIGER 16M
Crystal-to-detector distance (mm)	217.62
Rotation range per image (°)	0.15
Total rotation range (°)	360
Exposure time per image (s)	0.04
Space group	$P2_12_12_1$
<i>a</i> , <i>b</i> , <i>c</i> (Å)	55.13, 102.14, 106.61
$\alpha$ , $\beta$ , $\gamma$ (°)	90, 90, 90
Mosaicity (°)	0.301
Resolution range (Å)	48.97–1.90 (2.00–1.90)
Total No. of reflections	647348 (91387)
No. of unique reflections	91538 (13012)
Completeness (%)	99.9 (99.8)
Multiplicity	7.1 (7.0)
$\langle I/\sigma(I) \rangle^\dagger$	11.8 (1.34)†
$R_{\text{meas}}$	0.115 (1.686)
$CC_{1/2}$	0.99 (0.50)
Overall <i>B</i> factor from Wilson plot (Å <sup>2</sup> )	31.3

† The  $\langle I/\sigma(I) \rangle$  value in the outer shell is 1.34, but the other values are still satisfactory as  $CC_{1/2} = 50\%$ .

**Table 4**  
Structure refinement.

Resolution range (Å)	48.97–1.90 (1.95–1.90)
Completeness (%)	99.9
No. of reflections, working set	45747 (3280)
No. of reflections, test set	2409 (173)
Final $R_{\text{cryst}}$	0.193 (0.321)
Final $R_{\text{free}}$	0.238 (0.385)
Cruickshank DPI	0.167
No. of non-H atoms	
Protein	4926
Ion	1
ADP	27
PEG	19
Water	432
R.m.s. deviations	
Bond lengths (Å)	0.008
Angles (°)	1.486
Average <i>B</i> factors (Å <sup>2</sup> )	
Protein	36
Ion	31
ADP	41
PEG	65
Ramachandran plot	
Most favored (%)	96.93
Allowed (%)	3.07

prepared using *PyMOL* (version 1.8; Schrödinger). Refinement statistics are summarized in Table 4.

## 3. Results and discussion

The spliceosomal DEAH-box helicase Prp16 from the thermophilic ascomycete *C. thermophilum* (*ctPrp16*) was crystallized in complex with ADP to mimic a post-catalytic state. The ortholog from *C. thermophilum* was chosen as proteins from this organism exhibit high thermostability and therefore tend to crystallize better (Amlacher *et al.*, 2011). Several studies of the closely related DEAH-box helicases Prp43, Prp2 and Prp22 have shown that orthologs from *C. thermophilum* are highly suitable for structural investiga-

tions of spliceosomal helicases (Tauchert *et al.*, 2016, 2017; Schmitt *et al.*, 2018; Hamann *et al.*, 2019; Absmeier *et al.*, 2020). An N- and C-terminally truncated construct *ctPrp16* (302–920) was used for crystallization, as the removed parts are expected to be mainly disordered according to the *PredictProtein* server (Bernhofer *et al.*, 2021). The crystal belonged to the orthorhombic space group  $P2_12_12_1$ , with unit-cell parameters  $a = 55.13$ ,  $b = 102.14$ ,  $c = 106.61$  Å,  $\alpha = \beta = \gamma = 90^\circ$  (Table 2). The phase problem was solved via molecular replacement using the structure of ADP-bound *ctPrp2* as the search model (PDB entry 6fac). The asymmetric unit contains one *ctPrp16* molecule. The polder omit map clearly reveals the presence of ADP and a magnesium ion in the active center (Supplementary Figs. S2 and S3). The reported structure of *ctPrp16* in complex with ADP was determined at a resolution of 1.9 Å and was refined to an  $R_{\text{work}}$  of 19.3% and an  $R_{\text{free}}$  of 23.8%. According to the Ramachandran plot, 96.93% of all residues are in the most favored region, 3.07% are in the allowed region and 0% are outliers (Table 4). The atomic model comprises *ctPrp16* (302–920), one ADP molecule, one magnesium ion, one polyethylene glycol (PEG) molecule and 423 water molecules. The overall assembly of *ctPrp16* can be divided into six different entities: the N-terminal extension (302–311), two RecA-like domains [referred to as RecA1 (312–468) and RecA2 (467–661)], the WH domain (662–728), the HB domain (729–842) and the OB-fold domain (843–920) (Fig. 1). While the RecA domain architecture is shared by all members of the SF2 helicase family, the three C-terminal domains are characteristic of DEAH-box helicases. A prominent  $\beta$ -hairpin, which protrudes out of the RecA2

domain and comprises residues 601–620 in *ctPrp16*, is another typical feature of DEAH-box helicases.

