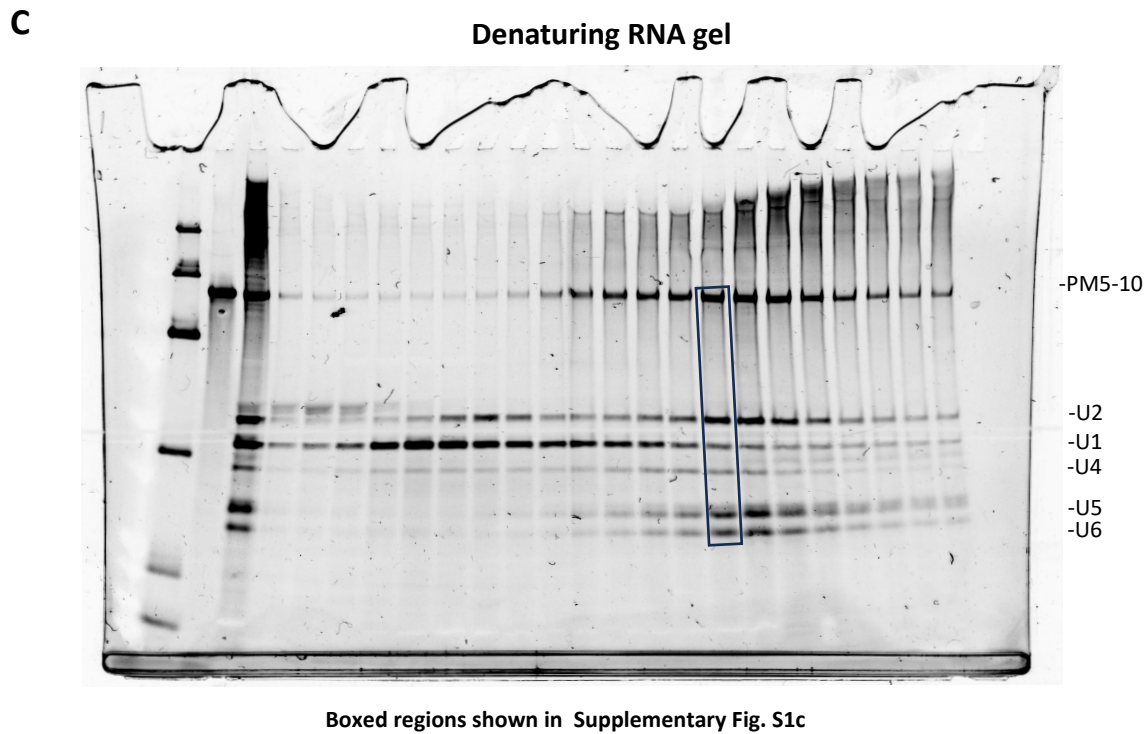
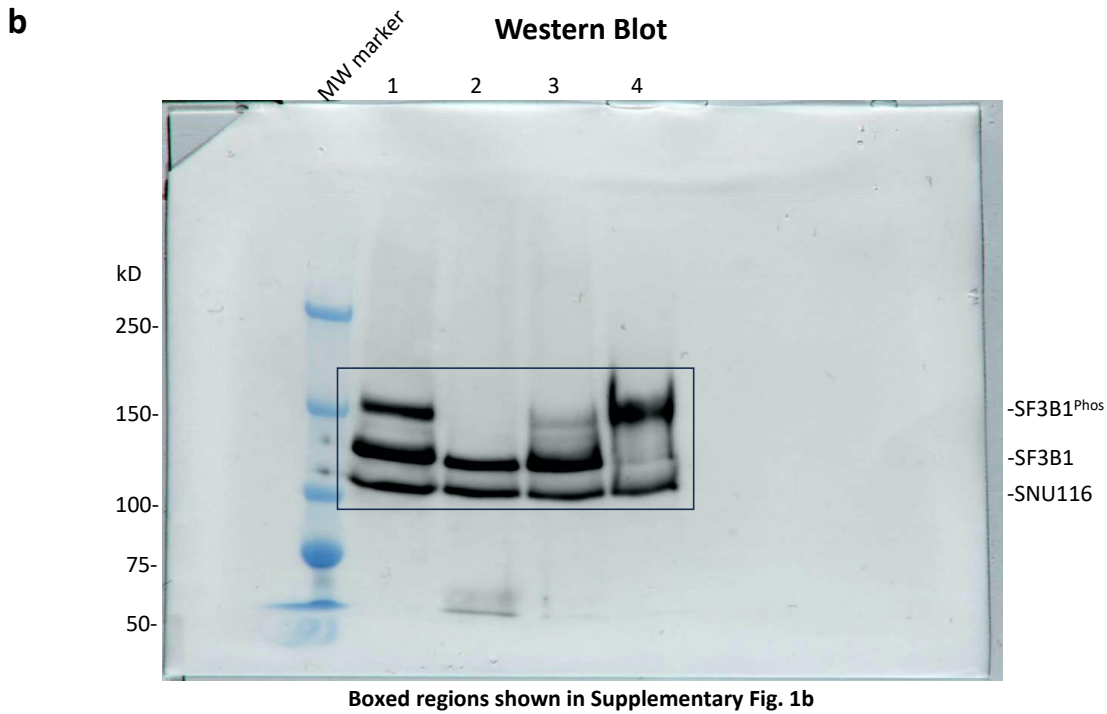
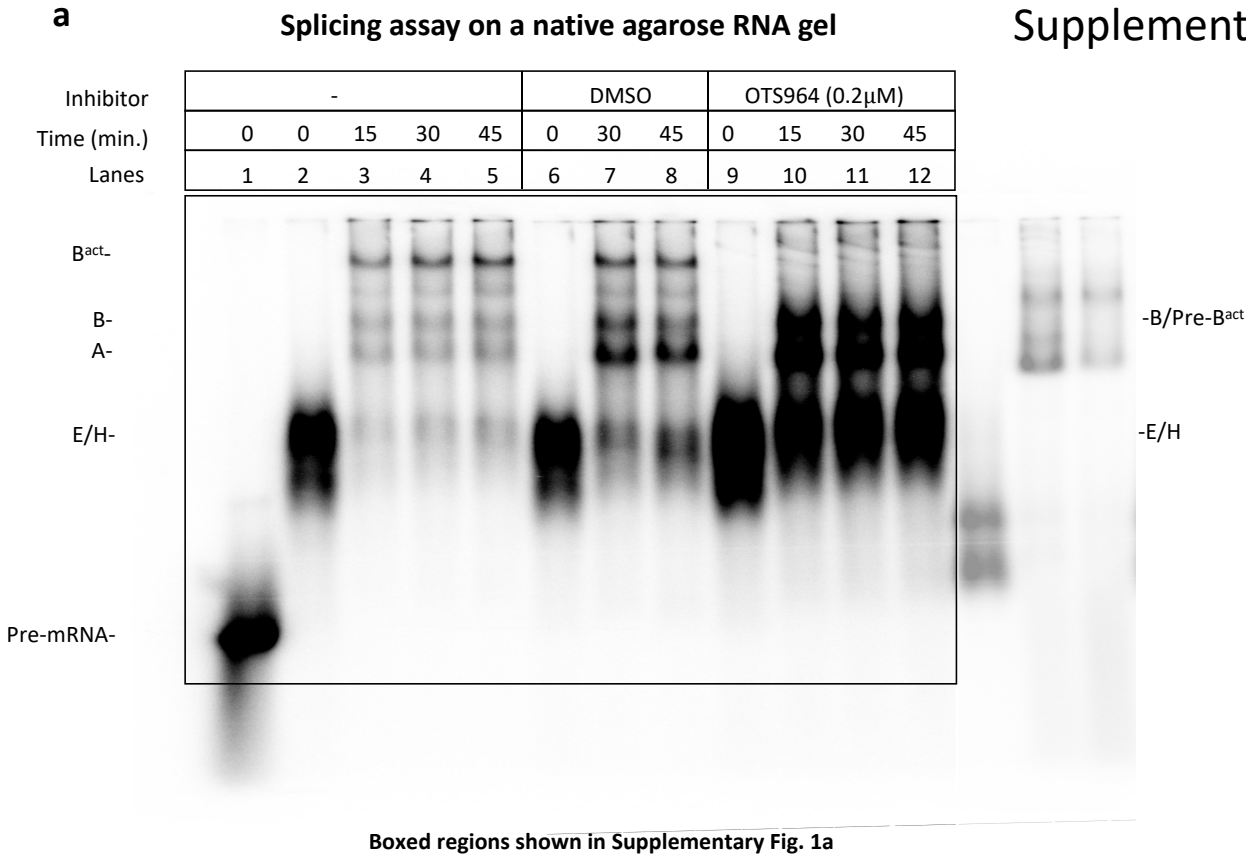
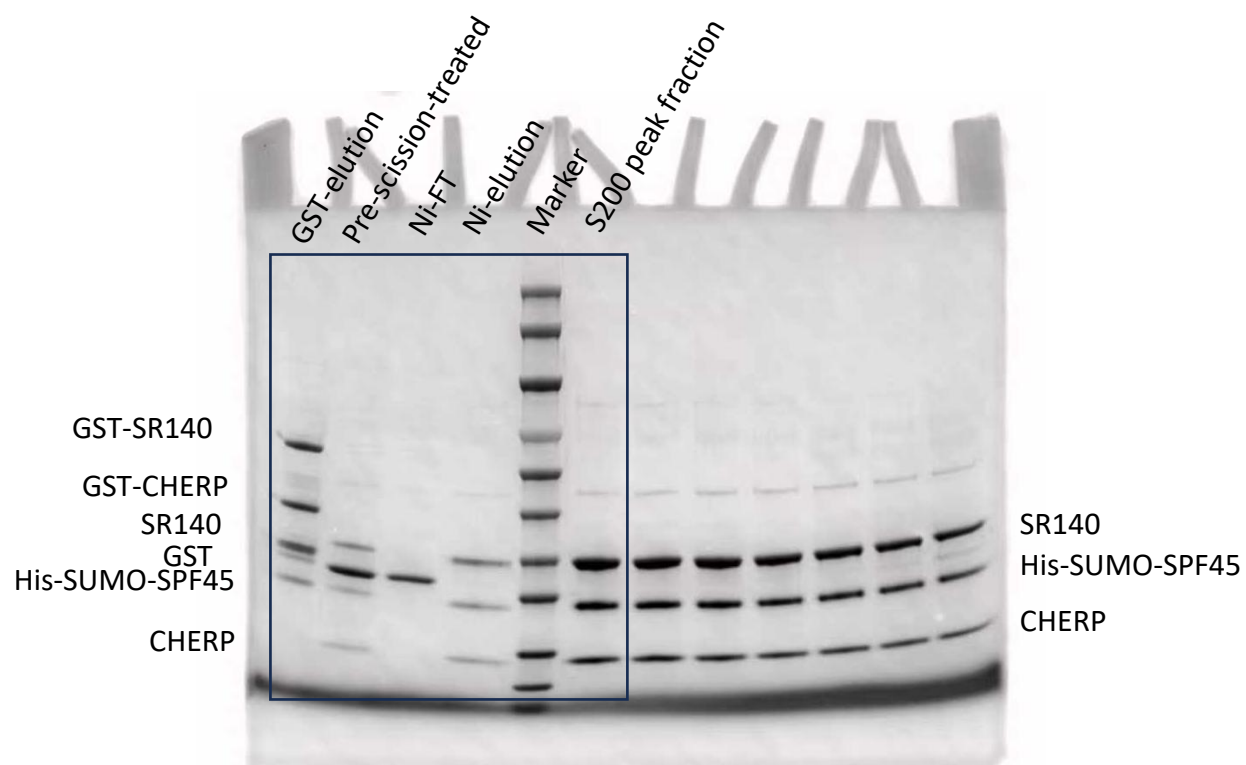
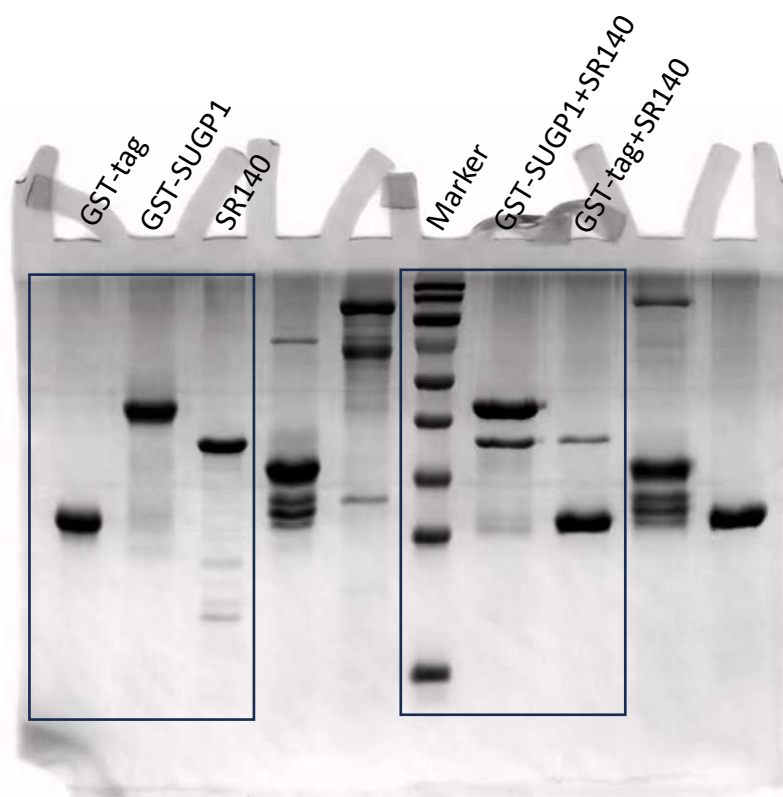






Boxed regions shown in Supplementary Fig. 9e



Boxed regions shown in Supplementary Fig. 10c

Reporting Summary

Nature Portfolio 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 Portfolio 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
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| ☒ | ☐ | 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
<i>Give P values as exact values whenever suitable.</i> |
| ☒ | ☐ | 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	Thermo Fischer EPU2.1, Thermo Fischer Foundation 3.1SP7 QF1, Thermo Scientific Orbitrap Exploris 480 4.2 362.16 and 3.0 261.13, Xcalibur v. 4.4.16.14
Data analysis	Max Quant v.2.4.13.0, pLink 3.0.11 and 2.3.11, MotionCor2, Gctf v.1.06, RELION 3.1, RELION 5, UCSF Chimera v.1.13.1, ChimeraX v.1.3, cryolo v.1.6.1, cryoSPARC v.2.1, COOT v.0.8.92, PHENIX v.1.2

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 Portfolio [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 description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The cryo-EM maps and coordinates have been deposited into the Electron Microscopy Data Bank (EMDB) and the PDB as follows: complete structure of the pre Bact OTS complex (EMD-53843: <https://www.ebi.ac.uk/emdb/EMD-53843>; PDB 9R8V: <https://www.rcsb.org/structure/9R8V>), core structure of the pre Bact OTS

complex (EMD-53554: <https://www.ebi.ac.uk/emdb/EMD-53554>; PDB 9R3D: <https://www.rcsb.org/structure/9R3D>). Source Data are provided with this paper. The B, pre-Bact-1, pre-Bact-2, Bact, and ILS' structures used in this study for structural comparisons are available under accession codes 8QO9, 7ABG, 7ABI, 6FF7, and 8RO1, respectively (<https://www.rcsb.org/structure/8RO9>, <https://www.rcsb.org/structure/7ABG>, <https://www.rcsb.org/structure/7ABI>, <https://www.rcsb.org/structure/6FF7>, <https://www.rcsb.org/structure/8RO1>). The AlphaFold3 predicted models of SPF45-SR140-CHERP, SR140-SUGP1, SR140-CTNNBL1, and SF3b core-SR140-DHX15 complexes are available in ModelArchive (modelarchive.org) with the accession codes ma-9luyy, ma-ire8e, ma-sd0ae, and ma-m2b6q, respectively (<https://www.modelarchive.org/doi/10.5452/ma-9luyy>, <https://www.modelarchive.org/doi/10.5452/ma-ire8e>, <https://www.modelarchive.org/doi/10.5452/ma-sd0ae>, and <https://www.modelarchive.org/doi/10.5452/ma-m2b6q>).

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender

Reporting on race, ethnicity, or other socially relevant groupings

Population characteristics

Recruitment

Ethics oversight

Note that full information on the approval of the study protocol must also be provided in the manuscript.

