The EMBO Journal (2008) 27, 2929–2940 | © 2008 European Molecular Biology Organization | All Rights Reserved 0261-4189/08
www.embojournal.org

THE EMBO JOURNAL

Structure and function of an RNase H domain at the heart of the spliceosome

Vladimir Pena[1,2], Alexey Rozov[1], Patrizia Fabrizio[1], Reinhard Lührmann[1,*] and Markus C Wahl[2,3,*]

[1]Abteilung Zelluläre Biochemie, Max-Planck-Institut für Biophysikalische Chemie, Göttingen, Germany, [2]Abteilung Zelluläre Biochemie, AG Röntgenkristallographie, Max-Planck-Institut für Biophysikalische Chemie, Göttingen, Germany and [3]Universitätsmedizin, Georg-August-Universität, Göttingen, Germany

Precursor-messenger RNA (pre-mRNA) splicing encompasses two sequential transesterification reactions in distinct active sites of the spliceosome that are transiently established by the interplay of small nuclear (sn) RNAs and spliceosomal proteins. Protein Prp8 is an active site component but the molecular mechanisms, by which it might facilitate splicing catalysis, are unknown. We have determined crystal structures of corresponding portions of yeast and human Prp8 that interact with functional regions of the pre-mRNA, revealing a phylogenetically conserved RNase H fold, augmented by Prp8-specific elements. Comparisons to RNase H–substrate complexes suggested how an RNA encompassing a 5′-splice site (SS) could bind relative to Prp8 residues, which on mutation, suppress splice defects in pre-mRNAs and snRNAs. A truncated RNase H-like active centre lies next to a known contact region of the 5′SS and directed mutagenesis confirmed that this centre is a functional hotspot. These data suggest that Prp8 employs an RNase H domain to help assemble and stabilize the spliceosomal catalytic core, coordinate the activities of other splicing factors and possibly participate in chemical catalysis of splicing.

The EMBO Journal (2008) **27,** 2929–2940. doi:10.1038/emboj.2008.209; Published online 9 October 2008
Subject Categories: RNA; structural biology
Keywords: pre-mRNA splicing; Prp8 protein; RNase H domain; splicing catalysis; X-ray crystallography

Introduction

Nuclear precursor-messenger RNA (pre-mRNA) splicing constitutes an essential step in the maturation of most eukaryotic primary transcripts, during which non-coding intervening sequences (introns) are removed and neighbouring coding regions (exons) are concatenated. Each round of splicing proceeds through two sequential transesterification reactions that are facilitated by the spliceosome, a dynamic molecular machine composed of five small nuclear RNAs (snRNAs) and more than 120 proteins (Will and Lührmann, 2006). In the first step, the 2′-hydroxyl of an invariant adenosine in the intronic branch point sequence (BPS) attacks the 5′SS (splice site), liberating the 5′-exon and an intron lariat-3′-exon. In the second step, the 3′-hydroxyl of the 5′-exon carries out a nucleophilic attack on the 3′SS, leading to exon ligation and excision of the branched intron (Moore *et al*, 1993).

Spliceosomes form only in the presence of a pre-mRNA substrate, on which they initially assemble as inactive particles that subsequently undergo catalytic activation (Brow, 2002). Spliceosome assembly and maturation are accompanied by profound changes in the snRNA–pre-mRNA network (Nilsen, 1998; Staley and Guthrie, 1998) and by the dynamic exchange of many proteins (Will and Lührmann, 2006; Bessonov *et al*, 2008). They proceed through landmark intermediates that are defined by the sequential incorporation and release of spliceosomal sn ribonucleoprotein particles (snRNPs; U1, U2, U4/U6 and U5). Initially, U1 and U2 snRNPs recognize, through base pairing, the 5′SS and BPS, respectively, forming complex A, which is then joined by the U4/U6–U5 tri-snRNP, generating complex B. Subsequently, the U1 and U4 snRNPs are expelled and the Prp19/Cdc5 multi-protein complex (the nineteen complex, NTC, in yeast) is stably integrated, generating the activated spliceosome (complex B*). During activation, the U4/U6 duplex is unwound, U6 snRNA base pairs with U2 and forms a functionally important intramolecular stem loop (ISL) and the 5′SS is handed over from the 5′-end of U1 to a conserved ACAGAGA-box of U6.

Several lines of evidence suggest that the two transesterification reactions are chemically facilitated by the spliceosomal RNA components through a two metal ion mechanism (Steitz and Steitz, 1993; Sontheimer *et al*, 1997). Phosphorothioate suppression experiments (Yean *et al*, 2000) and structural studies (Huppler *et al*, 2002) had demonstrated that the U6 ISL positions a catalytically or structurally important metal ion. Furthermore, fragments of the U2 and U6 snRNAs alone can mediate a reaction that resembles the first step of splicing (Valadkhan and Manley, 2001). However, spliceosomal proteins are indispensable for the assembly and stabilization of a fully active snRNA–pre-mRNA network; for example, some 35 proteins are integral components of a salt-stable step-one spliceosomal core (Bessonov *et al*, 2008). Thus, the question remains whether and how any of these proteins might contribute directly to splicing catalysis (Collins and Guthrie, 2000).

Prp8 is the largest (2413 and 2335 residues in yeast and human) and most highly conserved spliceosomal protein (62% identity between yeast and human) and is considered a master regulator of the spliceosome (Collins and Guthrie, 2000; Grainger and Beggs, 2005). It can be physically cross-

*Corresponding authors. R Lührmann, Department of Cellular Biochemistry, Max Planck Institute for Biophysical Chemistry, MPI of Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany. Tel.: +49 551 201 1405; Fax: +49 551 201 1197;
E-mail: reinhard.luehrmann@mpi-bpc.mpg.de or MC Wahl, Research Group X-Ray Crystallography, MPI of Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany.
Tel.: +49 551 201 1046; Fax: +49 551 201 1197;
E-mail: mwahl@gwdg.de

Received: 3 September 2008; accepted: 18 September 2008; published online: 9 October 2008

linked to U5 (Dix *et al*, 1998) and U6 snRNAs (Vidal *et al*, 1999), to a conserved GU dinucleotide at the 5′-end of an intron (Reyes *et al*, 1996, 1999), the BPS (MacMillan *et al*, 1994; McPheeters and Muhlenkamp, 2003) and the 3′SS (Teigelkamp *et al*, 1995a, b). Furthermore, it interacts genetically with U4 and U6 snRNAs (Kuhn *et al*, 1999), both splice sites, the BPS, and a polypyrimidine (PPy) tract (Umen and Guthrie, 1995a, 1996; Collins and Guthrie, 1999; Siatecka *et al*, 1999) located between the BPS and the 3′SS. Prp8 is also thought to regulate the activities of key protein splice factors. It forms a salt-stable complex (Achsel *et al*, 1998) with the Brr2 helicase that is required for spliceosome catalytic activation (Laggerbauer *et al*, 1998; Raghunathan and Guthrie, 1998) and disassembly (Small *et al*, 2006) and with the Snu114 GTPase that regulates Brr2 activity (Small *et al*, 2006). To understand the molecular mechanisms underlying the functions of Prp8 in the spliceosome, we started to characterize its 3D structure.

Results

A newly identified compact portion of Prp8 contains an RNase H domain expanded by unique elements

Apart from a putative RNA-recognition motif in the centre (Grainger and Beggs, 2005) and a Jab1/MPN domain at the very C terminus (Pena *et al*, 2007; Zhang *et al*, 2007), no known domains are discernable based on the Prp8 sequence. Guided by conservation patterns and disorder predictions, we expressed residues 1796–2092 of yeast (sc) Prp8 (scPrp8$^{1796-2092}$), located directly N-terminal of the Jab1/MPN domain. Limited proteolysis with chymotrypsin gave rise to a fragment, scPrp8$^{1827-2092}$, which could be crystallized directly or after recloning as scPrp8$^{1836-2092}$. We solved the structures of scPrp8$^{1836-2092}$ and scPrp8$^{1827-2092}$ at 2.0- and 2.1-Å resolution, respectively, and the structure of the corresponding human (hs) Prp8 protein (hsPrp8$^{1755-2016}$), obtained by a similar strategy, at 1.9-Å resolution (Table I). scPrp8$^{1836-2092}$ and hsPrp8$^{1755-2016}$ crystallized with one, scPrp8$^{1827-2092}$ with two protein molecules per crystallographic asymmetric unit.

The yeast and human Prp8 structures each comprise an ~160-residue N-terminal α/β domain (NTD; encompassing strands β1–β6 and β1a/β1b, α-helices α1–α3 and 3_{10}-helices η1–η5) and an ~95-residue C-terminal helical domain (CTD; encompassing α-helices α4–α7 and 3_{10}-helices η6 and η7), which are closely associated (ca. 2000 Å^2 of combined buried surface area; Figure 1A). The overall structures are reminiscent of a left-hand mitten (Figure 1A, inset), in which a central six-stranded mixed β-sheet and the surrounding α-helices of the NTD correspond to the palm, an extended β-hairpin of the NTD comprises the thumb and the α-helical CTD represents the fingers. The thumb and fingers frame a channel across the palm, the width of which at half height varies between a minimum of 16.5/24.0 Å in the hsPrp8$^{1755-2016}$ structure (distances between side chains/between Cα atoms of facing residues) and 22.2/29.0 Å in the scPrp8$^{1827-2092}$ structure (Figure 1B and C). The floor and the front opening of this channel are carpeted with extensive patches of positive surface potential (Figure 1B). The structures of scPrp8$^{1836-2092}$ and scPrp8$^{1827-2092}$ exhibit pairwise root-mean-square deviations (r.m.s.d.s) of 0.8–1.0 Å and align to hsPrp8$^{1755-2016}$ with r.m.s.d.s of 1.2–1.7 Å over ca. 230 Cα atoms, demonstrating that the structures of the yeast and human proteins are very similar. Superimposition of the independent structures showed that the palm regions are essentially invariant (Figure 1C). In contrast, the thumbs and the C-terminal portions of the fingers retain their intrinsic structures but adopt different relative orientations with respect to the palm, indicating that these elements are in principle mobile.

