
Accelerated Article Preview

Mechanism for the initiation of spliceosome disassembly

Received: 15 December 2023

Accepted: 19 June 2024

Accelerated Article Preview

Cite this article as: Vorländer, M. K. et al. Mechanism for the initiation of spliceosome disassembly. *Nature* <https://doi.org/10.1038/s41586-024-07741-1> (2024)

Matthias K. Vorländer, Patricia Rothe, Justus Kleifeld, Eric Cormack, Lalitha Veleti, Daria Riabov-Bassat, Laura Fin, Alex W. Phillips, Luisa Cochella & Clemens Plaschka

This is a PDF file of a peer-reviewed paper that has been accepted for publication. Although unedited, the content has been subjected to preliminary formatting. Nature is providing this early version of the typeset paper as a service to our authors and readers. The text and figures will undergo copyediting and a proof review before the paper is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers apply.

1 **Mechanism for the initiation of spliceosome disassembly**

2

3 Matthias K. Vorländer^{1,#}, Patricia Rothe^{1,2,#}, Justus Kleifeld^{1,2,#}, Eric Cormack³, Lalitha
4 Veleti^{1,2}, Daria Riabov-Bassat¹, Laura Fin¹, Alex W. Phillips¹, Luisa Cochella^{1,3,*}, and
5 Clemens Plaschka^{1,*}

6

7 *¹Research Institute of Molecular Pathology (IMP), Vienna BioCenter (VBC), Campus-*
8 *Vienna-Biocenter 1, 1030, Vienna, Austria.*

9 *²Vienna BioCenter PhD Program, Doctoral School of the University of Vienna and Medical*
10 *University of Vienna, 1030, Vienna, Austria*

11 *³Department of Molecular Biology and Genetics, Johns Hopkins University School of*
12 *Medicine, Baltimore, Maryland, USA.*

13 *#These authors contributed equally.*

14 **Correspondence to: mcochell1@jhmi.edu, clemens.plaschka@imp.ac.at*

15 **Pre-mRNA splicing requires the assembly, remodeling, and disassembly of the multi-**
16 **megadalton ribonucleoprotein complex called the spliceosome¹. Recent studies have**
17 **shed light on spliceosome assembly and remodeling for catalysis²⁻⁶, but the mechanism**
18 **of disassembly remains unclear. Here, we report 2.6 to 3.2 Å resolution cryo-electron**
19 **microscopy structures of nematode and human terminal intron-lariat spliceosomes**
20 **along with biochemical and genetic data. Our results uncover how four disassembly**
21 **factors and the conserved RNA helicase DHX15 initiate spliceosome disassembly. The**
22 **disassembly factors probe large inner and outer spliceosome surfaces to detect the**
23 **release of ligated mRNA. Two of these factors, TFIP11 and C19L1, and three general**
24 **spliceosome subunits, SYF1, SYF2 and SDE2, then dock and activate DHX15 on the**
25 **catalytic U6 snRNA to initiate disassembly. U6 thus controls both the start⁵ and end of**
26 **pre-mRNA splicing. Taken together, our results explain the molecular basis of**
27 **canonical spliceosome disassembly and provide a framework to understand general**
28 **spliceosomal RNA helicase control and the discard of aberrant spliceosomes.**

29
30 The spliceosome is a dynamic and multi-megadalton ribonucleoprotein complex that excises
31 introns from messenger RNA precursors (pre-mRNAs) to generate mRNA¹. This
32 macromolecular machine forms anew on each intron via the orchestrated assembly of five
33 small nuclear ribonucleoprotein particles (snRNPs; U1, U2, U4, U5, U6), the nineteen
34 complex (NTC), nineteen related complex (NTR), and other non-snRNP proteins⁷⁻⁹. After
35 several dramatic, ATP-driven changes the U6 snRNA forms the spliceosome's RNA-based
36 catalytic center. Splicing then proceeds, resulting in a post-catalytic (P) spliceosome, in
37 which the ligated mRNA and the excised intron-lariat are bound to the active site RNA
38 network. The ligated mRNA is then released, yielding an intron-lariat spliceosome (ILS) that
39 represents the terminal spliceosome state⁷⁻⁹. The ILS and its catalytic center must then be

40 disassembled to trigger the recycling of U2 and U5 snRNPs, U6 snRNA, NTC, NTR, and
41 non-snRNP proteins for the next round of splicing and to degrade or process the intron-lariat
42 RNA (Fig. 1a). Aberrant spliceosomes resulting from splicing errors are thought to be
43 discarded by a similar process¹⁰.

44 Spliceosome disassembly requires the essential RNA helicase DHX15 (yeast Prp43),
45 a member of the DExH-box ‘translocases’. Although DHX15 has been extensively studied^{11–}
46 ¹⁵, the mechanisms governing its regulation are poorly understood. Notably, DHX15 lacks
47 intrinsic RNA sequence specificity and its RNA target within the spliceosome remains
48 unclear, with recent studies suggesting either the intron-lariat RNA, U2 snRNA, or U6
49 snRNA^{10,13}. Four conserved proteins have additionally been implicated in metazoan ILS
50 disassembly: TFIP11 (yeast Ntr1), PAXBP1, C19L1 (yeast Drn1), and C19L2 (referred to by
51 their human names; Extended Data Table 1)^{16–22}. However, their functional roles in
52 disassembly are unclear. Owing to the incompleteness of available structural studies of
53 yeast^{19,23} and human²² ILS complexes, it further remains unknown how DHX15 and these
54 additional factors act together to orchestrate the disassembly of terminal spliceosomes but not
55 of earlier spliceosome intermediates.

56 Here we addressed these long-standing questions enabled by new high-resolution
57 cryo-electron microscopy (cryo-EM) structures of nematode and human ILS spliceosomes.
58 We combined our structural analysis with biochemical and genetic data, and a revised
59 structure of the human P complex spliceosome, to reveal the functions of all four disassembly
60 factors and the regulation of the disassembly helicase DHX15. Based on our results, we
61 derive a model for disassembly of the ILS via U6, the catalytic center snRNA, thus providing
62 critical insights into the stepwise dismantling of the terminal spliceosome and the discard of
63 aberrant spliceosomes.

64

65 ILS structures at high resolution

66 To understand how the four disassembly factors and DHX15 specifically trigger spliceosome
67 disassembly in metazoans, we determined the cryo-EM structure of a complete ILS complex.
68 To overcome past challenges in ILS structure determination in yeast and humans, we purified
69 spliceosomes from another metazoan, the nematode *C. elegans* (*Ce*). We chose *Ce*
70 spliceosomes as a model system because *Ce* introns are remarkably short, with a median
71 length of 65 nucleotides, which could help to increase spliceosome complex stability and the
72 steady-state abundance of specific states²⁴. Moreover, *Ce* spliceosomes contain an identical
73 protein composition to their human counterparts, unlike budding yeast spliceosomes, another
74 extensively used model system (Extended Data Fig. 1g). For spliceosome purification, we
75 endogenously tagged the NTC subunit PRP19 with an N-terminal 3xFLAG-tag using
76 CRISPR-Cas9 and obtained spliceosomes from the extract of ~12 million adult worms. *Ce*
77 spliceosomes sedimented around 35S in sucrose gradients and contained known spliceosome
78 components, including the disassembly factors, TFIP11, PAXBP1, C19L1, C19L2, and
79 DHX15 (Extended Data Fig. 1b,e). We collected cryo-EM data from this sample and
80 obtained four million cryo-EM single particle images. Analysis of these particles revealed
81 two ILS complexes (Supplementary Data Fig. 1). The general enrichment of the *Ce* ILS from
82 worm extract mirrors the enrichment of the fission yeast ILS from the extract of
83 logarithmically growing cells²⁵, suggesting that ILS disassembly could be a rate-limiting
84 splicing step in certain conditions. Extensive image classification and local refinements of the
85 *Ce* ILS data yielded twenty-seven cryo-EM maps, from which we assembled two composite
86 maps that capture the ILS in two major states. We arranged these states based on their
87 compositional complexity and named them ‘primed’ (ILS’, maps 1-7 and 8-15) and ‘double-
88 primed’ (ILS’’, maps 2-7 and 16-27), with the two structures differing in the bound
89 disassembly factors (Fig. 1b, Extended Data Fig. 1f, Extended Data Table 2, Video S1, S2,

90 Supplementary Data Fig. 1, 2, Supplementary Table 1). Local map resolutions reached 2.6 Å
91 in the spliceosome's core, with most densities resolved at a nominal resolution better than 3.5
92 Å (Fig. 1b, Supplementary Data Fig. 2). Together, these maps reveal a metazoan ILS in
93 unprecedented detail (Fig. 1b, Supplementary Data Fig. 1, 2). AlphaFold2 Multimer
94 predictions combined with manual building allowed us to prepare near-complete atomic
95 models of *Ce* ILS' and ILS'' complexes (Extended Data Fig. 2c, d). In those ILS regions,
96 where high-resolution human densities are available *Ce* and human maps are nearly identical
97 (Extended Data Fig. 2a, b), indicating that structural insights from *Ce* reveal features
98 common to metazoan spliceosomes.

99 The *Ce* ILS' and ILS'' models are substantially improved over human catalytic stage
100 spliceosome structures that share the 8-subunit NTC and 9-subunit NTR complexes
101 (Extended Data Fig. 2c, d). These improved models revealed the location of the essential
102 NTR protein CCDC12 and previously unresolved parts of SYF1, SYF2, SYF3, ISY1,
103 CDC5L, SDE2, PRP19, SPF27, SNW1, AQR, PPIE, PLRG1 (Extended Data Fig. 2d,
104 Supplementary Data Figs 1, 2), which collectively contribute to NTC and NTR formation. In
105 the ILS', we also resolved the disassembly factors TFIP11 and PAXBP1. In the ILS'', we
106 additionally resolved the disassembly factors C19L1, C19L2, and DHX15.

107 The ILS' and ILS'' states may thus be sequential on-pathway intermediates for
108 disassembly, consistent with evidence for the step-wise formation of the ILS'' (see
109 Supplementary Text 1). Alternatively, the ILS' could form through the partial breakdown of
110 the ILS'' during sample preparation, which would however not affect the conclusions drawn
111 from the structural comparisons of ILS' with ILS'' (see Supplementary Text 1). Taken
112 together, the high-resolution cryo-EM structures of the *Ce* spliceosome suggest that
113 recognition of the ILS for disassembly occurs in two steps.

114

115 **Recognition of outer ILS surfaces**

116 The two disassembly factors TFIP11 and PAXBP1 are essential for viability in human cells
117 (<https://depmap.org/portal/>) and in *C. elegans*²⁶⁻²⁹. While they are necessary for ILS
118 disassembly, their specific roles had been unclear. In the ILS' structure, TFIP11 and
119 PAXBP1 bind each other to form a ~250 Å long rod-like structure that spans the length of the
120 ILS exterior surface (Fig. 2a, b). In TFIP11 we observe five structured regions, the 'G-patch',
121 'hairpin', 'coiled-coil', a helical repeat 'GCFC' domain, and a 'C-terminal domain' (CTD)
122 (Fig. 2a, Extended Data Fig. 3b). In the ILS' complex, all these regions engage in specific
123 contacts (Fig. 2b, c), except for the TFIP11 G-patch, which is mobile. G-patch domains bind
124 DHX15 with high affinity^{30,31}, suggesting that DHX15 is already tethered to the ILS' but
125 remains mobile, which is also consistent with similar abundances of TFIP11 and DHX15 in
126 our mass spectrometry data of *Ce* spliceosomes (Extended Data Fig. 1e). The TFIP11 G-
127 patch is followed by a β -hairpin, which binds the PRP8 RNaseH domain. The TFIP11 hairpin
128 connects to a long helix (α 3), which leans against the PRP8 RNaseH domain, and forms an
129 anti-parallel coiled-coil with PAXBP1, consistent with a reported TFIP11–PAXBP1
130 interaction in *Ce* and humans^{20,26}. TFIP11 then continues with its α -helical repeat domain
131 (residues 391-723), which binds the U5 snRNP subunit SNU114 domains III (residues 696-
132 829), IV (residues 830-910), and V (residues 911-945) and connects via a linker to the
133 TFIP11 CTD (residues 759-830). This CTD binds between SNU114 domains II (residues
134 587-665) and III and CWC15 (residues 58-108). PAXBP1 contains an additional α -helical
135 'GCFC' domain (residues 382-809), which contacts the PRP8 JAB1/MPN domains (residues
136 2065-2329) as well as the BRR2 plug element (residues 97-173). Metazoan BRR2 could thus
137 aid disassembly by acting as a protein scaffold for TFIP11–PAXBP1, without using its RNA
138 translocase activity³². In contrast, budding yeast Brr2 was reported to aid ILS disassembly via
139 its ATPase activity³³. However, yeast Brr2 was also reported to be repressed by Ntr2 (yeast

140 PAXBP1) *in vitro*³⁴. The role of yeast Brr2 in spliceosome disassembly thus warrants further
141 investigation.

142 Taken together, the TFIP11–PAXBP1 dimer binds multiple exterior surfaces of the
143 ILS'. Through these interactions, TFIP11–PAXBP1 would deliver the disassembly helicase
144 DHX15 to the spliceosome. However, even though TFIP11 tethers DHX15 to the ILS', this
145 tethering is apparently not sufficient for DHX15 to dock onto the complex, suggesting that
146 this requires additional structural changes.

147

148 **Recognition of inner ILS surfaces**

149 In the ILS'', we could resolve the remaining disassembly factors C19L1, C19L2, and a
150 docked DHX15. Binding of these three proteins follows the dissociation of BRR2 from
151 PAXBP1 and a substantial movement of TFIP11–PAXBP1 and PRP8 regions. We note that
152 the location of BRR2 in the ILS' and its dissociation in the ILS'' may prime BRR2 for
153 binding of the recycling chaperone TSSC4 and thus aid the recycling of the 20S U5 snRNP
154 for the next splicing round^{35,36}. During the ILS' to ILS'' transition, TFIP11 bends at a hinge
155 located between the TFIP11–PAXBP1 coiled-coil and the TFIP11 helical domain (Fig. 2b, c),
156 swinging the TFIP11–PAXBP1 coiled-coil outwards from the ILS by ~20° degrees
157 (Extended Data Fig. 4a-c). As a result, the TFIP11–PAXBP1 coiled-coil helices swings
158 around the TFIP11 hinge, displacing the PRP8 JAB1/MPN–BRR2 complex and re-
159 positioning the connected TFIP11-hairpin–PRP8 RNaseH complex by ~25 Å from its ILS'
160 position. The new space generated by movement of the PRP8 RNaseH domain is now
161 occupied by C19L1 and C19L2 (Extended Data Fig. 4a-e). Based on the mobility of the
162 TFIP11–PAXBP1 in the ILS' (Fig. 1b, Supplementary Fig. 2), we speculate that TFIP11–
163 PAXBP1 may stochastically sample its ILS'' conformation, allowing for C19L2-binding to
164 selectively stabilize the ILS'' conformation.

165 C19L2 makes extensive contacts to the newly oriented PRP8 RNaseH domain, the
166 PRP8 Large domain as well as the RNA-based active site, similar to the C19L2–spliceosome
167 interaction network observed in the partial fission yeast¹⁹ and human ILS structures²²
168 (Extended Data Figs 3d; 4b, j). Compared to the *Ce* ILS', the RNA densities near the RNA
169 active site and the nearby U2/U6 snRNA helix II are much better defined in the ILS''
170 (Extended Data Fig. 4g,h). ILS recognition by C19L2 may thus conformationally lock the
171 ILS RNA active site, which is expected to be mobile immediately after the release of mRNA
172 and branching-specific splicing factors from the P complex. This conformational lock is thus
173 likely a structural consequence of C19L2-binding, which may also play a role in ILS
174 disassembly.

175 In the ILS'' structure, we could also assign C19L1, which comprises an N-terminal
176 metallophosphatase domain (MMP) and two C-terminal CwfJ domains (CWFJ) (Fig. 2b, c,
177 Extended Data Fig 5, Supplementary Data Fig. 2). The C19L1 CWFJ domain binds C19L2
178 helices $\alpha 1$ and $\alpha 2$ (Extended Data Fig. 5a-d) and connects through a linker to its N-terminal
179 MMP domain, which binds the now docked DHX15 through conserved interfaces (Extended
180 Data Fig. 5e-j). Notably, a conserved peptide of SYF1, which we name the 'tether', binds in
181 between C19L1 and DHX15 (Extended Data Fig. 5e), consistent with their *in vitro*
182 interaction using recombinant proteins (Extended Data Fig. 5h) and the yeast-two-hybrid
183 interaction of their yeast orthologs¹⁶. The SYF1 tether may thereby enhance the C19L1–
184 DHX15 affinity and facilitate DHX15 docking in the ILS''.

185 Taken together, C19L1–C19L2 are likely to be part of a two-factor authentication
186 mechanism to ensure that the RNA helicase DHX15 docks at the ILS only when both inner
187 and outer ILS surfaces have been recognized by all disassembly factors.

188

189 **Specificity of ILS recognition**

190 To prevent premature disassembly in the P complex spliceosome and earlier spliceosome
191 states, ILS recognition must be specific. However, our *Ce* spliceosome preparation did not
192 contain the P complex, precluding a meaningful comparison with an ILS from the same
193 species.

194 We therefore set out to determine the architecture of a complete human ILS and to
195 prepare a revised model of available, but incomplete human P complex structures^{22,37} (Fig.
196 3a, b, Supplementary Data Fig. 3). Since TFIP11 and PAXBP1 were missing from available
197 human ILS structures²², we overexpressed GFP-tagged TFIP11 in human K562 cells and
198 purified TFIP11-bound spliceosomes from nuclear extract (Extended Data Fig. 6). We
199 collected 40,043 micrographs of human TFIP11-containing spliceosomes from which we
200 obtained an overall density of a human ILS at 3.5 Å resolution (Extended Data Fig. 6,
201 Supplementary Data Fig. 4). Surprisingly, TFIP11–PAXBP1 had largely dissociated from
202 these complexes. However, further particle classification could reveal a subset of ~10,000
203 ILS particles that did contain TFIP11–PAXBP1 density with nominal resolution of 8 Å. This
204 human ILS architecture is in excellent agreement with its *Ce* ILS'' counterpart, including the
205 locations of TFIP11–PAXBP1, C19L2, and weak densities for DHX15 and C19L1 (Extended
206 Data Fig. 6). The high-resolution regions of the human ILS structure, including C19L2,
207 closely resembled cryo-EM densities of previously reported partial human ILS structures²²
208 (Extended Data Fig. 2b), which also showed weak densities for DHX15, but lacked the
209 disassembly factors TFIP11–PAXBP1 and C19L1. By combining our new cryo-EM data with
210 AlphaFold2 Multimer predictions and the *Ce* ILS'' model, we generated an integrative
211 architectural model of the human ILS'' (Fig. 3b, Supplementary Data Figs 5, 6, Video S3).
212 The quality of the human ILS densities was noticeably lower compared to *Ce* ILS densities in
213 peripheral regions, where disassembly factors bind. This could reflect differences in complex
214 stability between species or in complex heterogeneity due to the variable intron length in

215 humans (~7,000 nucleotides) compared to *Ce* (~65 nucleotides). Despite these differences in
216 quality, the locations of the disassembly factors are conserved between human and *Ce* ILS
217 complexes (Figs 2b, 3b).

218 To prepare a revised model of the human P complex, we next combined available
219 cryo-EM densities and models of human B^{act}, C, C*, and P complexes^{22,37-39}, and the
220 *Ce* ILS'' (solved here), with AlphaFold2 Multimer and manual building (Fig. 3a,
221 Supplementary Data Fig 3, Video S4). Compared to available P complex models, our revised
222 structure additionally contains CCDC12, PPWD1, TLS1, NOSIP, STEEP1, FAM50A, ESS2,
223 ISY1 and extensions in BRR2, SYF1, SYF2, SYF3, CDC5L, SDE2, PRP19, SPF27, and
224 SNW1, which together result in the most complete P complex model available to date.

225 Structural comparisons of human P complex and ILS revealed the basis for specific
226 ILS disassembly (Fig. 3, Extended Data Fig. 7). In the P complex, the mRNA
227 ribonucleoprotein complex (mRNP) 5'-end, which comprises mRNA 5'-exon and the
228 mRNA-bound exon junction complex, and the spliceosome subunits SRRM2, CWC22 would
229 clash with the TFIP11 α -helical domain. The re-oriented PRP8 JAB1/MPN-BRR2 complex
230 and BRR2-bound PPWD1 would further clash with PAXBP1 (Fig. 3c, Extended Data Fig. 7).
231 These clashing P complex proteins are also present in all earlier catalytic spliceosomes (B^{act}
232 to C* complex), explaining how binding of TFIP11-PAXBP1 discriminates against
233 spliceosome states before the ILS. In addition, the C* and P complex-specific proteins
234 CACTIN, FAM50A, SLU7, and TLS1 and the mRNP 3'-end containing mRNA 3'-exon
235 would clash with C19L2 (Fig. 3c, Extended Data Fig 7). Taken together, the disassembly
236 factors distinguish the ILS from earlier spliceosomal states by sensing the release of the
237 spliced mRNA, mRNA-associated proteins, and catalysis-specific spliceosome proteins.

238 To gain insights into the conservation of ILS recognition, we compared our *Ce* ILS'
239 and ILS'' structures with the budding yeast ILS structure²³. This revealed surprising

240 differences in how metazoan and budding yeast ILS complexes are recognized by their
241 disassembly factors, particularly among TFIP11–PAXBP1 and their yeast homologs, Ntr1–
242 Ntr2 (Extended Data Fig. 3). The yeast Ntr1 (ref.²³) and *Ce* TFIP11 CTDs bind similarly to
243 either yeast Snu114–Cwc23, where Cwc23 is yeast-specific, or the metazoan SNU114–
244 CWC15. However, the remaining interfaces between TFIP11 and the *Ce* ILSs are entirely
245 different from its yeast counterpart (Extended Data Fig. 3e). Further, although we identify
246 yeast Ntr2 as the homolog of *Ce* and human PAXBP1, Ntr2 binds the budding yeast ILS
247 through non-overlapping surfaces compared to *Ce* and human PAXBP1 (Extended Data. Fig.
248 3). Nevertheless, the alternate yeast Ntr1–Ntr2 interfaces also allow for specific ILS
249 recognition, compared to yeast catalytic spliceosome states²³. We speculate that owing to the
250 increased protein complexity of metazoan mRNPs (e.g. the presence of the exon junction
251 complex) and of the spliceosome, TFIP11–PAXBP1 may have evolved to recognize
252 metazoan mRNP and ILS features that discriminate metazoan catalytic-stage from terminal
253 spliceosomes through a larger surface area. Notably, while the C19L2–spliceosome
254 interactions are highly similar between fission yeast, *Ce*, and human complexes (Extended
255 Data Fig. 3d), C19L2 is absent from budding yeast, which instead contains Cwc23 that binds
256 at a different site of the budding yeast ILS. Drn1, the yeast homolog of C19L1, was not
257 observed in any yeast ILS structure, precluding a structural comparison to the *Ce* and human
258 ILS. Despite the apparent differences in ILS recognition between species, all employ multiple
259 independent factors to verify the release of mRNA 3'-exons from the spliceosome, while *Ce*
260 and human complexes in addition verify the release of the mRNA 5'-exon. Once release is
261 verified, the highly conserved disassembly helicase DHX15 (yeast Prp43) locates to the same
262 region in the ILS in various species, indicating conservation of the disassembly mechanism
263 (Extended Data Fig. 3, ref.^{19,22,23}). However, owing to the low resolution of these regions in

264 previous ILS structures^{19,22,23}, the mechanism of DHX15-mediated disassembly, and its RNA
265 target, remained unclear.

266

267 **DHX15 acts on U6 for disassembly**

268 In the *Ce* ILS'' structure, we resolved DHX15 to 3.9 Å resolution, revealing how DHX15 is
269 positioned and activated to disassemble the spliceosome (Fig. 4a, b). DHX15 is an RNA
270 helicase that translocates in a 3'-5' direction along an RNA substrate, and comprises RecA1
271 and RecA2 lobes, and a C-terminal domain (CTD). In the ILS'', DHX15 and the bound RNA
272 are in a relaxed, ATP-unbound conformation (Extended Data Fig. 8f). At 3.9 Å resolution,
273 we could unambiguously trace U6 snRNA from U2/U6 Helix II into the DHX15 active site
274 (Extended Data Fig. 4i), revealing that the spliceosome's catalytic center snRNA U6 (ref.⁴⁰)
275 is the target for metazoan spliceosome disassembly. This is consistent with biochemical data
276 in budding yeast¹⁰, and in contrast to reports that proposed that U2 snRNA or the intron-lariat
277 RNA play this role¹³. The TFIP11 G-patch bridges the DHX15 RecA2 and CTD domains,
278 similar to crystal structures of human RNA-unbound DHX15 in complex with the ribosomal
279 biogenesis factor NKRF1 (ref.³⁰) and the human splicing quality control factor SUGP1
280 (ref.⁴¹) (Extended Data Fig. 8a, b). Consistent with our structure, the recombinant *Ce* TFIP11
281 G-patch stimulates the *Ce* DHX15 helicase activity approximately 30-fold *in vitro* (Extended
282 Data Fig. 8c, d), akin to human NKRF1 and yeast Ntr1 (ref.^{12,30}). The DHX15 RecA2 lobe is
283 additionally bound by the C19L1 MMP domain and the SYF1 'tether'. The DHX15 CTD
284 domain is bound to the NTR subunits SYF2 and SDE2, which are both essential for viability
285 in humans (<https://depmap.org/portal/>) (Fig. 4b, c). These combined interactions, possibly
286 together with a peptide of SYF3 (Extended Data Fig. 5k-o), guide DHX15 onto the U6
287 snRNA 3'-end, from where DHX15 could translocate along U6 snRNA to disassemble the
288 catalytic center and the spliceosome.

