MOLECULAR BIOLOGY

Structure of the yeast spliceosomal postcatalytic P complex

Shiheng Liu,^{1,2*} Xueni Li,^{3*} Lingdi Zhang,³ Jiansen Jiang,^{1,2} Ryan C. Hill,³ Yanxiang Cui,¹ Kirk C. Hansen,³ Z. Hong Zhou,^{1,2}† Rui Zhao³†

The spliceosome undergoes dramatic changes in a splicing cycle. Structures of B, B^{act}, C, C*, and intron lariat spliceosome complexes revealed mechanisms of 5'-splice site (ss) recognition, branching, and intron release, but lacked information on 3'-ss recognition, exon ligation, and exon release. Here we report a cryo-electron microscopy structure of the postcatalytic P complex at 3.3-angstrom resolution, revealing that the 3' ss is mainly recognized through non-Watson-Crick base pairing with the 5' ss and branch point. Furthermore, one or more unidentified proteins become stably associated with the P complex, securing the 3' exon and potentially regulating activity of the helicase Prp22. Prp22 binds nucleotides 15 to 21 in the 3' exon, enabling it to pull the intron-exon or ligated exons in a 3' to 5' direction to achieve 3'-ss proofreading or exon release, respectively.

ntrons are spliced out through two sequential transesterification steps catalyzed by the spliceosome, a huge RNA-protein complex made of five small nuclear RNAs (snRNAs) and over 100 proteins. Spliceosomal remodeling by adenosine triphosphate (ATP)-dependent DExD/H box helicases during each splicing cycle generates at least seven distinct states known as the B (1), B^{act} (2), B*, C (3, 4), C* (5, 6), P, and intron lariat spliceosome (ILS) (7) complexes. After the first transesterification reaction (also referred to as the branching reaction) catalyzed by B*, the C complex emerges, encompassing the newly freed 3'-OH group of the 5' exon and a lariat intermediate. The C complex further rearranges to form the C* complex, primed for the second transesterification reaction (also referred to as the ligation reaction). In the ligation reaction, the two exons join, forming the postcatalytic P complex that contains the ligated exons and the lariat. The exons are then released through the action of the DEAH-box helicase Prp22, generating the ILS complex, which only contains the lariat.

High-resolution cryo–electron microscopy (cryo-EM) structures of almost all of the above complexes have been determined (*1–7*), leaving two missing pieces of the puzzle—atomic descriptions of the B* and P complexes. Several important questions all related to the P complex—remain to be answered to fully understand the molecular mechanism of splicing, including how the 3' splice site (ss) is recognized and docked to the proximity of the 5' exon, how exon ligation is catalyzed, and how Prp22 promotes 3'-ss proofreading and exon release.

Here we report a crvo-EM structure of the Saccharomyces cerevisiae (yeast) P complex at 3.3-Å resolution. Our atomic model of the P complex reveals that the 3'-ss recognition is driven by its interaction with the 5' ss and branch point (BP), likely facilitated by a stemlike structure formed between the intronic regions close to the branch site and 3' ss. Our structure reveals that one or more unidentified proteins become stably associated with the core components of the P complex around splicing factor Prp8 and Prp22, securing the 3' exon and potentially regulating the activity of Prp22. The structure demonstrates that Prp22 binds nucleotides 15 to 21 in the 3' exon, enabling it to pull the intron-exon spanning the 3' ss or the ligated exons in a 3' to 5' direction to fulfill its function in 3'-ss proofreading or exon release, respectively. By providing insights into the molecular mechanism of 3'-ss recognition, exon ligation, and the action of RNA helicase Prp22, the atomic model of the P complex fills a major gap in our understanding of the splicing cycle.

