

Volume 74 (2018)

Supporting information for article:

Crystal structure of the spliceosomal DEAH-box ATPase Prp2

Andreas Schmitt, Florian Hamann, Piotr Neumann and Ralf Ficner

Table S1Overview of interdomain interactions. Contacts between the helicase core andthe C-terminal domains are listed for each structure and residues are colorcoded according toFig. 1. Salt bridges (SB) are highlighted with a cross.

Аро	CF1	CF2	CF3	
HC - CTD SB	HC - CTD SB	HC - CTD SB	HC - CTD SB	
Arg380 - Pro719 Arg401 - Gly716 Arg423 - Glu654 Arg423 - Asp682 X Lys435 - Asp679 X Tyr461 - Asp679 X Asn596 - Arg657 Thr652 Thr652 - Pro653 Thr652 Thr652 - Glu654	Arg380 - (Pro719) Arg401 Gly716 Arg423 Glu654 Arg423 Asp682 X Lys435 Asp679 X Tyr461 - Asp679 X Asn596 Arg657 Thr652 Thr652 - Pro653 Thr652 Thr652 - Glu654	Arg380 Pro719 Arg401 Gly716 Arg423 Glu654 Arg423 Asp679 X Lys45 Lys45 Asp679 X Tyr461 Asn596 Arg657 Thr652 Pro653 Thr652 Glu654	Arg380 Pro719 Arg401 Gly716 Arg423 Glu654 Arg423 Asp682 Lys35 Asp679 X Tyr461 Asp657 Thr652 Thr652 Arg657 Thr652 Glu654	Conserved Interdomain Interactions
Glu595 - Arg657 X Thr597 - Arg657 Gln606 - Asn870 Ser607 - Arg657 Ser607 - Ser659 Leu609 - Ceu897 Leu609 - Val899 Val611 - Val899	Glu595 - Arg657 X Phe598 - Arg657 X Val601 - Asn870 Gly602 - Lys691 Thr603 - Asn694 Thr603 - Lys869 Gln606 - Asn870 Leu609 - Leu897 Leu609 - Val899	Gly605 - <u>Asn870</u>) Gly605 - <u>Lys691</u>) Gln606 - <u>Lys869</u> Gln606 - <u>Asn870</u>) Ser607 - <u>Arn870</u> Thr608 - <u>Leu897</u> Leu609 - Leu897 Leu609 - <u>Val899</u> Val611 - <u>Val899</u>	Glu595 - Arg657 X Asn596 - Glu654 Asn596 - Thr658 Thr597 - Arg657 Phe598 - Arg657 Phe598 - Arg657 Pro600 - Lys691 Pro600 - Tyr894 Thr603 - His895 Thr603 - His895 Thr604 - Ser662 Gln606 - Glu721 Gln606 - His895 Val611 - Val899	β-hairpin Interactions
Phe381 - Gin818 Arg401 - Gin666 Val425 - Glu654 Lys457 - Asp682 X Gin516 - Arg811 Arg521 - Arg756	Arg401 Gln666 Arg401 Glu717 Arg401 Pro719 Lys457 Asp682 Glu528 Lys780 Glu552 - Arg807 Thr652 - Gln656	Arg353 Arg811 Glu382 Arg811 Arg401 Gln666 Arg401 Glu717 Arg401 Glu654 Val4257 Glu654 Lys457 Asp682 Gln516 Gly748 Gln516 Gly748 Glu517 Arg905 Glu524 Pro757 Glu524 Lys757 Glu524 Pro757 Glu524 Lys748 Asn548 Glu749 Asn548 Arg811 Glu522 Yarg170	Phe387 - Gin818 Val425 - Giu654 Lys457 - Leu681 Giu524 - Lys758 Asn548 - Arg811 Pro651 - Pro653	Remaining Interactions
24 Contacts	26 Contacts	37 Contacts	32 Contacts	-
(RecA1) (RecA2) (WH)	(HB) (OB)			

Helicase Core C-T (HC) do

C-Terminal domains (CTD)

