Molecular Cell
Article

# 3D Cryo-EM Structure of an Active Step I Spliceosome and Localization of Its Catalytic Core

Monika M. Golas,[1,3,4,7] Bjoern Sander,[1,4,5,7] Sergey Bessonov,[2] Michael Grote,[2] Elmar Wolf,[2] Berthold Kastner,[2] Holger Stark,[1,6,*] and Reinhard Lührmann[2,*]
[1]Research Group of 3D Electron Cryomicroscopy
[2]Department of Cellular Biochemistry
Max Planck Institute for Biophysical Chemistry, D-37077 Göttingen, Germany
[3]The Water and Salt Research Center, Department of Anatomy
[4]Center for Stochastic Geometry and Advanced Bioimaging
[5]Stereology and EM Laboratory, Institute of Clinical Medicine
Aarhus University, DK-8000 Århus C, Denmark
[6]Göttingen Center for Molecular Biology, Georg August University Göttingen, D-37077 Göttingen, Germany
[7]These authors contributed equally to this work
*Correspondence: hstark1@gwdg.de (H.S.), reinhard.luehrmann@mpi-bpc.mpg.de (R.L.)
DOI 10.1016/j.molcel.2010.11.023

## SUMMARY

**The spliceosome excises introns from pre-mRNA in a two-step splicing reaction. So far, the three-dimensional (3D) structure of a spliceosome with preserved catalytic activity has remained elusive. Here, we determined the 3D structure of the human, catalytically active step I spliceosome (C complex) by cryo-electron microscopy (cryo-EM) in vitrified ice. Via immunolabeling we mapped the position of the 5′ exon. The C complex contains an unusually salt-stable ribonucleoprotein (RNP) core that harbors its catalytic center. We determined the 3D structure of this RNP core and also that of a post-step II particle, the 35S U5 snRNP, which contains most of the C complex core proteins. As C complex domains could be recognized in these structures, their position in the C complex could be determined, thereby allowing the region harboring the spliceosome's catalytic core to be localized.**

## INTRODUCTION

The excision of noncoding RNA ("introns") from pre-mRNA is catalyzed by the spliceosome, a complex molecular machine (Nilsen, 2003). The initial recognition of the 5′ splice site (ss) of an intron triggers the formation of the spliceosome by the stepwise assembly of its constituent small nuclear ribonucleoprotein particles (U snRNPs) and its many other proteins (Staley and Guthrie, 1998). The most important intermediate states in the splicing process are (1) the A complex, in which the ends of the intron/exon boundaries are defined, (2) the B complex, in which the complete set of U snRNPs is first present, (3) the B* complex, the catalytically activated form of the B complex, and (4) the C complex or step I spliceosome, which emerges from the first catalytic step of splicing and subsequently catalyzes the second (Wahl et al., 2009). Each transition from one of these complexes to the next involves radical changes in the spliceosome's composition, with corresponding rearrangements in its structure and patterns of internal protein-protein, protein-RNA, and RNA-RNA interaction (Bessonov et al., 2008; Makarov et al., 2002). Thus, the spliceosome exhibits an unusual complexity in terms of recruiting and destabilization events not encountered in comparable biomolecular machines such as large enzymes or the ribosome (Nilsen, 2003).

The size and structural variability of the spliceosome have made it a particularly challenging object for structural studies. For its smaller, individual components, X-ray crystallography and NMR are beginning to reveal detailed structures down to the atomic level (Kambach et al., 1999; Pomeranz Krummel et al., 2009; Ritchie et al., 2009). However, for the multimegadalton spliceosomal complexes such as A, B, B*, and C, cryo-electron microscopy (cryo-EM) is the method of choice (Behzadnia et al., 2007; Boehringer et al., 2004; Jurica et al., 2004), and indeed this approach has begun to shed light on the structural organization of large building blocks of the spliceosome (Golas et al., 2003; Sander et al., 2006; Stark and Lührmann, 2006). At the same time, mass spectrometry has been applied to identify the proteins involved—again revealing a pattern of high complexity, with altogether well over 170 different proteins associated with the various human spliceosomal complexes or their subunits (Behzadnia et al., 2007; Bessonov et al., 2008; Deckert et al., 2006; Jurica et al., 2002; Jurica and Moore, 2003).

Determining the structure of the C complex is a task that is particularly interesting and challenging. This is due to its position midway between the two catalytic steps of splicing and to the remarkable fact that ~35 proteins are lost during the transition from the B to the C complex, while about the same number are newly recruited to the spliceosome during this step (Bessonov et al., 2008), with accompanying structural rearrangements that are poorly understood (Brow, 2002; Konarska and Query, 2005; Wahl et al., 2009). Furthermore, the C complex possesses a relatively small subset of proteins (~35 of its full complement of ~110) that remain stably bound together with all of its RNAs—the U2, U5 and U6 snRNAs, as well as both pre-mRNA splicing

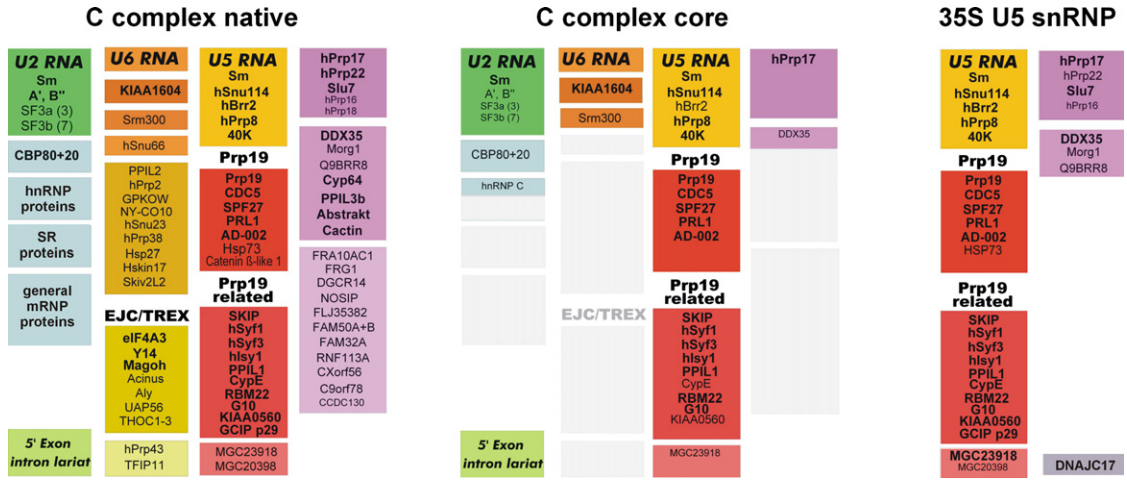


**Figure 1. Protein and RNA Compositions of the C Complex, Its Salt-Stable Core, and the 35S U5 snRNP**

The analyzed complexes include the step I spliceosome affinity purified under physiological salt conditions (C complex native, left), its salt-stable RNP core (C complex core, middle) (Bessonov et al., 2008), and the 35S U5 snRNP (right). The figure shows proteins, grouped according to function or complex association, along with the RNA content (Wahl et al., 2009). Note that the 35S U5 snRNP solely contains the U5 snRNA. The relative abundance of proteins is indicated by light (substoichiometric amounts) or dark (abundant) lettering and is based on the relative amounts of peptides sequenced.

intermediates (excised 5′ exon and intron lariat)—even at high salt concentrations at which many RNPs are unstable (Bessonov et al., 2008). This stable RNP core of the C complex contains, among others, most of the constituents of a U snRNP particle that is thought to be formed within the spliceosome during catalytic activation and then released after completion of splicing, namely the 35S U5 snRNP (Makarov et al., 2002).

Cryo-EM of transient complexes is challenging because such assemblies tend to disintegrate during purification and sample preparation. We recently presented the GraFix protocol, which involves mild fixation and thus overcomes the critical limitation of structural disintegration (Kastner et al., 2008). Here, we applied it to functionally active C complexes that contain all of the components required for performing the second step of splicing (Bessonov et al., 2008). We describe the three-dimensional (3D) architecture of this catalytically active spliceosome and of the postspliceosomal 35S U5 snRNP in vitrified ice. Moreover, we also determined the 3D structure of the C complex's salt-stable RNP core by negative-staining EM and examined the relationship between the structure of the C complex and that of its core components. Furthermore, we located the 5′ exon within the C complex by immunolabeling. From the combined data, we could localize the region harboring the catalytic core of the C complex.

## RESULTS

### 2D Representation of Functional C Complexes

To determine the structure of a catalytically active spliceosome, we assembled C complexes on a model pre-mRNA termed PM5 that contains a long polypyrimidine tract at the 3′ end of the intron, but lacks a 3′ ss and 3′ exon (Figure S1A, top, available online), so that assembled spliceosomes are stalled after performing the first step of the splicing reaction. However, they can catalyze the second step when a 3′ ss is added in *trans* (Bessonov et al., 2008). C complexes were affinity purified and found to contain both splicing intermediates (excised 5′ exon and intron-lariat), the U2, U5, and U6 snRNAs, and their known complement of ~110 proteins (Bessonov et al., 2008) (Figure 1, left, and Figure S1B, lanes 1–3).

