

ARTICLES

Isolation of an active step I spliceosome and composition of its RNP core

Sergey Bessonov¹, Maria Anokhina¹, Cindy L. Will¹, Henning Urlaub² & Reinhard Lührmann¹

Formation of catalytically active RNA structures within the spliceosome requires the assistance of proteins. However, little is known about the number and nature of proteins needed to establish and maintain the spliceosome's active site. Here we affinity-purified human spliceosomal C complexes and show that they catalyse exon ligation in the absence of added factors. Comparisons of the composition of the precatalytic versus the catalytic spliceosome revealed a marked exchange of proteins during the transition from the B to the C complex, with apparent stabilization of Prp19–CDC5 complex proteins and destabilization of SF3a/b proteins. Disruption of purified C complexes led to the isolation of a salt-stable ribonucleoprotein (RNP) core that contained both splicing intermediates and U2, U5 and U6 small nuclear RNA plus predominantly U5 and human Prp19–CDC5 proteins and Prp19-related factors. Our data provide insights into the spliceosome's catalytic RNP domain and indicate a central role for the aforementioned proteins in sustaining its catalytically active structure.

Splicing of pre-messenger RNA is catalysed by the spliceosome, which consists of the U1, U2, U4/U6 and U5 snRNPs, and numerous proteins¹. During spliceosome assembly, short conserved sequences of the pre-mRNA, including the 5' and 3' splice sites and the branch point sequence (BPS) are recognized sequentially¹. Initially, the U1 and U2 snRNPs bind, forming the A complex, followed by the U4–U6.U5 tri-snRNP, generating the pre-catalytic B complex. Subsequent activation of the spliceosome entails RNP rearrangements leading to the displacement of U1 and U4 (ref. 2). The activated B* spliceosome catalyses step I of splicing, during which the branch-point adenosine attacks the 5' splice site, generating the cleaved 5' exon and intron–3'-exon lariat intermediates. The spliceosomal C complex is formed at this time and catalyses the second step of splicing, during which the intron is excised and the 5' and 3' exons are ligated to form mRNA.

A complex RNA network involving the spliceosomal small nuclear RNAs (snRNAs) and the pre-mRNA is formed during spliceosome assembly³. Initially, U1 snRNA base-pairs with the 5' splice site and U2 with the BPS. U1 is then displaced by U6, which base-pairs with the 5' splice site and also U2, and the U5 snRNA contacts the 5' exon. The resulting RNA structure, involving pre-mRNA, U2, U5 and U6, has a central role in catalysing the first step of splicing. Before the second step, additional rearrangements occur^{4,5}, but they are still not well understood.

RNA is a major component of the spliceosome's catalytic core and there is good evidence that catalysis itself is RNA based⁶. Nonetheless, establishing a catalytically active RNA network in the spliceosome requires the assistance of proteins⁷. Several proteins seem to be core components of the spliceosome. These include the following: first, Prp8, which directly contacts the 5' splice site, the BPS and the 3' splice site of the pre-mRNA⁸; second, Prp19 and a group of proteins that associate tightly with it (the 'nineteen complex' (NTC) in yeast or the Prp19–CDC5 complex in humans), which are crucial during catalytic activation^{9–13}; and third, proteins of the heteromeric SF3a and SF3b complexes, which contact the pre-mRNA at and/or upstream of the BPS, and stabilize the U2–BPS interaction^{14,15}. However, little is known about the number and nature of proteins required for maintaining a catalytically active structure in the spliceosome's core.

Purification of the B and C complexes

To identify proteins essential to the spliceosome's RNP core, we set out to isolate spliceosomes that catalyse splicing in the absence of added factors. Previously, an adenovirus-derived pre-mRNA lacking a 3' splice site and 3' exon was shown to catalyse bimolecular exon ligation (that is, step II of splicing) in nuclear extract, when offered a 3' splice-site-containing RNA *in trans*^{16,17}. To isolate catalytically active C complexes, we performed splicing *in vitro* with PM5 pre-mRNA, which contains a 60-nucleotide polypyrimidine tract but no terminal AG dinucleotide and no 3' exon (Fig. 1a). When incubated under splicing conditions with HeLa nuclear extract, PM5 underwent 5' splice site cleavage and intron lariat formation but no exon ligation (Fig. 1b), and initially B complexes and later C complexes were efficiently formed (Fig. 1c). After allowing C complexes to form on the PM5 substrate, an

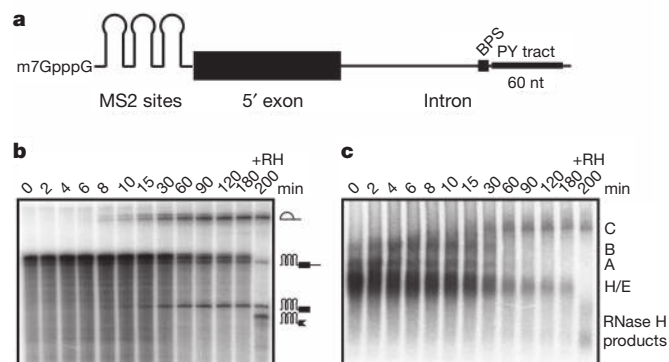


Figure 1 | Splicing of PM5 pre-mRNA lacking a 3' splice site and 3' exon. **a**, Diagram of PM5 pre-mRNA. PY, polypyrimidine tract; nt, nucleotides. **b**, **c**, Kinetics of *in vitro* splicing and splicing complex formation with PM5 pre-mRNA. Splicing was performed in HeLa nuclear extract for 0–180 min, followed by digestion with RNase H (RH) (right-hand lane). **b**, RNA was analysed by denaturing PAGE; **c**, splicing complexes were analysed on an agarose gel. The positions of the pre-mRNA, splicing intermediates and RH digestion product(s) and the H, A, B and C complexes are indicated on the right.

