

Published in final edited form as:

*Nature*. 2016 February 18; 530(7590): 298–302. doi:10.1038/nature16940.

## CryoEM structure of the yeast U4/U6.U5 tri-snRNP at 3.7 Å resolution

Thi Hoang Duong Nguyen<sup>#</sup>, Wojciech P. Galej<sup>#</sup>, Xiao-chen Bai, Chris Oubridge, Andrew J. Newman, Sjors H. W. Scheres, and Kiyoshi Nagai

MRC Laboratory of Molecular Biology Francis Crick Avenue Cambridge CB2 0QH UK

<sup>#</sup> These authors contributed equally to this work.

### Abstract

U4/U6.U5 tri-snRNP represents a substantial part of the spliceosome before activation. A cryoEM structure of *Saccharomyces cerevisiae* U4/U6.U5 tri-snRNP at 3.7Å resolution led to an essentially complete atomic model comprising 30 proteins plus U4/U6 and U5 snRNAs. The structure reveals striking interweaving interactions of the protein and RNA components including extended polypeptides penetrating into subunit interfaces. The invariant ACAGAGA sequence of U6 snRNA, which base-pairs with the 5'-splice site during catalytic activation, forms a hairpin stabilised by Dib1 and Prp8 while the adjacent nucleotides interact with the exon binding loop 1 of U5 snRNA. Snu14 harbours GTP but its putative catalytic histidine is held away from the  $\gamma$ -phosphate by hydrogen bonding to a tyrosine in Prp8's N-terminal domain. Mutation of this histidine to alanine has no detectable effect on yeast growth. The structure provides important new insights into the spliceosome activation process leading to the formation of the catalytic centre.

---

Pre-mRNA splicing is catalyzed by an intricate molecular machine called the spliceosome and proceeds by a two-step trans-esterification mechanism, analogous to group II intron self-splicing<sup>1</sup>. The spliceosome is assembled on pre-mRNA by the ordered addition of small nuclear ribonucleoprotein particles (snRNPs) and numerous proteins including nineteen nuclear complex (NTC) and, nineteen related (NTR) complex<sup>2-4</sup>. Initially U1 and U2 snRNPs recognise the pre-mRNA 5'-splice site (5'SS) and branch point (BP), respectively. Recruitment of U4/U6.U5 tri-snRNP produces the fully assembled but catalytically inactive complex B<sup>1</sup>. U1 snRNP is displaced from the 5'SS by Prp28 (ref. 5), the 5'SS pairs with the

---

Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use: [http://www.nature.com/authors/editorial\\_policies/license.html#terms](http://www.nature.com/authors/editorial_policies/license.html#terms)

Correspondence and requests for materials should be addressed to T.H.D.N. ([knguyen@mrc-lmb.cam.ac.uk](mailto:knguyen@mrc-lmb.cam.ac.uk)), W.P.G. ([wgalej@mrc-lmb.cam.ac.uk](mailto:wgalej@mrc-lmb.cam.ac.uk)) or K.N. ([kn@mrc-lmb.cam.ac.uk](mailto:kn@mrc-lmb.cam.ac.uk)).

#### Author contributions

T.H.D.N. purified yeast tri-snRNP and prepared EM grids. T.H.D.N., W.P.G and X.-C.B collected all EM images. T.H.D.N. processed data and calculated the maps with the help of X.-C.B and S.H.W.S. T.H.D.N., W.P.G. and C.O. built a model into the map and refined the structure. T.H.D.N. performed mutagenesis experiments with the help of A.J.N. T.H.D.N. and W.P.G prepared all illustrations. K.N. initiated and orchestrated the project. T.H.D.N., W.P.G. and K.N. analysed the structure and wrote the paper with invaluable contributions from all other authors.

The cryo-EM map have been deposited in the Electron Microscopy Data Bank with accession codes EMD-8006 to EMD-8014. The coordinates of the atomic models have been deposited in the Protein Data Bank under accession code 5GAN (Overall), 5GAP (Body domain), 5GAO (Head domain) and 5GAM (Foot domain).

The authors declare no competing financial interests.

ACAGAGA sequence in U6 snRNA, and Brr2 helicase unwinds the extensively base-paired U4/U6 snRNAs to release U4 snRNA with its associated proteins<sup>6,7</sup>. This allows U6 snRNA to base-pair with U2 snRNA generating the group II intron-like catalytic RNA core<sup>8-11</sup>. The 2'OH group of the BP adenosine attacks the 5'SS, producing exon1 and lariat intron-exon2 intermediates, and after further remodeling to complex C\*, U5 snRNA loop 1 aligns exons 1 and 2 for the second trans-esterification<sup>12,13</sup>. The spliced mRNA product is released and the residual Intron Lariat Spliceosome (ILS) is disassembled, recycling the snRNPs for subsequent rounds of splicing<sup>1</sup>.

Electron microscopic studies of spliceosomes at different stages of the splicing cycle revealed low-resolution pictures of these complexes<sup>14</sup>. Taking advantage of the recent revolution in cryoEM single particle analysis<sup>15</sup> we reported the organisation of the proteins, and U5 snRNA and U4/U6 snRNAs in *S. cerevisiae* U4/U6.U5 tri-snRNP based on cryoEM map at 5.9 Å<sup>16</sup>. In *Schizosaccharomyces pombe* cell extract, ILS complexes containing U2, U5 and U6 snRNAs accumulate through inefficient disassembly<sup>17,18</sup>. The structure of this endogenous U2.U6.U5 spliceosomal complex was determined by single particle cryoEM at 3.6 Å resolution<sup>19,20</sup>. This was an important breakthrough that revealed the overall architecture of a spliceosomal complex with the striking structures of NTC and NTR<sup>19,20</sup>. The absence of spliced mRNA and step 2 factors<sup>4</sup> from this complex<sup>19</sup> confirms that it is the post-splicing ILS<sup>18</sup>. The structure also revealed the important features of the well-established group II intron-like catalytic RNA core<sup>8-11</sup> remaining after spliced mRNA is released<sup>20</sup>.

Here we present an essentially complete atomic model of *S. cerevisiae* U4/U6.U5 tri-snRNP<sup>21,22</sup> based on a cryoEM density map at 3.7 Å overall resolution, revealing the architectural and mechanistic principles of spliceosome activation.

## Overall structure

We collected a new dataset on a Titan Krios microscope using the Gatan K2 Summit direct electron detector (Methods). The overall resolution of the tri-snRNP map was improved from 5.9 Å to 3.7 Å (Extended Data Fig. 1). Using a modified masked refinement with signal subtraction<sup>23</sup>, we obtained more homogeneous 3.6, 3.7 and 4.2 Å reconstructions for the Body, Foot and Head domains, respectively and improved the resolution of the Arm domain from 10 Å to 6-7.5 Å (Extended Data Fig. 2). The new maps enabled us to build a near-complete atomic model of the yeast tri-snRNP containing 30 proteins, U4/U6 and U5 snRNAs (Fig. 1) revealing an amazing web of interactions between components of the complex (Extended Data Fig. 3).

## Prp8

A complete atomic model of Prp8 is now built except the unstructured N-terminus and inter-domain linkers. The  $\alpha$ -helix ( $\alpha$ RT1) at the N-terminus of the reverse transcriptase (RT) domain in the crystal structure<sup>24</sup> extends further and forms a helix bundle (HB) with three additional long helices appended to the RT domain (Fig. 2a and 2b). Residues 108-733 form a predominantly  $\alpha$ -helical N-terminal domain. Stems I and II of U5 snRNA are coaxially

stacked<sup>16</sup> and an extra variable stem protrudes from the three-way junction (Extended Data Fig. 4). A long slightly bent C-terminal  $\alpha$ -helix (residues 703-735) of the N-terminal domain fits into the minor groove of the co-axially stacked stems I and II, which is tightly harnessed in the major groove by a polypeptide loop (residues 535-543) protruding from the N-terminal domain (Fig. 2c). The conserved loop 1 of U5 snRNA, which aligns the exons during the second trans-esterification reaction<sup>12,13</sup>, points towards the most positively charged and conserved surface of Prp8 in the Thumb/Linker domain, part of the active site cavity<sup>24</sup>. The BP+2 nucleotide cross-links in active spliceosomes between Prp8 residues 1585-1598, on the cavity surface (C. M. Norman and A.J.N., unpublished observations). This region is disordered in the Prp8-Aar2 complex<sup>24</sup> whereas in U4/U6.U5 tri-snRNP it forms a helix-turn-helix (the  $\alpha$ -finger) and contacts U54-U55 of U4 snRNA near the three-way junction (Fig. 2b).

The 5'-stem-loop of U6 snRNA interacts with the N-terminal domain of Prp8 and the adjacent single-stranded region pairs with the exon binding U5 snRNA loop 1 (Fig. 2d). The small highly conserved protein Dib1 (ref. 25) binds to the helix bundle and  $\alpha$ -finger of Prp8, and a long polypeptide of Prp31. U6 snRNA forms a short stem-loop, involving part of the ACAGAGA sequence, which is sandwiched between Dib1 and the Prp8 Large domain (residues 1648-1653) (Fig. 2d and e; Extended Data Fig. 4a)

## Snu114

We built a near complete atomic model of Snu114 comprising five domains (D1-D5) similar to EF-G/EF-2 (ref. 26,27). The relative arrangement of D1-D3 closely resembles that of EF-G/EF-2, whereas D4 and D5 pack more compactly (Fig. 3a). The guanine nucleotide density is consistent with GTP bound via canonical interactions with the surrounding residues (Fig. 3b; Extended Data Fig. 3a and 5a-e). In most GTPases the glutamine residue in switch II loop places a water molecule at the  $\gamma$ -phosphate of GTP and hydrolyses the phosphate ester<sup>28</sup>. As in EF-Tu, EF-G and their eukaryotic counterparts, the catalytic glutamine residue is replaced by histidine in Snu114 (ref. 26) (Extended Data Fig. 5e). In U4/U6.U5 tri-snRNP, His218 is hydrogen-bonded to Tyr403 of Prp8 preventing the His218 side chain from rotating towards the  $\gamma$ -phosphate of GTP and hence keeping the GTPase inactive (Fig. 3c). In EF-G and EF-Tu, GTP is hydrolysed when this histidine is repositioned by a hydrogen-bond with a phosphate in the sarcin-ricin loop of the ribosome<sup>29,30</sup> (Fig. 3d). The extensive interactions between Snu114 and the N-terminal domain of Prp8 are conserved between U4/U6.U5 tri-snRNP and the *S. pombe* ILS<sup>19</sup> (Extended Data Fig. 5f). The hydrogen-bond between His218 of Snu114 and Tyr403 of Prp8 is maintained by the equivalent residues in ILS<sup>19</sup> (Extended data Fig. 5d). The GTP binding site of Snu114 is at the interface with the N-terminal domain of Prp8 leaving insufficient room for U5 snRNA or any proteins to access the GTPase active site and act like the sarcin-ricin loop<sup>29,30</sup> or GTPase activating protein (GAP)<sup>28</sup>. Since the structure suggests no obvious mechanism for Snu114 GTPase activation we investigated the function of Snu114 by mutagenesis. With the His218Arg mutation yeast shows only a mild temperature-sensitive phenotype, confirming earlier results<sup>31</sup> (Extended Data Fig. 5g) whereas the equivalent mutation in EF-Tu reduces cognate tRNA-induced GTPase activity 10<sup>5</sup>-fold<sup>32</sup>. Surprisingly yeast containing the His218Ala mutant of Snu114 shows no apparent phenotype (Extended Data Fig. 5g) whilst the

equivalent mutation in EF-Tu reduces the rate of GTP hydrolysis more than  $10^6$ -fold<sup>32</sup>. Furthermore mutations of Tyr403 (Tyr403Phe and Tyr403Ala) in Prp8, which hydrogen-bonds with His218 in Snu114, have no apparent phenotype (Extended Data Fig. 5h). These results raise the possibility that Snu114-bound GTP may not be hydrolysed during splicing.

The guanine nucleotide in the post-splicing ILS is interpreted as GDP<sup>19</sup> but its conformation is distinct from that of GDP in other GTPases (Extended Data Fig. 5a). In contrast the conformation of the Snu114-bound GTP in tri-snRNP superimposes well with GTP or non-hydrolysable GTP analogues in other GTPases (Extended Data Fig. 5c). When we refined our structure of Snu114 with GDP the resulting GDP conformation is similar to that observed in ILS (Extended Data Fig. 5b). The proposed guanine nucleotide-dependent regulatory role of Snu114 is based on the effects of XDP, XTP and a non-hydrolysable XTP analogue on the XTP binding mutant (D271N) of Snu114 (ref. 33,34). Mutations in the GTPase domain prevent the interaction of Snu114 with Prp8, blocking U5 snRNP assembly<sup>35</sup>. The observed effect of XDP, XTP and non-hydrolysable XTP may be due to XTP-induced stabilization and association of Snu114 mutants with Prp8.