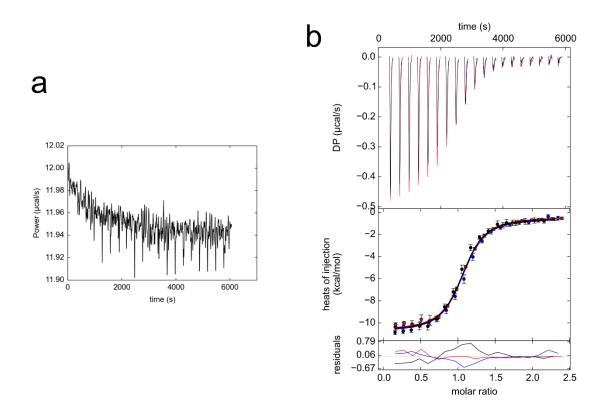
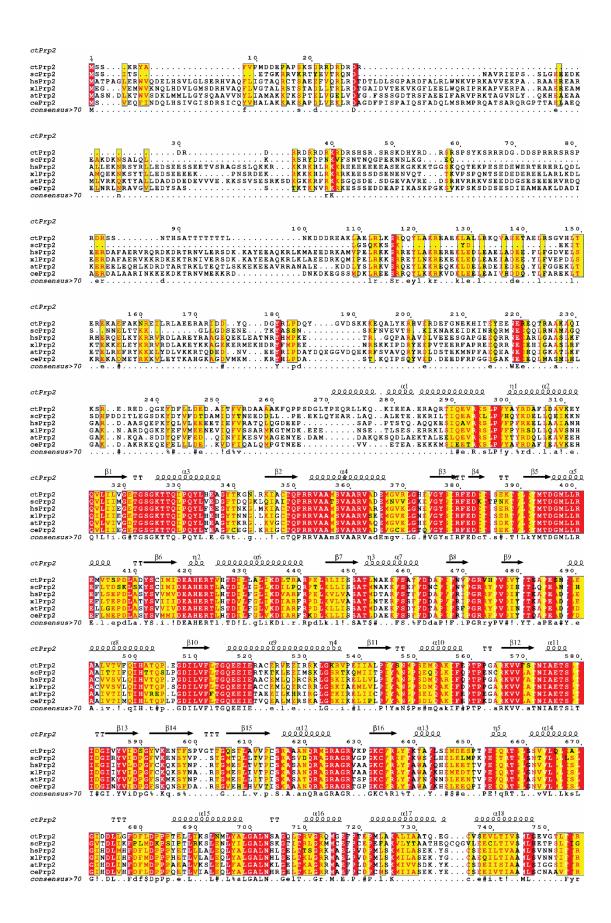


Figure S1 Isothermal titration calorimetry of Prp2 with AMPPCP and ADP. (a) The thermogram of the titration of AMPPCP to Prp2 did not result in thermic signals suited for integration, indicating there is no binding. (b) The binding of ADP was assessed and a dissociation constant (K_d) of 179 nM was determined (n= 1.01, Δ H= -10.63 kcal/mol, Δ S= - 5.06 cal/mol*K). Individual triplicates are depicted in red, blue and black. In the upper panel the thermograms are displayed, the middle panel shows the individual fittings with the correspondent residuals in the lower panel.



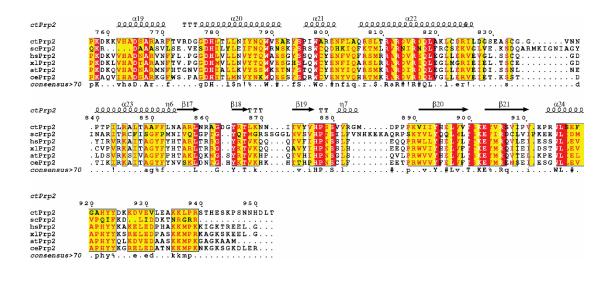


Figure S2 Sequence alignment of Prp2 from *Chaetomium thermophilum* (*ct*), *Saccharomyces cerevisiae* (*sc*), *Homo sapiens* (*hs*), *Xenopus laevis* (*xl*), *Arabidopsis thaliana* (*at*) and *Caenorhabditis elegans* (*ce*). Secondary structure elements from CF2 are displayed on top of the corresponding sequence.