### 3.1. ADP binding of *ctPrp16*

*ctPrp16*, which is a member of the SF2 helicase family, possesses a well conserved catalytic core comprising several highly similar structural motifs that are spread over both RecA-like domains. Motifs Ia (<sup>359</sup>TQPRRVAA<sup>366</sup>), Ib (<sup>405</sup>TDGVLLR<sup>411</sup>) and IV (<sup>520</sup>LVFMTG<sup>525</sup>) are reported to be necessary for interaction with the substrate nucleic acid, while motifs I (<sup>334</sup>GSGKT<sup>338</sup>), II (<sup>428</sup>DEAH<sup>431</sup>), V (<sup>581</sup>TNIAETSLT<sup>589</sup>) and VI (<sup>628</sup>QRAGRAGR<sup>635</sup>) are needed for binding and hydrolysis of the nucleoside triphosphate. Motif III (<sup>460</sup>SAT<sup>462</sup>) couples nucleotide hydrolysis to the unwinding activity. In the structure reported here, an ADP molecule as well as the catalytically required magnesium ion are present at the interface between the RecA1 and RecA2 domains, corresponding to *ctPrp16* in the post-catalytic state. The magnesium ion is hexavalently coordinated by four water molecules, Thr342 located within motif I and the  $\beta$ -phosphate of the ADP molecule. The  $\beta$ -phosphate is furthermore coordinated by interactions with the backbone amides of Gly334, Ser335, Gly336, Lys337 and Thr338 and with the side chains of Lys341 and Thr342. The former group of five amino acids together form motif I. Additionally, four water molecules could be identified in the structure which are involved in the interaction network of the  $\beta$ -phosphate, making a total of 13 interactions between this phosphate and *ctPrp16* (Fig. 2a). The O2' atom of the ribose interacts with the carbonyl backbone of

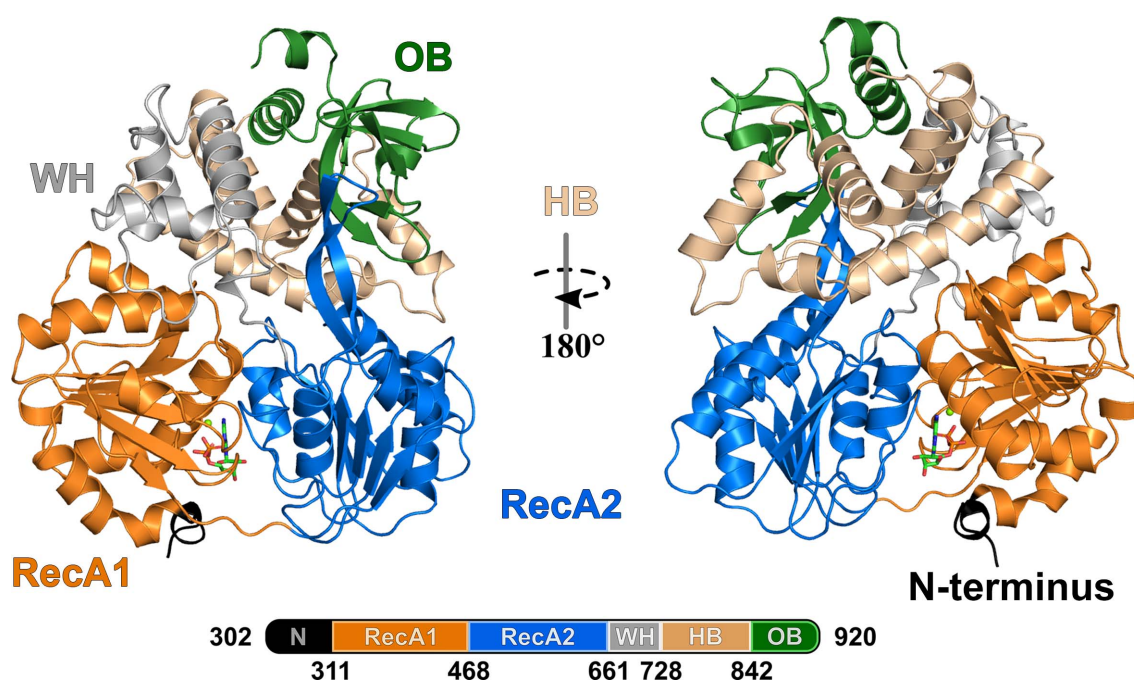


Figure 1

Crystal structure of *ctPrp16* (302–920) in complex with ADP at 1.9 Å resolution. The N-terminal extension (302–311) is depicted in black, the RecA1-like domain (312–468) in orange, the RecA2-like domain (469–661) in marine blue, the WH domain (662–728) in gray, the HB domain (729–842) in wheat and the OB-fold domain (843–920) in green. The structure is depicted in cartoon mode, while the bound ADP is represented as sticks and the magnesium ion is shown as a sphere.