¹Department of Cellular Biochemistry, and ²Bioanalytical Mass Spectrometry Group, Max Planck Institute of Biophysical Chemistry, D-37077 Göttingen, Germany.

RNA with the sequence 5'-GACAG-3', where AG represents the 3' splice site, followed by a 55-nucleotide adenovirus-derived 3' exon¹⁶ (Fig. 2a) was added to the splicing reaction. To control for proper 3' splice site recognition, an identical RNA with an AG to GG mutation (5'-GACGG-3' exon), which should lead to the use of the downstream, proximal-most AG dinucleotide¹⁶, was added alternatively. Northern blotting revealed the generation of mRNA products of the expected size with both 3' splice-site-containing RNAs, showing that spliceosomes formed on PM5 can catalyse bimolecular exon ligation (Fig. 2b). No mRNA product was observed if C complexes were not allowed to form, and sequencing after RT-PCR confirmed that splicing occurred at the expected 3' splice site in each case (data not shown).

The coliphage coat protein MS2 was used to affinity purify both B and C complexes formed on the PM5 substrate; purification was performed under identical, physiological conditions. Affinity-purified B complexes contained nearly equimolar amounts of U1, U2, U4, U5 and U6 snRNA together with unspliced pre-mRNA (Fig. 3a, lanes 1 and 3). Purified C complexes, in contrast, contained nearly equimolar amounts of U2, U5 and U6 plus first-step intermediates (that is, the excised 5' exon and the intron lariat) and only small amounts of unspliced pre-mRNA (Fig. 3a, lanes 2 and 4). Thus, highly pure complexes were isolated in each case.

Purified C complexes are active

To determine whether affinity-purified C complexes are active, we performed bimolecular exon ligation assays in the presence or

absence of micrococcal nuclease (MN)-treated nuclear extract. Exon ligation was observed at an efficiency comparable to that of non-purified complexes both in the presence of MN-treated nuclear extract (Fig. 2c, lanes 1 and 2) and, more strikingly, even in its absence (Fig. 2c, lanes 3 and 4). Mutation of the 3' splice site to GG also led to the expected switch in 3' splice site use. Note that the mRNA doublet seen with both 3' substrate RNAs is due to 3' end heterogeneity and not incorrect 3' splice site use (data not shown). Bimolecular exon ligation with purified C complexes was most efficient when performed in the presence of ATP in comparison with other NTPs, and less efficient in the presence of non-hydrolysable (AMP-PNP) or slowly hydrolyzable (ATP- γ S) ATP analogues (Fig. 2d). This represents the first instance where affinity-purified, non-immobilized spliceosomal complexes on their own could be shown to possess catalytic activity, and thus indicates that our affinity-purified C complexes contain all the factors necessary for step II of splicing.

Protein dynamics of the B to C transition

Proteins present in affinity-purified C, and also B, complexes were separated by SDS-PAGE (Fig. 3b) and identified by mass spectrometry (MS) (Supplementary Tables 1 and 2). MS consistently identified about 130 proteins in B complexes, the vast majority of which had previously been found in B complexes formed on an adenovirus-derived pre-mRNA¹⁸. Affinity-purified C complexes contained about 150 proteins, about 105 of which were also consistently found in B complexes. Proteins shared by both complexes include most proteins associated with U5 and U2 snRNPs, the Prp19-CDC5 protein complex and related factors, and components of the RES (retention and splicing) complex. Several proteins are lost during the B to C transition—that is, they are found solely or predominantly in B complexes—as shown by immunoblotting (see below) and the number of peptides sequenced by MS. Although not strictly quantitative, a comparison of the number of peptides identified is a good indication of the relative amounts of a protein present in two different samples when they are directly, sequentially analysed. Proteins lost or highly underrepresented in C complexes include almost all U1 and U4