Comparison of the present Prp8 structures to known structures in the Protein Data Bank (http://www.pdb.org) revealed a striking similarity of the palm region of the NTD to an RNase H fold (r.m.s.d. of 3.0 Å over 95 Cα atoms when compared with human RNase H1 (Nowotny *et al*, 2007); pdb entry 2QKK). RNase H-fold proteins exhibit a topologically identical and structurally very similar five-stranded mixed β-sheet and three corresponding α-helices surrounding the sheet (Figure 1A). The extended β-hairpin and the α-helical CTD are unique to the Prp8 protein, generating the mitten-like structure. The core of other RNase H-fold proteins can be likewise augmented by other elements (Supplementary Figure S1).

A 5′SS-contacting peptide has a variable sequence but a conserved structure

Residues 1894–1898 within the present portion of human Prp8 have been UV-crosslinked to the 5′SS in a trans-splicing system at the stage of U4/U6–U5 tri-snRNP addition and after establishment of the 5′SS-U6 ACAGAGA-box interaction (Reyes *et al*, 1996, 1999). The similarity of the protein to an RNase H fold suggests that its mode of RNA binding may resemble that of RNase H. To approximate how RNA may bind to this region of Prp8, we transplanted the RNA strand of an RNase H-DNA–RNA substrate complex (pdb ID 1ZBL) (Nowotny *et al*, 2005) on scPrp8$^{1827-2092}$ by superimposing the proteins. The nucleic acid came to lie on the central ridge between the thumb and the fingers (Figure 2A). At this position, the negatively charged sugar-phosphate backbone can favourably interact with the positively charged surfaces (Figure 1B).

We next asked whether we could use the scissile phosphate from the RNase H reaction as a landmark to assign the approximate positions of the 5′-exon and the intron around the 5′SS. Remarkably, considering this scissile phosphodiester bond as equivalent of the 5′-exon–intron junction, placed in the nucleotides, which would correspond to the conserved GU at positions +1 and +2 of the intron, directly next to the five amino-acid peptide that UV-crosslinks to this GU dinucleotide (Figure 2A) (Reyes *et al*, 1996, 1999). Thus, the simple model resulting from our structural comparison provides a facile explanation for the zero-length crosslink mapped between Prp8 and the 5′SS. In further agreement with this model, an 11-residue oligomer mimicking the 5′SS still specifically crosslinks to a recombinant portion of hsPrp8 comprising residues 1669–2034 (corresponding to scPrp8 residues 1741–2106; cited in Grainger and Beggs, 2005).

The peptide of hsPrp8 that crosslinks to the 5′-end of an intron and the corresponding yeast fragment designate the only portions of the present domains that significantly diverged in sequence (1894-QACLK-1898 and 1966-SAAMS-1970, respectively; Figure 3). Yet in both proteins, the peptides adopt an almost identical extended 3_{10}-helical structure (helix η5) at the centre of the pronounced channel (Figure 1C), suggesting that the unusual structure rather than the sequence of helix η5 may be functionally important.

Table I Crystallographic data

Data collection

	$ScPrp8^{1836–2092}$ Peak	Inflection	Remote	$scPrp8^{1827–2092}$	$hsPrp8^{1755–2016}$
Wavelength (Å)	0.97968	0.97987	0.91841	0.91841	0.91841
Temperature (K)	100	100	100	100	100
Space group	$P3_1$			$P2_12_12_1$	$P3_221$
Unit cell (*a*, *b*, *c*) (Å)	46.9, 46.9, 103.0			77.1, 84.2, 95.1	83.4, 83.4, 96.3
Resolution (Å)[a]	50.0–2.0 (2.07–2.00)	50.0–2.10 (2.18–2.10)	50.0–2.1 (2.18–2.10)	50.0–2.1 (2.18–2.10)	50.0–1.95 (2.02–1.95)
Reflections[b]					
Unique	33 148 (3005)	29 331 (1906)	29 376 (1929)	36 691 (3623)	28 624 (2828)
Completeness (%)	96.8 (88.2)	98.8 (98.2)	99.1 (98.5)	99.6 (99.9)	99.9 (100.0)
Redundancy	1.9 (1.7)	2.0 (1.9)	2.0 (2.0)	3.6 (3.7)	7.9 (8.0)
$I/\sigma(I)$	16.8 (2.1)	18.3 (3.9)	12.6 (2.6)	10.4 (1.9)	31.6 (4.9)
$R_{sym}(I)$[c]	3.2 (24.6)	3.8 (14.0)	8.2 (22.0)	10.4 (56.2)	4.9 (29.2)

Phasing

Resolution (Å)	Se sites	CC/CC_{weak}[d]	FOM[e]
50.0–2.5	8	35.3/25.0	0.42

Refinement

Resolution (Å)	30.0–2.0 (2.05–2.00)	30.0–2.1 (2.15–2.10)	30.0–1.95 (2.00–1.95)
Reflections			
Number	16 963 (1175)	36 392 (2696)	28 264 (2035)
Completeness (%)	99.5 (93.1)	98.9 (99.9)	98.7 (98.8)
Test set (%)	5	5	5
R_{work}[f]	17.7 (23.4)	24.0 (27.9)	20.5 (22.2)
R_{free}[f]	24.5 (32.0)	29.8 (30.7)	24.8 (32.1)
ESU (Å)[g]	0.13	0.20	0.10
Protein atoms	2107	4217	2126
Water oxygens	269	259	315
Ions	—	—	2 Na^+, 1 Cl^-
Mean B-factors (Å^2)			
Wilson	21.4	35.3	31.4
Protein	20.2	36.4	37.5
Water	31.4	41.1	47.2
Ions	—	—	38.7
ϕ/ψ favoured/allowed/ outliers[h] (%)	98.4/1.6/0	95.1/3.7/1.2	96.9/3.1/0
r.m.s.d. geometry			
Bond lengths (Å)	0.009	0.009	0.009
Bond angles (deg)	1.15	1.22	1.12
r.m.s.d. B-factors (Å^2)			
Main chain bonds	0.50	0.48	0.49
Main chain angles	0.87	0.85	0.91
Side chain bonds	1.70	1.28	1.45
Side chain angles	2.59	1.94	2.40
PDB ID	3E9O	3E9P	3E9L

[a]Data for the highest resolution shell in parentheses.
[b]Friedel pairs for anomalous data sets of $scPrp8^{1836–2092}$ were not merged during processing but were merged for refinement against the peak data set.
[c]$R_{sym}(I) = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i |I_i(hkl)|$; for n independent reflections and i observations of a given reflection; $\langle I(hkl)\rangle$—average intensity of the i observations.
[d]$CC = [\sum wE_oE_c \sum w - \sum wE_o \sum wE_c] / \{[\sum wE_o^2 \sum_w - (\sum wE_o)^2]\ [\sum wE_c^2 \sum w - (\sum wE_c)^2]\}^{1/2}$; w—weight (see http://shelx.uni-ac.gwdg.de/SHELX/shelx_de.pdf for full definitions).
[e]FOM—figure of merit $= |F(hkl)_{best}| / |F(hkl)|$; $F(hkl)_{best} = \sum_\alpha P(\alpha) F_{hkl}(\alpha) / \sum_\alpha P(\alpha)$.
[f]$R = \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$; R_{work}—$hkl \notin T$; R_{free}—$hkl \in T$; T—test set.
[g]ESU—estimated overall coordinate error based on maximum likelihood.
[h]Calculated with MolProbity (http://molprobity.biochem.duke.edu/).

