nature
structural &
molecular biology

Three-dimensional structure of C complex spliceosomes by electron microscopy

Melissa S Jurica[1–3], Duncan Sousa[1,3], Melissa J Moore[1] & Nikolaus Grigorieff[1]

The spliceosome is a multimegadalton RNA-protein machine that removes noncoding sequences from nascent pre-mRNAs. Recruitment of the spliceosome to splice sites and subsequent splicing require a series of dynamic interactions among the spliceosome's component U snRNPs and many additional protein factors. These dynamics present several challenges for structural analyses, including purification of stable complexes to compositional homogeneity and assessment of conformational heterogeneity. We have isolated spliceosomes arrested before the second chemical step of splicing (C complex) in which U2, U5 and U6 snRNAs are stably associated. Using electron microscopy, we obtained images of C complex spliceosomes under cryogenic conditions and determined a three-dimensional structure of a core complex to a resolution of 30 Å. The structure reveals a particle of dimensions 27 × 22 × 24 nm with a relatively open arrangement of three primary domains.

During eukaryotic gene expression, most nascent transcripts (pre-mRNAs) must be processed to excise noncoding intervening sequences (introns) and ligate coding sequences (exons). This processing, known as pre-mRNA splicing, is catalyzed by a large and dynamic macromolecular complex called the spliceosome. In size and complexity, spliceosomes are comparable to ribosomes. A major difference between them, however, is in the number of stable subcomplexes observable for each. Whereas ribosomes comprise a small and a large subunit that remain relatively unchanged when they join to form an 80S monosome, the spliceosome consists of five major subunits (the spliceosomal snRNPs U1, U2, U4, U5 and U6) that assemble *in vitro* via a complex series of interactions into at least four stable intermediates (E/CC, A, B and C). Separable by native gel electrophoresis or density gradient sedimentation, these subcomplexes probably represent key structural transitions that become stabilized as the splicing machinery proceeds through splice site identification and then catalysis *in vivo*[1]. Whereas the ribosome has been subject to intense structural analysis via electron microscopy and X-ray crystallography, the three-dimensional structure of the spliceosome remains largely undetermined. High-resolution crystal structures are available for only a handful of individual spliceosome components[2–8], and cryo-electron microscopy (cryo-EM) has been used to determine three-dimensional views of only U1 snRNP[9] and SF3b complex (a ~450-kDa protein complex associated with U2 snRNP)[10]. So far, however, no higher-order spliceosomal structures have been presented.

The complexity and dynamic nature of the spliceosome pose challenges to determining its three-dimensional structure. For example, although intermediate complexes have been characterized according to relative size and general makeup, none of them have been subjected to a quantitative compositional analysis. Thus, the stoichiometric presence of protein components in purified C complex identified using mass spectrometry has not yet been determined[11]. Furthermore, a homogeneous preparation of intermediate splicing complexes may not be feasible, even when great care is taken during their isolation. The intermediate states identified as distinct complexes *in vitro* may not occur in isolation *in vivo*[1,12]. Rather, they may manifest themselves as a mere tightening and loosening of interactions, as well as structural rearrangements, within a larger complex that never truly resolves into smaller subcomplexes. Thus, in contrast to the ribosome, the entire spliceosome may not encompass a single rigid structure, but instead may contain several loosely associated protein factors that gather around a more rigid core complex with a defined composition and structure. Our goal in the structural analysis of the spliceosome is, therefore, the identification of the core complex, the part of the structure that remains constant within a population of splicing complexes of a given intermediate functional state, such as the C complex we report on here.

RESULTS

Isolation of C complex spliceosomes

We recently described the purification of *in vitro*–assembled spliceosomes from HeLa cell nuclear extracts aimed at enriching the single structural stage known as C complex (**Fig. 1**)[11]. Using an affinity-tagged pre-mRNA substrate with a mutation at the 3′ splice site, we arrested spliceosomes in their dynamic progression at a point before the second chemical step of splicing. The complexes were then isolated

[1]Howard Hughes Medical Institute, Rosenstiel Basic Medical Sciences Research Center, Department of Biochemistry, Brandeis University, 415 South Street, Waltham, Massachusetts 02454, USA. [2]Center for the Molecular Biology of RNA, Department of Molecular, Cell and Developmental Biology, University of California Santa Cruz, 1156 High Street, Santa Cruz, California 95064, USA. [3]These authors contributed equally to this work. Correspondence should be addressed to M.J. (jurica@biology.ucsc.edu).