## The U4/U6 di-snRNP

The extensively base-paired U4/U6 snRNAs form a three-way helix junction (Extended Data Fig. 4a and 4b). Snu13, bound to the k-turn motif<sup>36</sup>, is wedged between U4 5'-stem-loop and the U4/U6 snRNA helix II and packed against the Prp4 WD40 domain<sup>37</sup> (Fig. 4a). Prp3 makes extensive interactions with the Prp4 WD40 domain, the basket handle-like structure and Snu13, and forms a long  $\alpha$ -helix sitting in the minor groove of U4/U6 helix II. After forming a short  $\alpha$ -helix, Prp3 folds back to form a long  $\alpha$ -helix binding across the major groove of U4/U6 helix II (Fig. 4b; Extended Data Fig. 3d). These latter two Prp3 helices and the connecting loop interact extensively with the RNaseH-like domain of Prp8 and Brr2 N-terminal domain. Prp3 further extends to form a ferredoxin-like domain<sup>38</sup>, which packs against the Prp4 WD40 domain<sup>37</sup>. Masked classification of the Arm domain reveals extra density for the N-terminus of Prp3 extending towards the LSm protein ring (Extended Data Fig. 6a and 6b). The 3'-end of U6 snRNA binds to the central hole of the LSm protein ring while the preceding single-stranded region binds to the ferredoxin-like domain of Prp3 (ref. 38). The Nop and coiled-coil domains of Prp31 interact with Snu13 whereas the k-turn motif of U4 5'-stem-loop is sandwiched between Snu13 and Prp31 (ref. 36,39) (Fig. 4c). The extended polypeptide chain of Prp31 runs between the phosphate backbone of U4 5'-stem and Dib1, and forms a small domain together with Prp6 which is surrounded by the three-way RNA helix junction and the  $\alpha$ -finger and helix bundle of Prp8 (Fig. 4d).

The C-terminus of the Prp6 TPR repeats<sup>40</sup> interacts with the Prp4 WD40 domain, Snu13, Prp31 and the tip of U4 5'-stem-loop (Fig. 4e) while an extended N-terminal polypeptide of Prp6 packs against the RNaseH-like domain of Prp8 and interacts with the small C-terminal domain of Prp31, the Prp8  $\alpha$ -finger, and U4/U6 snRNA three-way junction and then wraps around the Prp8 helix bundle (Fig. 4d; Extended Data Fig. 3d-f). The numerous interactions that Prp6 makes with U4/U6 snRNP components and Prp8 reflect its importance for tri-snRNP assembly<sup>41</sup>.

## Brr2

The single-stranded region of U4 snRNA (Extended Data Fig. 3b and 4a) extending from stem I, enters the active site of Brr2 N-terminal helicase cassette near the strand-separating  $\beta$ -hairpin and passes through the channel between the RecA1, RecA2, Ratchet and WH domains<sup>42</sup> (Extended Data Fig. 7a-c). The N-terminal domain (NTD) of Brr2 extends towards U4/U6 stem II and contacts the long helix of Prp3 running along the phosphate backbone of U4 snRNA. Brr2 inserts a loop of the NTD into the minor groove of U4/U6 stem II (Extended Data Fig. 7b and 7d). These interactions may guide U4/U6 stem II during unwinding. Snu13, Prp4 WD40, Prp3 ferredoxin, and Prp31 Nop and coiled-coil domains assemble together while the long  $\alpha$ -helices and stretched polypeptide chains of Prp3 and Prp31 extend from these domains and interact with U4/U6 stem II and U4 5'-stem-loop, respectively<sup>39</sup>. These long  $\alpha$ -helices and extended polypeptides may function like elastic bands to accommodate conformational changes and partial strand separation of the U4/U6 duplex as Brr2 translocates along U4 snRNA and unwinds U4/U6 stem I (ref. 16,43). Brr2 forms a stable complex with the Jab1/MPN domain of Prp8 (ref. 42), which is attached to the RNaseH-like domain of Prp8 via a long flexible linker, enabling both Brr2 and U4/U6 di-snRNP to detach from the main body of Prp8 during unwinding.

The improved map of the Head domain at 4.5-5Å resolution, obtained by masked refinement, enabled us to build most of the Snu66 structure as poly-Ala chains. Its N-terminal region forms a globular domain that interacts with Prp8 endonuclease-like and Brr2 N-terminal ratchet domains. This is followed by a long helix wedged between Prp8 Jab1/MPN and Brr2 N-terminal HLH domains while its C-terminus wraps around Brr2, forming extensive interactions with the Brr2 C-terminal cassette (Extended Data Fig. 7e), fully consistent with yeast two-hybrid and co-immunoprecipitation assays<sup>44</sup>. Interestingly, our global classification approach showed “open” and “closed” conformations of the Head and Foot domains (Extended Data Fig. 6c-e). In the “closed” conformation, the globular domain of Snu66 contacts the N-terminal domain of Prp8, which in turn interacts with Snu114.

## Insight into spliceosome activation

A comparison of the U4/U6.U5 tri-snRNP with B<sup>45</sup> and B<sup>U1</sup> complexes<sup>46</sup> shows that U2 snRNP docks with tri-snRNP where the LSm complex, Prp3 and Prp6 are located, whilst U1 snRNP sits on top of U2 snRNP (Fig. 5a). The components of NTC/NTR are also detected by mass spectrometry in complex B<sup>3</sup>. We compared the structures of our tri-snRNP and the post-splicing ILS<sup>19</sup> by overlaying the large domain of Prp8 together with Snu114 and the U5 core domain. This shows that NTC and NTR can associate with tri-snRNP without clashing and contact U2 snRNP (Fig. 5a and 5b). In complex B, U2 snRNP interacts with U4/U6.U5 tri-snRNP<sup>46</sup>, but when NTC and NTR dock with tri-snRNP, U2 snRNP is passed to NTC and NTR, and U2 Sm domain and U2B''/U2A' complex associate with Aquarius(*Cwf11*), Syf1(*Cwf3*) and Isy1(*Cwf12*)<sup>47</sup> as revealed in the *S. pombe* ILS<sup>19</sup> (*S. pombe* protein names are shown in italics in parenthesis).

The *S. cerevisiae* U4/U6.U5 tri-snRNP and *S. pombe* ILS structures reveal that the Foot domains of the two structures, containing the Prp8 N-terminal domain, U5 snRNA stem-loop 1 and Snu114, superpose very well showing that they form a stable structural unit (Extended Data Fig. 5f). Overlay of their Prp8 large domains shows that the Foot domain rotates as a rigid body, by 30° between the two structures, causing U5 loop 1 to move closer toward the Prp8  $\alpha$ -finger in the post-splicing ILS<sup>19</sup> (Fig. 5c). NTC forms extensive interfaces with both the N-terminal and Large domains of Prp8, hence the rotation of the Foot domain may be caused by NTC. When the Foot domain of the U4/U6.U5 tri-snRNP structure rotates by 30° (as in the post-splicing ILS) Prp8 residues 602-614 clash with Dib1 and ACAGAGA helix, forcing Dib1 to dissociate from the large domain of Prp8 and liberating the ACAGAGA sequence to bind the 5'SS. It is known from the U4 snRNA cs1 mutation and its suppressor in U6 snRNA (U6-Dup)<sup>48</sup> that the pairing between 5'SS and the ACAGAGA sequence is a checkpoint for the unwinding of the U4/U6 snRNA duplex by Brr2. Thus, conformational toggling of the Prp8 N-terminal domain could couple 5'SS recognition to U4/U6 unwinding. Suppressors of the U4-cs1 mutation<sup>49</sup> suggest the allosteric changes required for Brr2 activation. Interestingly these suppressors form four clusters in Prp8: one at the interface between the RT domain and D3 of Snu114, one at the interface between the helix bundle and Prp31, one on the surface of endonuclease-like domain near the ACAGAGA hairpin and one on the surface of the N-terminal domain where the 5'-stem-loop of U6 snRNA binds, near the interface with Snu66 that undergoes a transition between the Open and Closed forms (Extended Data Fig. 6c-e). It is possible that NTC/NTR play important roles in inducing allosteric changes that trigger the unwinding of U4/U6 snRNA by Brr2 (ref. 50).

The cryo-EM structures of *S. cerevisiae* U4/U6.U5 tri-snRNP and *S. pombe* ILS<sup>19</sup> have provided a wealth of new information about the architecture and conformational changes of these spliceosomal assemblies. Functional studies based on these new structural insights should greatly enhance our understanding of spliceosome activation and catalysis.

## Methods

### Statistics

No statistical methods were used to predetermine sample size.

### Sample preparation

Tri-snRNP sample was prepared as described in our published protocol<sup>16</sup>.

### Electron microscopy

Aliquots of 3.5  $\mu$ l of purified yeast tri-snRNP were applied to Quantifoil Cu R1.2/1.3, 400 mesh grids which were coated with 6 nm-thick homemade carbon film and glow-discharged in *N*-amylamine. The grids were blotted for 2s at 4°C, plunged into liquid ethane by an FEI Vitrobot MKIII at 100% humidity and loaded onto a Titan Krios transmission electron microscope operated at 300kV. Zero-loss-energy images were collected manually on a Gatan K2-Summit detector in super-resolution counting mode at a calibrated magnification of 35,714x (pixel size of 1.43Å) and a dose rate of  $\sim$ 2.5 electron per Å<sup>2</sup> per second (Extended

Data Figure 1A). We used a slit width of 20 eV on a GIF Quantum energy filter. Each image was exposed for a total of 16 seconds and dose-fractionated into 20 movie frames. A defocus range of 0.5-3.5  $\mu\text{m}$  was used.

## Image processing

MOTIONCORR<sup>51</sup> was used for whole-image drift correction of the movie frames of each micrograph, and contrast transfer function (CTF) parameters of the corrected micrographs were estimated using CTFFIND4 (ref. 52). All subsequent processing steps were done using RELION<sup>53</sup> unless otherwise stated. A subset of ~5000 particles was picked manually, extracted using a  $380^2$  pixel box and subjected to reference-free 2D classification. Some of the resulting 2D class averages were low-pass filtered to 20 $\text{\AA}$  and used as references for automatic particle picking of the whole dataset of 2477 micrographs. The automatically picked particles were screened manually to remove false positives, aggregation and ice contamination, resulting in an initial set of 473,827 particles for reference-free 2D classification. We selected 438,602 particles from good 2D classes for the 3D classification (Extended Data Fig. 1b and 1c), which was run for 25 iterations, using an angular sampling of  $7.5^\circ$ , a regularisation parameter  $T$  of 4 and a 60 $\text{\AA}$  low-pass filtered initial model from our previous reconstruction<sup>16</sup>. A subset of 140,155 particles was selected for the first 3D auto-refinement. Particle-based beam-induced motion correction and radiation-damage weighing (particle polishing) were performed on these particles<sup>54</sup>. Auto-refinement of the polished particles resulted in a reconstruction at 3.7 $\text{\AA}$  overall resolution with an estimated angular accuracy of  $1.1^\circ$ .

Local resolution analysis by Resmap<sup>55</sup> showed a range of resolution from 3.0  $\text{\AA}$  in the core to 10  $\text{\AA}$  in the Arm domain and part of the Head domain, indicating conformational heterogeneity within the complex. As previously observed, the four domains of the structure, particularly the Head and Arm domains, are flexible in our structure. We employed two classification/refinement approaches: a local approach to improve the local resolution of the domains and a global approach to allow global conformations of the domains relative to one another to be observed (Extended Data Fig. 1c). For the local approach, we used a masked refinement procedure with signal subtraction for each of the Head, Body and Foot domains<sup>23</sup> and a masked classification with signal subtraction followed by a masked refinement for the most flexible Arm domain<sup>23</sup>. Each of the four domains only makes up a third or less of the total mass of the complex. For each domain, we subtracted projections from the remaining three domains of the reconstruction in the experimental particle images using the relative orientation of each experimental image from the last auto-refinement run of all the polished particles. This resulted in four sets of new experimental particle images that only have signal from the domain of interest. For the Body, Foot and Head domains the subtracted experimental images were used in 3D auto-refinement with a soft mask for that domain, yielding 3.6, 3.7 and 4.2  $\text{\AA}$  reconstructions for the Body, Foot and Head domains, respectively (Extended Data Fig. 1b, 2a-c). The Arm domain is too small for masked refinement. Thus we performed 3D classification on the subtracted experimental images with a mask around the Arm domain and no alignments. We selected three classes with 23,760, 26,367 and 24,627 particles each with three distinct conformations for the Arm domain. Since the Arm domain is too small for accurate alignments of the particles, we

refined each of these classes together with the Body domain using a new set of modified experimental particle images that included both the Arm and Body domains by the same subtraction method used previously (Extended Data Fig. 1c, 6a and 6b). Class 1 with 23,760 particles yielded a 4.6Å overall resolution for both Body and Arm domains and 6.2Å resolution for the Arm domain alone. Class 2 with 26,363 particles yielded a 4.5Å overall resolution for both Body and Arm domains and 7.5Å resolution for the Arm domain alone. Class 3 with 24,627 particles yielded a 4.4Å overall resolution for both Body and Arm domains and 6.2Å resolution for the Arm domain alone.

For the global classification/refinement approach, we performed 3D classification of the polished particles for the whole complex with a finer angular sampling of 1.8° and local angular search range of 10°. Two of the sub-classes of 48,945 and 36,824 particles had significantly better angular accuracies and gave 4.2 Å and 4.3 Å reconstructions, respectively, after auto-refinement with more homogeneous conformations of the Head and Foot domains. We observed distinct “Open” and “Closed” conformations for the Head and Foot domains (Extended Data Figure 6c-e).