Thr589, one water molecule and the side chain of Asp595. The O3' atom interacts with the side chains of Asp595 and Arg635. An additional interaction could be observed between one water molecule and the O4' atom of the ribose ring. While Arg635 (motif VI) belongs to this conserved region of DEAH-box helicases, Thr593 and Asp595 do not. The adenine moiety does not form any polar contacts with *ctPrp16* residues, but the

adenine interacts with Phe567 via  $\pi$ - $\pi$  stacking (Fig. 2*b*). This interaction seems to be highly conserved as it can be found in all available X-ray structures of spliceosomal DEAH-box helicases bound to ADP, except for one structure (PDB entry 6faa), in which the adenine base is flipped over in a *syn* conformation (Tauchert *et al.*, 2016; Schmitt *et al.*, 2018; Hamann *et al.*, 2020). The eponymous motif II of this protein

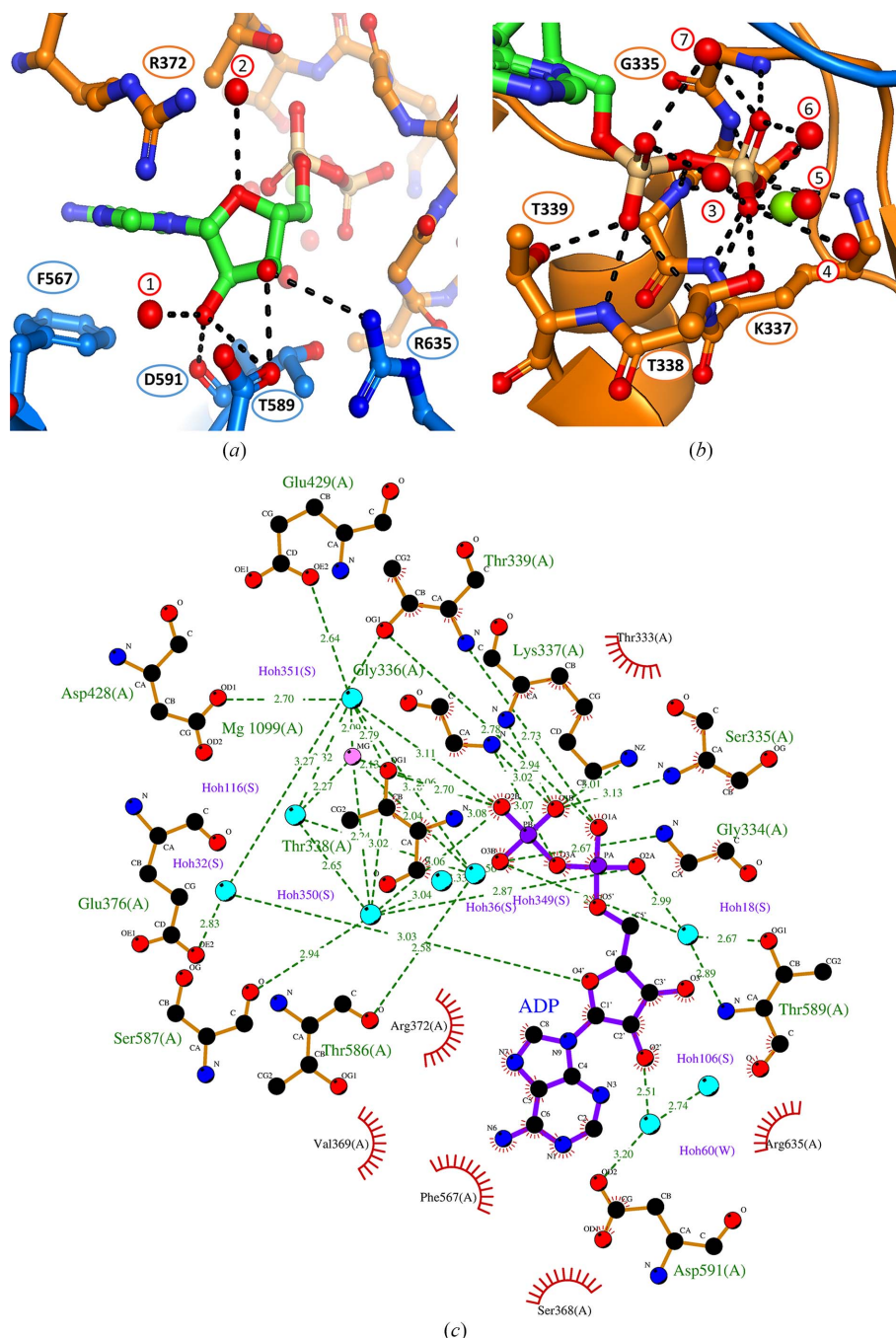


Figure 2

ADP-binding site of *ctPrp16*. The ADP molecule is sandwiched between the RecA1 (orange) and RecA2 (marine) domains (color scheme as in Fig. 1). C atoms are shown in orange/marine (protein) and green (ADP), N atoms are in blue, O atoms are in red, P atoms are in wheat, the  $Mg^{2+}$  ion is in light green and water molecules are shown as red spheres. Residues which are involved in ADP, water or  $Mg^{2+}$  binding are presented