

ACCEPTED MANUSCRIPT



A composite double-/single-stranded RNA-binding region in protein Prp3 supports tri-snRNP stability and splicing

Sunbin Liu, Sina Mozaffari-Jovin, Jan Wollenhaupt, Karine F Santos, Matthias Theuser, Stanislaw Dunin-Horkawicz, Patrizia Fabrizio, Janusz M Bujnicki, Reinhard Lührmann, Markus C Wahl

DOI: <http://dx.doi.org/10.7554/eLife.07320>

Cite as: eLife 2015;10.7554/eLife.07320

Received: 5 March 2015

Accepted: 9 July 2015

Published: 10 July 2015

This PDF is the version of the article that was accepted for publication after peer review. Fully formatted HTML, PDF, and XML versions will be made available after technical processing, editing, and proofing.

Stay current on the latest in life science and biomedical research from eLife.
Sign up for alerts at elifesciences.org

**A composite double-/single-stranded RNA-binding region in protein
Prp3 supports tri-snRNP stability and splicing**

Sunbin Liu¹, Sina Mozaffari-Jovin², Jan Wollenhaupt¹, Karine F. Santos¹, Matthias Theuser¹,
Stanislaw Dunin-Horkawicz³, Patrizia Fabrizio², Janusz M. Bujnicki^{3,4}, Reinhard Lührmann^{2,*},
Markus C. Wahl^{1,*}

¹ Laboratory of Structural Biochemistry, Freie Universität Berlin, Takustr. 6, D-14195 Berlin,
Germany

² Dept. of Cellular Biochemistry, Max Planck Institute for Biophysical Chemistry, Am
Fassberg 11, D-37077 Göttingen, Germany

³ International Institute of Molecular and Cell Biology, Laboratory of Bioinformatics and
Protein Engineering, 4 Ks. Trojdena Street, 02-109 Warsaw, Poland

⁴ Institute of Molecular Biology and Biotechnology, Faculty of Biology, Adam Mickiewicz
University, ul. Umultowska 89, PL-61-614 Poznan, Poland

*Corresponding authors:

Reinhard Lührmann (reinhard.luehrmann@mpi-bpc.mpg.de)

Markus C. Wahl (mwahl@zedat.fu-berlin.de)

Submitting author:

Prof. Markus C. Wahl, PhD

Freie Universität Berlin

Institute of Chemistry and Biochemistry

Laboratory of Structural Biochemistry

Takustr. 6

D-14195 Berlin

Germany

32

Tel: +49 30 838 53456

Fax: +49 30 8384 53456

Email: mwahl@zedat.fu-berlin.de

36

37

38

Abstract

Prp3 is an essential U4/U6 di-snRNP-associated protein whose functions and molecular mechanisms in pre-mRNA splicing are presently poorly understood. We show by structural and biochemical analyses that Prp3 contains a bipartite U4/U6 di-snRNA-binding region comprising an expanded ferredoxin-like fold, which recognizes a 3'-overhang of U6 snRNA, and a preceding peptide, which binds U4/U6 stem II. Phylogenetic analyses revealed that the single-stranded RNA-binding domain is exclusively found in Prp3 orthologs, thus qualifying as a spliceosome-specific RNA interaction module. The composite double-stranded/single-stranded RNA-binding region assembles cooperatively with Snu13 and Prp31 on U4/U6 di-snRNAs and inhibits Brr2-mediated U4/U6 di-snRNA unwinding in vitro. RNP-disrupting mutations in Prp3 lead to U4/U6•U5 tri-snRNP assembly and splicing defects in vivo. Our results reveal how Prp3 acts as an important bridge between U4/U6 and U5 in the tri-snRNP and comparison with a Prp24-U6 snRNA recycling complex suggests how Prp3 may be involved in U4/U6 re-assembly after splicing.

Impact statement

Structural and functional analyses show how the spliceosomal Prp3 protein concomitantly binds double- and single-stranded regions in U4/U6 di-snRNAs and serves to stabilize the U4/U6•U5 tri-snRNP for splicing.

Introduction

Pre-mRNA splicing is catalyzed by a multi-subunit RNA-protein (RNP) enzyme, the spliceosome, which facilitates two successive transesterification reactions (steps 1 and 2) that lead to the removal of an intron and the ligation of its flanking exons. For each splicing event, a spliceosome is newly formed via the stepwise recruitment of small nuclear (sn) RNPs (U1, U2, U4, U5 and U6 in the case of the major spliceosome) and numerous non-snRNP proteins to a pre-mRNA substrate (Wahl et al., 2009; Will & Lührmann, 2011). During canonical cross-intron spliceosome assembly, U1 and U2 snRNPs bind the 5'-splice site (5'SS) and the branch point region of an intron, respectively (Mount et al., 1983; Kramer et al., 1984; Parker et al., 1987; Wu & Manley, 1989; Zhuang & Weiner, 1989). Formation of this A complex is followed by the recruitment of a preformed U4/U6•U5 tri-snRNP, yielding the pre-catalytic B complex (Bindereif & Green, 1987; Cheng & Abelson, 1987; Konarska & Sharp, 1987; Deckert et al., 2006; Fabrizio et al., 2009), which requires extensive conformational and compositional rearrangements to form a catalytically active spliceosome. Catalytic activation includes the disruption of the U1/5'SS interaction (Konforti et al., 1993) and the separation of U4 snRNA from U6 snRNA (Konarska & Sharp, 1987; Yean & Lin, 1991), which are extensively base-paired in the tri-snRNP and in the B complex, leading to the dissociation of U1 snRNP, U4 snRNA and all U4/U6-associated proteins. These rearrangements give rise to the B^{act} complex (Fabrizio et al., 2009; Bessonov et al., 2010) and, upon further remodeling, to the B* complex (Warkocki et al., 2009), in which U6 snRNA forms a central part of the spliceosome's active site. The B* complex facilitates the first step of splicing, the ensuing C complex (Konarska et al., 2006; Bessonov et al., 2008; Fabrizio et al., 2009) catalyzes the second step of splicing, after which the spliceosome releases its products and the remaining snRNPs and non-snRNP factors are recycled.

To participate in further rounds of splicing, the U4/U6•U5 tri-snRNP must be re-assembled by initial dimerization of U4 and U6 snRNPs followed by association with U5 snRNP (Stanek & Neugebauer, 2006). Association of U4 and U6 snRNPs is mediated in part by base pairing between their respective snRNAs, which form two inter-molecular helices (stems I and II) that are separated by a U4 5'-stem-loop (5'SL; Figure 1A). U4/U6 base pairing is mutually exclusive with the U6 snRNA conformation in the activated spliceosome. Re-annealing of U4 and U6 snRNAs after splicing thus requires the Prp24 assembly chaperone in yeast (Raghunathan & Guthrie, 1998) or its SART3 ortholog in human (Bell et al., 2002), which transiently bind U6 snRNA, as well as the LSm proteins (Achsel et al., 1999; Rader & Guthrie, 2002; Verdone et al., 2004), which bind and remain at the 3'-end of U6 snRNA (Beggs, 2005; Zhou et al., 2014). In addition, the U4-specific Prp3 protein is required for U4/U6 di-snRNP and U4/U6•U5 tri-snRNP formation (Anthony et al., 1997) but molecular mechanisms underlying its functions are poorly understood.

Human (h) and yeast (y) Prp3 form a complex with the respective Prp4 proteins (Ayadi et al., 1998; Gonzalez-Santos et al., 2002). hPrp3 can be crosslinked to the U6 snRNA portion of a U4/U6 di-snRNA complex comprising the U4 5'SL, an intact stem II and a U6 3'-overhang (Nottrott et al., 2002) and can pull down U4 and U6 snRNAs from nuclear extract (Gonzalez-Santos et al., 2002), pointing towards direct Prp3-snRNA interactions. hPrp3 also interacts with the U5-specific proteins hPrp6 and hSnu66 (Liu et al., 2006).

Except for an N-terminal region, Prp3 is highly conserved from yeast to human and contains a C-terminal domain of unknown function (DUF1115; PFAM ID PF06544; Figure 1A; Figure 1 – figure supplement 1). Recent homology modeling has predicted a ferredoxin-like fold for the human Prp3 DUF1115 domain (Korneta et al., 2012). Here, we demonstrate that C-terminal regions of Prp3, including DUF1115, bind U4/U6 di-snRNA fragments containing stem II and a U6 3'-overhang and elucidated crystal structures of a yPrp3 C-terminal region alone and in complex with a U4/U6 di-snRNA fragment. Structure-guided mutations that led to reduced

U4/U6 interaction in vitro and to reduced cell viability also reduced U4/U6•U5 tri-snRNP levels and splicing in vivo. Our results indicate how Prp3 functionally bridges U4/U6 and U5 in the tri-snRNP by Prp3-RNA interactions on one side and Prp3-protein interactions on the other. Moreover, a comparison with the structure of a Prp24-U6 snRNA complex (Montemayor et al., 2014) suggests how Prp3 may initiate the handover of U6 snRNA from the Prp24 recycling factor to U4 snRNP during U4/U6 re-assembly.

Results

Characterization of a conserved C-terminal U4/U6 di-snRNA-binding region in Prp3

Previous studies have shown that human (h) Prp3 in the context of a hPrp3-hPrp4-hCypH complex contacts U6 snRNA in a region that forms stem II and a U6 single stranded 3'-overhang (Nottrott et al., 2002). To test whether hPrp3 alone is sufficient for stable RNA binding and to delineate hPrp3 elements required for complex formation, we produced full-length hPrp3 (hPrp3^{FL}) and fragments lacking ca. 200-250 residues from either end (hPrp3^N – residues 1-442; hPrp3^C – residues 195-683) as N-terminal maltose-binding protein [MBP] fusion proteins. We then tested binding of these proteins to a hU4/U6 construct containing stem II (fused by a GAAA tetraloop) and a 13 nucleotide [nt] U6 3'-overhang (hU4/U6^{stem II+13nt}) in electrophoretic mobility shift assays (EMSAs). Only the full-length protein and hPrp3^C, but not hPrp3^N, bound hU4/U6^{stem II+13nt} (Figure 1B, lanes 2-4). Trypsin treatment of the hPrp3^C-hU4/U6^{stem II+13nt} complex gave rise to a protein fragment containing residues 484 to 683 (hPrp3^{CTF}) as shown by mass spectrometric fingerprinting and N-terminal sequencing. hPrp3^{CTF} contains the predicted DUF1115 domain (residues 540-683) and a conserved, preceding peptide rich in basic amino acid residues (Figure 1A; Figure 1 – figure supplement 1). Binding of recombinant hPrp3^{CTF} (as a glutathione S-transferase [GST] fusion) to hU4/U6^{stem II+13nt} was comparable to hPrp3^{FL} or

hPrp3^C (Figure 1B, lane 5). These results show that the C-terminal ca. 200 amino acids of hPrp3 encompass the protein elements that mediate stable U4/U6 binding.

All following experiments were performed with yeast (y) factors. To test whether the C-terminal U4/U6 di-snRNA-binding region is conserved in yPrp3, we produced a protein comprising the 177 C-terminal residues of yPrp3 (residue 296-469; yPrp3^{CTF}). yPrp3^{CTF} bound a yU4/U6 duplex containing the complete stem II and a 13 nt yU6 3'-overhang (yU4/U6^{stem II+13nt}) with an apparent dissociation constant ($K_{d,app}$) of 0.92 μ M as determined by EMSA (Figure 1C, first gel and quantification) and with a K_d of 110 nM as determined by isothermal titration calorimetry (ITC; Figure 1D, top left). The lower apparent affinity estimated by EMSA is likely due to the presence of competing non-specific tRNA competitor in this assay. The yPrp3 DUF1115 domain (residues 325-469; yPrp3^{DUF1115}) lacking the preceding basic peptide showed ca. twofold reduced affinity for yU4/U6^{stem II+13nt} in EMSA ($K_{d,app}$ 1.83 μ M; Figure 1C, second gel and quantification) and a ca. 3.5-fold lower affinity in ITC (K_d 380 μ M; Figure 1D, top right). Exchange of conserved arginine residues at positions 304 and 322 in the N-terminal basic peptide of yPrp3^{CTF} (Figure 1 – figure supplement 1), which could directly interact with the negatively charged RNA backbone, reduced affinities for yU4/U6^{stem II+13nt} in both EMSA ($K_{d,app}$ 1.58 μ M and 2.63 μ M for yPrp3^{CTF,R304A} and yPrp3^{CTF,R322A}, respectively; Figure 1C, third and fourth gels and quantification) and ITC (K_d 260 nM and 400 nM for yPrp3^{CTF,R304A} and yPrp3^{CTF,R322A}, respectively; Figure 1D, bottom left and right) to a similar extent as removal of the entire preceding peptide. Furthermore, the thermodynamic signatures of the interactions involving yPrp3^{DUF1115}, yPrp3^{CTF,R304A} and yPrp3^{CTF,R322A} changed considerably compared to yPrp3^{CTF}. The DUF1115 domain as well as both arginine-to-alanine variants (in particular yPrp3^{CTF,R304A}) exhibited significantly less favorable interaction enthalpies and more favorable (or less unfavorable) interaction entropies compared to yPrp3^{CTF} (Figure 1D). One explanation for these observations could be that the N-terminal peptide is a flexible element in yPrp3^{CTF}, which becomes immobilized (loss in conformational entropy) by RNA contacts (gain in

interaction enthalpy) upon binding of yU4/U6^{stem II+13nt}. In any case, these observations show that both the DUF1115 domain and the preceding basic peptide contribute to the RNA binding.