All reported resolutions are based on the gold-standard Fourier Shell Correlation (FSC) = 0.143 criterion<sup>56</sup>. FSC curves were calculated using soft spherical masks and high-resolution noise substitution was used to correct for convolution effects of the masks on the FSC curves<sup>57</sup>. Prior to visualization, all maps were corrected for the modulation transfer function of the detector. Local resolution was estimated using Resmap<sup>55</sup>.

## Model building

The maps resulting from local masked refinements were first used for *de novo* model building using our previous protein placements<sup>15</sup> because they have the best local resolution for each of the domains separately. All model building was performed in COOT<sup>58</sup>. In our medium resolution structure, except for Brr2<sup>442-2163</sup>, Prp8<sup>885-2413</sup>, Snu13 and LSm proteins whose yeast structures are available, all other proteins are either human, homology models or idealized poly-Ala helices and only double-stranded RNA helices were modelled. Recently the structure of the ferredoxin domain of yeast Prp3 (residues 335-467) became available<sup>38</sup>, which replaced our homology model of this domain. We rebuilt and extended the yeast Brr2<sup>442-2163</sup>, Snu13, Prp8<sup>885-2413</sup> and Prp3<sup>335-467</sup> and all remaining components were built *de novo* first into the masked refinement maps and rigid-body fitted into the overall 3.7Å map. We identified a previously unassigned density as that of Snu66 based on previous yeast two-hybrid studies<sup>44</sup> and its interacting proteins in our structure. The LSm proteins were rigid-body fitted into the overall map and the improved maps of the 3 classes from masked classification and refinement. Extended Data Table 1 summarises all modeled components of the structure. The model was refined using REFMAC 5.8 (ref. 59) with secondary structure restraints provided by PROSMART<sup>60</sup> and RNA base-pair and stacking restraints provided by LIBG<sup>61</sup>. We first performed model refinement for the Body, Foot and Head domains separately against the corresponding masked refined maps (Extended Data Table 2a). The subunits of these three refined models were rigid-body fitted into the masked all map. To resolve the possible clashes in the domain interfaces, we refined this overall model against the overall map. Cross-validation of two half maps defined a REFMAC



refinement weights of 0.001. The Xmipp package<sup>62</sup> was used to calculate FSC model versus map. FSC curves of model versus map were calculated for the maps of the Body, Foot and Head domains, which were used for model building and refinement of the structure and also the overall map (Extended Data Fig. 2d-g). Extended Data Table 2 summarises refinement statistics for the overall structure and the domain structures and the deposited maps and their associated coordinates.

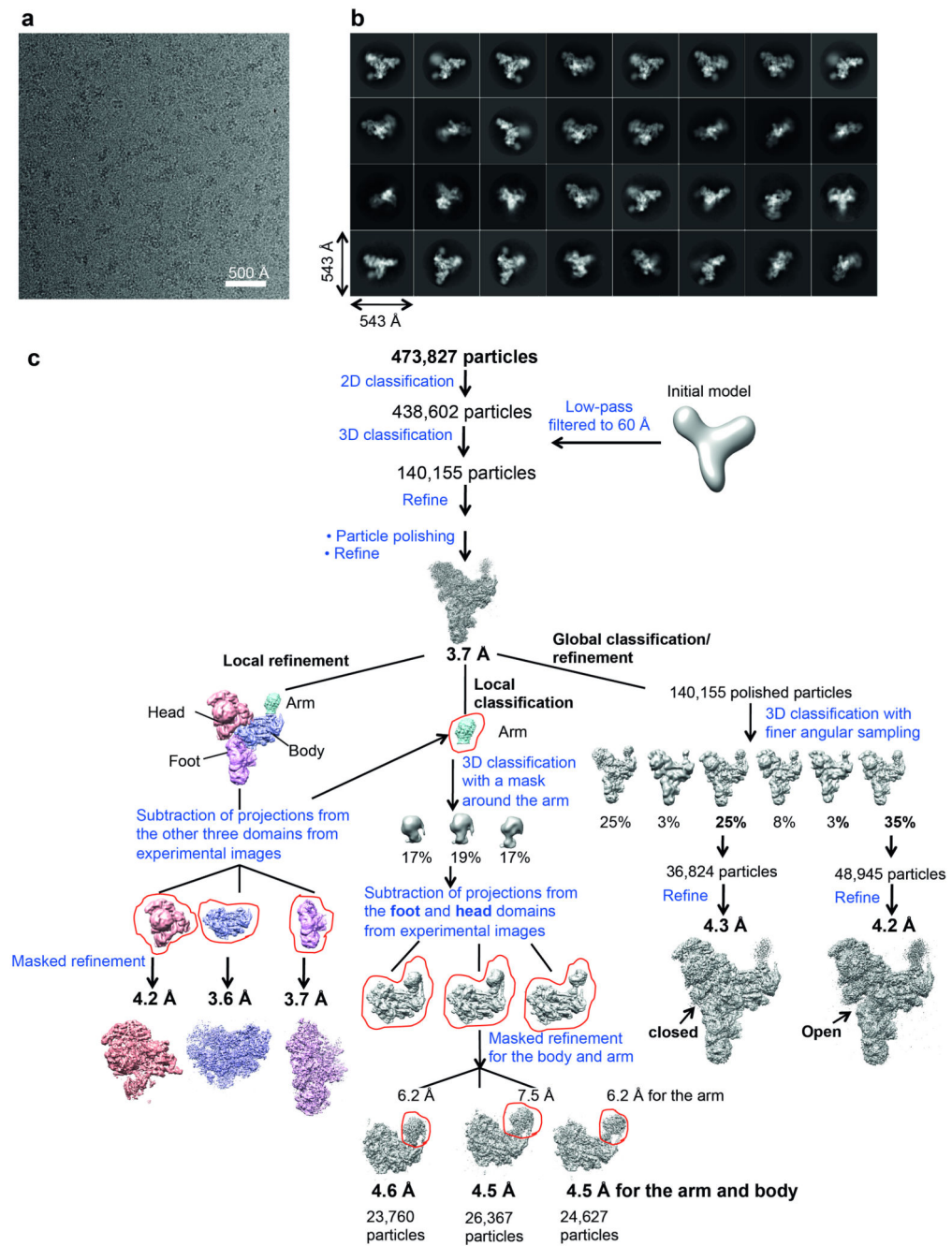
### Map visualisation

Maps were visualized in Chimera<sup>63</sup> and all figures were prepared using either Pymol ([www.pymol.org](http://www.pymol.org)) or Chimera.

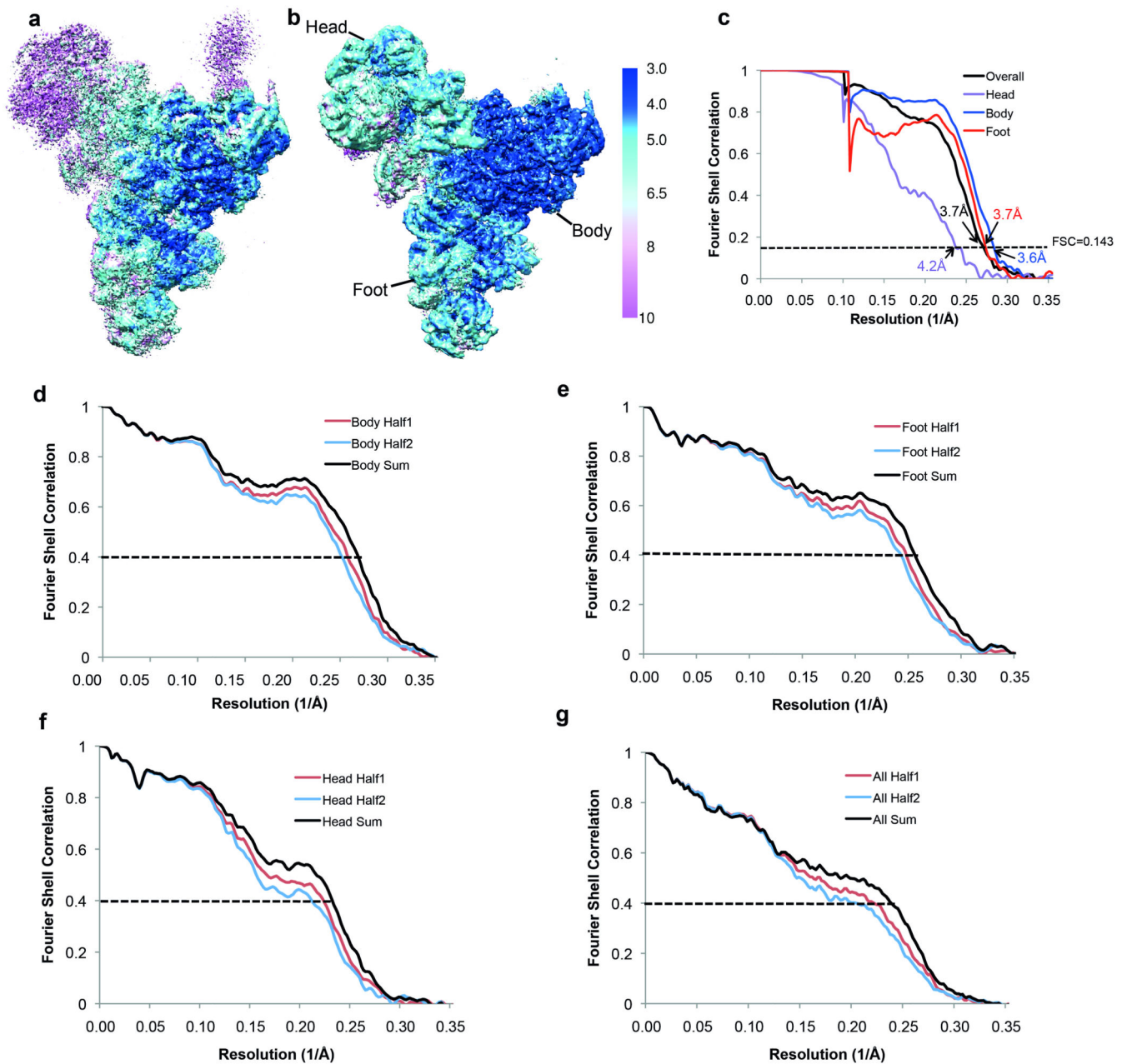
### Plasmid Shuffling

Mutations were introduced into *PRP8* and *SNU114* genes by the *dut<sup>-</sup> ung<sup>-</sup>* methods<sup>64</sup>. The viability of *Prp8* and *Snu114* mutants was assessed by plasmid shuffling analysis. The *Prp8* deletion strain SC261 8B1 (ref. 65) carrying wild type *PRP8* on pRS316 (*URA3*, centromeric replication origin) was transformed with mutant *Prp8* genes on pRS314 (*TRP1*, centromeric replication origin) and transformants were selected on plates lacking tryptophan. The *Snu114* deletion strain YSNU114KO1 (ref. 66) carrying wild type *SNU114* on pRS416 (*URA3*, centromeric replication origin) was transformed with mutant *Snu114* genes on pRS413 (*HIS3*, centromeric replication origin) and transformants were selected on plates lacking histidine. *Trp<sup>+</sup>* and *His<sup>+</sup>* cells were transferred onto plates containing 5-fluoro-orotic acid (5-FOA), to test cell growth at 30°C after loss of the *URA3*-marked plasmid. Plasmids were rescued from the 5-FOA-resistant strains and sequenced to confirm the presence of the appropriate mutation, and cell growth was assessed on YEPD plates incubated at various temperatures.