Published online 15 February 2004; doi:10.1038/nsmb728

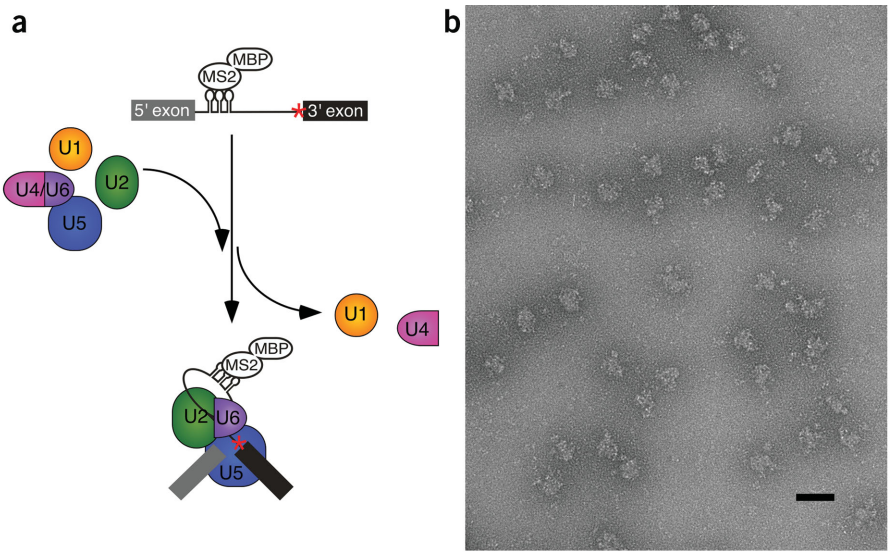

Figure 1 Purification and imaging of spliceosomes. (**a**) Spliceosomes were assembled *in vitro* on an affinity-tagged substrate with a mutation at the 3′ splice site. Splicing is thus arrested after 5′ exon cleavage and before exon ligation. The complexes contain U2, U5 and U6 snRNPs, but not U1 or U4 snRNP. (**b**) Raw images of C complex in cryo-negative stain. Bar, 50 nm.

via the affinity tag under native conditions. Splicing complexes at this stage contain the products of the first step of splicing (free 5′ exon and lariat intermediate) as well as the U2, U5 and U6 snRNPs. U1 and U4 snRNPs are not stably associated with C complex. Previously, we presented evidence for a largely homogeneous preparation of C complex that included direct staining of the RNA components in the sample showing approximately stoichiometric amounts of the substrate intermediates with U2, U5 and U6 snRNAs[11]. Furthermore, these complexes were resolved as a single ~40S peak upon glycerol gradient fractionation.

Cryo-electron microscopy

To obtain the structure of our purified C complex, available in only subpicomole quantities, we used cryo-EM and single particle reconstruction. Initially, we imaged spliceosomes on a thin carbon support in negative stain to ascertain the suitability of our preparation for image analysis[11]. These high-contrast images could be aligned and grouped into distinct classes, but a large fraction of images seemed to be of particles with a preferred orientation (**Fig. 2**). Although it was possible to calculate a structure from the different class averages by the angular reconstitution[13] (data not shown), the quality of the resulting structure could not be readily assessed. Of particular concern was the lack of relatively evenly spaced views from all directions when class averages were assigned angular orientations. The absence of certain views can cause severe artifacts that propagate through subsequent rounds of image processing. Furthermore, the spliceosome lacks symmetry, and, although our samples contained purified C complex, the amount of conformational and compositional heterogeneity in any of the spliceosome subcomplexes is unknown, as discussed in the previous section. Therefore, it is difficult to determine whether different class averages represent different views of complexes with the same conformation, or if the observed differences are due to conformational or compositional variability. These issues pose major challenges to structure determination by angular reconstitution, which depends on evenly spaced views and assumes conformational and compositional homogeneity when assigning orientation angles.