Next, we further probed the RNA requirements for stable binding. yPrp3^{CTF} efficiently bound a yU4/U6 constructs bearing full-length stem II and yU6 3'-overhangs of at least eight nts (Figure 1E, lanes 1-8), while further shortening of the yU6 3'-overhang led to progressively reduced binding (lanes 9-14). Removal of the first two A-U base pairs from the 5'-end of stem II had no consequence for binding of yPrp3^{CTF} (Figure 1E, lanes 15-18), but reduced binding was seen when seven base pairs, including a non-canonical C69^{U6}-C12^{U4} pair, were removed (Figure 1E, lane 20). A non-canonical pyrimidine-pyrimidine base pairs is conserved in the corresponding stem II regions of U4/U6 from other organisms (Figure 1 – figure supplement 2). When we converted yU6 C69 or yU4 C12 of the non-canonical C-C pair in stem II to G's to allow Watson-Crick base pairing at these positions, the affinity of yPrp3^{CTF} was reduced about twofold in EMSA titrations (Figure 1F, first and third gels and quantification). In contrast, the DUF1115 domain alone did not exhibit reduced affinity to the mutant RNAs compared to the wt (Figure 1F, second and fourth gels and quantification). These findings suggest that yPrp3^{CTF} recognizes portions of the U6 3'-overhang as well as parts of yU4/U6 stem II, including the non-canonical C69^{U6}-C12^{U4} pair.

Binding of Prp3 to U4/U6 stem II would be expected to stabilize the U4/U6 duplex, which is unwound by the Brr2 helicase during spliceosome catalytic activation. To test relative contributions of the yPrp3 DUF1115 domain and the preceding basic peptide to stabilization of yU4/U6, we therefore assessed the influence of yPrp3^{CTF} and yPrp3^{DUF1115} on Brr2-mediated U4/U6 unwinding. Addition of yPrp3^{CTF} reduced the rate of yBrr2-mediated yU4/U6 unwinding about five-fold, while yPrp3^{DUF1115} showed almost no effect (ca. 1.2-fold reduction; Figure 2A,B).

Together, these analyses show that Prp3 orthologs contain a conserved, C-terminal region (Prp3^{CTF}) that is necessary and sufficient for stable binding to U4/U6 di-snRNAs and that upper parts of stem II and at least eight nts of the U6 3'-overhang are required for stable binding.

Prp3^{CTF} comprises the DUF1115 domain and a preceding basic peptide, both of which contribute to RNA binding.

Cross-talk of U4/U6-bound proteins

Previous assembly and structural studies had shown that the U4/U6-specific proteins Snu13, Prp31 and Prp3 bind U4/U6 di-snRNAs in a cooperative manner (Nottrott et al., 2002; Liu et al., 2007). To see whether the C-terminal U4/U6-binding region of Prp3 is sufficient to sense pre-bound Snu13 and Prp31 proteins, we compared binding of yPrp3^{CTF} and yPrp3^{DUF1115} to a fused yU4/U6-like RNA (yU4/U6^{fused}; Figure 2C) alone or pre-bound to ySnu13 or ySnu13 and yPrp31¹⁻⁴⁶² under identical conditions. Both yPrp3 variants bound the ySnu13-yPrp31¹⁻⁴⁶²-yU4/U6^{fused} complex (Figure 2C, lanes 13-18, right) more efficiently than the ySnu13-yU4/U6^{fused} complex (lanes 7-12, middle), which in turn was bound better than the naked RNA (lanes 1-6, left). However, yPrp3^{CTF} interacted more stably than yPrp3^{DUF1115} with the naked RNA (lanes 4-6 vs. lanes 1-3) and with the ySnu13-yU4/U6^{fused} complex (lanes 10-12 vs. lanes 7-9). The sensitivity of the assay was insufficient to resolve possible binding differences to the ySnu13-yPrp31¹⁻⁴⁶²-yU4/U6^{fused} complex (lanes 13-18). These results show that yPrp3^{CTF} binds cooperatively with ySnu13 and yPrp31 to yU4/U6, in part based on the basic peptide preceding the DUF1115 domain.

Crystal structure of the C-terminal yU4/U6-binding region of yPrp3

Sequence analyses of the C-terminal U4/U6-binding portions of Prp3 proteins did not reveal any obvious similarities to known RNA-binding domains. We therefore determined the crystal structure of yPrp3^{CTF} at 2.7 Å resolution (Table 1). The protein crystallized with three monomers per asymmetric unit (Figure 3 – figure supplement 1A). The electron density allowed modeling of residues 335 to 467, corresponding to the predicted DUF1115 domain. Apart from the C-terminal two residues, the N-terminal 39 residues (296-334; comprising the preceding basic

peptide) could not be traced, indicating that they indeed constitute an intrinsically disordered or flexibly attached element, as surmised based on the ITC experiments.

The core of the yPrp3 DUF1115 domain resembles the homology model of the corresponding human domain (Korneta et al., 2012) (root-mean-square deviation [RMSD] of 2 Å for 87 common Cα atoms). The structure exhibits a five-stranded mixed β-sheet with two α-helices (α1 and α2) running parallel to the β-strands on one side of the sheet and one (α3) on the other (Figure 3A). The first 98 residues of the structure (residues 335–432) form a α/β sandwich with a ferredoxin-like topology and an additional, long β-hairpin (residues 403–421, comprising strands β3a and β3b) inserted between helix α2 and strand β4. The ferredoxin-like fold is further expanded by an extra β-strand (β5), a helix (α3) and a following loop at the C-terminus. In two of the three independent molecules, the β3a/β3b hairpins adopt very similar conformations with their long axes oriented perpendicular to the direction of the strands in the central β-sheet. Residues D405 and D418 at the bases of the hairpins coordinate an Yt³⁺ ion. The third molecule lacks electron density for large parts of the hairpin (residues 407–421), suggesting that crystal packing and a bound metal ion may have stabilized this normally flexible element in the first two molecules.

We also determined the crystal structure of yPrp3^{DUF1115}, lacking the preceding basic peptide, at 2.0 Å resolution (Table 1). While the ordered parts of yPrp3^{CTF} and yPrp3^{DUF1115} are very similar (RMSD of 0.48 Å for 131 common Cα atoms; Figure 3 – figure supplement 1A), the β3a/β3b hairpin in yPrp3^{DUF1115} is positioned closer to the globular part of the protein, showing that its relative positioning is indeed flexible.

Crystal structure of a yPrp3^{CTF}-yU4/U6^{stem II+10nt} complex

To elucidate the molecular mechanism underlying U4/U6 di-snRNA binding by yPrp3^{CTF}, we determined its crystal structure in complex with yU4/U6^{stem II+10nt} at 3.25 Å resolution (Table 1; Figure 3B). Residues 335–468 of yPrp3^{CTF} and all nts (G81–U90) of the yU6 3'-overhang were

well ordered in the two complexes contained in an asymmetric unit. The electron density for the yU4/U6 stem II regions was less well defined, in particular in the part of stem II distal to the duplex-to-single strand junction in one of the complexes. The electron density was consistent with the stem II regions adopting standard A-form geometry in both complexes but did not allow us to model in detail possible local deviations, e.g. around a non-canonical C69^{U6}-C12^{U4} base pair.

Protein-RNA interactions in the well-defined portions of the two crystallographically independent complexes were largely identical (Figure 3 – figure supplement 1B). In both complexes, the single-stranded (ss) yU6 3'-overhang arches across helix α 1 and strand β 5 of yPrp3^{CTF}, running below the long β 3a/ β 3b hairpin (Figure 3B). Its binding surface on the protein is carpeted by an electropositive surface potential (Figure 3C). The yU6 3'-overhang can be divided into three sections, between which its backbone changes directions on the surface of yPrp3^{CTF} (Figure 3B). Nts G81-A83 (section 1) stack on the terminal U80^{U6}-A1^{U4} base pair of stem II and run diagonally from the N-terminus of helix α 2 to the N-terminus of helix α 1; nts C84-G86 (section 2) are positioned perpendicular across the N-terminal end of helix α 1; nts U87-U90 (section 3) run along the exposed edge of strand β 5 and the C-terminus of helix α 1. As a consequence, the end of the U6 3'-overhang is directed via two ca. 90 ° kinks back towards stem II. A RNA structural similarity search using the RNA Bricks database (Chojnowski et al., 2014) failed to uncover a case of a similar RNA re-direction on the surface of a single protein domain. Notably, the unusual doubly-kinked RNA conformation is stabilized by protein elements that expand the core ferredoxin fold in yPrp3, i.e. the β 3a/ β 3b hairpin, strand β 5 and helix α 3 (details below).

All nts of the yU6 3'-overhang (G81-U90) contained in the present structure, except the first G81 residue, directly contact yPrp3^{CTF} in one or both complexes, consistent with the important role of this part of U6 for stable Prp3 binding in yeast and human. In one yPrp3^{CTF}-yU4/U6^{stem II+10nt} complex, the guanidinium group of R399 (helix α 2) forms ionic interactions to the backbone

265 phosphate of A82 and the side chain amino group of K355 (helix α 1) binds to the A83
266 phosphate (Figure 4A). In the complex lacking these interactions, the corresponding backbone
267 regions of yU6 and the stem II duplex are slightly pulled away from the protein, presumably by
268 crystal packing interactions. The exocyclic N6 group of A83 approaches the side chain carboxyl
269 of E362 (helix α 1) and its base stacks on the side chain of F354 (helix α 1; Figure 4A). The
270 extracyclic amino group (N4) of C84 is hydrogen bonded to the backbone carbonyl of E407
271 (β 3a/ β 3b hairpin) and the base is thereby held sandwiched between H409 (β 3a/ β 3b hairpin)
272 and P350 (helix α 1; Figure 4B). The A83 and C84 backbone phosphates approach the side
273 chain of K351 (helix α 1; Figure 4B). These interactions splay out A83 and C84, stabilizing the
274 first kink in the yU6 3'-overhang. The guanidinium group of R353 (helix α 1) engages in
275 hydrogen bonds to the Watson-Crick face of C85 (atoms O2 and N3) and to the Hoogsteen face
276 of G86 (atoms O6 and N7; Figure 4C). Furthermore, N1 of G86 is contacted by the side chain
277 carboxyl of D374 (β 1- β 2 loop), its base engages in cation- π stacking to the guanidium group of
278 R371 (strand β 2) and its 2'-hydroxyl group is hydrogen bonded to the backbone carbonyl of
279 M442 (strand β 5; Figure 4C). O4 of U87 is hydrogen bonded to the side chain amide of Q457
280 (helix α 3) and its base stacks on F441 (strand β 5; Figure 4D). The O2 atom of U88 hydrogen
281 bonds with the backbone NH of W440 (strand β 5), positioning the base laterally on the F441
282 side chain. The U87-U88 backbone region encircles the side chain of K357 (helix α 1), which
283 contacts the anionic phosphate oxygens of U87 and the O3 atom of U88 (Figure 4D). The
284 interactions involving K357, W440, F441 and Q457 stabilize the second kink in the yU6 3'-
285 overhang. The side chain amino group of K361 (helix α 1) contacts the U89 anionic phosphate
286 oxygens in one complex (Figure 4E) and interacts with O2, O2' and O3' of U88 as well as the
287 phosphate and O5' of U89 in the other complex (Figure 4F). The last nt, U90, adopts two
288 different conformations in the two complexes; in one case, its N3, O2 and O4 atoms are
289 hydrogen bonded to the side chain hydroxyl of S364 (α 3- β 2 loop; Figure 4E), while in the other

case it interacts with the N1 and N6 positions of nt A82 at the beginning of the yU6 3'-overhang (Figure 4F).