We first subjected the C complexes directly to negative staining, but this resulted in their partial dissociation, as observed previously for the spliceosomal B complexes (Kastner et al., 2008). This destabilization was also reflected by the observation of pre-mRNA-containing disintegration products in gradient profiles of radioactively labeled, uncrosslinked C complexes (Figure S1C). In contrast, the GraFix method yielded a homogeneous, monodisperse population of intact particles with maximum dimensions of >40 nm (Figure 2A). Unbiased single-particle image clustering indicated the preferential orientation of the particles on the carbon film; representative class averages show a particle comprising a central density and several surrounding domains attached to it (Figure 2B). The upper region appears to vary somewhat in size and shape, while the rest of the complex has a stable morphology.

Unstained cryo-EM using vitrification allows macromolecules to be visualized in a fully hydrated environment and requires that the buffer be compatible with the vitrification; e.g., a high concentration of glycerol (such as exists in our gradient fractions) results in poor specimen quality. Therefore, GraFix stabilized C complexes were placed in a glycerol-free physiological buffer to allow visualization of the particles under stain-free, hydrated conditions. Checks by scintillation counting of radioactively labeled RNA and by negative-staining EM confirmed the integrity of the particles after buffer-exchange. With unstained cryo-EM, a monodisperse population of particles >40 nm in size was observed (Figure 2C). Image clustering revealed a distribution of density elements similar to that seen with negative staining (Figure 2D, top row; the other class averages show different angular views).

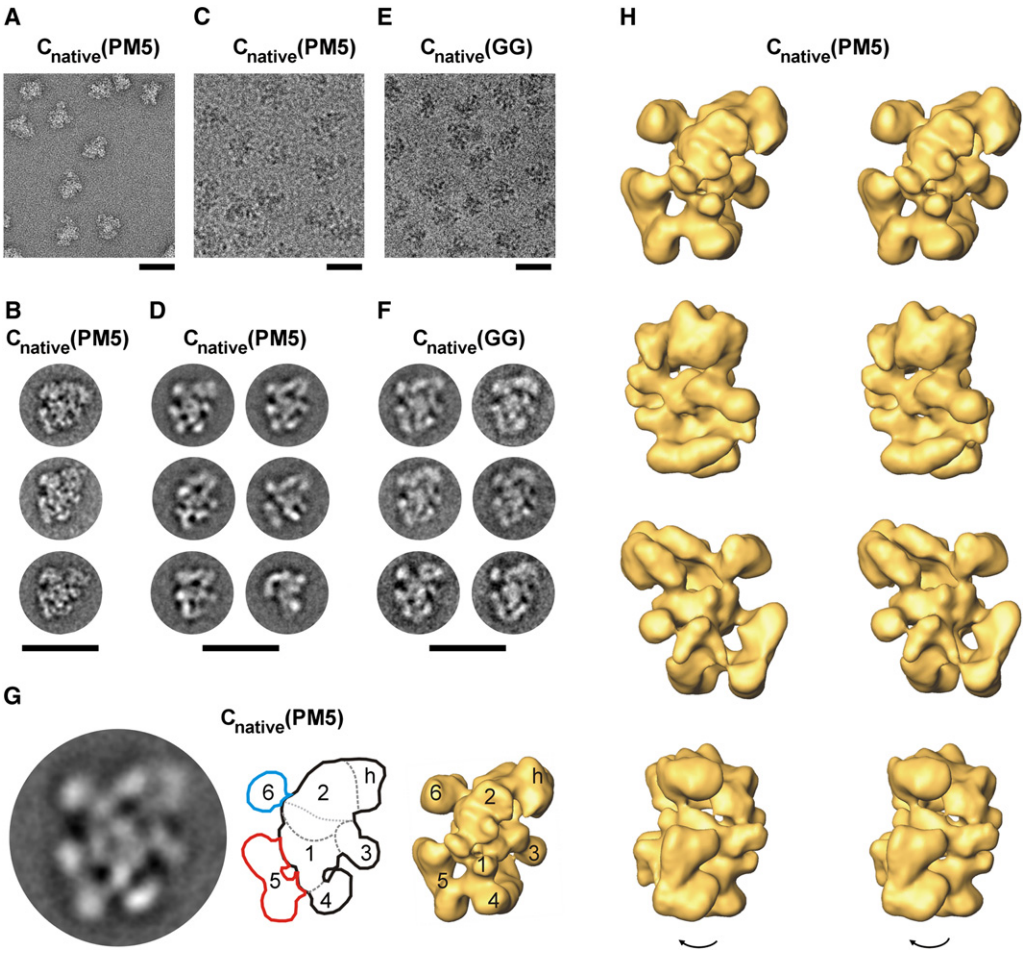


**Figure 2. Characterization of the Catalytically Active, Human Step I Spliceosome**

(A) Raw image of the step I spliceosome formed on the PM5 pre-mRNA with negative staining EM.

(B) Representative negative stain class averages of the step I spliceosome.

(C) Raw image of the step I spliceosome with unstained cryo-EM.

(D) Representative unstained class averages of the step I spliceosome in vitrified ice.

(E) Raw image of the step I spliceosome assembled on the $MINX_{GG}$ pre-mRNA with unstained cryo-EM in vitrified ice.

(F) Representative class averages of the step I spliceosome assembled on $MINX_{GG}$. Particles exhibit morphological features similar to those observed for the PM5 step I spliceosome (compare Figure 2D).

(G) Main structural domains of the step I spliceosome (middle) as exemplified by a characteristic unstained class average (left) and a 3D view (right).

(H) 3D structure of the step I spliceosome from vitrified, unstained cryo images. Shown are four stereo pairs separated by rotations of 90° as indicated.

All scale bars represent 50 nm. See also Figure S1 and Table S1.

## Structural Organization of the C Complex Assembled on another Pre-mRNA Substrate

Before undertaking a 3D structure determination, we tested whether the structural organization of the C complex is dependent upon the identity of the pre-mRNA. To this end, we purified C complexes formed on a different pre-mRNA construct ($MINX_{GG}$) that has a 3′ exon, but carries a GG mutation at the 3′ ss (Figure S1A, bottom); again, spliceosomes can assemble on this pre-mRNA and catalyze the first step of splicing, but thereafter splicing is stalled. The RNA and protein composition of purified C complexes formed on $MINX_{GG}$ are essentially identical to those formed on the PM5 substrate described above (Figure S1B, lanes 4–6). In unstained cryo-EM (Figure 2E), these particles again exhibited a maximum size of >40 nm, and their general appearance, fine-structural details, and morphological elements were undistinguishable from those of the PM5 C complex, except for an additional small domain at the bottom, which was observed in some class averages (Figure 2F). We conclude that the identity of the pre-mRNA does not have a substantial effect on the gross structure of the C complex for these two different pre-mRNAs.

## 3D Structure of Catalytically Active C Complexes

By combining data from 36,071 cryo-EM images in vitrified ice, we reconstructed a 3D map of the catalytically active C complex (Figures 2G and 2H and Figure S1D). The resulting structural map

is well-defined and has a resolution of 2.0–2.9 nm (Figure S1E). Several structural domains can be discerned (Figure 2G, middle and right). These include a central domain (domain 1) surrounded by several connected domains, the most prominent of which are numbered 2–6 (Figure 2G and Figure S1F), including a bipartite platform-like structure (domain 5). In addition, at the top of the complex to the right, domain h can be distinguished from domain 2. The structural landmarks seen in the 3D map are in line with the distribution of density elements already visible in the EM views (Figure S1D).

Domain 2 appears to form a bulge toward domain 1, which gives the complex a more compact appearance when viewed from the front (Figure 2H, top row). In contrast, when viewed from the back, a deep cleft between domains 1 and 2 in the upper part of the complex (third row) is visible, which divides the particle into an upper one-third and a lower two-thirds. The approximate position of the border between domains 1 and 2 in the front view is indicated in Figure 2G (middle) by the lower of two dashed lines, whereas the border observed in the back view is indicated by the upper dashed line. Overall, the C complex has a size of ~36 nm × 34 nm × 27 nm (oblique maximum diameter, ~42 nm, Figure S1F) and is estimated to enclose a molecular mass of 5–5.5 MDa. This estimate is in line with the molecular mass of the RNAs and the sum of the ~80 most abundant proteins of the C complex (Figure 1, left) and with measurements by quantitative scanning transmission electron microscopy (unpublished data).

## Annotation Strategy

Given its unusually large dimensions and complex multidomain architecture, the annotation of the building blocks of the step I spliceosome on the protein/RNA level represents a major challenge. We adopted a multifacetted strategy to obtain a global 3D understanding of the structural organization of the C complex. With the goal of mapping its catalytic core, we determined the structure of its salt-stable RNP core, which retains ~35 C complex proteins including all catalytic core components (Figure 1, middle), as well as the structure of the postspliceosomal 35S U5 snRNP, which shares U5-associated components with the native C complex and the salt-stable RNP core, but lacks the catalytic RNA-network of U2 and U6 snRNA and the pre-mRNA splicing intermediates (Figure 1, right). In addition, we determined the location of the 5′ exon of the pre-mRNA. By this combined approach, we were able to localize the region containing components of the C complex's catalytic core.

## Structure of the Salt-stable RNP Core of the C Complex

To purify the salt-stable core, we treated C complexes formed on the PM5 substrate with 1 M NaCl and isolated the resulting salt-stable RNP core. As summarized in Figure 1 (middle), analysis of the RNA and protein composition of the RNP core revealed that it contains all three snRNAs of the C complex (U2, U5, and U6), plus the pre-mRNA splicing intermediates and ~35 of its ~110 proteins; functional assays indicate that the RNA/RNP networks in this core complex are intact (Bessonov et al., 2008).