Extended Data



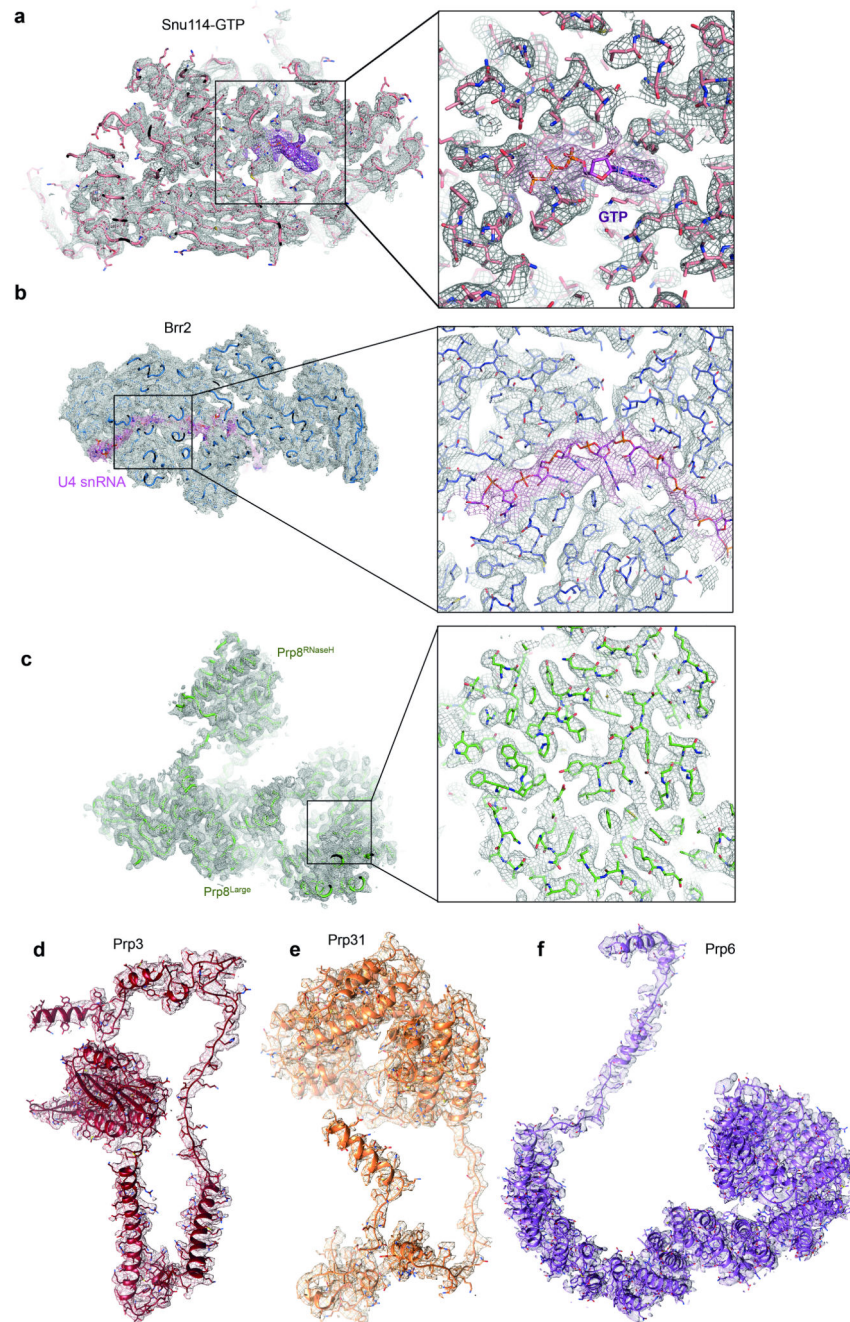
**Extended Data Figure 1. Image processing procedures**  
**a**, Representative micrograph. **b**, Representative 2D class averages obtained from reference-free 2D classification. **c**, Classification and refinement procedures used in this study.



**Extended Data Figure 2. Local and overall resolutions of tri-snRNP maps**

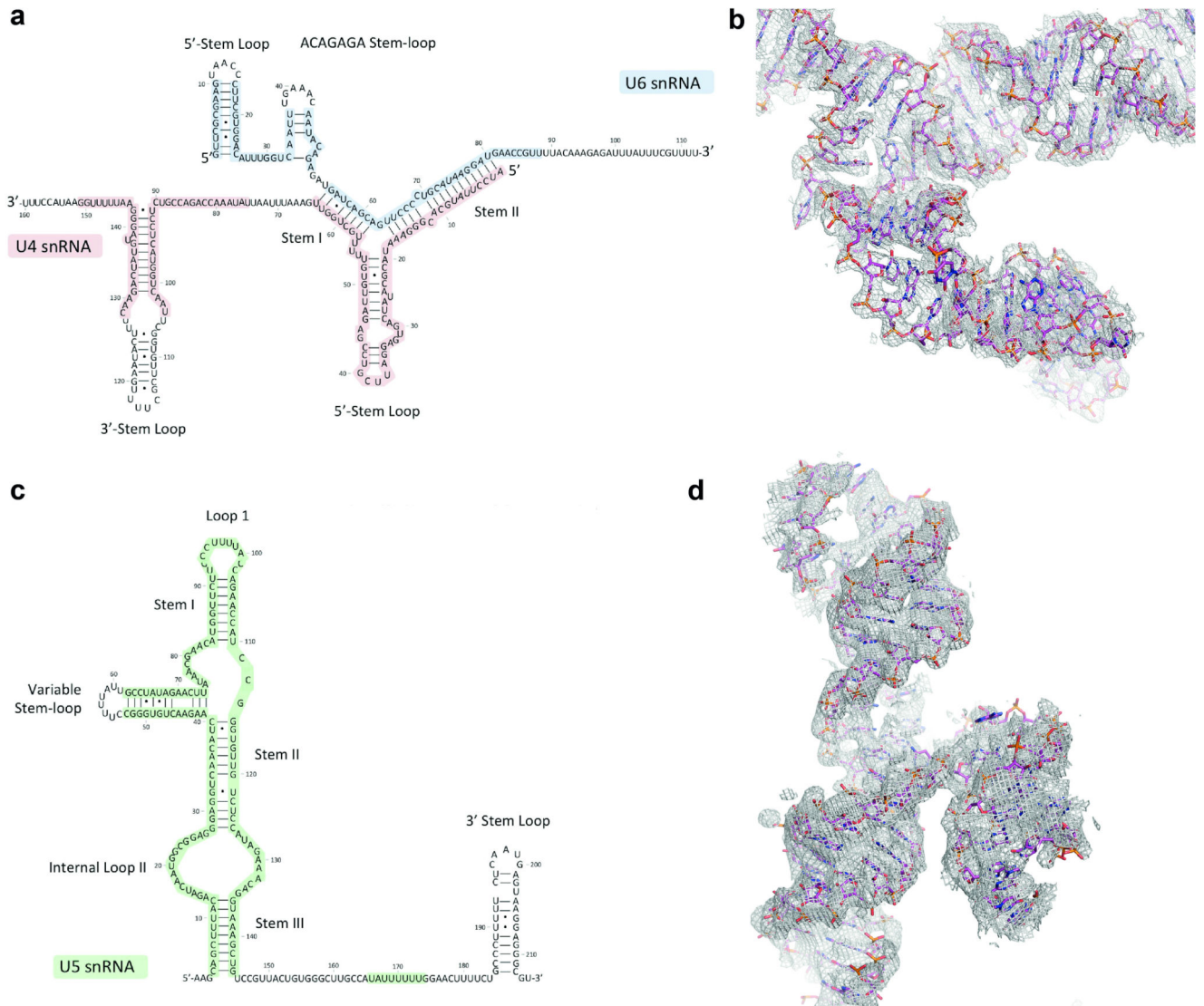
Local resolution estimation by Resmap<sup>55</sup> of **a**, the overall 3.7 Å map and **b**, maps of the head, body and foot domains obtained from masked refinements with signal subtraction<sup>23</sup>. **c**, Gold-standard FSC curves for the overall map and the maps of the head, body and foot domains obtained from masked refinements. Their resolutions are estimated at FSC=0.143. **d**, **e**, **f** and **g** FSC curves of model versus map and cross-validation of model refinement by half-maps for the Body, Foot, Head and Overall maps, respectively. The red curves show FSC between the atomic model and the half-map it was refined against (half1) and the blue curves show FSC between the atomic model and the other half-map (half2) it

was not refined against. The black curves show FSC between the atomic model and the sum map which the model was refined against.

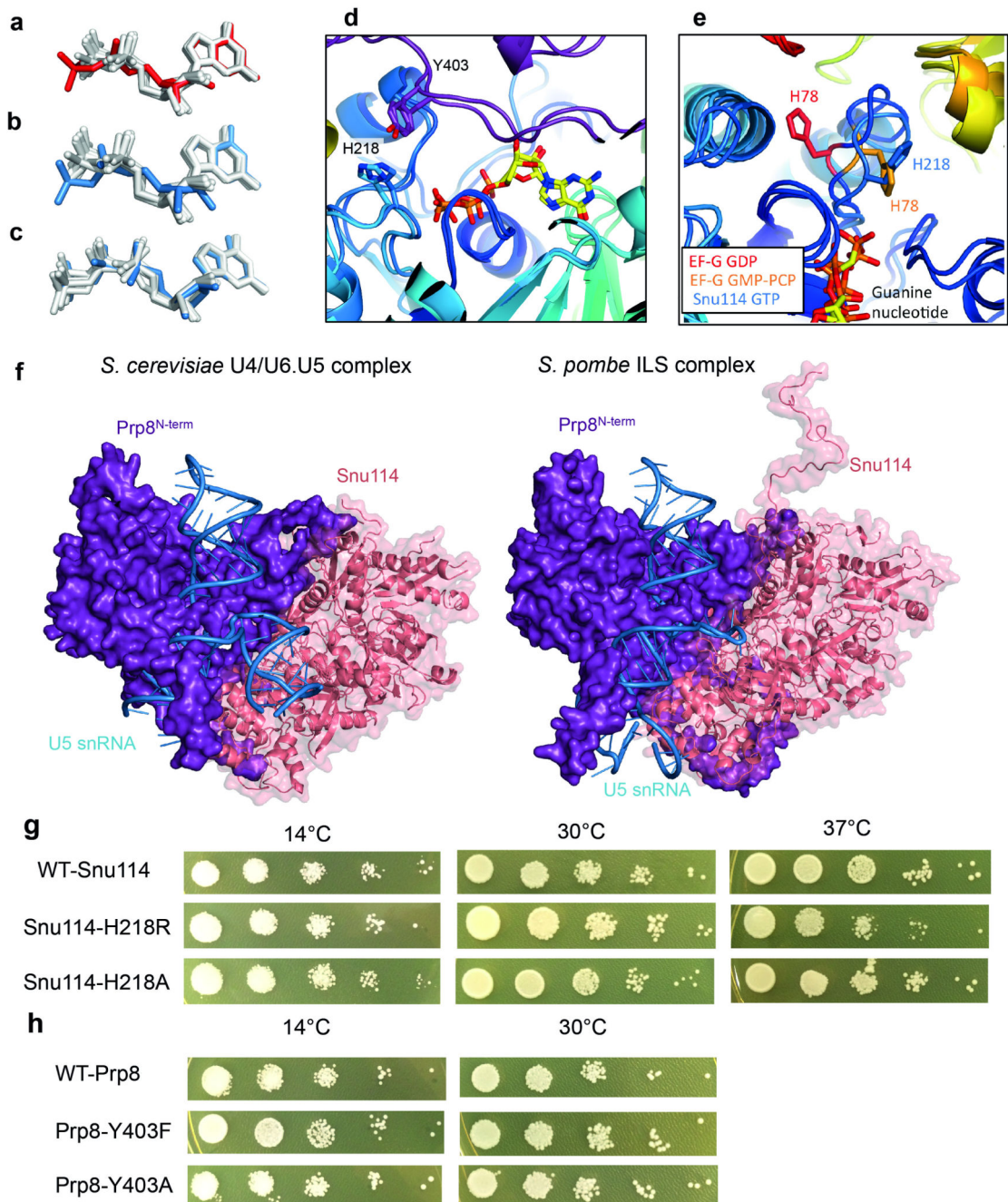


**Extended Data Figure 3. Representative EM density for different components of the map**  
**a**, Snu114 in the Foot domain with a bound GTP (magenta). The inset shows the GTP-binding pocket. **b**, Brr2 in the Head domain with a bound single-stranded region of U4 snRNA. The inset shows the density in the RNA binding tunnel. **c**, Density for Prp8 large

and RNase-like domains. The inset shows the density in the core of Prp8. **d, e** and **f**, Prp3, Prp31 and Prp6 densities, respectively, with extended polypeptides.

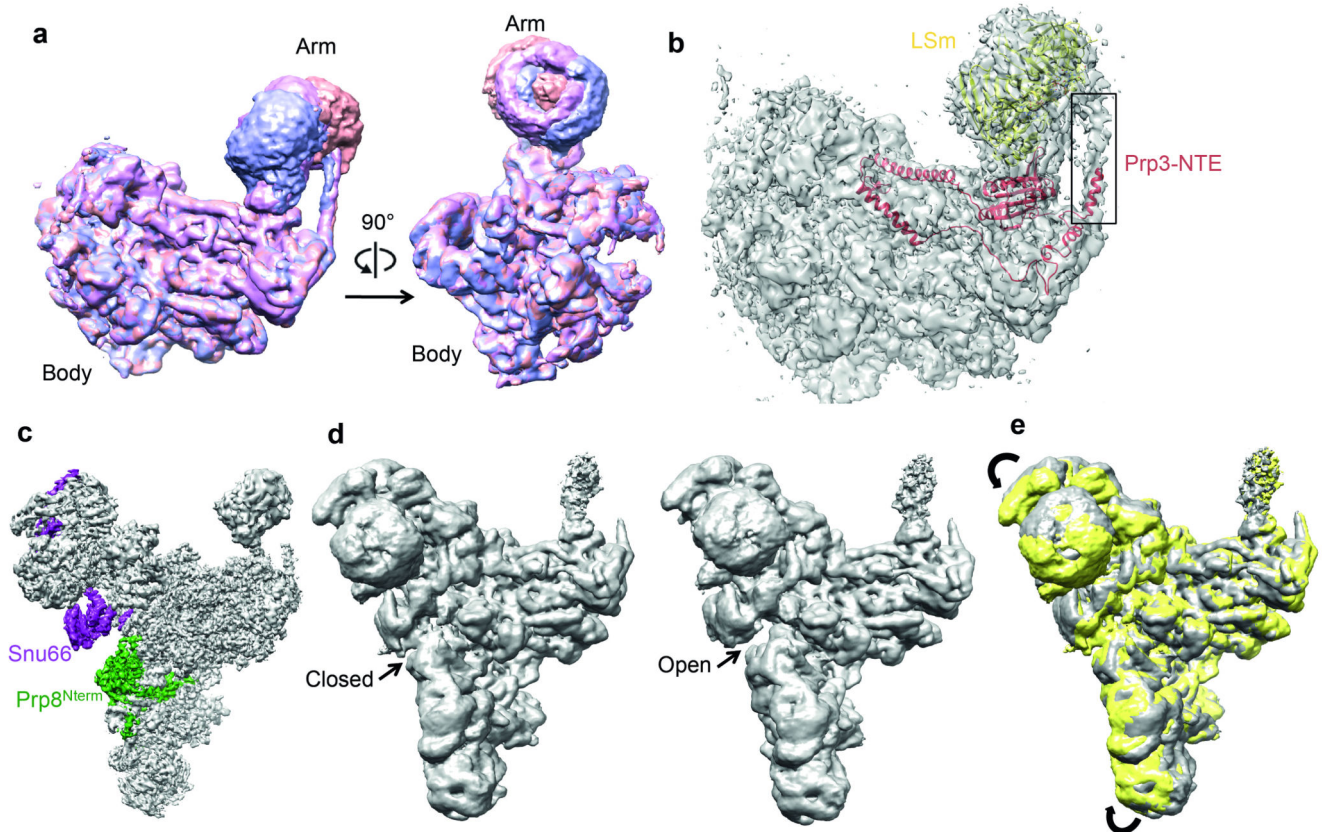


**Extended Data Figure 4. Secondary structure of the snRNAs in tri-snRNP**  
**a**, U4/U6 snRNA; **c**, U5 snRNA. The colored nucleotides with red, green and blue background were built *de novo* into our EM density. The region near the ACAGAGA sequence of U6 snRNA forms a stem-loop that was not predicted previously. **b, d**, Representative EM density for U4/U6 snRNA duplex and U5 snRNA, respectively.