The overall structure of the DUF1115 domain is globally unchanged compared to the structure of isolated yPrp3^{CTF} or yPrp3^{DUF1115} (RMSD of 0.80-0.86 Å for 133 common Cα atoms) but there are local adjustments accompanying RNA binding, which may contribute to the binding specificity. Upon RNA binding, the β3a/β3b hairpin moves closer to the yPrp3^{CTF} core domain to engage in direct interactions with the RNA and the H409 side chain in this element turns on top of the C84 base (Figure 4G). Within the yPrp3^{CTF} core domain, the side chain of F441 is rotated ca. 90 °C out of its intra-molecular position in the isolated protein to stack on U87. Instead, the U88 base moves into the original F441 position, from where it contacts the W440 backbone (Figure 4H). Furthermore, the K357 side chain adopts an alternative conformation upon RNA binding to interact with the anionic phosphate oxygens of U87 and the O3' atom of U88 (Figure 4H).

Mutational probing of the protein-RNA interface

Although not directly contacted by the protein, the bases of G81 and A82 mediate a continuous π-stack from the terminal base pair of stem II to the yPrp3-bound portion of the yU6 3'-overhang. Consistent with this stacking being important for yPrp3 binding, replacement of these purines in yU4/U6^{stem II+13nt} with a smaller pyrimidine (C) led to reduced affinity to yPrp3^{CTF} (Figure 4I, lanes 2 & 3). Sequence-specific interactions are seen between R353 and the Watson-Crick edge of C85 as well as the Hoogsteen edge of G86 (Figure 4C), suggesting that the protein reads out parts of the RNA sequence. Consistently, mutation of C85 to A or G86 to C in yU4/U6^{stem II+13nt} significantly reduced or essentially abrogated the interaction with yPrp3^{CTF}, respectively (Figure 4I, lanes 4 & 5). Correspondingly, a yPrp3^{CTF} variant bearing a R353A exchange showed essentially no binding to yU4/U6^{stem II+10nt} anymore (Figure 4J, lane 2; this and

all other yPrp3^{CTF} variants discussed below could be produced and purified like the wild type [wt] protein, suggesting that none of the tested mutations interfered with proper protein folding).

We also exchanged several additional yPrp3 residues that directly contact yU4/U6^{stem II+10nt} in our structure and tested binding of the corresponding yPrp3^{CTF} variants to this RNA (Figure 4J). R399A (lane 12) and F441H (lane 14) essentially abrogated the interaction with the RNA, while K357A (lane 6) strongly and F354A/H (lanes 3 & 4), K361A (lane 9) and S364R (lane 11) weakly destabilized the complex (indicated primarily by larger fractions of unbound yU4/U6^{stem II+10nt} in the corresponding lanes). Single residue exchanges K355A (lane 5), M358A/E (lanes 7 & 8), E362A (lane 10) and H409E (lane 13) of yPrp3^{CTF} showed no significant or only very mild reduction in RNA binding under the chosen conditions, indicating that individually the interactions involving these residues are not essential for stable RNA complex formation.

Nts G81, A82, C85 and G86, which upon mutation led to reduced yPrp3^{CTF} binding, are highly evolutionarily conserved in U6 snRNAs (Figure 1 – figure supplement 2). Likewise, R353, R399 and F441, where substitutions strongly affected RNA binding, are conserved in Prp3 orthologs (Figure 1 – figure supplement 1). We conclude that the mode of Prp3-U4/U6 interaction seen in our structure is present in all eukaryotes.

Evolutionary comparisons

The ferredoxin superfamily currently encompasses 59 subfamilies. Previous bioinformatics analyses indicated that within the superfamily the Prp3 DUF1115 domain is most homologous to acylphosphatase (AcyP) and BLUF domains (Korneta et al., 2012), which do not recognize nucleic acids. However, the sequence comparison did not reveal the evolutionary relationship between the Prp3 ssRNA-binding domain and other members of the ferredoxin superfamily. We therefore calculated a phylogenetic tree based on pairwise structural comparisons with representative ferredoxin fold structures (Figure 3 – figure supplement 2A). The tree is in agreement with sequence analyses, showing that the Prp3 ssRNA-binding domain is most

similar to AcyP and BLUF domains. For example, residues 335-467 of yPrp3^{CTF} (Figure 3 – figure supplement 2B) spatially align with cow AcyP (PDB ID 2ACY; Figure 3 – figure supplement 2C) with a RMSD of 2.5 Å over 90 common Cα atoms (Z-score 10.1) and show a comparable similarity to the BLUF domain of the BlrB photoreceptor from *Rhodobacter sphaeroides* (PDB ID 2BYC; RMSD of 2.0 Å for 84 common Cα atoms; Z-score of 10.0; Figure 3 – figure supplement 2D). On the other hand, the Prp3 domain is only distantly related to other ferredoxin fold proteins that bind RNAs, such as RRM (Figure 3 – figure supplement 2F,G) or ribosomal proteins S6, S10 and L10. The closest nucleic acid-binding relative of the Prp3 ssRNA-binding domain is the IS608 transposase domain, which recognizes DNA (Figure 3 – figure supplement 2H).

In vivo effects of Prp3 variants defective in U4/U6 di-snRNA binding

To test the importance of the observed yPrp3-yU4/U6 interactions for the function of yPrp3 in vivo, we used a haploid yeast strain, in which the chromosomal copy of the *PRP3* gene was deleted and the essential protein was produced from a counter-selectable *URA3*-marked plasmid. We then shuffled plasmids carrying mutant *prp3* genes into this strain, selected against the *URA3* plasmid and monitored the effects of yPrp3 variants on cell viability, snRNP levels and pre-mRNA splicing.

yPrp3 proteins bearing mutations R304A, R322A or both (in the basic peptide preceding DUF1115) as well as variants bearing R399A or F441A exchanges (yU6 3'-overhang-binding DUF1115 domain) led to mild temperature sensitive (ts) growth (Figure 5A). A R399A/F441A double mutant strongly exacerbated this effect (Figure 5A). These observations are consistent with the mutated residues contributing to functionally important yPrp3-yU4/U6 interactions, as seen in our binding assays and yPrp3^{CTF}-yU4/U6^{stem II+10nt} crystal structure. To trace the origin of the growth defects, we monitored the effects of the mutations on splicing in vivo at 37 °C, using endogenous U3 (pre-)snoRNA as a reporter. Slight accumulation of unspliced U3 pre-snoRNAs

was seen in strains producing the R304A, R399A, R304A/R322A or R399A/F441A variants of yPrp3 compared to the wt (Figure 5B), indicating that the observed growth defects likely root in inefficient splicing. Finally, monitoring snRNP levels by Northern blotting of cellular extracts spread out on a glycerol gradient showed that the levels of U4/U6 di-snRNP and of isolated U5 snRNP were increased at the expense of U4/U6•U5 tri-snRNP in strains producing Prp3 variants R304A/R322A and R399A/F441A (Figure 5C,D). At the same time, Western blotting revealed reduced levels of yPrp3 associated with U4/U6•U5 tri-snRNP in the mutant strains (Figure 5E). These results suggest that reduced tri-snRNP levels are the cause of the reduced splicing activity in the mutant strains and that efficient yPrp3 binding to U4/U6 di-snRNAs via its C-terminal region is important for U4/U6•U5 tri-snRNP stability.

Discussion

An evolutionarily conserved, composite ds/ssRNA-binding region in Prp3

Here, we showed that Prp3 orthologs contain a C-terminal U4/U6 di-snRNA binding region that encompasses a DUF1115 domain and a preceding peptide rich in basic residues. Our crystal structure and targeted mutational analyses demonstrate that the DUF1115 domain acts as a ssRNA-binding domain that specifically recognizes the first ten nts of the U6 3'-overhang. While our structural analysis did not reveal how the preceding peptide is involved in U4/U6 binding, several pieces of evidence indicate that it binds along U4/U6 stem II: The peptide is important for full binding of Prp3 C-terminal portions to U4/U6 (Figure 1C,D) but it is not a structural element of the DUF1115 domain and thus does not act via stabilizing the RNA-binding fold of DUF1115. Single Arg-to-Ala exchanges in the peptide lead to reduced RNA binding (Figure 1C,D), indicating that these residues may directly contact the RNA. RNA constructs with shortened stem II or in which a conserved non-canonical base pair was converted to a Watson-Crick pair show reduced binding to Prp3^{CTF} (Figure 1E,F), which cannot be explained by the

observed DUF1115 domain contacts (exclusively to the U6 3'-overhang). Finally, Prp3^{CTF}, but not the DUF1115 domain, inhibits Brr2-mediated U4/U6 unwinding (Figure 2A,B), suggesting that the peptide stabilizes contacts between U4 and U6, presumably by binding the stem II duplex.

Residues in both the DUF1115 domain and the preceding basic peptide of Prp3 as well as nt's in U4/U6, which we identified as being important for stable complex formation, are evolutionarily highly conserved, suggesting that the composite ds/ss U4/U6 di-snRNA-binding region of Prp3 is present in all eukaryotes. DUF1115 exhibits a ferredoxin-like core similar to AcyP/BLUF proteins but evolutionarily remote from other nucleic acid-binding domains in the ferredoxin superfamily. As we failed to detect a DUF1115-like domain by sequence comparisons in hundreds of other RNA-binding proteins recently identified in transcriptome-wide screens (Baltz et al., 2012; Castello et al., 2012), DUF1115 most likely represents a spliceosome-specific ssRNA-binding domain. Due to the pronounced re-orientation that the U6 3'-overhang experiences on the surface of the Prp3 DUF1115 domain, the domain likely is an important architectural element in the U4/U6 di-snRNP and/or the U4/U6•U5 tri-snRNP.

Prp3 binding helps explain defects associated with U6 snRNA variants

The mode of Prp3 binding to U4/U6 di-snRNAs revealed herein helps to rationalize a large body of mutational analyses on U6 snRNAs in yeast and human. Previously, the functional importance of various U6 snRNA regions was studied by site-directed mutagenesis. hU6 snRNA bearing a C62G mutation, which converts the C-U mismatch in U4/U6 stem II into a G-U wobble pair, supported only low levels of splicing (Wolff & Bindereif, 1993). Likewise, the corresponding C69G mutation in yU6 snRNA led to reduced splicing activity (Ryan & Abelson, 2002). We showed that a yU4/U6 stem II construct bearing the C69G mutation exhibits reduced binding to yPrp3 and according to our structural model this region of the U4/U6 di-snRNAs is recognized

by the peptide preceding the Prp3 DUF1115 domain. Therefore, a defect in Prp3 binding may underlie the splicing defects of these U6 snRNA mutations in yeast and human.

Residues of the highly conserved UGA motif located in the terminus of U6 stem II and beginning of the 3'-overhang were shown to be important for U4/U6•U5 tri-snRNP stability and splicing. Deletion of nts U74-G75-A76 in hU6 snRNA (corresponding to residues U80-G81-A82 in yU6), reduced the levels of U4/U6•U5 tri-snRNP and assembled spliceosomes to less than 10 % of wt and abrogated splicing activity (Wolff & Bindereif, 1992). An A76C variant of hU6 snRNA showed reduced interaction with U4 snRNA, concomitant with reduced spliceosome assembly and low splicing activity (Wolff & Bindereif, 1995). Furthermore, yU6 snRNA bearing a U80G exchange almost completely abolished formation of U4/U6 di-snRNPs and U4/U6•U5 tri-snRNPs and was defective in splicing (Fabrizio et al., 1989; Ryan & Abelson, 2002; Ryan et al., 2002). The G81C point mutation in yU6 snRNA strongly blocked splicing in vitro (Madhani et al., 1990; Ryan & Abelson, 2002) and caused a substantial accumulation of free U6 snRNP deficient in U4/U6 di-snRNP assembly (Ryan & Abelson, 2002). The effects of some of these mutations were suggested to root in a stabilization of intra-molecular base pairing in U6 snRNA. However, our observation that G81C and A82C exchanges in yU6 snRNA led to weak binding of yPrp3 in vitro suggests that aberrant Prp3 binding may also contribute to these phenotypes. The continuous stacking of nts in the transition region between stem II and the U6 3'-overhang seen in our structure may properly orient these elements for stable Prp3 binding.