in ball-and-stick model and are labeled according to the *ctPrp16* sequence. Polar interactions are visualized as dashed black lines. (a) The adenine moiety is bound via  $\pi$ - $\pi$  stacking and the ribose by hydrogen bonding. (b) The  $\alpha$ - and  $\beta$ -phosphates participate intensively in hydrogen bonding. The central  $Mg^{2+}$  ion is coordinated by four water molecules, the Thr342 side chain and an O atom of the  $\beta$ -phosphate. (c) A two-dimensional ligand-protein interaction diagram (created with *LigPlot+*) depicting the network of interactions between the ADP molecule and the RecA1 and RecA2 domains of Prp16.

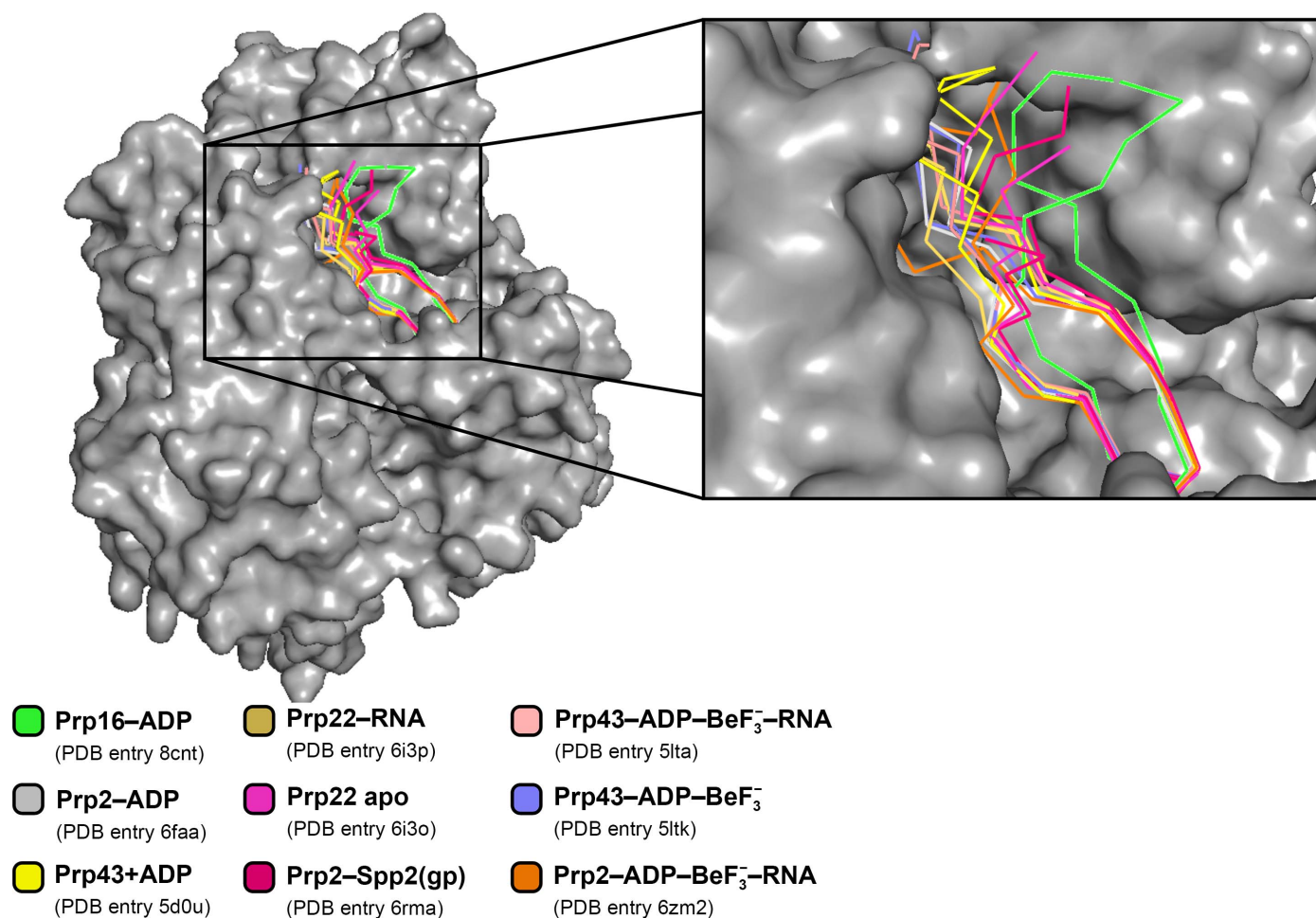
class, DEAH, is not directly involved in any interactions between the protein and the nucleotide. Instead, it is involved in the formation of the interaction network of the magnesium ion by interacting with two water molecules via the carbonyl groups of Asp428 and Glu429. Furthermore, Glu131 stabilizes the position of one water molecule near the magnesium ion. Overall, the interactions of the *C. thermophilum* spliceosomal DEAH-box helicases with ADP seem to be highly conserved as in other ADP-bound helicases: *ctPrp2* (PDB entry 6fac) and *ctPrp43* (PDB entry 5d0u) exhibit the same interaction pattern (Schmitt *et al.*, 2018; Tauchert *et al.*, 2016).