289 These observations suggest that the *Ce* ILS'' is poised for disassembly but lacks ATP
290 to initiate translocation (Extended Data Fig. 8a-f). To test this, we performed an *in vitro* ILS
291 disassembly assay. We purified spliceosomes on beads via a 3xFLAG-PRP19 pulldown and
292 then added ATP to bead-bound spliceosomes (Extended Data Fig. 8g). We would expect that
293 ATP addition dissociates the ILS'' into a complex comprising PRP19, the remainder of the
294 NTC, NTR, and the U5 snRNP, but lacking the disassembly factors, U2 snRNP, and U6
295 snRNA. Consistent with this expectation, we observed the depletion of DHX15, C19L1,
296 C19L2, TFIP11, PAXBP1, and U2 snRNP proteins, by quantitative mass spectrometry
297 (Extended Data Fig. 8f). We thus conclude that the ILS'' is competent for *in vitro*
298 disassembly, but we do not exclude the possibility that additional factors, such as DBR1
299 (ref.^{16,42}), might contribute to ILS disassembly *in vivo*. Since C19L1 is located near the
300 branch site in the ILS'' and binds DBR1 *in vitro*¹⁸, C19L1 could for example target DBR1 to
301 initiate intron-lariat RNA decay.

302 The choice of U6 snRNA as the target for DHX15 has implications for ILS
303 disassembly and spliceosome discard during quality control. After transcription by RNA
304 polymerase III, the U6 snRNA 3'-end is extended by oligo-uridylation^{43,44}. In the ILS'', the
305 U6 snRNA 3'-end is bound to DHX15, indicating that post-transcriptional oligo-uridylation
306 of U6 is not only required during spliceosome assembly⁴³, but might also be important for
307 spliceosome disassembly by providing a single-stranded RNA 3'-end that acts as a landing
308 pad for DHX15 (Fig. 4a, Extended Data Fig. 4f).

309 In our structure, the U6 snRNA 3'-end emerges from U2/U6 snRNA helix II (U2/U6
310 helix II; Extended Data Fig. 4h). U2/U6 helix II is embraced by the NTR subunits SYF2 and
311 SDE2 (Fig. 4c), which together may use three mechanisms to promote disassembly. First,
312 SYF2 and SDE2 spatially fix the position of U2/U6 helix II on the NTR complex, near
313 DHX15. Second, the positively charged SYF2 helix $\alpha 2$ ('wedge') and SDE2 helix $\alpha 2$ ('lid')

314 act as a wall that may protect the U2 snRNA 5'-end from the environment and guide the U6
315 snRNA 3'-end towards DHX15 (Fig. 4b, Extended Data Fig. 9a-b). The SYF2 residue W66
316 stacks onto the first nucleotide of the U2 snRNA, which together with the 'wedge' and 'lid'
317 domains may further facilitate separation of U2/U6 helix II upon the translocation of DHX15
318 (Fig. 4c). Third, SYF2 helix $\alpha 1$ and SDE2 helix $\alpha 3$, jointly named the 'anchor', and the
319 SYF1 'tether' may aid in docking DHX15 at the NTR, near the U6 snRNA 3'-end. We
320 propose that the NTR subunits SYF1, SYF2 and SDE2 contribute to priming the spliceosome
321 for eventual disassembly.

322 To probe the importance of SYF2, we mutated *Ce syf-2 in vivo*. Consistent with the
323 ILS'' structure, deletion of the combined *syf-2* helices $\alpha 1$ and $\alpha 2$ was lethal (Extended Data
324 Fig. 9d, e). Deletion of only *syf-2* helix $\alpha 1$ showed a partial loss of function. We could raise
325 homozygous animals carrying the *syf-2* helix $\alpha 1$ deletion at 20°C but these animals were
326 cold-sensitive, as has been reported for other splicing defective strains⁴⁵⁻⁴⁷ (Extended Data
327 Fig. 9f). Moreover, the partial loss of viability of this allele was enhanced by knock-down of
328 *sde-2* by RNAi, supporting a joint role for these proteins (Extended Data Fig. 9g).

329 Since SYF1, SYF2, and SDE2 join the spliceosome during its catalytic activation,
330 these proteins could also be used for splicing quality control, to assist in the discard of
331 aberrantly formed spliceosomes. Thus, we speculate that spliceosome disassembly and
332 discard pathways may not only share DHX15 as the disassembly helicase^{11,14,41,48}, but also a
333 common DHX15-binding site and RNA target¹⁰. Disassembly and discard pathways would
334 nevertheless differ depending on the respective G-patch protein⁴⁹ and putative accessory
335 factors for the specific multi-factor authentication of aberrant spliceosomes, before
336 committing to disassembly via the U6 snRNA.

337

338 **Conclusions**

339 Here we presented structures of the *Ce* ILS', *Ce* ILS'', human P complex, and human ILS'',
340 which together reveal the conserved architecture of the metazoan spliceosome and a
341 mechanism to initiate disassembly of the terminal spliceosome (Fig. 5, Video S5). We
342 establish *Ce* as a novel model organism for the structural study of pre-mRNA splicing,
343 yielding substantially improved cryo-EM densities over their human counterparts. Our results
344 reveal an elaborate 'multi-factor authentication' system that involves four disassembly factors
345 and three general spliceosome subunits, which collectively probe spatially distant surfaces to
346 identify the terminal spliceosome. These surfaces are inaccessible in earlier splicing steps,
347 revealing how specific spliceosome disassembly is achieved and pre-mature disassembly is
348 prevented. After successful authentication of the terminal state, these proteins guide DHX15
349 onto U6 snRNA and activate DHX15 for disassembly. DHX15 would then translocate along
350 U6 snRNA to unfold the spliceosome's catalytic RNA center. Thus, both the beginning⁵ and
351 the end of pre-mRNA splicing are orchestrated by the three-dimensional organization of U6
352 snRNA.

353 Our work also reveals how structural cues are read out and integrated in the terminal
354 spliceosome prior to irreversible remodeling by DHX15. We speculate that similar principles
355 govern the regulation of other DExH-box RNA translocases that act during splicing
356 progression, discard of aberrant spliceosomes, and other aspects of cellular RNA metabolism.

357

358 **REFERENCES**

- 359 1. Will, C. L. & Lührmann, R. Spliceosome structure and function. *Cold Spring Harb*
360 *Perspect Biol* **3**, a003707 (2011).
- 361 2. Bertram, K. *et al.* Cryo-EM Structure of a Pre-catalytic Human Spliceosome Primed for
362 Activation. *Cell* **170**, 701-713.e11 (2017).
- 363 3. Charenton, C., Wilkinson, M. E. & Nagai, K. Mechanism of 5' splice site transfer for
364 human spliceosome activation. *Science* **364**, 362–367 (2019).
- 365 4. Tholen, J., Razew, M., Weis, F. & Galej, W. P. Structural basis of branch site recognition
366 by the human spliceosome. *Science* **375**, 50–57 (2022).
- 367 5. Townsend, C. *et al.* Mechanism of protein-guided folding of the active site U2/U6 RNA
368 during spliceosome activation. *Science* **370**, eabc3753 (2020).
- 369 6. Zhan, X., Yan, C., Zhang, X., Lei, J. & Shi, Y. Structures of the human pre-catalytic
370 spliceosome and its precursor spliceosome. *Cell Res* **28**, 1129–1140 (2018).
- 371 7. Kastner, B., Will, C. L., Stark, H. & Lührmann, R. Structural Insights into Nuclear pre-
372 mRNA Splicing in Higher Eukaryotes. *Cold Spring Harb Perspect Biol* **11**, a032417
373 (2019).
- 374 8. Wan, R., Bai, R., Zhan, X. & Shi, Y. How Is Precursor Messenger RNA Spliced by the
375 Spliceosome? *Annu Rev Biochem* **89**, 333–358 (2020).
- 376 9. Wilkinson, M. E., Charenton, C. & Nagai, K. RNA Splicing by the Spliceosome. *Annu*
377 *Rev Biochem* **89**, 359–388 (2020).
- 378 10. Toroney, R., Nielsen, K. H. & Staley, J. P. Termination of pre-mRNA splicing requires
379 that the ATPase and RNA unwindase Prp43p acts on the catalytic snRNA U6. *Genes Dev*
380 **33**, 1555–1574 (2019).
- 381 11. Feng, Q., Krick, K., Chu, J. & Burge, C. B. Splicing quality control mediated by DHX15
382 and its G-patch activator SUGP1. *Cell Rep* **42**, 113223 (2023).

- 383 12. Fourmann, J.-B., Tauchert, M. J., Ficner, R., Fabrizio, P. & Lührmann, R. Regulation of
384 Prp43-mediated disassembly of spliceosomes by its cofactors Ntr1 and Ntr2. *Nucleic*
385 *Acids Res* **45**, 4068–4080 (2017).
- 386 13. Fourmann, J.-B. *et al.* The target of the DEAH-box NTP triphosphatase Prp43 in
387 *Saccharomyces cerevisiae* spliceosomes is the U2 snRNP-intron interaction. *Elife* **5**,
388 e15564 (2016).
- 389 14. Koodathingal, P., Novak, T., Piccirilli, J. A. & Staley, J. P. The DEAH box ATPases
390 Prp16 and Prp43 cooperate to proofread 5' splice site cleavage during pre-mRNA
391 splicing. *Mol Cell* **39**, 385–395 (2010).
- 392 15. Maul-Newby, H. M. *et al.* A model for DHX15 mediated disassembly of A-complex
393 spliceosomes. *RNA* **28**, 583–595 (2022).
- 394 16. Garrey, S. M. *et al.* A homolog of lariat-debranching enzyme modulates turnover of
395 branched RNA. *RNA* **20**, 1337–1348 (2014).
- 396 17. Hang, J., Wan, R., Yan, C. & Shi, Y. Structural basis of pre-mRNA splicing. *Science* **349**,
397 1191–1198 (2015).
- 398 18. Masaki, S. *et al.* Identification of the specific interactors of the human lariat RNA
399 debranching enzyme 1 protein. *Int J Mol Sci* **16**, 3705–3721 (2015).
- 400 19. Yan, C. *et al.* Structure of a yeast spliceosome at 3.6-angstrom resolution. *Science* **349**,
401 1182–1191 (2015).
- 402 20. Yoshimoto, R., Okawa, K., Yoshida, M., Ohno, M. & Kataoka, N. Identification of a
403 novel component C2ORF3 in the lariat-intron complex: lack of C2ORF3 interferes with
404 pre-mRNA splicing via intron turnover pathway. *Genes Cells* **19**, 78–87 (2014).
- 405 21. Yoshimoto, R., Kataoka, N., Okawa, K. & Ohno, M. Isolation and characterization of
406 post-splicing lariat-intron complexes. *Nucleic Acids Res* **37**, 891–902 (2009).

- 407 22. Zhang, X. *et al.* Structures of the human spliceosomes before and after release of the
408 ligated exon. *Cell Res* **29**, 274–285 (2019).
- 409 23. Wan, R., Yan, C., Bai, R., Lei, J. & Shi, Y. Structure of an Intron Lariat Spliceosome
410 from *Saccharomyces cerevisiae*. *Cell* **171**, 120–132.e12 (2017).
- 411 24. Howe, Kevin, Iast, Spieth, John, Lawson, Daniel, Davis, Paul & Williams, Gary.
412 Overview of gene structure in *C. elegans**. in *WormBook - The Online Review of C.*
413 *elegans Biology* (2005).
- 414 25. Chen, W. *et al.* Endogenous U2·U5·U6 snRNA complexes in *S. pombe* are intron lariat
415 spliceosomes. *RNA* **20**, 308–320 (2014).
- 416 26. Cao, W., Tran, C., Archer, S. K., Gopal, S. & Pocock, R. Functional recovery of the germ
417 line following splicing collapse. *Cell Death Differ* **29**, 772–787 (2022).
- 418 27. Ceron, J. *et al.* Large-scale RNAi screens identify novel genes that interact with the *C.*
419 *elegans* retinoblastoma pathway as well as splicing-related components with synMuv B
420 activity. *BMC Dev Biol* **7**, 30 (2007).
- 421 28. Kamath, R. S. *et al.* Systematic functional analysis of the *Caenorhabditis elegans* genome
422 using RNAi. *Nature* **421**, 231–237 (2003).
- 423 29. Simmer, F. *et al.* Genome-wide RNAi of *C. elegans* using the hypersensitive rrf-3 strain
424 reveals novel gene functions. *PLoS Biol* **1**, E12 (2003).
- 425 30. Studer, M. K., Ivanović, L., Weber, M. E., Marti, S. & Jonas, S. Structural basis for
426 DEAH-helicase activation by G-patch proteins. *Proc Natl Acad Sci U S A* **117**, 7159–
427 7170 (2020).
- 428 31. Tauchert, M. J., Fourmann, J.-B., Lührmann, R. & Ficner, R. Structural insights into the
429 mechanism of the DEAH-box RNA helicase Prp43. *Elife* **6**, e21510 (2017).

- 430 32. Fourmann, J.-B. *et al.* Dissection of the factor requirements for spliceosome disassembly
431 and the elucidation of its dissociation products using a purified splicing system. *Genes*
432 *Dev* **27**, 413–428 (2013).
- 433 33. Small, E. C., Leggett, S. R., Winans, A. A. & Staley, J. P. The EF-G-like GTPase
434 Snu114p regulates spliceosome dynamics mediated by Brr2p, a DExD/H box ATPase.
435 *Mol Cell* **23**, 389–399 (2006).
- 436 34. Wollenhaupt, J. *et al.* Intrinsically Disordered Protein Ntr2 Modulates the Spliceosomal
437 RNA Helicase Brr2. *Biophys J* **114**, 788–799 (2018).
- 438 35. Riabov Bassat, D. *et al.* Structural basis of human U5 snRNP late biogenesis and
439 recycling. *Nat Struct Mol Biol* (2024) doi:10.1038/s41594-024-01243-4.
- 440 36. Schneider, S. *et al.* Structure of the human 20S U5 snRNP. *Nat Struct Mol Biol* (2024)
441 doi:10.1038/s41594-024-01250-5.
- 442 37. Fica, S. M., Oubridge, C., Wilkinson, M. E., Newman, A. J. & Nagai, K. A human
443 postcatalytic spliceosome structure reveals essential roles of metazoan factors for exon
444 ligation. *Science* **363**, 710–714 (2019).
- 445 38. Bertram, K. *et al.* Structural Insights into the Roles of Metazoan-Specific Splicing
446 Factors in the Human Step 1 Spliceosome. *Mol Cell* **80**, 127-139.e6 (2020).
- 447 39. Dybkov, O. *et al.* Regulation of 3' splice site selection after step 1 of splicing by
448 spliceosomal C* proteins. *Sci Adv* **9**, eadf1785 (2023).
- 449 40. Fica, S. M. *et al.* RNA catalyses nuclear pre-mRNA splicing. *Nature* **503**, 229–234
450 (2013).
- 451 41. Zhang, J. *et al.* DHX15 is involved in SUGP1-mediated RNA missplicing by mutant
452 SF3B1 in cancer. *Proc Natl Acad Sci U S A* **119**, e2216712119 (2022).

- 453 42. Han, B. *et al.* Human DBR1 modulates the recycling of snRNPs to affect alternative
454 RNA splicing and contributes to the suppression of cancer development. *Oncogene* **36**,
455 5382–5391 (2017).
- 456 43. Didychuk, A. L., Butcher, S. E. & Brow, D. A. The life of U6 small nuclear RNA, from
457 cradle to grave. *RNA* **24**, 437–460 (2018).
- 458 44. Trippe, R. *et al.* Identification, cloning, and functional analysis of the human U6 snRNA-
459 specific terminal uridylyl transferase. *RNA* **12**, 1494–1504 (2006).
- 460 45. Cartwright-Acar, C. H. *et al.* A forward genetic screen in *C. elegans* identifies conserved
461 residues of spliceosomal proteins PRP8 and SNRNP200/BRR2 with a role in maintaining
462 5' splice site identity. *Nucleic Acids Res* **50**, 11834–11857 (2022).
- 463 46. Noble, S. M. & Guthrie, C. Identification of novel genes required for yeast pre-mRNA
464 splicing by means of cold-sensitive mutations. *Genetics* **143**, 67–80 (1996).
- 465 47. Strauss, E. J. & Guthrie, C. A cold-sensitive mRNA splicing mutant is a member of the
466 RNA helicase gene family. *Genes Dev* **5**, 629–641 (1991).
- 467 48. Koodathingal, P. & Staley, J. P. Splicing fidelity: DEAD/H-box ATPases as molecular
468 clocks. *RNA Biol* **10**, 1073–1079 (2013).
- 469 49. Bohnsack, K. E., Ficner, R., Bohnsack, M. T. & Jonas, S. Regulation of DEAH-box RNA
470 helicases by G-patch proteins. *Biol Chem* **402**, 561–579 (2021).
- 471 50. Plaschka, C., Lin, P.-C. & Nagai, K. Structure of a pre-catalytic spliceosome. *Nature* **546**,
472 617–621 (2017).

473

474

475 **METHODS**

476 ***C. elegans* strain maintenance**

477 All *C. elegans* strains were maintained on nematode growth media (NGM) plates seeded with
478 OP50 bacteria at 20 °C as described previously⁵¹. Three strains were generated for this study,
479 all of them were made using CRISPR/Cas9 starting with the reference background strain, N2
480 (Supplementary Data Table 3). Detailed information on their construction is provided below.

481

482 ***C. elegans* genome engineering for endogenous tagging**

483 For endogenous tagging of *prp-19* with 3xFLAG tag, animals were subjected to Cas9-
484 mediated genome engineering via RNP-microinjection following the approach described in
485 ref.⁵². Briefly, the Alt-R CRISPR/Cas9 system (Integrated DNA Technologies) was
486 employed and young adult animals were injected with a mixture containing 300 mM KCl, 20
487 mM HEPES, 4 µg/µl recombinant Cas9 (from *S. aureus*, purified in-house), 500 ng/µl
488 TracerRNA, 100 ng/µl crRNAs targeting the endogenous *prp-19* locus (Supplementary Data
489 Table 4), and 150 ng/µl repair template encoding the 3xFLAG tag with around 150 bp
490 flanking homology arms. Identification of candidates with the correct insertions was done by
491 genotyping with the primers (Supplementary Data Table 4) and verified by Sanger
492 sequencing. This yielded allele *luc205* (strain MLC2610).

493

494 **Large-scale cultivation and preparation of whole-worm extract**

495 For large-scale cultivation of worms, Peptone + Streptomycin plates were seeded with 2 mL
496 concentrated HB101 bacteria and left to dry overnight at 37 °C. The following day, 100 000
497 synchronized L1 larvae of strain MLC2610 (*prp-19::3xFLAG*) were seeded onto each plate
498 and grown at 25°C until they had reached adulthood (between 2-2.5 days). Young adult
499 animals were harvested by washing the 15 cm plates repeatedly with ice-cold M9 buffer (22

500 mM KH₂PO₄, 42 mM Na₂HPO₄, 86 mM NaCl, 1mM MgSO₄), collected in 50 mL tubes, and
501 washed three times in ice-cold M9 to remove residual bacteria. For extract preparation, the
502 collected worm pellet was resuspended in one volume of Lysis buffer (50 mM HEPES pH
503 7.9, 1 mM MgCl₂, 100 mM KCl, 10% (w/v) glycerol, 0.05% NP-40, 0.5 mM DTT, cOmplete
504 EDTA-free protease inhibitor cocktail (Roche)), frozen drop-by-drop in liquid Nitrogen, and
505 subjected to cryo-milling using the Spex SamplePrep™ Freezer/Mill™ Dual Chamber
506 Cryogenic Grinder (2 cycles of 2 minutes, each at 15 CPS). The resulting extract was cleared
507 of debris by two consecutive rounds of centrifugation (10 min 15 000 g, 4 °C), frozen in
508 liquid Nitrogen, and stored at -70 °C until use. For the large-scale purification that yielded
509 the sample used for cryoEM, we harvested 50-60 full plates of worms that yielded ~18 mL of
510 cleared extract that was used for immunoprecipitation.

511

512 ***C. elegans* genome engineering to generate *syf-2* alleles**

513 Truncations of *syf-2* in *C. elegans* were generated by CRISPR/Cas9 mediated genome
514 engineering. Cas9 ribonucleoprotein complexes (RNPs) were injected into synchronized
515 young adult hermaphrodites. Cas9 RNPs were assembled by incubating a mix containing 50-
516 100 ng/μL Alt-R crRNAs targeting the *syf-2* gene (two guides were simultaneously injected
517 in equal concentration, Supplementary Data Table 5), 500 ng/μl TracerRNA and 4 μg/μl
518 recombinant Cas9 (from *S. aureus*, purified in-house) in buffer containing 300 mM KCl, 20
519 mM HEPES. As repair templates we used 100-200 ng/μL dsDNA repair template encoding
520 the delta helix 1 truncation with ~150 bp flanking homology arms for homology directed
521 repair (Supplementary Data Table 5); or 75 ng/μL ssDNA repair template with ~20 bp
522 homology arms for the deletion of helices 1 and 2 (Supplementary Data Table 5) plus 75
523 ng/μL pBSK dsDNA. In all cases we added 2.5 ng/μL dsDNA encoding *myo-2^{prom}:mCherry*
524 as a co-injection marker. The complete mix was incubated at 37 °C for 15 minutes, followed

525 by centrifugation for 10 minutes at maximum speed in a tabletop centrifuge at room
526 temperature to pellet any precipitates in the injection mix. Following microinjection, P0
527 worms were moved to individual plates and grown at 15 °C. At 48 hours post injection, F1
528 progeny were screened for *myo-2^{prom}:mCherry* expression. Plates containing at least three
529 *myo-2^{prom}:mCherry* positive larval animals were separated. The injected P0 animals from
530 these plates were transferred to fresh plates and were grown at 15 °C. All F1 progeny laid
531 after 48 hours were moved to single plates, allowed to lay progeny, and were genotyped by
532 PCR of the *syf-2* locus. The *syf-2:Δ^{anchor}* (delta helix 1) mutant was isolated and is maintained
533 in homozygosity (strain MLC2722). The *syf-2:Δ^{anchor+wedge}* (delta helices 1-2) was isolated in
534 heterozygosity; it is sterile in homozygosity. It is maintained in heterozygosity with the
535 *qC1(qIs26)* chromosome III balancer (strain MLC2727).

536

537 **Viability, temperature Sensitivity, and RNAi**

538 Animals were maintained at 20 °C on standard NGM plates seeded with *E. coli* OP50, unless
539 otherwise noted. RNAi plates used were NGM plates supplemented with 50 µg/ml
540 Carbenicillin and 1 mM IPTG. RNAi plates were seeded with *E. coli* HT115 expressing the
541 corresponding dsRNA which elicits an RNAi response against target gene mRNAs. Genes
542 targeted were as follows: *sde-2* (ORF-ID F53F4.14), *syf-2* (ORF-ID K04G7.11), *mog-7*
543 (homolog of human PAXBP1, ORF-ID F43G9.12), *cwf-19L1* (ORF-ID F17A9.2), and empty
544 vector control (pL4440-dest-RNAi Destination vector). RNAi strains were streaked onto
545 Carbenicillin plates from glycerol stocks of the RNAi libraries (ORFeome-RNAiv1.1). Single
546 colonies were grown in Carbenicillin + Tetracycline containing LB, and the identity of the
547 target gene was confirmed by Sanger sequencing. RNAi plates were freshly made for each
548 experiment by growing 5 mL of RNAi culture O.N., concentrating the bacteria to 1 mL, and
549 pipetting 150 µL of concentrated culture into the center of an RNAi plate. Worms plated for

550 RNAi experiments were synchronized as follows: L4 larvae were picked and allowed to
551 mature into young adults at 20 °C overnight. Young adults were sliced open with a razor
552 blade in a drop of M9. Early-stage embryos were collected by mouth pipetting and washed in
553 M9. Washed embryos were allowed to hatch into L1 in 500 µL of M9 buffer O.N. Individual
554 L1 worms were transferred to RNAi plates by mouth pipetting. Plates were grown at their
555 respective temperatures (15 °C or 20 °C) until the empty vector control plates had laid a
556 sufficient number of embryos (~150-250) to allow reliable counting of progeny, at which
557 point all mother worms were removed from plates at that given temperature (note that 15 °C
558 and 20 °C RNAi experiments were conducted at different time intervals because of the slower
559 developmental time of *C. elegans* at lower temperatures – thus absolute viable worm counts
560 are only comparable within a single temperature, not across different temperatures). F1
561 worms were grown at their respective temperatures until they reached L4 stage. L4 worms
562 were removed by hand and counted manually to assess total number of viable worms from a
563 given plate. Plates were screened for multiple days in a row to ensure that all L4s were
564 counted and removed to avoid allowing animals to reach adulthood and start laying eggs
565 again.

566

567 **Generation of the human GFP-TFIP11 K562 cell line**

568 For endogenous purification of the human ILS, lentiviral particles carrying the GFP-3C-
569 TFIP11 constructs were generated in Lenti-X 293T cells (Takara) via polyethylenimine
570 transfection (Polysciences) of the viral carrier plasmid and helper plasmids pCMVR8.74
571 (Addgene #22036) and pCMV-VSV-G (Addgene #8454), according to standard procedures.
572 K562 (DSMZ) cells were infected at limiting dilutions and GFP-positive cells were isolated
573 using a BD FACSAria III cell sorter (BD Biosciences). Viral integration was confirmed by

574 immunoblotting for TFIP11 and GFP. Lenti-X and K562 cells tested negative for
575 mycoplasma.