3'-truncated yU6 snRNAs containing residues 1-94, 1-91 or 1-88 allowed reconstitution of 35-40 % of wt splicing activity, while only 20 % and 4 % of the wt splicing activity were regained with yU6 snRNA molecules further shortened to residues 1-86 or 1-85, respectively, and yU6 snRNAs containing only residues 1-81 or 1-80 fully blocked splicing (Madhani et al., 1990; Ryan et al., 2002). The phenotypes associated with the deletion of distal yU6 3'-overhang residues likely originate from the removal of binding sites for the Prp24 assembly chaperone and the LSm protein complex. However, the exacerbated effects when yU6 snRNA was shortened to 86

nts or less correlate very well with our finding of reduced yPrp3 binding to yU4/U6 constructs bearing sequentially shortened U6 3'-overhangs or the C85A and G86C point mutations. Reduced Prp3 binding may also contribute to the reduced hU4/U6 di-snRNP levels seen previously in the presence of hU6 snRNA bearing a C79U exchange in the 3'-overhang (Wolff & Bindereif, 1995), as we observed weak yPrp3 interaction with a yU4/U6 bearing a variant residue at the analogous position (C85A).

Mechanism of Prp3-dependent U4/U6•U5 tri-snRNP stability

The binding of the U4/U6-specific proteins Snu13, Prp31 and the Prp3-Prp4(-CypH) complex to U4/U6 di-snRNAs is highly interdependent. Snu13 binds a K-turn motif in the U4 5'SL (Nottrott et al., 1999; Vidovic et al., 2000), serving as an assembly initiating protein for the subsequent incorporation of the other components (Nottrott et al., 2002). An explanation for the ordered binding of Snu13 and Prp31 was provided by structural studies that showed how hSnu13 and the hU4 5'SL provide a composite binding platform for the hPrp31 Nop domain (Liu et al., 2007; Liu et al., 2011), while the mechanism by which hSnu13 leads to enhanced interaction of hPrp3 with U4/U6 di-snRNAs (Nottrott et al., 2002) remained enigmatic. Our results indicate that Prp3 contacts RNA elements along stem II, i.e. in the vicinity of the U4/U6 3-way junction. Snu13 binds one branch of the 3-way junction (the U4 5'SL), which could exert a stabilizing effect on the 3-way junction and stem II and lead to improved binding of Prp3. This effect could be further enhanced by Prp31, whose U4/U6 contacts extend into the lower part of stem II (Schultz et al., 2006). It is also conceivable that the Prp3 stem II-binding peptide directly contacts Snu13 and/or Prp31 on the RNAs and that the latter two proteins even modulate how the peptide interacts with stem II.

Our in vivo analyses of the effects of yPrp3 mutations that interfere with stable yU4/U6 binding in vitro, show that the yPrp3-yU4/U6 interactions we observe in our complex structure are important for U4/U6•U5 tri-snRNP stability and, likely as a consequence, for pre-mRNA

splicing. The role of Prp3 in mediating tri-snRNP stability can be understood from our present data, showing that the C-terminal portions of Prp3 maintain interactions with the U4/U6 di-snRNAs, and previous yeast 2-hybrid (Y2H) analyses, showing that Prp3 contacts the U5 snRNP proteins Prp6 and Snu66 (Liu et al., 2006). By combining these functions, Prp3 apparently acts as a bridge, linking the U4/U6 di-snRNP and the U5 snRNP via Prp3-RNA interactions on the U4/U6 side and Prp3-protein interactions on the U5 side (Figure 6A).

Possible role of Prp3 in U4/U6 re-assembly after splicing

After their release from the spliceosome, U4 and U6 snRNAs have to be re-decorated with proteins and re-assembled into a di-snRNP before they can participate in further rounds of splicing. This recycling requires the Prp24 protein in yeast and the related SART3 protein in human. Recently, a crystal structure of near full-length Prp24 in complex with a large portion of U6 snRNA was elucidated (Montemayor et al., 2014). While not contained in the Prp24-U6 structure, the distal 3'-end of U6 snRNA is available in that complex for binding the LSm proteins (Figure 6B), consistent with the observation of a stable Prp24-LSm-U6 snRNA complex in yeast (Vidal et al., 1999). In contrast, U6 regions that form stem I, stem II and the U6 3'-overhang in the assembled U4/U6 di-snRNP are sequestered in an internal U6 stem loop (ISL) and by Prp24-U6 interactions (Figure 6B). As Prp24 does not harbor an NTP-dependent RNA helicase activity, the question arises how the U6 ISL can be unwound and U6 handed over from Prp24 to U4 snRNP. Notably, regions of U6 snRNA that form stem II and the U6 3'-overhang in the U4/U6 di-snRNP (i.e. Prp3-binding elements), are largely exposed on the surface of the Prp24-U6 snRNA complex (Figure 6B). Furthermore, the Prp24-U6 structure represents an artificially stabilized situation, obtained after introducing a cold-sensitive A62G mutation in U6, which impedes unwinding of the U6 ISL (Fortner et al., 1994) by sequestering C85 in a non-natural Watson-Crick base pair (magenta in Figure 6B). We have shown here that C85 is a crucial contact residue of Prp3, which also latches onto the surrounding nts in the context of

U4/U6 di-snRNP via its DUF1115 domain. Thus in the wt situation, the U6 ISL on Prp24 can most likely “breathe” in the region surrounding C85, whereby this region could become available for Prp3 binding. Consistent with this idea, chemical probing studies showed that these Prp3-binding nts are only weakly protected in the Prp24-LSm-U6 snRNA particle (Karaduman et al., 2006). We therefore suggest that binding of Prp3 to its cognate U6 3'-overhang region traps and subsequently extends and stabilizes spontaneous local unwinding of the internal stem loop (ISL) on the Prp24-LSm-U6 snRNA complex (Figure 6C).

We envision that Prp24 displacement is completed by entry of a pre-assembled U4 snRNP (Figure 6C). Prp3 could also support this step by stabilizing the emerging U4/U6 stem II via its stem II-binding peptide. Furthermore, the cooperativity we detect in binding of Snu13, Prp31 and Prp3 to U4/U6 di-snRNA may help to complete assembly of the U4/U6 di-snRNP. Taken together, our results support an important role of Prp3 in the handover of U6 snRNA from Prp24 to U4 snRNP during U4/U6 di-snRNP re-assembly after splicing.

Materials and methods

Protein production and purification

DNA encoding full-length hPrp3 or the hPrp3^C fragment (residues 195-683) with C-terminal His₆-tags were PCR-amplified from pGADT7-*hPRP3* (Liu et al., 2006) and subcloned into vector pMAL-c2t, in which the Factor Xa cleavage site following the N-terminal MBP tag of pMAL-c2x (New England Biolabs) was replaced by a TEV cleavage site. The inserts of these and all other plasmids were verified by sequencing. A DNA fragment encoding hPrp3^N (residues 1-442) was cloned into pETM-41 (EMBL, Heidelberg) for production of the protein bearing an N-terminal His₆-MBP fusion. Plasmids were transformed into *Escherichia coli* BL21 (DE3)-RIL cells (Stratagene) and expressed at 16-18 °C using the auto-induction method (Studier, 2005). Target proteins were double affinity purified using Ni-NTA beads and amylose resin followed by

520 gel filtration using a Superdex 200 26/60 column (GE Healthcare) in 20 mM Tris-HCl, pH 7.5,
521 200 mM NaCl, 2 mM DTT. RNase A (Qiagen) was included in the initial purification steps for
522 digestion of host RNAs.

523 DNA fragments encoding hPrp3^{CTF} (residues 484-683) or yPrp31¹⁻⁴⁶² were introduced via
524 BamHI and XhoI restriction sites into pGEX-6P (GE Healthcare) for production as N-terminal,
525 PreScission-cleavable GST fusion proteins. GST-hPrp3^{CTF} and GST-yPrp31¹⁻⁴⁶² were produced
526 at 18 °C in *E. coli* BL21(DE3)-RIL cells using the auto-induction method and purified using
527 glutathione-Sepharose beads (GE Healthcare). The fusion proteins were eluted from beads in
528 buffer containing 10 mM reduced L-glutathione and further purified via a Superdex 75 16/60 gel
529 filtration column (GE Healthcare) in 20 mM Tris-HCl, pH 7.5, 200 mM NaCl, 2 mM DTT.

530 DNA constructs encoding yPrp3^{CTF} (residues 296-469), yPrp3^{DUF1115} (residues 325-469) or
531 ySnu13 were amplified from *PRP3* or *SNU13* synthetic genes (GENEART AG), respectively,
532 and subcloned into pETM-11 (EMBL, Heidelberg). Mutant versions of pETM-11-*PRP3*^{CTF} were
533 generated by site-directed mutagenesis using the QuikChange II kit (Agilent Technologies,
534 Germany). The plasmids were transformed into *E. coli* BL21 (DE3)-RIL cells and expressed at
535 18-25 °C for two days using the auto-induction method. N-terminally His₆-tagged proteins were
536 affinity purified using 1 ml or 5 ml HisTrap FF columns (GE Healthcare). 1 M LiCl buffer was
537 included in the wash step to remove unspecifically bound nucleic acids. His₆-tagged proteins
538 used in some EMSA and unwinding assays were directly applied to a Superdex 75 16/60
539 column in 10 mM Tris-HCl, pH 7.0, 150 mM NaCl, 2 mM DTT. For EMSA titrations, ITC and
540 crystallization, TEV protease was used to remove the N-terminal His₆-tag before the size-
541 exclusion chromatography step. Production of selenomethionine (SeMet)-substituted His₆-
542 yPrp3^{CTF} in BL21 (DE3)-RIL cells was carried out in M9 minimal medium supplemented with
543 trace metals and vitamins (van den Ent & Lowe, 2000). At an OD₅₉₅ of 0.6, 50 mg/l of SeMet
544 (Fisher Scientific), 100 mg/l of lysine, threonine and phenylalanine and 50 mg/l of leucine,
545 isoleucine and valine were added. After 15 min, the temperature was reduced to 20 °C and 0.32

mM IPTG was added for induction overnight. SeMet-substituted protein was purified in the same way as the native protein.

Full-length yBrr2 bearing an N-terminal His₆-tag was produced using a recombinant baculovirus in insect cells as described for hBrr2 (Santos et al., 2012). Briefly, a 800 ml infected High FiveTM cell pellet was resuspended in 50 mM HEPES, pH 8.0, 600 mM NaCl, 2 mM β -mercaptoethanol, 0.05 % NP-40, 20 % glycerol, 10 mM imidazole, supplemented with protease inhibitors (Roche) and lysed by sonication using a Sonopuls Ultrasonic Homogenizer HD 3100 (Bandelin). His₆-yBrr2 was captured from the cleared lysate on a 5 ml HisTrap FF column (GE Healthcare) and eluted with a linear gradient from 10 to 250 mM imidazole. The eluted protein was diluted to a final concentration of 80 mM sodium chloride and loaded on a Mono Q 10/100 GL column (GE Healthcare) equilibrated with 50 mM Tris-HCl, pH 8.0, 50 mM NaCl, 5 % glycerol, 2 mM β -mercaptoethanol. His₆-yBrr2 was eluted with a linear 50 to 600 mM sodium chloride gradient and further purified by gel filtration on a Superdex 200 26/60 column (GE Healthcare) in 40 mM Tris-HCl, pH 8.0, 200 mM NaCl, 20 % glycerol, 2 mM DTT.