ctPrp2 ctPrp2 scPrp2 ctPrp43 scPrp43 ctPrp16 scPrp16 scPrp16 ctPrp22 scPrp22 scPrp22 scPrp22 consensus>70	1 10 SSKRYAFVPMDDEPAP.S. SSITSETCK
ctPrp2	
ctPrp2 hsPrp2 ctPrp43 scPrp43 scPrp43 ctPrp16 scPrp16 hsPrp16 ctPrp22 scPrp22 hsPrp22 consensus>70	KRT
ctPrp2	
ctPrp2 scPrp2 tsPrp3 scPrp43 scPrp43 scPrp16 scPrp16 tsPrp16 ctPrp22 scPrp22 hsPrp22 consensus>70	YEVTRONDNA.VR. IEPSSLGEEEDKE. AK. KYARAAEREARAL. L EKNRSYRLEDSEESSE. ETV. SRAGSSLQKKRKKRKHLRKKREEBEEEEASEKGKKKTGGSK DSTCPKKI. APSINKRKKFK. KDDCEDKK. KSXVSSYRDWEESKDDQ. KDAEEEGGDQAGQNIRKD.RHYR. S KPPKSLEEKAAIFKGLAIPD. REIEPADDI.DDTLALLESLEPK. MKEEVSVLPSTKIPAK.IEAK.IE. EEVQKIESLD. KPKTEKEKLKELFPVLCQPDNPSVRTMLDEDDV.KVAVDVLKELEALMPSAAGQEKQRD.AEHR
ctPrp2	20 30
ctPrp2 scPrp2 thPrp3 scPrp43 scPrp43 tPrp16 scPrp16 scPrp16 ctPrp22 scPrp22 scPrp22 consensus>70	K.S. DRRDRDR DRDRDR. RD. R. QQTEKP ES E DEWERTERERLQDLEEEDAFAERVR Q. RD. K. KRAGT DGKDRDR. DRDREDRS.KD. R. K. K. KRAGT DGKDRDR. DRDREDRS.KD. R. K. KRAGT DGKDRDR. DRDREDRS.KD. R. KRAGT DYDKKR.D. R. DYDKKR.D. R. RADK DRPSRKR SRSPED Y. DGSYRSR. DRDRER.N. RDR. DRTKKKKKRSRS RDRNRDR. DRDRER.N. RDR DRDRER.N. RDR
ctPrp2	
ctPrp2 scPrp2 tsPrp2 ctPrp43 scPrp43 tsPrp16 scPrp16 scPrp16 hsPrp16 tsPrp16 tsPrp22 scPrp22 consensus>70	40 50 60 70 RDRKRDRSHSRSKD. HYRDRSRSKV SRRDGDDSP. VFSN. TNQGPE KNNLKGEQLG. DR.TRNVLERS
ctPrp2	
ctPrp2 scPrp2 tsPrp2 ctPrp43 scPrp43 ctPrp16 scPrp16 hsPrp16 tsPrp16 scPrp22 scPrp22 scPrp22 consensus>70	80, 90, 100. RRRSRS.PRDRSSNTHSA.TTTTTTTLN SQKKSS.K.YDEKITSNNELTT.
ctPrp2	
ctPrp2 scPrp2 hsPrp2 ctPrp43 scPrp43 scPrp43 ctPrp16 scPrp16 ctPrp22 scPrp22 hsPrp22	110 120 130 140 K.DDDREAKLAELRLKSRQQYLAKREA.EKLALLRKQVAEETAEKKGLLGDS D D D D D D D D D C C C C C C C C C C C C C
consensus>70	