### 3.2. Position of the $\beta$ -hairpin

Based on structural comparison with the Ski2-like helicase Hel308, the RecA2  $\beta$ -hairpin has been suggested to be involved in the double-strand separation of DEAH-box helicases (Büttner *et al.*, 2007; He *et al.*, 2010; Walbott *et al.*, 2010). In contrast, 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  $\beta$ -hairpin, positioning a

potential double strand on the opposite side of this structural element. Comparison with other helicases suggests that the conformation of the  $\beta$ -hairpin observed in Prp16 would support its role as a physical barrier, separating the 3' portion of the bound RNA that exhibits a stacked conformation from the 5' portion that mainly interacts with the C-terminal domains. Thereby, it is thought to prevent backsliding of the RNA strand during translocation (Hamann *et al.*, 2019).

Superposition of all spliceosomal DEAH-box helicase structures from *C. thermophilum* in different conformational stages reveals the unique conformation of the  $\beta$ -hairpin in the *ctPrp16* structure (Fig. 3). While all other X-ray structures harbor the  $\beta$ -hairpin between the WH and OB-fold domains, independent of the nucleotide- or substrate-loading state, in *ctPrp16* the upper part of the  $\beta$ -hairpin with its loop region is pushed out of the cleft and interacts exclusively with the OB-fold domain. However, this conformation appears to be an artifact caused by interaction with a symmetry mate within the crystal lattice (Supplementary Fig. S1). The bulky amino acids Glu306 and Phe311 of the symmetry mate, located at the N-terminus of the truncated *ctPrp16*, interfere with the loop