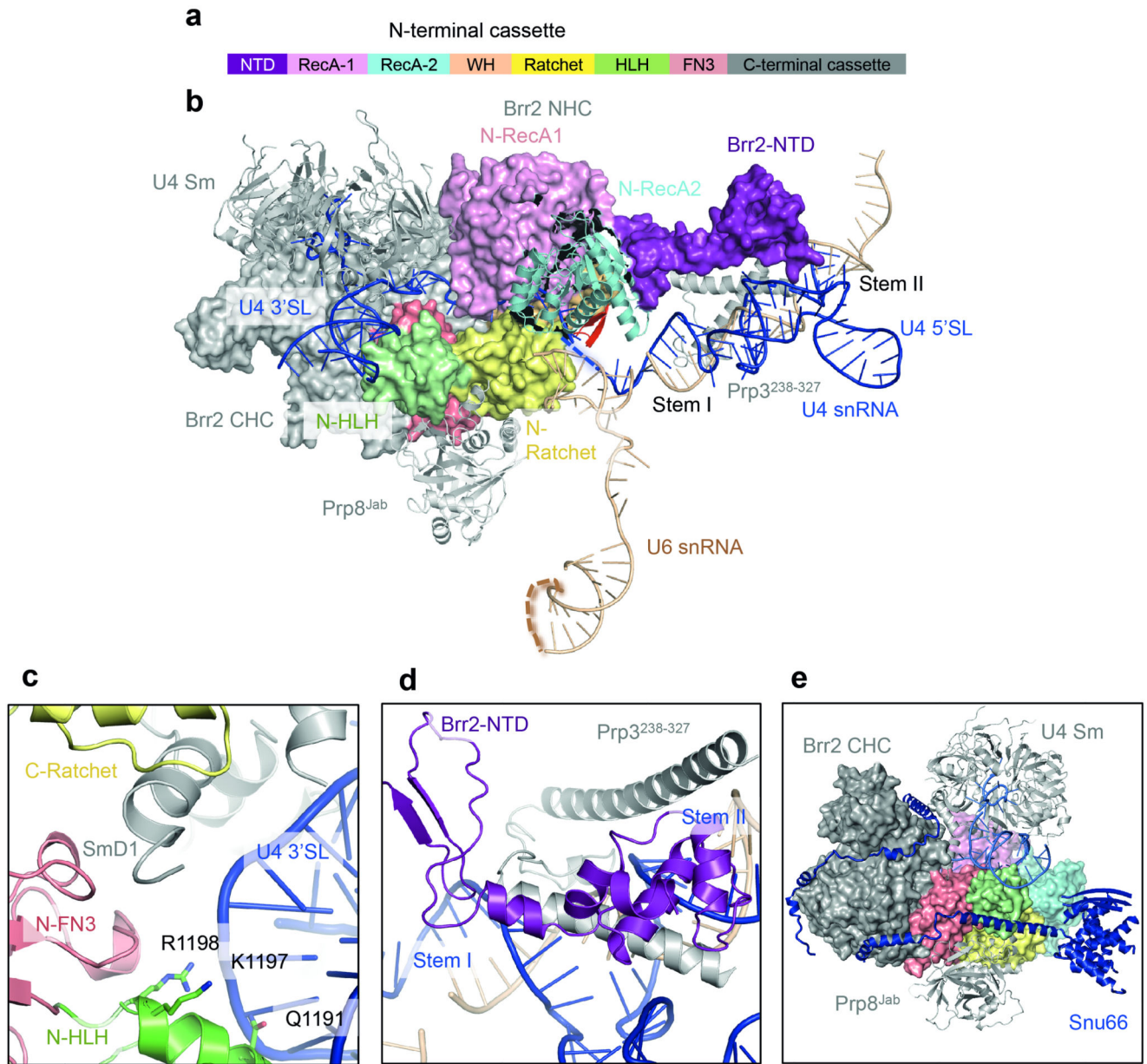
**Extended Data Figure 5. Interactions of Snu114 with guanine nucleotides and the N-terminal domain of Prp8 in the *S. cerevisiae* U4/U6.U5 tri-snRNP and *S. pombe* ILS complexes**  
**a**, Conformation of the Snu114(*Cwf10*)-bound GDP refined in the *S. pombe* ILS spliceosomal complex<sup>19,20</sup> (red, PDB 3JB9), was overlaid on GDPs found in other guanine-nucleotide binding proteins (grey, PDB coordinates: 1DAR, 2E1R, 2WRI, 1Z0I, 5CA8, 1XTQ, 4YLG, 1SF8, 5BXQ). **b**, Guanine nucleotide refined as GDP in Snu114 of the *S. cerevisiae* U4/U6.U5 tri-snRNP (blue) is overlaid on GDPs found in the PDB coordinates as in **a**. **c**, Conformation of guanine nucleotide refined as GTP in Snu114 of the *S. cerevisiae*

U4/U6.U5 tri-snRNP (blue) agrees well with GTP or GTP analogs in other guanine-nucleotide binding proteins (PDB code: 2BV3, 2DY1, 2J7K, 4YW9, 1ASO, 1LF0 (grey)). **d**, Superposition of the active site of Snu114-GTP and Cwf10-GDP. **e**, Superposition of the GDP-bound EF-G (2WRI), GMP-PCP bound EF-G (4JUW) and Snu114 (*S. cerevisiae* tri-snRNP) active sites. His218 (His78 in EF-G) positions water molecule crucial for GTP hydrolysis. **f**, Comparison of Prp8<sup>N-term</sup> domain, Snu114 and U5 snRNA in the *S. cerevisiae* U4/U6.U5 complex and *S. pombe* ILS complex. **g**, Growth of serial dilutions of yeast strains carrying wild-type Snu114, His218Arg or His218Ala Snu114 mutants at different temperatures. Cells were spotted on YPD plates and grown at 14°C for 10 days, 30°C and 37°C for 2 days. **h**, Growth of serial dilutions of yeast strains carrying wild-type Prp8, Tyr403Phe and Tyr403Ala mutants. Cells were spotted on YPD plates and grown at 14°C for 9 days, 30°C for 3 days. This yeast strain does not survive at 37°C and thus is not shown.



**Extended Data Figure 6. Conformational flexibility of tri-snRNP observed by classification**  
**a**, Different conformations of the Arm domain demonstrated by the unsharpened maps of the three major classes (purple, magenta and red) obtained from masked classification of the Arm domain alone followed by masked refinement with the Body and Arm domains. The Body domain was included in the refinement because the arm domain is too small for accurate alignments. **b**, The sharpened map of one of the three classes with Prp3 and LSM models shown. In the improved domain maps for the Arm domain, extra density for the N-terminal helix of Prp3 could be observed to extend to the LSM proteins. **c**, The sharpened

map of the tri-snRNP and the locations of Snu66 and Prp8. **d**, The open and closed conformations of the Head and Foot domains of the tri-snRNP observed by global classification. The unsharpened maps for the two major classes obtained from global classification with finer angular sampling (1.8°) followed by 3D auto-refinement are shown. The open and closed states are indicated. **e**, Superposition of the unsharpened maps of the open (grey) and closed (yellow) states shown in **d**. The arrows indicate the rotations of the head and foot domains.



**Extended Data Figure 7. Brr2 helicase and its U4/U6 snRNA substrate**  
**a**, domain structure of Brr2 helicase comprising the N-terminal domain and two helicase cassettes. Individual domains of N-terminal helicase cassette (NHC) are colour-coded. **b**,



Extensive interactions of Brr2 with U4/U6 snRNA and Prp3. The single-stranded region of U4 snRNA extending from stem I enters the active site near the  $\beta$ -finger (red). **c**, 3' stem of U4 snRNA interacts with the HLH domain of NHC. **d**, The N-terminal domain (NTD) of Brr2 interacts with a long helix of Prp3 and inserts a loop into U4/U6 Stem II. **e**, Snu66 has a long extended region that wraps around both helicase cassettes of Brr2.

**Extended Data Table 1**

**Summary of model building of tri-snRNP components**

	protein	total residues	M.W.	Modeled	Chain name	Local map	Human/S. pombe names	
U5 snRNP	Prp8	2413	279,299	110-2401	A	108-735: Foot 751-2104: Body 2147-2401: Head	220K/Spp42	
	Brr2	2163	246,125	364-2163	B	363-433: Body 439-2163: Head	200K/Brr2	
	Snu114	1008	114,025	102-989	C	Foot	116K/Cwf10	
	Dib1	143	16,774	2-137	D	Body	15K/Dim1	
	SmB	196	22,403	4-102	b		SmB/SmB	
	SmD3	110	11,229	1-109	d		SmD3/SmD3	
	SmD1	146	16,288	15-108	h		SmD1/SmD1	
	SmD2	110	12,856	4-85	i	Foot	SmD2/SmD2	
	SmE	94	10,373	4-92	e		SmE/SmE	
	SmF	96	9,659	12-83	f		SmF/SmF	
	SmG	77	8,479	2-76	g		SmG/SmG	
	U5 snRNA-L	214	68,847	4-173	U	4-173: Foot 88-107: Body		
U4/U6 snRNP	Snu13	126	13,570	3-126	K	Body	15.5K/Snu13	
	Prp31	494	56,305	43-457	F	Body	61K/Prp31	
	Prp3	469	55,877	150-467	G	Body	90K/Prp3	
	Prp4	465	52,425	109-465	H	Body	60K/Rna4	
	SmB	196	22,403	4-102	k		SmB/SmB	
	SmD1	146	16,288	1-118	l		SmD1/SmD1	
	SmD2	110	12,856	15-108	m		SmD2/SmD2	
	SmD3	110	11,229	4-85	n	Head	SmD3/SmD3	
	SmE	94	10,373	10-92	p		SmE/SmE	
	SmF	96	9,659	12-83	q		SmF/SmF	
	SmG	77	8,479	2-76	r		SmG/SmG	
		Lsm2	95	11,164	1-90	2		Lsm2/Lsm2
		Lsm3	89	10,020	3-79	3		Lsm3/Lsm3
	Lsm4	172	20,304	1-90	4		Lsm4/Lsm4	
	Lsm5	93	10,415	4-84	5	Arm	Lsm5/Lsm5	
	Lsm6	86	9,396	11-84	6		Lsm6/Lsm6	

	protein	total residues	M.W.	Modeled	Chain name	Local map	Human/S. pombe names
	Lsm7	115	13,010	26-105	7		Lsm7/Lsm7
	Lsm8	109	12,385	1-67	8		Lsm8/Lsm8
	U4 snRNA	160	51,390	1-152	V	1-67: Body 73-152: Head	
	U6 snRNA	112	36,088	1-112	!	1-26: Foot 26-88: Body 108-112: Arm	
	Prp6	899	104,234	155-898	J	Body	102K/Prp1
	Snu66	587	66,426	5-560 (poly-Ala)	E	Head	110K/Snu66
tri-snRNP specific	Prp38	242	27,957	Not modeled			hPrp38/Prp38
	Snu23	194	22,682	Not modeled			hSnu23/Snu23
	Spp381	291	33,764	Not modeled			