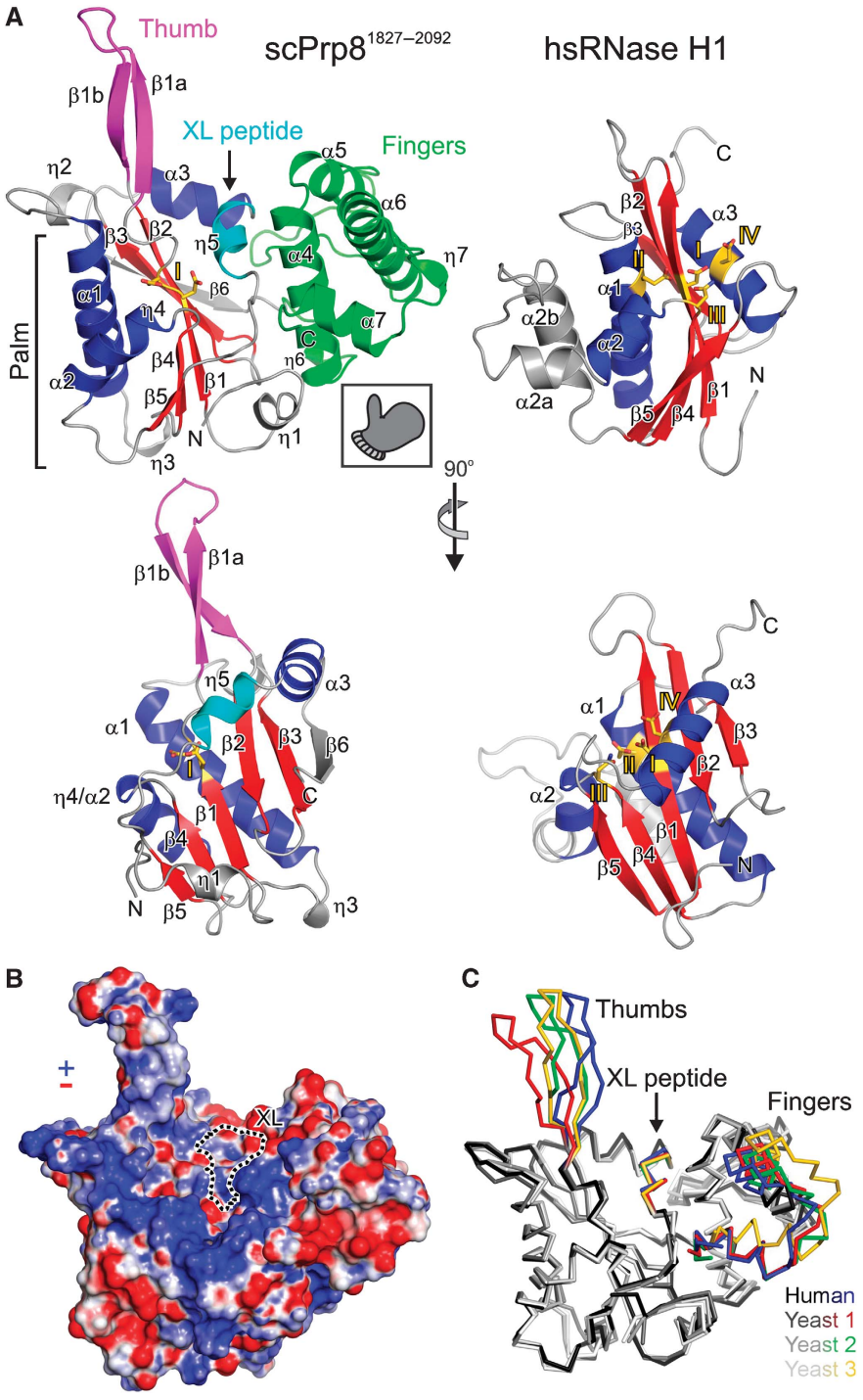

Figure 1 Structure of an RNase H domain in Prp8. (**A**) Ribbon diagram of scPrp8$^{1827–2092}$ (left) and human RNase H1 (right; pdb ID 2QKK; Nowotny *et al*, 2007) in the same orientations. Top panels: the shape of scPrp8$^{1827–2092}$ resembles a mitten (inset) with palm, thumb and finger regions. Secondary structure elements and termini are labelled. Equivalent α-helices and β-strands in Prp8 and RNase H are blue and red, respectively; β-hairpin comprising the thumb (strands β1a and β1b)—magenta; η5 encompassing the peptide crosslinked to the 5′SS in human Prp8—cyan; C-terminal domain of Prp8—green. Acidic residues I–IV involved in metal ion binding in RNase H (the asparagine at position III is an aspartate in the wild-type enzyme) as well as D1853 and D1854 in Prp8, which may correspond to carboxylate I, are shown as sticks (carbon—yellow; oxygen—red; nitrogen—blue). Bottom panels: view on the RNase H-like NTD of scPrp8$^{1827–2092}$ (left) and hsRNase H1 (right) rotated 90° about the vertical axis as compared with the top panels. The CTD of scPrp8$^{1827–2092}$ was removed for clarity. In scPrp8$^{1827–2092}$, the C-terminal helix α3 of hsRNase H1 is divided into helix η5 and the perpendicular helix α3. (**B**) Electrostatic potential mapped to the surface of scPrp8$^{1827–2092}$. Blue—positive charge; red—negative charge. The position of the crosslinked peptide is circled ('XL'). (**C**) Superimposition of scPrp8$^{1836–2092}$ (white/yellow) with two independent structure of scPrp8$^{1827–2092}$ (light grey/green and dark grey/red) and with hsPrp8$^{1755–2014}$ (black/blue). The thumbs, the peptides that crosslink to the 5′SS in the human system ('XL peptide') and the C-terminal portions of the fingers are shown in colours. Thumbs and fingers are evidently flexibly hinged to the central palms.

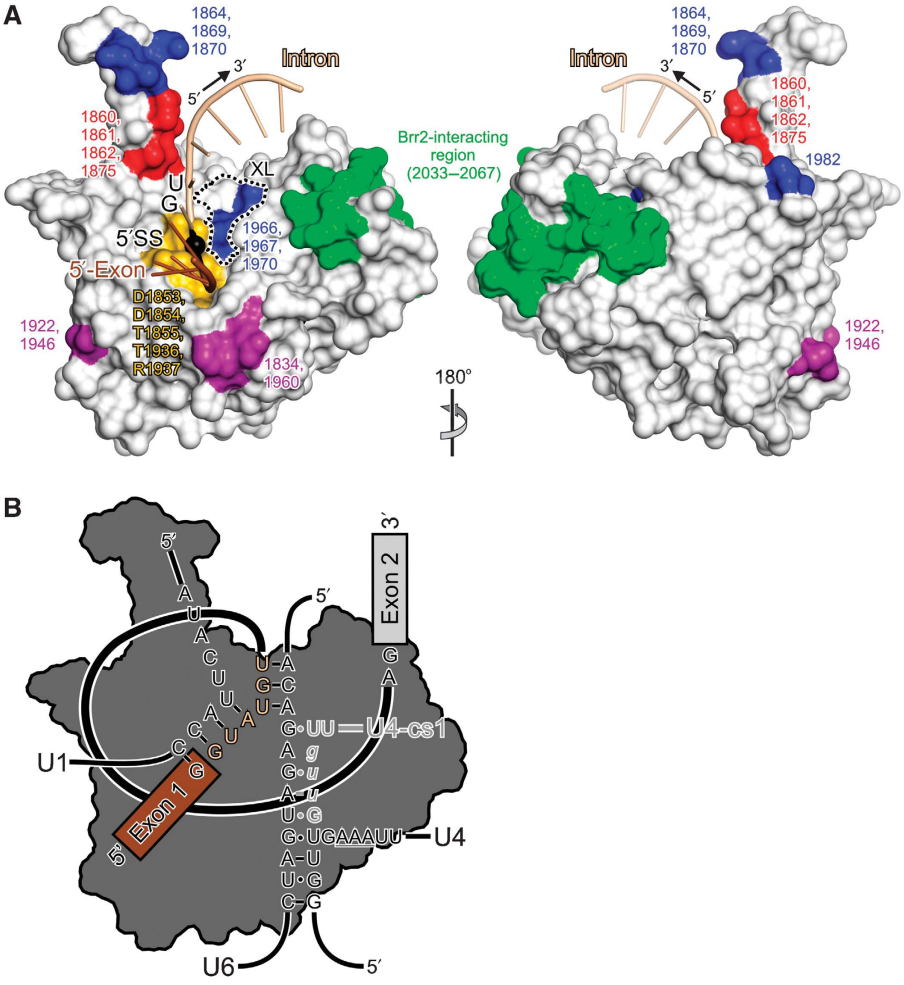

Figure 2 Functional interactions. (**A**) Diametrical views on the surface of scPrp8$^{1836-2092}$ with a modelled RNA mimicking a bound 5′SS region. The left orientation is identical to that in Figure 1A, top. 5′-exon—brown; 5′SS phosphate—black; intron—beige. The conserved GU at positions +1 and +2 of the intron are labelled. The surface patch corresponding to the peptide, which in hsPrp8 crosslinks to the 5′SS, is encircled. Positions that exhibit genetic interactions with the pre-mRNA or snRNAs are highlighted in colour. Interactions with 5′SS, 3′SS or BPS—blue; interactions with the PPy tract—purple (mutations at positions 1922 and 1946 always occur together with a mutation at 1834); interactions with *u4-cs1*—red. The region, in which mutations affect the physical interaction with Brr2, is in green. Residues changed in Prp8 mutants are listed. Conserved and invariant residues that lie in the region corresponding to the active site in RNase H enzymes (D1853, D1854, T1855, T1936 and R1937) are in gold. (**B**) RNA networks around the 5′SS shown on top of an outline of the present Prp8 domain. Before catalytic activation, U4 and U6 are base paired 3′ of the U6 ACAGAGA-box. U1 snRNA and U6 ACAGAGA-box interact at the 5′SS. In U4-cs1 (grey), an AAA sequence of U4 (underlined) is changed to UUG (lower case italics), which can form additional base pairs with U6 that involve the ACAGAGA-box as indicated.