EM after GraFix and negative staining revealed particles with a maximum diameter varying from ~28 nm to ~34 nm (in rare cases, up to ~40 nm; Figure 3A) in views similar to those of the intact C complex. Image clustering yielded well-resolved class averages that disclosed a central region where the elements of density are nearly always arranged in the same way (central and lower elements in Figure 3B), but also regions of structural variability between individual particle populations (red and yellow arrowheads). A semitransparent overlay of all observed main view shapes (Figure 3C, right) shows good correspondence in the central and lower part (domains 1 and 3–5), where the density patterns are highly reminiscent of the corresponding

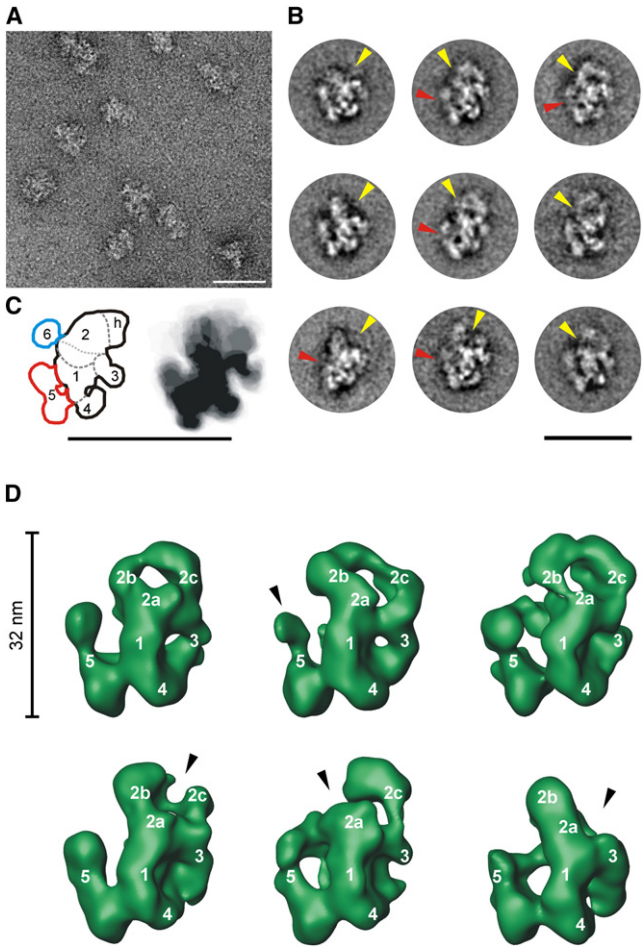


**Figure 3. Characterization of the Salt-Stable RNP Core of the Step I Spliceosome**

(A) Raw image of the salt-stable RNP core of the step I spliceosome obtained by negative staining EM.

(B) Representative class averages of the salt-treated RNP core. The underrepresentation of the upper domain 5 is indicated by a red arrowhead and of the domains 2 and 6 by a yellow arrowhead.

(C) Shape profiling of the class averages of the RNP core by grayscale-encoded overlays (right) showing the shared structural elements of the particle (domains 1, 3, 4, lower portion of 5), as well as the underrepresentation or lack of part of domains 2, 5, and 6 (left).

(D) Six different 3D class averages of the salt-stable RNP core. Domains 1–5 are labeled according to the corresponding domains of the native C complex; domain 2 is subdivided into the subdomains 2a–c. Variable regions are indicated with arrowheads.

All scale bars represent 50 nm. See also Figure S2.

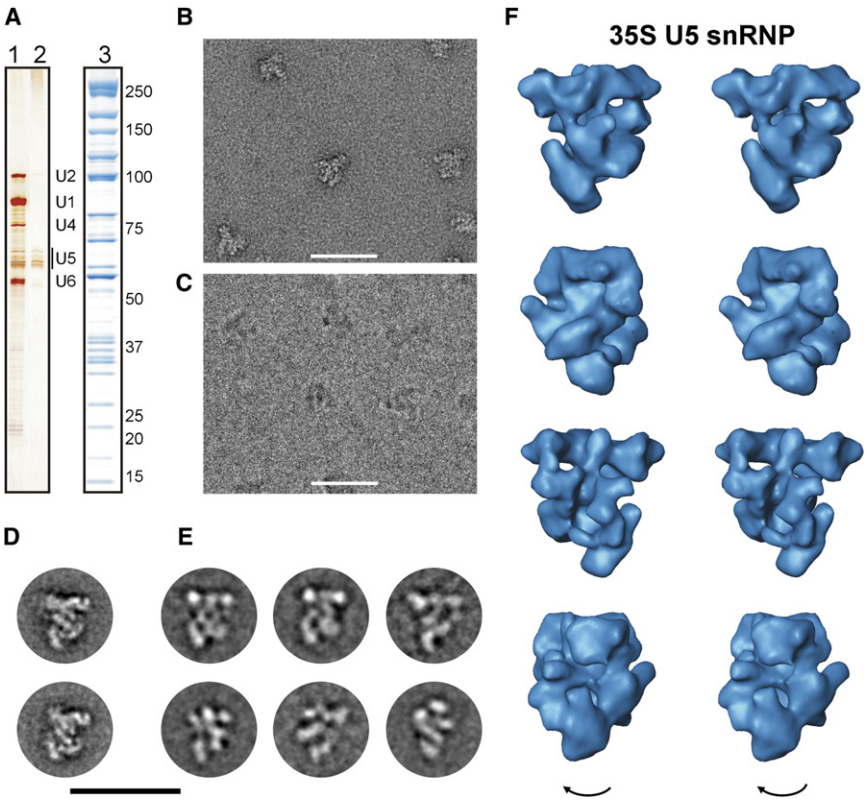


**Figure 4. Characterization of the Postspliceosomal 35S U5 snRNP**

(A) RNA and protein composition of the 35S U5 snRNP as determined by silver-stained denaturing PAGE (lane 2) and by Coomassie blue-stained SDS-PAGE (lane 3). Lane 1 shows a U snRNA marker. Identities of RNAs and the molecular weight marker (in kDa) are indicated on the right.

(B) Raw image of the 35S U5 snRNP with negative staining EM.

(C) Unstained raw cryo-EM image of the 35S U5 snRNP in vitrified ice.

(D) Representative class averages of the 35S U5 snRNP as obtained by negative staining EM.

(E) Representative class averages of the 35S U5 snRNP as obtained by unstained cryo-EM in vitrified ice.

(F) 3D structure of the 35S U5 snRNP determined from unstained, vitrified cryo images. Shown are four stereo pairs separated by 90° rotations as indicated.

All scale bars represent 50 nm. See also Figure S3.

portion of the native C complex (Figure 3C, left). In contrast, domain 6 and h are essentially absent, while some density at the upper part of domain 5 and also some density of domain 2 appear to be missing frequently.

The 2D analysis of the salt-stable RNP core complex indicated structural variability of the particles that is probably caused by variable dissociation of some salt-labile components. To determine the 3D structures of the different structural subgroups, we performed a 3D classification of individual 3D random conical tilt (RCT) maps subsequent to a 3D alignment of the RCT maps using a weighted average 3D structure (Figure S2) as reference volume. Six 3D class averages are shown in Figure 3D and they represent the main structure classes of the salt-stable RNP core complex. The largest 3D class average has dimensions of ~32 nm × 28 nm × 21 nm with an oblique diameter of ~34 nm. Despite the lower resolution (~3.2–4 nm) of the salt-stable RNP core, domains seen in the C complex, in particular domains 1, 3, 4, and 5 (the upper part of which is somewhat smaller and more variable in the salt-stable RNP core) can be recognized. In contrast, domains 6 and h are absent in these 3D structures. In addition, density comprising part of domain 2 is also missing, but to varying extents in the different 3D class averages. To describe these differences within domain 2 in more detail, we divided this domain further into three subdomains denoted 2a–2c (Figure 3D), of which solely subdomain 2a is present in all of the 3D reconstructions. Subdomains 2b and 2c are present in most, but not all, 3D reconstructions. Between 2b and 2c, a central hole is visible in most 3D reconstructions. This hole is caused by the absence of another portion of domain 2 (denoted 2d; see below) that is only present in the native C complex.

## 3D Structure of the 35S U5 snRNP

35S U5 snRNPs were purified by generation of a stable HeLa cell line that expresses one of the protein components of the 35S U5 snRNP (i.e., AD002) with an N-terminal tag, followed by affinity selection. Purified 35S U5 snRNPs are comprised nearly exclusively of the U5 snRNA (present in several isoforms; Figure 4A, lane 2) and ~37 proteins, including the U5 proteins hPrp8p/U5-220K, hBrr2p/U5-200K, hSnu114p/U5-116K, and U5-40K, as well as the Prp19 complex and its related factors (Figure 4A, lane 3, and Figure 1, right). All of these proteins are also part of the salt-stable C complex RNP core. Indeed, the 35S U5 snRNP contains only a handful of proteins—such as Slu7, hPrp22, DDX35, Q9BRR8, Morg1, and GCIPp29—that are not present in the salt-stable RNP core of the C complex. A major difference between the 35S U5 and the C complex's salt-stable RNP core particle is the absence of U2 snRNA and its associated proteins, most of the U6 snRNA, and the pre-mRNA splicing intermediates; notably, regions of these RNAs comprise the catalytically active RNA network (Nilsen, 1994). Thus, this part of the catalytic core of the C complex does not exist in the 35S U5 snRNP.