To circumvent these problems, we used random conical tilt[14,15], in which additional views of the complexes are generated during data collection by tilting the specimen. Thus, two of the three angular orientations relating particles to one another are determined experimentally rather than computationally. To obtain high-contrast images of spliceosomes preserved under cryogenic conditions, we froze particles sandwiched between two layers of thin carbon in an aqueous solution containing a small amount of uranyl formate (cryo-negative stain) (**Fig. 1b**)[10]. For structure determination, image pairs were taken of the sample first at 40° tilt and then at 0° tilt. We manually selected 2,247 image tilt pairs and 1,063 additional untilted images from 150 micrographs. Image processing was done primarily with SPIDER[16]. For each micrograph we determined the tilt angle and tilt axis. We then aligned the untilted images and generated 22 classes.

The averages of the two most populated classes, as well as of the three less populated classes (**Fig. 3**, classes 1–5), strongly resemble the predominant view present in negative stain (**Fig. 2a**). One of these classes is shown in **Figure 2b** for direct comparison. The average image of this view shows distinct structural features and has dimensions of ~27 × 25 nm. Most class member particles in the raw images also showed these features and dimensions. These classes also had the highest correlation among individual members and accounted for >40% of the total particles selected.

To further corroborate our assessment that the chosen classes represent a distinct complex with a well-defined structure, we also imaged unstained purified C complex in ice. Although the contrast was much weaker, and we were not able to obtain tilt pairs to carry out a random conical tilt reconstruction, the best class averages of these images showed the same features as class averages from negatively stained samples (**Fig. 2c**). We believe, therefore, that the complex that gives rise to the class averages shown in **Figure 2** and the first five panels of **Figure 3a** represents a stable core structure of C complex.

Of the remaining 17 classes, only two had comparable correlation coefficient statistics and distinct structural features. These two, as well as the previous five classes, were, therefore, chosen for three-dimensional reconstruction (**Fig. 3**). The remaining 15 classes could represent additional complexes that exist in our preparation, indicating heterogeneity. However, these classes have much greater intraclass variability and their occurrence is much less frequent than that represented by the first five classes. These classes and their member particles were, therefore, disregarded in our current analysis. Although we

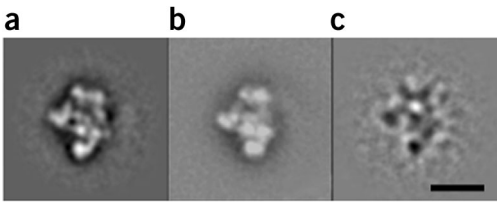

Figure 2 Predominant view of C complex spliceosomes. (**a**) In negative stain. (**b**) In cryo-negative stain. (**c**) In unstained cryogenic conditions. Bar, 20 nm.

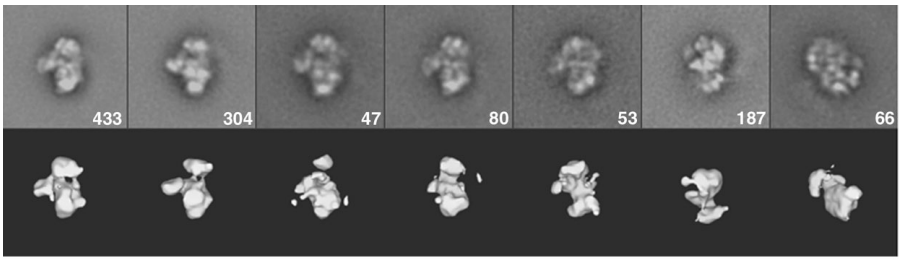

Figure 3 Class averages from untilted images (top) and the structures derived from the corresponding tilted images (bottom). The numbers in the lower left of the class averages indicate the number of class member images. The last two classes result in three-dimensional structures substantially different from the first two structures and were not included in the final structure.

cannot assign the identity of the particles in these classes, they may represent different conformations of C complex, partially disintegrated complexes or impurities in the preparation (for example, the material associated with the small amount (<10%) of unspliced pre-mRNA remaining after complex purification[11]).