**Extended Data Table 2**

**Refinement, model statistics and structure/map depositions**

**a. Statistic of tri-snRNP structure determination**

**Data collection**

EM	Titan Krios 300kV, K2 Gatan Summit			
Pixel size (Å)	1.43			
Defocus range (µm)	-0.5 to -3.5			
Reconstruction (RELION)	Overall	Body	Foot	Head
Accuracy of rotations (°)	1.13	1.15	1.73	2.42
Accuracy of translations (pixel)	0.65	0.67	0.89	1.28
Final resolution	3.7	3.6	3.7	4.2
Refinement (REFMAC)				
Refinement weight	0.001	0.001	0.001	0.001
Resolution limits	3.6	3.6	3.6	3.6
Residue numbers	9325	3728	2186	2922
Fourier Shell Correlation	0.75	0.85	0.82	0.6
R-factor (%)	29.7	27.8	28.7	31.5
Rms bond length (Å)	0.0078	0.0073	0.0073	0.011
Rms bond angle (°)	1.27	1.33	1.38	1.4
Ramachandran plot				
Favoured	8066 (91.4%)	3266 (91.9%)	1810 (90.9%)	2531 (89.3%)
Allowed	615 (6.9%)	237 (6.7%)	135 (7.2%)	238 (9.3%)
Outliers	152 (1.7%)	50 (1.4%)	37 (1.9%)	39 (1.4%)
Validation by Molprobit				
Geometry score (percentile)	2.52 (98 <sup>th</sup> )	2.41 (99 <sup>th</sup> )	2.79 (95 <sup>th</sup> )	2.62 (97 <sup>th</sup> )
Clashscore (percentile)	7.48 (97 <sup>th</sup> )	6.78 (100 <sup>th</sup> )	11.4 (97 <sup>th</sup> )	6.82 (100 <sup>th</sup> )

**a. Statistic of tri-snRNP structure determination****Data collection**

Good rotamer (%)	94.8	95.7	93.5	93.2
------------------	------	------	------	------

**b. Deposited maps and associated structures**

Maps	EMDB code	Associated ID	PDB
Overall map	EMD-8012	5GAN	
Body map	EMD-8014	5GAP	
Head map	EMD-8013	5GAO	
Foot map	EMD-8011	5GAM	
Global class 1 (closed state)	EMD-8007		
Global class 2 (open state)	EMD-8006		
Masked body/arm class 1	EMD-8008		
Masked body/arm class 2	EMD-8009		
Masked body/arm class 3	EMD-8010		

## Acknowledgements

We thank Christos Savva, Shaoxia Chen, Greg McMullan, Jake Grimmett and Toby Darling for smooth running of the EM and computing facilities, Alan Brown, Paul Emsley, Garib Murshudov for advice and help with model building and refinement, Ray O'Keefe for the Snu114 yeast strain, the members of the spliceosome group for help and advice throughout the project and Rafael Leiro for help with data processing. We thank Sebastian Fica for critical reading of the manuscript and Jan Löwe, Venki Ramakrishnan, and Richard Henderson for their continuing support and encouragements. The project was supported by the Medical Research Council (MC\_U105184330 to K.N. and MC\_UP\_A025\_1013 to S.H.W.S.).

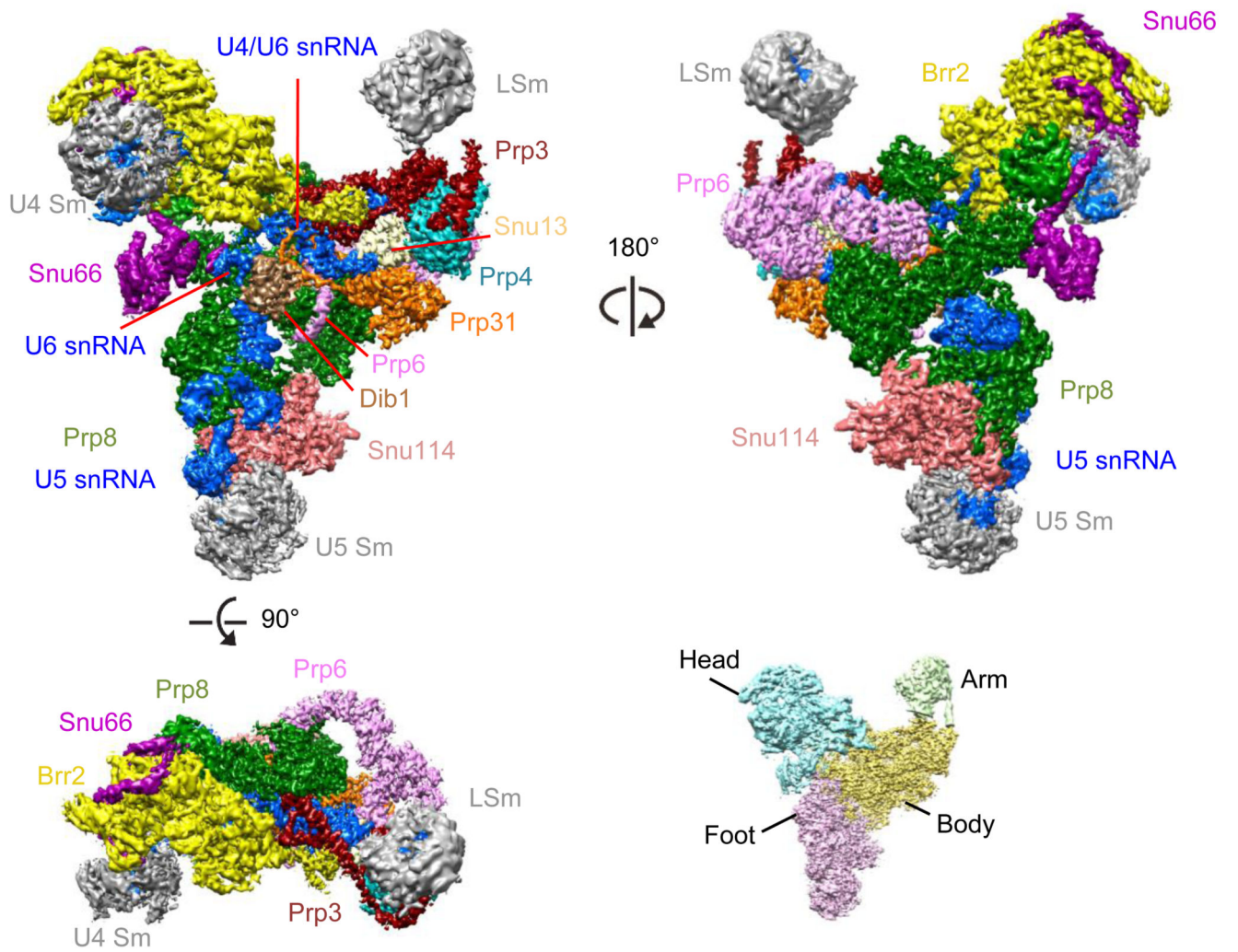
## References

- Will CL, Lührmann R. Spliceosome structure and function. *Cold Spring Harb. Perspect. Biol.* 2011; 3:pri: a003707.
- Chen CH, et al. Functional and physical interaction between components of the Prp19p-associated complex. *Nucleic Acids Res.* 2002; 30:1029–1037. [PubMed: 11842115]
- Fabrizio P, Dannenberg J, Dube P, Kastner B, Stark H, Urlaub H, Lührmann R. The evolutionarily conserved core design of the catalytic activation step of the yeast spliceosome. *Mol. Cell.* 2009; 36:593–608. [PubMed: 19941820]
- Ohr T, Odenwälder P, Dannenberg J, Prior M, Warkocki Z, Schmitzová J, Karaduman R, Gregor I, Enderlein J, Fabrizio P, Lührmann R. Molecular dissection of step 2 catalysis of yeast pre-mRNA splicing investigated in a purified system. *RNA.* 2013; 19:902–915. [PubMed: 23685439]
- Staley JP, Guthrie C. An RNA switch at the 5' splice requires ATP and the DEAD box protein Prp28p. *Mol. Cell.* 1999; 3:55–64. [PubMed: 10024879]
- Raghunathan PL, Guthrie C. RNA unwinding in U4/U6 snRNPs requires ATP hydrolysis and the DEIH-box splicing factor Brr2. *Curr. Biol.* 1998; 8:847–855. [PubMed: 9705931]
- Laggerbauer B, Achsel T, Lührmann R. The human U5-200kD DEXH-box protein unwinds U4/U6 RNA duplexes in vitro. *Proc. Natl. Acad. Sci. USA.* 1998; 95:4188–4192. [PubMed: 9539711]
- Fica SM, Tuttle N, Novak T, Li NS, Lu J, Koodathingal P, Dai Q, Staley JP, Piccirilli JA. RNA catalyses nuclear pre-mRNA splicing. *Nature.* 2013; 503:229–234. [PubMed: 24196718]
- Madhani HD, Guthrie C. Dynamic RNA-RNA interactions in the spliceosome. *Annu. Rev. Genet.* 1994; 28:1–26. [PubMed: 7534458]
- Toor N, Keating KS, Taylor SD, Pyle AM. Crystal structure of a self-spliced group II intron. *Science.* 2008; 320:77–82. [PubMed: 18388288]

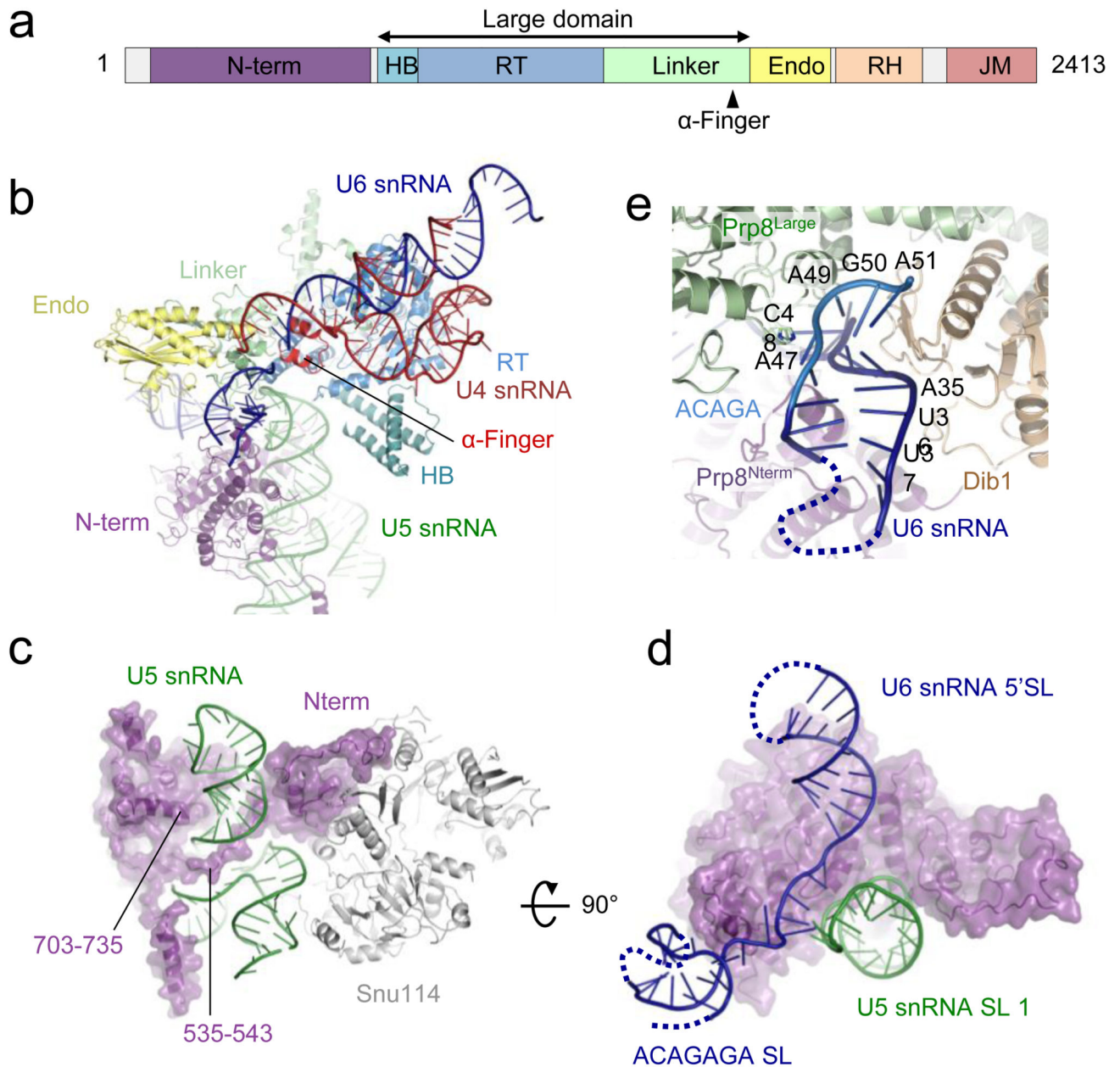
11. Fica SM, Mefford MA, Piccirilli JA, Staley JP. Evidence for a group II intron-like catalytic triplex in the spliceosome. *Nature Struct. Mol. Biol.* 2014; 21:464–471. [PubMed: 24747940]
12. Newman AJ, Norman C. U5 snRNA interacts with exon sequences at 5' and 3' splice sites. *Cell.* 1992; 68:743–754. [PubMed: 1739979]
13. Sontheimer EJ, Steitz JA. The U5 and U6 small nuclear RNAs as active site components of the spliceosome. *Science.* 1993; 262:1989–1996. [PubMed: 8266094]
14. Stark H, Lührmann R. Cryo-electron microscopy of spliceosomal components. *Annu. Rev. Biophys. Biomol. Struct.* 2006; 35:435–457. [PubMed: 16689644]
15. Bai XC, McMullan G, Scheres SH. How cryo-EM is revolutionizing structural biology. *Trends Biochem. Sci.* 2015; 40:49–57. [PubMed: 25544475]
16. Nguyen TH, Galej WP, Bai XC, Savva CG, Newman AJ, Scheres SHW, Nagai K. The architecture of the spliceosomal U4/U6.U5 tri-snRNP. *Nature.* 2015; 523:47–52. [PubMed: 26106855]
17. Ohi MD, Ren L, Wall JS, Gould KL, Walz T. Structural characterization of the fission yeast U5.U2/U6 spliceosome complex. *Proc. Natl. Acad. Sci. USA.* 2007; 104:3195–3200. [PubMed: 17360628]
18. Chen W, Shulha HP, Ashar-Patel A, Yan J, Green KM, Query CC, Rhind N, Weng Z, Moore MJ. Endogenous U2-U5-U6 snRNA complexes in *S. pombe* are intron lariat spliceosomes. *RNA.* 2014; 20:308–320. [PubMed: 24442611]
19. Yan C, Hang J, Wan R, Huang M, Wong CC, Shi Y. Structure of a yeast spliceosome at 3.6-angstrom resolution. *Science.* 2015; 349:1182–1191. [PubMed: 26292707]
20. Hang J, Wan R, Yan C, Shi Y. Structural basis of pre-mRNA splicing. *Science.* 2015; 349:1191–1198. [PubMed: 26292705]
21. Stevens SW, Barta I, Ge HY, Moore RE, Young MK, Lee TD, Abelson J. Biochemical and genetic analyses of the U5, U6, and U4/U6.U5 small nuclear ribonucleoproteins from *Saccharomyces cerevisiae*. *RNA.* 2001; 7:1543–1553. [PubMed: 11720284]
22. Gottschalk A, Neubauer G, Banroques J, Mann M, Lührmann R, Fabrizio P. Identification by mass spectrometry and functional analysis of novel proteins of the yeast [U4/U6.U5] tri-snRNP. *EMBO J.* 1999; 18:4535–4548. [PubMed: 10449419]
23. Bai X-C, Rajendra E, Yang G, Shi Y, Scheres SHW. Sampling the conformational space of the catalytic subunit of human  $\gamma$ -secretase. *Elife.* 2015; 4:pii: e11182.
24. Galej WP, Oubridge C, Newman AJ, Nagai K. Crystal structure of Prp8 reveals active site cavity of the spliceosome. *Nature.* 2013; 493:638–643. [PubMed: 23354046]
25. Reuter K, Nottrott S, Fabrizio P, Lührmann R, Ficner R. Identification, characterization and crystal structure analysis of the human spliceosomal U5 snRNP-specific 15 kD protein. *J. Mol. Biol.* 1999; 294:515–525. [PubMed: 10610776]
26. Fabrizio P, Laggerbauer B, Lauber J, Lane WS, Lührmann R. An evolutionarily conserved U5 snRNP-specific protein is a GTP-binding factor closely related to the ribosomal translocase EF-2. *EMBO J.* 1997; 16:4092–4106. [PubMed: 9233818]
27. Jorgensen R, Ortiz PA, Carr-Schmid A, Nissen P, Kinzy TG, Andersen GR. Two crystal structures demonstrate large conformational changes in the eukaryotic ribosomal translocase. *Nature Struct. Biol.* 2003; 10:379–385. [PubMed: 12692531]
28. Wittinghofer A, Vetter IR. Structure-function relationships of the G domain, a canonical switch motif. *Annu. Rev. Biochem.* 2011; 80:943–971. [PubMed: 21675921]
29. Tourigny DS, Fernández IS, Kelley AC, Ramakrishnan V. Elongation factor G bound to the ribosome in an intermediate state of translocation. *Science.* 2013; 340:1235490. [PubMed: 23812720]
30. Lin J, Gagnon MG, Bulkley D, Steitz TA. Conformational changes of elongation factor G on the ribosome during tRNA translocation. *Cell.* 2015; 160:219–227. [PubMed: 25594181]
31. Brenner TJ, Guthrie C. Genetic analysis reveals a role for the C terminus of the *Saccharomyces cerevisiae* GTPase Snu114 during spliceosome activation. *Genetics.* 2005; 170:1063–1080. [PubMed: 15911574]
32. Maracci C, Peske F, Dannies E, Pohl C, Rodnina MV. Ribosome-induced tuning of GTP hydrolysis by a translational GTPase. *Proc. Natl. Acad. Sci. USA.* 2014; 111:14418–1423. [PubMed: 25246550]