As with the C complex, we subjected the purified, GraFix-treated 35S U5 particle to negative staining as well as to unstained cryo-EM. EM images of negatively stained and unstained 35S U5 snRNPs revealed a compact particle with a maximum dimension of ~24–27 nm (Figures 4B and 4C). In accord with its reduced proteome, the 35S U5 snRNP complex is smaller than the C complex and shows fewer density elements. Particles appear to be typically trapezoidal or triangular, both in negatively stained and in unstained-cryo images (Figures 4D and 4E). The 3D map of the 35S U5 snRNP was

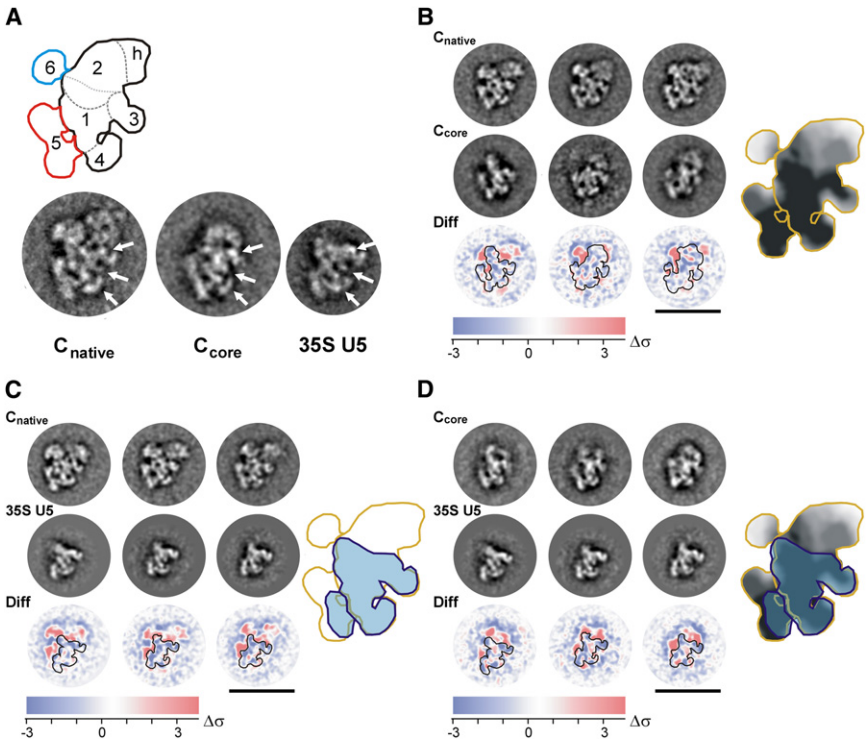


**Figure 5. Comparison of the Views of the Step I Spliceosome, Its Salt-Stable RNP Core, and the 35S U5 snRNP**

(A) Representative views of the PM5 step I spliceosome ($C_{native}$), the 1M salt-treated RNP core of the C complex ($C_{core}$), and the 35S snRNP (35S U5). For a better identification of shared structural elements, compare arrows.

(B) Difference mapping of the RNP core and the step I spliceosome. First row: class averages of the step I spliceosome. Second row: class averages of the salt-stable RNP core. Third row: difference maps (color scale at the bottom) computed by subtracting the RNP core from the C complex (Diff). Right: combined overlay of the outline of the C complex (golden) and the shape profile of its core (grayscale).

(C) Difference mapping (color scale at the bottom) of the step I spliceosome and the 35S U5 snRNP with resulting overlay (right; golden outline, step I spliceosome; blue area, 35S U5 snRNP).

(D) Difference mapping (color scale at the bottom) of the spliceosomal salt-stable RNP core and the 35S U5 snRNP with resulting overlay (right; golden outline, step I spliceosome; grayscale profile, salt-treated RNP core; blue area, 35S U5 snRNP).

The scale bars represent 50 nm.

reconstructed (Figure S3A), independently of the other particles, to a resolution of 2.4–2.8 nm using unstained cryo-EM (Figure S3B). Overall, the 35S U5 snRNP has dimensions of ~27 nm × 26 nm × 23 nm (Figure S3C) and appears to be built up around an elongated central domain with surrounding structural elements (Figure 4F). The 3D structure corresponds to a particle with an estimated enclosed molecular mass of ~2.4 MDa, in agreement with the composition of the particle.

## Difference Mapping of the 2D images of the C Complex, Its RNP Core and the 35S U5 snRNP

The salt-stable RNP core and the 35S U5 snRNP both contain fewer proteins than the native C complex (Figure 1). However, the recognition of shared structural elements is facilitated by the fact that the main orientations of the three complexes apparently differ only marginally from one another, as they are all adsorbed preferentially and in a very similar way on the EM carbon film. All three complexes thus show a core region with a remarkably similar density distribution in the class averages, including a hook-like element in domain 3 (Figure 5A, upper arrows) that is found at the same distance from the similarly shaped elements at and above domain 4 (lower and middle arrows). To corroborate this, we performed comprehensive pairwise comparisons of the different particles that included the computation of difference maps between two selected, aligned class average pairs.

First, class averages of the salt-stable core of the C complex were subtracted from class averages of the native C complex (examples in Figure 5B). This comparison revealed that, while the central parts of the two complexes appear structurally very similar, fewer regions of density are observed in the upper regions of the salt-treated C complex core (red regions in the difference maps), in particular those corresponding to domains h and 6 of the C complex, and in some regions of domain 2. Additionally, some views of the C complex core appear to lack density at the upper part of the bipartite domain 5 (e.g., view 3). This is also supported by docking the shape overlays of the C complex core into the outline of the native C complex (Figure 5B, right; overlays slightly blurred).

A comparison of the 35S U5 snRNP with the native C complex also disclosed remarkable similarities, both in the overall distribution of density elements and in structural details of the core domains (examples in Figure 5C). However, differences are seen in domains 2, h, and 6 and in the upper part of domain 5, which are missing in the 35S U5 snRNP (red regions in the difference maps). This is also summarized by combining the outlines of the 35S U5 snRNP and native C complex (Figure 5C, right). Accordingly, the upper border of the 35S U5 snRNP corresponds to the backmost border between domains 1 and 2 in the native C complex (Figure 5A, gray dashed line in domain 2).

Finally, a comparison of the salt-stable C complex core and the 35S U5 snRNP likewise revealed clear similarities (examples and sketch in Figure 5D). However, the upper part of the bipartite domain 5 found in both the native C and subsets of the salt-stable RNP core complex is missing in the 35S complex, and likewise density comprising domain 2 is missing in the 35S U5 snRNP (red regions in the difference maps).

## Location of the 5′ Exon in the C Complex

To locate the 5′ exon in the C complex, we made use of the MS2 hairpins at the 5′ end of the 5′ exon of the PM5 pre-mRNA. For