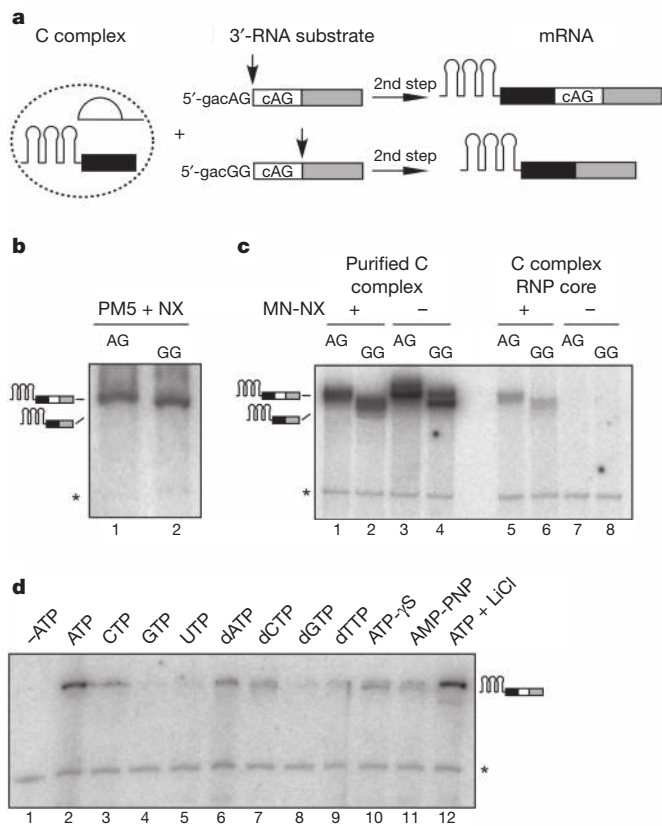


Figure 2 | Affinity-purified C complexes catalyse bimolecular exon ligation on their own. **a**, Diagram of 3' substrate RNAs and mRNA products of bimolecular exon ligation. Arrow, expected 3' splice site cleavage site. **b–d**, Bimolecular exon ligation with PM5 pre-mRNA in nuclear extract (NX) (**b**), with purified C complexes or salt-treated C-complex RNP cores with or without MN-treated NX, and AG or GG 3' RNAs (**c**), or with purified C complexes and indicated nucleotides plus AG 3' RNA (**d**). Exon ligation was monitored by northern blotting with a 3' RNA probe. Asterisk, ³²P-labelled, excised 5' exon (C complex loading control).

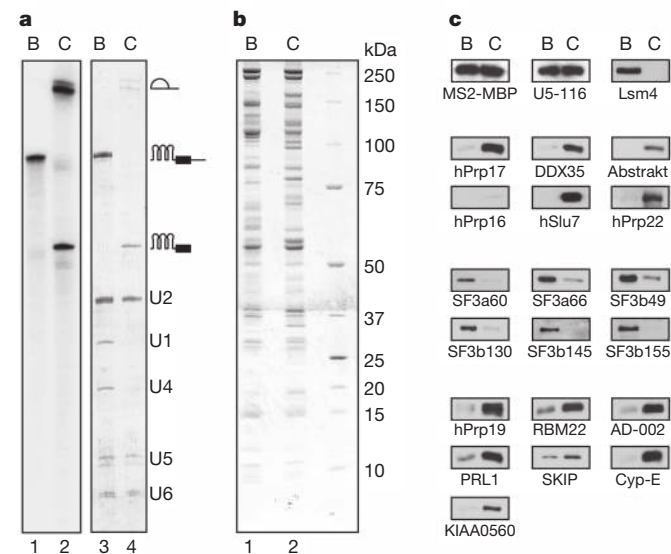


Figure 3 | Characterization of MS2 affinity-purified B and C complexes. **a**, RNA composition of MS2 affinity-purified B (lanes 1 and 3) and C (lanes 2 and 4) complexes. RNA was analysed by denaturing PAGE and detected by silver staining (lanes 1 and 2) or by autoradiography (lanes 3 and 4). RNA identities are indicated at the right. **b**, Protein composition of purified B (lane 1) and C (lane 2) complexes. Proteins were separated by SDS-PAGE and detected by staining with Coomassie blue. **c**, Protein composition of B or C complexes as revealed by immunoblotting with the indicated antibodies.

proteins, the U6 Lsm proteins, and several non-snRNP factors. Immunoprecipitation studies in yeast indicated that the Lsm proteins are destabilized during spliceosome activation, thereby allowing the 3' end of U6 to base pair with the 5' splice site¹². Our data provide direct evidence that human Lsm proteins are also destabilized or lost during the B to C transition, indicating that this important remodelling event is conserved among eukaryotes.

Several proteins were found solely or predominantly in C complexes and are thus recruited during catalytic activation or step I of splicing. These included most second-step factors, the DEAD-box helicases Abstrakt and DDX35, and the peptidyl-prolyl isomerases (PPIases) PPIL3b, PPWD1 and PPIG. Thus, one or more of these helicases or PPIases might facilitate conformational changes accompanying step II of splicing. Finally, most core components of the exon junction complex (for example, eIF4A3, Magoh and Y14) seemed to be enriched in C complexes.

To determine the absence or presence of selected proteins, we also performed immunoblotting (Fig. 3c). Antibodies against either U5–116K, a core component of B and C complexes, or the fusion protein of MS2 with maltose-binding protein (MBP) confirmed that comparable amounts of both complexes were analysed. Consistent with our MS results, Lsm4 was detected solely in B complexes, whereas the step II factors human Prp22 (hPrp22) and hSlu7, together with DDX35 and Abstrakt, were found exclusively in C complexes. Only trace amounts of hPrp16 were detected in C complexes. Because the latter catalyse exon ligation on their own, they might be stalled at a stage after hPrp16 has acted or, alternatively, the small amounts of Prp16 still present might be sufficient for step II.