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 exclusions

Replication

Randomization

Blinding

Behavioural & social sciences study design

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

Study description

Research sample

Sampling strategy

Data collection

Data collection	<i>computer, eye tracker, video or audio equipment) whether anyone was present besides the participant(s) and the researcher, and whether the researcher was blind to experimental condition and/or the study hypothesis during data collection.</i>
Timing	<i>Indicate the start and stop dates of data collection. If there is a gap between collection periods, state the dates for each sample cohort.</i>
Data exclusions	<i>If no data were excluded from the analyses, state so OR if data were excluded, provide the exact number of exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.</i>
Non-participation	<i>State how many participants dropped out/declined participation and the reason(s) given OR provide response rate OR state that no participants dropped out/declined participation.</i>
Randomization	<i>If participants were not allocated into experimental groups, state so OR describe how participants were allocated to groups, and if allocation was not random, describe how covariates were controlled.</i>

Ecological, evolutionary & environmental sciences study design

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

Study description	<i>Briefly describe the study. For quantitative data include treatment factors and interactions, design structure (e.g. factorial, nested, hierarchical), nature and number of experimental units and replicates.</i>
Research sample	<i>Describe the research sample (e.g. a group of tagged <i>Passer domesticus</i>, all <i>Stenocereus thurberi</i> within Organ Pipe Cactus National Monument), and provide a rationale for the sample choice. When relevant, describe the organism taxa, source, sex, age range and any manipulations. State what population the sample is meant to represent when applicable. For studies involving existing datasets, describe the data and its source.</i>
Sampling strategy	<i>Note the sampling procedure. Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.</i>
Data collection	<i>Describe the data collection procedure, including who recorded the data and how.</i>
Timing and spatial scale	<i>Indicate the start and stop dates of data collection, noting the frequency and periodicity of sampling and providing a rationale for these choices. If there is a gap between collection periods, state the dates for each sample cohort. Specify the spatial scale from which the data are taken</i>
Data exclusions	<i>If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.</i>
Reproducibility	<i>Describe the measures taken to verify the reproducibility of experimental findings. For each experiment, note whether any attempts to repeat the experiment failed OR state that all attempts to repeat the experiment were successful.</i>
Randomization	<i>Describe how samples/organisms/participants were allocated into groups. If allocation was not random, describe how covariates were controlled. If this is not relevant to your study, explain why.</i>
Blinding	<i>Describe the extent of blinding used during data acquisition and analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.</i>

Did the study involve field work? ☐ Yes ☐ No

Field work, collection and transport

Field conditions	<i>Describe the study conditions for field work, providing relevant parameters (e.g. temperature, rainfall).</i>
Location	<i>State the location of the sampling or experiment, providing relevant parameters (e.g. latitude and longitude, elevation, water depth).</i>
Access & import/export	<i>Describe the efforts you have made to access habitats and to collect and import/export your samples in a responsible manner and in compliance with local, national and international laws, noting any permits that were obtained (give the name of the issuing authority, the date of issue, and any identifying information).</i>
Disturbance	<i>Describe any disturbance caused by the study and how it was minimized.</i>

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	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Antibodies against human SF3B1 and SNU114 were used. SF3B1 antibody is from Proteintech (Cat No. 27684-1-AP); SNU114 antibody is from Fabrizio et al., 1997. The SF3B1 and SNU114 antibodies were used for Western blotting in 9Supplementary Fig 1.
Validation	SF3B1 was purchased commercially from: https://www.ptglab.com/products/SF3B1-Antibody-27684-1-AP.htm?srsId=AfmBOorzpe-7utxFR-tB7UKS9yxP7NqE6zEYoFuKM-J3VpUbLfBKS8k , and was used in Bessonov et al., 2010. SNU114 antibody was produced in our lab and was used previously in Bessonov et al., 2010.

Eukaryotic cell lines

Policy information about [cell lines and Sex and Gender in Research](#)

Cell line source(s)	HeLa S3 cells were obtained from the Helmholtz Zentrum für Infektionsforschung, Braunschweig.
Authentication	Cells were not authenticated
Mycoplasma contamination	HeLa S3 cells tested negative for mycoplasma.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified lines were used.

Palaeontology and Archaeology

Specimen provenance	<i>Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information). Permits should encompass collection and, where applicable, export.</i>
Specimen deposition	<i>Indicate where the specimens have been deposited to permit free access by other researchers.</i>
Dating methods	<i>If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.</i>
☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.	
Ethics oversight	<i>Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.</i>

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Animals and other research organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research, and [Sex and Gender in Research](#)

Laboratory animals	<i>For laboratory animals, report species, strain and age OR state that the study did not involve laboratory animals.</i>
Wild animals	<i>Provide details on animals observed in or captured in the field; report species and age where possible. Describe how animals were caught and transported and what happened to captive animals after the study (if killed, explain why and describe method; if released, say where and when) OR state that the study did not involve wild animals.</i>
Reporting on sex	<i>Indicate if findings apply to only one sex; describe whether sex was considered in study design, methods used for assigning sex. Provide data disaggregated for sex where this information has been collected in the source data as appropriate; provide overall</i>

numbers in this Reporting Summary. Please state if this information has not been collected. Report sex-based analyses where performed, justify reasons for lack of sex-based analysis.

Field-collected samples

For laboratory work with field-collected samples, describe all relevant parameters such as housing, maintenance, temperature, photoperiod and end-of-experiment protocol OR state that the study did not involve samples collected from the field.

Ethics oversight

Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about [clinical studies](#)

All manuscripts should comply with the ICMJE [guidelines for publication of clinical research](#) and a completed [CONSORT checklist](#) must be included with all submissions.

Clinical trial registration

Provide the trial registration number from ClinicalTrials.gov or an equivalent agency.

Study protocol

Note where the full trial protocol can be accessed OR if not available, explain why.

Data collection

Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.

Outcomes

Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures.

Dual use research of concern

Policy information about [dual use research of concern](#)

Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

No	Yes
☐	☐ Public health
☐	☐ National security
☐	☐ Crops and/or livestock
☐	☐ Ecosystems
☐	☐ Any other significant area

Experiments of concern

Does the work involve any of these experiments of concern:

No	Yes
☐	☐ Demonstrate how to render a vaccine ineffective
☐	☐ Confer resistance to therapeutically useful antibiotics or antiviral agents
☐	☐ Enhance the virulence of a pathogen or render a nonpathogen virulent
☐	☐ Increase transmissibility of a pathogen
☐	☐ Alter the host range of a pathogen
☐	☐ Enable evasion of diagnostic/detection modalities
☐	☐ Enable the weaponization of a biological agent or toxin
☐	☐ Any other potentially harmful combination of experiments and agents

Plants

Seed stocks	Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.
Novel plant genotypes	Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.
Authentication	Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosaicism, off-target gene editing) were examined.

ChIP-seq

Data deposition

- ☐ Confirm that both raw and final processed data have been deposited in a public database such as [GEO](#).
- ☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links <i>May remain private before publication.</i>	For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data.
Files in database submission	Provide a list of all files available in the database submission.
Genome browser session (e.g. UCSC)	Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.

Methodology

Replicates	Describe the experimental replicates, specifying number, type and replicate agreement.
Sequencing depth	Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and whether they were paired- or single-end.
Antibodies	Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and lot number.
Peak calling parameters	Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.
Data quality	Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.
Software	Describe the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community repository, provide accession details.

Flow Cytometry

Plots

Confirm that:

- ☐ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- ☐ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- ☐ All plots are contour plots with outliers or pseudocolor plots.
- ☐ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Describe the sample preparation, detailing the biological source of the cells and any tissue processing steps used.
Instrument	Identify the instrument used for data collection, specifying make and model number.
Software	Describe the software used to collect and analyze the flow cytometry data. For custom code that has been deposited into a community repository, provide accession details.

Cell population abundance

Describe the abundance of the relevant cell populations within post-sort fractions, providing details on the purity of the samples and how it was determined.

Gating strategy

Describe the gating strategy used for all relevant experiments, specifying the preliminary FSC/SSC gates of the starting cell population, indicating where boundaries between "positive" and "negative" staining cell populations are defined.

☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Magnetic resonance imaging

Experimental design

Design type

Indicate task or resting state; event-related or block design.

Design specifications

Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.

Behavioral performance measures

State number and/or type of variables recorded (e.g. correct button press, response time) and what statistics were used to establish that the subjects were performing the task as expected (e.g. mean, range, and/or standard deviation across subjects).

Acquisition

Imaging type(s)

Specify: functional, structural, diffusion, perfusion.

Field strength

Specify in Tesla

Sequence & imaging parameters

Specify the pulse sequence type (gradient echo, spin echo, etc.), imaging type (EPI, spiral, etc.), field of view, matrix size, slice thickness, orientation and TE/TR/flip angle.

Area of acquisition

State whether a whole brain scan was used OR define the area of acquisition, describing how the region was determined.

Diffusion MRI

☐ Used

☐ Not used

Preprocessing

Preprocessing software

Provide detail on software version and revision number and on specific parameters (model/functions, brain extraction, segmentation, smoothing kernel size, etc.).

Normalization

If data were normalized/standardized, describe the approach(es): specify linear or non-linear and define image types used for transformation OR indicate that data were not normalized and explain rationale for lack of normalization.

Normalization template

Describe the template used for normalization/transformation, specifying subject space or group standardized space (e.g. original Talairach, MNI305, ICBM152) OR indicate that the data were not normalized.

Noise and artifact removal

Describe your procedure(s) for artifact and structured noise removal, specifying motion parameters, tissue signals and physiological signals (heart rate, respiration).

Volume censoring

Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.

Statistical modeling & inference

Model type and settings

Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation).

Effect(s) tested

Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used.

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference

Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.

(See [Eklund et al. 2016](#))

Correction

Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo).

Models & analysis

n/a	Involvement in the study
☐	☐ Functional and/or effective connectivity
☐	☐ Graph analysis
☐	☐ Multivariate modeling or predictive analysis
Functional and/or effective connectivity	Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).
Graph analysis	Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).
Multivariate modeling and predictive analysis	Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.

Structural basis of the regulation by CDK11 kinase of early spliceosome activation and evidence for its proofreading by DHX15 helicase

Corresponding Author: Professor Reinhard Lührmann

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

Zhang, Kumar et al combine structural cryo-EM and biochemical crosslinking data to provide a detailed description of the molecular arrangements of human spliceosomes assembled on a model pre-RNA substrate upon inhibition of kinase CDK11 by the small molecule OTS964. Combined with previous pioneering structural work by the Lührmann & Stark teams and other groups, the results provide important new insights into one key step in the activation of the spliceosome, and into the molecular mechanisms that proofread the proper conformational changes that ensue the stepwise assembly of splicing complexes on splice site regions. The paper not only provides a snapshot of an important transition step but also novel models of general value to understand conformational triggers and mechanisms of proofreading.

Specifically, the structure of the OTS964-inhibited spliceosomal complex provides unexpected insights on how phosphorylation of a splicing factor (SF3B1 N-terminal domain) can 1) serve as a signal to trigger complex conformational transitions (including a striking relocation of SF3B6 required for progression to complex Bact formation, concomitant with major rearrangements in RNA:RNA, RNA:protein and protein:protein contacts) revealing novel driving forces in the process, and 2) compete with the association of proofreading helicase DHX15 and several co-factors, including G-patch-containing proteins and the scaffolding factor SR140, providing a novel model for proofreading activities by helicase complexes.

Collectively, the new structure illustrates the surprising complexity of what previously considered to be a single step for the activation of the fully assembled spliceosome, revealing key roles for post-translational modifications and their interplay with proofreading complexes that had not been implicated at this stage in the process. On this basis I am happy to support publication of the manuscript in Nature Communications. The study will certainly fuel new thinking and work on this and other steps of the complex process of spliceosome assembly and activation, as well as more generally on proofreading of ribonucleoprotein processes.

A potential issue with the manuscript is the high density of the data and information provided, as well as the abundance of sophisticated structural arguments made, which may be overwhelming for non-specialists. While the Discussion is an example of clarity, highlighting the key points of the paper in a clear and enticing manner, the Results section is indeed very dense and at points difficult to follow. In my opinion the following revisions would help to improve the text:

- 1) Summary: understanding some of the paper highlights requires background information that is not introduced earlier on in the summary, which may make it difficult to follow, specially to gather their general significance, for non-specialists. Examples include the sentences "This explains the stalled position...were not first to relocate" and "We also located the binding of the RNA helicase DHX15...and G-patch proteins". Non-specialist readers may wonder about the significance of "passing along the PRP8RT domain", or the significance of G-patch proteins for the recruitment of an RNA helicase.
- 2) Introduction: along the same lines, it may be helpful for the general reader to explain concepts such as "catalytic U6 snRNA triplex" (line 63) and clarify the difference in activities between upstream helicases that "remodel off-pathway" (line 98) and DHX15, which "subsequently dismantles" these complexes (line 99).
- 3) Line 127: it would be helpful for the general reader to know the rationale for the choice of the PM5-10 pre-mRNA (the word "construct" may not be necessary here).
- 4) Line 128: it may be useful to state in the main text the concentration used of OTS964 and its IC50.

- 5) Line 132: it will be helpful to indicate the identity of the tag used (MS2?) for affinity chromatography of the complexes.
- 6) Line 166: "5' domain of U2 snRNP" may require clarification for non-specialists the first time the term is used, e.g. proteins associated with the 5' x nucleotides of U2 snRNA.
- 7) Line 207: it may not be obvious why the extended U6 ISL "elicits" (rather than being a consequence of other conformational changes) a large-scale downward movement of the 5' domain of U2 snRNP, i.e. why it is causative of the movement.
- 8) While the extensive Extended Data Figures, Supplementary Tables and specially the supporting Movies are very helpful to get the message across, entire sections (e.g. those covering the "Coordination by SMU1/RED..." and "Location of NTC/NTR...") are illustrated through Supplementary Materials which the average reader may not be inclined to use extensively. I would recommend the authors to consider including a key summary Figure of each of these in the main set of Figures.
- 9) Line 294: RRM protein SF3B6: it may be helpful for the general reader to introduce the concept of the RRM fold.
- 10) Lines 312-313, lines 342-344, section starting in line 346 and Discussion: while the collective evidence supports a key role of CDK11-dependent phosphorylation of SF3B1N as the major driving force for triggering the repositioning of SF3B6, is there a clear model for how this occurs, i.e. what key molecular contacts between SF3B6 and HRs 12-14 would be broken (and/or contacts with RES protein SNIP1 facilitated) by the access of CDK11 complex / phosphorylation of SF3B1 T426/434/436?
- 11) Lines 442-450: the crosslinking data suggest a direct interaction between SR140 and SUGP1. However the groups of Manley, Rabadan and Liu have reported a direct interaction of SUGP1 with SF3B1 as an alternative way to recruit SUGP1 (and consequently DHX15), with the appealing added value of potentially explaining the mechanism by which cancer-associated mutations in SF3B1, by disrupting these interactions, cause cryptic splice site utilization: PMID: 31474574, PMID: 32332164, PMID: 40027711, PMID: 40714635, PMID: 36459648, PMID: 37977822. The authors could consider to extend the discussion of these findings.
- 12) Somewhat related to the previous point, recent work by Damianov and Black (both reference 48 as well as PMID: 40138349 -again potentially explaining the effects of cancer mutations in SF3B1-) indicates that association of SR140 (along with CHERP and RBM17/SPF45) occurs at early stages of branch point recognition on chromatin-associated nascent transcripts. Is this mode of DHX15 recruitment additional / overlapping / complementary to the model presented by the authors in the context of SF3B1N phosphorylation / SF3B6 translocation?
- 13) The authors could consider moving the final paragraph of Results (lines 480 – 491) to the Discussion section.
- 14) Lines 552 – 554: the absence of SF3B6 and of a conserved Nt region in the yeast SF3B1 counterpart are indeed striking and probably merit an extended discussion: how do yeast spliceosomes transition between B and Bact complexes, then? E.g. does CDK11-mediated phosphorylation of Hsh155 occur in *S. cerevisiae*? Does OTS964 inhibit spliceosome assembly in budding yeast?

Reviewer #2

(Remarks to the Author)

The manuscript by Zhang et al. uses OTS964 and a PM5-10 pre-mRNA construct to determine, by cryo-EM and mass spectrometry, the structure and composition of spliceosomes blocked by inhibition of CDK11, a kinase that phosphorylates the SF3B1 N-terminus and is essential for spliceosome activation. The authors identify a spliceosome intermediate termed PM5-10 pre-Bact-OTS, which proceeds a previously characterized intermediate, pre-Bact-1, identified by the same group (PMID: 33243851). The pre-Bact-OTS complex is characterized by dissociation of the U4 snRNP, the presence of NTC, NTR, and IBC proteins, and the formation of a previously unobserved step in the establishment of the U2/U6 snRNA catalytic core. The authors further identify roles for the SMU1/RED complex in early stages of spliceosome remodeling and suggest a crucial role for CDK11-mediated phosphorylation of the SF3B1 N-terminus in the transition of what they term the SF3B6/SF3B1(394–430) module from SF3B1 HEAT repeats 12–14 to repeats 1–2. To characterize the roles of phosphorylated SF3B1, they use crosslinking approaches and identify three threonine residues involved in SF3B6/SF3B1(394–430) module transition and SNIP1-mediated recruitment of the RES complex to the SF3B1 HEAT repeats 1–2. Finally, using their PM5-10 pre-Bact-OTS structures, crosslinking data, and AlphaFold3 modeling, the authors determine the location and propose functions of the DHX15 helicase and three G-patch-associated proteins (SUGP1, CHERP, and SPF45) in quality control during early stages of CDK11-dependent spliceosome activation. Overall, the manuscript provides new mechanistic insights into spliceosome activation, is well written and timely, and will be of interest to the readership of Nature Communications. However, before publication, I have the following questions, comments, and suggestions that need to be addressed:

1. Page 3, lines 70–77: This recent paper describing several structures of activated spliceosomes (PMID: 39068178) should also be cited, and these structures should be compared in the Discussion with those observed following CDK11 inhibition.
2. Page 3, line 84: The authors should additionally cite a recent paper characterizing SF3B1 phosphorylation (PMID: 41428738) and at page 4, line 88, this paper (PMID: 31511426) should be also cited.
3. Figure 1a and Extended Data Figure 1a, c: The PM5-10 pre-mRNA construct was used, but this substrate contains a very short intron (and a short distance between the 5' splice site and the branch site). Do the authors observe identical results when other splicing substrates with longer introns, such as MINX or AdML, are used? In addition, controls performed in the absence or at low concentrations of ATP and under low Mg^{2+} conditions should be included to allow a more rigorous comparison between B and Bact-OTS complexes.
4. Page 5, lines 132–147: The composition of OTS964-stalled complexes assembled on PM5-10 pre-mRNA differs from chromatin-associated spliceosome complexes in cells after OTS964 treatment (PMID: 36104565), specifically with respect to the absence of U4 snRNA and the presence of the IBC complex. There may be several reasons for these differences, and the authors should at least comment on them (see also points 3, 5, and 6).
5. SAD1 and SNU66 inhibit BRR2 function but are found as part of the Bact-OTS spliceosome (see Supplementary Table 1,

where SAD1 and SNU66 are abundant in Bact-OTS). Do the authors expect BRR2 to be active and capable of unwinding U4 snRNA when SAD1 and SNU66 are present? In addition, neither of these proteins is shown in any of the structures depicting the Bact-OTS complex.