Prp8 mutants that suppress pre-mRNA defects map on both sides of the putative path of the 5′SS RNA

Numerous Prp8 mutants (*prp8-D143* (K1864E), *prp8-151* (N1869D), *prp8-152* (N1869D/S1970R), *prp8-153* (T1982A), *prp8-154* (T1982A, SA1966/7AG), *prp8-155* (T1982A, V1987A), *prp8-162* (V1870N)) that suppress second-step defects (Query and Konarska, 2004) in the 5′SS, 3′SS or BPS (Umen and Guthrie, 1995a, 1996; Collins and Guthrie, 1999; Siatecka *et al*, 1999) map to the present Prp8 region. Residues 1966, 1967 and 1970 are part of the central peptide, the human equivalent of which crosslinks to the 5′SS (Figure 2A). The other mutants cluster in the flexible thumb and at its base (Figure 2A). According to our placement of RNA, the thumb is ideally positioned to directly interact with bound RNA. Similar hairpins that mediate RNA interactions are found in several ribosomal proteins, such as S5, S10, L4, L22 or L28 (Ban *et al*, 2000; Wimberly *et al*, 2000). However, unlike in Prp8, these modules are typically not stably structured off the ribosome (Worbs *et al*, 2000). These observations suggest that the above alleles exert an effect by affecting Prp8 regions that directly contact the 5′SS. As one possibility, the mutations may confer increased structural plasticity on the thumb and the central 5′SS-binding peptide, allowing them to accommodate deviant substrates.

Prp8 mutants that restore catalytic activation in the presence of a hyperstable U4-cs1/U6 duplex map to the flexible thumb

A variant U4 snRNA, U4-cs1, is thought to sequester the U6 ACAGAGA-box in base pairing and thereby inhibit catalytic activation (Kuhn *et al*, 1999) (Figure 2B). *prp8-201* (T1861P) and other *prp8-cat* alleles (F1851L, V1860D/N, V1862A/D/Y and I1875T) suppress *u4-cs1* phenotypes (Kuhn *et al*, 1999; Kuhn and Brow, 2000) and in their majority map at the base of the flexible thumb in direct vicinity of the positioned 5′SS RNA mimic (Figure 2A). On the basis of this observation, we

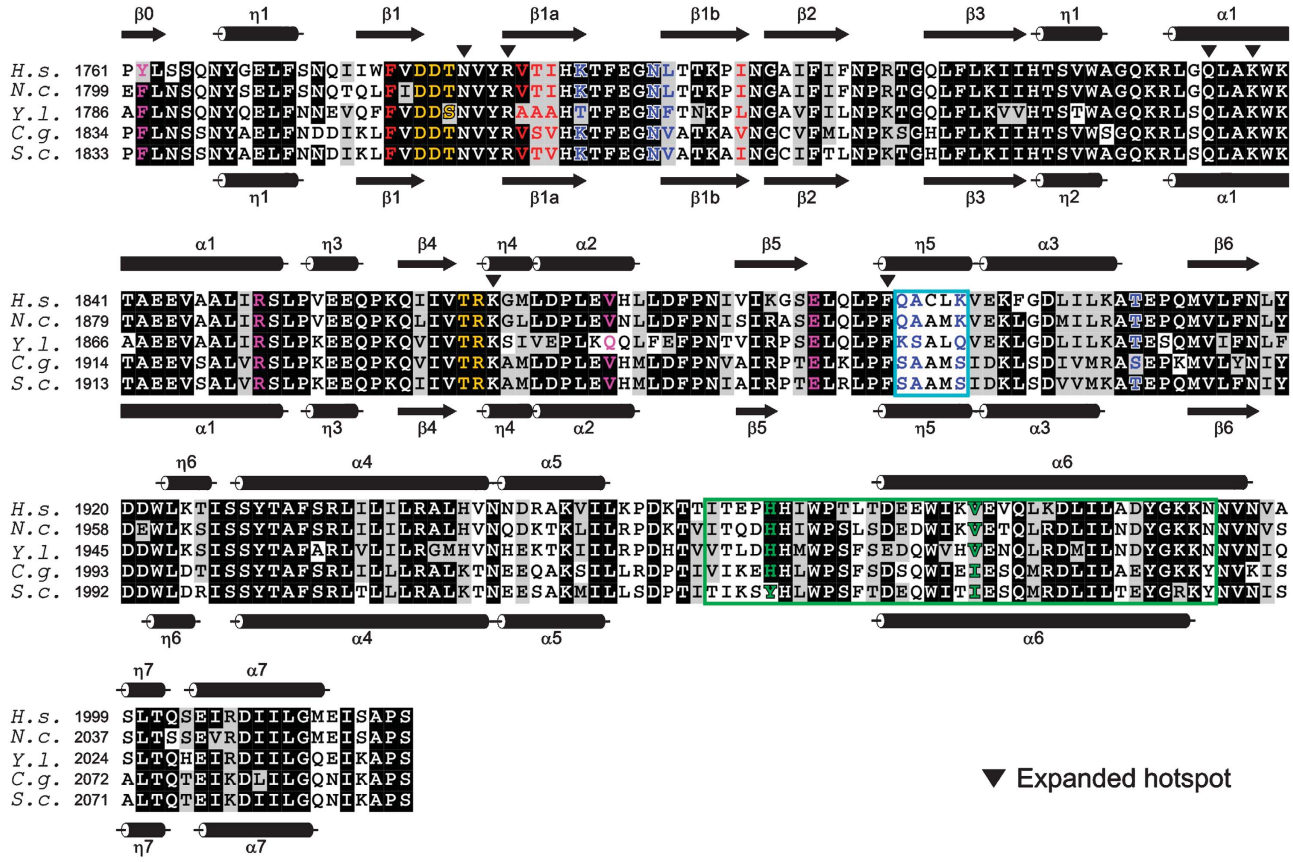

Figure 3 Multiple sequence alignment of Prp8 proteins. Darker background corresponds to higher conservation. Aligned orthologues are *H.s.—Homo sapiens*; *N.c.—Neurospora crassa*; *Y.l.—Yarrowia lipolytica*; *C.g.—Candida glabrata*; *S.c.—Saccharomyces cerevisiae*. Icons above/below the alignment indicate secondary structure elements of hsPrp8$^{1755-2016}$/scPrp8$^{1827-2092}$ as derived from the present crystal structures. Residues that exhibit genetic interactions with the 5′SS, 3′SS and BPS—blue; with *u4-cs1*—red; with the PPy tract—purple. Residues mutated in this work—yellow. Black triangles—conserved residues around the functional hotspot. The region of Prp8 crosslinked to the 5′SS and the region, in which mutations affect the interaction with Brr2, are boxed (cyan and green, respectively). Green residues in the Brr2-interacting region correspond to changes in the *prp8-52* allele.

suggest that the present Prp8 domain also interacts with U4/U6 and couples U4/U6 unwinding to U1-for-U6 strand exchange at the 5′SS (Figure 2B), consistent with a previous suggestion by Brow and coworkers (Kuhn *et al*, 1999).

Residues implicated in PPy tract recognition cluster on the Prp8 RNase H domain

Results from genetic suppression screens suggest that the present region of Prp8 is also involved in the recognition of the PPy tract (Umen and Guthrie, 1996). The corresponding alleles, *prp8-101* and *prp8-D102* to *prp8-D107*, exhibit E1960K/G or F1834L/S amino-acid exchanges; in *prp8-D103* and *prp8-D104*, a change at position 1834 occurs together with a V1946A and R1922G mutation, respectively. The affected residues are all contained in the current Prp8 structures. Strikingly, although they are some 130 residues apart in the protein sequence, the primary sites, 1834 and 1960, cluster immediately next to each other opposite to the flexible thumb and below the finger region (Figure 2A). This observation suggests that pre-mRNA regions close to the BPS and 3′SS may come to lie at the front entrance of the pronounced channel next to the 5′SS RNA. Consistent with this interpretation, crosslinking of the G-1 residue at the 3′SS to scPrp8 is reduced in the E1960K mutant of *prp8-101* (Umen and Guthrie, 1995b). Furthermore, binding of RNA at this location would be supported by the positive surface potential (Figure 1B).