More peptides were sequenced by MS for most proteins comprising the Prp19–CDC5 complex and for Prp19–CDC5-related proteins in the C versus the B complex (Supplementary Table 1). Consistent with this observation, more intense immunostaining of Prp19, RBM22, AD002, PRL1, SKIP, Cyp-E and KIAA0560 was observed with C complexes (Fig. 3c). Thus, Prp19 and related proteins seem to be enriched in C complexes, suggesting that they are more stably associated at this stage. In contrast, fewer peptides were consistently identified by MS for proteins comprising SF3a and SF3b in the C

versus the B complex (Supplementary Table 1). Immunoblotting with antibodies against various SF3a/b proteins confirmed that they are much less abundant in the C complex (Fig. 3c). Thus, SF3a/b components seem to be destabilized during the B to C transition (and are therefore lost more readily during purification). The apparent destabilization or loss of SF3a/b proteins, and stabilization of Prp19–CDC5 proteins, indicates that protein–protein and/or protein–RNA contacts involving these proteins are remodelled before or during step I of splicing.

Proteins in the spliceosome's RNP core

To determine which proteins are required for maintenance of the catalytically active RNA network in the step I spliceosome, purified C complexes were subjected to increasing concentrations of NaCl up to 1 M, followed by glycerol-gradient centrifugation. Complexes treated with 1 M salt sedimented as a single RNP peak, as determined by the distribution of ³²P-labelled RNA across the gradient (Fig. 4a). Only trace amounts of radiolabelled RNA were detected at the top of the gradient, indicating that nearly all of the excised 5' exon and lariet intron remained associated with this RNP complex. Gradient fractions 13–15 (RNP peak) contained nearly equimolar amounts of both first-step intermediates, as well as the U2, U5 and U6 snRNAs (Fig. 4b, c), suggesting that relatively homogeneous (at least at the RNA level), intact RNP particles are still present after salt treatment. In contrast, although native B and C complexes have almost identical S values (data not shown), the vast majority of B complexes treated with 1 M salt sedimented much more slowly than the salt-stable RNP cores of the C complex (Fig. 4a), indicating that the former dissociate. Thus, catalytic activation of the spliceosome seems to lead to a generally more stable core RNP conformation, suggesting that protein–RNA and protein–protein contacts in the B-complex and C-complex cores must differ.

Salt-treated C complexes had a slightly lower S value (Fig. 4e) than native complexes, which is consistent with the dissociation of a subset of proteins. We therefore analysed the protein composition of gradient fractions 13–15 containing the salt-resistant RNP complex and also gradient fractions 1–5, which should contain dissociated