33. Small EC, Leggett SR, Winans AA, Staley JP. The EF-G-like GTPase Snu114p regulates spliceosome dynamics mediated by Brr2p, a DExD/H box ATPase. *Mol. Cell.* 2006; 23:389–399. [PubMed: 16885028]
34. Bartels C, Urlaub H, Lührmann R, Fabrizio P. Mutagenesis suggests several roles of Snu114p in pre-mRNA splicing. *J. Biol. Chem.* 2003; 278:28324–2834. [PubMed: 12736260]
35. Brenner TJ, Guthrie C. Assembly of Snu114 into U5 snRNP requires Prp8 and a functional GTPase domain. *RNA.* 2006; 12:862–871. [PubMed: 16540695]
36. Liu S, Li P, Dybkov O, Nottrott S, Hartmuth K, Lührmann R, Carlomagno T, Wahl MC. Binding of the human Prp31 Nop domain to a composite RNA-protein platform in U4 snRNP. *Science.* 2007; 316:115–120. [PubMed: 17412961]
37. Ayadi L, Callebaut I, Saguez C, Villa T, Mornon JP, Banroques J. Functional and structural characterization of the prp3 binding domain of the yeast prp4 splicing factor. *J. Mol. Biol.* 1998; 284:673–687. [PubMed: 9826507]
38. Liu S, Mozaffari-Jovin S, Wollenhaupt J, Santos KF, Theuser M, Dunin-Horkawicz S, Fabrizio P, Bujnicki JM, Lührmann R, Wahl MC. A composite double-/single-stranded RNA-binding region in protein Prp3 supports tri-snRNP stability and splicing. *Elife.* 2015; 4:e07320. [PubMed: 26161500]
39. Nottrott S, Urlaub H, Lührmann R. Hierarchical, clustered protein interactions with U4/U6 snRNA: a biochemical role for U4/U6 proteins. *EMBO J.* 2002; 21:5527–5538. [PubMed: 12374753]
40. Makarov EM, Makarova OV, Achsel T, Lührmann R. The human homologue of the yeast splicing factor prp6p contains multiple TPR elements and is stably associated with the U5 snRNP via protein-protein interactions. *J. Mol. Biol.* 2000; 298:567–575. [PubMed: 10788320]
41. Galisson F, Legrain P. The biochemical defects of prp4-1 and prp6-1 yeast splicing mutants reveal that the PRP6 protein is required for the accumulation of the [U4/U6.U5] tri-snRNP. *Nucleic Acids Res.* 1993; 21:1555–1562. [PubMed: 8479905]
42. Nguyen THD, Li J, Galej WP, Oshikane H, Newman AJ, Nagai K. Structural basis of Brr2-Prp8 interactions and implications for U5 snRNP biogenesis and the spliceosome active site. *Structure.* 2013; 21:910–919. [PubMed: 23727230]
43. Hahn D, Kudla G, Tollervey D, Beggs JD. Brr2p-mediated conformational rearrangements in the spliceosome during activation and substrate repositioning. *Genes Dev.* 2012; 26:2408–2421. [PubMed: 23124065]
44. van Nues RW, Beggs JD. Functional contacts with a range of splicing proteins suggest a central role for Brr2p in the dynamic control of the order of events in spliceosomes of *Saccharomyces cerevisiae*. *Genetics.* 2001; 157:1451–1467. [PubMed: 11290703]
45. Wolf E, Kastner B, Deckert J, Merz C, Stark H, Lührmann R. Exon, intron and splice site locations in the spliceosomal B complex. *EMBO J.* 2009; 28:2283–2292. [PubMed: 19536130]
46. Boehringer D, Makarov EM, Sander B, Makarova OV, Kastner B, Lührmann R, Stark H. Three-dimensional structure of a pre-catalytic human spliceosomal complex B. *Nature Struct. Mol. Biol.* 2004; 11:463–468. [PubMed: 15098019]
47. De I, Bessonov S, Hofele R, dos Santos K, Will CL, Urlaub H, Lührmann R, Pena V. The RNA helicase Aquarius exhibits structural adaptations mediating its recruitment to spliceosomes. *Nature Struct. Mol. Biol.* 2015; 22:138–144. [PubMed: 25599396]
48. Li Z, Brow DA. A spontaneous duplication in U6 spliceosomal RNA uncouples the early and late functions of the ACAGA element in vivo. *RNA.* 1996; 2:879–894. [PubMed: 8809015]
49. Kuhn AN, Brow DA. Suppressors of a cold-sensitive mutation in yeast U4 RNA define five domains in the splicing factor Prp8 that influence spliceosome activation. *Genetics.* 2000; 155:1667–1682. [PubMed: 10924465]
50. Chan SP, Cheng SC. The Prp19-associated complex is required for specifying interactions of U5 and U6 with pre-mRNA during spliceosome activation. *J. Biol. Chem.* 2005; 280:31190–31199. [PubMed: 15994330]
51. Li X, Mooney P, Zheng S, Booth CR, Braunfeld MB, Gubbens S, Agard DA, Cheng Y. Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nature Methods.* 2013; 10:584–590. [PubMed: 23644547]

52. Mindell JA, Grigorieff N. Accurate determination of local defocus and specimen tilt in electron microscopy. *J. Struct. Biol.* 2003; 142:334–347. [PubMed: 12781660]
53. Scheres SH. A Bayesian view on cryo-EM structure determination. *J. Mol. Biol.* 2012; 415:406–418. [PubMed: 22100448]
54. Scheres SH. Beam-induced motion correction for sub-megadalton cryo-EM particles. *Elife.* 2014; 3:e03665. [PubMed: 25122622]
55. Kucukelbir A, Sigworth FJ, Tagare HD. Quantifying the local resolution of cryo-EM density maps. *Nature Methods.* 2014; 11:63–65. [PubMed: 24213166]
56. Scheres SH, Chen S. Prevention of overfitting in cryo-EM structure determination. *Nature Methods.* 2012; 9:853–854. [PubMed: 22842542]
57. Chen S, et al. High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy. *Ultramicroscopy.* 2013; 135:24–35. [PubMed: 23872039]
58. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Cryst.* 2010; D66:486–501.
59. Murshudov, et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D.* 2011; 67:355–367. [PubMed: 21460454]
60. Nicholls RA, Fischer M, McNicholas S, Murshudov GN. Conformation-independent structural comparison of macromolecules with PROSMART. *Acta Crystallogr. D.* 2014; 70:2487–2499. [PubMed: 25195761]
61. Brown A, Long F, Nicholls RA, Toots J, Emsley P, Murshudov GN. Tools for macromolecular model building and refinement into electron cryomicroscopy reconstructions. *Acta Crystallogr. D.* 2015; 71:136–153. [PubMed: 25615868]
62. Scheres SHW, Nuñez-Ramirez R, Sorzano COS, Carazo JM, Marabini R. Image processing for electron microscopy single-particle analysis using Xmipp. *Nature Protoc.* 2008; 3:977–990. [PubMed: 18536645]
63. Goddard TD, Huang CC, Ferrin TE. Visualizing density maps with UCSF Chimera. *J. Struct. Biol.* 2007; 157:281–287. [PubMed: 16963278]
64. Kunkel TA. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl. Acad. Sci. USA.* 1985; 82:488–492. [PubMed: 3881765]
65. Turner IA, Norman CM, Churcher MJ, Newman AJ. Dissection of Prp8 protein defines multiple interactions with crucial RNA sequences in the catalytic core of the spliceosome. *RNA.* 2006; 12:375–386. [PubMed: 16431982]
66. Frazer LN, Lovell SC, O’Keefe RT. Analysis of synthetic lethality reveals genetic interactions between the GTPase Snu114p and snRNAs in the catalytic core of the *Saccharomyces cerevisiae* spliceosome. *Genetics.* 2009; 183:497–515. [PubMed: 19620389]