RNA production and purification

Full-length yU4 and yU6 snRNAs were synthesized by in vitro transcription using T7 RNA polymerase and PCR-generated templates. The transcripts were purified using the RNeasy Midi Kit (Qiagen). RNA duplexes were prepared by combining 30 nM of [³²P]-5'-end labeled U4 snRNA with a 5-fold molar excess of unlabeled U6 snRNA in annealing buffer (40 mM Tris-HCl, pH 7.5, 100 mM NaCl). The solution was heated to 80 °C for 2 min and cooled to 25 °C over 90 min. 12.5 mM MgCl₂ were added to the solution at 70 °C. The annealed duplex was separated from ss U4 and U6 snRNAs by 6 % native PAGE. The duplex was eluted from the gel, phenol-chloroform extracted, ethanol precipitated and resuspended in annealing buffer. yU4/U6^{fused} was prepared by in vitro transcription and purified via a Mono Q column, followed by gel electrophoresis on an 8 % denaturing (7 M urea) polyacrylamide gel, eluted and precipitated by

isopropanol. All other RNAs were chemically synthesized (Dharmacon). Complementary oligonucleotides were annealed before use by resuspension in H₂O, mixing, heating to 95 °C for 2 min, slow cooling to room temperature followed by incubation on ice.

Electrophoretic mobility shift assays

Typically, [³²P]-5'-end labeled RNAs were mixed with recombinant proteins in 20 mM Tris-HCl, pH 7.0, 150 mM NaCl, 2 mM DTT, 0.5 µg/µl tRNA. For EMSAs involving hPrp3^{FL}, hPrp3^N and hPrp3^C, samples were incubated in 10 mM Tris-HCl, pH 7.5, 200 mM NaCl, 2 mM DTT, 0.33 µg/µl tRNA and 0.2 µg/µl heparin. For hPrp3^{CTF} the same buffer with 67 ng/µl heparin was used. RNP complexes were allowed to form for 30 min at 4 °C and then separated on a 5-6 % (60:1) polyacrylamide gel.

For EMSA titrations of yPrp3^{CTF} variants and yPrp3^{DUF1115}, proteins lacking the N-terminal His₆-tag were employed. EMSA experiments were performed as described above (pre-incubation in 20 mM HEPES, pH 6.8, 150 mM NaCl, 0.33 µg/µl tRNA) with the indicated concentrations of proteins. Radioactive bands were visualized by autoradiography and quantified with the Image Quant 5.2 software (GE Healthcare). Apparent K_d values were obtained by fitting the resulting data points to a single exponential Hill equation (fraction bound = $A[\text{protein}]^n / ([\text{protein}]^n + K_d^n)$; A, fit maximum of RNA bound; n, Hill coefficient) (Ryder et al., 2008) using GraphPad Prism (Graphpad Software, Inc.).

For monitoring binding cooperativity of yPrp3^{CTF} and yPrp3^{CTF} with Snu13 and Prp31¹⁻⁴⁶², [³²P]-5'-end labeled RNA oligonucleotides were mixed with recombinant proteins in 20 mM HEPES-NaOH, pH 7.5, 200 mM NaCl, 3 mM DTT, 0.33 µg/µl tRNA, 67 ng/µl acetylated BSA, 13 ng/µl heparin at 4 °C for 30 min and separated on a 4 % (30:1) polyacrylamide gel. Radioactive bands were visualized by autoradiography using a phosphorimager (Molecular Dynamics).

Isothermal titration calorimetry

Proteins without affinity tags were used for ITC experiments. Proteins and RNA were buffer exchanged to 20 mM HEPES, pH 6.8, 150 mM NaCl by dialysis and their concentrations were determined via the absorbances at 280 and 260 nm, respectively. yPrp3^{CTF} variants (50 μ M) or yPrp3^{DUF1115} (200 μ M) were used as titrants, yU4/U6^{stem II+13nt} (fused by a GAGA tetraloop; 300 μ l at 5 μ M or 25 μ M) as the analyte. Measurements were conducted at 20 °C on a MicroCalTM iTC200 (GE Healthcare), with 16 injections of 2.5 μ l each with 180 s intervals between injections. Titrant heats of dilution were subtracted and data were fit using MicroCal Origin 7 (GE Healthcare).

U4/U6 di-snRNA unwinding assays

1.75 nM yU4/U6 di-snRNAs without or with 20 μ M of yPrp3 variants were pre-incubated with RNA at 30 °C for 3 min in 40 mM Tris-HCl, pH 7.5, 50 mM NaCl, 8 % glycerol, 0.5 mM MgCl₂, 100 ng/ μ l acetylated BSA, 1 U/ μ l RNasin, 1.5 mM DTT before the addition of 50 nM yBrr2. The reactions were initiated by adding 1 mM ATP/MgCl₂ and further incubated at 30 °C for 0-90 min. Aliquots were taken at the indicated time points and reactions were stopped with one volume of 40 mM Tris-HCl, pH 7.4, 50 mM NaCl, 25 mM EDTA, 1 % SDS, 10 % glycerol, 0.05 % xylene cyanol, 0.05 % bromophenol blue. Samples were loaded on a 6 % native PAGE and run at 10 W for 2 hours. RNA bands were visualized by autoradiography and quantified with the Image Quant 5.2 software. Data were fit to a single exponential equation (fraction unwound = $A\{1-\exp(-k_u t)\}$; A, amplitude of the reaction; k_u , apparent first-order rate constant for unwinding; t, time) using GraphPad Prism (GraphPad Software, Inc.).

Plasmid shuffling, cell viability and in vivo splicing assays

The wt yPRP3 gene was PCR-amplified from plasmid EWB2235, cloned into vector pRS314 and used to introduce the desired mutations by the QuikChange site-directed mutagenesis

strategy (Stratagene). Wt and mutant plasmids were transformed into yeast strain EWY2845 (*prp3::LEU2; his3Δ200; leu2-3,112; lys2-810; trp1-1; ura3-52 [PRP3/YCp50]*; kindly provided by John L. Woolford, Jr., Carnegie Mellon University, Pittsburgh, USA) that harbors the *PRP3* gene on a counter-selectable *URA3* plasmid. Transformants were selected in a medium lacking tryptophan, followed by three times streaking on 5-FOA plates at 25 °C to counter-select the *URA3* plasmid. Growth phenotypes of the yeast cells were assessed by spotting about 5×10^5 cells and tenfold serial dilutions on YPD agar plates and incubating at 16 °C, 30 °C or 37 °C for 1-2 days. To analyze the effect of *prp3* mutations on splicing in vivo, the yeast cells producing wt or variant forms of yPrp3 were used to inoculate YPD medium at an OD₆₀₀ of 0.05 and the cultures were further incubated at 37 °C for 5 h. For monitoring the levels of unspliced and spliced U3 (pre-)snRNAs, eight µg of total RNA from each sample were used for primer extension as described (Mozaffari-Jovin et al., 2013).

Analysis of snRNP levels

Yeast cells expressing wt or ts variants of yPrp3 were grown in YPD medium at 30 °C. Whole cell extract prepared from each strain was incubated at 37 °C (the non-permissive temperature for the Prp3 ts variants) for 30 min, diluted with equal volume of G100 buffer (20 mM HEPES-KOH, pH 7.0, 100 mM KCl, 0.2 mM EDTA) and fractionated by ultracentrifugation on a 12 ml 10-30 % (v/v) glycerol gradient in G100 buffer. Subsequent to ultracentrifugation at 37,000 rpm for 15 h in a Sorvall TST41.14 rotor, the distribution of spliceosomal U4, U5 and U6 snRNPs across the gradient fractions was monitored by Northern blotting and quantified as described previously (Mozaffari-Jovin et al., 2013). The association of yPrp3 with tri-snRNP was analyzed by Western blotting of gradient fractions and immuno-staining using antibodies against yPrp3 and ySnu114 and the Amersham ECL detection kit (GE Healthcare).

Crystallographic procedures

Crystallization was performed using the sitting-drop vapor-diffusion method at 18°C. Protein yPrp3^{CTF} was concentrated to 15 mg/ml in 10 mM Tris-HCl, pH 7.5, 200 mM NaCl, 1 mM DTT. The best crystals were obtained by mixing 1 µl of protein solution with 0.2 µl of 0.1 M yttrium (III) chloride hexahydrate or praseodymium (III) acetate hydrate and 0.8 µl of reservoir solution (0.1 M succinic acid, pH 7.0, 16.4 % PEG 3350). Protein yPrp3³²⁵⁻⁴⁶⁹ was concentrated to 15 mg/ml in 10 mM HEPES, pH 6.8, 100 mM NaCl, 2 mM DTT. The best crystals were obtained using a 1:1 mixture of protein solution and reservoir buffer consisting of 0.1 M HEPES, pH 7.5, 10 % PEG 8000. Prior to data collection, both types of crystals were cryoprotected in reservoir solution supplemented with 27 % (v/v) ethylene glycol and flash-cooled in liquid nitrogen.

yU4/U6^{stem II+10nt} was slowly added to the SeMet-substituted yPrp3^{CTF} in a 1:1 molar ratio and incubated at 4 °C for 15 min. The mixture was then chromatographed on a Superdex 75 16/60 column in 10 mM Na HEPES, pH 6.8, 1 mM DTT. Peak fractions containing the yPrp3^{CTF}-yU4/U6^{stem II+10nt} complex were pooled and concentrated to 7.5 mg/ml. Crystals were grown by mixing 1 µl of complex solution with 0.2 µl of 0.5 M sodium fluoride and 0.8 µl of reservoir solution (0.22 M DL-malic acid, pH 6.8, 16.5 % PEG 3350). For diffraction data collection, crystals were transferred to a 1:1 mixture of Paratone-N and paraffin oil as a cryoprotectant and flash-cooled in liquid nitrogen.

All X-ray diffraction data were collected at 100 K on beamline BL14.2 at the BESSY II storage ring (Berlin, Germany). Data were processed with the HKL package (Minor et al., 2006) or XDS (Kabsch, 2010). The structure of yPrp3^{CTF} was solved by a praseodymium (III) multiple anomalous dispersion experiment as described (Puehringer et al., 2012) and refined against the higher resolution data from an yttrium (III)-derivatized crystal. The structures of yPrp3^{DUF1115} and of a yPrp3^{CTF}-yU4/U6^{stem II+10nt} complex were solved by molecular replacement using structure coordinates of yPrp3^{CTF} as search models with the programs MOLREP (Vagin & Teplyakov, 2010) and PHASER (McCoy, 2007), respectively. Models were manually completed and

adjusted with COOT (Emsley et al., 2010) and the structures were refined using REFMAC5 (Murshudov et al., 2011) and phenix.refine (Afonine et al., 2012). To preserve A-form geometry in the stem II portion of the yPrp3^{CTF}-yU4/U6^{stem II+10nt} structure, an A-form duplex was used as a reference model and the A-form geometry was restrained during the refinement. Electrostatic surface potentials were calculated with APBS (Unni et al., 2011). All structure figures were created using PyMOL (<http://www.pymol.org/>).

Phylogenetic analyses

For phylogenetic analyses, a representative set of protein structures containing ferredoxin-like folds was assembled using the SCOP database (Murzin et al., 1995). Pairwise structural similarity Z-scores were calculated for all structures and the yPrp3 ssRNA-binding domain using DaliLite (Holm & Park, 2000). The Z-score reciprocals were used to build a UPGMA tree with the neighbor tool of the PHYLIP package (Felsenstein, 1989). The tree was visualized with Evolview (Zhang et al., 2012) and annotated manually.

Data deposition

Coordinates and diffraction data have been deposited with the Protein Data Bank (www.pdb.org) under accession codes 4YHU (yPrp3^{CTF}), 4YHV (yPrp3^{DUF1115}) and 4YHW (yPrp3^{CTF}-yU4/U6^{stem II+10nt}) and will be released upon publication.

Acknowledgements

We thank Nicole Holton, Freie Universität Berlin, for help with ITC measurements and Grzegorz Chojnowski, International Institute of Molecular and Cell Biology, for help with RNA structure comparisons. We thank Jelena Jakovljevic and John L. Woolford, Jr., Carnegie Mellon University, Pittsburgh, USA, for yeast strain EWY2845 and plasmid EWB2235. We

acknowledge access to beamline BL14.2 of the BESSY II storage ring (Berlin, Germany) via the Joint Berlin MX-Laboratory sponsored by the Helmholtz Zentrum Berlin für Materialien und Energie, the Freie Universität Berlin, the Humboldt-Universität zu Berlin, the Max-Delbrück Centrum, and the Leibniz-Institut für Molekulare Pharmakologie. MT was supported by a fellowship from the Boehringer Ingelheim Fonds. This work was funded by the Deutsche Forschungsgemeinschaft (SFB 740 to MCW; SFB 860 to RL), the Bundesministerium für Bildung und Forschung (grant 05K10KEC to MCW), the European Research Council (grant RNA+P=123D to JMB) and the Foundation for Polish Science (to JMB).