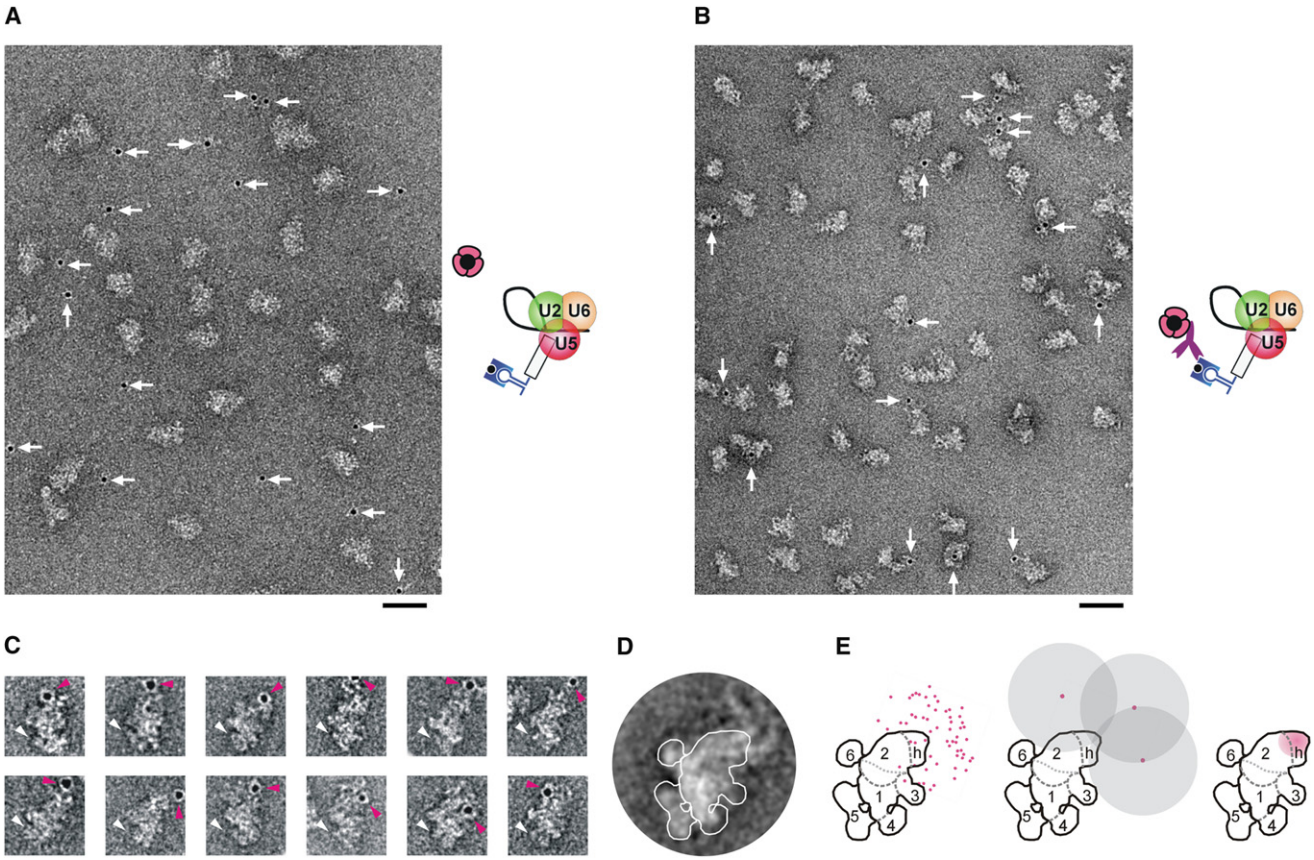


**Figure 6. Location of the 5′ Exon in the Step I Spliceosome**

(A) Overview image of negatively stained C complexes incubated with PA-gold in the absence of anti-MBP antibodies as negative control. C complexes were assembled on the PM5 pre-mRNA with three MS2 aptamers at the 5′ end of the 5′ exon. After affinity selection, the complexes were incubated in the absence of the antibody, fractionated in a GraFix gradient, and incubated with PA-gold.

(B) Same as (A), but with addition of the anti-MBP antibody. The scale bars represent 50 nm.

(C) Individual immunogold-labeled C complexes orientated according to the preferred view of the complexes (Figure 2G). For facilitation of orientation, the colloidal gold is marked with a pink arrowhead and the upper part of the platform density with a white arrowhead.

(D) Slightly rotated average of the immunolabeled particles shown in (C).

(E) The immunolabel maps the 5′ end of the 5′ exon. The observed positions of the PA-gold (pink dots) point toward domain h and the upper portion of domain 2 (left). The greatest possible distance between the center of the 5 nm colloidal gold and the antigen-binding sites of the antibody is ~17 nm. For three selected PA-gold positions, a 17 nm circle (gray) is shown (middle). The overlapping area of all labels indicates that the 5′ end of the 5′ exon is located in domain h or the adjacent, upper portion of domain 2 (right).

affinity purification, these MS2 hairpins are stably bound by the MS2 portion of the fusion protein MS2-maltose binding protein (MS2-MBP); the MBP can be specifically labeled with an anti-MBP antibody and Protein A (PA)-coated colloidal gold (Wolf et al., 2009). We therefore incubated C complexes with this anti-MBP antibody and subjected them to affinity purification and GraFix ultracentrifugation. Subsequently, PA-gold was added to the C complex-containing gradient fractions, and negative-stain EM specimens were prepared. When the anti-MBP antibody was omitted, the PA-gold was mostly observed in the EM images at varying distances away from the C complexes (Figure 6A; the electron-dense gold is indicated by arrows). In contrast, when the anti-MBP antibody was used, the PA-gold was typically found in close proximity to the C complexes (Figure 6B), indicating specific labeling of these complexes.

Selected examples of gold-labeled C complexes exhibiting the typical C complex features are shown in Figure 6C. The average of these images (Figure 6D) corresponds well with the C complex's outline shown in Figure 2, confirming that intact C complexes were imaged in their standard orientation. All of the observed positions of the gold were then located with respect to the typical structural features. Most of the PA-gold was found at the periphery of the upper part of domain 2 and in domain h (Figure 6E, left). As the distance between the center of the gold label and the antibody-binding site is at most ~17 nm (Wolf et al., 2009), the position of the 5′ end of the 5′ exon can be narrowed down by a superposition of 17 nm circles (for three typical examples, see Figure 6E, middle). If this constraint is taken into account, the overlapping area of these 17 nm circles indicates that the 5′ end of the 5′

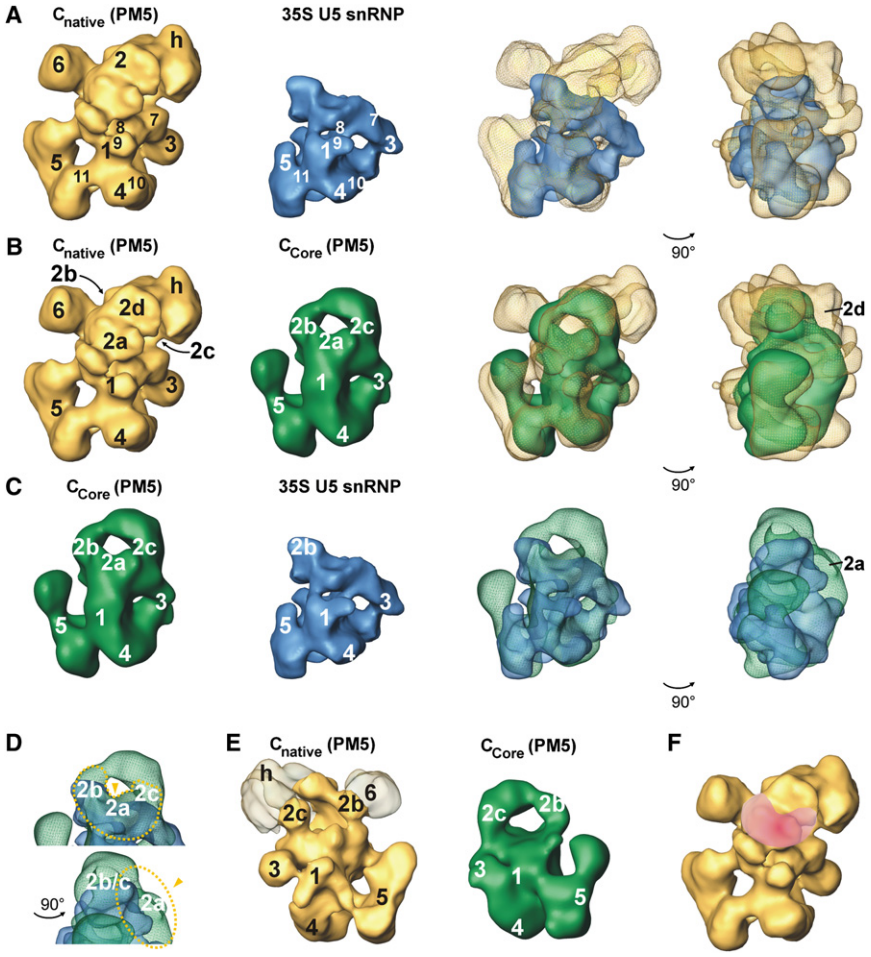


**Figure 7. Localization of the Catalytic Core of the Spliceosomal C Complex**

(A) The comparison of the step I spliceosome (golden) and the 35S U5 snRNP (blue) on the 3D level reveals that apart from domains 1, 3, 4, and the lower part of 5, additional fine structural details are consistently found in the complexes (labeled 7–11). In the 35S U5 snRNP, domain 2 is reduced to a portion of domain 2b. The fit (two rightmost panels) is shown from two different orientations separated by 90° as indicated.

(B) Same as (A), except the salt-stable RNP core (green) is fit into the native C complex (golden).

(C) Same as (A), but for the fit of the 35S U5 snRNP (blue) into the salt-stable RNP core of the C complex (green).

(D) Zoomed-in views separated by a 90° rotation of the fit shown in (C) reveal a volume consisting of 2a and parts of 2b and 2c that is missing in the 35S U5 snRNP and shows an interface to densities present in the 35S U5 snRNP.

(E) Comparison of the native C complex and C core complex seen from the back. Domains h and 6 are transparent in the native C complex (left). The remaining, opaque parts of the native C complex show very good correspondence to the C core complex (right).

(F) Proposed area of the catalytic core (pink).

See also Figure S4 and Movies S1–S3.

exon is located in the C complex in a region consisting mainly of the domain h plus a small region in the upper part of domain 2 (Figure 6E, right).

## Localization of the Catalytic Core by 3D Fitting of the C Complex, Its RNP Core, and the 35S U5 snRNP

We now fitted the structures of the C complex, its salt-stable RNP core, and 35S U5 snRNP on the 3D level. We first identified structural counterparts of the large domains 1–5 of the C complex (Figure 7A, gold) within the 35S U5 snRNP (blue) and aligned the 3D structures with respect to these coarse features. This allowed us to identify the C complex domains 1 and 3–5 three-dimensionally within the 35S U5 snRNP. The correspondence of domains 1 and 3–5 allowed us to dock the two complexes into each other, which resulted in an unambiguous fit (Figure 7A and Movie S1). After this docking further fine-structural elements shared by the two particles became recognizable: many fine-structural details—see, e.g., the small clefts and bumps labeled 7–11—were found in corresponding positions (Figure 7A).