6. TXNL4A (DIM1) and FBP21 (WBP4) are identified in the crosslinking dataset. The α -finger of PRPF8 interacts with PRP31 and DIM1 to stabilize U4/U6 stem I. Therefore, the presence of DIM1 is also inconsistent with U4 snRNA having been unwound in the Bact-OTS complex.

7. SMU1/RED were previously shown to be required for efficient splicing of short introns (PMID: 31409787). Are the newly identified roles of this complex described on page 9 restricted to spliceosomes assembled on short introns? See also point 3.

8. Page 19, line 520: "HR 1–2" should be used instead of "HR 1–3".

9. In the Discussion (page 19, line 529), the statement should omit the term "phosphorylated," as SF3B1 is not phosphorylated in the pre-Bact-OTS complex, but rather in the pre-Bact and Bact complexes. Similarly, in the Extended Data Figure 6e T426, T434 and T436 are labeled by the same color irrespectively of their phosphorylation status at different steps of the spliceosome activation (Bact-OTS, Bact-1/2 and Bact) which is misleading. This needs to be corrected and clarified. It is also not clear why amino acids 67 (SF3B6) and 394 (SF3B1) are highlighted in the Extended Data Figure 6e, please explain.

10. Page 5, line 118 and elsewhere: The statement "Aberrant spliceosomes would impair CDK11 recruitment (and thus SF3B1 phosphorylation)..." contradicts the data presented in Data S1, where CDK11-CCNL-SAP30BP are clearly crosslinked to the spliceosome in the pre-Bact-OTS complex. This indicates that aberrant spliceosomes do not impair CDK11 recruitment; rather, the recruited kinase is inhibited by OTS964 in the pre-Bact-OTS complex. In a similar vein, while SAP30BP may play a role in recruiting the CDK11 complex to certain spliceosome intermediates (as discussed on page 22, line 616), the mechanism proposed in reference 52 is difficult to reconcile with the current findings, since SAP30BP recruitment in that study depends on SF3B1 phosphorylation, which itself depends on CDK11 activity. In this context, it would also be important to discuss the role of SPF45, which has been shown in several studies to be recruited to unphosphorylated SF3B1 in earlier spliceosome intermediates and to function predominantly in the splicing of short introns or in alternative splicing regulation (PMID: 17589525, PMID: 34389706, PMID: 38065098). Once again, this raises the question of the extent to which the spliceosome structures described in this manuscript are specific to short introns due to the use of PM5-10 pre-mRNA, and to what extent the conclusions can be generalized, given that OTS964 blocks constitutive splicing in cells (PMID: 36104565, PMID: 38059508). These potential limitations should be discussed.

Reviewer #3

(Remarks to the Author)

Zhang et al. present a structural and functional analysis of a novel early-activated spliceosome intermediate, termed pre-Bact-OTS complex, stalled by the CDK11 inhibitor OTS964. By integrating single-particle cryo-EM, crosslinking mass spectrometry, biochemical assays, and AlphaFold3 modeling, the authors elucidate the mechanism through which CDK11-mediated phosphorylation of SF3B1 regulates spliceosome activation and reveal how the helicase DHX15 may proofread this process. While post-translational modifications (PTMs) such as SF3B1 phosphorylation have been recognized as important regulatory events in splicing, the precise structural and functional consequences of this phosphorylation have remained unclear. This study demonstrates that phosphorylation of this core splicing factor triggers a critical structural rearrangement, thereby establishing a checkpoint for the proofreading of spliceosome assembly. Importantly, the work identifies the recruitment of DHX15 and its associated G-patch proteins to this intermediate, with SR140 serving as a key binding hub. Although DHX15 is known for its role in dismantling the ILS complex at the end of the splicing cycle and has recently been implicated in proofreading early assembly stages, this study significantly expands its functional repertoire, positioning DHX15 as a general proofreading factor active at multiple points in the splicing cycle.

Overall, this work provides crucial insights into how PTMs regulate the splicing cycle and how the spliceosome in higher eukaryotes may take advantage of such modifications as checkpoints to ensure fidelity. The implications of this research could extend beyond the splicing field, offering valuable perspectives also for other RNP biogenesis processes. Thus, I am happy to support publication of the manuscript in Nature Communications.

A few comments:

1. The Results section is generally dense and may be challenging for readers not deeply familiar with spliceosome structural biology. Given the broad readership of this journal, providing additional context and simplifying some of the detailed descriptions would greatly improve accessibility and broaden the impact of the work.

2. The discussion on the absence of this kinase-switch system in *S. cerevisiae* is interesting and highlights evolutionary complexity in humans/higher eukaryotes. It raises the question of whether this mechanism is conserved in other organisms such as *S. pombe* or *C. elegans*, which possess an SF3B6 counterpart. A brief bioinformatic analysis commenting on the conservation of SF3B6, the N-terminal region of SF3B1, and CDK11 across species would be a valuable addition.

3. The mechanism of CDK11 complex recruitment to the spliceosome remains unclear. Specifically, why is CDK11 recruited specifically to this intermediate, and not to others where the SF3B1 N-terminus might also be accessible? Expanding the discussion on this point would be beneficial.

4. The Results section includes a segment on the NTC/NTR complex and notes that IBC proteins are loosely associated. The implications of these observations are not discussed later. A brief comment on the potential significance of these

associations in the Discussion would be helpful.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

The authors have adequately addressed the issues raised in my previous report and I am happy to recommend publication of the paper, which I expect to have an important impact in the field of spliceosome function and, more generally, the function of complex ribonucleoprotein machines.

Reviewer #2

(Remarks to the Author)

The authors responded to my questions and made the appropriate adjustments to the text. I recommend publication in Nature Communications

Reviewer #3

(Remarks to the Author)

We appreciate the authors' detailed responses to our comments. Based on the revisions and rebuttal, we now agree that the manuscript is suitable for publication.

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Point-by-Point Response to Reviewer Comments

Reviewer 1

Zhang, Kumar et al combine structural cryo-EM and biochemical crosslinking data to provide a detailed description of the molecular arrangements of human spliceosomes assembled on a model pre-RNA substrate upon inhibition of kinase CDK11 by the small molecule OTS964. Combined with previous pioneering structural work by the Lührmann & Stark teams and other groups, the results provide important new insights into one key step in the activation of the spliceosome, and into the molecular mechanisms that proofread the proper conformational changes that ensue the stepwise assembly of splicing complexes on splice site regions. The paper not only provides a snapshot of an important transition step but also novel models of general value to understand conformational triggers and mechanisms of proofreading.

Specifically, the structure of the OTS964-inhibited spliceosomal complex provides unexpected insights on how phosphorylation of a splicing factor (SF3B1 N-terminal domain) can 1) serve as a signal to trigger complex conformational transitions (including a striking relocation of SF3B6 required for progression to complex Bact formation, concomitant with major rearrangements in RNA:RNA, RNA:protein and protein:protein contacts) revealing novel driving forces in the process, and 2) compete with the association of proofreading helicase DHX15 and several co-factors, including G-patch-containing proteins and the scaffolding factor SR140, providing a novel model for proofreading activities by helicase complexes.

Collectively, the new structure illustrates the surprising complexity of what previously considered to be a single step for the activation of the fully assembled spliceosome, revealing key roles for post-translational modifications and their interplay with proofreading complexes that had not been implicated at this stage in the process. On this basis I am happy to support publication of the manuscript in Nature Communications. The study will certainly fuel new thinking and work on this and other steps of the complex process of spliceosome assembly and activation, as well as more generally on proofreading of ribonucleoprotein processes.

A potential issue with the manuscript is the high density of the data and information provided, as well as the abundance of sophisticated structural arguments made, which may be overwhelming for non-specialists. While the Discussion is an example of clarity, highlighting the key points of the paper in a clear and enticing manner, the Results section is indeed very dense and at points difficult to follow. In my opinion the following revisions would help to improve the text:

1) Summary: understanding some of the paper highlights requires background information that is not introduced earlier on in the summary, which may make it difficult to follow, specially to gather their general significance, for non-specialists. Examples include the sentences “This explains the stalled position...were not first to relocate” and “We also located the binding of the RNA helicase DHX15...and G-patch proteins”. Non-specialist readers may wonder about the significance of “passing along the PRP8RT domain”, or the significance of G-patch proteins for the recruitment of an RNA helicase.

We have provided more background information in the rewritten Summary

2) Introduction: along the same lines, it may be helpful for the general reader to explain concepts such as “catalytic U6 snRNA triplex” (line 63) and clarify the difference in activities between upstream helicases that “remodel off-pathway” (line 98) and DHX15, which “subsequently dismantles” these complexes (line 99).

As suggested, we have rewritten those sentences as follows

“catalytic U6 snRNA triplex”: Tertiary interactions between a bulged-out nucleotide of the U6 ISL (U74 in humans) and the U6 catalytic triad AGC nucleotides, which are part of U2/U6 helix Ib, as well as the U6 GA dinucleotide directly downstream of its ACAGA box, lead to formation of the catalytic U6 snRNA triplex and enables U6 to position two Mg²⁺ ions that participate directly in splicing catalysis ²⁻⁴. (lines 62-66 rev. ms)

“remodel off-pathway”: Several RNA helicases that drive the rearrangements required during spliceosome assembly have also been proposed as establishing fidelity in splicing by a kinetic proofreading mechanism²⁴⁻²⁷ in which the DEAD/H-box ATPase promotes an optimum substrate by acting after a productive step and by rejecting a suboptimal substrate (or incorrectly structured complex) and consequently driving the substrate into an “off-pathway” state^{8,27,28}. Such aberrant, off-pathway spliceosomes may then be dismantled by the RNA helicase DHX15 (also termed PRP43). DHX15 is well known for its role in dismantling the terminal ILS spliceosome, in cooperation with its ATPase-activating G-patch protein TFIP11²⁹⁻³¹ (lines 96-104)

3) Line 127: it would be helpful for the general reader to know the rationale for the choice of the PM5-10 pre-mRNA (the word “construct” may not be necessary here).

We have expanded the sentence as follows:

To elucidate the molecular architecture of an OTS964-stalled spliceosome, we used PM5-10 pre-mRNA, which lacks the second exon (Fig. 1a). It allows efficient B-to-B^{act} complex transformation in HeLa splicing extract¹⁸ (Extended Data Fig. 1a), but it stalls spliceosome formation at the activated (B^{act}) stage. In the presence of 0.2 μ M OTS964; however, PM5-10 B^{act} complex formation is completely inhibited, and a significant fraction of stalled spliceosomes migrates in native gels in a manner similar to that of B/pre-B^{act} stage complexes (Extended Data Fig. 1a). The IC₅₀ of OTS964 for CDK11 is less than 100 nM^{7,23}. (lines 127-134)

4) Line 128: it may be useful to state in the main text the concentration used of OTS964 and its IC₅₀.

Done; see point 3)

5) Line 132: it will be helpful to indicate the identity of the tag used (MS2?) for affinity chromatography of the complexes.

Done

6) Line 166: “5’ domain of U2 snRNP” may require clarification for non-specialists the first time the term is used, e.g. proteins associated with the 5’ x nucleotides of U2 snRNA.

In the revised version of the manuscript, we have clarified the term “5’ domain of U2 snRNP” in the new Summary. (lines 33-40)

7) Line 207: it may not be obvious why the extended U6 ISL “elicits” (rather than being a consequence of other conformational changes) a large-scale downward movement of the 5’ domain of U2 snRNP, i.e. why it is causative of the movement.

In the revised version we provide additional reasons why we believe that the formation of the U6 ISL is the likely driving force for moving the U2 5’ domain closer to the spliceosome core:

The ISL structure in our pre-B^{act-OTS} complex is similar to yeast U6 snRNA within the U6 snRNP complex, which also exhibits an extended ISL structure that includes the invariant AGC triad at the top of the ISL stem^{37,38}. Thus, the extended ISL appears to be a default structure into which this part of U6 RNA folds immediately after BRR2-helicase-mediated unwinding of the U4/U6 stems I and II within the B complex. (lines 188-193; see also lines 210-212)

8) While the extensive Extended Data Figures, Supplementary Tables and specially the supporting Movies are very helpful to get the message across, entire sections (e.g. those covering the “Coordination by SMU1/RED...” and “Location of NTC/NTR...” are illustrated through Supplementary Materials which the average reader may not be inclined to use extensively. I would recommend the authors to consider including a key summary Figure of each of these in the main set of Figures.

We agree with the referee. We have now added a new Figure 4 in the revised version, entitled “Structural organization of SMU1/RED and NTC proteins in pre-B^{act-OTS} complexes”; it shows a new schematic, depicting the coordinated movements of the U2 snRNP 5’ domain, the SMU1/RED tetrameric complex and the BRR2 helicase during the B-to-pre-B^{act-OTS} transition (panels a and b). In the new Figure 4c we show the intermediate docking sites of the WD40 domain of the key NTC protein PLRG1 and of the U6 LSm complex at the PRP8-HB domain in the pre-B^{act-OTS} complex. In the Extended Data Figures 4 and 5 of the revised manuscript version we have omitted panels e and a, respectively, from the old Figures. The old Figures 4–6 now become Figures 5–7, respectively.

9) Line 294: RRM protein SF3B6: it may be helpful for the general reader to introduce the concept of the RRM fold.

In the revised version we explain the term RRM: ...we noticed that the position of the SF3B6 (p14) protein, which contains an RNA recognition motif (RRM), at SF3B1^{HEAT} is strikingly different.... (lines 296-298)

10) Lines 312-313, lines 342-344, section starting in line 346 and Discussion: while the collective evidence supports a key role of CDK11-dependent phosphorylation of SF3B1N as the major driving force for triggering the repositioning of SF3B6, is there a clear model for how this occurs, i.e. what key molecular contacts between SF3B6 and HRs 12-14 would be broken (and/or contacts with RES protein SNIP1 facilitated) by the access of CDK11 complex / phosphorylation of SF3B1 T426/434/436?

Unfortunately, the resolution of the U2 5' domains in the cryo-EM structures of our pre-B^{act-OTS} and later pre-B^{act} complexes are not good enough to allow us to answer this interesting question at the atomic level. However, we consider that the positioning of the three threonines T426/434/436 at or close to the interfaces between the SF3B6 module and SF3B1^{HEAT} in pre-B^{act-OTS} (in their unphosphorylated states) and in pre-B^{act-1} and later activated stages (in their phosphorylated states) is intriguing and raises the interesting possibility that phosphorylation of those threonines may play a dual role in the relocation of the SF3B6 model, as discussed in our manuscript (see lines 370-377). Clearly, higher resolution structures will be needed to clarify this point in the future.

11) Lines 442-450: the crosslinking data suggest a direct interaction between SR140 and SUGP1. However the groups of Manley, Rabadan and Liu have reported a direct interaction of SUGP1 with SF3B1 as an alternative way to recruit SUGP1 (and consequently DHX15), with the appealing added value of potentially explaining the mechanism by which cancer-associated mutations in SF3B1, by disrupting these interactions, cause cryptic splice site utilization: PMID: 31474574, PMID: 32332164, PMID: 40027711, PMID: 40714635, PMID: 36459648, PMID: 37977822. The authors could consider to extend the discussion of these findings.