576

577 **Preparation of human nuclear extract**

578 To prepare nuclear extract (NE), 30 L of human K562 cells overexpressing GFP-3C-TFIP11
579 were grown to a density of 1.5×10^6 cells mL^{-1} at 37 °C, 5% CO₂, stirred at 70 rpm. The NE
580 was prepared as previously described⁵³ and dialysed against buffer F (20 mM HEPES pH 7.9,
581 100 mM KCl, 20% (w/v) glycerol, 0.2 mM EDTA, 2 mM DTT).

582

583 **Cryo-EM sample preparation**

584 ***C. elegans* spliceosome anti-FLAG affinity purification and grid preparation**

585 Whole-worm extract prepared from *prp-19::3xFLAG* animals and incubated for 2 h at 4 °C
586 with anti-FLAG M2 resin previously equilibrated with Equilibration buffer C
587 (20 mM HEPES pH 7.9, 50 mM KCl, 2 mM MgCl₂, 0.05% NP-40, 8% (w/v) glycerol,
588 1 mM TCEP). After washing, samples were eluted by incubation with 300 ng/μl FLAG
589 peptide dissolved in TBS (Tris-buffered saline) for 2 hours at 4 °C. The eluate was loaded
590 onto a GraFix⁵⁴ 15–30% (w/v) sucrose density gradient containing 0.05% glutaraldehyde in
591 buffer D (25 mM HEPES pH 7.9, 50 mM KCl, 15% sucrose, 1 mM TCEP) and spun at
592 22000 rpm for 16 h in a SW60 Ti rotor (Beckman coulter). The sedimentation coefficients
593 were simulated using the CowSuite software (<https://www.cow-em.de>). Peak spliceosome
594 fractions were quenched for 15 minutes using a final concentration of 50 mM lysine, pooled,
595 concentrated in a 0.5 mL 30 kDa MWCO Amicon Ultra concentrator (Sigma) and the buffer
596 exchanged to buffer E (20 mM HEPES pH 7.9, 100 mM KCl, 2 mM MgCl₂, 1 mM TCEP)
597 and immediately used for EM grid preparation. Briefly, 4 μL of concentrated and crosslinked
598 samples was applied to glow discharged R2/1 200 holey carbon grids (Quantifoil) coated

599 with a home-made 2 nm continuous carbon layer. Grids were blotted at 4 °C and 80%
600 humidity and plunged into liquid ethane using a Leica EM GP2.

601

602 **Purification and grid preparation of endogenous human ILS complexes**

603 30 ml of GFP-3C-TFIP11 K562 NE were incubated with GFP-Trap Agarose resin
604 (Chromotek) pre-equilibrated with binding buffer G (20 mM HEPES pH 7.9, 100 mM KCl,
605 2 mM MgCl₂, 0.05% (w/v) NP-40, 8% (v/v) Glycerol, 1 mM TCEP, cOmplete EDTA-free
606 protease inhibitor cocktail (Roche)) for two hours at 4 °C under constant rotation. After five
607 washes with six times the bead volume of binding buffer G, the human ILS complexes were
608 eluted by cleavage using 3C PreScission Protease diluted in elution buffer H (20 mM HEPES
609 pH 7.9, 100 mM KCl, 2 mM MgCl₂, 8% (v/v) glycerol, 0.04 µg/µl protease, 1 mM TCEP) for
610 1.5 hours.

611 The eluate was loaded onto a GraFix⁵⁴ 15–30% w/v sucrose density gradient
612 containing 0.05% glutaraldehyde in buffer I (20 mM HEPES pH 7.9, 50 mM KCl, 1 mM
613 TCEP) and centrifuged at 21000 rpm for 16 h in 4 °C in a SW60 Ti rotor (Beckman coulter).
614 The sedimentation coefficients were simulated using the CowSuite software
615 (<https://www.cow-em.de>). Fractions containing the human ILS were quenched for 15 minutes
616 using a final concentration of 50 mM lysine, pooled, concentrated in a 0.5 mL 100kDa
617 MWCO Amicon Ultra concentrator (Sigma) and exchanged to buffer J (20 mM HEPES pH
618 7.9, 100 mM KCl, 2 mM MgCl₂, 1 mM TCEP) and immediately used for EM grid
619 preparation. Briefly, concentrated, and crosslinked human ILS complex were incubated with
620 a home-made 2 nm continuous carbon layer for 20 minutes and picked up with glow
621 discharged R2/1 200 holey carbon grids (Quantifoil). Grids were blotted at 4 °C and 68%
622 humidity and plunged into liquid ethane using a Leica EM GP2.

623

624 **Cryo-EM data acquisition**

625 ***C. elegans* spliceosomes.**

626 We collected two datasets of the *C. elegans* spliceosome sample, encompassing 17,600 and
627 25,358 micrographs, respectively. Both datasets were imaged on the same microscope (Titan
628 Krios G3 at IST Austria, equipped with a Gatan K3 direct detector). For both datasets, we
629 collected movies with a total dose of $60 \text{ e}^-/\text{\AA}^2$ fractionated over 40 frames at a pixel size of
630 1.06 \AA using ThermoFisher's EPU software. The target defocus range was set to $-0.9 \text{ }\mu\text{m}$ to -
631 $2.1 \text{ }\mu\text{m}$ for dataset 1 and -0.75 to $-1.9 \text{ }\mu\text{m}$ for dataset 2. The electron filter was set to a filter
632 width of 20 eV, and we used a $50 \text{ }\mu\text{m}$ C2 aperture and no objective aperture.

633

634 **Human TFIP11-containing spliceosomes.**

635 We collected two datasets of the human TFIP11-spliceosome sample, encompassing 12,344
636 and 27,699 micrographs, respectively. Both datasets were imaged on the same microscope
637 (Titan Krios G4 at Vienna BioCenter, equipped with a Falcon 4i direct detector). The datasets
638 had a defocus range of $-0.75 \text{ }\mu\text{m}$ to $2 \text{ }\mu\text{m}$. For both datasets, we collected movies with a total
639 dose of $50 \text{ e}^-/\text{\AA}^2$ in EER format at a pixel size of 0.945 \AA using ThermoFisher's EPU
640 software. The electron filter was set to a filter width of 10 eV, and we used a $50 \text{ }\mu\text{m}$ C2
641 aperture and no objective aperture.

642

643 **Cryo-EM data processing**

644 ***C. elegans* spliceosomes**

645 **Pre-processing.** Data was pre-processed with cryoSPARC live⁵⁵. Movies were gain- and
646 motion corrected using the 'Patch Motion' program and the defocus was estimated using
647 'Patch CTF'. Particles were picked using the 'Blob picker', with the minimal and maximal
648 particle diameters set to 320 \AA and 370 \AA , respectively. This yielded 1,782,017 and

649 2,358,078 particle coordinates from dataset 1 and 2, respectively. Particles were initially
650 extracted with a box size of 583 Å and binned to a pixel size of 2.45 Å for dataset 1 (box size
651 of 256 px). For dataset 2, particles were binned to a pixel size of 1.3 Å/px and a box size of
652 448 pixels. Datasets 1 and 2 were subjected to initial cleaning and classification
653 independently, and merged only after high-quality particle sets were identified.

654 **Initial particle cleaning.** To generate an initial model, we selected 173,220 particles from
655 dataset 1 and cleaned the set using 2D classification, yielding 48,465 particles. From these,
656 we calculated three initial volumes using the cryoSPARC ‘Ab-initio reconstruction’
657 algorithm, yielding one high-quality spliceosome class containing 76% of the input particles,
658 and two ‘junk’ classes. These volumes were then used as reference volumes to classify all
659 extracted particles from dataset 1 using ‘Heterogenous refinement’, yielding a dominant class
660 of 924,617 dataset 1 ILS particles. Attempts to identify other spliceosome states in the data
661 using reference volumes of human Bact, C, C*, or P complex maps as additional references
662 were unsuccessful. Dataset 1 ILS particles were re-extracted at a pixel size of 1.3 Å/px and
663 refined using ‘Homogenous Refinement’, enabling particle scale optimization. This yielded a
664 map at a global resolution of 3.1 Å. CTF refinement using global or local CTF refinement
665 strategies did not improve resolution and was not further pursued.

666 For dataset 2, we split the 2,358,078 auto-picked particles into four batches and
667 subjected each batch to ‘Heterogenous refinement’ using the same reference volumes as for
668 dataset 1. ILS particles from each batch were combined and subjected to a second round of
669 ‘Heterogenous refinement’. The best class contained 1,339,014 particles and was refined
670 using ‘Homogenous Refinement’, enabling particle scale optimization. This yielded a map at
671 a global resolution of 2.84 Å (dataset 2 ILS particles).

672 We next classified particles of each dataset into ILS’ and ILS’’ states, using global 3D
673 classification without image alignment. We set the target resolution to 15 Å and the initial

674 low-pass filter to 40 Å and used 50 classes (dataset 1) or 40 classes (dataset 2). We chose a
675 slightly lower number of classes for dataset 2 because we were facing run-time issues when
676 attempting to use fifty classes using the larger particle number in dataset 2.

677 Classes from 3D classification runs were inspected and grouped into either ILS' (87%
678 and 86% for datasets 1 and 2, respectively) or ILS'' (13% / 14%), based on the presence or
679 absence of DHX15 density. From ILS' classes we excluded a large number of particles
680 (360,899 / 433,898 for dataset 1 and 2, respectively) that showed no or poor density for the
681 U2 snRNP, presumably due to mobility. These particles were otherwise indistinguishable
682 from ILS' and yielded high-resolution reconstructions (~3 Å).

683 We also separately pooled all particles that showed BRR2 density (155,690 / 299,666
684 particles). We exclusively observed BRR2 density in ILS' classes.

685 To further improve homogeneity within the ILS' and ILS'' populations, we subjected ILS'
686 particles with strong U2 density and ILS'' classes to another round of global 3D
687 classification, discarding classes with residual DHX15 density from the ILS' particle set and
688 discarding classes with weaker DHX15 density from the ILS'' particle set. These stringently
689 classified particle sets contained 321,346 and 558,116 particles for the ILS' state, 155,690
690 and 299,66 for the ILS' subset with BRR2 density, and 105,996 and 151,874 particles for the
691 ILS'' state. At this stage, particles from both datasets were combined, and we calculated
692 consensus refinement using 'Non uniform refinement', yielding an ILS' consensus map at a
693 global resolution of 2.9 Å (Map 1) and an ILS'' consensus map at a global resolution of
694 3.06 Å (Map 16).

695 **Local refinements for peripheral regions.** While the core of the consensus refinements
696 reached a resolution of 2.6 Å, peripheral regions remained less well-defined due to molecular
697 motion. To overcome this, we optimized local masks and local refinement parameters for
698 different regions of the ILS that appeared to behave approximately as rigid bodies, guided by

699 ‘cryoFlex’ analysis as implemented in CryoSPARC. While we initially subjected ILS’ and
700 ILS’’ particle sets to independent local refinements, we noticed six regions that appeared
701 indistinguishable at high resolution in the ILS’ and ILS’’. For these regions, we combined
702 ILS’ and ILS’’ particles for local refinements. Please refer to Supplementary Table 1 for a
703 detailed list of map boundaries. Regions common between ILS’ and ILS’’ included i) the
704 U5 snRNA 5’ end and the associated sm ring (map 2), ii) the intron binding complex (AQR
705 and associated proteins) (map 3), the NTC core (map 4) and the PRP19 WD-40 domain
706 (map 5), and the most peripheral region of the NTR (containing the SYF1 and SYF3 C-
707 termini and associated proteins) (map 6), and the U2 snRNP 5’ end and associated sm ring
708 (map 7).

709 To generate map 2, we performed an initial local refinement with a wider mask
710 focused on the U5 snRNA sm ring using gaussian priors and setting the search ranges for
711 translations and rotation to 21 pixels and 45°, followed by a second local refinement with a
712 tighter mask and limiting the search ranges to 1 pixels /45°, yielding a map at a nominal
713 resolution of 3.29 Å.

714 To generate map 3, we performed an initial local refinement with a mask around AQR
715 and associated proteins using gaussian priors and setting the search ranges for translations
716 and rotation to 21 pixels and 45°, followed by a second local refinement with a tighter mask
717 and limiting the search ranges to 1 pixels /45°, yielding a map at a nominal resolution of
718 2.99 Å.

719 To generate map 4, we performed an initial local refinement with a mask around the
720 NTC core, limiting the search ranges to 5 pixels /12°, yielding a map at 3.14 Å.

721 To generate map 5 and better resolve the PRP19 WD 40 (which was visible in only
722 one of the four PRP19 subunits), we subjected Map 4 particles to 3D classification using four
723 classes and a 20Å low-pass filtered reference volume, yielding two classes with 599,066

724 particles with improved WD40 density. Local refinement with gaussian priors and search
725 ranges for translations and rotation of 21 pixels and 45° yielded a 5.57 Å nominal resolution
726 density.

727 To generate map 6 and better resolve the NTR periphery, we first used local
728 refinement with a mask around the SYF1 and SYF3 C-termini and associated proteins,
729 limiting the search ranges to 5 pixels /5°, followed by 3D classification using four classes and
730 a 20 Å low-pass filtered reference volume. Class 1 contained 345,902 particles and yielded a
731 3.31 Å nominal resolution map after reconstruction.

732 To generate map 7, we performed masked 3D classification of combined ILS particles
733 using 20 classes and a 20 Å low-pass filtered reference volume. This revealed substantial
734 mobility, and we selected class 5 containing 60,582 particles that yielded a 6.35 Å nominal
735 resolution map upon reconstruction.

736 **Local refinements in the ILS'.**

737 Next, we performed local refinements on ILS' particles to generate focused refinement in
738 regions that differ in the ILS' and ILS''.

739 We first generated map 8 and 9, that together with map 6 completed the NTR lobe.

740 To generate map 8, we subjected ILS' particles to local refinement using a mask around the
741 SYF1 central region, limiting the search ranges to 3 pixels /3°, yielding a 3.09 Å nominal
742 resolution density.

743 To generate map 9, we subjected map 8 particles to local refinement using a mask
744 focused on a central region of SYF1 adjacent to the region resolved in map 8 and limiting the
745 search ranges to 3 pixels /3°, yielding a 3.69 Å nominal resolution density. While a map
746 calculated with the same mask from ILS'' particles showed clear and high-resolution density
747 for SDE2 bound to SYF1, SDE2 was entirely absent, indicating that SDE2 which is recruited

748 to the spliceosome during catalytic activation, is prone to dissociation in the absence of
749 catalytic-stage splicing proteins or disassembly factors.

750 We next calculated local refinement maps to better resolve disassembly factors and
751 their binding sites, starting at the base of the U5 snRNP where the TFIP11 CTD binds and
752 moving towards the periphery, where the highly mobile BRR2 binds, yielding maps 10 to 14.
753 To generate map 10, we subjected ILS' particles to local refinement using a mask around the
754 TFIP11-CTD and U6S1 domains III+IV, limiting the search ranges to 3 pixels /3°, yielding a
755 2.73 Å nominal resolution density.

756 To generate map 11, we subjected ILS' particles to local refinement using a mask
757 around the TFIP11 helical region, limiting the search ranges to 3 pixels /3°, yielding a 3.03 Å
758 nominal resolution density.

759 To generate maps 12 to 14, we used the ILS' subset containing BRR2 density, as
760 BRR2 binding appeared to conformationally lock the mobile PRP8 RNaseH and as and
761 PAXBP1. We initially performed focused 3D classification using a mask on BRR2 and a
762 20 Å low-pass filtered reference map and 10 classes, from which we selected a subset of
763 148,075 particles with improved density. We next performed focused refinement with a
764 search range of 5 px / 5° to pre-aligned particles, followed by a second refinement using
765 gaussian priors and a search range of 21 px and 30°, yielding a 4.24 Å density of BRR2 and
766 the PRP8 JAB domain (map 12). Starting from these particles, we performed local refinement
767 using a wider mask encompassing the PRP8-RNaseH and JAB domains, PAXBP1 and
768 BRR2, using a search range of 10 px /20° to yield map 12.

769 To generate map 13, map 14 particles were pre-aligned using local refinement with a
770 similar mask as used for map 12, but larger, and a search range of 10 px/20°. A second local
771 refinement with a mask around the PAXBP1-CTD yielded map 13 at a nominal resolution of
772 6.4 Å.

773 Finally, we noticed that density around the U2/U6 helix was poor in the ILS',
774 precluding conclusions on whether the SYF2 'wedge helix' is bound on top of the U2-U6
775 helix as a consequence of DHX15 binding, or prior to it. To address this, we further classified
776 ILS' particles, using masked 3D classification around U2/U6 Helix 10 with 10 classes and a
777 20 Å low-pass filtered reference volume. We selected class 7, which showed clearly
778 improved U2/U6 Helix II density and contained 127,957 particles. We also noticed very
779 weak and blurry density in the region where DHX15 binds the ILS''. To exclude that
780 stabilized Helix II density was only observed because of a small proportion of ILS'' particle
781 in the set, we further classified the 127,957 particle set using masked classification with a
782 DHX15 density. From this we removed all classes that showed any residual DHX15 density,
783 yielding a set of 78,633 ILS'' particles with improved U2/U6 Helix II density at a nominal
784 resolution of 3.61 Å (map 15).

785 **Local refinements in the ILS''.**

786 We followed a similar refinement strategy for the ILS'', starting from a 'Non-uniform
787 refinement' of ILS'' particles that yielded a density at a nominal resolution of 3.06 Å
788 (map 16). First, we pre-aligned particles on the NTR-IBC-DHX15 region using a local
789 refinement with a wide mask encompassing these regions, using gaussian priors and a search
790 range of 21 pixels 45° (NTR-IBC pre-aligned particles). To improve density for DHX15 and
791 its interactors, performed a local refinement with a mask around DHX (search range 3 px/ 3°)
792 to generate pre-aligned particles. These particles were imported into RELION 5 (ref. ^{56,57})
793 and subjected to 3D classification using 10 classes without image alignment and using 'blush
794 regularization'. While all classes showed clear density for DHX15, a major class of 31%
795 (84,004 particles) showed improved density for the DHX15 RecA2 lobe and the C19L1 N-
796 terminus. These particles were re-imported into cryoSPARC and subjected to local
797 refinement using a mask on DHX15 (search range 3 px/ 3°) to generate a map 25 at a

798 nominal resolution at 3.94 Å/px or a local refinement on the DHX15-RecA2 lobe and the
799 C19L1 NTD (search range 1 px/ 1°) to generate map 27 at a nominal resolution of 3.88 Å.

800 To improve density for the DHX15-NTR interaction site, we used DHX15 pre-aligned
801 particles to perform a local refinement with a mask focused on SYF1 central region and the
802 SYF2 and SDE2 ‘anchor helices’ (search range 1 px/1°) to generate map 23 at a nominal
803 resolution of 3.62Å.

804 In the ILS’’, we noticed additional weak density near the IBC. To better resolve this,
805 we used masked 3D classification of NTR-IBC pre-aligned particles using 10 classes and a
806 6 Å low-pass filtered reference volume and a mask focused on the additional density, and
807 selected the class with the strongest density (57,951 particles), which we refined using a local
808 refinement (search range 3px/ 3°) to generate map 17, showing the ISY1 N-terminus at a
809 nominal resolution of 5.7 Å. An AlphaFold2 Multimer^{58,59} prediction supported the density
810 assignment. We also performed the same processing strategy with ILS’ particles, however
811 ISY N-terminus could not be resolved, indicating it might be stabilized in the ILS’.

812 We next sought to better resolve TFIP11 and PAXBP1 and its interactors. Starting from
813 Mapp 16 particles, we performed local refinement (search range 3 px/3°) with wide mask
814 focused on the TFIP11 helical domain, followed by a second local refinement with a tighter
815 mask (search range 1px/1°) to generate map 19 at a nominal resolution od 3.14 Å.

816 Map 19 particles were further refined using a local refinement with a mask focused on the
817 TFIP11-CTD and U5S1 domains III and IV (search range 3px/3°) to generate map 18 at a
818 nominal resolution od 2.82 Å.

819 To generate map 20 and focus on the PRP8 RNaseH domain and PAXBP1, we
820 performed focused refinement on map 19 particles with a mask focused on the PRP8-RNase
821 H and (search range 3px/3°) to generate map 20 at a nominal resolution of 3.13 Å.

822 To generate map 21 and focus on PAXBP1-CTD, we performed focused refinement on
823 map 20 particles with a mask focused on the PAXBP1-CTD and (search range $3\text{px}/3^\circ$) to
824 generate map 21 at a nominal resolution of 6.33 \AA .

825 To better resolve U2-U6 Helix II and its interactors, we first aligned map 16 particles
826 using local refinement (search range $3\text{px}/3^\circ$) with a mask focused on U2-U6 Helix II to
827 generate map 22.

828 To further improve density for U2/U6 Helix II and the SYF2 ‘wedge helix’ and SDE2
829 ‘anchor helix’, we classified map 22 particles using the same mask as for map 22 and
830 selected a subset of 69,968 particles that improved connectivity for the U6 snRNA between
831 Helix II and DHX15 and SDE2 density. The selected particles were refined with a tight mask
832 around U2-U6 (search range $1\text{px}/1^\circ$), yielding map 24 at a nominal resolution of 3.92 \AA and
833 clearly revealing the U6 snRNA as the target for DHX15.

834 Finally, to improve the mobile C19L1 CWFJ density and generate map 26, we pre-
835 aligned map 16 particles using a local refinement with a mask focused on C19L2 and the
836 C19L1 CWFJ (search range $1/1$). Focused 3D classification using 10 classes and a 15 \AA low-
837 pass filtered reference volume revealed a class of 26,170 particles with improved C19L1
838 CWFJ density, which were reconstructed to generate a nominal resolution map of 3.19 \AA .
839 Note that this resolution value reflects overall resolution of the reconstruction, and that local
840 resolution of the C19L1 CWFJ is lower.

841

842 **Model building.**

843 To generate an initial model, we downloaded AlphaFold2 models for all subunits present in
844 the human ILS2, except for subunits CDC5L, SNW1, whose highly extended conformation
845 was better captured by Phyre2 homology models (ref.⁶⁰). We further generated an
846 AlphaFold2 Multimer model for a complex containing four copies of PRP19, SPF27, and

847 CDC5L, which fitted Map 4 almost perfectly, and AlphaFold2 Multimer models for the
848 TFIP11–PAXBP1, TFIP11-G-patch–DHX15, and DHX15C–C19L1-CTD, and C19L2–
849 C19L1. These starting models were aligned to the human ILS2 (PDB 61D1), truncated as
850 appropriate and manually adjusted into their corresponding local refinement maps in
851 ISOLDE⁶¹ and ChimeraX⁶², and confidence-weighted reference restraints and secondary
852 structure restraints from AlphaFold2 models where side-chains were not unambiguously
853 resolved, as implemented in ISOLDE.

854 To identify the SYF1 ‘tether’, we performed an AlphaFold2 Multimer screen of
855 DHX15-C19L1(1-277) against the top 200 most abundant in *Ce* spliceosome sample
856 identified by MS, which identified the SYF1 C-terminus as protein element interacting with
857 both DHX15 and C19L1. The AF model was then manually adjusted into the density.

858 To model snRNAs, used the human ILS structure (PDB 61D1) as a template and
859 adjusted the sequence. Compared to their human counterparts, the U6 snRNA and U5
860 snRNAs are highly conserved (94% and 93 % sequence identity, respectively), and contain
861 only minor differences. The *C. elegans* U5 snRNA contains a three-nucleotide insertion after
862 position 20, which were resolved in the cryo-EM map and could be unambiguously modelled,
863 and a three-nucleotide deletion in the 3’ stem loop. The U6 snRNA contains two deletions
864 (three nucleotides after position +3 and one nucleotide after position +10, using the human
865 numbering), making the 5’ stem loop slightly shorter. The U2 snRNA is slightly more
866 divergent (61% sequence identity), but all nucleotides that pair with U6 or the intron RNA
867 branch site are perfectly conserved. Due to a divergent sequence of the U2 snRNA 3’ stem
868 loop, we generated a model using RNAcomposer⁶³ and fitted it into the cryo-EM density.
869 Due to limited resolution, we truncated bases and refrained from assigning a sequence
870 register for the U2 snRNA 5’ stem loop.

871 After adjusting the model in ISOLDE, the model geometry was refined in phenix
872 (1.20.1-4487) against the composite map. We generated base-pair and stacking restraints in
873 phenix to stabilize nucleic acid geometry, and further used the input model to generate
874 reference model restraints. The nonbonded weight parameter was set to 2000 and rotamers
875 were fitted for sidechains with poor density that were also outliers. These settings gave
876 yielded excellent refinement statistics and real-space correlation values (Extended Data Table
877 2a-c).

878 **Composite map generation.**

879 Composite maps were generated using UCSF ChimeraX⁶². To optimally preserve high-
880 resolution information from each focused refinement, we first fitted each focused refinement
881 into the locally filtered consensus map of the corresponding state. We then manually
882 segmented our atomic model and visually identified the map that shows the best quality for
883 each model region (Supplementary Table 1). Maps were then zoned around these atoms with
884 a 10 Å distance cutoff. Finally, we combined all zoned focused maps and scaled them
885 relative to each other using the ChimeraX ‘volume max’ command.

886

887 **Human spliceosome cryo-EM data analysis.**

888 **Pre-processing.** Both human ILS datasets were pre-processed using cryoSPARC³⁷ live with
889 default settings for gain and motion correction using ‘Patch Motion’ and CTF estimation
890 using ‘Patch CTF’. Particle picking was done in WARP v1.09 using a custom BoxNet2Mask
891 neural network. For the first dataset 303,126 particles were picked from 12,344 micrographs.
892 For the second dataset 1,388,112 particles were picked from 27,699 micrographs. The
893 particles were extracted using a box size of 550 pixels and binned to 2.65 Å/px for initial
894 classification.