The unambiguous placement of the 35S U5 snRNP structure into the C complex allowed the identification of the additional C complex domains. The most obvious difference is present in the upper region of the C complex. Domains h and 6 and a large portion of domain 2 are missing in the 35S U5 snRNP. In the lower region, the upper part of domain 5 is present only in the C complex. Some smaller features also differed between the two complexes; e.g., in the 35S particle, the remaining part of the bipartite platform is rotated toward the central domain 1. Our analysis thus indicates that the 35S U5 snRNP is located in the lower and central region of the C complex (Figure 7A, the two rightmost panels).

Despite its lower resolution, the fitting of the salt-stable C core complex (green) into the C complex proved to be equally straightforward. Here, in addition to domains 1 and 3–5, several corresponding structural features present in the upper part of the particle are also visible (Figure 7B). The salt-stable RNP core complex clearly lacks domain 6 and h, as well as some parts of domain 5 (Figures 7B and 7E and Movie S2). In the salt-stable C complex core, part of domain 2 is also missing. As described above (Figure 3), domain 2 is present to varying extents in the main 3D reconstructions of the salt-stable C complex core. In Figure 7, the 3D reconstruction of the salt-stable core that represents the main structural subpopulation of all particles is shown. Superpositioning of the native C complex and its salt-stable core reveals that subdomain 2a is present in both complexes, while subdomain 2d is lacking in the latter (Figure 7B). In the standard view of the native C complex, subdomain 2d obscures most of subdomains 2b and 2c. The latter subdomains are best seen in the view of the back side of the C complex

(Figure 7E). A comparison with the back side of the salt-stable core complex reveals that these subdomains have very similar morphologies in both complexes. Since the salt-stable RNP core particle contains the major components of the catalytic core (which comprises largely regions harboring the pre-mRNA-U2-U6 RNA network), it follows that the catalytic core cannot be located in any of the C complex regions that are absent in the salt-stable RNP core (i.e., domain 6, h, and 2d).

Fitting of the 35S U5 snRNP into the salt-stable RNP core (Figure 7C) discloses clear similarities in the central and lower domains. Indeed, the lower part of the salt-stable core is nearly completely encompassed by the 35S U5 snRNP structure. Only very small domains of the 35S U5 snRNP are not covered by the densities of the salt-stable RNP core complex, in line with there being only a few additional proteins in the 35S U5 snRNP (Figure 1 and Movie S3). Regions present in the salt-stable RNP core but absent in the 35S U5 snRNP are the upper part of domain 5 and nearly all of the subdomains of domain 2 that are still present in the salt-stable core RNP; only the lower part of subdomain 2b is also present in the 35S U5 snRNP (Figure 7C, the two rightmost panels; Movie S3). Subdomain 2a is not present in the 35S U5 snRNP, as best seen in the superposition of it and the salt-stable RNP core (the rightmost panel of Figure 7C; Figure 7D). The U2-U6 RNA network of the catalytic core should be located in those regions of the salt-stable RNP core not found in the 35S U5 structure but that also share a substantial interface with regions of the 35S U5 snRNP, as components of the 35S U5 snRNP also contribute to the catalytic core of the C complex. Because of its proximity to the site where the 5′ exon was located, portions of domain 2 are the most likely candidates. Of these, subdomain 2a has the most extensive interface with the region containing the 35S U5 snRNP and it is the only subdomain of domain 2 that is present in all of the 3D reconstructions obtained for the salt-stable RNP core. We thus conclude that major parts of the U2-U6-pre-mRNA network that comprises part of the catalytic core are primarily located in domain 2a (see below).

## DISCUSSION

### Architecture of the Step I Spliceosome

Here, we have reported the 3D structure, obtained by unstained cryo-EM, of a catalytically active spliceosomal complex purified under physiological conditions. The human step I spliceosome was found to exhibit dimensions of ~36 nm × 34 nm × 27 nm (oblique maximum diameter, ~42 nm), which is clearly larger than the prokaryotic 70S ribosome (~25 nm × 23 nm × 20 nm; Figure S4) (Gao et al., 2009). Thus, the spliceosomal C complex is to date the largest structurally defined RNP machine. With maximum dimensions exceeding 40 nm and a molecular mass of ~5–5.5 MDa, our C complexes are significantly larger than C complexes described earlier, which had maximum dimensions of ~27 nm × 24 nm × 22 nm and a molecular mass of ~2.6 MDa (Jurica et al., 2004) (Figure S4). The dimensions of the latter are even considerably smaller than our structure of the salt-stable RNP core of the C complex (~32 nm × 28 nm × 21 nm). They fit more closely to those of the 35S U5 snRNP (~27 nm × 26 nm × 23 nm).

It could be argued that the difference in size is due to the choice of pre-mRNA used for the experiments. However, this seems unlikely: C complexes that we assembled on two different pre-mRNAs (PM5 and $MINX_{GG}$) were essentially identical, and the $MINX_{GG}$ substrate is very similar to the one used in the earlier study (Jurica et al., 2004). Moreover, the dimensions of our C complex accords with the size of the human BΔU1 complex (Boehringer et al., 2004) and also with the 2D EM views of the *Saccharomyces cerevisiae* B and C complexes (Fabrizio et al., 2009), both of which exhibited a maximum dimension of ~40 nm. We suggest that the aforementioned size difference was due to the fact that our spliceosomes were purified under physiological conditions and further stabilized by mild fixation to prevent dissociation of the complexes after their isolation.

### Comparison with the Spliceosomal BΔU1 Complex

Despite their similar sizes, the human C complex and the BΔU1 particle have very different architecture (Figure S4). Although the structure of the BΔU1 complex is at relatively low resolution (4 nm), it is readily apparent that the global distribution of density elements differs between the two spliceosomal particles. The BΔU1 complex has a compact, rhombic shape and is composed of two main substructures (a triangular main body and a globular head domain), while the C complex has a multimodular appearance. This is consistent with the extensive RNP remodeling known to occur during the transition from the B to the C complex, during which more than 35 proteins are lost and more than 35 other proteins are recruited (Bessonov et al., 2008). A structural rearrangement is also observed between yeast B and C complexes (Fabrizio et al., 2009). Thus, the spliceosome has no universal structure during its assembly and catalytic cycle; rather, its structure varies substantially according to its stage of assembly, catalytic activation, and function.

### Structural Model of the C Complex

Determining the structural organization of such a large and dynamic molecular machine is a major challenge. As a first step toward this end, we located the 5′ exon by immunolabeling and determined the 3D structures of the salt-stable RNP core of the C complex and the postspliceosomal 35S U5 snRNP. This allowed us to map the region of the C complex containing its catalytic core.

A fit of the 35S U5 snRNP into the C complex structure revealed that the 35S U5 snRNP resides in the lower 2/3 of the latter. Thus, this region of the C complex contains the U5 proteins hPrp8p/U5-220K, hBrr2/U5-200K, and hSnu114/U5-116K, which are known to function in unwinding U4/U6 during catalytic activation (Grainger and Beggs, 2005; Staley and Guthrie, 1998; Wahl et al., 2009), the Prp19/CDC5L complex and its related factors, plus the U5 snRNA. These and other 35S U5 proteins are needed to maintain the catalytically active RNP core of the C complex (Bessonov et al., 2008). The 35S U5 snRNP also contains components that are thought to comprise part of the C complex's active site, namely loop 1 of the U5 snRNA, which orients the exons for their subsequent ligation (McConnell and Steitz, 2001; Newman et al., 1995), and very probably also the RNase H-like domain of Prp8 (Pena et al., 2008; Ritchie et al., 2008; Yang et al., 2008). Thus, a portion of

the catalytic core of the step I spliceosome should be located in the lower two-thirds of the 3D structure, encompassed by the 35S U5 snRNP. The unambiguous fit of the 35S U5 snRNP into the lower part of the C complex particle shows also that this part of the C complex must consist more or less only of components of 35S U5.

In contrast to the 35S U5 snRNP, the salt-treated C complex core contains, not only U5 snRNA and nearly all other proteins found in the 35S U5 snRNP, but additionally components of the RNA network of the catalytic core, including the U2 snRNP, the U6 snRNA, and the pre-mRNA splicing intermediates (Bessonov et al., 2008; Figure 1). In spite of the relatively low resolution of the high-salt C complex structure, a clear correlation between this structure and that of the 35S U5 snRNP could be established. In this way the regions of the high-salt C complex that are not represented in the 35S U5 snRNP could be recognized. One or more of these regions, which included the upper parts of domain 5 and 2b as well as domains 2a and 2c, can be expected to contain the RNAs of the catalytic center. Furthermore, as components of the 35S U5 snRNP (e.g., hPrp8 and U5 snRNA) also form part of the catalytic core, a portion of the latter must be located at an interface between a 35S U5 containing region and one or more of the aforementioned domains.

That the remainder of the upper part of domain 5 present in the salt-stable RNP core does not contain the catalytic core is supported by the fact that its size is too small to encompass the entire RNP network at the core of the C complex. Furthermore, based on our localization of the 5′ end of the 5′ exon via immunolabeling at and adjacent to domain h, which is located completely opposite of domain 5, it is unlikely that the 3′ end of the 120 nucleotides long 5′ exon might reach this domain and interact there with U5 snRNA loop 1 within the catalytic core. Indeed, in preliminary experiments (data not shown), in which we have removed most of the 5′ exon by RNase H digestion, only domain h and parts of domain 2 were lost, indicating that most of the 5′ exon is located in the upper part of the C complex.