We agree with reviewer 1 that the direct interaction of SUGP1 with HEAT repeats 2–7 of SF3B1 plays a major role in the stable complex formation between SUGP1, SF3B1^{HEAT} and DHX15, as indicated by the combinatorial structural biology approach of the groups of Manley, Galej and Liu. However, our observation that SUGP1 interacts with SR140 does not contradict the studies by Zhang *et al.* (PMID: 37977822; ref. 36). First of all, SUGP1 binds to SR140 through its N-terminally located SURP domains, far upstream of the SUGP1 regions that directly interact with SF3B1^{HEAT}; that is, the binding of SUGP1 to SR140 and its binding to SF3B1^{HEAT} are not mutually exclusive. In our view, SR140 functions as a hub for several G-patch proteins, including SUGP1, facilitating kinetically the interaction of SUGP1 with SF3B1^{HEAT}, and consequently also the recruitment of DHX15 to the spliceosome. In addition, we would like to point out that Zhang *et al.* (PMID: 37977822, ref. 36) propose in their study that SUGP1 plays a major role in the recruitment of U2 snRNP to the branch site of the pre-mRNA intron, and would thus interact with the spliceosome already very early, directly after E-complex formation, while in our study, SUGP1 (in complex with SR140) interacts with the spliceosome much later, during spliceosome activation. It is therefore possible that the requirements for complex formation between SUGP1, the U2 5' domain and DHX15 may differ somewhat at different stages of spliceosome assembly. In the revised version of our manuscript, we have referred to this point as follows:

Recent studies have revealed that SUGP1 co-operates with DHX15 in proofreading and dismantling defective spliceosomes, although without evidence for the participation of the SR140-CHERP-SPF45 complex³²⁻³⁵. While in some cases it has been suggested that such proofreading may take place during very early stages of spliceosome assembly, such as in the pre-A or A complex³⁵, in our study, SUGP1 (in complex with SR140) interacts with the spliceosome much later, i.e., during spliceosome activation. It is therefore possible that the requirements for complex formation between SUGP1, the U2 5' domain and DHX15 may differ at different stages of spliceosome assembly (see also below). (lines 652-660)

12) Somewhat related to the previous point, recent work by Damianov and Black (both reference 48 as well as PMID: 40138349 -again potentially explaining the effects of cancer mutations in SF3B1-) indicates that association of SR140 (along with CHERP and RBM17/SPF45) occurs at early stages of branch point recognition on chromatin-associated nascent transcripts. Is this mode of DHX15 recruitment additional / overlapping / complementary to the model presented by the authors in the context of SF3B1N phosphorylation / SF3B6 translocation?

In the work of Damianov, Black and coworkers (reference 61), U2 snRNP was purified from chromatin following digestion of the isolated chromatin with DNase and RNase. As the authors discuss in their paper, it is therefore difficult to decide whether their isolated U2 snRNP-branch site RNA-DHX15-SR140/CHERP/SPF45 complex is derived from an early, A-stage spliceosome or from a later assembly stage such as pre-B or B complexes. However, as Damianov *et al.* have purified their U2 snRNP complex with a tagged RBM5 (or RBM10) G-patch protein, and both proteins are essentially absent from our pre-B^{act-OTS} complex preparation, this would indicate that their U2 snRNP complex is not derived from an early, activated spliceosome. Therefore, it will be interesting to see whether the SR140/CHERP/SPF45 complex may also play a role in tethering the G-patch proteins RBM10 and/or RBM15 to SF3B1^{HEAT}. If so, this would indeed raise the possibility that SR140/CHERP/SPF45 may act repeatedly during different spliceosome remodeling steps, *en route* to spliceosome activation. We have included the following sentences in the revised version of the manuscript:

Notably, other G-patch proteins, such as RBM5 and RBM10, have also been implicated in supporting DHX15-mediated quality control of early spliceosome assembly^{61,62}. For example, Damianov *et al.*⁶¹ purified U2 snRNP/branch site complexes from RNAase-treated chromatin, together with DHX15 and the SR140-SPF45-CHERP protein complex, raising the possibility that the latter complex may also play a role in tethering the G-patch proteins RBM10 and/or RBM5 to SF3B1^{HEAT}. As RBM5/10 proteins are absent in our pre-B^{act-OTS} complex, it appears likely that their U2 snRNP complex is derived from an earlier spliceosome assembly stage, suggesting that SR140-SPF45-CHERP may act repeatedly during different spliceosome remodeling steps *en route* to spliceosome activation. (lines 666-677)

13) The authors could consider moving the final paragraph of Results (lines 480 – 491) to the Discussion section.

As we have expanded significantly the Discussion in the revised version, we would prefer to keep this paragraph at the end of the Results section, though in a slightly shortened form. (lines 483-486)

14) Lines 552 – 554: the absence of SF3B6 and of a conserved Nt region in the yeast SF3B1 counterpart are indeed striking and probably merit an extended discussion: how do yeast spliceosomes transition between B and Bact complexes, then? E.g. does CDK11-mediated phosphorylation of Hsh155 occur in *S. cerevisiae*? Does OTS964 inhibit spliceosome assembly in budding yeast?

We would like to thank reviewer 1 and reviewer 3 (point 2) for addressing this interesting point. As suggested by reviewer 3, we have carried out a brief bioinformatic analysis of the evolutionary conservation of SF3B6, CDK11 and numerous TP/SP dipeptide repeats in the N-terminal region of SF3B1 (SF3B1^{N-TP/SP}). As indicated in the new Extended Data Figures 9 and 10, SF3B1^{N-TP/SP}, SF3B6 and CDK11 kinase are highly conserved across metazoans and plants. In contrast, in *S. cerevisiae* and numerous other fungi, the SF3B6-SF3B1^{N-TP/SP}-CDK11 axis is absent. Only few organisms such as *S. pombe* (new Extended Data Fig. 9) contain SF3B6, but lack a TP/SP rich N-terminus in the *S. pombe* SF3B1 and a genuine homolog of CDK11 (PMID: 39245599; ref. 57), suggesting that the relocation of SF3B6 at SF3B1^{HEAT} during early activation of the *S. pombe* spliceosome requires a distinct regulatory mechanism. Alternatively, the SF3B6 module may be less stably attached to the C-terminal region of SF3B1^{HEAT} in *S. pombe*, so that it may already be destabilized and relocated at SF3B1^{HEAT} during the large-scale translocation of the U2 5' domain during early activation of the spliceosome. Interestingly, the intron-rich yeast *Cryptococcus neoformans* does contain an SF3B6 homolog, an SF3B1^N domain with a short TP-rich region, and putative homologs of CDK11 kinase, cyclin and SAP30BP⁵⁶ (Extended Data Fig. 9 and 10). In summary, there appears to be a strong correlation between the conservation of a TP/SP dipeptide-rich N-terminal domain of SF3B1, the CDK11 kinase and SF3B6 across species during evolution; this further suggests that the CDK11-mediated SF3B6 "protein switch" we describe in the human spliceosome is a conserved regulatory mechanism in organisms with complex splicing architectures.

In the revised manuscript, we describe the evolutionary conservation of the SF3B6-SF3B1^{N-TP/SP}-CDK11 axis in the Discussion (lines 560-579). The results of our bioinformatic analysis are shown in the new Extended Data Figures 9 and 10. The old Extended Data Figure 9 becomes the new Extended Data Figure 11. We also include this point in the Summary of the new version of the manuscript.

Interestingly, our bioinformatic analysis also revealed that the **SMU1/RED proteins**, which guide the U2 snRNP 5' domain to the intermediate checkpoint state in pre-B^{act-OTS}, are absent in all organisms that lack the CDK11-mediated SF3B6 protein switch, including *S. cerevisiae* (included in the new Extended Data Figure 9). Consistently, the yeast *C. neoformans*, however, also expresses conserved SMU1 and RED homologs (Extended

Data Fig. 9). These observations underpin the idea that the role of SMU1/RED in coordinating the sophisticated remodeling events of U2 and U6 snRNPs as well as BRR2, is evolutionarily connected to the conserved SF3B6-SF3B1^{N-TP/SP}-CDK11 axis. The absence of this regulatory module in *S. cerevisiae* suggests that the early activation phase of the yeast spliceosome might follow a less complex remodeling pathway than that of humans. We have described this point in the Discussion of the revised manuscript (lines 580-594).

Reviewer 2

The manuscript by Zhang et al. uses OTS964 and a PM5-10 pre-mRNA construct to determine, by cryo-EM and mass spectrometry, the structure and composition of spliceosomes blocked by inhibition of CDK11, a kinase that phosphorylates the SF3B1 N-terminus and is essential for spliceosome activation. The authors identify a spliceosome intermediate termed PM5-10 pre-Bact-OTS, which proceeds a previously characterized intermediate, pre-Bact-1, identified by the same group (PMID: 33243851). The pre-Bact-OTS complex is characterized by dissociation of the U4 snRNP, the presence of NTC, NTR, and IBC proteins, and the formation of a previously unobserved step in the establishment of the U2/U6 snRNA catalytic core. The authors further identify roles for the SMU1/RED complex in early stages of spliceosome remodeling and suggest a crucial role for CDK11-mediated phosphorylation of the SF3B1 N-terminus in the transition of what they term the SF3B6/SF3B1(394–430) module from SF3B1 HEAT repeats 12–14 to repeats 1–2. To characterize the roles of phosphorylated SF3B1, they use crosslinking approaches and identify three threonine residues involved in SF3B6/SF3B1(394–430) module transition and SNIP1-mediated recruitment of the RES complex to the SF3B1 HEAT repeats 1–2. Finally, using their PM5-10 pre-Bact-OTS structures, crosslinking data, and AlphaFold3 modeling, the authors determine the location and propose functions of the DHX15 helicase and three G-patch-associated proteins (SUGP1, CHERP, and SPF45) in quality control during early stages of CDK11-dependent spliceosome activation.

Overall, the manuscript provides new mechanistic insights into spliceosome activation, is well written and timely, and will be of interest to the readership of Nature Communications. However, before publication, I have the following questions, comments, and suggestions that need to be addressed:

1. Page 3, lines 70–77: This recent paper describing several structures of activated spliceosomes (PMID: 39068178) should also be cited, and these structures should be compared in the Discussion with those observed following CDK11 inhibition.

We have included the requested references in the revised version of the manuscript. We have also referred to the cryo-EM structures described in PMID: 39068178 in the following way:

Recently, two studies, using small-molecule splicing inhibitors, revealed cryo-EM structures of three human activated-spliceosome precursors, termed pre-B^{act-1} and pre-B^{act-2}⁵, or just pre-B^{act}¹⁶, the latter representing an intermediate assembly stage that arises after pre-B^{act-2} but before the formation of mature B^{act} complexes. (lines 69-73)

As the only pre-B^{act} structure captured in this paper is positioned further downstream of pre-B^{act-2}, while our pre-B^{act-OTS} structure is positioned upstream of pre-B^{act-1}, we do not think that it is necessary to compare the two structures directly.

2. Page 3, line 84: The authors should additionally cite a recent paper characterizing SF3B1 phosphorylation (PMID: 41428738) and at page 4, line 88, this paper (PMID: 31511426) should be also cited.

Done

3a. Figure 1a and Extended Data Figure 1a, c: The PM5-10 pre-mRNA construct was used, but this substrate contains a very short intron (and a short distance between the 5' splice site and the branch site). Do the authors observe identical results when other splicing substrates with longer introns, such as MINX or AdML, are used?

The PM5-10 pre-mRNA contains a 151-nt-long intron (distance between the 5'SS and BS) and is thus significantly longer than the 96-nt-long intron of MINX pre-mRNA. Moreover, the length of the polypyrimidine tract downstream of the intron's BS-A is also comparable between PM5-10 and MINX pre-mRNAs (16 nt vs. 21nt; not counting the 3'SS of the MINX intron, as this is missing in PM5-10). Finally, both poly-Py tracts share a run of 7 or 8 consecutive uridine nts, respectively, and should thus exhibit a similar affinity towards U2AF⁶⁵. Thus, we did not push for the structural analysis of spliceosomes assembled in the presence of OTS964 on other pre-mRNAs with similar sizes.

3b. In addition, controls performed in the absence or at low concentrations of ATP and under low Mg²⁺ conditions should be included to allow a more rigorous comparison between B and Bact-OTS complexes.

The rationale for these experiments is not obvious to us. We have recently published the currently most comprehensive and concise 3D cryo-EM structure of a human B complex that was stalled in the presence of ATP_γS (PMID: 38383864, ref. 41). We showed previously (PMID: 27411562) that low Mg²⁺ and/or ATP concentrations also stall spliceosome formation at the B stage level, however, these complexes are biochemically more heterogeneous than those stalled in the presence of ATP_γS. Thus, the gold standard for a rigorous comparison of our pre-B^{act-OTS} structure with a B complex cryo-EM structure is that of the B^{ATP_γS} complex. To rule out possible structural ambiguities, we have repeated the reconstruction of the PM5-10 pre-B^{act-OTS} cryo-EM structure three times, using three independent spliceosome purifications, which revealed essentially the same structures.

4. Page 5, lines 132–147: The composition of OTS964-stalled complexes assembled on PM5-10 pre-mRNA differs from chromatin-associated spliceosome complexes in cells after OTS964 treatment (PMID: 36104565), specifically with respect to the absence of U4 snRNA and the presence of the IBC complex. There may be several reasons for these differences, and the authors should at least comment on them (see also points 3, 5, and 6).

These apparent inconsistencies may be at least partially explained by the significant differences in the experimental design of the two studies. That is, while we discuss the near-atomic structure of an OTS964 stalled, defined spliceosome population, assembled *in vitro*, Hluchy *et al.* (PMID: 36104565, ref. 7) investigate the global composition of thousands of distinct pre-mRNA spliceosomes that remain associated with chromatin upon treatment of HTC116 cells with OTS964 *in vivo*. On the basis of retention of U4 snRNA in chromatin and reduced amounts of IBC proteins, they concluded that the chromatin-associated spliceosomes are stalled at a B-like stage, before BRR2-mediated activation. We note that chromatin-associated spliceosomes are assembled co-transcriptionally and are therefore very likely still to be tethered to the Pol II transcription machinery. Thus, it is conceivable that one or more proteins in chromatin-associated spliceosomes might require CDK11-mediated phosphorylation before BRR2-mediated dissociation of the U4/U6 RNA helices occurs, and subsequent phosphorylation of SF3B1. Alternatively, or additionally, the release of destabilized factors such as U4 snRNP may also follow different kinetics in chromatin-associated – and thus Pol II-tethered – spliceosomes, compared with spliceosomes assembled *in vitro*.

The apparent inconsistency regarding the different abundance of AQR in the spliceosomes of the two studies can also be explained. In the OTS964 stalled endogenous spliceosomes AQR is not detected by iCLIP. However, in the immunoblot analysis AQR is still present in reduced amounts (Fig. 4b; Hluchy *et al.* (PMID: 36104565). In our purified PM5-10 pre-B^{act-OTS} preparation, AQR is clearly present, as determined by MS analysis; however, it is only flexibly tethered by protein–protein interactions to the peripheral part of the spliceosome complex and not positioned in the intron-bound position that AQR adopts at later activation stages. Thus, AQR would not be detectable in our preparation via iCLIP.

Interestingly, in a more recent proteomic study, Blazek and coworkers (PMID: 41904131, ref. 53) show that certain NTR proteins, such as SKIP and RBM22, are also present in the chromatin-associated spliceosomes isolated from HTC116 cells that had been incubated with OTS964. As the binding of SKIP to spliceosomes is mutually exclusive with that of the U4 snRNA-associated PRP31 protein, this would indicate that a fraction of the chromatin-associated spliceosomes had undergone at least partial activation. In view of the evolutionary conservation of the SF3B6-SF3B1^{N-TP/SP}-CDK11 axis, which we reveal in our study, we are confident that the CDK11-mediated molecular SF3B6 protein switch is one of the major structural rearrangements that is driven by CDK11 kinase during activation of human spliceosomes. We summarize the above points in the Discussion of our revised manuscript as follows:

In a recently published study, Hluchy *et al.*⁷ blocked CDK11 kinase with OTS964 *in vivo* in HTC116 cells and investigated the effect of this on the global composition of thousands of distinct pre-mRNA spliceosomes that remain associated with chromatin. On the basis of retention of U4 snRNA in chromatin and reduced amounts of IBC proteins, they concluded that the chromatin-associated spliceosomes are stalled at a B-like stage, before BRR2-mediated activation. We note that chromatin-associated spliceosomes are assembled co-transcriptionally and are therefore very likely still to be tethered to the Pol II transcription machinery^{59,60}. Thus, it is conceivable that one or more proteins in chromatin-associated spliceosomes might require CDK11-mediated phosphorylation before BRR2-mediated dissociation of the U4/U6 RNA helices occurs, and subsequent phosphorylation of SF3B1. Alternatively, or additionally, the release of destabilized factors such as U4 snRNP

may also follow different kinetics in chromatin-associated – and thus Pol II-tethered – spliceosomes, compared with spliceosomes assembled in vitro^{59,60}. Interestingly, in a more recent proteomic study, Blazek and coworkers⁵³ show that certain NTR proteins, such as SKIP and RBM22, are also present in the chromatin-associated spliceosomes isolated from HTC116 cells that had been incubated with OTS964. As the binding of SKIP to spliceosomes is mutually exclusive with that of the U4 snRNA-associated PRP31 protein, this would indicate that a fraction of the chromatin-associated spliceosomes had undergone at least partial activation. (lines 607-625)

5. SAD1 and SNU66 inhibit BRR2 function but are found as part of the Bact-OTS spliceosome (see Supplementary Table 1, where SAD1 and SNU66 are abundant in Bact-OTS). Do the authors expect BRR2 to be active and capable of unwinding U4 snRNA when SAD1 and SNU66 are present? In addition, neither of these proteins is shown in any of the structures depicting the Bact-OTS complex.