Author contributions

Author contributions: SL – protein and RNA production, EMSA, crystal structure analyses; SMJ, PF – generation of yeast strains, yeast viability and splicing assays, analysis of snRNP levels; JW, MT, KS – protein and RNA production, EMSA, ITC, unwinding, binding cooperativity studies; SDH, JMB – bioinformatics analyses; SL, RL, MCW – writing of the manuscript. All authors participated in the analysis and interpretation of the data.

Conflict of interest

The authors declare that they have no conflict of interest.

Funding

Deutsche Forschungsgemeinschaft Collaborative Research Center 740

Deutsche Forschungsgemeinschaft Collaborative Research Center 860

Bundesministerium für Bildung und Forschung grant 05K10KEC

European Research Council grant RNA+P=123D

Foundation for Polish Science

724 Boehringer Ingelheim Fonds

725

References

- Achsel T, Brahms H, Kastner B, Bachi A, Wilm M, Lührmann R. 1999. A doughnut-shaped heteromer of human Sm-like proteins binds to the 3'-end of U6 snRNA, thereby facilitating U4/U6 duplex formation in vitro. *EMBO J* **18**:5789-802. doi:10.1093/emboj/18.20.5789
- Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH, Adams PD. 2012. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D* **68**:352-67. doi:10.1107/S0907444912001308
- Anthony JG, Weidenhammer EM, Woolford JL, Jr. 1997. The yeast Prp3 protein is a U4/U6 snRNP protein necessary for integrity of the U4/U6 snRNP and the U4/U6.U5 tri-snRNP. *RNA* **3**:1143-52. doi:
- Ayadi L, Callebaut I, Saguez C, Villa T, Mornon JP, Banroques J. 1998. Functional and structural characterization of the prp3 binding domain of the yeast prp4 splicing factor. *J Mol Biol* **284**:673-87. doi:10.1006/jmbi.1998.2183
- Baltz AG, Munschauer M, Schwanhauser B, Vasile A, Murakawa Y, Schueler M, Youngs N, Penfold-Brown D, Drew K, Milek M, Wyler E, Bonneau R, Selbach M, Dieterich C, Landthaler M. 2012. The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. *Mol Cell* **46**:674-90. doi:10.1016/j.molcel.2012.05.021
- Barton GJ. 1993. ALSCRIPT - A tool for multiple sequence alignments. *Protein Eng* **6**:37-40. doi:10.1093/protein/6.1.37
- Beggs JD. 2005. Lsm proteins and RNA processing. *Biochem Soc Trans* **33**:433-8. doi:10.1042/BST0330433
- Bell M, Schreiner S, Damianov A, Reddy R, Bindereif A. 2002. p110, a novel human U6 snRNP protein and U4/U6 snRNP recycling factor. *EMBO J* **21**:2724-35. doi:10.1093/emboj/21.11.2724

751 Bessonov S, Anokhina M, Krasauskas A, Golas MM, Sander B, Will CL, Urlaub H, Stark H,
 752 Lührmann R. 2010. Characterization of purified human B act spliceosomal complexes
 753 reveals compositional and morphological changes during spliceosome activation and first
 754 step catalysis. *RNA* **16**:2384-403. doi:10.1261/Rna.2456210
 755 Bessonov S, Anokhina M, Will CL, Urlaub H, Lührmann R. 2008. Isolation of an active step I
 756 spliceosome and composition of its RNP core. *Nature* **452**:846-50.
 757 doi:10.1038/nature06842
 758 Bindereif A, Green MR. 1987. An ordered pathway of snRNP binding during mammalian pre-
 759 mRNA splicing complex assembly. *EMBO J* **6**:2415-24. doi:
 760 Bochkarev A, Bochkareva E, Frappier L, Edwards AM. 1998. The 2.2 angstrom structure of a
 761 permanganate-sensitive DNA site bound by the Epstein-Barr virus origin binding protein,
 762 EBNA1. *J Mol Biol* **284**:1273-8. doi:10.1006/jmbi.1998.2247
 763 Castello A, Fischer B, Eichelbaum K, Horos R, Beckmann BM, Strein C, Davey NE, Humphreys
 764 DT, Preiss T, Steinmetz LM, Krijgsveld J, Hentze MW. 2012. Insights into RNA biology from
 765 an atlas of mammalian mRNA-binding proteins. *Cell* **149**:1393-406.
 766 doi:10.1016/j.cell.2012.04.031
 767 Cheng SC, Abelson J. 1987. Spliceosome assembly in yeast. *Genes Dev* **1**:1014-27.
 768 doi:10.1101/gad.1.9.1014
 769 Chojnowski G, Walen T, Bujnicki JM. 2014. RNA Bricks-a database of RNA 3D motifs and their
 770 interactions. *Nucleic Acids Res* **42**:D123-D31. doi:10.1093/Nar/Gkt1084
 771 Deckert J, Hartmuth K, Boehringer D, Behzadnia N, Will CL, Kastner B, Stark H, Urlaub H,
 772 Lührmann R. 2006. Protein composition and electron microscopy structure of affinity-
 773 purified human spliceosomal B complexes isolated under physiological conditions. *Mol Cell*
 774 *Biol* **26**:5528-43. doi:10.1128/MCB.00582-06
 775 Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high
 776 throughput. *Nucleic Acids Res* **32**:1792-7. doi:10.1093/Nar/Gkh340

777 Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. *Acta*
778 *Crystallogr D* **66**:486-501. doi:10.1107/S0907444910007493

779 Fabrizio P, Dannenberg J, Dube P, Kastner B, Stark H, Urlaub H, Lührmann R. 2009. The
780 evolutionarily conserved core design of the catalytic activation step of the yeast
781 spliceosome. *Mol Cell* **36**:593-608. doi:10.1016/j.molcel.2009.09.040

782 Fabrizio P, McPheeters DS, Abelson J. 1989. In vitro assembly of yeast U6 snRNP: a functional
783 assay. *Genes Dev* **3**:2137-50. doi:10.1101/gad.3.12b.2137

784 Felsenstein J. 1989. PHYLIP - Phylogeny Inference Package. *Cladistics* **5**:164-6. doi:
785 Fortner DM, Troy RG, Brow DA. 1994. A stem/loop in U6 RNA defines a conformational switch
786 required for premessenger RNA splicing. *Genes Dev* **8**:221-33. doi:10.1101/Gad.8.2.221

787 Gonzalez-Santos JM, Wang A, Jones J, Ushida C, Liu J, Hu J. 2002. Central region of the
788 human splicing factor Hprp3p interacts with Hprp4p. *J Biol Chem* **277**:23764-72.
789 doi:10.1074/jbc.M111461200

790 Holm L, Park J. 2000. DaliLite workbench for protein structure comparison. *Bioinformatics*
791 **16**:566-7. doi:10.1093/bioinformatics/16.6.566

792 Jung A, Domratcheva T, Tarutina M, Wu Q, Ko WH, Shoeman RL, Gomelsky M, Gardner KH,
793 Schlichting L. 2005. Structure of a bacterial BLUF photoreceptor: Insights into blue light-
794 mediated signal transduction. *Proc Natl Acad Sci USA* **102**:12350-5.
795 doi:10.1073/pnas.0500722102

796 Kabsch W. 2010. XDS. *Acta Crystallogr D* **66**:125-32. doi:10.1107/S0907444909047337

797 Karaduman R, Fabrizio P, Hartmuth K, Urlaub H, Lührmann R. 2006. RNA structure and RNA-
798 protein interactions in purified yeast U6 snRNPs. *J Mol Biol* **356**:1248-62.
799 doi:10.1016/j.jmb.2005.12.013

800 Konarska MM, Sharp PA. 1987. Interactions between small nuclear ribonucleoprotein particles
801 in formation of spliceosomes. *Cell* **49**:763-74. doi:10.1016/0092-8674(87)90614-3

802 Konarska MM, Vilardeell J, Query CC. 2006. Repositioning of the reaction intermediate within the
 803 catalytic center of the spliceosome. *Mol Cell* **21**:543-53. doi:10.1016/j.molcel.2006.01.017
 804 Konforti BB, Koziolkiewicz MJ, Konarska MM. 1993. Disruption of base pairing between the 5'
 805 splice site and the 5' end of U1 snRNA is required for spliceosome assembly. *Cell* **75**:863-
 806 73. doi:10.1016/0092-8674(93)90531-T
 807 Korneta I, Magnus M, Bujnicki JM. 2012. Structural bioinformatics of the human spliceosomal
 808 proteome. *Nucleic Acids Res* **40**:7046-65. doi:10.1093/nar/gks347
 809 Kramer A, Keller W, Appel B, Lührmann R. 1984. The 5' terminus of the RNA moiety of U1 small
 810 nuclear ribonucleoprotein particles is required for the splicing of messenger RNA
 811 precursors. *Cell* **38**:299-307. doi:10.1016/0092-8674(84)90551-8
 812 Laskowski RA, Hutchinson EG, Michie AD, Wallace AC, Jones ML, Thornton JM. 1997.
 813 PDBsum: a Web-based database of summaries and analyses of all PDB structures. *Trends*
 814 *Biochem Sci* **22**:488-90. doi:10.1016/S0968-0004(97)01140-7
 815 Liu S, Ghalei H, Lührmann R, Wahl MC. 2011. Structural basis for the dual U4 and U4atac
 816 snRNA-binding specificity of spliceosomal protein hPrp31. *RNA* **17**:1655-63.
 817 doi:10.1261/rna.2690611
 818 Liu S, Li P, Dybkov O, Nottrott S, Hartmuth K, Lührmann R, Carlomagno T, Wahl MC. 2007.
 819 Binding of the human Prp31 Nop domain to a composite RNA-protein platform in U4
 820 snRNP. *Science* **316**:115-20. doi:10.1126/science.1137924
 821 Liu S, Rauhut R, Vornlocher HP, Lührmann R. 2006. The network of protein-protein interactions
 822 within the human U4/U6.U5 tri-snRNP. *RNA* **12**:1418-30. doi:10.1261/rna.55406
 823 Madhani HD, Bordonne R, Guthrie C. 1990. Multiple roles for U6 snRNA in the splicing
 824 pathway. *Genes Dev* **4**:2264-77. doi:10.1101/gad.4.12b.2264
 825 McCoy AJ. 2007. Solving structures of protein complexes by molecular replacement with
 826 Phaser. *Acta Crystallogr D* **63**:32-41. doi:10.1107/S0907444906045975

827 Minor W, Cymborowski M, Otwinowski Z, Chruszcz M. 2006. HKL-3000: the integration of data
 828 reduction and structure solution--from diffraction images to an initial model in minutes. *Acta*
 829 *Crystallogr D* **62**:859-66. doi:10.1107/S0907444906019949
 830 Montemayor EJ, Curran EC, Liao HH, Andrews KL, Treba CN, Butcher SE, Brow DA. 2014.
 831 Core structure of the U6 small nuclear ribonucleoprotein at 1.7-angstrom resolution. *Nat*
 832 *Struct Mol Biol* **21**:544-51. doi:10.1038/Nsmb.2832
 833 Mount SM, Pettersson I, Hinterberger M, Karmas A, Steitz JA. 1983. The U1 small nuclear
 834 RNA-protein complex selectively binds a 5' splice site in vitro. *Cell* **33**:509-18.
 835 doi:10.1016/0092-8674(83)90432-4
 836 Mozaffari-Jovin S, Wandersleben T, Santos KF, Will CL, Lührmann R, Wahl MC. 2013.
 837 Inhibition of RNA helicase Brr2 by the C-terminal tail of the spliceosomal protein Prp8.
 838 *Science* **341**:80-4. doi:10.1126/science.1237515
 839 Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F,
 840 Vagin AA. 2011. Refmac5 for the refinement of macromolecular crystal structures. *Acta*
 841 *Crystallogr D* **67**:355-67. doi:10.1107/S0907444911001314
 842 Murzin AG, Brenner SE, Hubbard T, Chothia C. 1995. SCOP - a Structural Classification of
 843 Proteins Database for the Investigation of Sequences and Structures. *J Mol Biol* **247**:536-
 844 40. doi:10.1016/S0022-2836(05)80134-2
 845 Nottrott S, Hartmuth K, Fabrizio P, Urlaub H, Vidovic I, Ficner R, Lührmann R. 1999. Functional
 846 interaction of a novel 15.5kD [U4/U6.U5] tri-snRNP protein with the 5' stem-loop of U4
 847 snRNA. *EMBO J* **18**:6119-33. doi:10.1093/emboj/18.21.6119
 848 Nottrott S, Urlaub H, Lührmann R. 2002. Hierarchical, clustered protein interactions with U4/U6
 849 snRNA: a biochemical role for U4/U6 proteins. *EMBO J* **21**:5527-38.
 850 doi:10.1093/emboj/cdf544