895 **Processing.** The cryo-EM data was processed as indicated in Supplementary Data Fig. 4.
896 Briefly, the two acquired datasets were individually classified using heterogenous
897 refinements. As reference volumes we chose two ab initio models derived after initial 2D
898 classification obtained from the first 100,000 particles of which 32,295 particles were used to
899 generate two ab-initio models. As second reference we supplied low pass filtered maps of the
900 *C.elegans* spliceosome homogenous refinement before classification and the human ILS2
901 density (emdb:9647).

902 After three rounds of heterogenous refinements without a mask we combined the
903 highest quality particles from each dataset and applied a non-uniform refinement using a
904 dynamic mask, yielding a map at resolution 3.53 Å with 116 770 particles. A second
905 application of three consecutive rounds of heterogenous refinement using the low pass
906 filtered published ILS structures (PDB: 6ID1 and 6ID0) as reference volumes yielded a
907 second particle set (particle set II), which after non-uniform refinement yielded Map 1, at
908 3.38 Å resolution with 87,951 particles.

909 **Local refinements.** We subsequently applied several masks for focused classifications and
910 refinements to improve local densities. Particle set II was used for focused refinements with
911 masks around the core, PRP8, the U5snRNP and the IBC/AQR to generate Maps 2,3,7,11. To
912 yield better results for the sub-regions of PRP8, the U5 Sm ring and U5 associated proteins
913 two focused refinements with a wider and then smaller mask around the region of interest
914 were applied to generate Maps 4,6,8. For the peripheral more flexible region of the NTC, we
915 additionally applied 3D classification of particle set I without a mask and chose only the
916 classes showing best densities for the region of interest. Following a focused refinement of
917 the remaining 52,490 particles with a mask around the NTC yielded Map 9 at 6.08 Å
918 resolution.

919 Similarly, we used the particles of Map 10 (IBC) as input for 3D classification with a
920 mask around DHX15. Subsequent homogenous reconstruction of the best class with 9,321
921 particles yielded Map 13 at 8.06 Å resolution.

922 3D Variability analysis (3DVAR) with a mask around the heterodimer of
923 TFIP11/PAXBP1 was applied to particle set I. The 23,887 particles of the best class (showing
924 density for the disassembly factors) were further classified using 3D classification, yielding
925 14,635 particles. Focused refinement with a mask encompassing TFIP11/ PAXBP1 yields
926 Map 12 at 4.78 Å. The particles were further used for focused refinement of TFIP11 only.
927 Therefore, we applied 3D classification with a second smaller mask focused on TFIP11 on
928 particle set II and combined these particles with the particles of Map 12 for focused
929 refinement of TFIP11 only, yielding Map 14 from 26,103 particles at 5.79 Å resolution. For
930 more details see Extended Data Table 2, Supplementary Table 1, and Supplementary Fig. 5.

931 **Model building.** To prepare the integrative model of the human ILS, we started out with the
932 already published model of the human ILS (PDB:6ID1) (see Supplementary Fig. 6). From
933 this model we deleted the U2 snRNP, part of the U2 snRNA (residues 54-184), the intron
934 (residues 25-30, 118-135), part of the U5 snRNA (residues 7,8,70-84) and SYF2 (residues
935 114-126, 196-243), which were not resolved in our densities. Additionally, we deleted chains
936 corresponding to SYF1, SYF3, SNU114, AQR, SPF27, PRP19, CDC5L (residues 251-270),
937 CWC15, the U5 Sm ring and rebuilt them using AlphaFold2 Multimer prediction. In addition,
938 we built extensions for CWC15 (residues 36-105), SNW1 (residues 1-46), C19L2 (residues
939 54-75), SNRPB (residues 99-117), PLRG1 (residues 10-52) and SYF2 (residues 20-116).

940 Finally, we added chains for CCDC12, ESS2, PAXBP1, TFIP11, ISY1, DHX15, SDE2,
941 C19L1. All newly modelled chains are based on AlphaFold2 Multimer predictions. To see
942 which subunits are best resolved in which map, see Supplementary Table 1.

943 For the NTC we predicted combinations of SPF27, PRP19, CDC5L, PLRG1 and rigid
944 body fitted them into the density of Map 6, followed by adjustments using ChimeraX and
945 ISOLDE. For the IBC we predicted combinations of SYF1, SYF3, CCDC12, ISY1, AQR,
946 SYF2, SDE2. We used ChimeraX and ISOLDE to fit SYF1 and SYF3 into Map 10. We used
947 the low pass filtered consensus refinement Map 1 for rigid body fitting of the predicted
948 models of CCDC12, SYF2, ISY1, AQR and SDE2.

949 TFIP11 (residues 726-837) was predicted with SNU114 and fitted into Map 14. The
950 TFIP11/PAXBP1 dimer was predicted using AF2 and fitted into Map 12. PAXBP1 was
951 trimmed according to the homologues domains which were resolved in the *C. elegans* ILS''
952 model. The helicase DHX15 was predicted with the G-patch domain of TFIP11 (residues
953 146-211) and placed in Map 13; its orientation was compared to the *C. elegans* ILS'' model.
954 In the integrative model of the human ILS, we modelled the RNA network based on the PDB
955 6ID1 and trimmed the U2 snRNA, the intron and the U5 snRNA. However, we observe that
956 the RNA in the human structure is very mobile. The final coordinate model was real space
957 refined using Phenix as described for the *C. elegans* ILS models (Extended Data Table 2).
958 After refinement, we trimmed lowly resolved protein chains to a minimal backbone model for
959 subunits SYF1, SYF3, C19L2 (54-75), CWC15, CCDC12, CDC5L, SYF2, ESS2, AQR,
960 PAXBP1, PRP19, ESS2, DHX15, TFIP11, ISY1, C19L1, SPF27, SNW1.

961 **Modelling of a revised human P complex structure**

962 To generate the revised model of the human P complex, we used the published P complex
963 model (PDB 6QDV), mass spectrometry data and cryo-EM densities from ref.³⁷ (EMD-4530,
964 EMD-4525, EMD-4532, EMD-4526, EMD-4539). We first compared the deposited EM
965 densities with PDB 6QDV to identify unmodelled densities, and then used AlphaFold2
966 models of PPWD1, TLS1, ESS2, CDC5L, Cxorf56, FAM50A, NKAP, and NOSIP to obtain
967 candidate fits to these densities. Subsequently we used AlphaFold2 Multimer to predict

968 combinations of the new factors with spliceosome proteins adjacent to unmodelled densities
969 (PRP8, BRR2, SNW1, DHX8, CACTIN). AlphaFold2 Multimer models were aligned to
970 PDB 6QDV and adjusted into the maps using rigid body fitting and ISOLDE in UCSF
971 ChimeraX (Supplementary Data Fig. 1). In addition, we used models of SRRM2 and PPIE
972 which have been assigned in previous structures of the human B^{act}, C, and C* complexes^{22,37-}
973 ³⁹, and extended them using respective AlphaFold2 models. In addition, we used the
974 NTR/NTC AF2 models, which we prepared for the *Hs* ILS'' to improve models of PRP19,
975 SPF27, CDC5L, SYF3, SYF2, SDE2, ISY1, CCDC12, SNW1). The final coordinate model
976 was refined in real space using Phenix as described for the *C. elegans* ILS models, with the
977 exception that no composite map was used. Instead, refinement was performed with the
978 consensus map EMD-4525. After refinement, we removed the sidechains located in lower
979 resolution areas of the map. These include sidechains for: ESS2, FAM50A, Cxorf56, PRP8
980 (residues 2067-2335), ISY1, SYF1, SYF3, CDC5L (residues 518-801), SYF2, NOSIP,
981 SRRM2 (residues 56-99), DHX8, SDE2, PRP19, NKAP (residues 376-415), EIF4A3,
982 MAGOH, RBM8A, CWC22 (residues: 149-406), AQR, SPF27, PPWD1, TSSC4, CCDC12,
983 PPIE. The revised P complex coordinate model extends the previously deposited structure by
984 over 2,100 residues.

985

986 **Protein purification**

987 **DHX15**

988 Recombinant *Ce* DHX15 was expressed in insect cells using a pGB10 plasmid containing
989 10x-His-FLAG-3C-ceDXH15. The plasmid was electroporated into DH10EMBacY cells to
990 generate bacmids⁶⁴ that were then transfected into *Spodoptera frugiperda* Sf9 cells to
991 generate a V0 virus. The V0 virus was further amplified in Sf9 cells to yield V1 virus. Then,
992 we expressed this new construct in Hi5 insect cells using baculovirus. Insect cell pellets were

993 resuspended in buffer A (50 mM HEPES pH 7.0, 400 mM NaCl, 20 mM Imidazol, 10% (v/v)
994 glycerol, 2 mM MgCl₂, 2 mM beta-mercaptoethanol), and lysed by sonication. The lysate
995 was cleared by centrifugation (first for 30 min at 18,500 rpm, then for 1 h at 40,000 rpm in a
996 Ti45 rotor). The supernatant was filtered through 0.45 µm filters and applied to a HisTrap HP
997 5 ml column, previously equilibrated in buffer A. The column was washed with buffer A and
998 washed with 5 % Buffer B (50 mM HEPES pH 7.9, 400 mM NaCl, 10% Glycerol, 2 mM
999 MgCl₂, 500 mM imidazole, 2 mM beta-mercaptoethanol) and eluted with 60% buffer B.
1000 Peak fractions were pooled, diluted to 100 mM NaCl, and further purified via a Heparin
1001 chromatography column, using buffer C (50 mM HEPES pH 7.9, 10% Glycerol, 2 mM
1002 MgCl₂, 2 mM DTT) and buffer D (50 mM HEPES pH 7.9, 1 M NaCl, 10% Glycerol,
1003 2 mM MgCl₂, 2 mM DTT). After loading, the column was washed with 5 column volumes
1004 (CV) of 5% buffer D and then eluted with a linear gradient from 5% to 100% buffer D over
1005 20 CV. Peak-fractions were concentrated and further purified via gel filtration, using a
1006 HiLoad S200 16-60 column in gel filtration buffer (25 mM HEPES pH 7.9, 250 mM NaCl,
1007 5% Glycerol, 2 mM DTT). Peak fractions were concentrated to 8.6 mg/ml and flash-frozen in
1008 liquid nitrogen.

1009 **C19L1 full-length protein**

1010 Recombinant *Ce* C19L1 was expressed in insect cells using a pGB10 plasmid containing
1011 10x-His-MBP-3C-C19L1. Virus generation, expression, lysis and Ni-NTA chromatography
1012 were performed as described for DHX15. Peak fractions from Ni-NTA chromatography were
1013 diluted to 100 mM NaCl and applied to a 5 ml HiTrapQ FF anion exchange column, using
1014 buffer C and buffer D as described for Heparin chromatography for DHX15. The column was
1015 washed with 5% buffer D (100 mM NaCl) for 5 CV and developed with a linear gradient
1016 from 5 to 40% buffer D over 30 CV, followed by a step gradient at 100% buffer D (2M
1017 NaCl). Peak fractions were either directly concentrated and further purified via SEC on a

1018 HiLoad 200 16-60 column in gel filtration buffer (25 mM HEPES pH 7.9, 250 mM NaCl, 5%
1019 Glycerol, 2 mM DTT), of first cleaved with 3C protease for 2h. 3C protease and the His-
1020 MBP tag were removed via Ni-NTA chromatography prior to SEC.

1021 Peak fractions were concentrated to ~20 mg/mL and flash-frozen in liquid nitrogen.

1022 **C19L1(1-277)**

1023 The *Ce* C19L1(1-277) was cloned into pOPINB+ vector with an N-terminal 10x-His-MBP-
1024 3C tag and expressed in LB medium overnight at 18 °C. Pellets were frozen and resuspended
1025 in buffer A (25 mM HEPES pH 7.9, 500 mM NaCl, 5% Glycerol, 20 mM imidazole), lysed
1026 by sonication and cleared by centrifugation (18 000 rpm for 45 minutes at 4 degrees). The
1027 lysate was filtered through 0.45 µm pores and loaded on HisTrap column. The column was
1028 washed with 5% buffer B (25 mM HEPES pH 7.9, 500 mM NaCl, 5% Glycerol, 500 mM
1029 imidazole) and then eluted with a linear gradient to 60% buffer B over 10 CV. The protein
1030 was further purified using anion exchange chromatography as described for the C19L1 full-
1031 length protein. Tag cleavage and SEC were performed as described for C19L1 full-length
1032 protein but using a Superdex 75 16/60 column. Proteins were concentrated to ~20 mg/mL and
1033 flash-frozen in liquid nitrogen.

1034

1035 **TFIP11 G-patch domain**

1036 The *Ce* TFIP11 G-patch domain (residues 117-221) was cloned into pOPINB+ vector with an
1037 N-terminal 10x-His-MBP-3C tag and expressed in LB medium for 3h at 37 °C. Pellets were
1038 frozen and resuspended in buffer A (25 mM HEPES pH 7.9, 500 mM NaCl, 5% Glycerol, 20
1039 mM imidazole), lysed by sonication and cleared by centrifugation (18 000 rpm for 45
1040 minutes at 4 degrees). The lysate was filtered through 0.45 µm pores and loaded on HisTrap
1041 column. The column was washed with 5% buffer B (25 mM HEPES pH 7.9, 500 mM NaCl,
1042 5% Glycerol, 500 mM imidazole) and then eluted with a linear gradient to 60% buffer B over

1043 10 CV. Peak fractions were pooled and digested with 3C protease to cleave the tag for 3h and
1044 then diluted to 50 mM NaCl and further purified via cation exchange chromatography on a
1045 MonoS column, concentrated to 2 mg/ml and flash frozen.

1046

1047 **Peptide synthesis for pulldowns.**

1048 Peptides were synthesized in-house on a Liberty Blue peptide synthesizer (CEM) using
1049 standard Fmoc chemistry. For each amino acid cycle, 4 min coupling with DIC/Oxyma was
1050 performed. N-term was acetylated on resin with 5% Acetic anhydride/2,5% DIPEA in DMF
1051 for 20min. Peptides were purified on a Phenomenex Luna C18(2) using a 2-45% in 45 min
1052 0,1%TFA/ACN+0,1%TFA gradient. Peptides were synthesized with an N-terminal
1053 fluorescein and a C-terminal biotin modification. The identity and quality of peptides were
1054 confirmed using MALDI-MS (4800 MALDI TOF/TOF, Sciex). The following peptides were
1055 synthesized for pull-downs (Fluo = Fluorescein):

1056 ceC19L2(75-106)-wildtype -Fluo-EDEKNKLSAKILKAEMKGDSDLVKKLKRKLESM-
1057 biotin;

1058 ceC19L2(75-106)-scrambled: Fluo-SKVMKEELKLEDASRNGKATEILDKLLKMLDKK-
1059 biotin;

1060 ceSYF1(788-818)-wildtype: Fluo-SMNKGNISFVRGAGKTVQQNTTENPDEIDLD-biotin;

1061 ceSYF1(788-818)-scrambled Fluo-QAIITSDMLETRENDPNDQTNVNKVGKSGSF-biotin.

1062 Peptides were dissolved in 100 mM HEPES pH 7.9 to a concentration of 1 mM.

1063

1064 **Peptide pulldown assay.**

1065 **C19L2(α 1- α 2) versus C19L1:**

1066 30 μ l of High Capacity Neutravidin Agarose (ThermoScientific) beads were equilibrated in
1067 binding buffer (50 mM KCl, 25 mM HEPES pH 7.9, 2 mM MgCl₂, 10 % glycerol). Next, we

1068 added saturating amounts of either wildtype or scrambled *Ce* C19L2(75-106)-biotin peptide
1069 (12.5 μ L of 1 mM peptide solution) and incubated at 4 °C for 1h on a rotating wheel.
1070 Complete saturation of the neutravidin beads with peptides was indicated by strongly visible
1071 fluorescence in the supernatant after 1h.
1072 Beads were washed three times with 500 μ L wash buffer (binding buffer + 0.05% NP-40)
1073 and the supernatant removed. We then added 15 μ L of full-length, untagged C19L1 at a
1074 concentration of 20 μ M in 50 mM NaCl, 25 mM HEPES, 10% Glycerol and allowed to bind
1075 for 1h at 4 °C. Unbound protein was removed by washing 5 times in 200 μ L wash buffer. The
1076 beads were eluted with 30 μ L elution buffer (200 mM Glycine, pH 2.5) and the eluted
1077 proteins visualized with SDS-PAGE.

1078 ***Ce* DHX15, C19L1(1-277) versus SYF1(788-818):**

1079 The pulldown was performed as described for C19L2(α 1- α 2)-C19L1 but with the following
1080 modifications. The binding buffer contained 25 mM HEPES pH 7.9, 100 mM NaCl, 0.05 mM
1081 ZnCl₂, 20 % glycerol, 0.05% NP-40 and 2 mg/mL BSA. The wash buffer contained 25 mM
1082 HEPES pH 7.9, 100 mM NaCl, 0.05 mM ZnCl₂, 20 % glycerol, 0.05% NP-40. Fluo-
1083 SYF1(788-818)-biotin (wildtype or scrambled) were immobilized at saturating concentrations
1084 as before. Untagged C19L1(1-277) and FLAG-DHX15 were used at a final concentration of
1085 10 μ M each.

1086

1087 ***In vitro* ILS disassembly assay**

1088 For the *in vitro* spliceosome disassembly assay, 500 μ l of extract were used for a PRP19-
1089 3xFLAG IP using 20 μ L of magnetic M2-FLAG beads. The beads were equilibrated prior to
1090 extract addition in wash buffer 1 (20 mM Hepes pH 7.9, 100 mM KCl, 2 mM MgCl₂, 0.05%
1091 NP-40, 10% glycerol, 1 mM TCEP). After extract addition, the beads were incubated for 1
1092 hour at 20 °C with gentle agitation. After this incubation, the beads were washed three times

1093 in 500 μ L wash buffer 1. The washed beads were then resuspended in 300 μ l of disassembly
1094 buffer (wash buffer containing either 10 mM ATP and 20 mM additional $MgCl_2$ or no ATP)
1095 and incubated for 1h at 20 °C with agitation. The beads were then washed 3 times with 500
1096 μ L wash buffer containing 250 mM KCl, and further washed four times with 1 mL of
1097 detergent-free wash buffer (20 mM HEPES, pH 7.5, 150 mM KCl), changing tubes midway
1098 to prevent detergent carry-over. Three triplicates for each condition were performed in
1099 parallel.

1100 **Mass spectrometry analysis**

1101 Co-immunoprecipitated proteins coupled to magnetic beads from the ILS disassembly assay
1102 were digested with LysC on the beads, eluted with glycine followed by trypsin digestion. The
1103 nano HPLC system (UltiMate 3000 RSLC nano system, Thermo Fisher Scientific) was
1104 coupled to an Exploris 480 mass spectrometer equipped with a FAIMS pro interface and a
1105 Nanospray Flex ion source (Thermo Fisher Scientific).

1106 Peptides were loaded onto a trap column (PepMap Acclaim C18, 5 mm \times 300 μ m ID,
1107 5 μ m particles, 100 Å pore size, Thermo Fisher Scientific) at a flow rate of 25 μ l/min using
1108 0.1% TFA as mobile phase. After loading, the trap column was switched in line with the
1109 analytical column (PepMap Acclaim C18, 500 mm \times 75 μ m ID, 2 μ m, 100 Å, Thermo Fisher
1110 Scientific). Peptides were eluted using a flow rate of 230 nl/min, starting with the mobile
1111 phases 98% A (0.1% formic acid in water) and 2% B (80% acetonitrile, 0.1% formic acid)
1112 and linearly increasing to 35% B over the next 120 min. This was followed by a steep
1113 gradient to 95% B in 5 min, stayed there for 5 min and ramped down in 2 min to the starting
1114 conditions of 98% A and 2% B for equilibration at 30 °C.

1115 The mass spectrometer was operated in data-dependent mode, performing a full scan
1116 (m/z range 350-1200, resolution 60,000, normalized AGC target 300%) at 3 different
1117 compensation voltages (CV-45, -60, -75), followed each by MS/MS scans of the most

1118 abundant ions for a cycle time of 0.9 (CV -45, -60) or 0.7 (CV -75) seconds per CV. MS/MS
1119 spectra were acquired using HCD collision energy of 30%, isolation width of 1.2 m/z,
1120 Orbitrap resolution of 30.000, normalized AGC target of 200% and minimum intensity
1121 threshold of 2.5E4. Precursor ions selected for fragmentation (include charge state 2-6) were
1122 excluded for 45 seconds. The monoisotopic precursor selection (MIPS) filter and exclude
1123 isotopes feature were enabled.

1124

1125 **Proteomics data analysis of the ILS disassembly assay and of gradient-purified *Ce***
1126 **spliceosomes ILS**

1127 Raw MS data was loaded into Proteome Discoverer (PD, version 2.5.0.400, Thermo
1128 Scientific). All MS/MS spectra were searched using MS Amanda v2.0.0.16129 (ref.⁶⁵).
1129 Trypsin was specified as a proteolytic enzyme cleaving after lysine and arginine (K and R)
1130 without proline restriction, allowing for up to 2 missed cleavages. Mass tolerances were set to
1131 ± 10 ppm at the precursor and fragment mass level. Peptide and protein identification was
1132 performed in two steps. An initial search was performed against the databases
1133 ID1242_Flag.fasta (1 sequences; 22 residues), ID1242_PRP19.fasta (1 sequences; 492
1134 residues), tags_v11.fasta (28 sequences; 2,153 residues), uniprot_reference_C_elegans_2023-
1135 05-15.fasta (19,823 sequences; 8,134,158 residues) and
1136 PD_Contaminants_TAGs_v20_tagsremoved.fasta. Here, Beta-methylthiolation of cysteine
1137 was searched as fixed modification, whereas oxidation of methionine, deamidation of
1138 asparagine and glutamine were defined as variable modifications. Results were filtered for a
1139 minimum peptide length of 7 amino acids and 1% FDR at the peptide spectrum match (PSM)
1140 and the protein level using the Percolator algorithm⁶⁶ integrated in Proteome Discoverer.
1141 Additionally, an Amanda score of at least 150 was required.

1142 A sub-database of proteins identified in this search was generated and used for a second
1143 search, where the RAW-files were searched using the same settings as above plus considering
1144 additional variable modifications: Beta-methylthiolation on cysteine was set as a fixed
1145 modification, oxidation on methionine, phosphorylation on serine, threonine and tyrosine,
1146 deamidation on asparagine and glutamine, pyro-glu from q on peptide N-terminal glutamine,
1147 acetylation on protein N-Terminus were set as variable modifications. The localization of the
1148 post-translational modification sites within the peptides was performed with the tool ptmRS,
1149 based on the tool phosphoRS⁶⁷. Identifications were filtered using the filtering criteria
1150 described above, including an additional minimum PSM-count of 2 per protein in at least one
1151 sample. The identifications were subjected to label-free quantification using IMP-apQuant⁶⁸.
1152 Proteins were quantified by summing unique and razor peptides and applying intensity-based
1153 absolute quantification (iBAQ⁶⁹) with subsequent normalization of each replicate using
1154 PRP19, based on the MaxLFQ algorithm⁷⁰. Identified proteins were filtered to contain at least
1155 3 quantified peptide groups. Statistical significance of differentially expressed proteins was
1156 determined using limma⁷¹.

1157

1158 **RNA helicase assay**

1159 We used a fluorogenic RNA duplex for the helicase assay, as previously described⁷². For this,
1160 an AlexFluor-488 labeled RNA oligo (5'-AF488-
1161 UAGUACCGCCACCCUCAGAACCUUUUUUUUUUUUUUU-3) was mixed with an
1162 equimolar amount of quenching strand (5'-GGUUCUGAGGGUGGCCCUACUA-BHQ-3')
1163 containing a 'black hole quencher' modification at a concentration of 500 nM in annealing
1164 buffer (50 mM NaCl, 5% glycerol, 1 mM TCEP, 50 mM HEPES) and heated to 95 °C for
1165 five minutes and then cooled to 12 °C over 2 hours to anneal the strands. We then added a
1166 fivefold molar excess of unlabeled competitor DNA (5'-

1167 TAGTACCGCCACCCTCAGAACC-3'). For the helicase assay, the RNA duplex and
1168 competitor strand (50 nM scaffold and 250 nM competitor) were mixed with DHX15
1169 (0.5 μ M), TFIP11 G-patch (1 μ M) or both DHX15 and TFIP11-G-patch in helicase buffer
1170 (50 mM NaCl, 5% glycerol, 1 mM TCEP, 50 mM HEPES, with or without 2 mM ATP, 2
1171 mM MgCl₂) and incubated for 30 minutes at 25 °C.

1172 The reaction was then analyzed using a PheraStar plate reader. For this, we first
1173 digested proteins with proteinase K at 37 °C for 40 minutes to mitigate fluorescence
1174 quenching effects, and then measured fluorescence in 384 well black bottom plates (Greiner),
1175 using 10 μ L sample per well.

1176 Fluorescence values were background subtracted and normalized to the highest
1177 fluorescence value in the experiment. Statistical significance was tested using unpaired t-tests
1178 in the GraphPad Prism software.

1180 **Data availability**

1181 Three-dimensional cryo-EM composite density maps of the *Ce* ILS' and ILS'' have been
1182 deposited in the Electron Microscopy Data Bank under the accession numbers EMD- 19397
1183 to EMD- 19398. The individual maps 1-27 have been deposited under the accession numbers
1184 EMD-50447, EMD-50449 to EMD-50569, and EMD-50471 to EMD-40475. Three-
1185 dimensional cryo-EM composite density map of the human ILS'' have been deposited in the
1186 Electron Microscopy Data Bank under the accession numbers EMD-19399. The individual
1187 human ILS'' maps 1-14 have been deposited under the accession numbers EMD-50477 to
1188 EMD-50490. The coordinate files of the *Ce* ILS', *Ce* ILS'', the revised human P complex,
1189 and the human ILS'' have been deposited in the Protein Data Bank under the accession
1190 numbers 8RO0, 8RO1, 9FMD, and 8RO2.