**Figure 3**

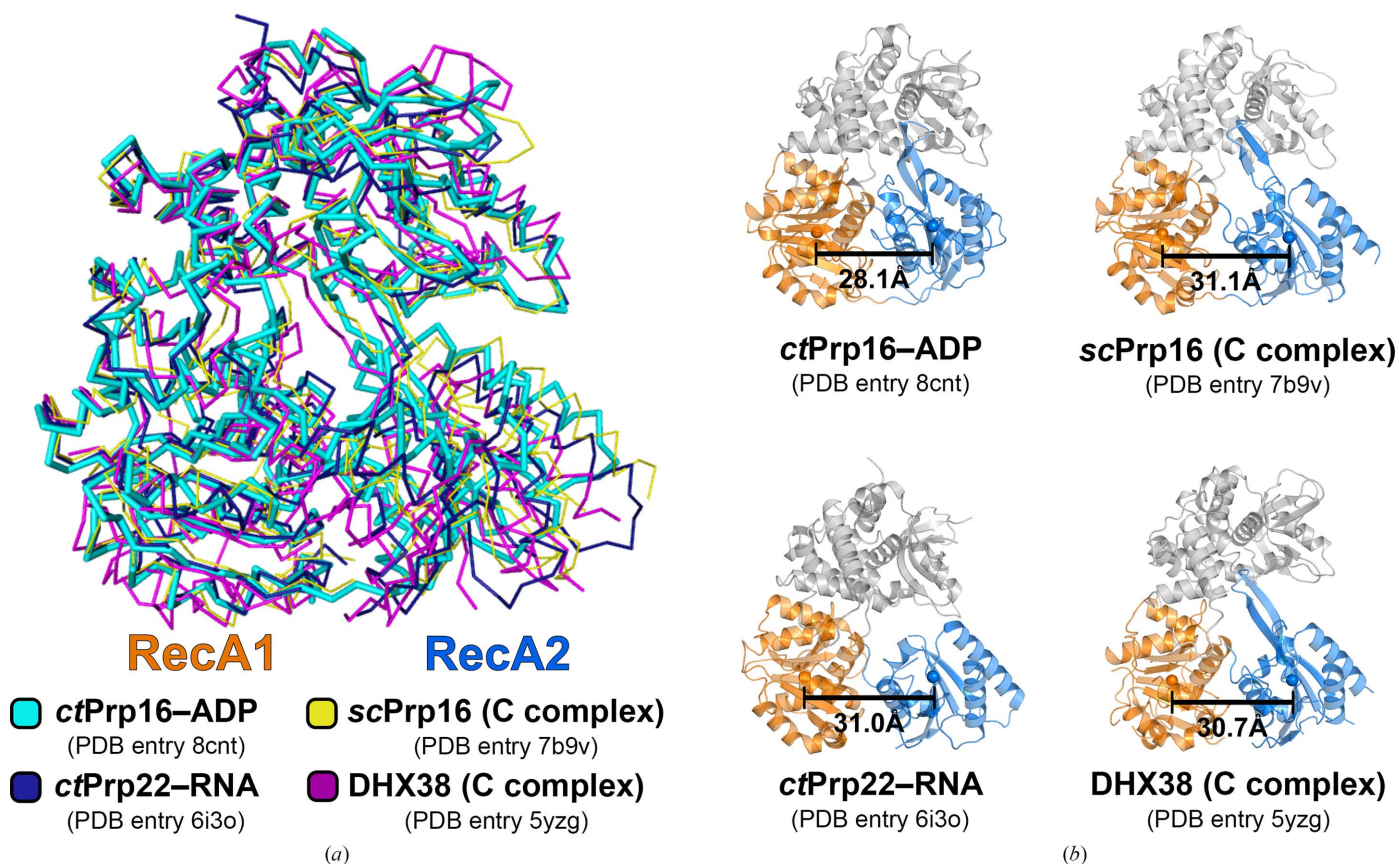
Comparison of the different  $\beta$ -hairpin conformations. The structure of *ctPrp16* was superposed with different spliceosomal DEAH-box helicases from *C. thermophilum* in different catalytic states via their RecA1 domains. The different  $\beta$ -hairpins are represented in cartoon mode, while the remaining part of the protein is shown as a gray surface (*ctPrp16*-ADP only). Parts of the OB-fold domain were omitted for clarity. The superposition reveals that the  $\beta$ -hairpin of *ctPrp16*-ADP adapts a more distant conformation compared with the other structures.

which connects the two  $\beta$ -strands forming the  $\beta$ -hairpin. Arg610 in particular, which is in the middle of the loop region, would clash with the previously mentioned N-terminal residues if the  $\beta$ -hairpin adopts a position between the WH and OB-fold domains. In the known complex structures of DEAH-box helicases with bound RNA, there is only one conserved Lys residue of the  $\beta$ -hairpin that interacts with the RNA, for example Lys835 in the *ctPrp22*–RNA complex (Hamann *et al.*, 2019). Superposition of the *ctPrp16* structure with the structure of the *ctPrp22*–RNA complex shows that the conserved Lys is located virtually in the same position in both structures, meaning that RNA binding seems to be unaffected by the unique conformation of the  $\beta$ -hairpin in *ctPrp16*.

### 3.3. Comparison with Prp16 structures emerging from cryo-EM 3D reconstructions

In recent years, five different structural models of spliceosomal C\* or C complexes from *Saccharomyces cerevisiae* and *Homo sapiens* containing Prp16 have been obtained by means of single-particle cryo-EM (Wilkinson *et al.*, 2021; Galej *et al.*, 2016; Yan *et al.*, 2017; Bertram *et al.*, 2020; Zhan *et al.*, 2018). In