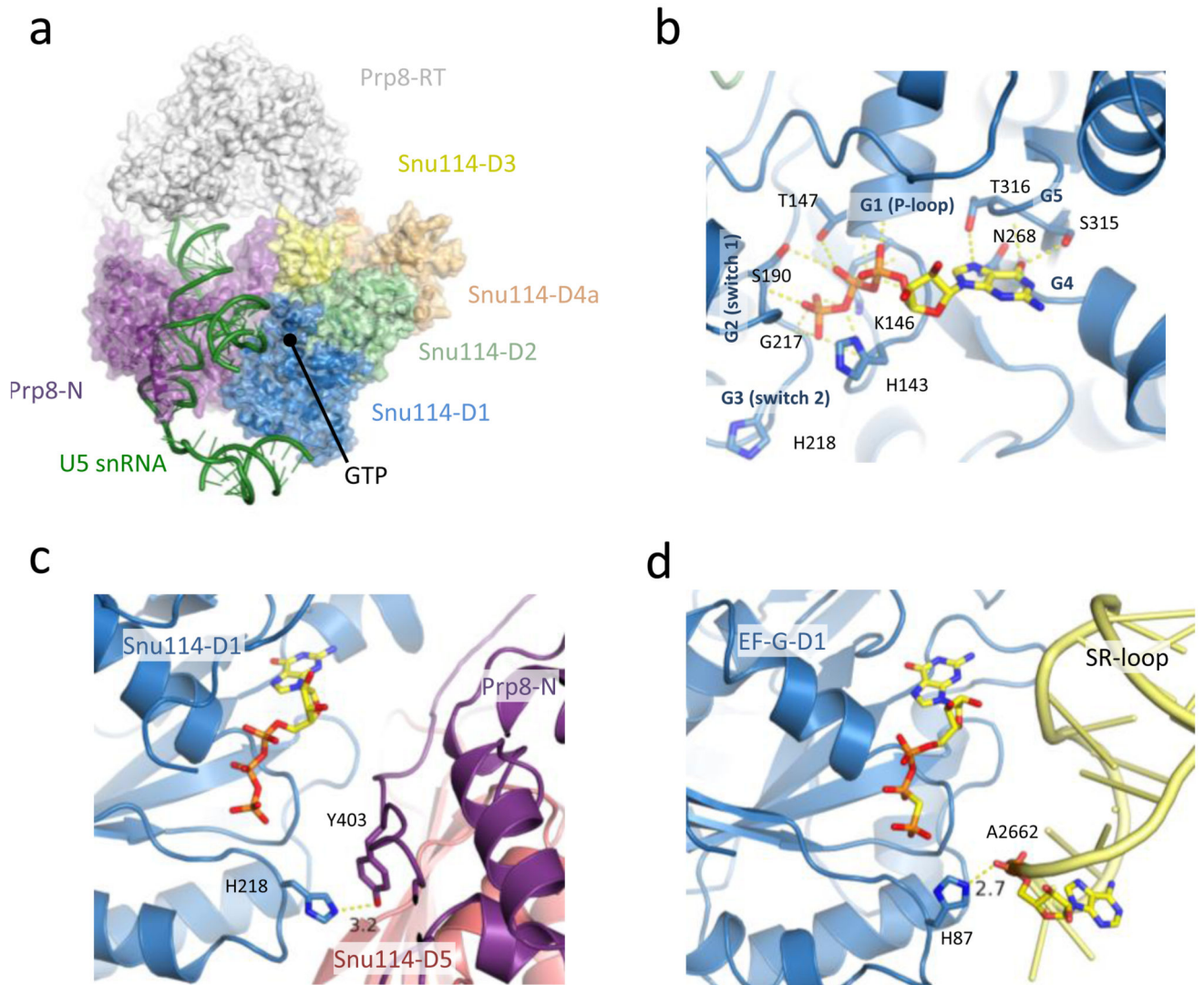
**Figure 1.** Three orthogonal views of a near-complete atomic model of the *Saccharomyces cerevisiae* U4/U6.U5 tri-snRNP  
Inset shows four subdomains.



**Figure 2. Prp8 and U4/U6 and U5 snRNAs**

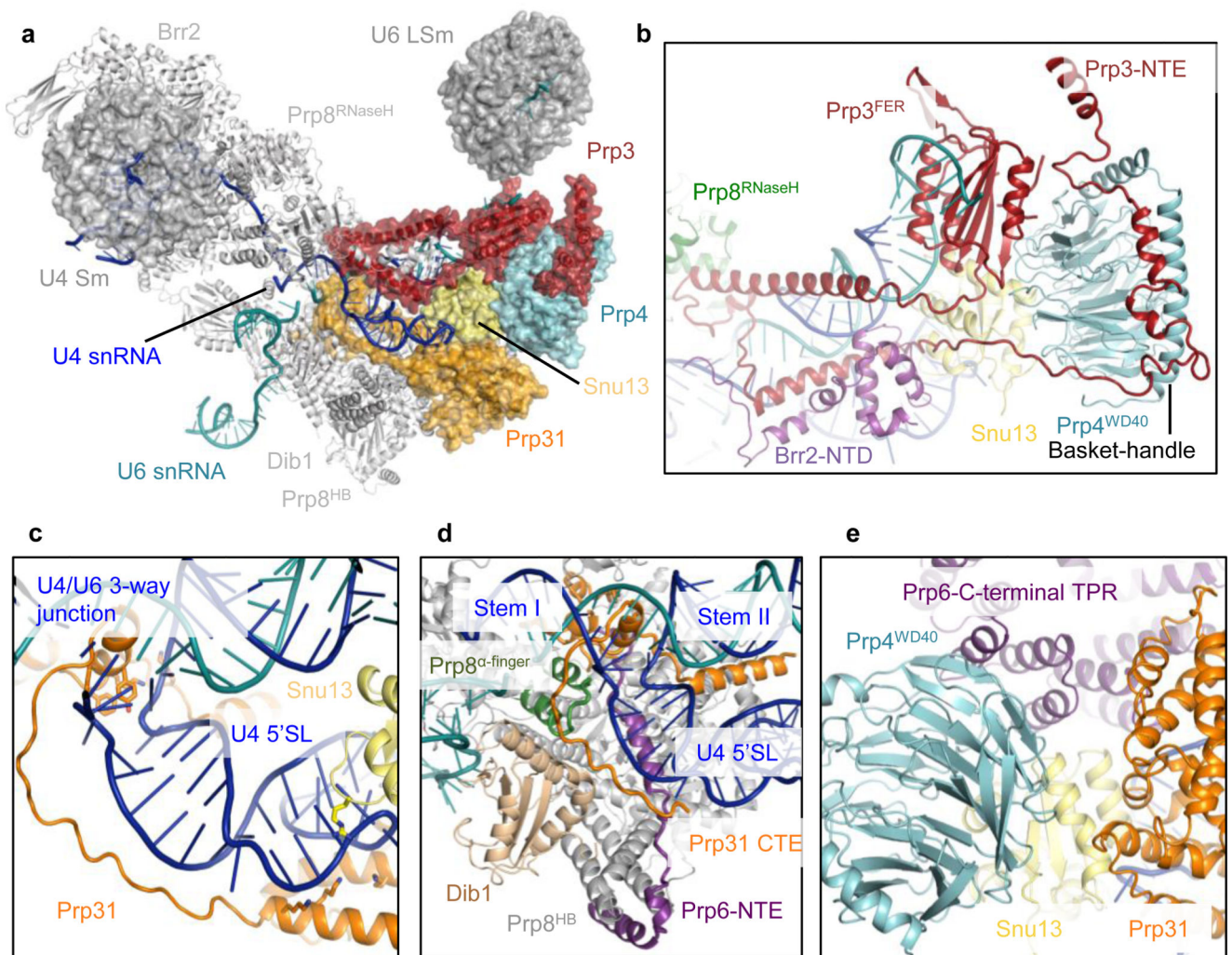
**a**, Domain structure of Prp8<sup>24</sup>: HB, helix bundle; RT, Reverse Transcriptase-like; Endo, Endonuclease-like; RH, RNaseH-like; JM, Jab1/MPN. **b**, Prp8 makes extensive interactions with U4/U6 and U5 snRNAs. **c**, The α-helix (residues 703-735) of the N-terminal domain fits into the minor groove of U5 snRNA and an extended polypeptide (residue 535-543) fits into the major groove on the opposite face, harnessing the RNA helix firmly in place. **d**, orthogonal view. U5 snRNA loop 1 interacts with the single-stranded region of U6 snRNA. **e**, The region around the ACAGAGA sequence forms a hairpin and is sandwiched between the Large and N-terminal domains of Prp8 and Dib1.





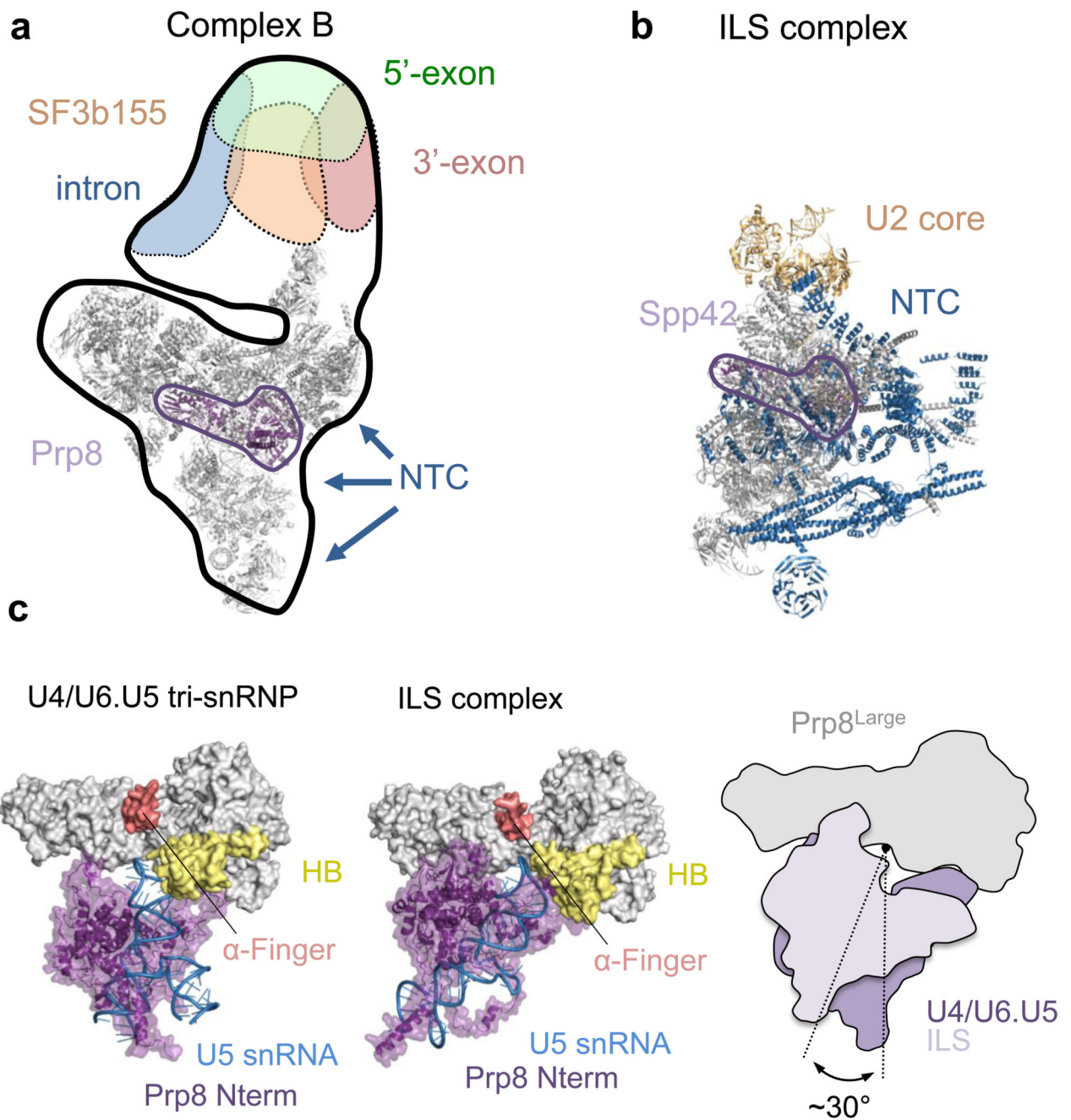
**Figure 3. Snu114 and its interaction with Prp8 and U5 snRNA**

**a**, Snu114, the N-terminal domain of Prp8 and U5 snRNA stems I and II form a stable domain in the Foot domain in U4/U6.U5 tri-snRNP. GTP is bound in the GTPase active site at the interface with the Prp8 N-terminal domain. **b**, Canonical interactions of GTP with surrounding residues in Snu114. **c**, The catalytic His218, hydrogen bonded to Tyr403 in Prp8, points away from the GTP  $\gamma$ -phosphate. **d**, Activation of EF-G GTPase upon binding to the sarcinricin (SR) loop in the ribosome. His87 moves closer to the  $\gamma$ -phosphate and places a water molecule<sup>29</sup>.



**Figure 4. Interactions of U4/U6 snRNAs with proteins**

**a**, Overview of U4/U6 di-snRNP. **b**, the extraordinary structure of Prp3 and its multiple interactions with U4/U6 snRNA, Prp4, Snu13, the RNaseH-like domain of Prp8, Brr2 N-terminal domain and the LSm core domain. **c**, The C-terminal region of Prp31 extends along U4 snRNA 5'-stem towards the three-way junction. **d**, The C-terminal extension of Prp31 makes multiple interactions with U4/U6 snRNAs, Dib1, Prp8  $\alpha$ -finger and the N-terminal extension of Prp6. **e**, the Prp4 WD40 domain and Prp31 interact with the C-terminal TPR domain of Prp6.



**Figure 5. B complex formation and activation mechanism**

**a**, U4/U6.U5 tri-snRNP fits into the EM envelope of human complex B<sup>45</sup> (reproduced from ref. 45 with permission), showing that U2 snRNP binds near the LSm core domain, Prp6 and Prp3. **b**, Overlay of the Prp8 large domain between tri-snRNP and the ILS<sup>19</sup> shows how NTC/NTR might bind to complex B and interact with U2 snRNP so that U2 snRNP can be passed to the NTC/NTR complex. **c**, A comparison of the tri-snRNP and the ILS<sup>19</sup> structures shows rotation of the Foot domain with respect to the Prp8 large domain. Upon rotation,

Prp8 residues 602-614 will clash with Dib1 and ACAGAGA helix, causing them to dissociate thus liberating the ACAGAGA sequence to bind the 5'-splice site.