Indeed, in the first version we only briefly described the biochemical composition of our purified spliceosomes that were stalled in the presence of OTS964, as we focused mainly on describing the cryo-EM structure of pre-B^{act-OTS} assembly intermediate. To address the referee's point: First of all, it is important to note that our pre-B^{act-OTS} spliceosome preparation still contains reduced amounts of U4 snRNA (as stated in line 139 of our original manuscript and shown in Ext. Fig. 1c.). Thus our pre-B^{act-OTS} preparation still contains spliceosome populations that likely contain U4/U6.U5 tri-snRNPs. This explains why our MS analysis also reveals the presence of U4/U6.U5 tri-snRNP proteins such as SAD1, SNU66 and DIM1. This is also consistent with the observed crosslinks of WBP4. Whether the tri-snRNP is part of a B-stage or pre-B stage complex, or both, is difficult to decide in the absence of cryo-EM structures of this tri-snRNP-containing spliceosome population(s). In our original manuscript we note that the presence of OTS964 in splicing extracts also leads to an accumulation of A complexes (Ext. Fig 1a and as stated in lines 1213-1216 of our orig. manuscript). This may indicate that the stable incorporation of U4/U6.U5 tri-snRNP into a B complex during the pre-B to B complex transformation may be hindered/slowed down, because one or more early splicing factors such as SR proteins may require phosphorylation by CDK11^{78,79}. In the revised version of our manuscript, we have added the following sentence to the legend of Extended Data Fig. 1c:

The reduced amounts of U4 snRNA indicates that some U4/U6.U5 tri-snRNP-containing populations of spliceosomes are also present in our pre-B^{act-OTS} preparation, consistent with the presence of U4/U6.U5 tri-snRNP proteins such as SAD1, SNU66 and DIM1 in our MS analysis (Supplementary Table 1). (lines 1377-1381)

6. TXNL4A (DIM1) and FBP21 (WBP4) are identified in the crosslinking dataset. The α -finger of PRPF8 interacts with PRP31 and DIM1 to stabilize U4/U6 stem I. Therefore, the presence of DIM1 is also inconsistent with U4 snRNA having been unwound in the Bact-OTS complex.

The explanations given under point 5 also apply here.

7. SMU1/RED were previously shown to be required for efficient splicing of short introns (PMID: 31409787). Are the newly identified roles of this complex described on page 9 restricted to spliceosomes assembled on short introns? See also point 3.

In our previous paper (PMID: 31409787), we showed that SMU1/RED play a particularly important role for the activation of spliceosomes that are assembled on introns with a short distance (about 50 nt) between the 5'SS and BS, while depletion of SMU1/RED did not abolish the activation of PM5 or MINX pre-mRNAs, but significantly affected the efficiency and kinetics of splicing. While at that time we could only speculate about the possible role(s) of SMU1/RED in the transformation of B to B^{act} complexes, our new pre-B^{act-OTS} structure now reveals that SMU1/RED coordinate the contra-rotating movements of U2 snRNP and BRR2 during early activation. As these RNP modules belong to the core of the spliceosome, their choreography should also be similar in the presence of large introns, and hence SMU1/RED likely facilitate the remodeling of any human pre-mRNA spliceosome. This idea is further supported by our observation that the role of SMU1/RED in spliceosome activation appears to be evolutionarily connected to the SF3B6-SF3B1^{N-TP/SP}-CDK11 axis, as in the absence of SF3B6 and CDK11 in various organisms such as *S. cerevisiae*, SMU1/RED is also missing (see our more detailed discussion of this point above).

8. Page 19, line 520: "HR 1–2" should be used instead of "HR 1–3".

Done

9. In the Discussion (page 19, line 529), the statement should omit the term “phosphorylated,” as SF3B1 is not phosphorylated in the pre-Bact-OTS complex, but rather in the pre-Bact and Bact complexes.

Done

Similarly, in the Extended Data Figure 6e T426, T434 and T436 are labeled by the same color irrespectively of their phosphorylation status at different steps of the spliceosome activation (Bact-OTS, Bact-1/2 and Bact) which is misleading. This needs to be corrected and clarified. It is also not clear why amino acids 67 (SF3B6) and 394 (SF3B1) are highlighted in the Extended Data Figure 6e, please explain.

We thank the reviewer for pointing out these errors. In the revised version we have changed the colors of T426, T434, and T436 in the pre-B^{act-OTS} state. We have also deleted the labels for SF3B6 aa 67 from the figure to avoid any possible confusion. Amino acid 394 of SF3B1 is still labelled as we mention the SF3B6/SF3B1³⁹⁴⁻⁴³⁰ module in the text, and therefore it is necessary to indicate where this amino acid is located.

We note that in this context we have also referred to a recent paper by Blazek and co-workers (reference 53) the revised version of the manuscript:

Consistent with our structural analysis, it has recently been shown that CDK11-mediated SF3B1 phosphorylation is functionally coupled to the SNIP1-dependent recruitment of the RES complex during B-to-B^{act} transition⁵³.

10a. Page 5, line 118 and elsewhere: The statement “Aberrant spliceosomes would impair CDK11 recruitment (and thus SF3B1 phosphorylation)...” contradicts the data presented in Data S1, where CDK11-CCNL-SAP30BP are clearly crosslinked to the spliceosome in the pre-Bact-OTS complex. This indicates that aberrant spliceosomes do not impair CDK11 recruitment; rather, the recruited kinase is inhibited by OTS964 in the pre-Bact-OTS complex.

“Aberrant spliceosomes would impair....”: We would like to point out that the term “Aberrant spliceosomes” refers in our model to spliceosomes that have undergone aberrant remodeling events during early activation, which would prevent the docking of the N-terminal domain of SF3B1 at the PRP8-RT domain. In this sense, our pre-B^{act-OTS} complex does not really represent an aberrantly remodeled intermediate assembly stage (our cryo-EM structure testifies too the native state of the remodeled early activated spliceosome) – and we show docking of SF3B1^N to PRP8^{RT} (Extended Data Fig. 6b,c); only owing to the inhibition of CDK11, pre-B^{act-OTS} cannot progress further towards pre-B^{act-1}, and therefore it mimics an aberrant assembly state, allowing the binding of DHX15 to SF3B1^{HEAT}.

“..where CDK11-CCNL-SAP30BP are clearly crosslinked to the spliceosome ...”:

While Supplementary Table 2 indeed shows some crosslinks of the CDK11-CCNL1/2-SAP30BP complex to pre-B^{act-OTS}, we would like to emphasize that they are minor. First of all, the amounts of the CDK11 kinase proteins that co-purify with our pre-B^{act-OTS} complexes are very low (they do not appear among the 100 most abundant proteins). The crosslinks in Supplementary Table 2 were detected by carrying out the p-Link searches against the top 303 proteins. Yet, most of the crosslinks at FDR 1% are observed between the three protein subunits of the CDK11 complex. A few, very weak, crosslinks are observed at FDR 5% to spliceosomal proteins. However, several of the later proteins, such as hnRNP D, SYF1, TRA2B, PABP1 or ROA1, are not parts of our pre-B^{act-OTS} cryo-EM structure. Thus, in our view, these crosslinks do not really allow conclusions about the binding site of the kinase complex at the pre-B^{act-OTS} snapshot.

10b. In a similar vein, while SAP30BP may play a role in recruiting the CDK11 complex to certain spliceosome intermediates (as discussed on page 22, line 616), the mechanism proposed in reference 52 is difficult to reconcile with the current findings, since SAP30BP recruitment in that study depends on SF3B1 phosphorylation, which itself depends on CDK11 activity.

We have referred to the paper by Mayeda and coworkers (PMID: 38065098, ref. 47), primarily because they showed for the first time that SAP30BP can in principle interact with SPF45 via their ULM and UHM modules, respectively. Otherwise, the roles discussed for SPF45 and SAP30BP in the study by Mayeda and coworkers are significantly different from those in our study. For example, they propose that SPF45 interacts via its G patch with the poly-Py tract of very short introns at an early stage of spliceosome formation, thereby likely facilitating the deposition of an unknown factor X at the 3'SS of the pre-mRNA (or *vice versa*). In their study, SAP30BP

appears to play a chaperone role for SPF45, controlling the timely interaction of SPF45 with SF3B1. In our view, the role the authors discuss for the phosphorylation of SF3B1's N-terminus in the recruitment of SPF45 to SF3B1 remains enigmatic, as it would require phosphorylation of SF3B1N at an early spliceosome assembly state, for which there is no evidence.

In contrast, in our study, the proposed SPF45-SAP30BP interaction, would happen at a later stage of spliceosome assembly, shortly after activation, but prior to SF3B1 phosphorylation. Moreover, we would like to point out that in the pre-B^{act-OTS} state, the poly-pyrimidine tract downstream of PM5-10 would not be available for interaction with SPF45 at this state, as it is occluded by the inner ring of the SF3B1-HEAT domain. Finally, it is very likely that in the study by Mayeda and coworkers, SPF45 and SAP30BP function as individual proteins, while in our study, they are part of the SPF45-SR140-CHERP complex and the CDK11 kinase heterotrimer, respectively; thus, no interaction of SPF45 with SF3B1 needs to be envisaged for the recruitment of SPF45.

Therefore, it was (and still is) tempting to suggest that upon initial docking of CDK11 kinase to SF3B1^N, SAP30BP may be tethered to SPF45 to facilitate efficient phosphorylation of the numerous TP and SP phosphorylation motifs. However, we agree that in the absence of any experimental evidence this specific interaction model is premature. Therefore, we would prefer to present this tethering model in the revised manuscript in more general terms:

While the mechanism of the recruitment of CDK11 to SF3B1 remains unclear, we propose that the tethering of the N-terminal domain of SF3B1 to PRP8^{RT} in pre-B^{act-OTS} may re-orient its TP/SP-rich central region, thereby presenting a high-affinity interaction site for the CDK11 kinase. It remains to be seen whether, upon initial docking to SF3B1^N, the CDK11 complex is further stabilized by additional interactions with one or more other spliceosomal proteins, thus facilitating efficient (and possibly processive) phosphorylation of the numerous TP/SP phosphorylation sites. For example, SAP30BP and SPF45 have been shown in principle to interact through their ULM and UHM modules, respectively (though in a spliceosomal context very different from ours)⁴⁷. It is therefore tempting to suggest that the binding of CDK11 to the spliceosome could be stabilized by interactions between those two proteins. Alternatively, or additionally, cyclin L, which contains an RS-rich region, may also play a role in tethering CDK11 to the spliceosome. We anticipate that elucidating the structural basis of the targeting of CDK11 to SF3B1 will prove important for attaining a full understanding of the proposed kinase-dependent proofreading mechanism. (lines 690-708).

10c. In this context, it would also be important to discuss the role of SPF45, which has been shown in several studies to be recruited to unphosphorylated SF3B1 in earlier spliceosome intermediates and to function predominantly in the splicing of short introns or in alternative splicing regulation (PMID: 17589525, PMID: 34389706, PMID: 38065098). Once again, this raises the question of the extent to which the spliceosome structures described in this manuscript are specific to short introns due to the use of PM5-10 pre-mRNA, and to what extent the conclusions can be generalized, given that OTS964 blocks constitutive splicing in cells (PMID: 36104565, PMID: 38059508). These potential limitations should be discussed.

As outlined above, the role SPF45 appears to play in the splicing of very short pre-mRNAs (PMID: 3438970, PMID: 38065098), is different from possible roles that we envisage for SPF45 in the context of our studies.

...or in alternative splicing regulation (PMID: 17589525, PMID: 34389706, PMID: 38065098).

In the paper by Sattler, Valcarcel and coworkers (PMID: 17589525), it has been shown that overexpression of SPF45 in cells induces exon-6-skipping in the FAS receptor pre-mRNA, and that the alternative splicing regulation activity of SPF45 is dependent on the interaction of its UHM motif with a ULM motif of constitutive splicing factors such as SF1 and/or SF3B1. While the exact mechanism of how SPF45 may cause exon-6-skipping of FAS pre-mRNA remains an open question, it is evident that this regulation event will happen during early spliceosome formation at intron 5, before commitment-complex formation, and not after spliceosome activation, where the SPF45-SR140-CHERP complex comes into play in our study. It is interesting to note, though, that SPF45 regulates alternative splicing of exon 6, although the flanking introns 5 and 6, which are relevant for this regulation, are not members of the short intron type, studied by Mayeda and coworkers (PMID: 34389706, PMID: 38065098, and see above). Thus, SPF45 can in principle play various distinct roles in pre-mRNA splicing, and interact with the spliceosome at distinct stages of its assembly. However, as the molecular basis of those functions remains poorly understood, we do not think that it is helpful to discuss these roles in our manuscript in detail. We will, of course, refer to the papers discussed above in our revised manuscript as follows:

We note that our crosslink dataset also shows crosslinks between SPF45 and DHX15 (Supplementary Table 2), but not with SF3B1^{HEAT}. Whether SPF45 may recruit DHX15 to SF3B1 in the context of an aberrantly remodeled early activated spliceosome remains to be seen. As discussed below, SPF45 could also play a role in facilitating the productive interaction of CDK11 kinase with pre-B^{act-OTS} stage spliceosomes. Previous studies have shown that SPF45 exerts additional functions in pre-mRNA splicing, such as promoting spliceosome formation on very short introns^{46,47}, or inducing exon 6 skipping in the FAS receptor pre-mRNA⁴⁸, roles in which SPF45 has to act at early stages of spliceosome formation, before the activation step. (lines 410-419)

Reviewer 3:

Zhang et al. present a structural and functional analysis of a novel early-activated spliceosome intermediate, termed pre-B^{act-OTS} complex, stalled by the CDK11 inhibitor OTS964. By integrating single-particle cryo-EM, crosslinking mass spectrometry, biochemical assays, and AlphaFold3 modeling, the authors elucidate the mechanism through which CDK11-mediated phosphorylation of SF3B1 regulates spliceosome activation and reveal how the helicase DHX15 may proofread this process. While post-translational modifications (PTMs) such as SF3B1 phosphorylation have been recognized as important regulatory events in splicing, the precise structural and functional consequences of this phosphorylation have remained unclear. This study demonstrates that phosphorylation of this core splicing factor triggers a critical structural rearrangement, thereby establishing a checkpoint for the proofreading of spliceosome assembly. Importantly, the work identifies the recruitment of DHX15 and its associated G-patch proteins to this intermediate, with SR140 serving as a key binding hub. Although DHX15 is known for its role in dismantling the ILS complex at the end of the splicing cycle and has recently been implicated in proofreading early assembly stages, this study significantly expands its functional repertoire, positioning DHX15 as a general proofreading factor active at multiple points in the splicing cycle.

Overall, this work provides crucial insights into how PTMs regulate the splicing cycle and how the spliceosome in higher eukaryotes may take advantage of such modifications as checkpoints to ensure fidelity. The implications of this research could extend beyond the splicing field, offering valuable perspectives also for other RNP biogenesis processes. Thus, I am happy to support publication of the manuscript in Nature Communications.

A few comments:

1. The Results section is generally dense and may be challenging for readers not deeply familiar with spliceosome structural biology. Given the broad readership of this journal, providing additional context and simplifying some of the detailed descriptions would greatly improve accessibility and broaden the impact of the work.

As requested also by the other referees, we have provided additional background information at several places in the revised manuscript (see also our response to points 1 and 2 of reviewer 1). Moreover, we have further streamlined the main text by moving some of the more detailed descriptions of the cryo-EM structure to the legends of the corresponding extended figures.

2. The discussion on the absence of this kinase-switch system in S. cerevisiae is interesting and highlights evolutionary complexity in humans/higher eukaryotes. It raises the question of whether this mechanism is conserved in other organisms such as S. pombe or C. elegans, which possess an SF3B6 counterpart. A brief bioinformatic analysis commenting on the conservation of SF3B6, the N-terminal region of SF3B1, and CDK11 across species would be a valuable addition.

Please see our response to a very similar request by reviewer 1 (point 14 above).

3. The mechanism of CDK11 complex recruitment to the spliceosome remains unclear. Specifically, why is CDK11 recruited specifically to this intermediate, and not to others where the SF3B1 N-terminus might also be accessible? Expanding the discussion on this point would be beneficial.

We agree with reviewer 3 that this is an important issue, which is still poorly understood. As the reviewer correctly states, the N-terminal domain of SF3B1 should be accessible in earlier stages of spliceosome assembly, such as the 17S U2 snRNP, A complex, pre-B complex and B complex, yet the SF3B1^N domain is not phosphorylated. We were therefore excited to see that in our pre-B^{act-OTS} complex, the N-terminal region of

SF3B1^N is docked to the RT domain of PRP8, which positions the central, TP/SP rich domain of SF3B1^N between SF3B1^{HEAT} and PRP8^{RT} (Extended Data Fig 6b). As the tethering of SF3B1^N to the PRP8^{RT} domain is exclusively observed at the pre-B^{act-OTS} stage of spliceosome assembly, while in the B complex and later stages of spliceosome activation this region of PRP8 interacts in a mutually exclusive manner with other spliceosomal proteins (see Extended Data Fig. 6c), we propose that it facilitates the recruitment of the CDK11 kinase to SF3B1^N, and thus its efficient phosphorylation. For example, the tethering of the N-terminal domain of SF3B1 to PRP8^{RT} in pre-B^{act-OTS} may reorient its TP/SP rich central domain, thereby presenting a high-affinity binding site for the CDK11 kinase. At least, this would explain, why SF3B1^N is phosphorylated first at the pre-B^{act-OTS} stage, as it requires the significant movement of the U2 5' domain from its peripheral position in the B complex closer to the core of the spliceosome, that is to PRP8. Moreover, as SAP30BP, the stabilizing component of CDK11 (ref. 6), has been shown to interact in principle with SPF45 via their ULM and UHM motifs, respectively (ref. 47), we speculate that upon initial docking of CDK11 kinase to SF3B1^N, SAP30BP may be tethered to SPF45 to facilitate efficient, possibly processive, phosphorylation of the numerous TP and SP phosphorylation motifs. In the revised version, we discuss this point in the context of a potential facilitating role of the CDK11 kinase subunit SAP30BP at the end of the Discussion (lines 690-708; see also response to point 10b of reviewer 2)

4. The Results section includes a segment on the NTC/NTR complex and notes that IBC proteins are loosely associated. The implications of these observations are not discussed later. A brief comment on the potential significance of these associations in the Discussion would be helpful.