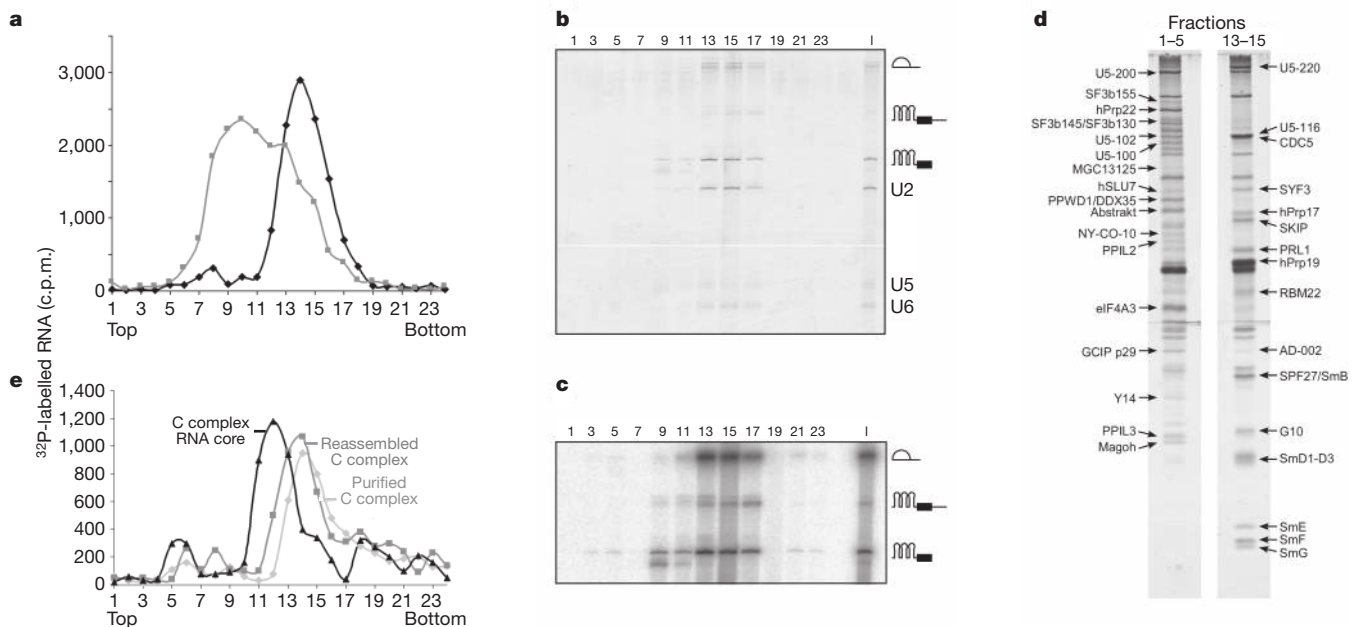


Figure 4 | U5- and Prp19–CDC5-proteins remain associated with C-complex RNP cores. **a**, Affinity-purified C complexes (black diamonds) or B complexes (grey squares) treated with 1 M NaCl and separated on glycerol gradients containing 1 M NaCl. **b, c**, RNA from salt-treated C complex gradient fractions, detected by silver staining (**b**) or autoradiography (**c**). I, 10% of input. **d**, Salt-treated C-complex proteins from gradient fractions 1–5 or 13–15, separated by SDS–PAGE and stained with silver. A subset of proteins exclusively or predominantly dissociated (right) or present in RNP cores (left) are indicated. **e**, Sedimentation of purified C complexes, C-complex RNP cores, and reassembled C complexes on glycerol gradients.

10% of input. **d**, Salt-treated C-complex proteins from gradient fractions 1–5 or 13–15, separated by SDS–PAGE and stained with silver. A subset of proteins exclusively or predominantly dissociated (right) or present in RNP cores (left) are indicated. **e**, Sedimentation of purified C complexes, C-complex RNP cores, and reassembled C complexes on glycerol gradients.

proteins (Fig. 4d). Proteins were separated by SDS-PAGE and analysed by MS (Supplementary Table 1). Visual comparison of the band patterns of the RNP peak with top gradient fractions reveals clear differences in their compositions. The peak contains predominantly Prp19 complex proteins and Prp19-related factors (for example, hPrp19, CDC5, hSyf3 SKIP, Prl1, RBM22, AD002 and G10) as well as the U5–220, U5–116 and U5–40K proteins, which form a highly stable heteromeric complex in higher eukaryotes¹⁹. Altogether, 37 proteins were predominantly or exclusively present in the salt-resistant RNP core complex as determined by band intensity and/or the number of peptides sequenced by MS (Supplementary Table 1), suggesting that they are tightly incorporated components of the C complex and are crucial for maintaining its catalytically active RNA network. Other proteins were found in apparently similar amounts in both fractions, whereas most, such as proteins of the exon junction complex and step II factors such as hPrp22 and hSlu7, or components of SF3a/b, were found exclusively or predominantly at the top of the gradient. These proteins therefore do not seem to be required for maintaining the RNP core of the C complex.

In contrast with C complexes isolated under physiological conditions, 1 M salt-treated C-complex cores did not support bimolecular exon ligation on their own (Fig. 2c, lanes 7 and 8). However, exon ligation was observed when the C-complex cores were supplemented with MN-digested nuclear extract, although at a lower efficiency than that observed with native C complexes (Fig. 2c, compare lanes 1 and 2 with lanes 5 and 6). Thus, the RNP core probably lacks one or more proteins required for exon ligation, but its RNA and RNP networks nonetheless remain relatively intact. C complexes incubated with MN-treated nuclear extract (reassembled C complexes) sedimented faster than salt-treated RNP cores on glycerol gradients (Fig. 4e), with an S value only slightly smaller than that of native C complexes, indicating that several proteins reassociated.

Concluding remarks

Here we have purified mammalian spliceosomes that are catalytically active on their own. This was a key step towards identifying proteins involved in stabilizing the spliceosome's catalytically active RNA network. We show that there is a large exchange of proteins during the B to C complex transition when the active site of the spliceosome is formed. Proteins comprising the heteromeric SF3a and SF3b complexes were largely underrepresented in purified C versus B complexes, suggesting that they are destabilized before or during the first catalytic step of splicing. Before the second step, the branch adenosine is probably removed from the active site of the spliceosome and replaced with the 3' splice site, which serves as the nucleophile for the second step of splicing. A recent report indicates that, in yeast, step II of splicing can proceed in the absence of U2–BPS base pairing²⁰. Thus, whereas SF3a and SF3b are essential early in the splicing reaction, they may not be required after the first catalytic step. SF3a and SF3b were for the most part absent from the salt-resistant RNP core of the C complex, raising the question of what stabilizes the association of U2 in this complex.

In contrast with SF3a/b, Prp19–CDC5 and related proteins seemed to be more stably associated in C complexes. Indeed, they, together with U5 proteins, remained associated with the pre-mRNA splicing intermediates and the U2, U5 and U6 snRNAs on treatment of purified C complexes with high salt concentrations. The presence of most human homologues of NTC proteins in the salt-stable C complex RNP core is consistent with studies in yeast demonstrating that the NTC is required for stable association of U5 and U6 with the pre-mRNA after U1 and U4 have dissociated¹². Our data additionally point to potential contributions by other Prp19-related proteins that are not part of the core Prp19–CDC5 complex, in stabilizing the catalytically active RNA structures in step I spliceosomes. The presence of hPrp8 in the salt-resistant C-complex core is also consistent with a wealth of data demonstrating that hPrp8 contacts all reactive groups of the pre-mRNA and supports the idea that it acts as a

cofactor during splicing catalysis²¹. The protein composition of the salt-stable C-complex core is highly similar to that of the human 35S U5 snRNP. This supports our previous hypothesis that the latter represents a remodelled form of U5 that is generated during spliceosome activation and is released after exon ligation as a post-spliceosomal complex²².