the area of the spliceosome which was predicted to harbor Prp16, the authors claimed a local resolution ranging between 8 and 15 Å. Interestingly, in none of the structures was an additional cryo-EM map observed for the 300 N-terminal residues. In fact, the cryo-EM map implicated the presence of the helicase core of Prp16 or only parts of it. The N-terminal region of Prp16 could not be traced despite its function in the spliceosomal context (Wang & Guthrie, 1998). For a direct structural comparison with *ctPrp16* in complex with ADP, the cryo-EM 3D reconstructions for which a subnanometre local resolution of Prp16 was reported were chosen (PDB entries 7b9v and 5yzg). For further comparison, the structure of the closely related spliceosomal DEAH-box helicase Prp22 in its respective apo state was also used (PDB entry 6i3o). The overall alignment shows that the RecA1 and C-terminal domains are quite similar in their arrangement, while the location of the RecA2 domain differs dramatically. In contrast to the other structures, the RecA2 domain of ADP-bound *ctPrp16* is clearly shifted towards the RecA1 domain (Fig. 4*a*). To compare the exact distances between the RecA-like domains of each analyzed helicase, the center of mass of each RecA domain was determined and the distance between them



**Figure 4** Movement of RecA2 upon nucleotide binding. (a) Comparison of Prp16 models originating from X-ray diffraction and cryo-EM and Prp22 (in a nucleotide-free state). The different structures are shown as ribbons and are aligned via their RecA1-like domains. The RecA2 domain of Prp16–ADP is shifted towards the RecA1 domain compared with the other structures. (b) All structures are depicted as semi-transparent cartoon models. The RecA1 domain is colored orange, the RecA2 domain marine and the C-terminal part gray. The centers of mass of the RecA-like domains are displayed as spheres and colored accordingly. In order to calculate the centers of mass for the same sets of atoms, the RecA1 and RecA2 domains were superimposed and the centers of mass of these superimposed domains were determined. Prp16 structures derived from a cryo-EM model adopt a more open conformation that is comparable to the distances of Prp22 in a nucleotide-free state (PDB entry 6i3o) and the mammalian ortholog DHX38 of yeast Prp16 (bottom right).

was measured (Fig. 4b). While the structure of ADP-bound *ctPrp16* exhibits a center-of-mass distance of 28.1 Å between the two RecA domains, the distances calculated for Prp16 molecules derived from the cryo-EM structures of spliceosomes are increased by about 3 Å (31.1 Å in PDB entry 7b8v and 30.7 Å in PDB entry 5yzy). Therefore, it can be concluded that Prp16 in the spliceosome structures is unlikely to exhibit the ADP-bound post-catalytic state. In fact, the increased RecA distance fits better to the center-of-mass distance measured in the structure of Prp22 in its apo state (PDB entry 6io3, 31.0 Å). This is in line with the suggestion of Wilkinson and coworkers, who stated that they used the structure of scPrp43 in complex with ADP as the starting model but the corresponding map was more similar to the conformation of Prp22 bound to RNA found in PDB entry 6i3p (Wilkinson *et al.*, 2021; PDB entry 7b9v). Zhan and coworkers also used the structure of scPrp43 bound to ADP as the starting model but recognized that the RecA2 domain is shifted by a distance of 5–8 Å (Zhan *et al.*, 2018; PDB entry 5yzy). Therefore, it can be concluded that the model of Prp16 presented in this work is the first model of Prp16 with subnanometre resolution bound to a nucleotide.

## Acknowledgements

We are grateful for beamtime allocation at EMBL beamline P13, PETRA III, DESY, Hamburg, Germany. Open access funding enabled and organized by Projekt DEAL.

## Funding information

This work was supported by a grant from the Deutsche Forschungsgemeinschaft (DFG) to RF (SFB 860, TPA02).