We have added the following sentences to the Discussion of the revised manuscript:

That is: Our pre-B^{act-OTS} snapshot visualizes the handover of two outgoing proteins (SMU1/RED, which were recruited at the B complex stage), for new, incoming proteins – such as NTC and IBC proteins – that are positioned in pre-B^{act-OTS} at intermediate spliceosomal docking sites (Fig. 4, Extended Data Fig. 4 and 5)). The provisional docking of NTC/IBC proteins will presumably enhance the kinetics of their final integration into the activated spliceosome – a principle that may be expected to apply to the exchange of many other proteins during the entire activation phase. (lines 511-518).

Description of Additional supplementary files

Supplementary Data 1: Protein composition of the pre-Bact-OTS complex.

Supplementary Data 2: Protein–protein crosslinks identified in the pre-Bact-OTS complex.

Supplementary Data 3: Summary of modelled proteins and RNAs in the pre-Bact-OTS complex.

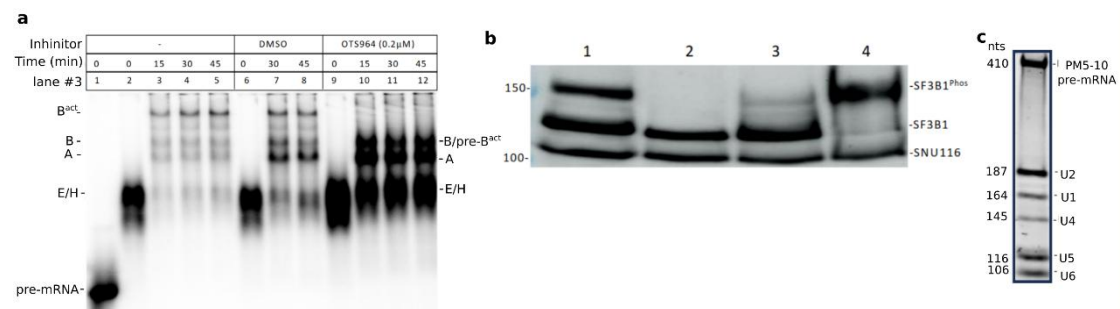
Supplementary Movie 1: Structural remodeling from B to pre-Bact-OTS complex After BRR2-mediated removal of U4 snRNA and the associated proteins, the now single-stranded U6 snRNA folds into the extended ISL conformation, docking to PRP8NTD. The folding of U6 snRNA drives the large-scale translocation of the U2 snRNP. Concomitantly, the proteins hNTR and hNTC are recruited.

Supplementary Movie 2: Remodeling of the interface between U2 snRNP and BRR2 during the B-to-pre-Bact-OTS transition The interface between U2 SF3B3WD40 and SMU1/RED is largely maintained during the B-to-pre-B act-OTS transition. The movement of U2 is coordinated with the rotation of BRR2 by the SMU1/RED complex, and coupled with the docking of SMU1GAC(B) domain to BRR2CC , where it replaces SMU1WD40(A), forming a new interaction between SMU1GAC(B) and BRR2.

Supplementary Movie 3: Coordinated movements of U2 snRNP and BRR2 The 5' domain of the U2 snRNP, together with the SMU1-RED tetramer, rotates clockwise by ca. 45° towards PRP8. Concurrently, BRR2 rotates by ca. 100°.

Supplementary Movie 4: Structural remodeling from the pre-Bact-OTS to the pre-Bact-1 complex During the pre-B act-OTS - to-pre-B act-1 transition, the 5' domain of U2 snRNP undergoes a translocation and docks to PRP8RH. The translocation of SF3B6 from HRs 12–14 to HRs 1–2 is coordinated with the docking of the U2 SF3b complex to PRP8RH. In pre-B act-1 , RES proteins (including SNIP1) are recruited, making contact with the translocated SF3B6, and they further stabilize the new position of the latter.

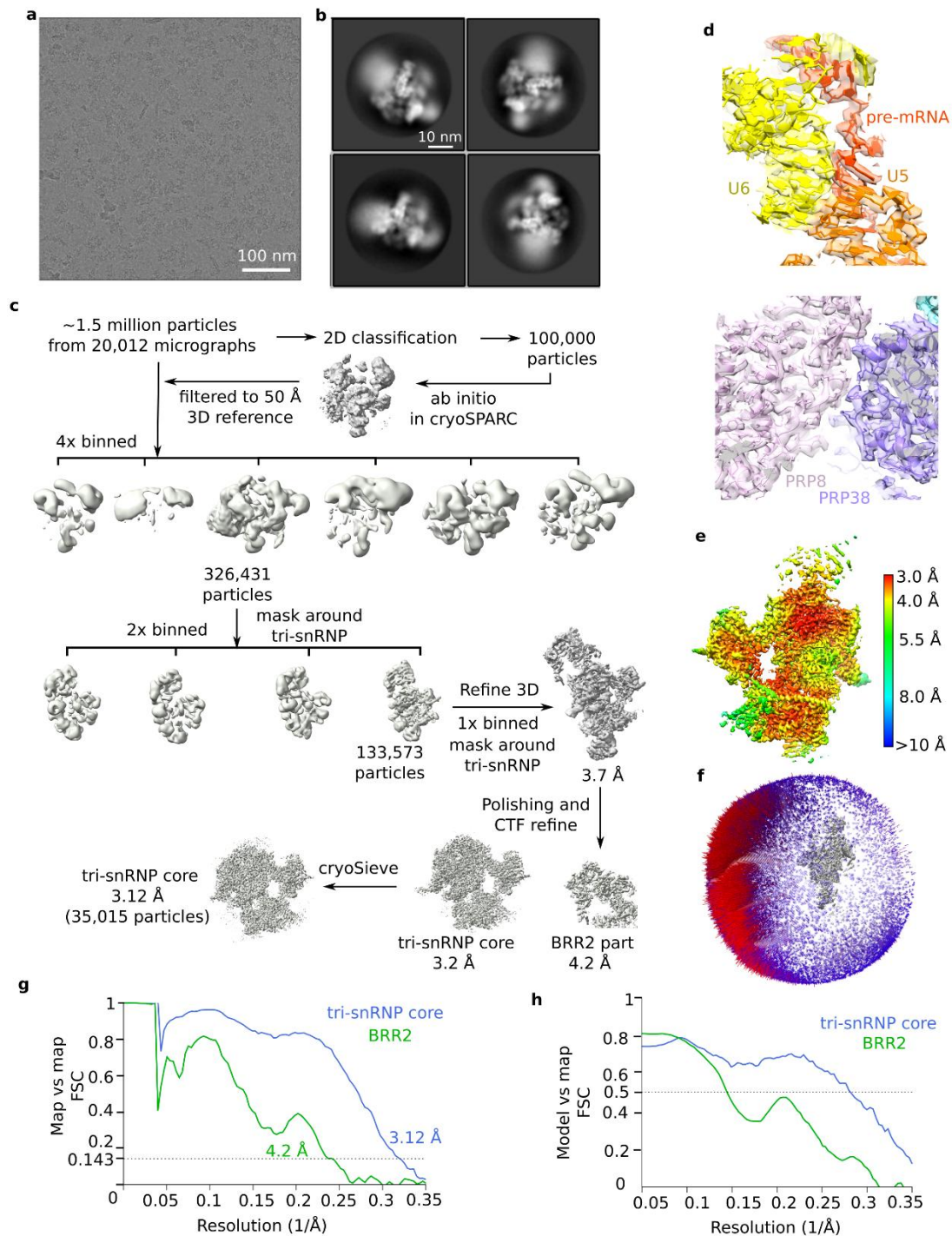
Supplementary Movie 5: Translocation of SF3B6 as a prerequisite for the docking of SF3b to PRP8RH The translocation of SF3B6 is required for docking of the U2 SF3b complex to PRP8RH. If SF3B6 were to remain bound to HRs 12–14, it would clash with PRP8RT and block the translocation of the U2 snRNP.



Supplementary Fig. 1 OTS964 blocks splicing during the transformation of the spliceosomal B complex into B^{act}

a, Inhibition of pre-mRNA splicing by OTS964. Splicing was performed with [³²P]-PM5-10 pre-mRNA in the presence of 0.2 μM OTS964 in HeLa nuclear extract for the times indicated. Spliceosomal complex formation was analyzed on an agarose gel. We note that pre-B^{act} complexes, comigrate in native gels together with B complexes¹. Bands were visualized by autoradiography. DMSO, control reaction with the solvent for OTS964. Spliceosome assembly studies were repeated at least three times with similar results. Lanes 10–12 show that a significant fraction of stalled spliceosomes migrate in native gels in a manner similar to that of B/pre-B^{act} stage complexes. However, A-stage complexes also accumulate, consistently with the idea that inhibition of phosphorylation of SR proteins (such as SRSF1) by CDK11 kinase^{2,3}, may hinder efficient integration of U4/U6.U5 tri-snRNP into spliceosomes⁴ **b**, Western blots reveal that OTS964 blocks phosphorylation of the U2 SF3B1 protein. Proteins from various purified spliceosomes were analyzed by western blotting, using antibodies against the human U2 SF3B1 protein. Lane 1, mixture of affinity-purified PM5-10 A, B and B^{act} complexes; lane 2, purified OTS964-stalled PM5-10 complexes; lane 3, same complex as used in lane 2, but incubated for 15 min with 2 mM ATP before protein extraction; lane 4, purified PM5-10 B^{act} complexes. SF3B1^{Phos}, phosphorylated SF3B1 protein. Antibodies against SNU114 were used to ensure equal loading. Western blotting was analyzed from three independent preparations of the various spliceosomal complexes with similar results. **c**, RNA composition of purified PM5-10 pre-B^{act}-OTS complexes. OTS964-stalled spliceosomes, formed on PM5-10 pre-mRNA, were affinity-purified and subjected to gradient centrifugation; RNA from fractions of the fastest sedimenting

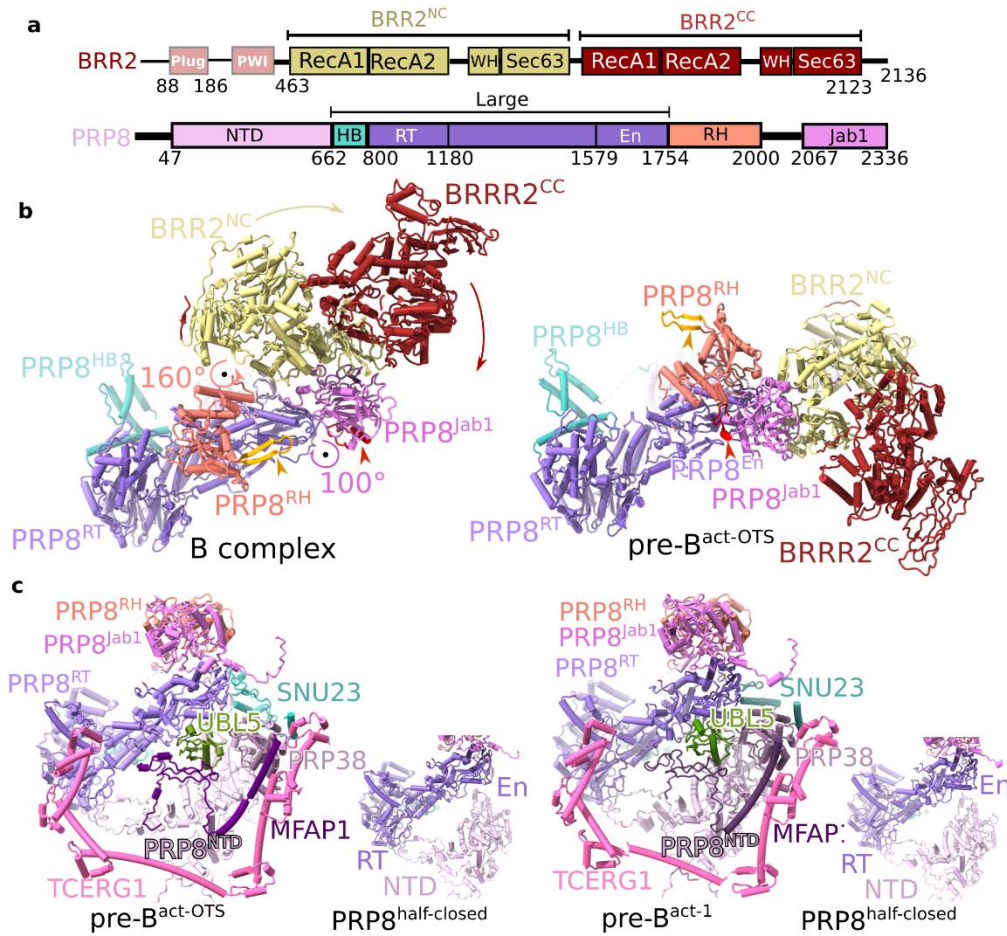
peak was isolated, separated on a denaturing NuPAGE gel, and visualized by staining with SyBr gold. The RNA composition was analyzed from three independent OTS964-stalled complex purifications with similar results. The remaining amounts of U4 snRNA indicate that some U4/U6.U5 tri-snRNP-containing populations of spliceosomes are also present in our pre-B^{act-OTS} preparation, consistent with the presence of U4/U6.U5 tri-snRNP proteins such as SAD1, SNU66 and DIM1 in our MS analysis (Supplementary Data 1).



Supplementary Fig. 2 Cryo-EM and image-processing of the OTS964-stalled PM5-10 spliceosomes (PM5-10 pre-B^{act}-OTS).

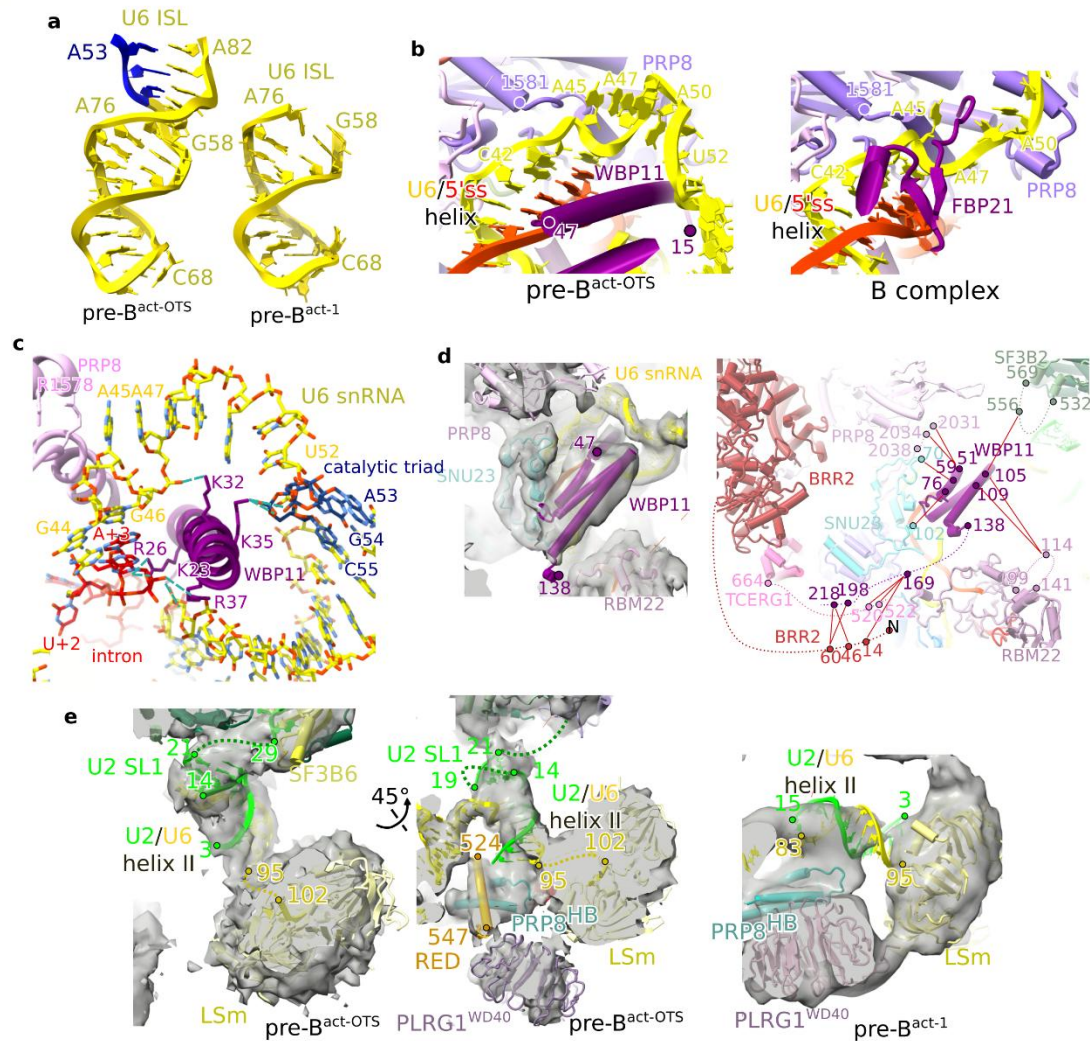
a, A representative micrograph after motion correction. **b**, Representative cryo-EM 2D class averages of the pre-B^{act}-OTS complex. Class averages typical of B complexes were absent. **c**, Cryo-EM computation sorting scheme. All major image-processing steps are depicted. For a more detailed explanation, see the EM data-processing section in

Methods. **d**, Representative regions of the EM density fit with a stick model. **e**, Local resolution estimation of the tri-snRNP region of the pre-B^{act-OTS} complex. **f**, Orientation distribution plot for the particles contributing to the reconstruction of the tri-snRNP core region. Red depicts a higher relative number of particles at a certain angle. **g**, Fourier shell correlation (FSC) values for the pre-B^{act-OTS} complex revealed a resolution of 3.12 Å for the tri-snRNP core and of 4.2 Å for the BRR2 region. **h**, Map versus model FSC curves for the tri-snRNP core and the BRR2 region.