1191

1192 **METHODS REFERENCES**

- 1193 51. Brenner, S. The genetics of *Caenorhabditis elegans*. *Genetics* **77**, 71–94 (1974).
- 1194 52. Paix, A., Folkmann, A., Rasoloson, D. & Seydoux, G. High Efficiency, Homology-
1195 Directed Genome Editing in *Caenorhabditis elegans* Using CRISPR-Cas9
1196 Ribonucleoprotein Complexes. *Genetics* **201**, 47–54 (2015).
- 1197 53. Mayeda, A. & Krainer, A. R. Preparation of HeLa cell nuclear and cytosolic S100
1198 extracts for in vitro splicing. *Methods Mol Biol* **118**, 309–314 (1999).
- 1199 54. Kastner, B. *et al.* GraFix: Sample preparation for single-particle electron cryomicroscopy.
1200 *Nature Methods* (2008) doi:10.1038/nmeth1139.
- 1201 55. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. CryoSPARC: Algorithms
1202 for rapid unsupervised cryo-EM structure determination. *Nature Methods* **14**, 290–296
1203 (2017).
- 1204 56. Scheres, S. H. W. RELION: Implementation of a Bayesian approach to cryo-EM
1205 structure determination. *Journal of Structural Biology* **180**, (2012).
- 1206 57. Kimanius, D. *et al.* Data-Driven Regularisation Lowers the Size Barrier of Cryo-EM
1207 Structure Determination. <http://biorxiv.org/lookup/doi/10.1101/2023.10.23.563586>
1208 (2023) doi:10.1101/2023.10.23.563586.
- 1209 58. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**,
1210 583–589 (2021).
- 1211 59. Varadi, M. *et al.* AlphaFold Protein Structure Database: massively expanding the
1212 structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids*
1213 *Res* **50**, D439–D444 (2022).
- 1214 60. Kelly, L. A., Mezulis, S., Yates, C., Wass, M. & Sternberg, M. The Phyre2 web portal for
1215 protein modelling, prediction, and analysis. *Nature Protocols* **10**, 845–858 (2015).

- 1216 61. Croll, T. I. ISOLDE: A physically realistic environment for model building into low-
1217 resolution electron-density maps. *Acta Crystallographica Section D: Structural Biology*
1218 **74**, 519–530 (2018).
- 1219 62. Goddard, T. D. *et al.* UCSF ChimeraX: Meeting modern challenges in visualization and
1220 analysis. *Protein Science* **27**, 14–25 (2018).
- 1221 63. Biesiada, M., Purzycka, K. J., Szachniuk, M., Blazewicz, J. & Adamiak, R. W.
1222 Automated RNA 3D Structure Prediction with RNAComposer. *Methods Mol Biol* **1490**,
1223 199–215 (2016).
- 1224 64. Trowitzsch, S., Bieniossek, C., Nie, Y., Garzoni, F. & Berger, I. New baculovirus
1225 expression tools for recombinant protein complex production. *J Struct Biol* **172**, 45–54
1226 (2010).
- 1227 65. Dorfer, V. *et al.* MS Amanda, a universal identification algorithm optimized for high
1228 accuracy tandem mass spectra. *J Proteome Res* **13**, 3679–3684 (2014).
- 1229 66. Käll, L., Canterbury, J. D., Weston, J., Noble, W. S. & MacCoss, M. J. Semi-supervised
1230 learning for peptide identification from shotgun proteomics datasets. *Nat Methods* **4**,
1231 923–925 (2007).
- 1232 67. Taus, T. *et al.* Universal and confident phosphorylation site localization using
1233 phosphoRS. *J Proteome Res* **10**, 5354–5362 (2011).
- 1234 68. Doblmann, J. *et al.* apQuant: Accurate Label-Free Quantification by Quality Filtering. *J*
1235 *Proteome Res* **18**, 535–541 (2019).
- 1236 69. Schwanhäusser, B. *et al.* Global quantification of mammalian gene expression control.
1237 *Nature* **473**, 337–342 (2011).
- 1238 70. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized
1239 p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotech* **26**,
1240 1367–1372 (2008).

- 1241 71. Smyth, G. K. Linear models and empirical bayes methods for assessing differential
1242 expression in microarray experiments. *Stat Appl Genet Mol Biol* **3**, Article3 (2004).
- 1243 72. Tani, H. *et al.* Real-time monitoring of RNA helicase activity using fluorescence
1244 resonance energy transfer in vitro. *Biochem Biophys Res Commun* **393**, 131–136 (2010).
- 1245 73. Yariv, B. *et al.* Using evolutionary data to make sense of macromolecules with a ‘face-
1246 lifted’ ConSurf. *Protein Sci* **32**, e4582 (2023).
- 1247 74. Adams, P. D. *et al.* PHENIX: A comprehensive Python-based system for macromolecular
1248 structure solution. *Acta Crystallographica Section D: Biological Crystallography* **66**,
1249 213–221 (2010).
- 1250

1251 **Acknowledgments**

1252 We thank the Plaschka and Cochella groups for their help and discussions; the Protein
1253 Technologies facility at the Vienna BioCenter Core Facilities GmbH (VBCF), a member of
1254 the Vienna BioCenter (VBC), for assistance with protein production; the VBCF Electron
1255 Microscopy Facility, in particular T. Heuser and H. Kotisch, for support, data collection and
1256 maintaining facilities; V.-V. Hodirnau and L. Lovicar at the Institute of Science and
1257 Technology Austria EM facility for cryo-EM data collection. The computational results
1258 presented were obtained using the CLIP cluster (<https://clip.science>) and we thank R.
1259 Zimmermann and his team for computation support. We thank K. Mechtler and his team for
1260 mass spectrometry; the in-house Molecular Biology Service for reagents; M. Madalinski for
1261 peptide synthesis; M. Kostic (Life Science Editors), C. Bernecky, and A. Stark for critical
1262 reading of the manuscript. M.K.V was supported by a Marie Skłodowska-Curie Postdoctoral
1263 Fellowship (101022449), P.R. was supported by a Boehringer-Ingelheim Fonds fellowship,
1264 L.C. was supported by the National Science Foundation (NSF CAREER Award 2238425).
1265 Research in the laboratory of C.P. is supported by Boehringer Ingelheim, the European
1266 Research Council under the Horizon 2020 research and innovation programme (ERC-2020-
1267 STG 949081 RNApaxport) and by the Austrian Science Fund (FWF) doc.funds program
1268 DOC177-B (RNA@core: Molecular mechanisms in RNA biology).

1269

1270 **Author contributions**

1271 M.K.V. designed research, carried out cryo-EM data analysis and structure determination of
1272 *Ce* spliceosomes, performed biochemical assays, and drafted the initial manuscript. P.R.
1273 carried out cryo-EM data analysis and structure determination of the human ILS, aided by
1274 M.K.V., and prepared the revised human P complex structure. J.K. generated the 3xFLAG-
1275 PRP19 *Ce* strains and prepared *Ce* spliceosomes and cryo-EM grids. E.C. generated *Ce* syf-2

1276 mutant strains, performed RNAi experiments, and analyzed the data with L.C.. L.V. prepared
1277 the human ILS and cryo-EM grids. D.R.-B. generated the human GFP-TFIP11 K562 cell line.
1278 M.K.V., P.R., and L.F. purified proteins. A.W.P grew large GFP-TFIP11 K562 cell cultures
1279 and prepared nuclear extract. M.K.V., P.R., L.C. and C.P. analyzed data and prepared the
1280 manuscript with input from all authors. L.C designed research, supervised E.C., and co-
1281 supervised J.K. with C.P.. C.P. supervised M.K.V, P.R., L.V., D.R-B., L.F., A.W.P.,
1282 designed research, and initiated the spliceosome project.

1283

1284 **Competing interests**

1285 The authors declare no competing interests.

1286

1287 **Additional information**

1288 Supplementary Information is available for this paper. Correspondence and requests for
1289 materials should be addressed to Clemens Plaschka.

1290

1291 **Figure Legends**

1292 **Figure 1 | Structures of a metazoan intron-lariat spliceosome (ILS) in two states.**

1293 **a.** Cartoon schematic of specific ILS disassembly. The ILS ‘prime’ (ILS’) and ‘double-prime’
1294 (ILS’’) states were identified in this study.

1295 **b.** Composite ILS’ and ILS’’ cryo-EM densities from *C. elegans* (*Ce*) are shown from a front
1296 view⁵⁰. The maps range from 2.6 Å to 8.0 Å resolution (U2 3’ domain) and were generated
1297 from 15 (ILS’) or 18 (ILS’’) local three-dimensional refinements. Subunits are colored
1298 according to snRNP identity (U2, green; U5, blue; U6, red; disassembly factors, shades of
1299 purple). A protein color code for each ILS subunit is shown underneath and is used
1300 throughout.

1301

1302 **Figure 2 | Disassembly factors recognize inner and outer ILS surfaces.**

1303

1304 **a.** Domain organization of the disassembly factors TFIP11, PAXBP1, C19L1, C19L2, and

1305 DHX15. Solid lines indicate regions included in the atomic model. CTD, C-terminal domain;

1306 MMP, Matrix Metalloproteinase.

1307 **b.** TFIP11–PAXBP1 recognize the ILS' (left) and ILS'' (right) exterior, whereas C19L1,

1308 C19L2, and DHX15 recognize the ILS'' interior (right). Spliceosome regions not in contact

1309 with TFIP11–PABP1 are shown as transparent surfaces, except for the RNA active site,

1310 which is shown for reference. The black outline indicates the regions shown in panel **c**.

1311 **c.** Interfaces between TFIP11–PAXBP1 and ILS' (left) and ILS'' (right) subunits. On the

1312 ILS'' (right), the numbers 1, 2, and 3 mark regions of change during the ILS' and ILS''

1313 transition: 1, movements at the TFIP11 'Hinge'; 2, movements of the PRP8 RNaseH (RH)

1314 domain and TFIP11 'Hairpin'; and 3, the newly liberated site in 2 is bound by C19L1–

1315 C19L2. See main text for details.

1316

1317

1318 **Figure 3 | Human P complex and ILS'' structures reveal determinants of state-specific**
1319 **disassembly.**

1320 **a.** The revised P complex coordinate model shown from the front. Subunits are colored
1321 according to snRNP identity (U2, green; U5, blue; U6, red; stage-specific proteins, shades of
1322 purple). The '#' indicates that the P complex structure was generated by combining previous
1323 cryo-EM densities and models of human B^{act}, C*, and P complex spliceosomes. Below,
1324 regions of the human P complex that clash with the disassembly factors are shown as
1325 cartoons, the remainder is rendered as a transparent surface. The numbers 1, 2, and 3
1326 highlight regions of the P complex that are used to discriminate P from ILS'' complexes by
1327 the ILS disassembly factors.

1328 **b.** The integrative human ILS'' structure is shown from a front view. Below,
1329 human disassembly factors are highlighted, revealing that they bind the ILS'' similar to their
1330 *Ce* counterparts (compare Fig. 2b). Colors as for the *Ce* ILS'' in Fig. 1.

1331 **c.** Structural comparisons of the human P and ILS'' structures elucidate specific recognition
1332 of the human ILS. P-ILS clashes 1 (left): Shows the P complex-bound mRNP (mRNA 5'-
1333 exon, Exon Junction Complex), and the subunits SRRM2, NOSIP, CWC22 and SLU7 that
1334 clash with the ILS subunits TFIP11–PAXBP1. Structures were aligned on SNU114
1335 (transparent surface). P-ILS clashes 2 (middle): Shows clashes between the P complex
1336 subunits SLU7, PPWD1, and the PRP8 JAB1/MPN domain with the ILS subunits TFIP11–
1337 PAXBP1. Structures were aligned on the PRP8 L domain (transparent surface). P-ILS clashes
1338 3 (right): Shows clashes between the P complex-bound mRNA 3'-exon, the subunits
1339 FAM50A, CACTIN, and the PRP8 RH domain with the ILS subunit C19L2. Structures were
1340 aligned on the PRP8 L domain (transparent surface); the U5 snRNA Loop 1 is shown for
1341 reference to panels **a** and **b**.

1342

1343 **Figure 4 | DHX15 is primed for spliceosome disassembly via U6 snRNA.**

1344 **a.** Interactions of DHX15 (surface) with U6 snRNA and proteins of the *Ce* ILS'' and
1345 disassembly factors (cartoons) are shown from the front. Non-interacting ILS regions are
1346 shown as a transparent surface. DHX15 is rendered as a sliced-through surface to highlight
1347 the U6 snRNA segment bound in its active site.

1348 **b.** DHX15 is positioned on the ILS'' by the NTR subunits SYF1, SYF2 and SDE2, and the
1349 disassembly factor C19L1 MMP domain. SYF2 and SDE2 act as a wall to protect U2/U6
1350 helix II and guide the U-rich U6 snRNA 3'-end into the DHX15 active site. The TFIP11 G-
1351 patch activates DHX15 (Extended Data Fig. 8 and ref.⁴⁹).

1352 **c.** SYF2 and SDE2 use a network of positively charged amino acids to guide the path of the
1353 U6 snRNA 3'-end towards DHX15 and possibly assist the separation of U2/U6 helix II upon
1354 the ATP-dependent translocation of DHX15 on U6 snRNA, from 3' - to 5' - ends. On the
1355 right, a cartoon schematic visualizes the key interactions of SDE2 and SYF2 with U2 and U6
1356 snRNAs.

1357

1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377

Figure 5 | Model for terminal spliceosome disassembly.

The disassembly factors act together with the NTR subunits SYF1, SYF2 and SDE2 to initiate the specific dismantling of the ILS. After ligated mRNP release, the disassembly factors TFIP11–PAXBP1–DHX15 may recognize the ILS first, yielding the ILS'. This may license the binding of C19L1–C19L2 to form the ILS''. This multi-factor authentication would prime DHX15 to unwind the U6 snRNA-based active site, (iv) initiating ILS disassembly for spliceosome recycling and intron-lariat degradation (see Movie S5).

ACCELERATED ARTICLE PREVIEW

1378 **Extended Data Figure Legends**

1379 **Extended Data Figure 1 | Cryo-EM analysis of *C. elegans* spliceosomes.**

- 1380 **a.** Schematic of purification of spliceosomes from *C. elegans*. The endogenous locus of
1381 PRP19 was tagged with an N-terminal FLAG-tag using CRISPR/Cas9 and extract was
1382 prepared from ~12 million adult worms. After immunopurification (IP) and elution with
1383 FLAG peptide, spliceosomes were further purified via a sucrose gradient.
- 1384 **b.** Coomassie-stained SDS-Poly-Acrylamide Gel (SDS-PAGE) of gradient-purified *Ce*
1385 spliceosomes. This experiment was performed seven times.
- 1386 **c.** Denoised cryo-EM micrograph of gradient-purified and crosslinked *Ce* spliceosomes
1387 imaged on a Titan Krios with a K3 detector.
- 1388 **d.** 2D class averages from the dataset.
- 1389 **e.** Abundance of ILS subunits in gradient-purified sample measured by mass spectrometry.
1390 For this analysis we quantified absolute protein abundances by integrating the protein
1391 peptide peaks and normalizing to the protein length using iBAQ⁶⁹, which were then
1392 normalized to PRP8. The labels next to the bars indicate how many peptides were
1393 identified for each subunit and which percent of the sequence was covered.
- 1394 **f.** Schematic of the data analysis pipeline. Stringent classification of ~4 million single
1395 particle images revealed the ILS' (~85-90% of ILS particles) and ILS'' (~10-15% of ILS
1396 particles, see Supplementary Data Fig. 1) as the major PRP19-containing spliceosomes
1397 populations in *Ce* extract. Extensive focused refinements of each state yielded a total of
1398 27 maps, revealing the ILS' and ILS'' in unprecedented detail and facilitating the
1399 building of high-quality structural models. For details, see Extended Data Fig. 2 and
1400 Supplementary Data Figs 1,2.

1401 **g.** Sequence conservation plot of ILS subunits between human and *C. elegans* (*Ce*), and
1402 human and *Saccharomyces cerevisiae* (*Sc*) shows a highly conserved ILS protein
1403 composition between human and *Ce*.

1404

1405 **Extended Data Figure 2 | Comparison of the complete *C. elegans* ILS'' to a partial**
1406 **human ILS2.**

1407 **a. -b.** Side-by-side comparison of the *Ce* ILS'' cryo-EM density map with the deposited
1408 human ILS2 map. Top: Overview, with cryo-EM density colored by subunits. For the
1409 human ILS2 (EMD-9647), a low pass filtered map (gaussian filter with a width of three
1410 standard deviations) is shown in addition (transparent white surface). Bottom: Zoom-ins
1411 to the spliceosomes core reveal nearly indistinguishable densities where high-resolution
1412 density is available for both *Ce* and *Hs* ILS.

1413 **c.** Coordinate model statistics for *Ce* ILS'' and *Hs* ILS2, listing number of residues included
1414 as full sidechain models or backbone models, respectively. Numbers in brackets indicate
1415 completeness relative to the sum of all residues calculated from deposited sequences for
1416 the full-length proteins.

1417 **d.** ILS subunit diagrams indicating which residues are included in coordinate models of the
1418 *Ce* ILS'' or *Hs* ILS2 as full side chain models (solid fill), backbone models (semi-
1419 transparent fill with stripes), or not modelled (transparent fill). Asterisks indicate severe
1420 register error in deposited human ILS2 models in SYF1 (register error of up to 120
1421 residues) and SYF3 (register error of ~20 residues).

1422

1423 **Extended Data Figure 3 | Yeast and metazoan ILS architectures are poorly conserved.**

1424 **a.** Comparison of disassembly factors observed in available baker's yeast (*S. cerevisiae*, *Sc*),
1425 fission yeast (*S. pombe*, *Sp*), human (*Hs*) or nematode (*Ce*) ILS structures.

- 1426 **b.** Cartoon representation of the *Ce* TFIP11-PAXBP1 heterodimer.
- 1427 **c.** Cartoon representation of the *Sc* TFIP11-PAXBP1 homolog Ntr1-Ntr2, with the G-patch
1428 factor Ntr1 aligned to its homolog TFIP11.
- 1429 **d.** Side views of the ILS from *Sc*, *Sp*, *Hs* and *Ce*, with disassembly factors shown in ribbon
1430 representations and the ILS core shown as a transparent white surface.
- 1431 **e.** Yeast *Sc* Ntr1-Ntr2 (transparent ribbons) overlaid on the *Ce* ILS'', revealing
1432 substantially different binding sites on the ILS. Structures were aligned on PRP8.
- 1433
- 1434 **Extended Data Figure 4 | Conformational and compositional changes from ILS' to**
1435 **ILS''.**
- 1436 **a.** Close-up view of protein-protein interactions between the PRP8 RNase H (RH) domain,
1437 TFIP11, PAXBP1 and BRR2 in the ILS'. Protein elements that are mobile in the ILS'-to-
1438 ILS''-transition are shown as ribbons, whereas elements that are static are shown in
1439 addition as transparent surfaces.
- 1440 **b.** Close-up view of protein-protein interactions between the PRP8 RNase H (RH) domain,
1441 TFIP11, PAXBP1 and C19L2 in the ILS''. C19L2 binding requires repositioning of the
1442 PRP8 RH domain and TFIP11-PAXBP1, which displaces the BRR2-PRP8 JAB1/MPN
1443 domains from PAXBP1. C19L2 recruits C19L1 by binding its C-terminal CWFJ domain.
- 1444 **c.** Overlay of TFIP11-PAXBP1 in the ILS' and ILS'' in a 90° rotated view. Yellow arrows
1445 connect identical residues in both states.
- 1446 **d.** Overview of the ILS'. DHX15 (transparent) is likely tethered via the TFIP11 G-patch
1447 domain but cannot dock onto its target.
- 1448 **e.** Overview of the ILS''. C19L1 and C19L2 binding allows docking of DHX15, and the
1449 associated conformational change in TFIP11-PAXBP1 and PRP8 displaces BRR2.
- 1450 Circled numbers 1 and 2 indicate regions of zoom-ins in panels f, i and j.

- 1451 **f.** Close-up view of the ILS'' U6 snRNA 3' end, with DHX15 and the TFIP11 G-patch
1452 removed for clarity. The oligo-uridylated and single-stranded U6 snRNA 3' end is the
1453 ideal substrate for DHX15, and SDE2 and SYF2 shield the U2/U6 helix II.
- 1454 **g.** RNA cryo-EM density in the ILS'. After dissociation of ligated mRNA and catalysis-
1455 specific splicing proteins, the RNA active site is more mobile.
- 1456 **h.** RNA cryo-EM density in the ILS''. Compared to the ILS', RNA densities are better
1457 defined in the ILS'', presumably due to binding of C19L2 (see panel **j**).
- 1458 **i.** Continuous cryo-EM density between U2-U6 helix II and the DHX15 active site reveals
1459 U6 snRNA as the target for ILS disassembly.
- 1460 **j.** C19L2 binds the active site RNA network near the branch helix and contacts U2 snRNA,
1461 intron-lariat RNA, U5 snRNA, and U6 snRNA.

1462

1463 **Extended Data Figure 5 | AlphaFold2 Multimer predictions support disassembly factor**
1464 **interactions.**

1465 **a.** AlphaFold2 Multimer prediction of full-length *Ce* C19L2–C19L1. The prediction is
1466 shown colored by subunit (left) or AlphaFold2 confidence score (per residue local
1467 difference distance test, pLDDT, right). The prediction supports the binding of the C19L1
1468 CWFJ domain to C19L2. The C19L1 MMP domain (rendered transparent) is predicted to
1469 be collapsed onto the structure, however our experimental cryo-EM density shows that in
1470 the ILS'' the C19L1 MMP domain is distant from the C19L1 CWFJ–C19L2 complex.
1471 Note that the C19L1 CWFJ–C19L2 interaction is predicted with low confidence and in
1472 only 2 of 5 models (panel c).

1473 **b.** pLDDT scores of the 5 models plotted over the amino acid number. Scores for the 5
1474 models are overlaid.

- 1475 **c.** Predicted aligned error (PAE) plot of the 5 models, sorted from highest ranked prediction
1476 (left) to lowest ranked prediction (right).
- 1477 **d.** Pull-down experiment with immobilized C19L2(α 1- α 2) peptide and recombinant C19L1.
1478 C19L1 binds the wildtype C19L2(α 1- α 2) peptide but not a C19L2(α 1- α 2) scrambled
1479 peptide control. Small insets underneath the lanes show fluorescent images of the beads
1480 with immobilized fluorescently labelled C19L2 peptides in the fluorescein channel to
1481 show equal loading of the wildtype and scrambled C19L2 peptides. This experiment was
1482 performed once.
- 1483 **e.** Overview (left) and close-up (right) view of the *Ce* C19L1–DHX15–SYF1 interfaces in
1484 the ILS”. C19L1 and SYF1 jointly bind a conserved hydrophobic pocket in the DHX15
1485 CTD. The close-up panel shows the DHX15 surface colored by molecular hydrophobicity
1486 potential, with white colors indicating hydrophobic surfaces.
- 1487 **f.** The same close up as in panel e (right), but colored by sequence conservation. Residue
1488 conservation scores were obtained from the ConSurf server⁷³.
- 1489 **g.** AlphaFold2 Multimer prediction of *Hs* DHX15 with C19L1 and SYF1 suggests a
1490 conserved binding mode and conserved hydrophobic residues in the DHX15 CTD, the
1491 C19L1 ‘loop 2’, and the SYF1 ‘tether’.
- 1492 **h.** Pull-down experiment with immobilized SYF1(788–818) peptide and recombinant
1493 C19L1 and DHX15. C19L1 and DHX15 both bind the SYF1(788–818) ‘tether’ peptide
1494 but not a scrambled peptide control. C19L1 and DHX15 can also simultaneously bind to
1495 the wildtype but not the scrambled peptide. Small insets underneath the lanes show
1496 images of the beads with immobilized fluorescently labelled SYF1 peptides in the
1497 fluorescein channel to show equal loading of the wildtype and scrambled SYF1 peptide.
1498 This experiment was three times.

- 1499 **i.** Predicted aligned error plots of the *Ce* and *Hs* DHX15–C19L1–SYF1 AlphaFold2
1500 Multimer predictions.
- 1501 **j.** Predicted local distance difference test (pLDDT) plots of the AlphaFold2 Multimer
1502 predictions from panel **i**.
- 1503 **k.** SYF3 might assist with positioning the C19L1 MMP domain in the ILS''. Zoom-in of an
1504 AlphaFold2 Multimer prediction between *Hs* SYF3 and C19L1, highlighting the interface
1505 between a SYF3 C-terminal β -hairpin (residues 780-805) that extends the C19L1 MMP
1506 domain central β -sheet. The SYF3 β -hairpin might be flexible relative to the HAT (half a
1507 tetratricopeptide repeat) domain through movement around a hinge residue (indicated
1508 with an arrow).
- 1509 **l.** As panel **k**, but for the *Ce* proteins.
- 1510 **m.** The *Ce* SYF3–C19L1 AlphaFold2 Multimer prediction overlaid onto C19L1 in the *Ce*
1511 ILS'' cryo-EM structure. Cryo-EM density is shown as a transparent surface. Weak
1512 density is visible at the position predicted for SYF3 by AlphaFold2 Multimer, and also
1513 near the end of the SYF3 HAT domain (circled with a dashed line), indicating that the
1514 putatively assigned SYF3 β -hairpin might alternate between a C19L1-bound and C19L1-
1515 unbound conformation.
- 1516 **n.** and **o.** AlphaFold2 Multimer PAE and pLDDT plots for the predictions shown in **k** and **l**.
1517

1518

1519 **Extended Data Figure 6 | Cryo-EM analysis of human ILS spliceosomes.**

- 1520 **a.** Schematic of purification of TFIP11-bound spliceosomes from human cells. GFP-
1521 TFIP11 was overexpressed in K562 suspension cells and TFIP11-bound spliceosomes
1522 were purified from 30 L of suspension cell culture. After immunoprecipitation (IP)
1523 and elution with 3C protease, spliceosomes were further purified via a sucrose
1524 gradient.
- 1525 **b.** Coomassie-stained SDS-Poly-Acrylamide Gel (SDS-PAGE) of the TFIP11-GFP IP.
1526 Bands in the gel are labelled according to the molecular weight of the ILS subunits.
1527 This experiment was performed four times.
- 1528 **c.** Denoised micrograph of gradient-purified and crosslinked *Hs* ILS, imaged on a Titan
1529 Krios G4 with a Falcon 4i detector.
- 1530 **d.** 2D class averages from the dataset.
- 1531 **e.** Composite cryo-EM density, obtained from 14 local refinements and filtered by local
1532 resolution. Transparent density in the background shows a local refinement map
1533 (focused on PAXBP1) low-pass filtered with a gaussian filter with a sigma of three
1534 standard deviations.
- 1535 **f.** Model of the human ILS'', with disassembly factors shown as ribbons and
1536 spliceosome core proteins shown in addition as a transparent surface. A difference
1537 density, calculated by subtracting simulated model density (low-pass filtered to 20Å
1538 resolution) from experimental density (ILS consensus refinement map low-pass
1539 filtered with a gaussian filter with a sigma of three standard deviations) reveals
1540 additional density at the *Ce* ILS'' C19L1 CWFJ position.