We cannot exclude that components of the catalytic core reside in the lower portions of subdomains 2b and 2c, which are immediately adjacent to the U5-related portion of the native C complex (see golden dashed outline in Figure 7D). However, it is unlikely that the catalytic core is only present in subdomains 2b or 2c, as each of these domains is absent in one or more of the 3D reconstructions of the salt-stable RNP core (Figure 3), indicating that a certain fraction of the purified salt-stable RNP core particles lack each of these regions. Gradient analysis of RNA released from the native C complex upon treatment with high salt and subsequent purification of the salt-stable RNP core indicates that that there is little or no loss of U5, U2, or U6 snRNA or splicing intermediate. Thus, we conclude that domain 2a, which was present in all 3D reconstructions of the salt-stable RNP core, contains the major parts of the U2/U6/splicing intermediate RNA network. Given that the catalytic core is present in domain 2a, key functional components of the 35S U5 snRNP, such as the U5 snRNA and domains of Prp8, likely approach the RNA network of the catalytic core from the neighboring upper region of domain 1 (Figure 7F, catalytic core highlighted in pink).

Overall, the spliceosomal C complex assembled on the PM5 pre-mRNA construct offers a tool to study the structural transitions of the spliceosome during its assembly and two-step catalysis in detail. Our 3D cryo-EM maps of the catalytically active C complex and the postspliceosomal 35S U5 snRNP, together with the 3D structures of the salt-stable RNP core, provide a detailed structural description of the step I spliceosome. We believe that future labeling and high-resolution studies to locate its functional components can be based on this structural framework.

## EXPERIMENTAL PROCEDURES

### Purification of the C Complex and Its RNP Core

C complexes were assembled in HeLa nuclear extract on $^{32}$P-labeled PM5 pre-mRNA and purified according to a previously published protocol (Bessonov et al., 2008; Deckert et al., 2006), except that 2 mM ATP was added during elution of the complexes. The purification includes an initial size-fractionation step with glycerol-gradient ultracentrifugation (10%–30% in buffer G150: 20 mM HEPES [pH 7.9], 150 mM NaCl, 1.5 mM $MgCl_2$), followed by affinity selection on amylose beads. For the isolation of C complexes assembled on $MINX_{GG}$, pre-mRNA was incubated with a 20-fold molar excess of purified MS2-MBP fusion protein (Deckert et al., 2006), and a standard splicing reaction containing 15 nM of [$^{32}$P]-labeled $MINX_{GG}$ pre-mRNA was incubated for 70 min. A 30-fold molar excess of DNA oligonucleotide complementary to nucleotides +6 to −6 and −15 to −26 relative to the 5′ ss was added, and the reaction was incubated at 30°C for an additional period of 20 min. Salt-stable RNP core particles were prepared according to a described protocol (Bessonov et al., 2008). RNA was recovered from the eluted complexes, analyzed on a denaturing polyacrylamide gel, and visualized by silver staining and autoradiography. For analysis of their protein composition, spliceosomes were pelleted at 195,000 g for 5 hr. Proteins were recovered, separated by 8/14% SDS-PAGE, and stained with Coomassie blue followed by mass spectrometry (Bessonov et al., 2008). For EM, eluted complexes were subjected to an additional GraFix gradient (Kastner et al., 2008).

### Purification of the 35S U5 snRNP

35S U5 snRNPs were purified from HeLa nuclear extract prepared from a cell line that stably expresses an N-terminally FLAG-tagged version of the protein AD002 (Grote et al., 2010), and particles were isolated by immunochromatography essentially according to a protocol described previously, except that here anti-FLAG antibodies were used instead of the anti-SKIP antibodies used in our previous report (Makarov et al., 2002). For the determination of the protein composition, the purified particles were separated by SDS-PAGE, stained with Coomassie blue, and subjected to mass spectrometry (Bessonov et al., 2008).

### Negative Staining and Unstained Cryo-EM Preparation in Vitrified Ice

For negative staining, glycerol-gradient fractions were used directly for adsorption on a carbon film and stained with ~2% uranyl formate according to the sandwich carbon procedure (Golas et al., 2005), except that specimens were stored and imaged at ambient temperature (Häcker et al., 2008). For cryo-EM in vitrified ice, the buffer of glycerol-gradient fractions was exchanged for a glycerol-free G150 buffer with a Zeba spin column (Pierce, Rockford, IL). Purified particles were adsorbed on a small piece of carbon film for 90–360 min and subsequently mounted on an EM copper grid covered with a perforated carbon film. The specimen was carefully blotted and plunge-frozen in liquid ethane (Adrian et al., 1984).

### Imaging

Using a side-entry cryoholder (Gatan, Pleasanton, CA), tilted image pairs were taken under cryogenic conditions in a CM200FEG (Philips/FEI, Eindhoven, Netherlands) electron microscope on a 4k × 4k TemCam-F415 charge-coupled device (TVIPS, Gauting, Germany) in tile mode (Sander et al., 2005). Tilt pairs were first recorded at a tilt angle of 45° and subsequently at 0°. Negatively stained images were taken in a CM200FEG with the TemCam, except for

the tilt pairs (45°/0°) of the RNP core, which were recorded with a Tecnai T12 (Philips/FEI) electron microscope equipped with a MultiScan 794 CCD camera (Gatan). Details of the data sets are listed in Table S1.

### 2D Image Processing

All single-particle image processing steps at the 2D and 3D level were performed independently and de novo; i.e., no information from one data set was used to refine the respective other data sets. For negative-staining EM, 4,618, 29,382, and 59,380 single-particle images of the PM5 C complex, the salt-stable RNP core, and the 35S U5 snRNP were selected, respectively. For the C complexes assembled on the $MINX_{GG}$ pre-mRNA, 56,202 unstained cryo-EM images taken in vitrified ice were used. All data sets were subjected to iterative rounds of multireference alignment with resampling to polar coordinates (Sander et al., 2003b) and multivariate statistical classification (van Heel, 1984).

### Difference Mapping

Class averages of the respective data sets to be compared were subtracted from each other, the standard deviation of the resulting difference map was determined, and the difference map was color-coded to represent a range from $-3\sigma$ (blue) to $4\sigma$ (red). A contour line was manually drawn around the smaller of the two particles and transferred to the same position in the difference map to indicate its position.

### 3D Structure Determination of the Salt-Stable RNP Core Complex

Negatively stained zero-tilt images (63,045) were subjected to five iterative rounds of alignment and classification; images from 12,913 particles yielded classes of highest visual quality. Of these, 6,331 tilted counterparts were collected. The others were discarded owing to insufficient overlap of the tilt pairs, to their being too close to the edges of the CCD image, or to too high or too low defocus of the tilted images as a consequence of the focus gradient. Two hundred and eleven RCT structures (Radermacher, 1988) were calculated with an average of 30 members per 3D map on the basis of the classification of the untilted images. The best 77 RCT maps were subjected to iterative 3D alignment and weighted averaging, with the missing cone of the RCTs being taken into account (Sander et al., 2010). The aligned 3D maps were finally classified into 3D class averages (Golas et al., 2009; Sander et al., 2006). On average, 5.5 RCT maps contributed to a 3D class average.

### 3D Image Processing of the Cryo-EM Data

Unstained tilt pairs were selected manually (18,152 PM5 C complexes and 7,890 35S U5 snRNPs), and iterative rounds of multireference alignment with resampling to polar coordinates (Sander et al., 2003b) and multivariate statistical analysis (van Heel, 1984) were performed in the context of IMAGIC-5 (van Heel et al., 1996). 3D maps of each class average were calculated according to the RCT method (Radermacher, 1988), and a weighted average of these was determined as for the C core complex, but with unstained cryo tilt pairs. For refinement with unstained cryo-EM, 36,071 images of the PM5 C complex and 21,476 images of the 35S U5 snRNP were selected manually, defocus corrected (Sander et al., 2003a), and subjected to projection matching. Following our corrim-based alignment (Sander et al., 2003b), the iteratively refined reference projections had an angular separation of 0.72° (C complex) and 0.5° (35S U5 snRNP) in the final alignment round. 3D maps are presented as isosurface views (Amira, TGS Europe, Merignac Cedex, France).

### Immunolabeling of the 5′ Exon

Immunolabeling was performed essentially as described previously (Wolf et al., 2009), except that spliceosomes were incubated with the anti-MBP antibodies before loading of the C complexes onto the amylose column.

## ACCESSION NUMBERS

EM maps of the C complex, the 35S U5 snRNP, and its salt-stable RNP core have been deposited in the Electron Microscopy Data Bank under the accession numbers EMD-1846, EMD-1847, and EMD-1848, respectively.

## SUPPLEMENTAL INFORMATION

Supplemental Information includes four figures, one table, and three movies and can be found with this article online at doi:10.1016/j.molcel.2010.11.023.