Results Overall structure

We purified the P complex using a protein A tag on splicing factor Cef1 and a calmodulin-binding peptide (CBP) tag on Prp22 from yeast carrying the Prp22^{H606A} mutant (H606A, alanine residue substituted for histidine at position 606), a mutant in the DEAH box that blocks exon release after ligation at nonpermissive temperature (8). We observed a substantial accumulation of ligated ACTI (actin) exons in spliceosome purified from the Prp22^{H606A} strain under nonpermissive temperature compared to that in the wild-type (WT) strain, whereas intron lariats were observed at similar amounts in both strains (fig. S1A). We determined the cryo-EM structure of the P complex to an average resolution of 3.3 Å on the basis of the "gold-standard" Fourier shell correlation 0.143 cutoff criterion (9, 10). We built models for ligated exons [21 nucleotides (nt) downstream of the exon-exon junction] and interacting splicing factors, including Prp22, Prp8, Slu7, and Prp18 (figs. S1 to S4 and table S1). We also modeled U2, U5, and U6 snRNAs, the intron lariat, and other proteins that are nearly identical to those in the C* complex structures (5, 6), though with clearly better-resolved density features. In addition, there are well-defined densities for multiple helices around Prp22 and Prp8, although the identity of their corresponding protein(s) cannot be established based on possible sequences, known spliceosomal structures, and Prp22- or Prp8-interacting partners. We will refer to this unidentified protein as UNK, noting that these helices can belong to multiple proteins or be connected to a known protein in the P complex through a flexible linker.

The overall architecture of the P complex is similar to the structure of the C^* complex (5, 6) (Fig. 1 and movie S1), but with several major differences. Of the two near-atomic-resolution structures of the yeast C^* complex, one (6) was assembled using a precursor mRNA (pre-mRNA) substrate with the 2'-OH group of the G nucleotide at the 3' ss removed, which inhibits exon ligation (11), whereas the other (5) was based on computational classification of WT spliceosome complexes. We therefore will make most of our detailed structural comparisons with the former structure. First, the most notable difference between the two complexes is that the 3' intron-exon (see below for definition) is not visible in the C* complex, whereas exons are clearly ligated and the 3' intron (see below for definition) remains docked in the active site in the P complex (Fig. 2, A and B). To facilitate our description, we refer to the intron region upstream of the BP as the 5' intron, the region downstream of BP as the 3' intron, and the 3' intron covalently linked (at the 3' ss) to the 3' exon as the 3' intronexon. Second, the 1585 loop of Prp8 [residues 1576 to 1599, also referred to as the α finger] (12-14) and the C-terminal tail of Prp22 (residues 1106 to 1145) become ordered in the P complex and interact with both the 3' intron and 3' exon. Third, several regions in the Slu7-Prp18 heterodimer (residues 30 to 165 and 241 to 292 in Slu7 and residues 189 to 221 in Prp18) are ordered in the P complex, but absent in the C* complex structure. Finally, an unknown protein, or proteins, made of mostly helices interacts with Prp22, Prp8, and nineteen complex (NTC) components Cef1 and Syf1 in the P complex.

RNA components in the active site

Comparison of the RNA components in the active site of the C* and P complexes reveals how the 3' intron-exon is loaded. In the C* complex, the 3' intron-exon spanning the 3' ss is not docked to the active site yet and is, in fact, entirely missing in the structure. The region that would accommodate the presumably disordered 3' intron-exon is exposed, consistent with biochemical observations that the 3' intron-exon is susceptible to ribonuclease H (RNase H) cleavage after DEAHbox RNA helicase Prp16-mediated conformational changes but before exon ligation (i.e., the C* complex) (15). In the P complex, the 3' intron-exon is

¹Electron Imaging Center for Nanomachines, University of California, Los Angeles (UCLA), Los Angeles, CA 90095, USA.
²Department of Microbiology, Immunology, and Molecular Genetics, UCLA, Los Angeles, CA 90095, USA.
³Department of Biochemistry and Molecular Genetics, University of Colorado Denver (UCD), Anschutz Medical Campus, Aurora, CO 80045, USA.

^{*}These authors contributed equally to this work.

⁺Corresponding author. Email: rui.zhao@ucdenver.edu (R.Z.); hong.zhou@ucla.edu (Z.H.Z.)

docked into the active site through the recognition of both the 3' ss and 3' exon (Fig. 2, A and B).