1541

1542

1543 **Extended Data Figure 7 | Release of mRNP and spliceosome proteins from the post-**
1544 **catalytic spliceosome unmask binding sites for the disassembly factors.**

- 1545 **a.** Overview cartoon, placing the depicted structures into context of the spliceosome
1546 disassembly pathway.
- 1547 **b.** Structural comparison of the structures of the P complex (Model of a *Ce* P complex
1548 based on the updated human P complex, this work), the intron lariat spliceosome
1549 immediately after mRNP release (modelled) and the ILS' (*Ce* structure, this work).
1550 Proteins that are exchanged in the transition are labelled. Numbers indicate regions
1551 for zoom ins in panels **c-e**.
- 1552 **c.** Overlay and close-up view of the P-complex structure with the ILS' reveals a clash
1553 of TFIP11 with the EJC (EIF4A3 subunit) and with CWC22, NOSIP, and SRRM2 in
1554 the P complex. This clash would occur both in the ILS' and ILS''. Clashing proteins
1555 are outlined in black.
- 1556 **d.** Overlay and close-up view of the P complex structure with the ILS' reveals a clash
1557 between PPWD1 and PAXBP1 on BRR2.
- 1558 **e.** Overlay and close-up view of the P complex structure with the ILS'' reveals a clash
1559 between C19L2 and the path of the ligated exons in the P-complex.

1560

1561 **Extended Data Figure 8 | The ILS'' is competent for disassembly upon ATP addition.**

- 1562 **a.** Comparison between the structures of DHX15 bound to the G-patch domains of
1563 TFIP11 (this study), NKRF1 (ref.³⁰) (PDB 6SH7), and SUGP1 (ref.⁴¹) (PDB 8EJM).
1564 All G-patch domains show an identical binding mode, but additional residues are
1565 observed in the TFIP11 G-patch in the *Ce* ILS''.
- 1566 **b.** Sequence alignments of the G-patch domains shown in **a**.

- 1567 **c.** Schematic of a fluorescence-based helicase assay as described in ref.⁷² in which an
1568 RNA substrate with 3' overhang and a 5' fluorophore label (AlexaFluor588) is
1569 annealed to a complementary RNA that carries a fluorescence quencher (black hole
1570 quencher, BHQ) at its 3' end. When the RNA strands are annealed, fluorescence is
1571 quenched. Upon separation of the RNA duplex by a helicase, fluorescent signal is
1572 increased. To prevent re-annealing, an excess of an unlabeled DNA strand
1573 complementary to the AF588-labeled RNA is added (not shown in schematic).
1574 Sequences for RNA and DNA used are as in ref.⁷².
- 1575 **d.** The *Ce* TFIP11 G-patch stimulates DHX15 helicase activity. Helicase assay was
1576 performed as shown in panel **c**. N=4 replicates were measured. Fluorescence
1577 intensities were measured in a plate reader, background corrected, and normalized to
1578 the highest value. Error bars show the standard deviation of the mean. P-values from
1579 pairwise two-sided t-tests are indicated.
- 1580 **e.** Denaturing PAGE analysis of the RNAs from the helicase assay shown in panel **d**
1581 after incubation of the RNAs with proteins and ATP. No degradation of the RNA was
1582 observed. This experiment was performed three times.
- 1583 **f.** DHX15 is not bound to ATP in the *Ce* ILS'' cryo-EM structure. Comparison of
1584 DHX15-RNA structure in the ILS'' (left) and a crystal structure of the highly
1585 conserved *Chaetomium thermophilium* (*Ct*) PRP43 (PDB ID 5LTA , ref.³¹, 60 %
1586 sequence identity between *Ce* DHX15 and *Ct* Prp43) bound to RNA and the ATP
1587 mimic ADP-BeF₃ (middle). In presence of ADP-BeF₃, DHX15 adopts a closed
1588 conformation, compressing the RNA so that nucleotide +5 (n+5) is flipped outwards
1589 and no longer forms a stacking interaction with the neighboring bases. Right:
1590 Overlays of the *Ce* ILS'' RNA density (transparent red) with the modelled U6 snRNA
1591 3' end, or the poly-U RNA conformation of Prp43 in presence of ADP-BeF₃ indicates

1592 that in the *Ce* ILS'' the RNA is relaxed and DHX15 is ATP-unbound. This is further
1593 confirmed by the lack of density in the DHX15 ATP binding pocket in the *Ce* ILS''
1594 (not shown).

1595 **g.** *In vitro* ILS disassembly assay. Left: schematic of the assay. Spliceosomes were
1596 immobilized on beads via the PRP19-3xFLAG tag, washed, and incubated with ATP.
1597 Upon ILS disassembly, components of the Nineteen core complex (NTC core), the
1598 Nineteen related complex (NTR) and the U5 snRNP should remain immobilized,
1599 while the disassembly factors and U2 snRNP proteins should be depleted from the
1600 beads. The bead bound fraction was then analyzed by mass spectrometry. Right:
1601 Volcano plot showing differential abundance of proteins with or without ATP
1602 treatment. Consistent with the ILS'' structure, the disassembly factors TFIP11,
1603 PAXBP1, DHX15, C19L1, C19L2, the NTR subunit SDE2, and the U2 snRNP
1604 subunits U2A' (RU2A) and U2B'' (RU2B) are depleted upon ATP treatment,
1605 suggesting that the *Ce* ILS'', which constitutes a minor fraction of spliceosomes in *Ce*
1606 extract according to cryo-EM particle classification (Supplementary Data Fig. 1), is
1607 competent for *in vitro* disassembly. ILS subunits are indicated by large circles and
1608 color-coded according to subcomplex. The horizontal line at $p=0.05$ indicates the
1609 commonly used statistical significance cutoff.

1610 **h.** Fold reduction of ILS subunit abundance after incubation with ATP and PRP19-
1611 3xFLAG IP as determined by mass spectrometry in panel **g**.

1612

1613 **Extended Data Figure 9 | Genetics in *C. elegans* support roles of SYF2.**

1614 **a.** *Ce* SYF2 and *Ce* SDE2 bind the U2-U6 helix II in the ILS''. U2 snRNA, U6 snRNA,
1615 SYF2 and SDE2 are shown as ribbons and DHX15 is shown as an outline.

- 1616 **b.** An AlphaFold2 Multimer prediction of human SDE2 with SYF2 and SYF1 suggests
1617 an identical binding mode of *Hs* SDE2, however it was not observed in the
1618 experimental density due to limited local resolution.
- 1619 **c.** Sequence alignments of *Ce* and *Hs* SDE2 and SYF2.
- 1620 **d.** Schematic of *syf-2* mutant alleles generated by CRISPR-Cas9 in *C. elegans*.
- 1621 **e.** Viability of *syf-2* mutant animals. Single worms of the indicated genotypes were
1622 placed on individual plates at the L3/L4 stage and grown at 20 °C for 96 hours.
1623 Animals with deletion of helix 1 (*Δanchor*) are viable as homozygotes, but animals
1624 with deletion of helices 1 and 2 (*Δanchor+wedge*) are only viable as heterozygotes;
1625 homozygous mutants are thus progeny of heterozygous mothers. Sterility was scored
1626 as the inability to produce numerous progeny that developed into L4 larvae. A few
1627 sterile animals still produced <10 embryos or early larvae but these did not develop
1628 further.
- 1629 **f.** Viability of wild-type or *syf-2 Δanchor* mutant strains treated with
1630 empty vector (e.v.) or anti-*syf-2* RNAi, at standard (20 °C) or low (15 °C) culture
1631 temperatures. Worms were synchronized as L1 larvae, placed on RNAi plates and
1632 grown at the corresponding temperatures. Viability was assessed as the total number
1633 of F1 progeny that reached the L4 stage. N=3 animals were analyzed for assays at
1634 15 °C and N=5 animals were analyzed for assays at 20 °C. P-values from pairwise
1635 two-sided t-tests are indicated.
- 1636 **g.** Measurement of the synthetic effect of RNAi against *sde-2* on *syf-2 Δanchor* mutant
1637 viability. Viability was measured as described in **e** at both 15 °C and 20 °C. RNAi
1638 against *mog-7/PAXBP1* was used as a positive control as an essential splicing protein.
1639 N=3 animals were analyzed for assays at 15 °C and N=5 animals were analyzed for
1640 assays at 20 °C. P-values from pairwise two-sided t-tests are indicated.

1641 **EXTENDED DATA TABLES**

1642 **Extended Data Table 1 | Proteins contained in human P complex and ILS spliceosomes**
1643 **and their homologs in *Ce* and *S. cerevisiae* yeast.**

1644 Orthologs of splicing proteins relevant to this work across *S. cerevisiae*, *C. elegans*, and *H.*
1645 *sapiens*. The protein color code as used throughout. Orthologs were assigned using The
1646 Alliance of Genome Resources (<https://www.alliancegenome.org>), as well as The
1647 Spliceosome Database (<http://spliceosomedb.ucsc.edu>). Though no discrepancies were
1648 identified between these two sources, not all orthologs were predicted with equal confidence.
1649 The table above does not provide information on the confidence of ortholog prediction for
1650 each protein. For additional details, please see Supplementary Data Table 2.

1651

1652

1653 **Extended Data Table 2 | Cryo-EM model refinement, data collection, and map statistics.**

1654 **a.** Cryo-EM data collection and focused refinement map statistics for the *Ce* ILS' and
1655 *Ce* ILS''.

1656 **b.** Cryo-EM data collection and focused refinement map statistics for the *Hs* ILS''.

1657 **c.** Refinement statistics for the coordinate models of the *Ce* ILS', *Ce* ILS'', *Hs* ILS'',
1658 and the revised *Hs* P complex. Refinement statistics were calculated using Phenix⁷⁴.

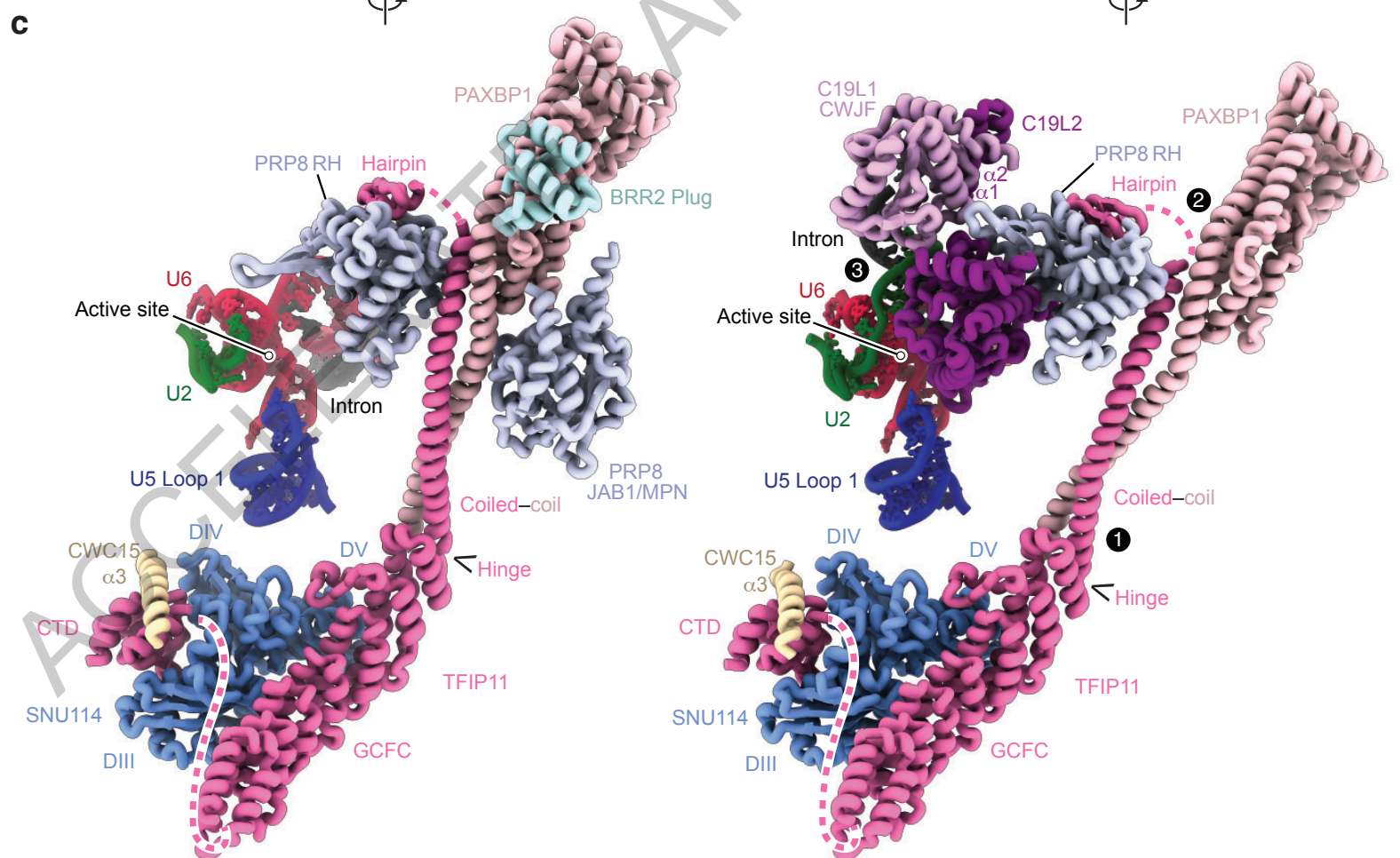
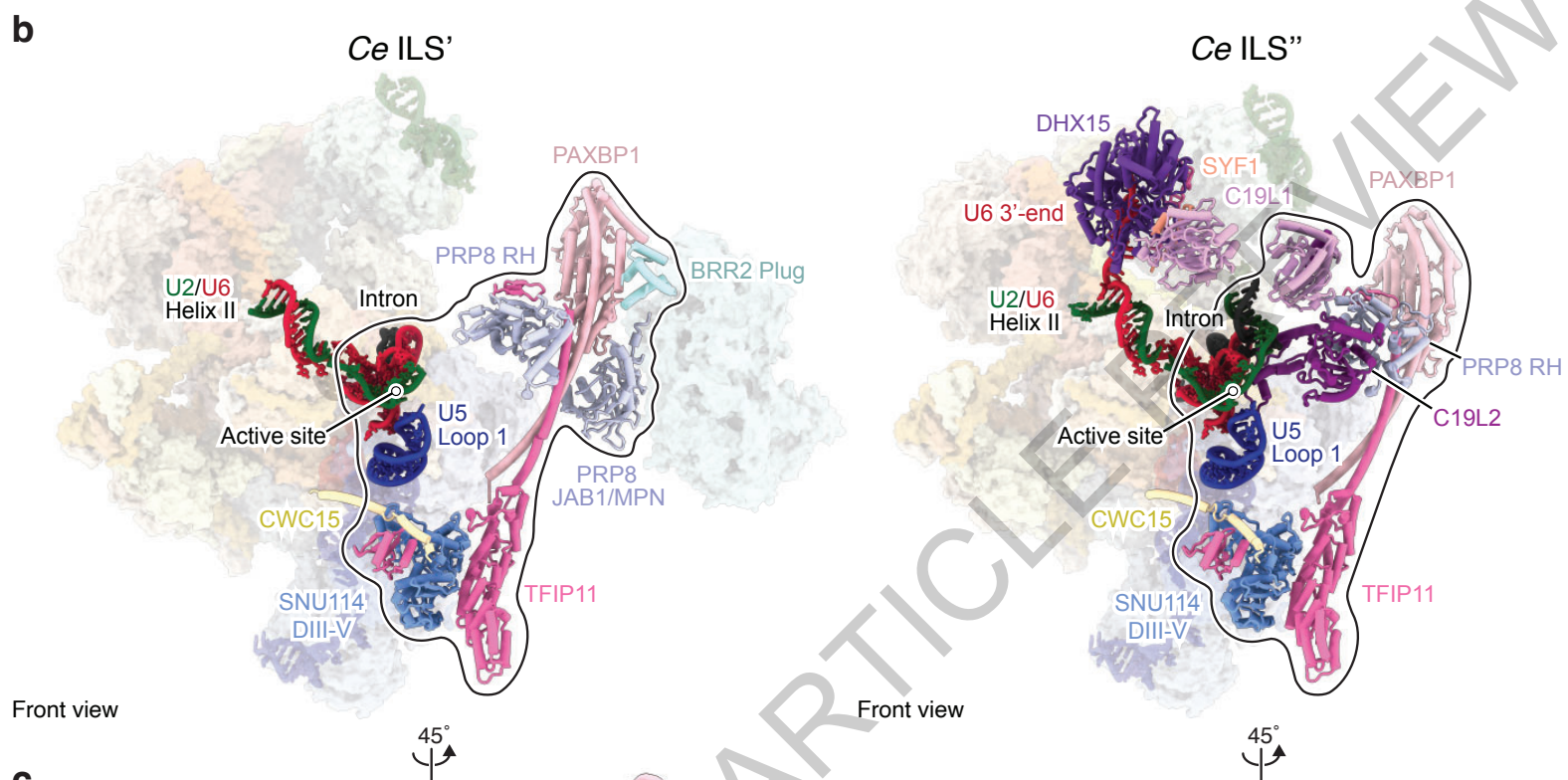
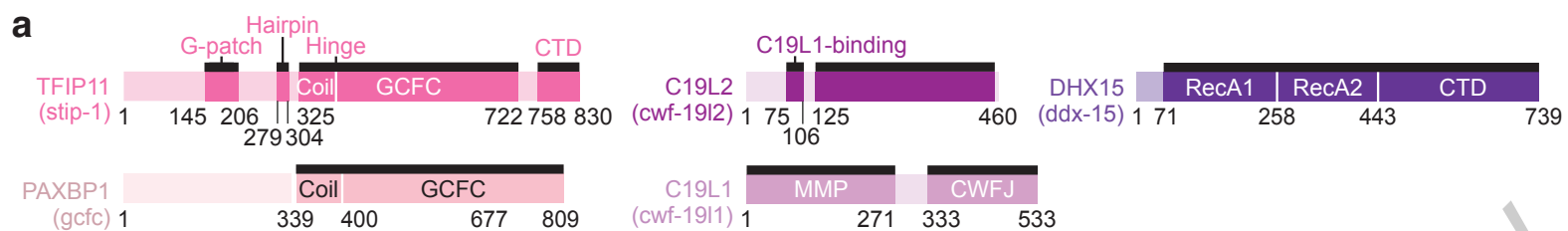
1659

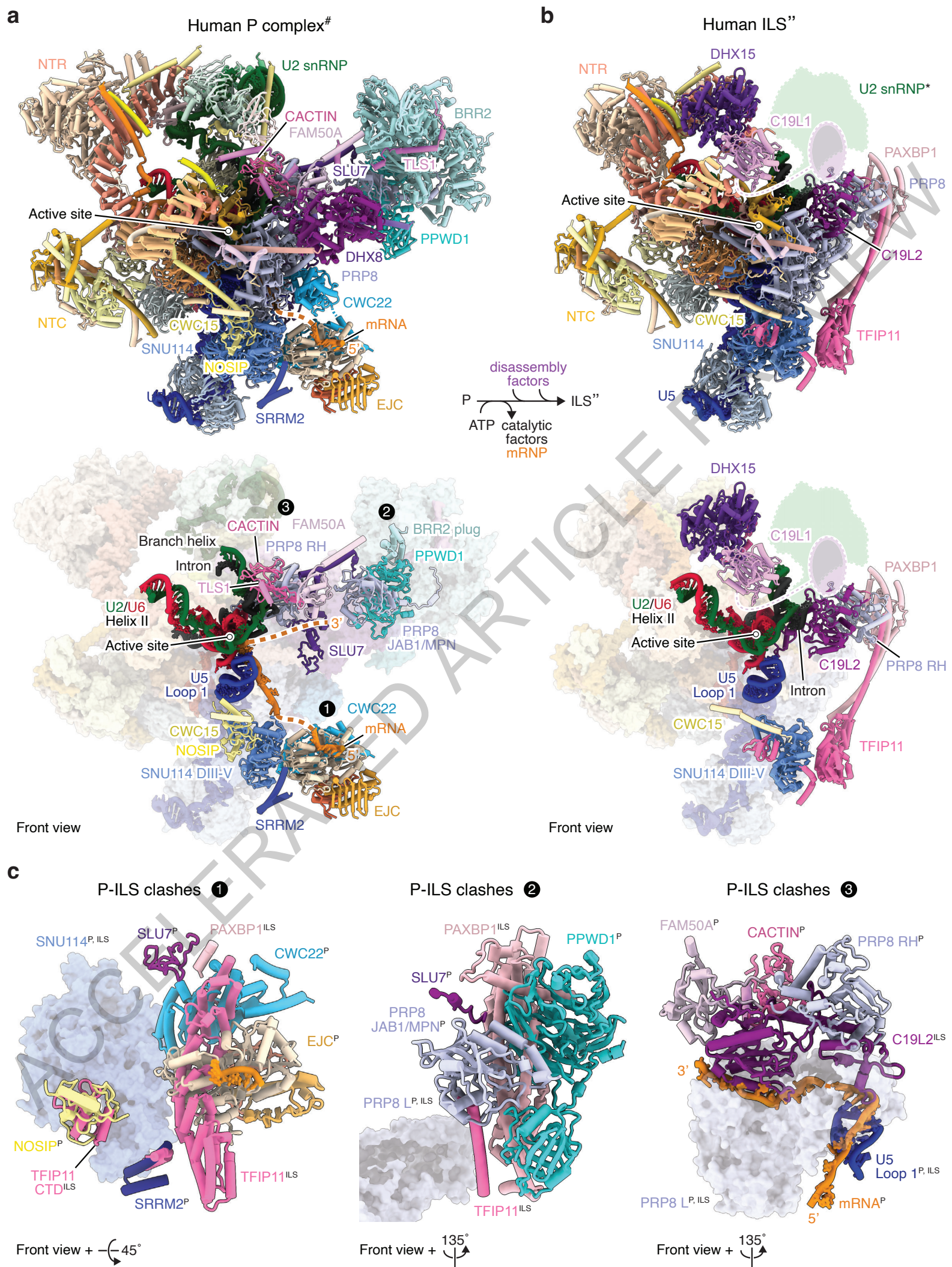
1660

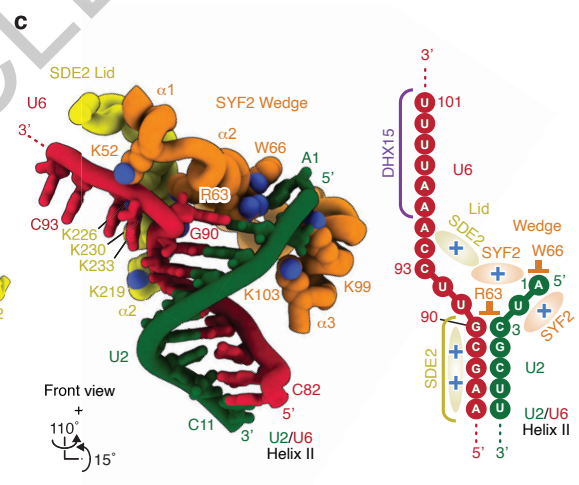
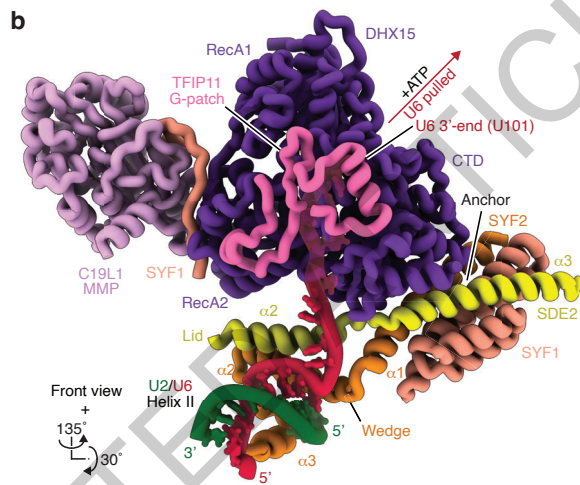
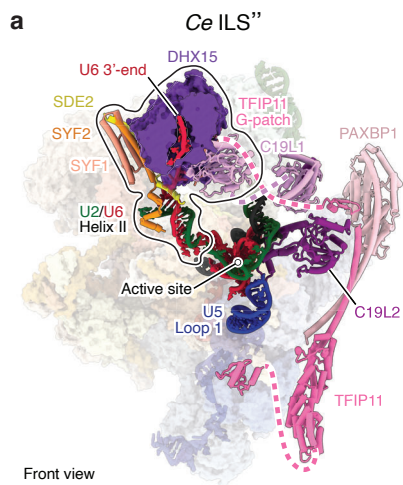
1661