Supplementary Fig. 3 Conformational state of PRP8 and BRR2 in pre-B^{act-OTS}

a, Schematic of the domain organization of BRR2 and PRP8. PWI, domain with a PWI tripeptide located within its N-terminal region; WH, winged helix; NC, the N-terminal helicase cassette; CC, the C-terminal helicase cassette; RT, reverse-transcriptase-like domain; HB, helical bundle; RH, RNase H-like domain; En, endonuclease-like domain; NTD, N-terminal domain; Jab1, Jab1/MPN domain. **b**, Structural rearrangements of the PRP8 and BRR2 region during the transition from B to pre-B^{act-OTS}. The BRR2 helicase, together with its associated PRP8^{Jab1}, is rotated clockwise by 100°. Concomitantly, PRP8^{RH} is rotated counterclockwise by 160°, forming a new contact with PRP8^{Jab1} in the pre-B^{act-OTS} complex. The PRP8-NTD domain is not shown. Orange arrow, β -hairpin of PRP8^{RH}; red arrow, Jab1²²³⁰⁻²²⁵⁰. **c**, In the pre-B^{act-OTS} complex, PRP8 adopts the half-closed conformation, in which the B-specific proteins (SNU23, MFAP1, PRP38 and UBL5) are bound to PRP8 in the same manner as in the pre-B^{act-1} complex¹. TCERG1 also adopts the same conformation as in the pre-B^{act-1} complex.

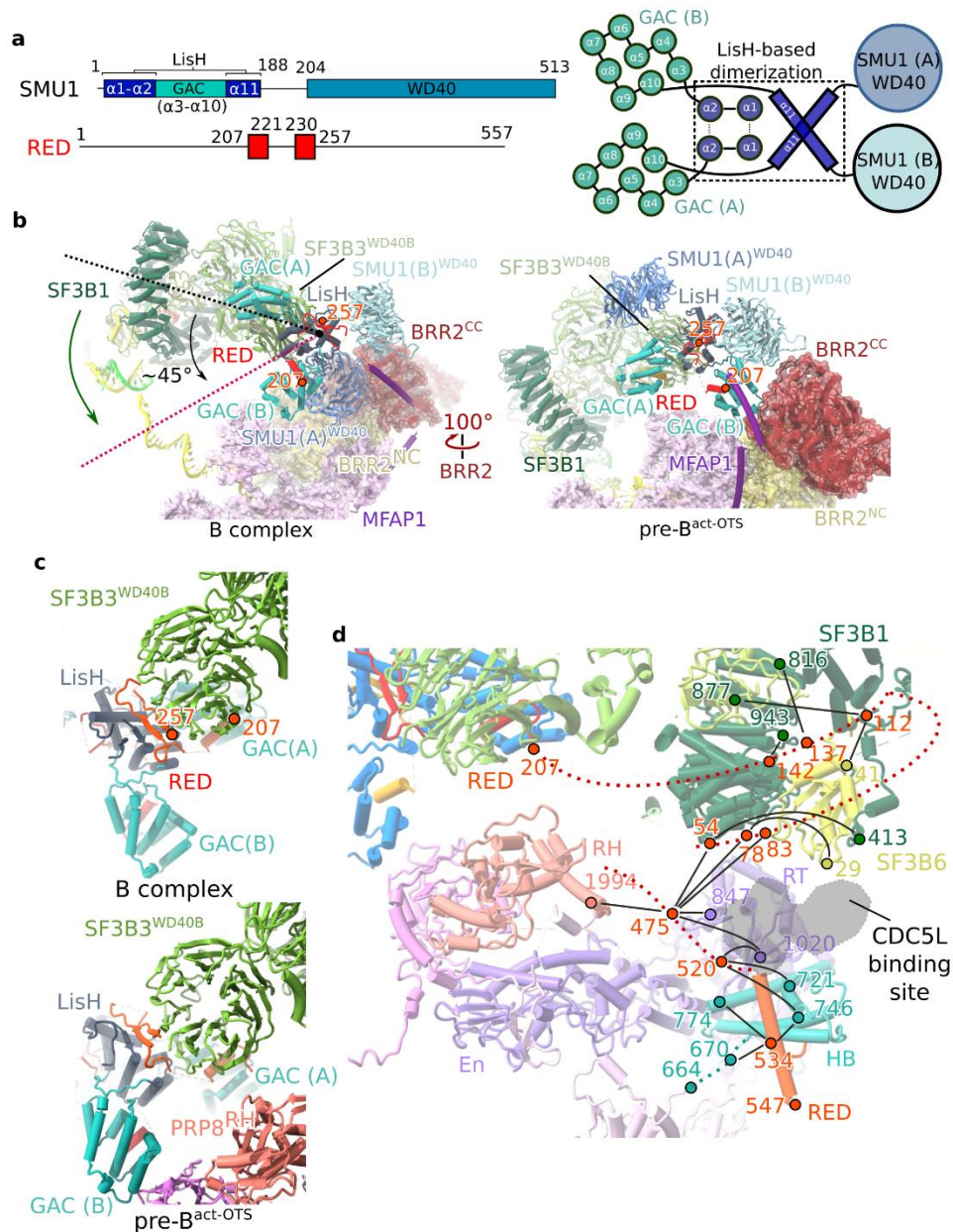


Supplementary Fig. 4 Structural organization of the U2 and U6 snRNAs in pre-B^{act-OTS}

a, Structural comparison of the U6 ISL of the pre-B^{act-OTS} and the pre-B^{act-1} complexes. The extended U6 ISL in the pre-B^{act-OTS} complex contains six additional base-pairings at its base, which includes the U6 catalytic triad A53-G54-C55 (colored in blue). We note that the base-pairings in the upper stem also differ slightly from those in pre-B^{act-1}.

b, In pre-B^{act-OTS}, the N-terminal α helix of the B-specific protein WBP11 is inserted into the recess created by the extended U6/5'ss helix, the U6 ISL and the connecting single-stranded U6 nts A42–U52. Comparison of the pre-B^{act-OTS} and the B complexes shows that the positions of WBP11 and FBP21 in the pre-B^{act-OTS} and B complexes, respectively, are mutually exclusive. **c**. Atomic interactions between the long α helix of

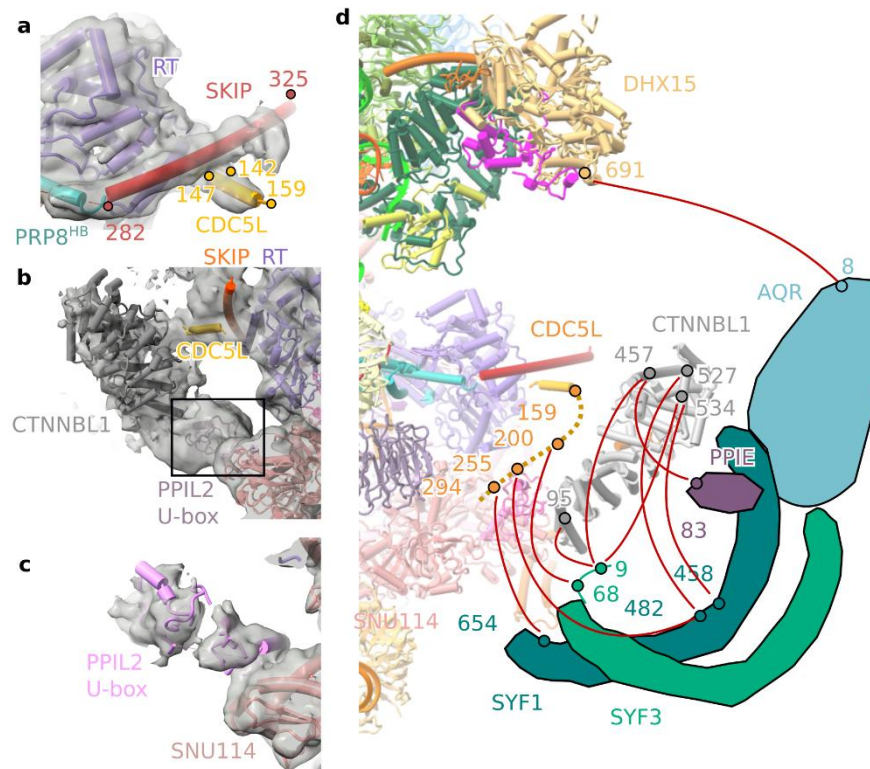
WBP11 and U6 snRNA and pre-mRNA. WBP11's α helix runs along the U6/5'ss helix, contacting the phosphate backbone of the 5'-terminal region of the pre-mRNA intron and with the backbone of U6 A47 (K32) and G54-G55 (K35). Dashed blue lines, hydrogen bonds. **d**, AF3 predicts a helical bundle for WBP11 (aa 48–138) directly downstream of its N-terminal α helix, which can be fitted into the EM density near SNU23, guided by our crosslinks. The crosslinks of WBP11 and its neighboring proteins are shown alongside. This helical bundle possibly facilitates the docking of WBP11's N-terminal α helix along the extended 5'ss/U6 RNA helix. Crosslinked residues are depicted as colored circles connected by red lines. Dashed, colored lines indicate the potential position of unstructured regions of the proteins. **e**, Fitting of U2/U6 helix II, U2 stem-loop1 (SL1), RED (aa 524–547) and U6 LSm into the EM density. In pre-B^{act-OTS}, the U2/U6 helix II is docked perpendicularly to PRP8^{HB}. However, in pre-B^{act-1}, it runs almost parallel to the HB domain. In addition, in pre-B^{act-OTS}, a short α helix of RED (aa 524–547) binds PRP8^{HB} at a site, mutually exclusive with PLRG1^{WD40} in pre-B^{act-1}. This may explain why in pre-B^{act-OTS}, the U6 LSm complex and PLRG1^{WD40} occupy an intermediate position compared to pre-B^{act-1} (see also Fig. 4c).



Supplementary Fig. 5 Structural comparison of SMU1/RED in B and pre-B^{act}-OTS complexes.

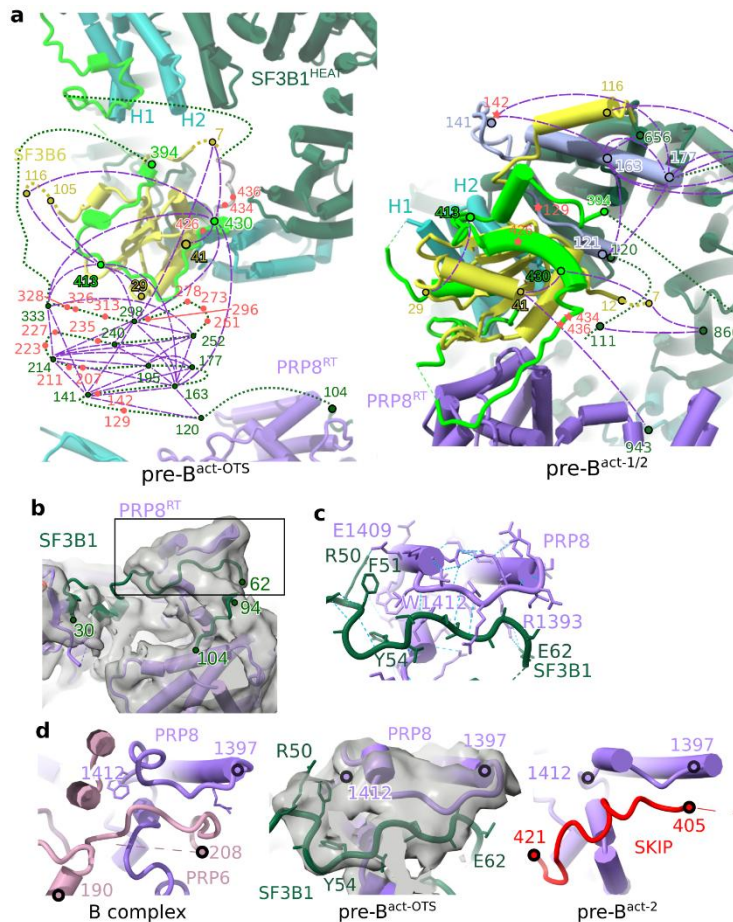
a, Schematic of the domain organization of SMU1 and RED, and schematic depiction of the SMU1 dimerization domain; see also Figure 4a. **b**, Coordinated movements of the U2 snRNP 5' domain, the SMU1-RED tetrameric complex and the BRR2 helicase during the B-to-pre-B^{act}-OTS transition. In the B complex, the LisH (Lis homology) and GAC^A (globular α-helical core A) domains of the SMU1 dimer form a bridge to the U2 snRNP 5' domain through interactions mediated by residues 207–257 of one RED copy (RED^{207–257}). The 5' domain of the U2 snRNP, together with the SMU1-RED

tetramer, rotates clockwise by ca 45° towards PRP8. Concurrently, BRR2 rotates by ca. 100°. As the interface between U2 SF3B3^{WD40} and SMU1/RED's LisH and GAC(A) domains is maintained during the B to pre-B^{act-OTS} transition, the movement of U2 is coupled with the docking of SMU1's GAC(B) domain to BRR2^{CC}, where it replaces the WD40 domain of SMU1(A), leading to a new interaction between SMU1^{GAC(B)} and BRR2. This contact between SMU1^{GAC(B)} and BRR2 probably prevents any further movement of the U2 snRNP (prior to CDK11-mediated phosphorylation of SF3B1^N). Thus, among others, SMU1-RED appears to orchestrate the movement of U2 snRNP and BRR2 during the B to pre-B^{act-OTS} transition and stabilize the structure of the latter complex. **c**, RED²⁰⁷⁻²⁵⁷ mediates a stable interaction between the U2 SF3B3^{WD40B} domain and SMU1. **d**, Tentative paths of the C- and N-terminal regions of RED in pre-B^{act-OTS}. RED's paths are based on crosslinks and are shown as dashed red-orange lines. Crosslinked amino acid residues are depicted as colored circles, connected by black lines. RED's C-terminal α helix (aa 524–547) binds PRP8^{HB}, and its upstream region extends towards PRP8^{RH}. PRP8 residues K847 and K1020 that crosslink to RED, are at the interface where CDC5L binds. The N-terminal region of RED, upstream of the SMU1-binding region, extends towards the SF3b complex and shows strong crosslinks to both SF3B1 and SF3B6. Its region closer to the N terminus may circle back to RED's C-terminal region, as indicated by the intramolecular crosslinks.



Supplementary Fig. 6 Structural organization of NTC and NTR proteins in pre-B^{act-OTS} complexes

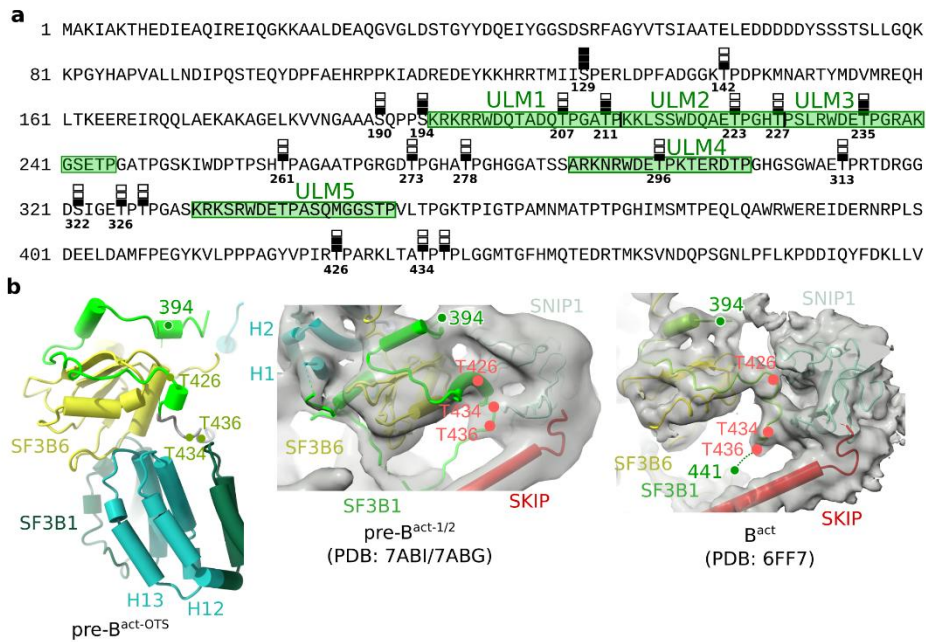
a, Based on the interactions observed in pre-B^{act-2}, a short α helix of CDC5L can be docked in the EM density near SKIP1. **b**, Based on its position in pre-B^{act-21}, CTNNBL1 can be docked within a flexible density at the corresponding position. **c**, The U-box domain of PPIL2 can be docked within the EM density, connecting SNU114 and CTNNBL1. **d**, Crosslinks between the IBC proteins, including SYF3, and the modeled parts of the pre-B^{act-OTS} complex suggest that the IBC proteins are loosely associated at the periphery of the complex. Crosslinked residues are depicted as colored circles connected by red lines.