The 3' ss is recognized through two main structural features. The first is the interaction between the 3' ss and 5' ss (enlarged view in Fig. 2B). The 3' ss is in close proximity to the 2'-5' linkage formed between the 5' ss and the BP A nucleotide in the P complex. The last nucleotide in the intron (the underlined G in the UAG sequence of 3' ss) forms a G-G non-Watson-Crick base pair with the first nucleotide in the intron (GU in the 5' ss), consistent with previous biochemical and genetic analyses that indicate the existence of an essential non-Watson-Crick interaction between the first and last nucleotide of the intron (16). The second to last nucleotide in the intron (UAG in 3' ss) forms an A-A non-Watson-Crick base pair with the BP A nucleotide, as well as a π - π stacking with the 5'-ss G nucleotide. The third to last nucleotide in the intron (UAG in the 3' ss) forms a π - π stacking with the BP A nucleotide (enlarged view in Fig. 2B). In addition, the last nucleotide in the intron (UAG in the 3' ss) interacts with nucleotides A51, C58, and A59 of U6 snRNA. Furthermore, the last several nucleotides in the intron interact extensively with the 1585 loop of Prp8, which likely facilitates the recognition of the 3' ss. For example, nucleotides U (-2) and A (-1) (UAG in the 3' ss) both form hydrogen bonds with $Prp8 \ Gln^{1594}$. The fourth to the last nucleotide in the intron also forms a π - π interaction with Phe¹⁵⁸¹ of Prp8.

A second structural feature that could facilitate 3'-ss recognition is a stemlike structure formed close to the BP and 3' ss (Fig. 2C). The density of the stem is relatively weak and not defined enough for individual-base assignment, potentially because multiple introns with different sequences form a similar stemlike structure around this region. An examination of yeast intron sequences indicates that many of them could form stemlike secondary structures close to the BP and 3' ss (fig. S5). These stemlike structures could bring the 3' ss close to the BP and the active site, facilitating the recognition of the 3' ss through interactions with the 5' ss.

The 3' exon is recognized through interactions with U5 snRNA. In the C* structure, the 5' exon forms Watson-Crick base pairs and additional hydrogen bonds with loop 1 (nucleotides 96 to 100) of U5 snRNA. In our structure of the P complex, the 3' exon mainly interacts with U5 snRNA through π - π stacking between the first nucleotide in the 3' exon and nucleotide U96 in U5 snRNA, consistent with previous studies that suggest that U5 loop 1 interacts with both the 5' and 3' exons and facilitates the alignment of both exons (17, 18).

In the P complex, the 3' ss is close to the ligated exons (Fig. 2D) even though the second transesterification reaction has already occurred, allowing us to envision how exons are ligated (Fig. 2E). The last nucleotide of the 5' exon is in close proximity to the 3' ss. The catalytic Mg²⁺ (M1) ion is coordinated by G78 and U80 of U6 snRNA as well as the first nucleotide of the 3'

exon and is in close proximity to the 3'-OH at the end of the 5' exon. M1 is in a perfect position to activate the 3'-OH group for nucleophilic attack on the 3' ss, ligating the two exons and cleaving off the 3' intron, as suggested by biochemical analyses (19, 20). Another Mg^{2+} ion, M2, that is proposed to stabilize the leaving group in exon ligation (19) is absent in the P complex, which could be reflective of a postcatalytic state.

Protein UNK, Prp8, Prp22, and Slu7-Prp18 jointly stabilize the 3' intron and 3' exon

The major differences in protein components (UNK, Prp22, Prp8, and the Slu7-Prp18 heterodimer) between the C* and P complex structures are found around the 3' intron and 3' exon. Protein UNK is composed mostly of helices and is located near Prp22 and Prp8 (Fig. 3). Helix A in protein UNK is positioned like a "door bolt" between the linker domain of Prp8 and RecA2 domain of Prp22, locking the 3' exon in place (Fig. 3). Helices B and C in UNK start from the RecA2 domain of Prp22, contact the RNase H domain, go under the stemlike region of the 3' intron, and end around the reverse transcriptase-like (RT) domain of Prp8, securing the 3' intron (Fig. 3) [helices C and D of UNK are present in the C* complex, but with less-well-defined densities, and were not connected (6)]. Therefore, protein UNK likely plays a major role in ensuring the correct positioning of the 3' intron and 3' exon, potentially for both exon ligation and release.