ctPrp2	150 160
ctPrp2 sePrp2 thPrp3 sePrp43 thPrp43 ttPrp16 sePrp16 hsPrp16 ctPrp22 sePrp22 sePrp22 consensus>70	LRSGVHLTEREKAEFAKNREIL. R. LRSGVHLTEREKAEFAKNREIL. R. KKAY. EEQQRLKMAEEDRKAMVP. SK. RVK. TNGDKMDAKNPYLAH QA. AEI AEELSKQHPLPSEEP RA STN A ERERA. LAEKKMN.MLSRGMMNARQLQKQKDVDAW. ETN R. LPVTRNIDNDALR. NTV Q RREQH. LHKQKQK. RISAQRQINEDNEW ETN R. SRRGIGGS. DGAANGP.PLSG.ATRQCKRLTSPERW EIR Q RNDDRPTH. LSLVSAP.EVEDSLERKRLTQISDPEKW EIK Q
ctPrp2	170
ctPrp2 scPrp2 tsPrp2 ctPrp43 tsPrp43 tsPrp43 tsPrp16 scPrp16 ctPrp16 ctPrp22 scPrp22 hsPrp22 consensus>70	LLEERARIDD QLRNAM.AGQSDHPDDITLE.GS LEERKSRE.YL LEERKPE LY.HHDA M.LTSAG M.LTSAG M.LTSAG LYPIPLKOR LYPIPLKOR M.LTSGVURREV LYPIPLKOR M.LTSGVURREV LYPIPLKOR M.LTSGVURREV LYPIPLKOR M.LTSGVURREV LYPIPLKOR M.LTSGVURREV LORDFED.N.AAKVHLWH.NLVPPLCGRUV.FTKQ. M.LTSGVURREV LORDFED.N.AAKVHLWH.NLVPPLCGRUV.FTKQ. M.LTSGVURREV M.L
ctPrp2	180 190 200 210 220
ctPrp2 ssPrp2 tsPrp3 ssPrp43 tsPrp43 tsPrp46 ssPrp16 ssPrp16 tsPrp16 ctPrp22 ssPrp22 consensus>70	DGYRLPD.QYGVDSKKKEQALY.K HVER.DEFONEKHI.TEYEEWERE.QTRAAKA EEKLQYEAR KKUQAHATSAAGE.QEKLE.ATNAYHMPKETRGQPARAVDLVEEESGAPG.EEQRRWEEA.RLGAASL ESLLDKLE RRNTT.AK KGLQAHATS.AE HSAHSTHSTBSA.HSTH.S.AE VDPVPAVRDFTSDMAVFSRKGSKVV.RERAQR.ER VDPVPAVRDFTSDMAVFSRKGSKVV.RERAQR.ER LSSLDKUPFNPDSEFSANAKRGSKLV.ALRRINM.EH LSPIRVVKAPDGSLNRAAMTGAQLA.RE.REAR.QQ KSPIRVVKAPDGSLNRAAMTGAQLA.RE.REAR.QQ MSPIKIVKNPDGSLSQAAMMQSALA.KE RELK.QA OR. EAEMDSTPMGLNKHWV.
ctPrp2	الًا المَّذَ مَنْ مَنْ مَنْ مَنْ مَنْ مَنْ مَنْ مَن
ctPrp2 scPrp2 tsPrp2 ctPrp43 tsPrp43 tsPrp43 scPrp16 scPrp16 scPrp16 scPrp22 scPrp22 consensus>70	QIKSRERE DQGEVDFLLDEDAI.TFVRDA.AAAKFQPPSDGLTPEQRL.LKQNIEEAERAQAT QEVIKS PYAYHDAFID AVKEY KFGARDAASQEPKYQLVLEEEETTEFVRATQLQG.DEEPSAPTSTQAQQKES QAVKRS DFPFPEELIAIIAN VEDSGI.NPFTSTQAQQKES QAVKRS DFPFPEELIAIIAN LEDGKI.NPFTCGEFTPKVD LKKRE DFNACHDEFKILVGN LP.QCI.NPFTNLPHTPRVD LKKRE DFNACHDEFKILVGN LNGVKEDDTDSAVFI.ASE.DASKAQSSS.GTSKFS.DAIK.CODTALFTPSKDIKKTIG.DFVEDLKVIEDL LNGVKEDDTDSAVFI.ASE.DASKAQSSS.GTSKFS.DAIK.KSCASSFS.DAIK.LQCD DFNC
ctPrp2	$\begin{array}{cccccccccccccccccccccccccccccccccccc$
ctPrp2 scPrp2 tsPrp2 ctPrp43 tsPrp43 tsPrp43 tsPrp45 scPrp16 scPrp16 ctPrp16 ctPrp22 scPrp22 tsPrp22 consensus>70	QVLIVGETGSGKTTOIP OVLHEASYTKGNRKACTOPRRVAASVA VABSVASVASVASVASVASVASVASVASVASVASVASVASVAS
ctPrp2	$\begin{array}{cccccccccccccccccccccccccccccccccccc$
ctPrp2 scPrp2 tsPrp2 ctPrp43 scPrp43 tsPrp43 ctPrp16 scPrp16 hsPrp16 ctPrp22 scPrp22 hsPrp22 consensus>70	400 410 420 420 430 440 460 470 460 470 480. I REM TS PDIA DY CIMPEAHERTY HOIT ILA ING TRADELERIIS SATANAKT DIA THNYE GYNEV VYY SATESN I REF TDSKISKYCIM DEAHERTIA DIILGILKO VART DELTINISSATANAKT DIA THNYE GYNEV VYY SATESN I REF SEDDIA SY VYMYDEAHERTIA DIILGILKO VART DELTINISSATANAKT DIA THNYE GYNEV VIHYYI GHEAN I REF SEDDIA SY VYMYDEAHERTIA DIILGIKO VART DELTINISSATANAKT DIA TYDA THYE GYNEV VIHYYI GHEAN I REF SEDDIA SY VYMYDEAHERTIA DIILGIKO VART DELTINISSATANAKT DIA TYDA THYE GYNEV VIHYYI GHEAN I REF SEDDIA SY VYMYDEAHERTIA DIILGIKO VART DELTINISSATANAKT DIA SY TA DA LA VEGTI YVIHYI GHEAN I REF SEDDIA SY VYMYDEAHERTIA DIILGIKO VART DELTINISATAN DIA SY TA DA LA VEGTI YVIHY EF I REA MEDHISSY CIIIDEAHERTIA DIILGIKOV VK SEDUNIN SATADIA SY TA DA LA VEGTI YVIHY EF I REA MEDHISSY CIIIDEAHERTIA DIILGIKOV VK SEDUNIN SATADIA SY TA DA LA VEGTI YVIHY EF I REA SATEDHISSY CIIIDEAHERTIA DIILGIFKI VR SEDUNIN SATADIA SY TA DA LA VEGTI YVIHY EF I RES SATAN DILE SY CIIIDEAHERTIA DIILGIFKI LA SATANA I REF YT DIA SY VIIDEAHERTIA DIILGIFKI LA SATANAKT DI SATANAKT DI SY TA DA LA VEGTI YVIHY EF I RES SATANAY I REF SATANAY SATANAKT SATANAY SA
ctPrp2	$ \xrightarrow{\beta 10} \qquad \xrightarrow{\beta 11} \qquad \xrightarrow{\beta 10} \qquad \beta$
ctPrp2 sePrp2 ctPrp43 sePrp43 hsPrp43 ctPrp16 scPrp16 ctPrp16 ctPrp22 scPrp22 consensus>70	490 500 510 520 530 540 550 560 MERARIY TYPOINATO, PEGDITUSEIT SEET RACERVENTRERAL