Three-dimensional reconstruction

For each of the classes chosen for reconstruction, a structure was calculated from the corresponding tilted images using the known tilt angle and tilt axis (**Fig. 3**). The two most populated classes gave structures with features similar to those produced by angular reconstitution from negatively stained particles. Three additional classes also led to structures that could be aligned with the structures from the first two classes, although they were rather noisy owing to the smaller number of members in each. The remaining two classes together contained only 11% of the total particle images collected and yielded structures similar to each other but completely different from the primary structure. This indicates the presence of a second minor complex that has either a different composition or a different conformation from the primary structure. In the subsequent analysis, we consider only the first complex, which is represented by the first five class averages.

To further improve image alignment and reconstruction, we combined the five similar models and refined their alignment parameters with FREALIGN[17]. Currently, we estimate the resolution of the structure to be 3 nm, based on the Fourier shell correlation of 0.5 between two halves of the data (**Fig. 4**). We contoured the density map of the refined model (**Fig. 5**) so that the densities are continuous and the molecular mass of the entire structure is 2.6 MDa (gray contour).

Contouring the map at a particular density threshold is another challenge in the structural analysis of the spliceosome, as an exact molecular mass of the complex currently cannot be determined. Although

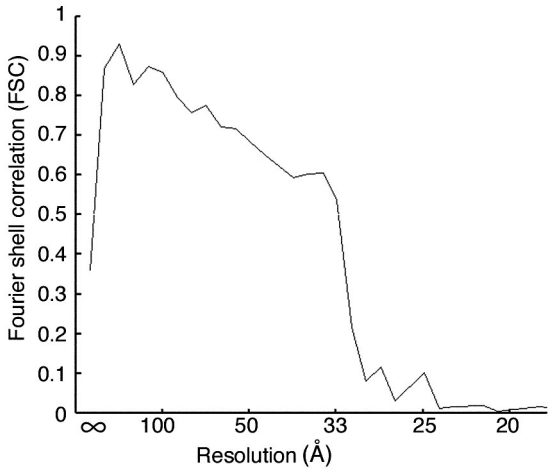

Figure 4 Fourier shell correlation curve indicating a resolution of ~3 nm.

the protein composition of C complex has been analyzed by mass spectrometry[11], those results are not quantitative, as proteins present in substoichiometric amounts are also identified. Indeed, summing the molecular masses of all proteins identified by mass spectrometry yields a much larger particle than would be expected based on either glycerol gradient analysis or our images (on the order of 6 MDa). This discrepancy further supports our assertion that the functional intermediate corresponding to C complex is represented by a core complex, and many additional, more loosely associated protein factors. Furthermore, even for the proteins associated with every complex, only those in the structurally ordered regions of the spliceosome will be visible in the averaged structure. We therefore chose the 2.6-MDa threshold for the structure in **Figure 5** based on the summed molecular masses of U2, U5 and U6 snRNPs and the PRP19 complex.

The PRP19 complex is a set of non-snRNP proteins that joins the spliceosome coincident with B complex formation. The stoichiometric presence of these three snRNPs was verified by staining of the snRNAs. Furthermore, the proteins in all four complexes were consistently identified in mass spectrometry analysis of diverse spliceosome preparations[11,12,18–22], and they are probably associated with the core structure of C complex. Undoubtedly, additional protein factors are also present in stoichiometric amounts and contribute to the averaged structure. In fact, the map can be contoured to include more density that is largely continuous up to ~3.0 MDa. Beyond that, discontinuous noise peaks begin to appear in the map. To more accurately estimate the molecular mass of the spliceosome, a method to quantitatively analyze its many components must be developed.