Prp22 wraps around the exon and contacts the intron together with Prp8 (Fig. 3). The 1585 loop of Prp8 and the C-terminal tail of Prp22 both become ordered in the P complex. The 1585 loop of Prp8 interacts with U2 snRNA, the 3' exon, the 3' intron, and the C-terminal tail of Prp22. The C-terminal end of Prp22 inserts into the cavity formed by the RT domain of Prp8, the 1585 loop of Prp8, U2 snRNA, the 3' intron, and the 3' exon, and interacts with all of these components. Both the 1585 loop of Prp8 and the C-terminal tail of Prp22 are likely important for stabilizing the conformation of the 3' intron and 3' exon for exon ligation and/or release. As a consequence, the



Fig. 1. Overall structure of the P complex. (A to **C**) Three different views of the P complex, with each subunit colored according to subunit identity. Sm, Sm ribonucleoproteins. The view orientation in (C) is the same as those used in (5) to facilitate comparison with the structures of other spliceosomal complexes. The inset in (C) depicts the RNA elements only and is used as a landmark to orient subsequent figures. (**D**) List of modeled subunits in different functional subcomplexes. NTC, nineteen complex; NTR, NTC-related complex.



Fig. 2. RNA components in the P complex. (A) Pre-mRNA (ligated exons and intron lariat), U2, U5, and U6 snRNAs in the P complex, with each RNA colored differently. Gray represents pre-mRNA, U2, U5, and U6 snRNAs in the C* complex, noting that the 3' intron and 3' exon of pre-mRNA are not present in the C* complex. The inset shows the orientation of the figure. **(B)** A different view of the RNA components in (A). The enlarged view

shows the interaction between the 3' ss and the 5' ss with non–Watson-Crick base pairing and π - π stacking. (**C**) The intron forms a stemlike structure that brings the 3' ss close to the 5' ss. This panel is in the same orientation as (A). (**D**) The structure of the active site in the P complex. (**E**) A model of the active site right before exon ligation, illustrating the mechanism of exon ligation.

3' exon in the P complex makes extensive interactions with Prp8 and extends all the way to the core of Prp22, interacting with its C-terminal domain (CTD) and the two RecA domains. This is consistent with biochemical studies demonstrating that the 3' exon is completely accessible to RNase H digestion in the C* complex; however, more than 20 nt of the 3' exon are protected immediately after exon ligation, but before exon release (15).

Several regions of Prp18-Slu7 (residues 189 to 221 in Prp18 as well as residues 30 to 165 and 241 to 292 in Slu7) become ordered in the P complex (Fig. 3). Loop 189 to 221 in Prp18 inserts deeply into the N domain and RNase H domain of Prp8 and interacts with the nucleotides close to the 3' ss. Slu7 directly interacts with the 3' intron through residues Arg²⁸⁷ and Arg²⁸⁹, consistent with cross-linking observed between Slu7 and the 3' ss (21). In addition, the region from residues 30 to 165 and 241 to 292 in Slu7 wraps around multiple domains of Prp8 (including the linker domain where the 1585 loop resides) (Fig. 3), po-

tentially helping the 1585 loop become ordered in the P complex. The combined effect of protein UNK, the 1585 loop of Prp8, the C-terminal tail of Prp22, as well as the loops in Prp18 and Slu7, help secure the 3' intron and 3' exon in the right position for nucleophilic attack of the 3' ss by the 3'-OH group at the end of the 5' exon.

The structural basis of Prp22-mediated exon release and 3'-ss proofreading

Prp22 plays at least two distinct roles in the splicing cycle. The first is to release ligated exons in an ATPase- and helicase-dependent manner (22). The second is to proofread the 3' ss for exon ligation, allowing for alternative 3' ss choice through an ATP-dependent mechanism (23, 24). There have been some discrepancies about whether Prp22 plays a third, ATP-independent role in facilitating the ligation reaction (22, 25).

Our structure illustrates how Prp22 participates in exon release. In the P complex, Prp22 attaches to the spliceosome at the periphery through interaction with Prp8, protein UNK, and the 3' exon (Figs. 1 and 4A). Prp22 binds to nucleotides +15 to +21 in the 3' exon [upstream (+) with respect to the exon-exon junction] through its two RecA domains and the CTD. It can perceivably use a winching motion to pull the exon out of the spliceosome, consistent with previous biochemical observations (24). This also explains why at least 18 nt are required for efficient exon release, even though only 3 nt in the 3' exon are required for exon ligation (15). The interaction between Prp22 and the 3' exon is likely important for Prp22 to stay on the spliceosome, so that when the exon is pulled out, Prp22 consequently falls off the spliceosome. Therefore, Prp22 binds to single-stranded RNA, and RNA translocation, rather than unwinding, drives the release of the ligated exons, consistent with the observation that exon release by Prp22 is a less-demanding requirement than unwinding (8). The ligated exons are mostly surrounded by the N, RT, thumb, and linker domains of Prp8, with additional interactions with helix A of protein UNK, the C-terminal tail of Prp22, and a loop