851 Oubridge C, Ito N, Evans PR, Teo CH, Nagai K. 1994. Crystal structure at 1.92 Å resolution of
 852 the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin.
 853 *Nature* **372**:432-8. doi:10.1038/372432a0

854 Parker R, Siliciano PG, Guthrie C. 1987. Recognition of the TACTAAC box during mRNA
 855 splicing in yeast involves base pairing to the U2-like snRNA. *Cell* **49**:229-39.
 856 doi:10.1016/0092-8674(87)90564-2

857 Puehringer S, Hellmig M, Liu S, Weiss MS, Wahl MC, Mueller U. 2012. Structure determination
 858 by multiple-wavelength anomalous dispersion (MAD) at the Pr LIII edge. *Acta Crystallogr F*
 859 **68**:981-4. doi:10.1107/S1744309112025456

860 Rader SD, Guthrie C. 2002. A conserved Lsm-interaction motif in Prp24 required for efficient
 861 U4/U6 di-snRNP formation. *RNA* **8**:1378-92. doi:10.1017/S1355838202020010

862 Raghunathan PL, Guthrie C. 1998. A spliceosomal recycling factor that reanneals U4 and U6
 863 small nuclear ribonucleoprotein particles. *Science* **279**:857-60.
 864 doi:10.1126/science.279.5352.857

865 Ryan DE, Abelson J. 2002. The conserved central domain of yeast U6 snRNA: Importance of
 866 U2-U6 helix I-a in spliceosome assembly. *RNA* **8**:997-1010.
 867 doi:10.1017/S1355838202025013

868 Ryan DE, Stevens SW, Abelson J. 2002. The 5' and 3' domains of yeast U6 snRNA: Lsm
 869 proteins facilitate binding of Prp24 protein to the U6 telestem region. *RNA* **8**:1011-33.
 870 doi:10.1017/S1355838202026092

871 Ryder SP, Recht MI, Williamson JR. 2008. Quantitative analysis of protein-RNA interactions by
 872 gel mobility shift. *Methods Mol Biol* **488**:99–115. doi:10.1007/978-1-60327-475-3_7

873 Santos KF, Mozaffari-Jovin S, Weber G, Pena V, Lührmann R, Wahl MC. 2012. Structural basis
 874 for functional cooperation between tandem helicase cassettes in Brr2-mediated remodeling
 875 of the spliceosome. *Proc Natl Acad Sci USA* **109**:17418-23. doi:10.1073/pnas.1208098109

876 Schultz A, Nottrott S, Hartmuth K, Lührmann R. 2006. RNA structural requirements for the
877 association of the spliceosomal hPrp31 protein with the U4 and U4atac snRNP. *J Biol*
878 *Chem* **281**:28278-86. doi:10.1074/jbc.M603350200

879 Stanek D, Neugebauer KM. 2006. The Cajal body: a meeting place for spliceosomal snRNPs in
880 the nuclear maze. *Chromosoma* **115**:343-54. doi:10.1007/s00412-006-0056-6

881 Studier FW. 2005. Protein production by auto-induction in high-density shaking cultures. *Protein*
882 *Expres Purif* **41**:207-34. doi:10.1016/j.pep.2005.01.016

883 Thunnissen MMGM, Taddei N, Liguri G, Ramponi G, Nordlund P. 1997. Crystal structure of
884 common type acylphosphatase from bovine testis. *Structure* **5**:69-79. doi:Doi
885 10.1016/S0969-2126(97)00167-6

886 Unni S, Huang Y, Hanson RM, Tobias M, Krishnan S, Li WW, Nielsen JE, Baker NA. 2011. Web
887 servers and services for electrostatics calculations with APBS and PDB2PQR. *J Comput*
888 *Chem* **32**:1488-91. doi:10.1002/Jcc.21720

889 Vagin A, Teplyakov A. 2010. Molecular replacement with MOLREP. *Acta Crystallogr D* **66**:22-5.
890 doi:10.1107/S0907444909042589

891 van den Ent F, Lowe J. 2000. Crystal structure of the cell division protein FtsA from *Thermotoga*
892 *maritima*. *EMBO J* **19**:5300-7. doi:10.1093/emboj/19.20.5300

893 Verdone L, Galardi S, Page D, Beggs JD. 2004. Lsm proteins promote regeneration of pre-
894 mRNA splicing activity. *Curr Biol* **14**:1487-91. doi:10.1016/j.cub.2004.08.032

895 Vidal VPI, Verdone L, Mayes AE, Beggs JD. 1999. Characterization of U6 snRNA-protein
896 interactions. *RNA* **5**:1470-81. doi:10.1017/S1355838299991355

897 Vidovic I, Nottrott S, Hartmuth K, Lührmann R, Ficner R. 2000. Crystal structure of the
898 spliceosomal 15.5kD protein bound to a U4 snRNA fragment. *Mol Cell* **6**:1331-42.
899 doi:10.1016/S1097-2765(00)00131-3

900 Wahl MC, Will CL, Lührmann R. 2009. The spliceosome: design principles of a dynamic RNP
901 machine. *Cell* **136**:701-18. doi:10.1016/j.cell.2009.02.009

902 Warkocki Z, Odenwälder P, Schmitzova J, Platzmann F, Stark H, Urlaub H, Ficner R, Fabrizio
 903 P, Lührmann R. 2009. Reconstitution of both steps of *Saccharomyces cerevisiae* splicing
 904 with purified spliceosomal components. *Nat Struct Mol Biol* **16**:1237-43.
 905 doi:10.1038/nsmb.1729
 906 Will CL, Lührmann R. 2011. Spliceosome structure and function. *Cold Spring Harb Perspect*
 907 *Biol* **3**:1-24. doi:10.1101/cshperspect.a003707
 908 Wolff T, Bindereif A. 1992. Reconstituted mammalian U4/U6 snRNP complements splicing - a
 909 mutational analysis. *EMBO J* **11**:345-59. doi:
 910 Wolff T, Bindereif A. 1993. Conformational changes of U6 RNA during the spliceosome cycle:
 911 an intramolecular helix is essential both for initiating the U4-U6 interaction and for the first
 912 step of slicing. *Genes Dev* **7**:1377-89. doi:10.1101/gad.7.7b.1377
 913 Wolff T, Bindereif A. 1995. Mutational analysis of human U6 RNA: stabilizing the intramolecular
 914 helix blocks the spliceosomal assembly pathway. *Biochim Biophys Acta* **1263**:39-44.
 915 doi:10.1016/0167-4781(95)00085-U
 916 Wu J, Manley JL. 1989. Mammalian pre-mRNA branch site selection by U2 snRNP involves
 917 base pairing. *Genes Dev* **3**:1553-61. doi:10.1101/gad.3.10.1553
 918 Yean SL, Lin RJ. 1991. U4 small nuclear RNA dissociates from a yeast spliceosome and does
 919 not participate in the subsequent splicing reaction. *Mol Cell Biol* **11**:5571-7.
 920 doi:10.1128/MCB.11.11.5571
 921 Zhang HK, Gao SH, Lercher MJ, Hu SN, Chen WH. 2012. EvolView, an online tool for
 922 visualizing, annotating and managing phylogenetic trees. *Nucleic Acids Res* **40**:W569-W72.
 923 doi:10.1093/Nar/Gks576
 924 Zhou L, Hang J, Zhou Y, Wan R, Lu G, Yin P, Yan C, Shi Y. 2014. Crystal structures of the Lsm
 925 complex bound to the 3' end sequence of U6 small nuclear RNA. *Nature* **506**:116-20.
 926 doi:10.1038/nature12803

927 Zhuang Y, Weiner AM. 1989. A compensatory base change in human U2 snRNA can suppress
928 a branch site mutation. *Genes Dev* **3**:1545-52. doi:10.1101/gad.3.10.1545
929

Figure legends

Figure 1.

Protein and RNA requirements for Prp3 binding to U4/U6 di-snRNA.

- (A) Schematic presentation of yeast U4/U6 di-snRNA (yU4 – gold; yU6 – orange) and domain organizations of yeast and human Prp3. Regions corresponding to the C-terminal U4/U6-binding fragments (CTF) of the proteins are indicated by black lines above the schemes.
- (B) EMSA monitoring binding of hPrp3 protein variants (as MBP or GST fusions; 25 μ M) to hU4/U6^{stem II+13nt} (scheme on the left). hPrp3 constructs are indicated above the lanes. FL – full-length; N – residues 1-442; C – residues 195 – 683; CTF – residues 484-683. Bands are identified on the right; RNA – unbound RNA; RNPs – RNA-protein complexes
- (C) EMSA titrations monitoring binding of yPrp3^{CTF} variants and yPrp3^{DUF1115} (proteins indicated at the left of the gels) to yU4/U6^{stem II+13nt} (scheme on the top). Protein concentrations in each lane are indicated above the first gel. Bottom: Quantification of the data above. The data were fit to a single exponential Hill equation (fraction bound = $A[\text{protein}]^n/([\text{protein}]^n + K_d^n)$; A, fit maximum of RNA bound; n, Hill coefficient) (Ryder et al., 2008). Errors indicate standard errors of the mean of at least two independent experiments.
- (D) Isothermal titration calorimetry monitoring interactions between yPrp3^{CTF} variants or yPrp3^{DUF1115} (proteins indicated above and on the left of each panel) to yU4/U6^{stem II+10nt} (scheme in the first panel). The proteins, in particular yPrp3^{DUF1115} and yPrp3^{CTF,R322A}, tended to aggregate when added in excess of available RNA binding sites, giving rise to background signals at the ends of some runs. Data points in gray in the DUF1115 analysis were omitted during the fitting. Deduced binding stoichiometries (N), K_d 's, enthalpies (ΔH) and entropies (ΔS) of the interactions are listed in the lower parts of the panels.

(E) Binding of yPrp3^{CTF} (20 μM) to the indicated fragments of yU4/U6 (schemes above the gels; arrows indicate sequential shortening of RNA elements). Lanes 1-14 – shortening of the yU6 3'-overhang. Lanes 15-20 – shortening of stem II.

(F) EMSA titrations monitoring binding of binding of yPrp3^{CTF} or yPrp3^{DUF1115} to yU4/U6^{stem II+13nt} bearing yU6 C69G (top two gels) or yU4 C12G (bottom two gels) exchanges that restore Watson-Crick base pairing (schemes on the far left; mutant residues highlighted in green). Proteins are indicated at the left of the gels. Protein concentrations in each lane are indicated above the first gel. Bottom: Quantification as in panel 1C. CTF-wt U4/U6 – reference copied from panel 1C. Errors indicate standard errors of the mean of at least two independent experiments.

Figure 1 – figure supplement 1.

Protein sequence comparisons.

Multiple sequence alignments of Prp3 orthologs, acylphosphatases (AcyP) and blue light photoreceptors containing the BLUF domain. Sequences were aligned using the program MUSCLE (Edgar, 2004), manually adjusted and displayed using ALSCRIPT (Barton, 1993). Symbols below the alignments denote secondary structure elements assigned using PDBsum (Laskowski et al., 1997) based on the crystal structures of yPrp3^{CTF} (this study), *Bos taurus* AcyP (PDB ID 2ACY) (Thunnissen et al., 1997) and the *R. sphaeroides* BLUF domain (PDB ID 2BYC) (Jung et al., 2005). Cyan triangles – yPrp3²⁹⁶⁻⁴⁶⁹ residues in contact with yU6 snRNA; red diamonds – residues in the flavin-binding pocket of the *R. sphaeroides* BLUF domain, yellow circles – phosphate-binding and active site residues in *B. taurus* AcyP. Residue numbers refer to the yPrp3 sequence.

Figure 1 – figure supplement 2.

RNA sequence comparisons.