The ability to isolate active human C complexes and their stable RNP cores is an important step towards establishing the three-dimensional architecture of the spliceosome's RNP core. Our purified C complexes can be used to determine the structure of catalytically active spliceosomes poised for step II of splicing, either by electron microscopy or by structure probing methods. Salt-stable C-complex cores may be used in the future to identify proteins involved in catalytic step II, and also to study their interaction partners within the spliceosome. Similarly, as a result of their apparent high stability, they may ultimately be suitable for high-resolution structural studies, including X-ray crystallography.

METHODS SUMMARY

In vitro splicing. ³²P-labelled PM5 pre-mRNA was synthesized *in vitro*, and splicing was performed *in vitro* with HeLa nuclear extract²³. RNA was separated on an 8.0 M urea–10% polyacrylamide gel. Splicing complexes were analysed on agarose gels²⁴.

MS2 affinity selection and immunoblotting. MS2 affinity selection was performed as described previously¹⁸ except that an excess of DNA oligonucleotide complementary to nucleotides –6 to –18 relative to the 5' splice site was added after splicing and the reaction was incubated at 30 °C for a further 2 min (B complex) or 20 min (C complex) before glycerol-gradient centrifugation. Proteins were separated by SDS-PAGE, transferred to Hybond P membrane and immunostained with an enhanced chemiluminescence detection kit.

Mass spectrometry. Proteins were identified by liquid chromatography–tandem mass spectrometry (LC–MS/MS) after separation by SDS-PAGE as described previously¹⁸.

Bimolecular exon ligation. An excess of 3' substrate RNA was added after splicing for 180 min with PM5 substrate, and the mixture was incubated for 30 min. Alternatively, purified C complex or C-complex RNP cores were incubated for 30 min under standard splicing conditions with 3' substrate RNA. Generation of a ³²P-labelled DNA probe against the 3' substrate RNA and subsequent northern blotting were performed as described²⁵.

Isolation of C-complex RNP cores and reassembled C complexes. Purified C complexes were incubated in buffer containing 1 M NaCl and subjected to glycerol-gradient centrifugation in the presence of 1 M NaCl. For functional assays, C complexes bound to amylose beads were washed with buffer containing first 1 M NaCl, followed by 150 mM KCl (G-150). Complexes were then eluted with maltose (to generate C-complex RNP cores) or incubated with MN-treated nuclear extract and then eluted to yield reassembled C complexes. Complexes were subjected to glycerol-gradient centrifugation in the presence of G-150 buffer.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 11 December 2007; accepted 19 February 2008.

Published online 5 March 2008.

- Will, C. L. & Lührmann, R. in *The RNA World* (eds Gesteland, R. F., Cech, T. R. & Atkins, J. F.) 369–400 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2006).
- Staley, J. P. & Guthrie, C. Mechanical devices of the spliceosome: motors, clocks, springs, and things. *Cell* **92**, 315–326 (1998).
- Nilsen, T. W. in *RNA Structure and Function* (eds Grundber-Manago, M. & Simons, R. W.) 279–307 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1998).
- Umen, J. G. & Guthrie, C. The second catalytic step of pre-mRNA splicing. *RNA* **1**, 869–885 (1995).
- Konarska, M. M., Vilardeell, J. & Query, C. C. Repositioning of the reaction intermediate within the catalytic center of the spliceosome. *Mol. Cell* **21**, 543–553 (2006).
- Valadkhan, S. & Manley, J. L. Splicing-related catalysis by protein-free snRNAs. *Nature* **413**, 701–707 (2001).
- Will, C. L. & Lührmann, R. Protein functions in pre-mRNA splicing. *Curr. Opin. Cell Biol.* **9**, 320–328 (1997).
- Grainger, R. J. & Beggs, J. D. Prp8 protein: at the heart of the spliceosome. *RNA* **11**, 533–557 (2005).