Fig. 3. Protein UNK, Prp8, Prp22, and Slu7-Prp18 jointly stabilize the 3' intron and 3' exon in the P complex. The bulk of the proteins are shown in surface models. Four helices (green, labeled A to D) in protein UNK and regions in each of the other proteins that are present in the P complex structure, but absent in the C* complex structure, are shown in ribbons, including the 1585 loop of Prp8, the C-terminal tail (designated as C-tail) of Prp22, the residue 189-to-221 loop of Prp18, and residues 30 to 165 and 241 to 292 of Slu7. The enlarged view better illustrates the relationship among the 1585 loop of Prp8, the C-tail of Prp22, the residue 189-to-221 loop of Prp18, the 3' intron, and the 3' exon.

region from Prp18. The Prp8 domains are highly modular and are connected by linkers, which presumably allow the domains to loosen up and let the exon pass.

The structures of Prp22 bound to the 3' exon in the P and C* complexes also shed light on how Prp22 proofreads the 3' ss. Although the 3' intronexon is not visible in the C* complex, there is density for 3 nt that is likely part of the 3' exon in the center of Prp22 (6). This suggests that Prp22 could pull the 3' intron-exon after the 3' intron-exon is docked close to the end of the 5' exon, but before exon ligation. This pulling destabilizes the 3' ss and spliceosome interaction, which competes against the exon ligation event, and provides an opportunity to discard suboptimal 3' ss or to select alternative 3' ss (24). The branch formation at the intron lariat, the interaction between the 5' ss and U6 snRNA, and the interaction between the BP sequence and U2 snRNA, together with the proteins that stabilize these interactions, potentially make it hard for Prp22 to completely pull out the 3' intron-exon. In addition, our density of the CTD of Prp22 is much-better defined than its density in the C^* complex (6),



Fig. 4. Structure and function of Prp22. (**A**) Prp22 attaches to the spliceosome through its interaction with Prp8, protein UNK, and the 3' exon. (**B**) Prp22 has a similar structure as Prp43; both are DEAH-box helicases involved in splicing. (**C**) The structure of Prp22 resembles DNA helicase Hel308 (left). The ratchet helix of Hel308 superimposes well with the helix composed of residues 988 to 1006 of Prp22 (right), and multiple residues (labeled) on this helix may be important for the translocation activity of Prp22 that is critical for exon release and 3'-ss proofreading. R, arginine; F, phenylalanine; K, lysine.

suggesting that Prp22 is more stably associated with the core components in the P complex than with those in the C* complex. Indeed, helices A and B (absent in the C* complex) in protein UNK both serve to fasten Prp22 onto the P complex by binding to Prp22 at one end of the helices and binding to Prp8 at the other (Fig. 3). An additional short helix-loop-helix from protein UNK binds to the oligonucleotide-binding (OB) domain within the CTD of Prp22. Protein UNK may stimulate the helicase activity of Prp22, just as two G-patch proteins (Spp2 and Ntr1) stimulate the helicase activity of Prp2 and Prp43, presumably by binding to their OB domains (*26–30*).

The well-defined density of Prp22 in our P complex structure allowed us to compare Prp22 with other helicases. The structure of Prp22 is similar to that of Prp43 (Fig. 4B), both of which resemble DNA helicase Hel308 (*31, 32*) (Fig. 4C). Nucleotide strands in all three structures go into the same channel formed by the two RecA domains and the CTD. The crystal structure of Hel308 in complex with a partially unwound DNA duplex and biochemical analyses revealed a ratchet helix (residues 586 to 611) critical for DNA translocation (32) (Fig. 4C). Residues Arg⁵⁹² and Trp⁵⁹⁹ on the ratchet helix interact with single-stranded DNA, whereas the N terminus of the ratchet helix interacts with motif IVa of RecA2. ATP-dependent movement of RecA2 is therefore thought to modulate the position of the ratchet helix and facilitate strand translocation. The residue 987-to-1006 helix in Prp22 is well superimposed with the ratchet helix in Hel308 (Fig. 4C). Prp22 Arg⁹⁹² is equivalent to Hel308 Arg⁵⁹². Although there is no tryptophan equivalent to Hel308 Trp⁵⁹⁹, Prp22 Phe^{1004} may form a comparable π - π stacking with the RNA strand. There are several other basic residues (Prp22 Arg^{988} , Lys¹⁰⁰⁵, and Lys¹⁰⁰⁶) on the ratchet helix. The residue 987-to-1006 helix and multiple residues on this helix (Prp22 Arg⁹⁹², Arg⁹⁸⁸, Phe¹⁰⁰⁴, Lys¹⁰⁰⁵, and Lys¹⁰⁰⁶) are likely important for Prp22-mediated exon release and possibly 3'-ss proofreading.