The final structure consists of three large domains that are more easily visualized as the density threshold is raised slightly (**Fig. 5**). The top domain is roughly cylindrical, with a diameter of 12.6 nm, a height of 8.0 nm and a mass of ~600 kDa. It is clearly separated from the rest of the volume, creating a large cavity in the center of the structure. The rest of the density creates a larger bottom domain with an armlike extension under the central cavity. The bottom domain is ovoid, has dimensions of 19 × 13 × 17 nm and encompasses ~1,500 kDa of density, whereas the arm domain contains ~500 kDa. This open domain arrangement generates a structure very different from that of the densely packed ribosome, perhaps allowing for the structural flexibility necessary for the many dynamic rearrangements that take place during splicing. Such flexibility may also account for the limited resolution attainable in the refinement of the spliceosome structure.

DISCUSSION

The C complex spliceosome analyzed here consists of >100 different protein factors and 4 RNA molecules, and still represents only a transient intermediate in the working cycle of an unusually dynamic molecular machine. Clearly, more work must be done to fully characterize the conformational and compositional variability that may be present in our purified C complex. Nevertheless, we determined the three-dimensional structure of a stable core of C complex at a resolution that

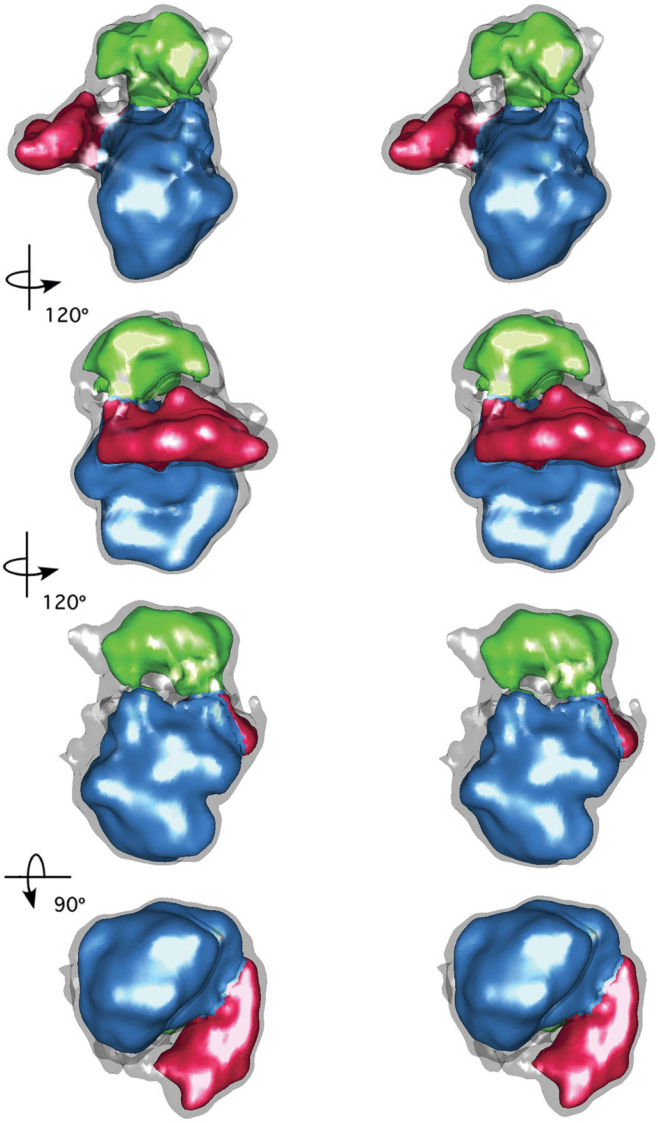

Figure 5 Stereo views of the surface representation of the refined three-dimensional structure. Orientations correspond to rotations (120° or 90°) around the axes indicated. The gray contour represents a complex of ~2.6 MDa. To accentuate the division between the three major domains of the structure (colored regions) the contour level was increased slightly.

allows the identification of three domains. This structure is the first three-dimensional image of a mature splicing complex, and, therefore, is an important step toward a more complete understanding of the inner workings of this machine.