1662

ACCELERATED ARTICLE PREVIEW

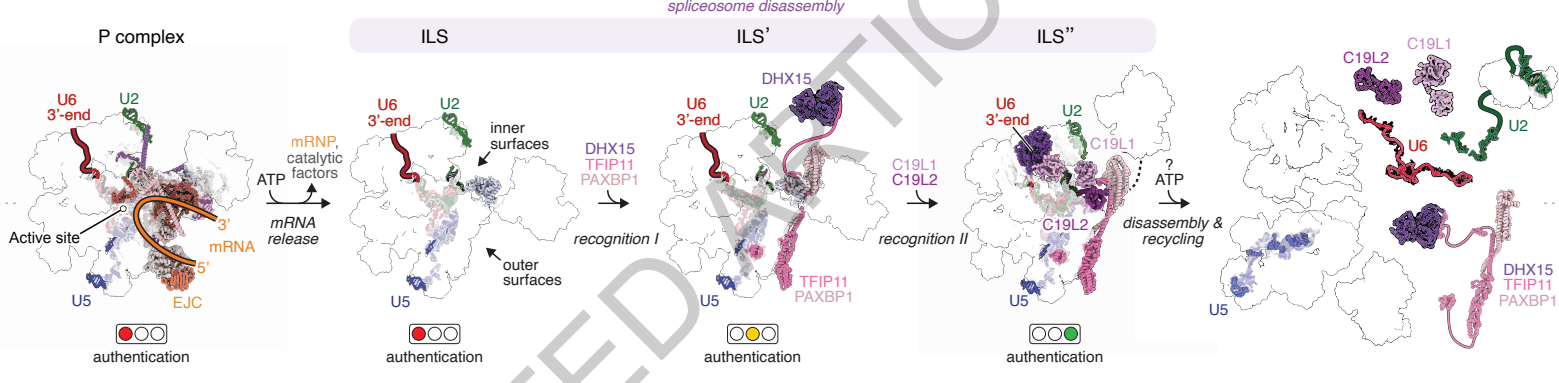


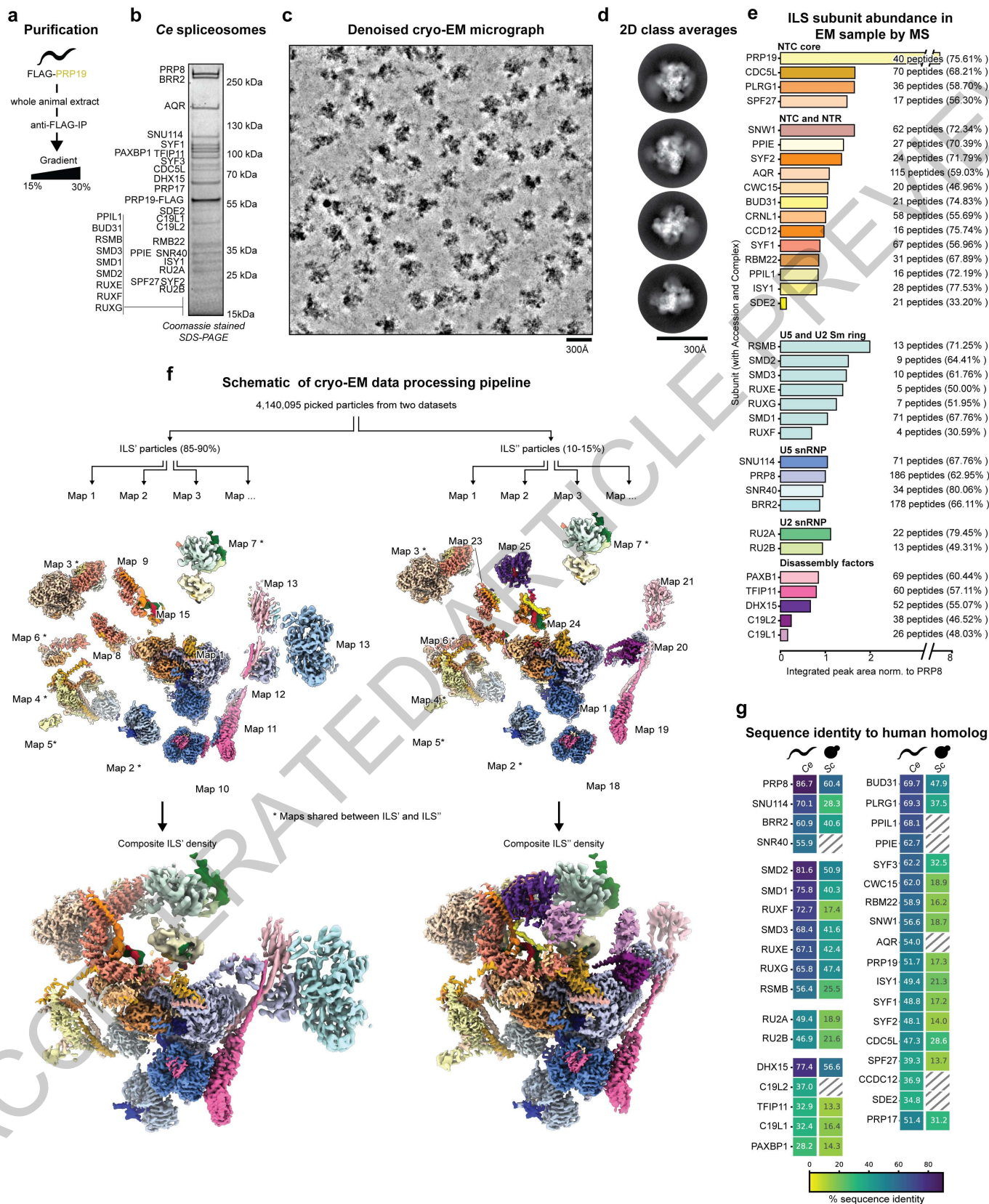




ACCELERATE YOUR MANUSCRIPT PREVIEW

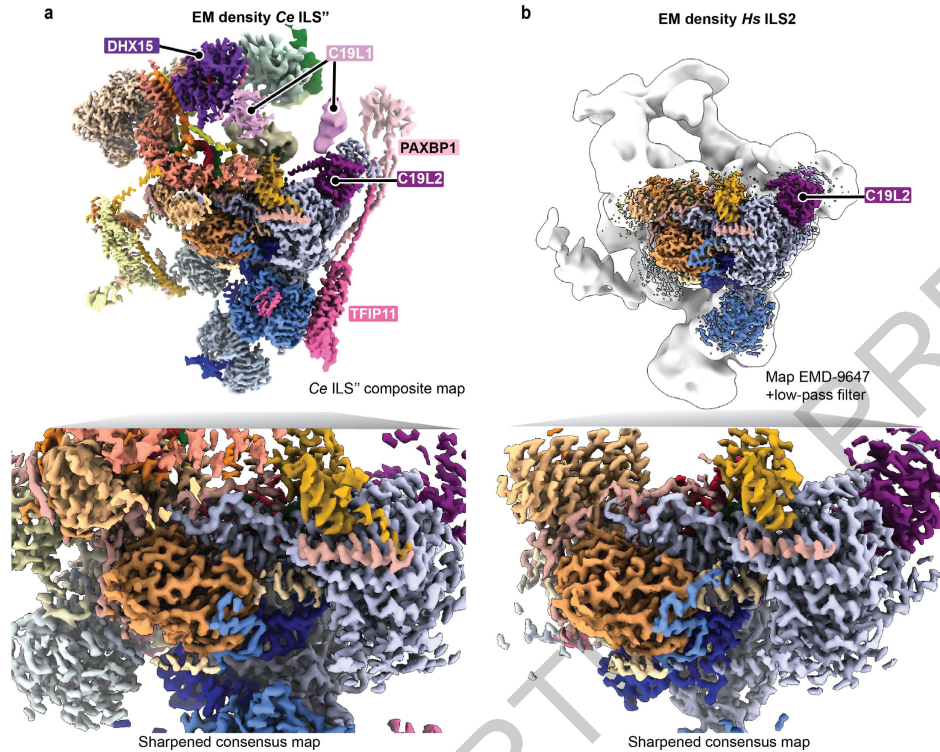
ACCELERATED ARTICLE PREVIEW





Extended Data Fig. 1

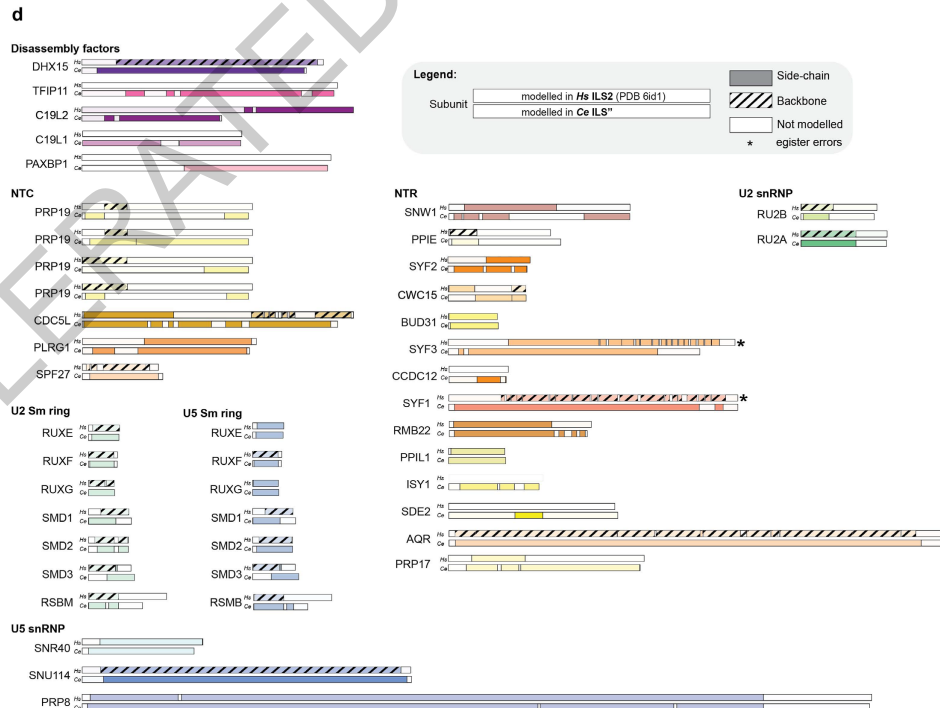
Ce ILS^{''} and Hs ILS2 densities are highly similar at high resolution but the Ce ILS^{''} densities are dramatically improved in peripheral regions



c

Model completeness

	Ce ILS ^{''}	Hs ILS2 (PDB 6id1)
Residues (with sidechains)[%]	14,075 [73.5%]	4,496 [28.9%]
Residues (backbone only)[%]	0 [0%]	5,903 [37.9%]

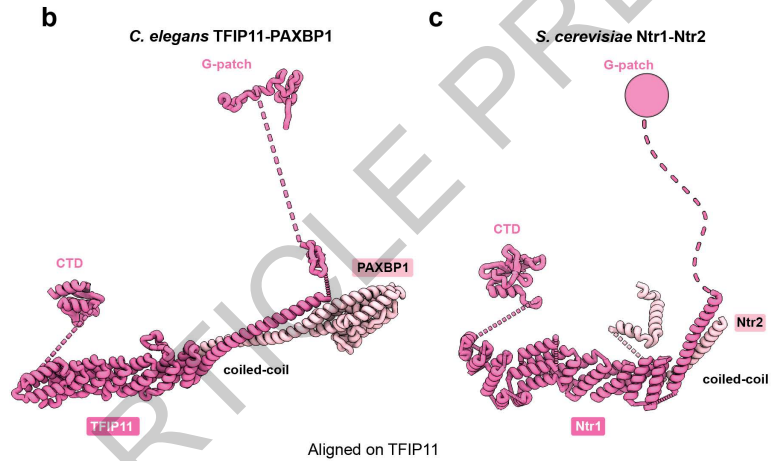


Disassembly factors resolved in ILS cryo-EM structures

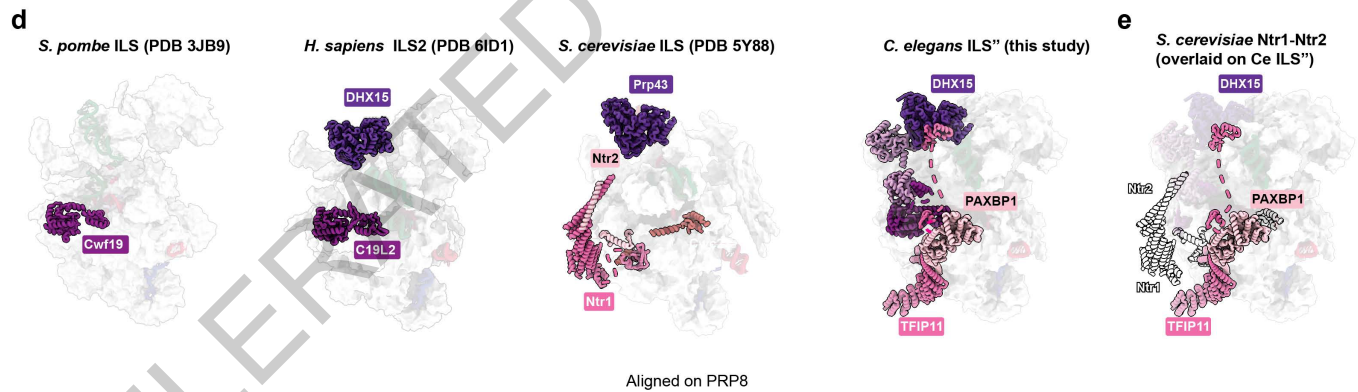
a

	Sc (5Y88)	Sp (3JB9)	Hs (6ID1)	Ce ILS" (this study)
TFIP11 (Ntr1)	✓	-	-	✓
PAXBP1(Ntr2)	✓	-	-	✓
C19L2 (Cwf19)	-	✓	✓	✓
C19L1 (Drm1)	-	-	-	✓
DHX15 (Prp43)	✓	-	✓	✓

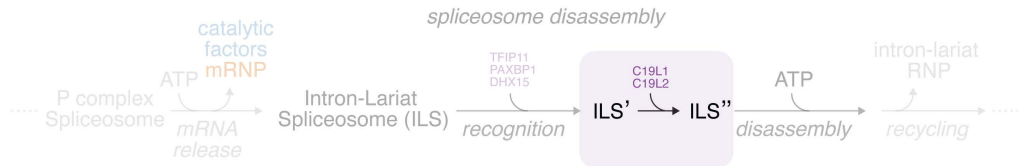
TFIP11-PAXBP1 are structural homologs of bakers yeast Ntr1-Ntr2



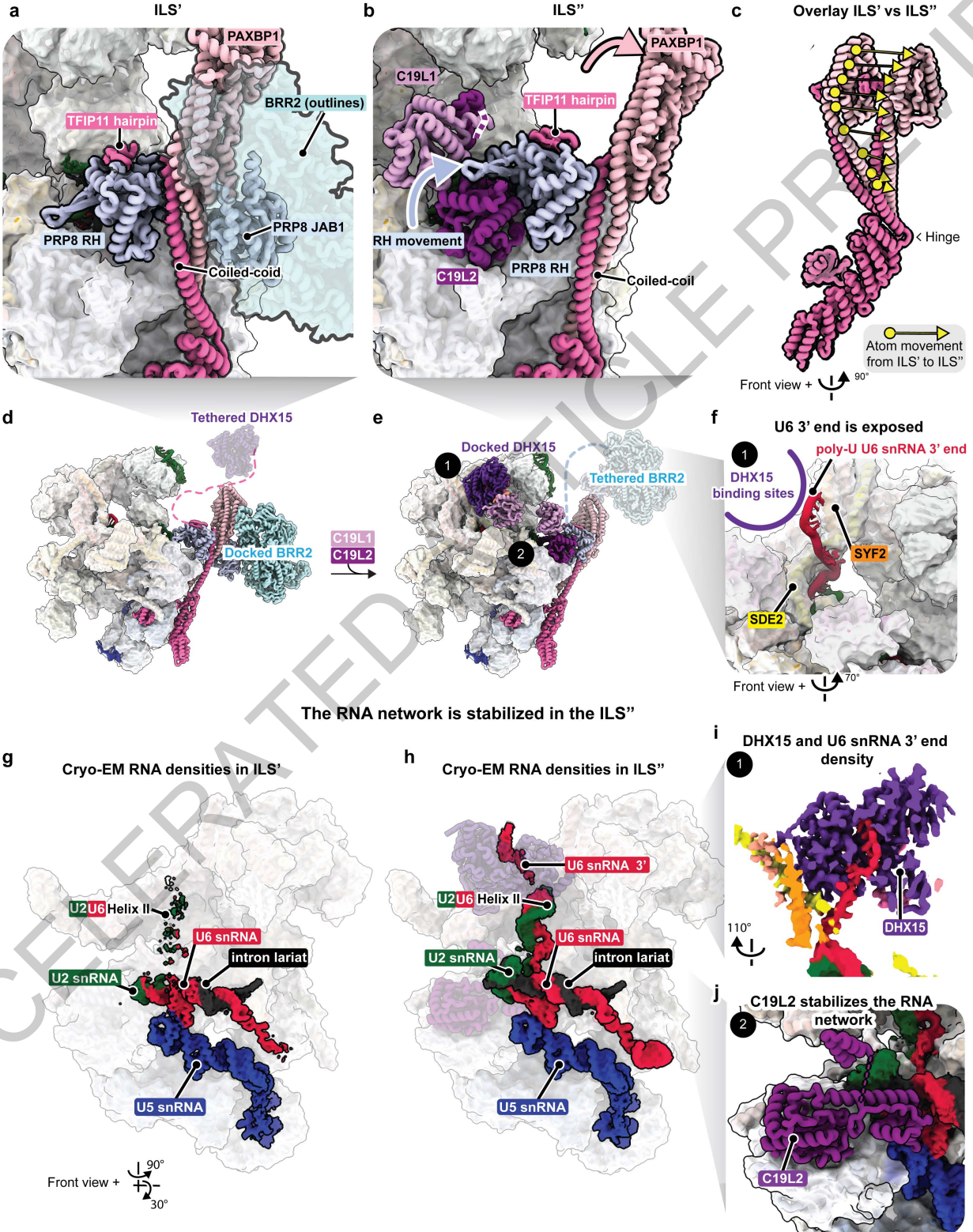
Bakers yeast Ntr1-Ntr2 show a divergent recognition of the ILS



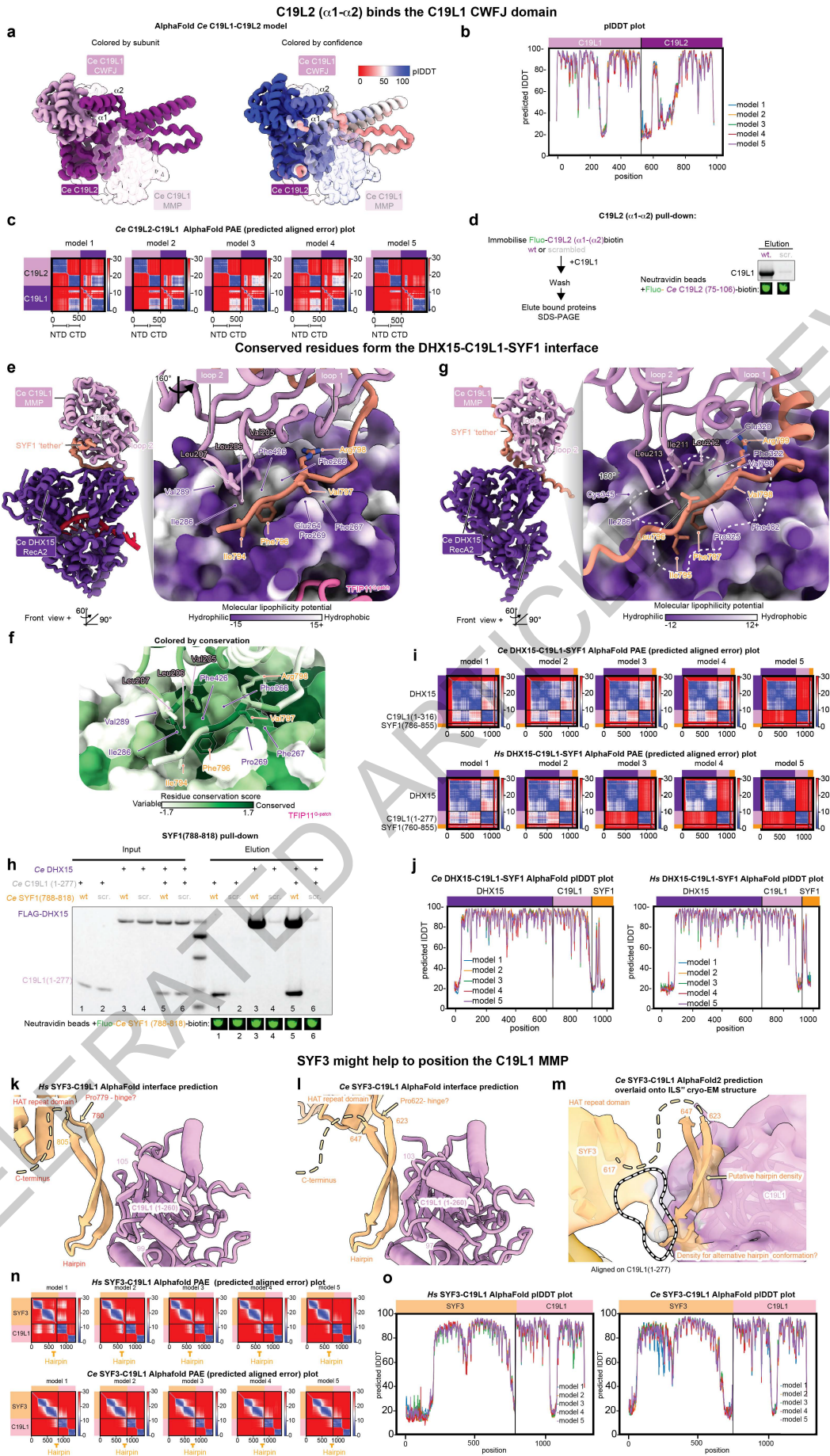
Extended Data Fig. 3



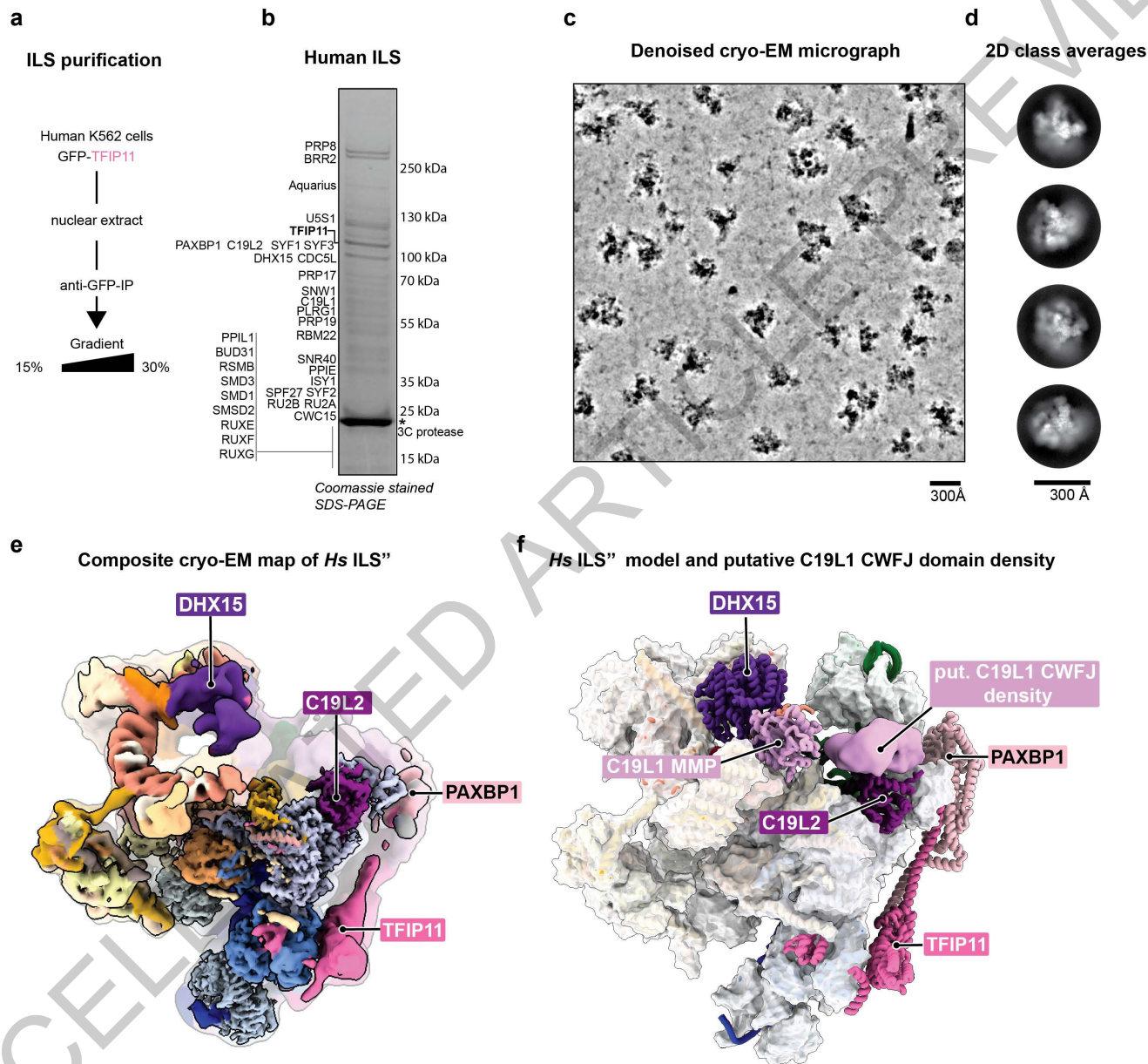
Conformational changes in the transition from ILS' to ILS''