Prp8 contains an incomplete RNase H-like active site

Our discovery of an RNase H fold in the region of Prp8 investigated herein naturally led to the question whether an RNase H-like active centre is present in Prp8. Canonical RNase H enzymes contain four conserved carboxylates in their active sites, used to position two catalytic metal ions (Nowotny *et al*, 2005) (Figure 4A). Two aspartates (positions I and III; Figure 1A; Supplementary Figure S1) are found in the first β-strand and C-terminal of the fourth β-strand of the RNase H core (corresponding to strands β1 and β4, respectively, of the present Prp8 domain); two additional acidic residues (positions II and IV; Figure 1A; Supplementary Figure S1) originate from the first and last helix of the RNase H core domain (equivalent to Prp8 helices α1 and η5/α3, respectively). In Prp8, the invariant D1853/1781 and D1854/1782 (yeast/human numbering; Figure 3) are found directly C-terminal of strand β1 in topologically equivalent positions to the first aspartate (I) of RNase H (Figure 4A and B), the carboxylate of which bridges the two metal ions in the RNase H catalytic situation (Nowotny and Yang, 2006). We

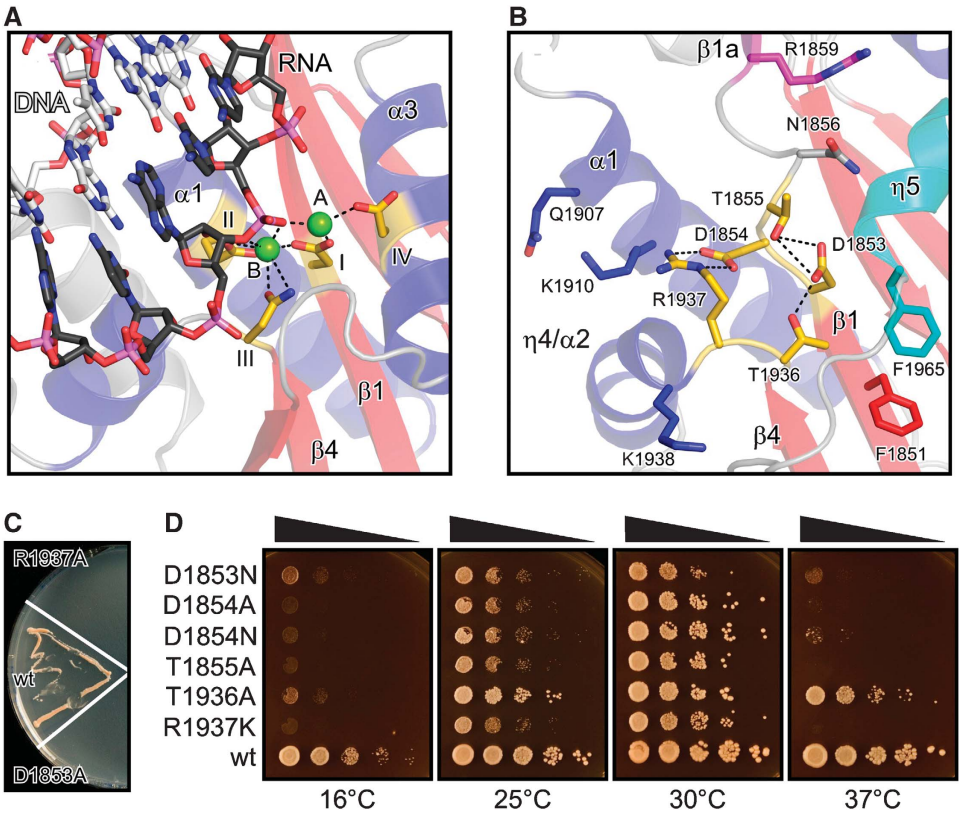

Figure 4 Identification of a novel functional hotspot. (**A**) Active site of hsRNase H1 in complex with two divalent metal ions (A and B; green spheres) and a DNA–RNA substrate duplex (carbons—light/dark grey; phosphorus—violet; other colours as in Figure 1). The four active site carboxylates (I–IV) are shown as sticks and are colour-coded as before (the asparagine at position III is an aspartate in the wild-type enzyme). Landmark secondary structure elements are labelled. The view is rotated 80° about the vertical axis and 20° about the horizontal axis as compared with the view in Figure 1A, top. (**B**) Close-up of the region of scPrp8$^{1827-2092}$, which is spatially equivalent to the RNase H active site, in the same orientation as in (A). The invariant D1853, D1854, T1855, T1936 and R1937 are shown as sticks and are colour-coded as above. Other conserved residues are labelled. Dashed lines indicate hydrogen bonds and salt bridges. Selected secondary structure elements are labelled. D1853 occupies the same topological and spatial position as carboxylate I of hsRNase H1, the carboxylate of D1854 spatially coincides with the carboxylate II of hsRNase H1. (**C**) Failure of D1853A and R1937A mutants of scPrp8 to grow on 5-FOA plates. (**D**) Cell viability assay monitoring the effects of exchanging invariant Prp8 residues D1853, D1854, T1855, T1936 and R1937 as indicated. After selection of clones, the culture and serial dilutions were spotted and grown at the indicated temperatures for 2 days.

did not find equivalents in Prp8 of the other RNase H catalytic residues, which would be topologically and spatially conserved. There is no acidic residue C-terminal of strand β4. Although two glutamates (E1915/1843 and E1916/1844) originate from helix α1, they lie on the backside of the helix and extend away from D1853/1781 and D1854/1782. The C-terminal helix of RNase H enzymes is represented by two elements in Prp8, helix η5, which in human Prp8 cross-links to the 5′SS, and helix α3, positioned at a right angle to η5 (Figure 1A). No acidic residue is found in helix η5. D1972/E1900 and D1976/1904 originate from helix α3 but are again remote from D1853/1781 and D1854/1782. Large conformational changes would be required to recruit any of these residues to a common active site centred on D1853/1781 or D1854/1782. Thus, the present Prp8 domain contains an incomplete set of metal-chelating residues or adopts a structure, in which these residues occupy non-productive spatial positions.

We tested whether divalent metal ions could bind to any of the Prp8 structures by soaking the crystals in solutions containing up to 20 mM Mg^{2+}, Mn^{2+} or Zn^{2+}. No metal ions were observed close to D1853/1781 or D1854/1782 (searching, for example, the anomalous difference Fourier maps of Mn^{2+}- and Zn^{2+}-soaked crystals). In one structure, a Mn^{2+} ion bound fortuitously between neighbouring molecules in the crystal lattice, remote from these aspartates. Although the failure to bind metal ions may be due to the lack of appropriately positioned carboxylates, any metal ion coordination would most likely be also strongly dependent on the presence of RNA, as observed earlier in other systems (Song *et al*, 2004).

A novel functional hotspot in Prp8

On transplanting the RNA strand from RNase H onto Prp8, the presumed 5′-exon–intron junction came to rest directly next to D1853/1781 and D1854/1782. These two aspartates are surrounded by additional invariant Prp8 residues, T1855/1783, T1936/1864 and R1937/1865 (Figures 3 and 4B). We investigated whether this cluster of residues was functionally important by mutagenesis. Single-residue mutants were brought into a genetic background, in which we could test whether they supported cell viability. All mutants were associated with dramatic growth defects (Figure 4C and D). D1853/1781 was lethal on conversion to alanine (Figure 4C) and produced both cold- and temperature-sensitive phenotypes on mutation to asparagine (Figure 4D). R1937/1865

was similarly lethal on mutation to alanine (Figure 4C) and conditionally lethal at low and high temperatures on change to a lysine (Figure 4D). D1854/1782, T1855/1783 and T1936/1864 were both cold- and temperature-sensitive on mutation to asparagine or alanines, respectively (Figure 4D). As some changes at each of the positions supported cell growth at moderate temperatures, we exclude a general folding problem as a source of the observed phenotypes. Remarkably, in a previous screen of 300 mutants covering the C-terminal three-quarters of the Prp8 sequence less than 10% (20 alleles) were found temperature sensitive and no cold-sensitive mutation was detected (Umen and Guthrie, 1996). Therefore, our findings mark a novel functional hotspot on Prp8 that coincides with a region spatially equivalent to an RNase H active centre.

Discussion

U2, U5 and U6 snRNAs and a single protein, Prp8, have been implicated as active site components of the spliceosome by the evolutionary conservation of regions that physically or genetically interact with the functional sites on the pre-mRNA (Collins and Guthrie, 2000). Here, we have investigated one such conserved region of yeast and human Prp8 by a combination of crystal structure analysis and site-directed mutagenesis. We find that this portion of Prp8 adopts an RNase H fold with additional elements specific to Prp8 orthologues. This observation motivated us to derive an approximate 5′SS-RNA-binding model from comparison with the structure of an RNase H–substrate complex, which was consistent with the surface topology and charge distribution of the Prp8 domain. Strikingly, on this crude model, biochemical and genetic interaction data quite naturally 'fall into place'. Numerous Prp8 mutants that genetically interact with the pre-mRNA, snRNAs or other protein splice factors mapped close to each other and around a peptide that can be UV-crosslinked to the 5′SS on the structure of the Prp8 RNase H domain (Figure 2A), strongly suggesting that the genetic interactions represent direct physical contacts. Our results provide a structural basis for expected Prp8 activities (Kuhn *et al*, 1999; Collins and Guthrie, 2000), suggesting how the protein mediates and proofreads RNA transactions on the way to a functional catalytic core (5′SS-U1 to 5′SS-U6 strand exchange, U4/U6 unwinding) and how it stabilizes RNA networks during catalysis (5′SS-U6 ACAGAGA-box, apposition of 5′SS and BPS). While this paper was under review, a paper describing a structure consistent with ours of a very similar fragment of yeast Prp8 appeared in press (Yang *et al*, 2008).

An RNase H domain as an ideal device to mediate nucleic acid transactions at the 5′SS on the way to the activated spliceosome

Genetic and physical interaction studies show that the Prp8 RNase H domain is involved in RNA rearrangements at the 5′SS during catalytic activation. At that time, the 5′SS-U1 snRNA duplex is unwound and the 5′SS-U6 ACAGAGA-box duplex is concomitantly formed. The RNase H domains of different enzymes are tuned to interact with diversely structured nucleic acids, such as DNA–RNA hybrids in the case of *bona fide* RNase H's (Nowotny *et al*, 2005), RNA duplexes in the case of Argonaute proteins (Song *et al*, 2004) or DNA duplexes in the case of retroviral integrases, DNA transposases or Holliday junction resolvases (Rice and Baker, 2001). Furthermore, the latter domains handle more than one nucleic acid strand or duplex at the same time as is also required during formation of the spliceosomal active sites and during splicing catalysis. Therefore, an RNase H domain seems to be the ideal choice for a device that needs to control coordinated structural rearrangements involving multiple RNA strands and duplexes (Figure 2B).