What might the three domains represent? Assuming that the individual snRNP components are largely intact in the complex, one interpretation is based on the molecular mass and the approximate shape of snRNPs seen in previous EM studies. Possible candidates for the three domains include U5 snRNP, with a molecular mass of ~1,000 kDa, U2 snRNP, at ~850 kDa (including the 450 kDa SF3b subunit and 350 kDa SF3a subunit), and the PRP19 complex, at ~550 kDa. U6 snRNP is unlikely to form a separate recognizable domain because it has a minimal molecular mass of only ~113 kDa. In addition, U6 snRNP is in close contact with U2 snRNP, because the U2 and U6 snRNAs are highly base-paired at this stage of spliceosome dynamics[23].

Images of 20S U5 snRNP in negative stain indicate a particle of 21 × 12 nm (ref. 24). The large bottom domain in the spliceosome has a similar size and shape, and could contain the calculated molecular mass of U5 snRNP. Recently, a 35S complex containing primarily U5 snRNP and PRP19 complex proteins has been identified[25]. This complex has been proposed to be the form of the U5 snRNP released after splicing and suggests a direct interaction between the two complexes. Thus, the PRP19 complex could account for the remaining density in the large bottom domain. Alternatively, the PRP19 complex is also of an appropriate molecular mass to constitute the arm density.

Negatively stained images of the 17S U2 snRNP reveal a two-lobed particle, with each lobe of diameter ~10 nm (refs. 24,26). We do not observe a two-lobed structure of similar size for U2 snRNP in the C complex structure, but rearrangements to incorporate base pairing with U6 snRNA could change the overall conformation of this particle[23]. One lobe of the 17S U2 snRNP was proposed to be the protein complex SF3b[26], of which a three-dimensional structure has recently been determined[10]. That structure reveals a particle of dimensions 15 × 12.5 × 16 nm, including several protuberances extending from a globular core. The globular core of SF3b has dimensions very similar to the top domain of the C complex. The second lobe of the 17S U2 snRNP has been proposed to contain the SF3a complex and an Sm core. In the micrographs, the two lobes are connected by a very thin RNase-sensitive filament, presumably U2 snRNA[26]. If this RNA tether is flexible then the mass of SF3a and the Sm core could easily be placed into the C complex arm domain. Indeed, the arm is connected to the top domain by a very thin density. Alternatively, this part of the U2 snRNP may extend into the bottom domain and thus be interacting with U5 snRNP. Specific labels for the individual spliceosome components will be necessary to confirm the identity of these domains.

Although our preparation of purified C complex may contain other, less populated conformations of this intermediate that reflect the native heterogeneity of this machine, our structure represents the major conformation of C complex that can be observed in negative stain, cryo-negative stain and unstained preparations in ice. Further structural and mechanistic studies of this complex, as well as other stable subcomplexes, will be necessary for a more detailed understanding of this complicated macromolecular machine.

METHODS

Sample preparation. C complex spliceosomes were purified as described[11]. Grids were prepared using the cryo-stain sandwich technique[10]. Briefly, thin carbon was floated for 1 min on a sample of 10 nM spliceosomes in 150 mM KCl, 20 mM Tris, pH 7.5, 5 mM EDTA, 10 mM maltose and 5% (v/v) glycerol. The carbon was transferred onto a solution of 1% (w/v) uranyl formate, sample side down, and then picked up with a holey carbon grid (Quantifoil). A second layer of thin carbon was sandwiched on top of the sample and the grid was frozen and stored in liquid nitrogen.

Data collection. Tilt pair images of C complex at 40° and 0° tilt were obtained using a Philips CM12 transmission electron microscope operating at 120 kV and using 1 μm average defocus, 60,000× magnification, and a dose of >10 e^{-} Å^{-2}. Micrographs were digitized using a Zeiss SCAI scanner at a 7-μm step size. 4 × 4 pixel averaging was applied to the images, resulting in a pixel size of 4.67 Å at the specimen level. CTFFIND3 (ref. 27) was used to identify astigmatic images, which were discarded.

Image processing. Single particles were selected manually from the untilted and tilted images using XIMDISP[28] and WEB[16]. The images were processed using SPIDER[16]. First, particles were band-pass-filtered and centered. The untilted particles were further aligned using an iterative process consisting of multivariate statistical analysis, k-means classification, and multireference alignment (MRA), and resulted in 22 classes. The three largest classes, and four smaller classes that clearly resembled at least one of the large classes

(for example, the first and fourth class in **Fig. 2**), were selected for subsequent three-dimensional reconstruction.