## References

- Absmeier, E., Santos, K. F. & Wahl, M. C. (2020). *Structure*, **28**, 236–243.
- Amlacher, S., Sarges, P., Flemming, D., van Noort, V., Kunze, R., Devos, D. P., Arumugam, M., Bork, P. & Hurt, E. (2011). *Cell*, **146**, 277–289.
- Bernhofer, M., Dallago, C., Karl, T., Satagopam, V., Heinzinger, M., Littmann, M., Olenyi, T., Qiu, J., Schütze, K., Yachdav, G., Ashkenazy, H., Ben-Tal, N., Bromberg, Y., Goldberg, T., Kajan, L., O'Donoghue, S., Sander, C., Schafferhans, A., Schlessinger, A., Vriend, G., Mirdita, M., Gawron, P., Gu, W., Jarosz, Y., Trefois, C., Steinegger, M., Schneider, R. & Rost, B. (2021). *Nucleic Acids Res.* **49**, W535–W540.
- Bertram, K., El Ayoubi, L., Dybkov, O., Agafonov, D. E., Will, C. L., Hartmuth, K., Urlaub, H., Kastner, B., Stark, H. & Lührmann, R. (2020). *Mol. Cell*, **80**, 127–139.
- Burgess, S. M. & Guthrie, C. (1993). *Cell*, **73**, 1377–1391.
- Büttner, K., Nehring, S. & Hopfner, K. P. (2007). *Nat. Struct. Mol. Biol.* **14**, 647–652.
- Campodonico, E. & Schwer, B. (2002). *Genetics*, **160**, 407–415.
- Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst. D* **66**, 12–21.
- Cordin, O. & Beggs, J. D. (2013). *RNA Biol.* **10**, 83–95.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst. D* **66**, 486–501.
- Ficner, R., Dickmanns, A. & Neumann, P. (2017). *Methods*, **125**, 63–69.
- Galej, W. P., Wilkinson, M. E., Fica, S. M., Oubridge, C., Newman, A. J. & Nagai, K. (2016). *Nature*, **537**, 197–201.
- Hamann, F., Enders, M. & Ficner, R. (2019). *Nucleic Acids Res.* **47**, 4349–4362.
- Hamann, F., Schmitt, A., Favretto, F., Hofele, R., Neumann, P., Xiang, S., Urlaub, H., Zweckstetter, M. & Ficner, R. (2020). *Proc. Natl Acad. Sci. USA*, **117**, 2948–2956.
- He, Y., Andersen, G. R. & Nielsen, K. H. (2010). *EMBO Rep.* **11**, 180–186.
- Jankowsky, E. & Fairman, M. E. (2007). *Curr. Opin. Struct. Biol.* **17**, 316–324.
- Kabsch, W. (2010). *Acta Cryst. D* **66**, 125–132.
- Koodathingal, P., Novak, T., Piccirilli, J. A. & Staley, J. P. (2010). *Mol. Cell*, **39**, 385–395.
- Koodathingal, P. & Staley, J. P. (2013). *RNA Biol.* **10**, 1073–1079.
- Kudlinzki, D., Schmitt, A., Christian, H. & Ficner, R. (2012). *Biol. Chem.* **393**, 1131–1140.
- Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L.-W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V., Stockwell, D. H., Terwilliger, T. C., Urzhumtsev, A. G., Videau, L. L., Williams, C. J. & Adams, P. D. (2019). *Acta Cryst. D* **75**, 861–877.
- Lorson, C. L., Hahnen, E., Androphy, E. J. & Wirth, B. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 6307–6311.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst. D* **67**, 355–367.
- Schmitt, A., Hamann, F., Neumann, P. & Ficner, R. (2018). *Acta Cryst. D* **74**, 643–654.
- Schneider, S., Campodonico, E. & Schwer, B. (2004). *J. Biol. Chem.* **279**, 8617–8626.
- Schwer, B. & Guthrie, C. (1992). *EMBO J.* **11**, 5033–5039.
- Semlow, D. R., Blanco, M. R., Walter, N. G. & Staley, J. P. (2016). *Cell*, **164**, 985–998.
- Takeshima, Y., Yagi, M., Okizuka, Y., Awano, H., Zhang, Z., Yamauchi, Y., Nishio, H. & Matsuo, M. (2010). *J. Hum. Genet.* **55**, 379–388.
- Tauchert, M. J., Fourmann, J.-B., Christian, H., Lührmann, R. & Ficner, R. (2016). *Acta Cryst. F* **72**, 112–120.
- Tauchert, M. J., Fourmann, J.-B., Lührmann, R. & Ficner, R. (2017). *eLife*, **6**, e21510.
- Tseng, C. K., Liu, H. L. & Cheng, S. C. (2011). *RNA*, **17**, 145–154.
- Wahl, M. C., Will, C. L. & Lührmann, R. (2009). *Cell*, **136**, 701–718.
- Walbott, H., Mouffok, S., Capeyrou, R., Lebaron, S., Humbert, O., van Tilbeurgh, H., Henry, Y. & Leulliot, N. (2010). *EMBO J.* **29**, 2194–2204.
- Wang, Y. & Guthrie, C. (1998). *RNA*, **4**, 1216–1229.
- Wilkinson, M. E., Fica, S. M., Galej, W. P. & Nagai, K. (2021). *Mol. Cell*, **81**, 1439–1452.
- Will, C. L. & Lührmann, R. (2011). *Cold Spring Harb. Perspect. Biol.* **3**, a003707.
- Yan, C., Wan, R., Bai, R., Huang, G. & Shi, Y. (2017). *Science*, **355**, 149–155.
- Zhan, X., Yan, C., Zhang, X., Lei, J. & Shi, Y. (2018). *Science*, **359**, 537–545.