Extended Data Fig. 4

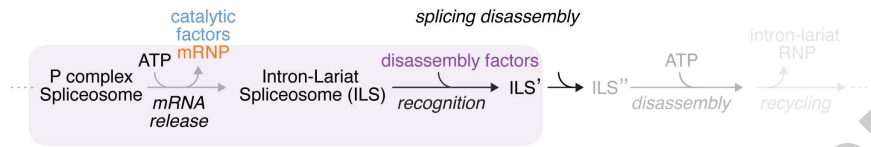


Extended Data Fig. 5



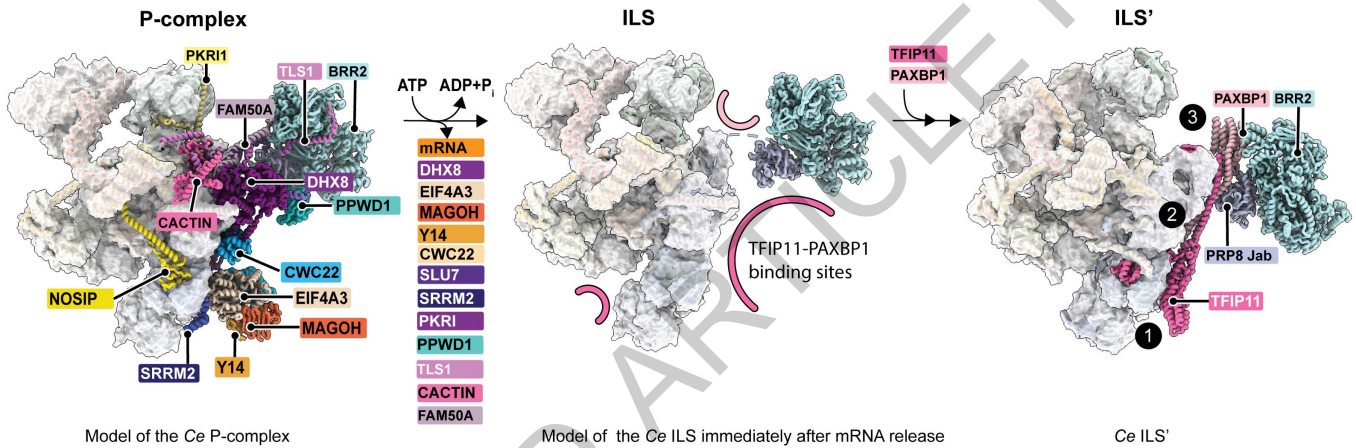
Extended Data Fig. 6

a



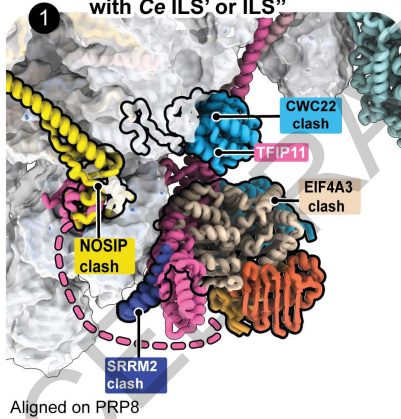
b

Model of the transition from the P-complex to the ILS'



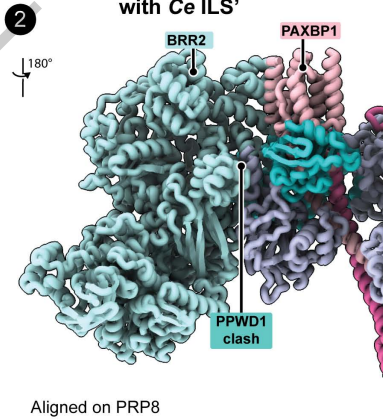
c

Clashes of *Hs* P-complex with Ce ILS' or ILS''



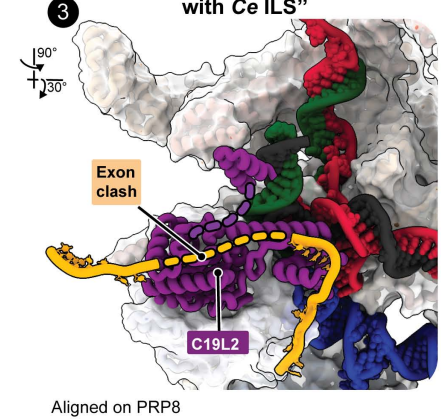
d

Clashes of *Hs* P-complex with Ce ILS'

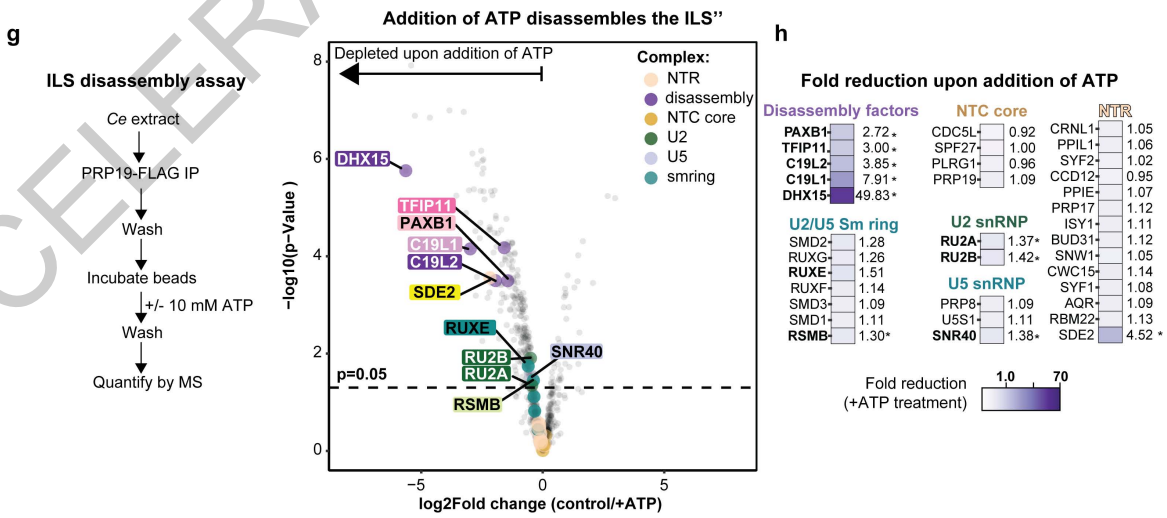
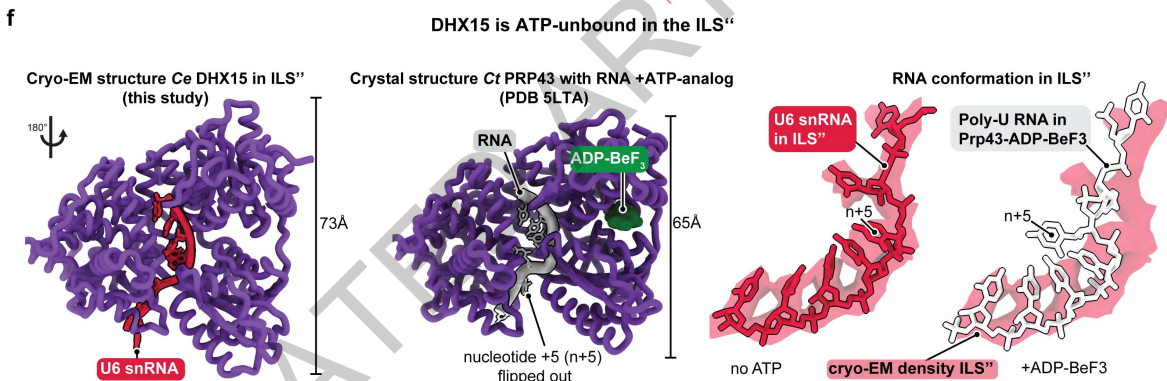
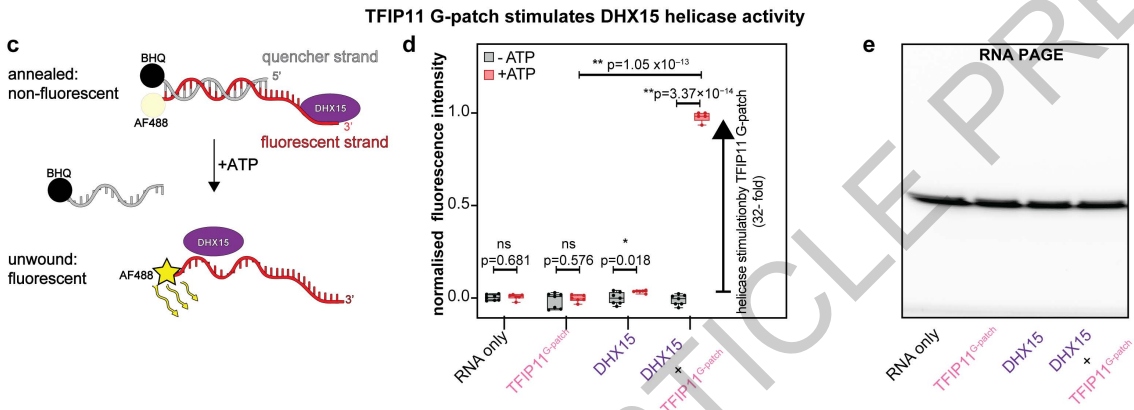
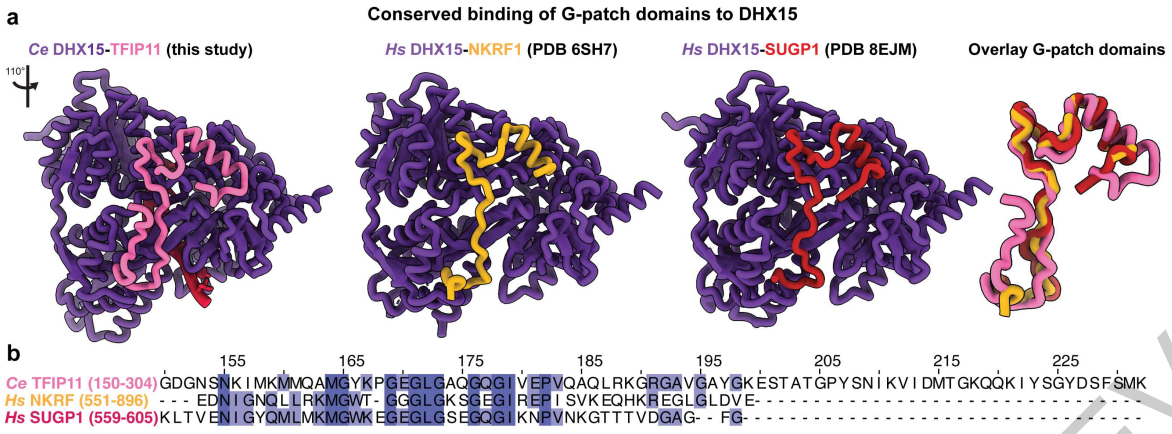


e

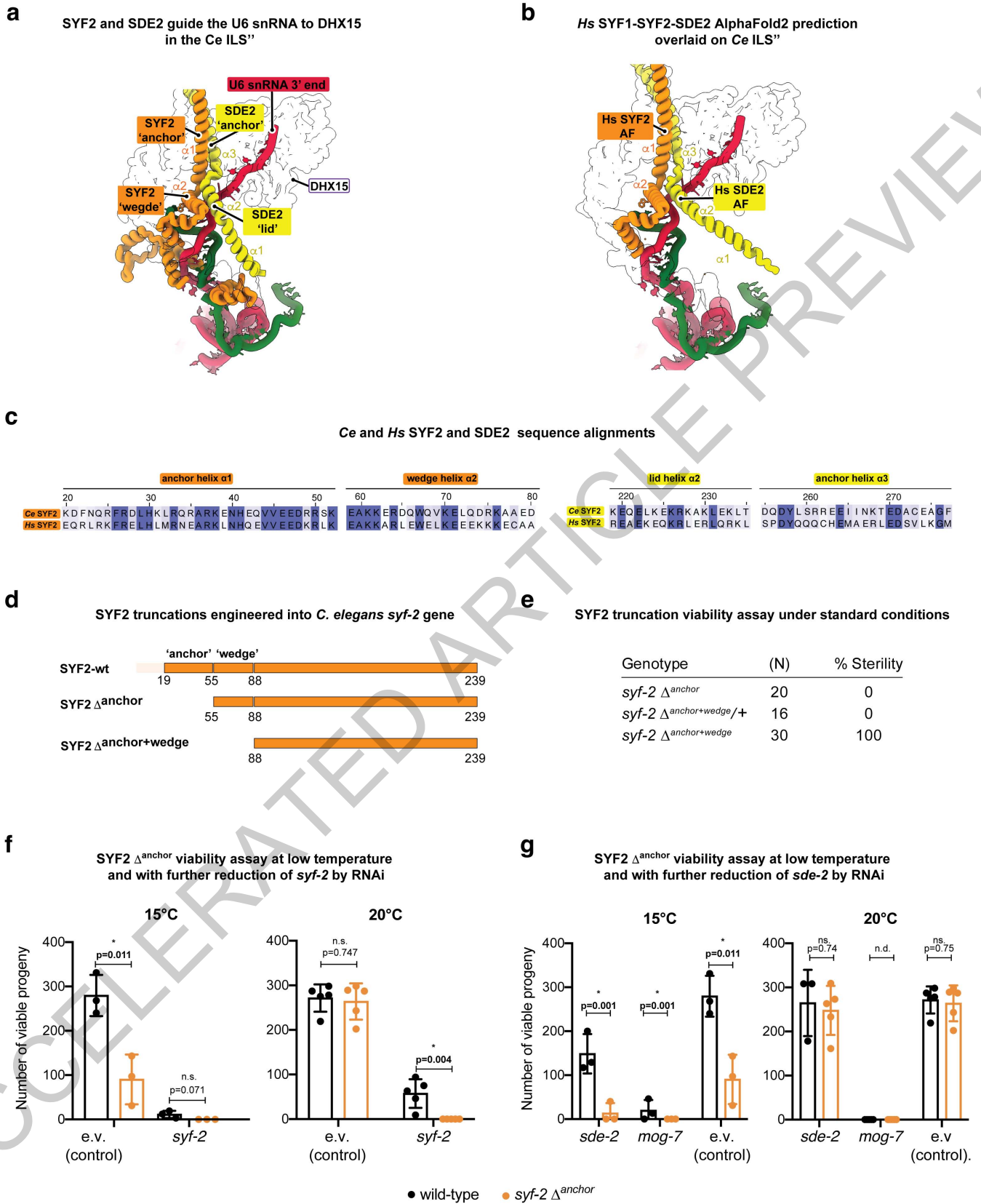
Clashes of *Hs* P-complex with Ce ILS''












































Extended Data Fig. 7



Extended Data Fig. 8



Extended Data Fig. 9

	Bakers yeast	Human	C. elegans	Color		Bakers yeast	Human	C. elegans	Color
U5 snRNP	Snu114	SNU114	eftu-2		NTC	Syf1	SYF1	syf-1	
	Prp8	PRPF8	prp-8			Syf2	SYF2	syf-2	
	Brr2	BRR2	snrp-200			Cif1	SYF3	syf-3	
	unknown	SNR40	snrp-40.1			Isy1	ISY1	isy-1	
U2 snRNP	Lea1	RU2A	mog-2		Cef1	CDC5L	cdc-5L		
	Msl1	RU2B	rnp-2/rnp-3		unknown	SDE2	sde-2		
U5/U2 smring	Sm D3	SMD3	snr-1		Prp19	PRPF19	prp-19		
	Sm B1	RSMB	snr-2		Snt309	SPF27	bcas-2		
	Sm D1	SMD1	snr-3		NTR	Ntc20	CCDC12	ccdc-12	
	Sm D2	SMD2	snr-4			Ecm2/Cwc2	RBM22	rbm-22	
	Smx3	RUXF	snr-5			Prp46	PLRG1	plrg-1	
	Sm E1	RUXE	snr-6			Cwc15	CWC15	cwc-15	
	Smx2	RUXG	snr-7			Bud31	BUD31	bud-31	
P Complex	Cwc22	CWC22	let-858			unknown	AQR	emb-4	
	eIF4A/Fal1	EIF4A3	eif-4A3			unknown	PPI1	cyn-12	
	unknown	RBM8A/Y14	rnp-4		Cdc40	CDC40	prp-17		
	Mnh1	MAGOH	mag-1		Prp45	SNW1	skp-1		
	unknown	SRRM2	rsr-2		Prp17	CDC40			
	unknown	PRKRIP1/PKRI1	F37A4.2		Dissassembly	Ntr1/SPP382	TFIP11	stip-1	
	unknown	FAM32A	KO1G5.8			Ntr2	PAXBP1	mog-7	
	unknown	CACTIN	cacn-1			Prp43	DHX15	ddx-15	
	Slu7	SLU7	sluh-7			Drn1	C19L1	cwf-19L1	
	Prp22	HRH1/DHX8	mog-5		unknown	C19L2	cwf-19L2		
CPR1/CPR3	PPWD1	cyn-15							

Extended Data Table 1

a

Cryo-EM map statistics

Cryo-EM maps (Ce ILS²)

Map number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
State	ceILS ²	ceILS	ceILS	ceILS	ceILS	ceILS	ceILS	ceILS ²	ceILS ²	ceILS ²	ceILS ²	ceILS ²	ceILS ²	ceILS ²	ceILS ²
Data collection															
Defocus range (µm)	-0.72 to -2.1	-0.72 to -2.2	-0.72 to -2.3	-0.72 to -2.4	-0.72 to -2.5	-0.72 to -2.6	-0.72 to -2.7	-0.72 to -2.8	-0.72 to -2.9	-0.72 to -2.10	-0.72 to -2.11	-0.72 to -2.12	-0.72 to -2.13	-0.72 to -2.14	-0.72 to -2.15
Voltage (kV)	300	300	300	300	300	300	300	300	300	300	300	300	300	300	300
Electron dose (e-/Å ²)	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60
Reconstruction (CryoSPARC)															
Resolution	2.9	3.29	2.99	3.14	5.51	3.31	6.35	3.09	3.69	2.73	3.03	4.77	6.4	4.24	3.61
Map-sharpening B-factor (Å ²)	79.9	102	92	75.1	537.7	70.5	559.5	66.3	111.1	68.9	67.6	225	656.4	131.2	23.8
Pixel size (Åpx)	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3
Number of Particles	879,523	1,131,513	1,131,513	1,131,513	599,066	349,802	60,258	879,523	879,523	879,523	879,523	148,075	148,075	148,075	78,663
Data deposition (EMDB)															
	EMD-50447	EMDB-50449	EMDB-50450	EMDB-50451	EMDB-50452	EMDB-50453	EMDB-50454	EMDB-50455	EMDB-50456	EMDB-50457	EMDB-50458	EMDB-50459	EMDB-50460	EMDB-50461	EMDB-50462

Cryo-EM maps (Ce ILS²), continued)

Map number	16	17	18	19	20	21	22	23	24	25	26	27
State	ceILS ²	ceILS ²	ceILS ²	ceILS ²	ceILS ²	ceILS ²	ceILS ²	ceILS ²	ceILS ²	ceILS ²	ceILS ²	ceILS ²
Data collection												
Defocus range (µm)	-0.72 to -2.1	-0.72 to -2.2	-0.72 to -2.3	-0.72 to -2.4	-0.72 to -2.5	-0.72 to -2.6	-0.72 to -2.7	-0.72 to -2.8	-0.72 to -2.9	-0.72 to -2.10	-0.72 to -2.11	-0.72 to -2.12
Voltage (kV)	300	300	300	300	300	300	300	300	300	300	300	300
Electron dose (e-/Å ²)	60	60	60	60	60	60	60	60	60	60	60	60
Reconstruction (CryoSPARC)												
Resolution	3.06	2.97	2.82	3.14	3.13	6.33	3.13	3.62	3.92	3.94	3.91	3.88
Map-sharpening B-factor (Å ²)	62.6	63.1	67.9	69.8	67.6	671.1	76.6	104.6	102.4	103.9	10.7	132.9
Pixel size (Åpx)	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3
Number of Particles	247,908	57,951	247,908	247,908	247,908	247,908	247,908	247,908	69,968	84,004	26,170	84,004
Data deposition												
	EMD-50463	EMD-50464	EMD-50465	EMD-50466	EMD-50467	EMD-50468	EMD-50469	EMD-50471	EMD-50472	EMD-50473	EMD-50475	EMD-50474

b

Cryo-EM maps (Hs ILS²)

Map number	1	2	3	4	5	6	7	8	9	10	11	12	13	14
State	hs ILS ²	hs ILS ²	hs ILS ²	hs ILS ²	hs ILS ²	hs ILS ²	hs ILS ²	hs ILS ²	hs ILS ²	hs ILS ²	hs ILS ²	hs ILS ²	hs ILS ²	hs ILS ²
Data collection														
Defocus range (µm)	0 to 3.57	0 to 3.57	0 to 3.57	0 to 3.57	0 to 3.57	0 to 3.57	0 to 3.57	0 to 3.57	0 to 3.57	0 to 3.57	0 to 3.57	0 to 3.57	0.01 to 1.78	0 to 3.57
Voltage (kV)	300	300	300	300	300	300	300	300	300	300	300	300	300	300
Electron dose (e-/Å ²)	50	50	50	50	50	50	50	50	50	50	50	50	50	50
Reconstruction (CryoSPARC)														
Resolution	3.41	3.22	3.13	3.39	3.23	3.51	3.14	3.24	6.08	7.32	5.7	4.78	8.06	5.79
Map-sharpening B-factor (Å ²)	12	49.3	63.1	68.4	56.3	56.7	58.4	50.7	299.7	817	266.1	37.9	566.8	274.1
Pixel size (Åpx)	0.945	0.945	0.945	0.945	0.945	0.945	0.945	0.945	0.945	0.945	0.945	0.945	0.945	0.945
Number of Particles	87951	87951	87951	87951	87951	87951	87951	87951	52490	87951	87951	14635	9321	26103
Data deposition														
	EMD-50477	EMD-50478	EMD-50479	EMD-50480	EMD-50481	EMD-50482	EMD-50483	EMD-50484	EMD-50485	EMD-50486	EMD-50487	EMD-50488	EMD-50489	EMD-50490

c

Coordinate model statistics

	Ce ILS ²	Ce ILS ²	Hs ILS ²	HsP
Model composition				
Protein residues	14323	14055	11764	16748
Nucleotide residues	308	323	277	365
Ligands	3	3	2	3
Refinement (PHENIX)				
Map CC (aroundatoms)	0.71	0.76	0.62	0.54
RMS deviations				
Bond lengths (Å)	0.007	0.007	0.007	0.010
Bond angles	1.448	1.417	1.458	1.457
Validation				
MolProbity score	1.42	1.29	1.38	1.91
All-atom clash score	1.67	1.63	1.29	4.37
Rotamer outliers (%)	1.8	1.36	1.75	2.97
C-beta deviations (%)	0.04	0.04	0.03	0.02
Ramachandran plot				
Outliers (%)	0.041	0.31	0.28	0.24
Allowed (%)	4.06	3.73	4.61	4.51
Favoured (%)	95.53	95.96	95.10	95.25
Data deposition				
PDB-ID	8R00	8R01	8R02	9FMD
EMDB deposition (composite map)	EMD-19397	EMD-19398	EMD-19399	ref. [37]

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection Cryo-EM data were collected with SerialEM 4 and EPU 3.

Data analysis Cryo-EM data were analysed with WARP 1, RELION 5, cryoSPARC 4, Coot 1, Phenix 1.2, and ISOLDE 1.6. Biochemical and genetic data was analyzed with GraphPad Prism 8.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Three-dimensional cryo-EM composite density maps of the Ce ILS' and ILS'' have been deposited in the Electron Microscopy Data Bank under the accession numbers EMD- 19397 to EMD- 19398. The individual maps 1-27 have been deposited under the accession numbers EMD-50447, EMD-50449 to EMD-50569, and EMD-50471 to EMD-40475. Three-dimensional cryo-EM composite density map of the human ILS'' have been deposited in the Electron Microscopy Data Bank under the accession numbers EMD-19399. The individual human ILS'' maps 1-14 have been deposited under the accession numbers EMD-50477 to EMD-50490. The coordinate files of the Ce ILS', Ce ILS'', the revised human P complex, and the human ILS'' have been deposited in the Protein Data Bank under the accession numbers 8ROO, 8RO1, 9FMD, and 8RO2.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Data set sample sizes were chosen to achieve the desired cryo-EM density map quality and resolution.
Data exclusions	No data were excluded.
Replication	Details on experiments replicates are indicated in each respective figure legends. The data analysis was highly reproducible.
Randomization	For 3D refinement the cryo-EM data were split randomly into two halves for gold-standard FSC determination. Biochemical experiments were not randomized, but contained appropriate controls.
Blinding	Blinding is not relevant for these types of experiments, due to rigorous experiment quality criteria and the experiment type.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

Methods

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	Leukemia cell line K562 (source: ATCC), and Lenti-X 293T (source: Takara)
Authentication	The K562 cell line was authenticated by short tandem repeat analysis (see Muhar et al. 2018, Science, DOI: 10.1126/science.aao2793).
Mycoplasma contamination	The cell line was confirmed to be negative for mycoplasma.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used in the study.