Coordination of activities and events during catalytic activation

Apart from Prp8, the activity of the DEAD-box protein Prp28 is also required for strand exchange at the 5′SS (Staley and Guthrie, 1999). Interestingly, the *prp8-201* allele (mapped at the base of the thumb of the Prp8 RNase H domain) and the *prp8-501* allele (I1825K and L1835F; coinciding with the PPy tract recognition region) interact genetically with the *prp28* locus (cited in Grainger and Beggs, 2005). In the framework of a dysfunctional U1-C protein, Prp28 is dispensable for catalytic activation (Chen *et al*, 2001) but becomes essential again in the presence of *prp8-201* or *prp8-501*. Thus, the Prp8-201 or Prp8-501 mutant proteins may interfere with U1-5′SS unwinding or with pairing of the U6 ACAGAGA-box to the 5′SS. It is possible, for example, that they stabilize the U1-5′SS duplex or inhibit positioning of the 5′SS.

The above genetic interactions suggest that Prp28 and the Prp8 RNase H domain exert an effect in close physical neighbourhood during catalytic activation. Indeed, a 5′SS RNA can be crosslinked to a position in Prp28 close to its ATP-binding site in the same spliceosomal assembly stage (Ismaili *et al*, 2001), in which the 5′SS GU dinucleotide crosslinks to the Prp8 RNase H domain (Reyes *et al*, 1996). Taken together, these observations show that the RNase H domain of Prp8 is involved in the recognition of the 5′SS early during the splicing cycle when it is still base-paired with U1 snRNA, and that it closely cooperates with Prp28 to provide the trans-helicase activity required for hand-over of the 5′SS from U1 to U6 during catalytic activation. In yeast Prp8, regions different from the RNase H domain can be crosslinked to the 5′SS (Turner *et al*, 2006). Thus, it is possible that the 5′SS binds only transiently at the present domain during early assembly stages of the spliceosome and is then transferred to other domains of Prp8.

The mapping of *prp8-cat* alleles (that suppress *u4-cs1* phenotypes) at the base of the thumb, close to the 5′SS interaction channel, indicates that the present domain of Prp8 also exerts an effect to couple U4/U6 unwinding to the rearrangements at the 5′SS, a previously postulated function of Prp8 (Kuhn *et al*, 1999). The exact nature of this coupling is not clear. Positioning of U4/U6 appropriately for unwinding could be affected by binding of the U6 ACAGAGA-box by the present Prp8 domain, which in turn may be dependent on correct positioning of the 5′SS RNA. In this picture, the *prp8-cat* suppressors may destabilize the non-natural U4-cs1/U6 duplex to allow catalytic activation.

Prp8 has also been implicated in coordinating the activities of other proteins that mediate spliceosome assembly and remodelling steps, including those of the DEXD/H-box protein Brr2 that is required for U4/U6 unwinding (Laggerbauer *et al*, 1998; Raghunathan and Guthrie, 1998). Interestingly, substitutions in region 2033–2067 of Prp8 (e.g. Y2037H and

I2051T in the *prp8-52* allele), which encompasses a long α-helix and the preceding loop on the backside of the finger region (Figure 2A), enhance the physical interaction with Brr2 in yeast two-hybrid and pull-down assays (van Nues and Beggs, 2001). Thus, apparently the Prp8 RNase H domain also coordinates RNA structural transitions by either recruiting the RNA remodelling enzyme Brr2 (directly or indirectly by interacting with a Brr2-binding site), by allosterically influencing its activities or by competing for its substrates. We have shown earlier that the Jab1/MPN domain of Prp8, which lies immediately C-terminal of the RNase H domain, also serves as a binding module for Brr2 (Pena *et al*, 2007). In the light of their close proximity, it appears feasible that Brr2 jointly recognizes epitopes on both of these Prp8 domains.

The Prp8 RNase H domain may also organize pre-mRNA regions close to the 3′SS

Fe-BABE-directed hydroxyl radical probing has demonstrated that the 5′-end of U2 snRNA is in close proximity of pre-mRNA functional regions and of U1 snRNP in early spliceosomal complexes (Dönmez *et al*, 2007). U2 snRNA base pairs with the BPS of the intron close to the PPy tract. The close proximity of U1, U2 and the functional pre-mRNA regions they bind would favour concomitant recognition of the 5′SS-U1 snRNA duplex and regions close to the U2-BPS duplex by Prp8, consistent with the mapping of PPy tract recognition mutants close to the 5′SS-binding region on our structures (Figure 2A). As the PPy tract marks the approximate position of the BPS, which attacks the 5′SS in the first step, we suggest that the Prp8 RNase H domain is involved in correct positioning of the two pre-mRNA regions that react in the first step of splicing. In activated spliceosomes, portions of the U6 ISL also approach the U2 BP-binding sequence and the 5′SS (Rhode *et al*, 2006). These data place the Prp8 RNase H domain in close proximity of the U6 ISL that is expected to constitute another active site component. Thus, the U6 ISL and the Prp8 RNase H domain may cooperate to achieve splicing catalysis.

The Prp8 RNase H domain facilitates structural transitions between two or more competitive conformations of the spliceosome

In analogy to the ribosome, the spliceosome may toggle between competing pre-step-one, step-one and step-two conformations, where the first and third states may resemble each other (Query and Konarska, 2004). According to this model, alleles that suppress step-two defects destabilize a step-one conformation (or interaction network), thus favouring adoption of the step-two conformation. Step-one mutants can be rationalized by a negative effect on a conformation (or interaction network) of the second step or of a pre-step-one stage. Significantly, both step-one (*prp8-101/D102* and *prp8-D103/4/5/6/7*) and step-two alleles (suppressors of all other pre-mRNA defects) map to the present structures. Combining our data with the toggling model, the step-one suppressor mutations could impair recognition of 3′-portions of the intron by the Prp8 RNase H domain in a pre-step-one stage (affecting assembly) or in the second step (affecting conversion of step one to step two or catalysis of step two). The step-two suppressor alleles could exert an effect by destabilizing Prp8–5′SS interactions required for the first step (affecting stabilization of a step-one situation or catalysis of step one). Thus, the Prp8 RNase H domain provides a platform that may exert an effect during catalytic activation and throughout both catalytic steps.

Possible roles of the Prp8 RNase H domain in the catalysis of pre-mRNA splicing

Pre-mRNA splicing chemically recapitulates self-splicing by group II introns and several snRNA elements resemble catalytic portions of group II introns. A recent crystal structure of a hydrolytic group IIC intron in the post-catalytic state (Toor *et al*, 2008) revealed that the U6-like ISL binds two metal ions 3.9 Å apart, in agreement with a two-metal-ion mechanism mediated by that site, which may also apply to the spliceosome. However, details of the spliceosomal catalytic strategies may have diverged from the RNA-based mechanisms of group II introns (Michel and Ferat, 1995; Collins and Guthrie, 2000). Most obviously, catalysis by the spliceosome depends on many protein splice factors but their precise roles are normally not known. Our results show how an RNase H domain of the Prp8 protein is involved in the correct preparation and positioning of substrates during pre-mRNA splicing. As positioning is a crucial catalytic function, this finding alone attributes to Prp8 a direct role in splicing catalysis.

Similar to group II introns, RNase H-fold enzymes employ a two-metal-ion mechanism that invariably leads to retention of the phosphate group on the 3′-portion of the product and a free 3′-hydroxyl on the 5′-portion, as observed in both steps of splicing. Our mutational analysis has identified candidate residues on the RNase H domain in Prp8 (Figure 4B), through which the protein might even chemically facilitate splicing catalysis. For example, Prp8 may contribute D1853/1781 or D1854/1782 (yeast/human numbering; equivalent to the RNase H carboxylate I) to a composite RNA–protein active site, in which they could assist RNA residues in coordinating a catalytic metal ion.

Previous data indicated a particular importance for the recognition of the base moiety at position +2 of the intron (Reyes *et al*, 1996). Residues in the functional hotspot and the neighbouring Q1907/1835, which is conserved and recognizes substrate bases in other RNase H-fold enzymes (Supplementary Figure S2), are appropriately positioned to mediate such interactions (Figure 4B). Residues in this region may also recognize backbone functionalities. They could sequester the 2′-hydroxyl group of a scissile phosphate, which in an irregular RNA structure, such as around the 5′SS in the first step, could otherwise function as a nucleophile and lead to aberrant 5′-OH and 2′–3′-cyclic phosphate products. However, none of the 2′-OH groups is absolutely required for trans-splicing of a 5′SS oligo (Konarska, 1998) and the 2′-OH at the 5′SS is not essential for the first step of splicing (Moore and Sharp, 1992), in general agreement with the conditional phenotypes we observe with many mutants. It is also conceivable that R1937/1865, which is notably conserved in resolvases and transposases (Supplementary Figure S2), could carry out functions normally associated with the catalytic metal ions (activation of the nucleophile, stabilization of the transition state or stabilization of the negative charge on the leaving group) in one of the steps. An analogous situation has been observed in *Eco*RV, in which a lysine cooperates with metal ions (Horton *et al*, 1998).