(A,B) Multiple sequence alignments of U6 regions (B) forming stem II and the 3'-overhang and the U4 region (C) forming stem II in U4/U6 di-snRNAs. Residues in red form non-Watson-Crick pyrimidine-pyrimidine base pairs in the lower parts of stem II. The U6 sequences for *H. sapiens* (gi: 669633321), *M. musculus* (gi: 90855862), *D. rerio* (gi: 159248718), *X. tropicalis* (gi: 284925142), *D. melanogaster* (gi: 8768), *C. elegans* (gi: 6894), *S. pombe* (gi: 159883886) and *S. cerevisiae* (gi: 4512) as well as the U4 sequences for *H. sapiens* (gi: 36174), *M. musculus* (gi: 175433), *D. rerio* (gi: 111306314), *D. melanogaster* (gi: 45557156), *C. elegans* (gi: 6893), *S. pombe* (gi: 5084) and *S. cerevisiae* (gi: 172644) were obtained from NCBI (<http://www.ncbi.nlm.nih.gov>). The U4 sequence for *X. tropicalis* was extracted from the *X. tropicalis* 8.0 Genome (<http://www.xenbase.org/entry>) by using *H. sapiens* U6 as a query sequence. Sequences were aligned using the program MUSCLE (Edgar, 2004), manually adjusted and displayed using Geneious (<http://www.geneious.com>). The relevant U6 (B) and U4 regions (C) were extracted according to the full-length sequence alignments and secondary structure predictions.

(C) Schemes of U4/U6 stem II from selected organisms, highlighting (red) the non-Watson-Crick pyrimidine-pyrimidine base pairs in the lower parts of stem II.

Figure 2.

Effects on Brr2-mediated U4/U6 unwinding and interplay with other U4/U6 proteins.

(A) Native gels monitoring yU4/U6 di-snRNA unwinding by yBrr2 in the absence of other proteins (top), in the presence of yPrp3^{CTF} (middle) or yPrp3^{DUF1115} (bottom). Asterisks indicate radioactive label on yU4 snRNA.

(B) Quantification of the data in (A). The data were fit to a single exponential equation: % duplex unwound = $A\{1-\exp(-k_u t)\}$; A – amplitude of the reaction; k_u – apparent first-order rate constant; t – time. Amplitudes and rate constants are listed. Errors indicate standard errors of the mean of four independent experiments.

(C) Binding of yPrp3^{DUF1115} (left three lanes of each panel) or yPrp3^{CTF} (right three lanes of each panel) to yU4/U6^{fused} (lanes 1-6; left), yU4/U6^{fused} pre-bound to Snu13 (lanes 7-12; middle) or yU4/U6^{fused} pre-bound to Snu13 and Prp31¹⁻⁴⁶² (lanes 13-18; right) under otherwise identical conditions. All three panels are from the same gel and were re-grouped. Schemes of yU4/U6^{fused} alone or pre-bound by proteins are shown on the top.

Figure 3.

Structural overviews.

(A) Structure of a yPrp3^{CTF}. Secondary structure elements and termini are labeled. Dashed line indicates residues contained in the construct but not visible in the electron density.

(B) Structure of a yPrp3^{CTF}-yU4/U6^{stem II+10nt} complex with the protein in the same orientation as in Figure 3A. yPrp3^{CTF} – cyan; yU4 – gold; yU6 – orange. Dashed line indicates residues contained in the construct but not visible in the electron density. Sections 1-3, between which the yU6 3'-overhang changes direction on the surface of yPrp3^{CTF}, are indicated by black lines. Residues in the yU6 3'-overhang are numbered.

(C) Electrostatic surface potential of yPrp3^{CTF} in complex with yU4/U6^{stem II+10nt}. Blue – positive; red – negative. Units kT/e with k – Boltzmann's constant, T – absolute temperature, e – charge of an electron.

Figure 3 – figure supplement 1.

Structural comparisons.

(A) Ribbon plots of all crystallographically independent, superimposed structures of the yPrp3 DUF1115 domain. Isolated yPrp3^{DUF1115} – white; isolated yPrp3^{CTF} – gray colors; RNA-bound yPrp3^{CTF} – cyan and blue.

(B) Structural superimposition of the two crystallographically independent yPrp3^{CTF}-yU4/U6^{stemII+10nt} complexes based on the protein portions. yPrp3²⁹⁶⁻⁴⁶⁹ – cyan (complex 1) and blue (complex 2); yU4 – gold (complex 1) and light gray (complex 2); yU6 – orange (complex 1) and dark gray (complex 2).

Figure 3 – figure supplement 2.

Phylogenetic analysis.

(A) Structure-based phylogenetic tree of ferredoxin-like fold proteins. Red and gray boxes indicate structures that recognize and do not recognize nucleic acids, respectively. The position of the Prp3 ssRNA-binding domain is indicated by a cyan box.

(B-H) Comparison of the yPrp3 DUF1115 domain alone (B) and in complex with RNA (F) to *B. taurus* AcyP (C; PDB ID 2ACY) (Thunnissen et al., 1997), the *R. sphaeroides* BLUF domain (D; PDB ID 2BYC) (Jung et al., 2005), a structural model of the hPrp3 DUF1115 domain (E) (Korneta et al., 2012), the C-terminal RRM of the human spliceosomal U1A protein in complex with stem-loop I of U1 snRNA (G; PDB ID 1URN) (Oubridge et al., 1994), and Epstein-Barr viral DNA-binding domain in complex with DNA (H; PDB ID 1B3T) (Bochkarev et al., 1998). All structures are shown with the ferredoxin-like core elements in the same orientation. The cores are shown in color, additional elements in light gray. Nucleic acids – orange.

Figure 4.

Details of the yPrp3-yU4/U6 di-snRNA interaction.

- (A-H) Close-up views of yPrp3-RNA contacts. The protein and the RNAs are shown as semi-transparent cartoons (yPrp3^{CTF} – cyan; yU6 – orange; yU4 – gold) with interacting residues as sticks (colored by element; carbon or phosphorus – as the respective molecule; nitrogen – blue; oxygen – red, sulfur – yellow). Black dashed lines – hydrogen bonds or salt bridges. Panels (G) and (H) show overlays of unbound (protein – gray) and complex (protein – cyan; yU6 snRNA – orange) situations. Rotation symbols indicate the views relative to Figure 3B.
- (I) Binding of yPrp3^{CTF} (20 μM) to wt and mutant versions (indicated above the gel) of U4/U6^{stem II+13nt} (scheme on the left; mutated region in green).
- (J) Variants of yPrp3^{CTF} (20 μM; indicated above the gel) binding to yU4/U6^{stem II+10nt} (scheme on the left).

Figure 5.

Consequences of yPrp3 variants in vivo.

- (A) Growth of yeast strains producing the indicated yPrp3 variants at various temperatures. Serial dilutions of liquid cultures were spotted on YPD agar plates and incubated at the indicated temperatures for 1-2 days.
- (B) In vivo splicing assays, monitoring levels of U3 pre-snoRNAs in yeast strains producing wt yPrp3 or the indicated yPrp3 variants. Cells were grown at 37 °C for 5 h before total RNA was extracted and U3 mature or pre-snoRNAs were detected by primer extension using a radiolabeled DNA oligonucleotide complementary to a region in U3 exon 2.
- (C,D) Northern blotting of cellular extracts of wt (top) or the indicated *prp3* mutant yeast strains (middle and bottom). Cellular extracts were separated on 10-30 % (v/v) glycerol gradients. Even-numbered gradient fractions (indicated above the blots) were probed with

radiolabeled DNA oligomers complementary to snRNA regions. Positions of various snRNPs on the gradients are indicated below the blots. (D) Quantification of the data shown (C).

(E) Western blots of the same gradients. Odd numbered gradient fractions (indicated above the blots) were probed with anti-ySnu114 (top) and anti-yPrp3 (bottom) antibodies. Positions of various snRNPs on the gradients are indicated below the blots. Dotted boxes – yPrp3 signals in the U4/U6•U5 tri-snRNP fractions.

Figure 6.

Models for splicing-associated functions of Prp3.

(A) Function of Prp3 as a bridge in the U4/U6•U5 tri snRNP. U4/U6 di-snRNP – orange; U5 snRNP – red; snRNAs – black sticks; U4/U6 stem II and U6 3'-overhang – gold and orange sticks. Prp3 uses its C-terminal region to bind U4/U6 stem II and the U6 3'-overhang on the U4/U6 di-snRNP side (this work) and interacts with proteins Prp6 or Snu66 on the U5 snRNP side (Liu et al., 2006).

(B) Structure of yPrp24 in complex with yU6 snRNA (Montemayor et al., 2014). yPrp24 RRM 1-4 domains – white, light gray, dark gray and black, respectively. Regions of yU6 forming stem I, stem II and the 3'-overhang in yU4/U6 are shown in different blue colors. Cold-sensitive A62G mutation and its paired nucleotide, C85 – magenta.

(C) Model for the function of Prp3 during U4/U6 di-snRNP re-assembly after splicing. (i) Upon release from the spliceosome, U6 snRNA is bound by Prp24 and LSm proteins. (ii) Recruitment of Prp3 and association with its cognate U6 regions, which are partially exposed in the Prp24-U6 complex. Emerging Prp3-U6 interactions initiate detachment of U6 snRNA from Prp24. (iii) Incorporation of pre-assembled U4 snRNP may complete Prp24 displacement and assembly of U4/U6 di-snRNP. U4 incorporation may be aided by Prp3

1099 stabilizing the emerging stem II and by the cooperative binding of Snu13, Prp31 and Prp3
1100 to U4/U6 di-snRNAs.

1101

Tables

Table 1.

Crystallographic data.

Data set	yPrp3 ²⁹⁶⁻⁴⁶⁹	yPrp3 ³²⁵⁻⁴⁶⁹	yPrp3 ²⁹⁶⁻⁴⁶⁹ -yU4/U6 ^{stem II-2}
Data collection			
Wavelength (Å)	0.91840	0.91841	0.97968
Space group	C222 ₁	P6 ₅	C2
Unit cell parameters			
a,b,c (Å)	87.7, 161.2, 105.2	56.1, 56.1, 86.8	144.7, 59.6, 109.8
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 120.0	90.0, 118.5, 90.0
Resolution (Å)	50.0-2.70 (2.80-2.70) ^a	50.0-2.00 (2.12-2.00)	30.0-3.25 (3.33-3.25)
Reflections			
Unique	20036 (1993)	10459 (1622)	24895 (1844)
Completeness (%)	97.3 (99.0)	99.3 (96.2)	97.9 (98.5)
Redundancy	3.4 (3.3)	11.4 (11.3)	2.0 (1.9)
R _{meas} ^b	0.066 (0.791)	0.072 (0.459)	0.017 (0.146)
I/σ(I)	13.31 (1.11)	24.96 (5.94)	5.38 (0.90)
CC(1/2) ^c	-	99.9 (95.5)	99.4 (42.0)
Refinement			
Resolution (Å)	30.00-2.70 (2.77-2.70)	24.28-2.00 (2.20-2.00)	29.78-3.25 (3.38-3.25)
Reflections			
Number	18932 (1146)	10454 (2424)	24894 (2631)
Completeness (%)	95.8 (80.5)	99.3 (97.0)	98.3 (98.0)
Test set (%)	5.2	5.0	5.0
R factors ^d			
R _{work} (%)	20.7 (37.0)	16.1 (16.6)	24.9 (36.9)
R _{free} (%)	26.2 (42.7)	22.1 (21.9)	29.9 (38.9)
RMSD ^e			
Bond lengths (Å)	0.010	0.013	0.007
Bond angles (°)	1.316	1.278	1.006
Ramachandran plot ^f (%)			
Favored	96.57	98.56	90.87
Allowed	2.64	1.44	7.98
Outlier	0.79	0.00	1.14

^a Values for the highest resolution shell in parentheses.

^b $R_{meas} = \sum_h [n/(n-1)]^{1/2} \sum_i |I_h - I_{h,i}| / \sum_h \sum_i I_{h,i}$, where I_h is the mean intensity of symmetry-equivalent reflections and n is the redundancy.

^c CC(1/2) is the percentage of correlation between intensities from random half-data sets.

^d $R = \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$; $R_{work} - hkl \notin T$; $R_{free} - hkl \in T$; T – test set.

^e Root-mean-square deviation from target geometries.

^f Calculated with MolProbity (<http://molprobity.biochem.duke.edu/>).