## ACKNOWLEDGMENTS

We are grateful to T. Conrad, P. Kempkes, H. Kohansal, and Karen Thomsen for expert technical assistance, and we thank H. Urlaub, M. Raabe, and U. Plessmann for their help in MS analysis. We are very grateful to C.L. Will for comments on the manuscript. This work was supported by grants from the Ernst Jung Stiftung and FP6 of the European Union via the NoE EURASNET to R.L., and by grants from the BMBF (0311899) and FP6 of the European Union via "3D Repertoire" to H.S. M.M.G. and B.S. are grateful for support by Danish Center for Scientific Computing, the Agnes og Poul Friis' Foundation, and the Villum Foundation. M.M.G. is supported by the Danish Council for Independent Research and the Danish National Research Foundation. B.S. is grateful for support by the Novo Nordisk Foundation, the A.P. Møller Foundation, and the Fru Astrid Thaysens Foundation. M.G. was supported by the Stiftung Stipendien-Fonds des Verbandes der Chemischen Industrie and E.W. by the Studienstiftung des Deutschen Volkes.

Received: February 10, 2010
Revised: July 28, 2010
Accepted: October 7, 2010
Published: December 21, 2010

## REFERENCES

Adrian, M., Dubochet, J., Lepault, J., and McDowall, A.W. (1984). Cryo-electron microscopy of viruses. Nature *308*, 32–36.

Behzadnia, N., Golas, M.M., Hartmuth, K., Sander, B., Kastner, B., Deckert, J., Dube, P., Will, C.L., Urlaub, H., Stark, H., and Lührmann, R. (2007). Composition and three-dimensional EM structure of double affinity-purified, human prespliceosomal A complexes. EMBO J. *26*, 1737–1748.

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

Boehringer, D., Makarov, E.M., Sander, B., Makarova, O.V., Kastner, B., Lührmann, R., and Stark, H. (2004). Three-dimensional structure of a pre-catalytic human spliceosomal complex B. Nat. Struct. Mol. Biol. *11*, 463–468.

Brow, D.A. (2002). Allosteric cascade of spliceosome activation. Annu. Rev. Genet. *36*, 333–360.

Deckert, J., Hartmuth, K., Boehringer, D., Behzadnia, N., Will, C.L., Kastner, B., Stark, H., Urlaub, H., and Lührmann, R. (2006). Protein composition and electron microscopy structure of affinity-purified human spliceosomal B complexes isolated under physiological conditions. Mol. Cell. Biol. *26*, 5528–5543.

Fabrizio, P., Dannenberg, J., Dube, P., Kastner, B., Stark, H., Urlaub, H., and Lührmann, R. (2009). The evolutionarily conserved core design of the catalytic activation step of the yeast spliceosome. Mol. Cell *36*, 593–608.

Gao, Y.G., Selmer, M., Dunham, C.M., Weixlbaumer, A., Kelley, A.C., and Ramakrishnan, V. (2009). The structure of the ribosome with elongation factor G trapped in the posttranslocational state. Science *326*, 694–699.

Golas, M.M., Sander, B., Will, C.L., Lührmann, R., and Stark, H. (2003). Molecular architecture of the multiprotein splicing factor SF3b. Science *300*, 980–984.

Golas, M.M., Sander, B., Will, C.L., Lührmann, R., and Stark, H. (2005). Major conformational change in the complex SF3b upon integration into the spliceosomal U11/U12 di-snRNP as revealed by electron cryomicroscopy. Mol. Cell *17*, 869–883.

Golas, M.M., Böhm, C., Sander, B., Effenberger, K., Brecht, M., Stark, H., and Göringer, H.U. (2009). Snapshots of the RNA editing machine in trypanosomes captured at different assembly stages in vivo. EMBO J. *28*, 766–778.

Grainger, R.J., and Beggs, J.D. (2005). Prp8 protein: at the heart of the spliceosome. RNA *11*, 533–557.

Grote, M., Wolf, E., Will, C.L., Lemm, I., Agafonov, D.E., Schomburg, A., Fischle, W., Urlaub, H., and Lührmann, R. (2010). Molecular architecture of the human Prp19/CDC5L complex. Mol. Cell. Biol. *30*, 2105–2119.

Häcker, I., Sander, B., Golas, M.M., Wolf, E., Karagöz, E., Kastner, B., Stark, H., Fabrizio, P., and Lührmann, R. (2008). Localization of Prp8, Brr2, Snu114 and U4/U6 proteins in the yeast tri-snRNP by electron microscopy. Nat. Struct. Mol. Biol. *15*, 1206–1212.

Jurica, M.S., and Moore, M.J. (2003). Pre-mRNA splicing: awash in a sea of proteins. Mol. Cell *12*, 5–14.

Jurica, M.S., Licklider, L.J., Gygi, S.R., Grigorieff, N., and Moore, M.J. (2002). Purification and characterization of native spliceosomes suitable for three-dimensional structural analysis. RNA *8*, 426–439.

Jurica, M.S., Sousa, D., Moore, M.J., and Grigorieff, N. (2004). Three-dimensional structure of C complex spliceosomes by electron microscopy. Nat. Struct. Mol. Biol. *11*, 265–269.

Kambach, C., Walke, S., and Nagai, K. (1999). Structure and assembly of the spliceosomal small nuclear ribonucleoprotein particles. Curr. Opin. Struct. Biol. *9*, 222–230.

Kastner, B., Fischer, N., Golas, M.M., Sander, B., Dube, P., Boehringer, D., Hartmuth, K., Deckert, J., Hauer, F., Wolf, E., et al. (2008). GraFix: sample preparation for single-particle electron cryomicroscopy. Nat. Methods *5*, 53–55.

Konarska, M.M., and Query, C.C. (2005). Insights into the mechanisms of splicing: more lessons from the ribosome. Genes Dev. *19*, 2255–2260.

Makarov, E.M., Makarova, O.V., Urlaub, H., Gentzel, M., Will, C.L., Wilm, M., and Lührmann, R. (2002). Small nuclear ribonucleoprotein remodeling during catalytic activation of the spliceosome. Science *298*, 2205–2208.

McConnell, T.S., and Steitz, J.A. (2001). Proximity of the invariant loop of U5 snRNA to the second intron residue during pre-mRNA splicing. EMBO J. *20*, 3577–3586.

Newman, A.J., Teigelkamp, S., and Beggs, J.D. (1995). snRNA interactions at 5′ and 3′ splice sites monitored by photoactivated crosslinking in yeast spliceosomes. RNA *1*, 968–980.

Nilsen, T.W. (1994). RNA-RNA interactions in the spliceosome: unraveling the ties that bind. Cell *78*, 1–4.

Nilsen, T.W. (2003). The spliceosome: the most complex macromolecular machine in the cell? Bioessays *25*, 1147–1149.

Pena, V., Rozov, A., Fabrizio, P., Lührmann, R., and Wahl, M.C. (2008). Structure and function of an RNase H domain at the heart of the spliceosome. EMBO J. *27*, 2929–2940.

Pomeranz Krummel, D.A., Oubridge, C., Leung, A.K., Li, J., and Nagai, K. (2009). Crystal structure of human spliceosomal U1 snRNP at 5.5 A resolution. Nature *458*, 475–480.

Radermacher, M. (1988). Three-dimensional reconstruction of single particles from random and nonrandom tilt series. J. Electron Microsc. Tech. *9*, 359–394.

Ritchie, D.B., Schellenberg, M.J., Gesner, E.M., Raithatha, S.A., Stuart, D.T., and Macmillan, A.M. (2008). Structural elucidation of a PRP8 core domain from the heart of the spliceosome. Nat. Struct. Mol. Biol. *15*, 1199–1205.

Ritchie, D.B., Schellenberg, M.J., and MacMillan, A.M. (2009). Spliceosome structure: piece by piece. Biochim. Biophys. Acta *1789*, 624–633.

Sander, B., Golas, M.M., and Stark, H. (2003a). Automatic CTF correction for single particles based upon multivariate statistical analysis of individual power spectra. J. Struct. Biol. *142*, 392–401.

Sander, B., Golas, M.M., and Stark, H. (2003b). Corrim-based alignment for improved speed in single-particle image processing. J. Struct. Biol. *143*, 219–228.

Sander, B., Golas, M.M., and Stark, H. (2005). Advantages of CCD detectors for de novo three-dimensional structure determination in single-particle electron microscopy. J. Struct. Biol. *151*, 92–105.

Sander, B., Golas, M.M., Makarov, E.M., Brahms, H., Kastner, B., Lührmann, R., and Stark, H. (2006). Organization of core spliceosomal components U5 snRNA loop I and U4/U6 Di-snRNP within U4/U6.U5 Tri-snRNP as revealed by electron cryomicroscopy. Mol. Cell *24*, 267–278.

Sander, B., Golas, M.M., Lührmann, R., and Stark, H. (2010). An approach for de novo structure determination of dynamic molecular assemblies by electron cryomicroscopy. Structure *18*, 667–676.

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

Stark, H., and Lührmann, R. (2006). Cryo-electron microscopy of spliceosomal components. Annu. Rev. Biophys. Biomol. Struct. *35*, 435–457.

van Heel, M. (1984). Multivariate statistical classification of noisy images (randomly oriented biological macromolecules). Ultramicroscopy *13*, 165–183.

van Heel, M., Harauz, G., Orlova, E.V., Schmidt, R., and Schatz, M. (1996). A new generation of the IMAGIC image processing system. J. Struct. Biol. *116*, 17–24.

Wahl, M.C., Will, C.L., and Lührmann, R. (2009). The spliceosome: design principles of a dynamic RNP machine. Cell *136*, 701–718.

Wolf, E., Kastner, B., Deckert, J., Merz, C., Stark, H., and Lührmann, R. (2009). Exon, intron and splice site locations in the spliceosomal B complex. EMBO J. *28*, 2283–2292.

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