Three-dimensional spliceosome structures were generated via random conical tilt[14,15] for the seven selected classes. Each of the structures was refined iteratively by centering the tilted images against projections of their respective density maps and then generating a new refined structure. Five of the seven refined structures were then merged after three-dimensional alignment based on their similarity. The images associated with the remaining structures were excluded from further analysis.

A combined structure was generated in FREALIGN[17] for contrast transfer function (CTF) correction assuming 20% amplitude contrast. The combined structure was calculated from both tilted particles and the corresponding untilted particles. The defocus for each particle image was determined using CTFTILT[27]. Initial refinement of particle shift and in-plane rotation using FREALIGN was done using data extending to a resolution of 50 Å. The resolution of the data included in subsequent refinement cycles was incrementally increased until no further improvement in the resolution of the reconstruction was observed. The resolution of the reconstruction was determined by measuring the Fourier shell correlation between reconstructions from two halves of the data (**Fig. 4**). The resolution at a Fourier shell correlation of 0.5 was 3 nm. For volume rendering, the volume enclosed at a certain density threshold was converted to a molecular mass assuming a protein density of 0.81 Da $Å^{-3}$ (ref. 29).

Coordinates. A density map of the structure has been deposited in the Macromolecular Structure Database at the European Bioinformatics Institute (accession code EMD-1062).

ACKNOWLEDGMENTS

We thank T. Walz and M. Ohi for discussion and H. Stark for advice on sample preparation. N.G. is an assistant investigator and M.J.M is an associate investigator with the Howard Hughes Medical Institute; M.S.J. is a Paul Sigler/Agouron Institute fellow of the Helen Hay Whitney Foundation; D.R.S. is supported by a US National Science Foundation integrative graduate education and research training grant. This work was supported by US National Institutes of Health grant 1 P01 GM-62580 (N.G.) and GM 53007 (M.J.M.) and funding from the Keck Foundation.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Received 2 September; accepted 29 December 2003
Published online at http://www.nature.com/natstructmolbiol/