Additional invariant residues (N1856/1784, R1859/1787, K1910/1838 and K1938/1866) surround the functional hotspot (Figure 4B) and are suitably positioned for RNA binding. Furthermore, the invariant F1965/1893 rests on another conserved phenylalanine, F1851/1779, and fixes the N terminus of the η5 helix. Extended 3_{10}-helices are rarely found in proteins and may interconvert with α-helices (Topol *et al*, 2001). Such transitions at η5 could be important for the remodelling around the 5′SS during spliceosome assembly and catalysis. For example, structural transitions on RNA binding could expose F1965/1893 and allow it to stack with nucleobases. Interestingly, F1865L is the only *prp8-cat* mutation that maps to the present domain outside of the thumb. A leucine at position 1851 provides a smaller platform for F1965, which could in turn more easily rearrange to interact with an RNA ligand.

Taken together, our results point to various mechanisms by which the spliceosome may function as a veritable ribonucleoprotein enzyme ('RNPzyme').

Truncation of the Prp8 RNase active site may protect the pre-mRNA during assembly

The activation of substrates in the two transesterification reactions of pre-mRNA splicing has to be timed with exquisite precision. An RNase H domain that handles functional pre-mRNA regions during assembly and catalysis poses the danger of activating a water molecule that could initiate an inappropriate hydrolytic attack on the pre-mRNA. Thus, mechanisms must be at work, which inactivate the Prp8 RNase H domain as a hydrolytic enzyme. In the conformation seen in the present structures, the Prp8 RNase-like active site is incomplete and does not support metal ion binding on its own. It is conceivable that conformational changes on substrate binding could lead to rearrangements that bring about a complete active site. Perhaps more likely, the contraction of multiple RNA elements on the surface of the present Prp8 domain in proximity of the 5′SS could lead to complementation of the truncated protein active site by RNA elements. Similar mechanisms have been observed before in proteins that were originally presumed to be pseudoenzymes (Mukherjee *et al*, 2008). The expected molecular crowding may additionally help to exclude water molecules from the active site environment.

Materials and methods

Details of the cloning, expression, purification and the crystallographic procedures are given in the Supplementary data.

Limited proteolysis

For analytical digestion, 10 μg of scPrp8$^{1796-2092}$ was mixed with increasing amounts of chymotrypsin (Sigma-Aldrich) and incubated for 1 h on ice. The reactions were stopped by the addition of SDS–PAGE loading buffer and incubation at 95°C. A stable band resulting from treatment with chymotrypsin was analysed by tryptic peptide mass fingerprinting (Griffin *et al*, 1995) and indicated scPrp8$^{1827-2092}$ as a stable subfragment of scPrp8$^{1796-2092}$.

For preparative digestion of scPrp8$^{1796-2092}$ with chymotrypsin, 3.5 mg of scPrp8$^{1796-2092}$ was mixed with 0.075 mg chymotrypsin and incubated for 19 h at room temperature. The reaction was stopped by the addition of phenylmethylsulphonylfluoride (PMSF) to 1 mM and dialysis against 10 mM Tris–HCl, pH 7.5, 150 mM NaCl, 1 mM DTT and 5% glycerol. The digested protein was concentrated to 7 mg/ml.

For preparative digestion of hsPrp8$^{1747-2016}$ with chymotrypsin, 10 mg of the protein was mixed with 0.4 mg of chymotrypsin and incubated for 2 h at 16°C. The reaction was stopped by the addition of PMSF to 0.1 mM and the buffer was exchanged to 10 mM Tris–HCl, pH 7.5, 150 mM NaCl, 1 mM DTT, 5% glycerol on a NAP-25 column (GE Healthcare). A stable band that resulted from the proteolysis was analysed by tryptic peptide mass fingerprinting and identified as hsPrp8$^{1755-2016}$. The truncated protein was concentrated to 37 mg/ml.

Site-directed mutagenesis and plasmid shuffling

The desired substitutions were introduced into the *PRP8* gene of plasmid pJU186 (Umen and Guthrie, 1995a) (wild-type *PRP8* on plasmid pSE362 (*CEN*, *HIS*) was a kind gift of Catherine Collins and Christine Guthrie, University of California, San Francisco, USA) by the QuikChange site-directed mutagenesis strategy (Stratagene) and verified by sequencing. Wild-type and mutant plasmids were transformed into yeast strain JDY8.06 (kindly provided by Richard Grainger and Jean Beggs, University of Edinburgh, UK), containing wild-type *PRP8* on a counter-selectable *URA3*-marked plasmid (Brown and Beggs, 1992). Before plasmid shuffling (Boeke *et al*, 1987), cells were selected at 25°C in a medium lacking histidine. Transformants were streaked once on medium lacking histidine and grown at 25°C and patches were streaked three times on 5-fluoroorotic acid (5-FOA) plates to select for cells lacking the *URA3* plasmid. Cells that survived on 5-FOA plates were streaked on rich medium, and their growth phenotypes were analysed by incubating cells (ca. 5×10^4 cells) and serial eight-fold dilutions at 37, 30, 25 and 16°C for 1–2 days. Cold- and temperature-sensitive strains failed to grow on −Ura plates at any of the temperatures and, in addition, did not grow in rich medium at 16 or 37°C, respectively, but grew at 25 or 30°C.

Supplementary data

Supplementary data are available at *The EMBO Journal* Online (http://www.embojournal.org).

Acknowledgements

We are grateful to Catherine Collins and Christine Guthrie for providing wild-type *PRP8* in pSE362 and to Richard Grainger and Jean Beggs for providing yeast strain JDY8.06. We thank Klaus Hartmuth for critical reading of the paper and helpful discussions, Berthold Kastner for advice during cloning and expression, Henning Urlaub for mass spectrometric analyses, Elke Penka for technical assistance and the teams of beamlines 14.2 (BESSY, Berlin, Germany) and PXII (SLS, Villigen, Switzerland) for support during diffraction data collection. This study was supported by the Max-Planck-Gesellschaft (RL and MCW), the Volkswagen Stiftung (RL and MCW), the Bundesministerium für Bildung und Forschung (RL and MCW), the Fonds der Chemischen Industrie (RL) and the Ernst-Jung-Stiftung (RL).

References

Achsel T, Ahrens K, Brahms H, Teigelkamp S, Lührmann R (1998) The human U5-220kD protein (hPrp8) forms a stable RNA-free complex with several U5-specific proteins, including an RNA unwindase, a homologue of ribosomal elongation factor EF-2, and a novel WD-40 protein. *Mol Cell Biol* **18:** 6756–6766

Ban N, Nissen P, Hansen J, Moore PB, Steitz TA (2000) The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* **289:** 905–920

Bessonov S, Anokhina M, Will CL, Urlaub H, Lührmann R (2008) Isolation of an active step I spliceosome and composition of its RNP core. *Nature* **452:** 846–850

Boeke JD, Trueheart J, Natsoulis G, Fink GR (1987) 5-Fluoroorotic acid as a selective agent in yeast molecular genetics. *Methods Enzymol* **154:** 164–175

Brow DA (2002) Allosteric cascade of spliceosome activation. *Annu Rev Genet* **36:** 333–360

Brown JD, Beggs JD (1992) Roles of PRP8 protein in the assembly of splicing complexes. *EMBO J* **11:** 3721–3729

Chen JY, Stands L, Staley JP, Jackups Jr RR, Latus LJ, Chang TH (2001) Specific alterations of U1-C protein or U1 small nuclear RNA can eliminate the requirement of Prp28p, an essential DEAD box splicing factor. *Mol Cell* **7:** 227–232

Collins CA, Guthrie C (1999) Allele-specific genetic interactions between Prp8 and RNA active site residues suggest a function for Prp8 at the catalytic core of the spliceosome. *Genes Dev* **13:** 1970–1982

Collins CA, Guthrie C (2000) The question remains: is the spliceosome a ribozyme? *Nat Struct Biol* **7:** 850–854

Dix I, Russell CS, O'Keefe RT, Newman AJ, Beggs JD (1998) Protein–RNA interactions in the U5 snRNP of *Saccharomyces cerevisiae*. *RNA* **4:** 1239–1250

Dönmez G, Hartmuth K, Kastner B, Will CL, Lührmann R (2007) The 5′ end of U2 snRNA is in close proximity to U1 and functional sites of the pre-mRNA in early spliceosomal complexes. *Mol Cell* **25:** 399–411

Grainger RJ, Beggs JD (2005) Prp8 protein: at the heart of the spliceosome. *RNA* **11:** 533–557

Griffin PR, MacCoss MJ, Eng JK, Blevins RA, Aaronson JS, Yates III JR (1995) Direct database searching with MALDI-PSD spectra of peptides. *Rapid Commun Mass Spectrom* **9:** 1546–1551

Horton NC, Newberry KJ, Perona JJ (1998) Metal ion-mediated substrate-assisted catalysis in type II restriction endonucleases. *Proc Natl Acad Sci USA* **95:** 13489–13494

Huppler A, Nikstad LJ, Allmann AM, Brow DA, Butcher SE (2002) Metal binding and base ionization in the U6 RNA intramolecular stem-loop structure. *Nat Struct Biol* **9:** 431–435