Discussion

Our structure of the P complex reveals that the 3' ss is recognized mainly through non-Watson-Crick base pairing between the 3' ss and the 5' ss, as well as between the 3' ss and the BP (Fig. 2B),



Fig. 5. A schematic representation of the splicing cycle from the C* complex to the P complex stage based on our structure and the structures of the C* complex. In the C* complex (left), the branching reaction has occurred, but the 3' intron-exon has not been loaded to the active site. In the spliceosomal complex immediately before ligation (middle), protein UNK and several other proteins help load the 3' intron-

exon into the active site, where the 3' ss is recognized through interactions with the 5' ss facilitated by the stemlike structure formed in the intron. Prp22 pulls the 3' intron-exon to fulfill its function in 3'-ss proofreading. In the P complex (right), the two exons are ligated, but the intron lariat remains in the complex. Prp22 pulls the ligated exons to release them from the spliceosome.

which is further stabilized by their interactions with surrounding proteins. Although Watson-Crick base pairs are crucial for RNA structures, non-Watson-Crick base pairs are often involved in protein-RNA or RNA-RNA interactions (*33, 34*). For example, non-Watson-Crick base pairs are used for ss recognition in self-splicing group I introns (*35*). The 5' ss (GU) and 3' ss (AG) are conserved among all species, testifying to the importance of non-Watson-Crick base pairing in 3'-ss recognition.

Conversely, what brings the 3' ss close to the 5' ss and BP to form these non-Watson-Crick base pairs has remained elusive. It is usually the first YAG sequence (where Y is any pyrimidine) after the BP that is recognized as the 3' ss, leading to the scanning hypothesis, which assumes that the spliceosome reads the intron from the BP in a 5'-to-3' direction until it encounters the first YAG. However, the only moving part in the preligation C* complex is the pulling of the 3' exon by Prp22 in a 3'-to-5' direction. Therefore, the YAG further away from the BP (instead of the first YAG) would have been encountered by the spliceosome active site first. Instead, the stemlike structure we observed close to the BP and 3' ss (Fig. 2C) suggests that secondary structures formed within the 3' intron play an important role in bringing the 3' ss close to the active site for base pairing with the 5' ss. This is consistent with previous bioinformatics and experimental observations (36-38). The presence and stability of these secondary structures can potentially be a way to modulate splicing efficiency at the 3' ss. The observation of the stemlike structure also implies that intronic secondary structures may play a much more prevalent role in 3' ss recognition and alternative splicing in higher eukaryotes than has been previously appreciated.

The Prp22 structure bound to the 3' exon in our structure (made possible by the high-quality map in this region), together with the C* complex structure, provide a vivid image of how Prp22 fulfills its function in exon release and 3'-ss proofreading (Fig. 5). Prp22 binds to nucleotides +15 to +21 (with respect to the exon-exon junction, Fig. 4A), enabling it to pull on the 3' intron-exon (in C*) or ligated exons (in P) through a winching motion from this remote location, achieving 3'-ss proofreading or exon release activity, respectively, in agreement with previous biochemical data (24).