1. Nilsen, T.W. The spliceosome: no assembly required? *Mol. Cell* **9**, 8–9 (2002).
2. Jiang, J., Horowitz, D.S. & Xu, R.M. Crystal structure of the functional domain of the splicing factor Prp18. *Proc. Natl. Acad. Sci. USA* **97**, 3022–3027 (2000).
3. Kambach, C. *et al*. Crystal structures of two Sm protein complexes and their implications for the assembly of the spliceosomal snRNPs. *Cell* **96**, 375–387 (1999).
4. Kielkopf, C.L., Rodionova, N.A., Green, M.R. & Burley, S.K. A novel peptide recognition mode revealed by the X-ray structure of a core U2AF35/U2AF65 heterodimer. *Cell* **106**, 595–605 (2001).
5. Oubridge, C., Ito, N., Evans, P.R., Teo, C.H. & Nagai, K. Crystal structure at 1.92 Å resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature* **372**, 432–438 (1994).
6. Price, S.R., Evans, P.R. & Nagai, K. Crystal structure of the spliceosomal U2B″-U2A′ protein complex bound to a fragment of U2 small nuclear RNA. *Nature* **394**, 645–650 (1998).
7. Vidovic, I., Nottrott, S., Hartmuth, K., Luhrmann, R. & Ficner, R. Crystal structure of the spliceosomal 15.5kD protein bound to a U4 snRNA fragment. *Mol. Cell* **6**, 1331–2342 (2000).
8. Reuter, K., Nottrott, S., Fabrizio, P., Luhrmann, R. & Ficner, R. Identification, characterization and crystal structure analysis of the human spliceosomal U5 snRNP-specific 15 kD protein. *J. Mol. Biol.* **294**, 515–525 (1999).
9. Stark, H., Dube, P., Luhrmann, R. & Kastner, B. Arrangement of RNA and proteins in the spliceosomal U1 small nuclear ribonucleoprotein particle. *Nature* **409**, 539–542 (2001).
10. Golas, M.M., Sander, B., Will, C.L., Luhrmann, R. & Stark, H. Molecular architecture of the multiprotein splicing factor SF3b. *Science* **300**, 980–984 (2003).
11. Jurica, M., Licklider, L., Gygi, S., Grigorieff, N. & Moore, M. Purification and characterization of native spliceosomes suitable for three-dimensional structural analysis. *RNA* **8**, 426–439 (2002).
12. Stevens, S.W. *et al*. Composition and functional characterization of the yeast spliceosomal penta-snRNP. *Mol. Cell* **9**, 31–44. (2002).
13. Van Heel, M. Angular reconstitution: a posteriori assignment of projection directions for 3D reconstruction. *Ultramicroscopy* **21**, 111–123 (1987).
14. Radermacher, M., Wagenknecht, T., Verschoor, A. & Frank, J. Three-dimensional reconstruction from a single-exposure, random conical tilt series applied to the 50S ribosomal subunit of *Escherichia coli*. *J. Microsc.* **146 (Pt 2)**, 113–136 (1987).
15. Frank, J. & Radermacher, M. Three-dimensional reconstruction of single particles negatively stained or in vitreous ice. *Ultramicroscopy* **46**, 241–262 (1992).
16. Frank, J. *et al*. SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J. Struct. Biol.* **116**, 190–199 (1996).
17. Grigorieff, N. Structure of the respiratory NADH:ubiquinone oxidoreductase (complex I). *Curr. Opin. Struct. Biol.* **9**, 476–483 (1999).
18. Ajuh, P. *et al*. Functional analysis of the human CDC5L complex and identification of its components by mass spectrometry. *EMBO J.* **19**, 6569–6581. (2000).
19. Neubauer, G. *et al*. Mass spectrometry and EST-database searching allows characterization of the multi-protein spliceosome complex. *Nat. Genet.* **20**, 46–50. (1998).
20. Ohi, M.D. *et al*. Proteomics analysis reveals stable multiprotein complexes in both fission and budding yeasts containing Myb-related Cdc5p/Cef1p, novel pre-mRNA splicing factors, and snRNAs. *Mol. Cell. Biol.* **22**, 2011–2024. (2002).
21. Rappsilber, J., Ryder, U., Lamond, A.I. & Mann, M. Large-scale proteomic analysis of the human spliceosome. *Genome Res.* **12**, 1231–1245. (2002).
22. Zhou, Z., Licklider, L.J., Gygi, S.P. & Reed, R. Comprehensive proteomic analysis of the human spliceosome. *Nature* **419**, 182–185. (2002).
23. Staley, J.P. & Guthrie, C. Mechanical devices of the spliceosome: motors, clocks, springs, and things. *Cell* **92**, 315–326 (1998).
24. Kastner, B., Bach, M. & Luhrmann, R. Electron microscopy of small nuclear ribonucleoprotein (snRNP) particles U2 and U5: evidence for a common structure-determining principle in the major U snRNP family. *Proc. Natl. Acad. Sci. USA* **87**, 1710–1714 (1990).
25. Makarov, E.M. *et al*. Small nuclear ribonucleoprotein remodeling during catalytic activation of the spliceosome. *Science* **298**, 2205–2208 (2002).
26. Behrens, S.E., Tyc, K., Kastner, B., Reichelt, J. & Luhrmann, R. Small nuclear ribonucleoprotein (RNP) U2 contains numerous additional proteins and has a bipartite RNP structure under splicing conditions. *Mol. Cell. Biol.* **13**, 307–319 (1993).
27. Mindell, J.A. & Grigorieff, N. Accurate determination of local defocus and specimen tilt in electron microscopy. *J. Struct. Biol.* **142**, 334–347 (2003).
28. Crowther, R.A., Henderson, R. & Smith, J.M. MRC image processing programs. *J. Struct. Biol.* **116**, 9–16 (1996).
29. Matthews, B.W. Solvent content of protein crystals. *J. Mol. Biol.* **33**, 491–497 (1968).