9. Tarn, W. Y., Lee, K. R. & Cheng, S. C. Yeast precursor mRNA processing protein PRP19 associates with the spliceosome concomitant with or just after dissociation of U4 small nuclear RNA. *Proc. Natl Acad. Sci. USA* **90**, 10821–10825 (1993).
10. Tarn, W. Y. *et al.* Functional association of essential splicing factor(s) with PRP19 in a protein complex. *EMBO J.* **13**, 2421–2431 (1994).
11. Ajuh, P. *et al.* Functional analysis of the human CDC5L complex and identification of its components by mass spectrometry. *EMBO J.* **19**, 6569–6581 (2000).
12. Chan, S. P., Kao, D. I., Tsai, W. Y. & Cheng, S. C. The Prp19p-associated complex in spliceosome activation. *Science* **302**, 279–282 (2003).
13. Makarova, O. V. *et al.* A subset of human 35S U5 proteins, including Prp19, function prior to catalytic step 1 of splicing. *EMBO J.* **23**, 2381–2391 (2004).
14. Gozani, O., Feld, R. & Reed, R. Evidence that sequence-independent binding of highly conserved U2 snRNP proteins upstream of the branch site is required for assembly of spliceosomal complex A. *Genes Dev.* **10**, 233–243 (1996).
15. Will, C. L. *et al.* A novel U2 and U11/U12 snRNP protein that associates with the pre-mRNA branch site. *EMBO J.* **20**, 4536–4546 (2001).
16. Anderson, K. & Moore, M. J. Bimolecular exon ligation by the human spliceosome. *Science* **276**, 1712–1716 (1997).
17. Anderson, K. & Moore, M. J. Bimolecular exon ligation by the human spliceosome bypasses early 3' splice site AG recognition and requires NTP hydrolysis. *RNA* **6**, 16–25 (2000).
18. Deckert, J. *et al.* Protein composition and electron microscopy structure of affinity-purified human spliceosomal B complexes isolated under physiological conditions. *Mol. Cell. Biol.* **26**, 5528–5543 (2006).
19. Achsel, T., Ahrens, K., Brahms, H., Teigelkamp, S. & Lührmann, R. 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 (1998).
20. Smith, D. J., Query, C. C. & Konarska, M. M. Trans-splicing to spliceosomal U2 snRNA suggests disruption of branch site-U2 pairing during pre-mRNA splicing. *Mol. Cell* **26**, 883–890 (2007).
21. Collins, C. A. & Guthrie, C. The question remains: is the spliceosome a ribozyme? *Nature Struct. Biol.* **7**, 850–854 (2000).
22. Makarov, E. M. *et al.* Small nuclear ribonucleoprotein remodelling during catalytic activation of the spliceosome. *Science* **298**, 2205–2208 (2002).
23. Dignam, J. D., Lebovitz, R. M. & Roeder, R. G. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* **11**, 1475–1489 (1983).
24. Das, R. & Reed, R. Resolution of the mammalian E complex and the ATP-dependent spliceosomal complexes on native agarose mini-gels. *RNA* **5**, 1504–1508 (1999).
25. Hartmuth, K. *et al.* Protein composition of human prespliceosomes isolated by a tobramycin affinity-selection method. *Proc. Natl Acad. Sci. USA* **99**, 16719–16724 (2002).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank T. Conrad, P. Kemkes, H. Kohansal and I. Öchsner for technical assistance; M. Raabe and U. Plessmann for their help in MS analysis; E. Makarov for raising anti-peptide antibodies; and C. Smith for providing the PYP pre-mRNA. This work was supported by grants from the Deutsche Forschungsgemeinschaft, EURASNET, the Fonds der Chemischen Industrie and the Ernst Jung Stiftung to R.L., and a Young Investigator Programme grant from EURASNET to H.U.

Author Contributions S.B., M.A., H.U. and R.L. designed experiments; S.B., M.A. and H.U. performed the experiments; S.B., M.A., C. L.W., H.U. and R.L. analysed the data; and S.B., C. L.W., H.U. and R.L. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to R.L. (reinhard.luehrmann@mpi-bpc.mpg.de).

METHODS

In vitro splicing and native gel analysis. MS2-binding sites were introduced at the 5' end of the PYP pre-mRNA substrate²⁶ by standard PCR techniques, and uniformly ³²P-labelled, m7G(5')ppp(5')G-capped PM5 premRNA was synthesized *in vitro* by SP6 runoff transcription. Standard *in vitro* splicing reactions contained 50% (v/v) HeLa nuclear extract prepared as described previously²³, 65 mM KCl, 3 mM MgCl₂, 2 mM ATP, 20 mM creatine phosphate and 30 nM ³²P-labelled PM5 pre-mRNA, and were incubated for 0–180 min at 30 °C. RNA was recovered and separated on an 8.0 M urea–10% polyacrylamide gel. For analysis of spliceosome assembly, 0.5 μl of heparin (5 mg ml⁻¹) was added to 20-μl aliquots of the splicing reaction at a given time, and splicing complexes were separated on a 2% native agarose gel²⁴.

MS2 affinity selection of spliceosomal B and C complexes. PM5 pre-mRNA was incubated with a 20-fold molar excess of purified MS2-MBP fusion protein¹⁸ and a 12-ml standard splicing reaction containing 10 nM (B complex) or 30 nM (C complex) ³²P-labelled pre-mRNA was incubated for 6 min (B complex) or 180 min (C complex). A 30-fold molar excess of DNA oligonucleotide complementary to nucleotides -6 to -18 relative to the 5' splice site was added and the reaction was incubated at 30 °C for a further 2 min (B complex) or 20 min (C complex). Complexes were separated on 10–30% glycerol gradients in the presence of 150 mM KCl, and 40–45S gradient fractions were subjected to affinity selection on amylose beads as described previously¹⁸. RNA was recovered from the eluted B or C complexes, analysed on a 8.0 M urea–10% polyacrylamide gel, and detected by silver staining and autoradiography. The stoichiometry of the snRNAs and pre-mRNA/splicing intermediates was determined by visual inspection of the silver-stained bands, taking into consideration the length of the individual RNAs. Protein was recovered, separated by 8%/14% step-gel SDS-PAGE and stained with Coomassie blue.