Ismaili N, Sha M, Gustafson EH, Konarska MM (2001) The 100-kda U5 snRNP protein (hPrp28p) contacts the 5′ splice site through its ATPase site. *RNA* **7:** 182–193

Konarska MM (1998) Recognition of the 5′ splice site by the spliceosome. *Acta Biochim Pol* **45:** 869–881

Kuhn AN, Brow DA (2000) Suppressors of a cold-sensitive mutation in yeast U4 RNA define five domains in the splicing factor Prp8 that influence spliceosome activation. *Genetics* **155:** 1667–1682

Kuhn AN, Li Z, Brow DA (1999) Splicing factor Prp8 governs U4/U6 RNA unwinding during activation of the spliceosome. *Mol Cell* **3:** 65–75

Laggerbauer B, Achsel T, Lührmann R (1998) The human U5-200kD DEXH-box protein unwinds U4/U6 RNA duplices *in vitro*. *Proc Natl Acad Sci USA* **95:** 4188–4192

MacMillan AM, Query CC, Allerson CR, Chen S, Verdine GL, Sharp PA (1994) Dynamic association of proteins with the pre-mRNA branch region. *Genes Dev* **8:** 3008–3020

McPheeters DS, Muhlenkamp P (2003) Spatial organization of protein–RNA interactions in the branch site-3′ splice site region during pre-mRNA splicing in yeast. *Mol Cell Biol* **23:** 4174–4186

Michel F, Ferat JL (1995) Structure and activities of group II introns. *Annu Rev Biochem* **64:** 435–461

Moore MJ, Query CC, Sharp PA (1993) Splicing of precursors to mRNAs by the spliceosome. In *The RNA World First edition*, Gesteland RF, Atkins JF (eds), pp 303–357. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press

Moore MJ, Sharp PA (1992) Site-specific modification of pre-mRNA: the 2′-hydroxyl groups at the splice sites. *Science* **256:** 992–997

Mukherjee K, Sharma M, Urlaub H, Bourenkov GP, Jahn R, Südhof TC, Wahl MC (2008) CASK functions as a Mg^{2+}-independent neurexin kinase. *Cell* **133:** 328–339

Nilsen TW (1998) RNA–RNA interactions in nuclear pre-mRNA splicing. In *RNA Structure and Function*, Simons RW, Grunberg-Manago M (eds), pp 279–308. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press

Nowotny M, Gaidamakov SA, Crouch RJ, Yang W (2005) Crystal structures of RNase H bound to an RNA/DNA hybrid: substrate specificity and metal-dependent catalysis. *Cell* **121:** 1005–1016

Nowotny M, Gaidamakov SA, Ghirlando R, Cerritelli SM, Crouch RJ, Yang W (2007) Structure of human RNase H1 complexed with an RNA/DNA hybrid: insight into HIV reverse transcription. *Mol Cell* **28:** 264–276

Nowotny M, Yang W (2006) Stepwise analyses of metal ions in RNase H catalysis from substrate destabilization to product release. *EMBO J* **25:** 1924–1933

Pena V, Liu S, Bujnicki JM, Lührmann R, Wahl MC (2007) Structure of a multipartite protein–protein interaction domain in splicing factor prp8 and its link to retinitis pigmentosa. *Mol Cell* **25:** 615–624

Query CC, Konarska MM (2004) Suppression of multiple substrate mutations by spliceosomal prp8 alleles suggests functional correlations with ribosomal ambiguity mutants. *Mol Cell* **14:** 343–354

Raghunathan PL, Guthrie C (1998) RNA unwinding in U4/U6 snRNPs requires ATP hydrolysis and the DEIH-box splicing factor Brr2. *Curr Biol* **8:** 847–855

Reyes JL, Gustafson EH, Luo HR, Moore MJ, Konarska MM (1999) The C-terminal region of hPrp8 interacts with the conserved GU dinucleotide at the 5′ splice site. *RNA* **5:** 167–179

Reyes JL, Kois P, Konforti BB, Konarska MM (1996) The canonical GU dinucleotide at the 5′ splice site is recognized by p220 of the U5 snRNP within the spliceosome. *RNA* **2:** 213–225

Rhode BM, Hartmuth K, Westhof E, Lührmann R (2006) Proximity of conserved U6 and U2 snRNA elements to the 5′ splice site region in activated spliceosomes. *EMBO J* **25:** 2475–2486

Rice PA, Baker TA (2001) Comparative architecture of transposase and integrase complexes. *Nat Struct Biol* **8:** 302–307

Siatecka M, Reyes JL, Konarska MM (1999) Functional interactions of Prp8 with both splice sites at the spliceosomal catalytic center. *Genes Dev* **13:** 1983–1993

Small EC, Leggett SR, Winans AA, Staley JP (2006) The EF-G-like GTPase Snu114p regulates spliceosome dynamics mediated by Brr2p, a DExD/H box ATPase. *Mol Cell* **23:** 389–399

Song JJ, Smith SK, Hannon GJ, Joshua-Tor L (2004) Crystal structure of Argonaute and its implications for RISC slicer activity. *Science* **305:** 1434–1437

Sontheimer EJ, Sun S, Piccirilli JA (1997) Metal ion catalysis during splicing of premessenger RNA. *Nature* **388:** 801–805

Staley JP, Guthrie C (1998) Mechanical devices of the spliceosome: motors, clocks, springs, and things. *Cell* **92:** 315–326

Staley JP, Guthrie C (1999) An RNA switch at the 5′ splice site requires ATP and the DEAD box protein Prp28p. *Mol Cell* **3:** 55–64

Steitz TA, Steitz JA (1993) A general two-metal-ion mechanism for catalytic RNA. *Proc Natl Acad Sci USA* **90:** 6498–6502

Teigelkamp S, Newman AJ, Beggs JD (1995a) Extensive interactions of PRP8 protein with the 5′ and 3′ splice sites during splicing suggest a role in stabilization of exon alignment by U5 snRNA. *EMBO J* **14:** 2602–2612

Teigelkamp S, Whittaker E, Beggs JD (1995b) Interaction of the yeast splicing factor PRP8 with substrate RNA during both steps of splicing. *Nucleic Acids Res* **23:** 320–326

Toor N, Keating KS, Taylor SD, Pyle AM (2008) Crystal structure of a self-spliced group II intron. *Science* **320:** 77–82

Topol IA, Burt SK, Deretey E, Tang TH, Perczel A, Rashin A, Csizmadia IG (2001) alpha- and 3(10)-helix interconversion: a quantum-chemical study on polyalanine systems in the gas phase and in aqueous solvent. *J Am Chem Soc* **123:** 6054–6060

Turner IA, Norman CM, Churcher MJ, Newman AJ (2006) Dissection of Prp8 protein defines multiple interactions with crucial RNA sequences in the catalytic core of the spliceosome. *RNA* **12:** 375–386

Umen JG, Guthrie C (1995a) A novel role for a U5 snRNP protein in 3′ splice site selection. *Genes Dev* **9:** 855–868

Umen JG, Guthrie C (1995b) Prp16p, Slu7p, and Prp8p interact with the 3′ splice site in two distinct stages during the second catalytic step of pre-mRNA splicing. *RNA* **1:** 584–597

Umen JG, Guthrie C (1996) Mutagenesis of the yeast gene PRP8 reveals domains governing the specificity and fidelity of 3′ splice site selection. *Genetics* **143:** 723–739

Valadkhan S, Manley JL (2001) Splicing-related catalysis by protein-free snRNAs. *Nature* **413:** 701–707

van Nues RW, Beggs JD (2001) Functional contacts with a range of splicing proteins suggest a central role for Brr2p in the dynamic control of the order of events in spliceosomes of *Saccharomyces cerevisiae*. *Genetics* **157:** 1451–1467

Vidal VP, Verdone L, Mayes AE, Beggs JD (1999) Characterization of U6 snRNA–protein interactions. *RNA* **5:** 1470–1481

Will CL, Lührmann R (2006) Spliceosome structure and function. In *The RNA World Third Edition*, Gesteland RF, Cech TR, Atkins JF

(eds), pp 369–400. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press

Wimberly BT, Brodersen DE, Clemons WM, Morgan-Warren RJ, Carter AP, Vonrhein C, Hartsch T, Ramakrishnan V (2000) Structure of the 30S ribosomal subunit. *Nature* **407:** 327–339

Worbs M, Huber R, Wahl MC (2000) Crystal structure of ribosomal protein L4 shows RNA-binding sites for ribosome incorporation and feedback control of the S10 operon. *EMBO J* **19:** 807–818

Yang K, Zhang L, Xu T, Heroux A, Zhao R (2008) Crystal structure of the β-finger domain of Prp8 reveals analogy to ribosomal proteins. *Proc Natl Acad Sci USA* **105:** 13817–13822

Yean SL, Wuenschell G, Termini J, Lin RJ (2000) Metal-ion coordination by U6 small nuclear RNA contributes to catalysis in the spliceosome. *Nature* **408:** 881–884

Zhang L, Shen J, Guarnieri MT, Heroux A, Yang K, Zhao R (2007) Crystal structure of the C-terminal domain of splicing factor Prp8 carrying retinitis pigmentosa mutants. *Protein Sci* **16:** 1024–1031