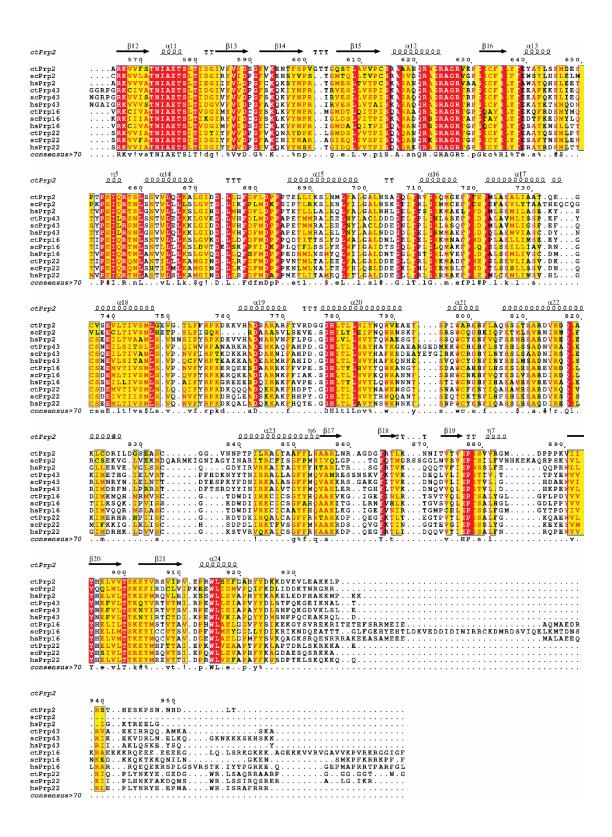


Figure S3 Sequence alignment of Prp2, Prp43, Prp16 and Prp22 from *Chaetomium thermophilum* (*ct*), *Saccharomyces cerevisiae* (*sc*) and *Homo sapiens* (*hs*). Secondary structure elements from CF2 are displayed on top of the corresponding sequence.

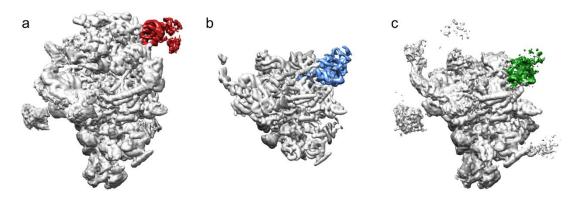


Figure S4 Position of DEAH-box helicases in the spliceosome. The cryo-EM maps of the *Saccharomyces cerevisiae* B^{act} complex (a, Yan *et al.*, 2016), *Saccharomyces cerevisiae* post-catalytic P complex (b, Liu *et al.*, 2017) and *Saccharomyces cerevisiae* spliceosome complex remodelled for exon ligation (c, Fica *et al.*, 2017) superimposed and are shown in grey and DEAH-box helicases of each complex are depicted in red (Prp2), blue (Prp22) and green (Prp16), respectively. All DEAH-box helicases are found in the periphery of their respective spliceosomal complex.



Figure S5 Secondary structure prediction of N-terminal extension of *Chaetomium thermophilum* Prp2. The secondary structure prediction was performed using the GeneSilico metaserver (Kurowski & Bujnicki, 2003). Used prediction server for the metasearch are stated the left followed by the corresponding prediction on the right. Residues predicted to be part of a helical secondary structure are highlighted by "H" whereas residues predicted to be part of beta strands are displayed as "E".