Immunoblotting. Proteins were separated by SDS-PAGE on an 8% or 12% polyacrylamide gel, transferred to Hybond P membrane and immunostained with an enhanced chemiluminescence detection kit (Pierce). Antibodies against the following proteins were used: MS2-MBP (Abcam); U5-116K (ref. 13), Lsm4 (ref. 13), hPrp19 (ref. 13), AD002 (ref. 13), KIAA0560 (ref. 13), SKIP (ref. 13), SF3a66 (ref. 15), SF3b155 (ref. 15), SF3b145 (ref. 27), SF3b130 (ref. 27), SF3b49 (ref. 27) and SF3a60 (ref. 27); RBM22 (raised against residues 179–195), hPrp22 (residues 1116–1130), hSlu7 (residues 562–577), hPrp17 (residues 92–107), hPrp16 (residues 193–207), DDX35 (residues 392–409), Abstrakt (residues 608–622), Cyp-E (residues 104–118) and Prl1 (residues 60–74).

Mass spectrometry. For MS, proteins recovered from purified B and C complexes (about 4 pmol) were separated by SDS-PAGE and stained with Coomassie blue. Entire lanes were cut into 60–70 slices and proteins were digested *in-gel* with trypsin and extracted as described previously¹⁸. For MS of salt-treated core complexes, proteins from the top gradient fractions or RNP peak fractions were separated by SDS-PAGE and stained with silver. For direct comparison, single bands and gel regions of the same height in both lanes of the gel were excised and the proteins therein were digested with trypsin and extracted. Peptides were analysed in an LC-coupled electrospray ionization quadrupole time-of-flight

mass spectrometer (QToF Ultima; Waters) and/or a linear ion-trap mass spectrometer (4000 QTrap; Applied Biosystems) under standard conditions. Proteins were identified by searching fragment spectra against the National Centre for Biotechnology Information non-redundant database, using Mascot as search engine. Mass accuracy was 200 p.p.m. for the parent ion (MS) and 200 milli-mass units (mmu) for the fragment ions (MS/MS) when samples were analysed by Q-ToF MS, and 1.4 Da and 400 mmu, respectively, for samples analysed on the linear ion trap. Peptides were constrained to be tryptic with a maximum of one missed cleavage site. Carbamidomethylation of cysteine residues and oxidation of methionine residues were considered as variable modifications. The highest-scoring peptide from each protein, as well as single-hit peptides, were inspected manually to eliminate any false positives in the data set. The overall numbers of sequenced peptides were counted and only fragment spectra with an ion score above 30 (Mascot) were included.

Bimolecular exon ligation. 3' substrate RNAs used in bimolecular exon ligation assays were generated by *in vitro* SP6 runoff transcription. To assay for bimolecular exon ligation in nuclear extract, 10 pmol of 3' substrate RNA were added after splicing for 180 min with 1.5 pmol of PM5 substrate (at this time point about 0.15 pmol of excised 5' exon are formed), and the reaction (50 μl) was incubated for a further 30 min. Alternatively, 0.3 pmol of either purified C complex or C-complex core RNP were incubated for 30 min under standard splicing conditions (see above) with 10 pmol of 3' substrate RNA. MN-treated nuclear extract (20% of total reaction) was added as indicated. RNA was recovered and separated on an 8.0 M urea–10% polyacrylamide gel and transferred to Hybond XL membrane. Northern blotting and the generation of a ³²P-labelled DNA probe against the 3' substrate RNA were performed as described previously²⁵.

Isolation of salt-stable C-complex RNP cores and reassembled C complexes. Purified C complexes were incubated on ice for 15 min in G buffer (20 mM HEPES pH 7.9, 1.5 mM MgCl₂) containing 1 M NaCl and then subjected to 10–30% glycerol-gradient centrifugation (150 min at 490,000g; Sorvall TH660 rotor) in the presence of 1 M NaCl. The distribution of ³²P-labelled RNA across the gradient was determined by Cherenkov counting in a scintillation counter. RNA and protein were analysed as described above. For functional assays, C complexes bound to amylose beads were washed with 20 column volumes of G buffer containing 1 M NaCl at 4 °C, followed by G buffer containing 150 mM KCl (G-150). Complexes were then either eluted with maltose (to generate C-complex RNP cores) or were first incubated at 4 °C for 30 min with MN-treated nuclear extract in G buffer containing 125 mM KCl and 0.1 mM EDTA, washed with 75 volumes of G-150 buffer, and then eluted (to yield reassembled C complexes). To compare their S values, complexes were subjected to 10–30% glycerol-gradient centrifugation (90 min at 490,000g; TH660 rotor) in the presence of G-150 buffer.

26. Wollerton, M. C., Gooding, C., Wagner, E. J., Garcia-Blanco, M. A. & Smith, C. W. Autoregulation of polypyrimidine tract binding protein by alternative splicing leading to nonsense-mediated decay. *Mol. Cell* **13**, 91–100 (2006).
27. Dybkov, O. *et al.* U2 snRNA-protein contacts in purified human 17S U2 snRNPs and in spliceosomal A and B complexes. *Mol. Cell. Biol.* **26**, 2803–2816 (2006).