Our structure reveals a yet to be identified protein, UNK, mostly made of helices that is likely important for the function of Prp22 (Fig. 5). None of the known second-step factors present in our sample fit the density. The density suggests that UNK could be composed of multiple proteins that await further experimental confirmation, including the N-terminal domain of Prp22 and FYV6, a protein with unknown function. UNK provides a door bolt between Prp22 and Prp8 to lock the 3' exon in position (Fig. 3). It also contacts the OB domain of Prp22 and potentially stimulates Prp22 activity similar to G-patch protein stimulation, offering a possible way to regulate Prp22 activity so that it efficiently pulls out the ligated exons, but not the 3' intron-exon. With the atomic models of nearly all other major complexes already in the literature, the P complex structure reported here highlights the exciting prospect that a full mechanistic understanding of the entire splicing cycle may just be around the corner.

REFERENCES AND NOTES

- 1. C. Plaschka, P. C. Lin, K. Nagai, Nature 546, 617–621 (2017).
- C. Yan, R. Wan, R. Bai, G. Huang, Y. Shi, Science 353, 904–911 (2016).
- R. Wan, C. Yan, R. Bai, G. Huang, Y. Shi, Science 353, 895–904 (2016).
- 4. W. P. Galej et al., Nature 537, 197-201 (2016).
- C. Yan, R. Wan, R. Bai, G. Huang, Y. Shi, Science 355, 149–155 (2017).
- 6. S. M. Fica et al., Nature 542, 377-380 (2017).
- 7. R. Wan, C. Yan, R. Bai, J. Lei, Y. Shi, Cell 171, 120-132.e112 (2017).
- 8. E. Campodonico, B. Schwer, Genetics 160, 407-415 (2002).
- 9. P. B. Rosenthal, R. Henderson, J. Mol. Biol. 333, 721-745 (2003).
- 10. S. H. Scheres, J. Struct. Biol. 180, 519-530 (2012).

- 11. M. J. Moore, P. A. Sharp, Science 256, 992–997 (1992).
- T. H. D. Nguyen et al., Nature 530, 298–302 (2016).
 W. P. Galej, T. H. Nguyen, A. J. Newman, K. Nagai, Curr. Opin. Struct. Biol. 25, 57–66 (2014).
- 14. Y. Shi, *J. Mol. Biol.* **429**, 2640–2653 (2017).
- 15. B. Schwer, Mol. Cell 30, 743-754 (2008).
- 16. R. Parker, P. G. Siliciano, Nature 361, 660-662 (1993).
- 17. A. J. Newman, C. Norman, Cell 68, 743-754 (1992).
- 18. E. J. Sontheimer, J. A. Steitz, Science 262, 1989-1996 (1993).
- 19. S. M. Fica et al., Nature 503, 229-234 (2013).
- S.-L. Yean, G. Wuenschell, J. Termini, R.-J. Lin, *Nature* 408, 881–884 (2000).
- 21. J. G. Umen, C. Guthrie, RNA 1, 584-597 (1995).
- 22. B. Schwer, C. H. Gross, EMBO J. 17, 2086-2094 (1998).
- R. M. Mayas, H. Maita, J. P. Staley, *Nat. Struct. Mol. Biol.* 13, 482–490 (2006).
- D. R. Semiow, M. R. Blanco, N. G. Walter, J. P. Staley, *Cell* 164, 985–998 (2016).
- 25. T. Ohrt et al., RNA 19, 902-915 (2013)
- 26. E. J. Silverman *et al.*, *Mol. Cell. Biol.* **24**, 10101–10110 (2004). 27. J. B. Fourmann, M. J. Tauchert, R. Ficner, P. Fabrizio,
- R. Lührmann, *Nucleic Acids Res.* **45**, 4068–4080 (2017). 28. Z. Warkocki *et al.*, *Genes Dev.* **29**, 94–107 (2015).
- Walkock et al., Genes Dev. 23, 54–107 (2013).
 N. Tanaka, A. Aronova, B. Schwer, Genes Dev. 21, 2312–2325
- (2007). 30. H. Walbott *et al.*, *EMBO J.* **29**, 2194–2204 (2010).
- 31. Y. He, J. P. Staley, G. R. Andersen, K. H. Nielsen, *RNA* 23,
- 1110–1124 (2017). 32. K. Büttner, S. Nehring, K. P. Hopfner, *Nat. Struct. Mol. Biol.* **14**,
- 647–652 (2007). 33. N. B. Leontis, J. Stombaugh, E. Westhof, *Nucleic Acids Res.* **30**,
- N. B. Leonius, J. Stornbaugh, E. Westhol, Nucleic Acids Res. 30, 3497–3531 (2002).
 T. Hermann, E. Westhof, Chem. Biol. 6, R335–R343 (1999).
- 34. 1. Hermann, E. Westnor, *Chem. Biol.* **6**, R335–R343 (1999) 35. F. Michel, M. Hanna, R. Green, D. P. Bartel, J. W. Szostak,
- Nature **342**, 391–395 (1989).
- O. Gahura, C. Hammann, A. Valentová, F. Půta, P. Folk, Nucleic Acids Res. 39, 9759–9767 (2011)
- M. Plass, C. Codony-Servat, P. G. Ferreira, J. Vilardell, E. Eyras, RNA 18, 1103–1115 (2012).
- M. Meyer, M. Plass, J. Pérez-Valle, E. Eyras, J. Vilardell, Mol. Cell 43, 1033–1039 (2011).

ACKNOWLEDGMENTS

We thank B. Schwer for the p358-PRP22 and p358-H606A plasmids, as well as yeast strain YBSTI. This work was supported by NIH grants GMI14178 (R.Z.) and GMO71940 and Al094386 (Z.H.Z.). We acknowledge the use of instruments at the CryoEM Structural Biology Shared Resource Facility at UCD and at the Electron Imaging Center for Nanomachines [supported by UCLA and grants from the NIH (1S100D018111 and 1U24GM116792) and NSF (DBI-1338135 and DMR-1548924)]. Molecular graphics and analyses were performed with the University of California–San Francisco (UCSF) Chimera package. Chimera is developed by the Resource for Biocomputing, Visualization, and Informatics at UCSF (supported by the NIH National Institute of General Medical Sciences P41-GM103311). The model has been deposited in the Protein Data Bank with ID 6BK8 and the EM map has been deposited in the EM Data Bank with ID EMD-7109. X.L., R.Z., and Z.H.Z. conceived the project; X.L. and L.Z. prepared, optimized, and performed biochemical analyses of the sample; S.L., J.J., Y.C., and Z.H.Z. recorded and processed the EM data; R.C.H. and K.C.H. performed mass spectrometry analyses; S.L. built the atomic models;

S.L., R.Z., and Z.H.Z. analyzed and interpreted the models; S.L., X.L., and R.Z. prepared the illustrations; R.Z., S.L., and Z.H.Z. wrote the paper; and all authors contributed to the editing of the manuscript.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/358/6368/1278/suppl/DC1 Materials and Methods Figs. S1 to S5 Table S1 References (*39–54*) Movie S1

28 October 2017; accepted 9 November 2017 Published online 16 November 2017 10.1126/science.aar3462

Science

Structure of the yeast spliceosomal postcatalytic P complex

Shiheng Liu, Xueni Li, Lingdi Zhang, Jiansen Jiang, Ryan C. Hill, Yanxiang Cui, Kirk C. Hansen, Z. Hong Zhou and Rui Zhao

Science **358** (6368), 1278-1283. DOI: 10.1126/science.aar3462originally published online November 16, 2017

Understanding splicing from the 3' end

The spliceosome removes introns from eukaryotic mRNA precursors and yields mature transcripts by joining exons. Despite decades of functional studies and recent progress in understanding the spliceosome structure, the mechanism by which the 3' splice site (SS) is recognized by the spliceosome has remained unclear. Liu *et al.* and Wilkinson *et al.* report the high-resolution cryo-electron microscopy structures of the yeast postcatalytic spliceosome. The structures reveal that the 3 'SS is recognized through non-Watson-Crick base pairing with the 5'SS and the branch point, stabilized by the intron region and protein factors. *Science*, this issue p. 1278, p. 1283

ARTICLE TOOLS	http://science.sciencemag.org/content/358/6368/1278
SUPPLEMENTARY MATERIALS	http://science.sciencemag.org/content/suppl/2017/11/15/science.aar3462.DC1
RELATED CONTENT	http://science.sciencemag.org/content/sci/358/6368/1283.full
REFERENCES	This article cites 54 articles, 15 of which you can access for free http://science.sciencemag.org/content/358/6368/1278#BIBL
PERMISSIONS	http://www.sciencemag.org/help/reprints-and-permissions

Use of this article is subject to the Terms of Service

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. 2017 © The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. The title *Science* is a registered trademark of AAAS.