Supplementary Fig. 7 Structural organization of the U2 SF3B1 N-terminal region in pre-B^{act}-OTS

a, Intramolecular crosslinks within SF3B1^N and between SF3B1^N and SF3B6 in pre-B^{act}-OTS and pre-B^{act}-1/2. In pre-B^{act}-OTS the SF3B6/SF3B1³⁹⁴⁻⁴³⁰ is associated with HRs 12–14 of SF3B1^{HEAT}. SF3B1^N (aa 394–430) that wraps around SF3B6 in the model is shown in light green. The central region of SF3B1^N (aa 104–394) is depicted as a virtual, meandering stippled green line located between the SF3B6/SF3B1³⁹⁴⁻⁴³⁰ module and PRP8^{RT}. (The N-terminal region of SF3B1^N (aa 30–104) is bound to PRP8^{RT}, see below.) Crosslinked lysine residues within the modeled part of SF3B1 (aa 394–430) are depicted in light green numbers (with black outline). Dark green numbers indicate the virtual position of crosslinked lysine residues within the central region of SF3B1^N. Yellow numbers indicate the likely positions of crosslinked lysine residues in SF3B6. Red numbers indicate the serine and threonine residues within SF3B1^N that were found to be phosphorylated in purified PM5-10 B^{act} complexes (see also Supplementary Fig.

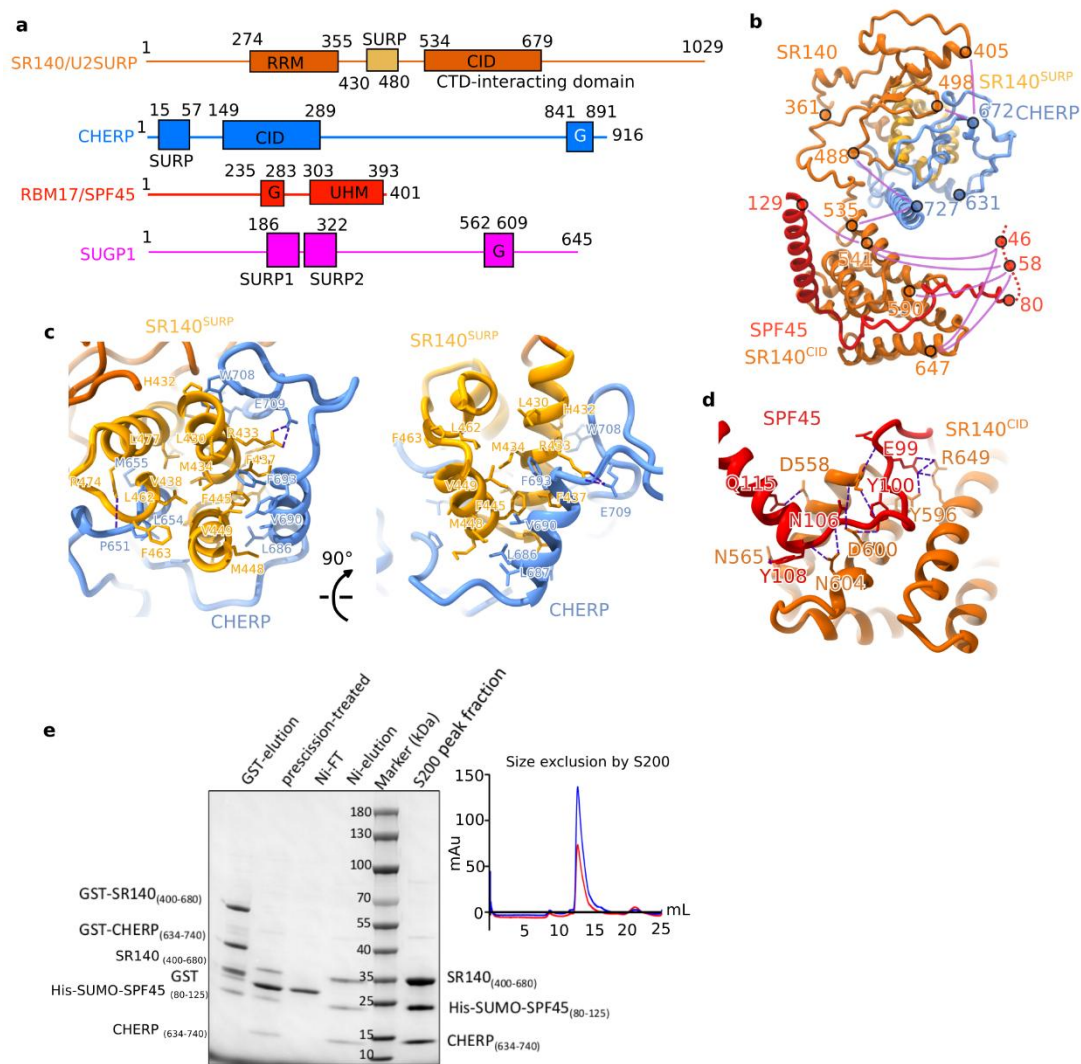
8a), and which thus become phosphorylated by CDK11 during the transition of pre-B^{act-OTS} to B^{act}. In contrast, in pre-B^{act1/2} complexes (pre-B^{act-2} complex is shown), after CDK11-mediated phosphorylation of SF3B1^N (phosphorylated amino acids are shown by red asterisks), significantly fewer crosslinks were identified (taken from Townsend *et al.*,¹). **b**, AF3-predicted high-confidence interaction between the N-terminal region of SF3B1 (aa 30–104) and PRP8^{RT} residues around 1400–1500, which can be fit into the corresponding density of the pre-B^{act-OTS} map. **c**, Depiction of AF3-predicted molecular contacts between SF3B1⁵⁰⁻⁶² and PRP8¹³⁹⁰⁻¹⁴³⁰. **d**, The PRP8^{RT} region comprising residues 1390 to 1430, sequentially interacts with PRP6¹⁹⁰⁻²⁰⁸ in the B complex⁵, with SF3B1⁵⁰⁻⁶² in pre-B^{act-OTS}, and with SKIP⁴⁰⁵⁻⁴²¹ in pre-B^{act-2} ¹.



Supplementary Fig 8. Mapping of phosphorylated serine and threonine residues within SF3B1's N-terminal region in B^{act}

a, Phosphorylation pattern of SF3B1^N in the PM5-10 B^{act} complex. Phosphor-sites were identified and semi-quantified by MS. The black boxes above the SF3B1 sequence depict the degree of phosphorylation as reported by PEAKS (% modified precursor area of total, schematically one filled box corresponds to 2–30%, two filled boxes to 30–60%, three filled boxes to >60%). ULMs 1–5⁶ are marked with light-green boxes. Except for ULM5, all other ULMs are either flanked or interspersed with phosphor-sites, raising the possibility that CDK11-mediated phosphorylation may also regulate the strength of interaction of UHM proteins, such as U2AF2, with their cognate ULMs.

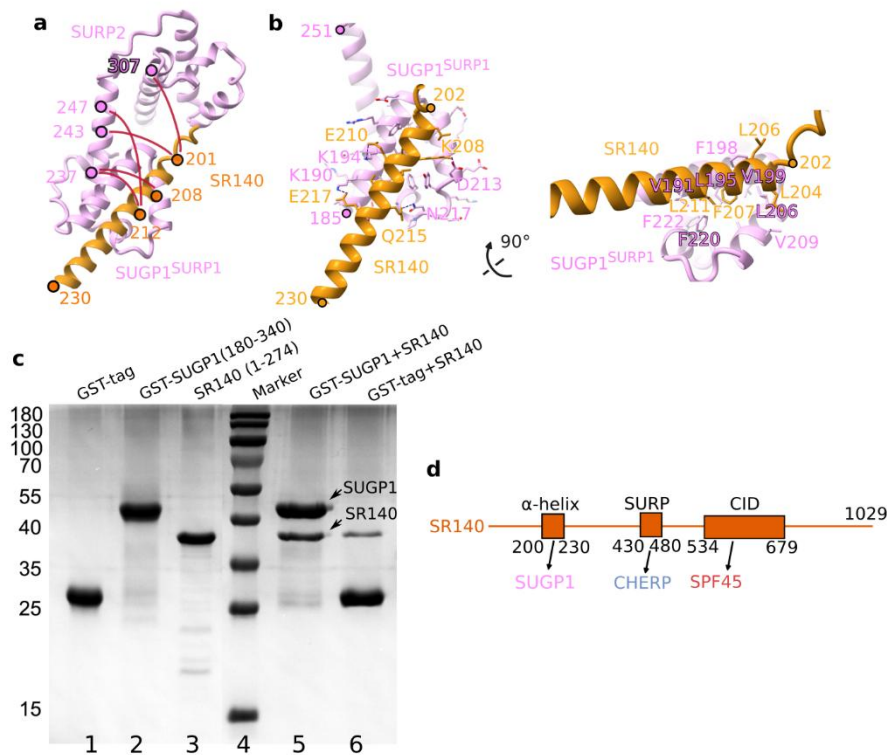
b, Potential role of the phosphorylation of threonine residues 426, 434 and 436, in the repositioning of the SF3B6/SF3B1³⁹⁴⁻⁴³⁰ module at SF3B1^{HEAT} during the pre-B^{act}-OTS to B^{act} transition. In pre-B^{act}-OTS, the SF3B6/SF3B1³⁹⁴⁻⁴³⁰ module is attached to the HRs 12–14 of SF3B1^{HEAT} with T426, T434 and T436 (yellow-green color; unphosphorylated state) being located close to/at the interface of HR12. We showed previously that in pre-B^{act}-1/2, the RES complex protein SNIP1 is also recruited to the spliceosome and interacts with the repositioned SF3B1/SF3B1³⁹⁴⁻⁴³⁰ module¹. Notably, the T426, T434 and T436 residues are now located at, or near to, the interface with SNIP1; they maintain their position also in the mature B^{act} complex.⁷



Supplementary Fig. 9 Characterization of interaction sites of the G-patch proteins CHERP and SPF45 at the SR140 protein within pre-B^{act-OTS}

a, Schematic of the domain organization of SR140/U2 SURP, CHERP, SPF45/RBM17, and SUGP1. RRM, RNA-binding motif; CTD, C-terminal domain; UHM, U2AF homology motif; G, G-patch motif, a ~45-residue glycine-rich motif that directly binds and activates DEAH helicases⁸⁻¹⁰. SURP, also known as SWAP (suppressor-of-white-apricot) domain. **b**, Crosslinks between SR140 and G-patch proteins CHERP and SPF45, respectively, as observed within the pre-B^{act-OTS} complex; consistently with AF3-mediated structure predictions (see **c**). Crosslinked residues are depicted as colored circles connected by pink lines. **c**, AF3-predicted molecular interactions between SR140 and CHERP. CHERP wraps around the SURP domain of SR140, and

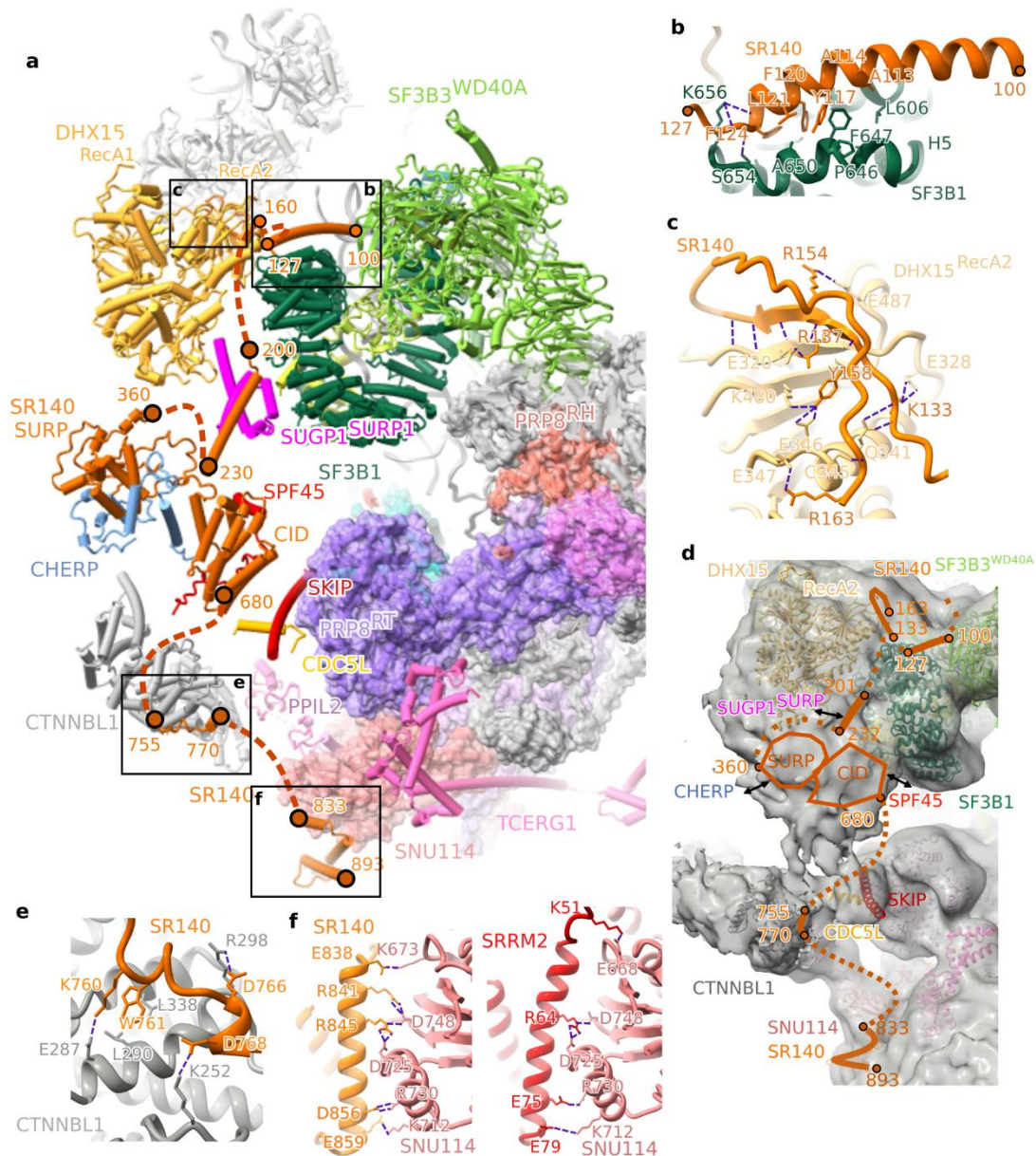
the interactions are mediated mainly by hydrophobic interactions. Dashed blue lines represent hydrogen bonds. **d**, AF3-predicted molecular interactions between part of SR140's CID domain and SPF45. **e**, Co-purification of GST-SR140 (aa 400–680), GST-CHERP (aa 634–740) and His-SUMO-SPF45 (aa 80–125). The plasmids encoding the three protein fragments were co-transformed into *Escherichia coli* strain BL2, and protein complexes, formed by the recombinantly expressed proteins, were purified and subsequently separated by SDS-PAGE as follows. Lane 1, protein complex as purified by GST pull-down; lane 2, protein fragments after removal of GST-tag by digestion with PreScission protease; lane 3, the PreScission-treated sample was then subjected to Ni-NTA resin (the flow-through contains only the GST-tag); lane 4, the eluate from the Ni-NTA resin contains the three protein fragments SR140 (aa 400–680), His-SUMO-SPF45 (aa 80–125), and CHERP (aa 634–740); lane 5, marker; lane 6, the eluate from Ni-NTA was subjected to Superdex 200 (S200) size exclusion column, and the peak fraction was visualized, showing that the three protein fragments form a stable complex and can be co-purified after two-step affinity purification followed by a size-exclusion purification. The purification profile by S200 is shown on the right. Blue and red lines indicate UV absorption at 280 nm and 260 nm, respectively. mAu, milli absorbance unit.



Supplementary Fig. 10 Characterization of the interaction site of the G-patch protein SUGP1 at the SR140 protein within pre-B^{act}-OTs

a, Crosslinks between SUGP1's SURP domains and an α helix of SR140 (aa 201–230). Crosslinked residues are depicted as colored circles connected by red lines. **b**, AF3 further predicts a high-confidence interaction between an α helix of SR140 (aa 200–230), just upstream of the N-terminal side of SR140's RRM domain, and the SURP domains of SUGP1 (SUGP1^{SURP}), which agrees well with our crosslinks, and with the interactions *in vitro* between recombinant proteins (see panel c, below). The two proteins interact mainly through salt bridges and hydrophobic interactions. We note that the predicted molecular interactions between the latter proteins are very similar to those observed by X-ray crystallography of the complex between an α helix of SF1 and the SURP-1 domain of SF3A1¹¹. **c**, Purified recombinant GST-tagged SUGP1 fragment (aa 180–340) and an N-terminal SR140 fragment (aa 1–274) were used for pull-down assays. The purified protein fragments are separated by SDS-PAGE. Lanes 1–3 show GST-tag, GST-SUGP1 (aa 180–340) and SR140 (aa 1–274), respectively; lane 4, marker proteins; lane 5, 20 μ g of purified GST-SUGP1 (aa 180–340) and 70 μ g of

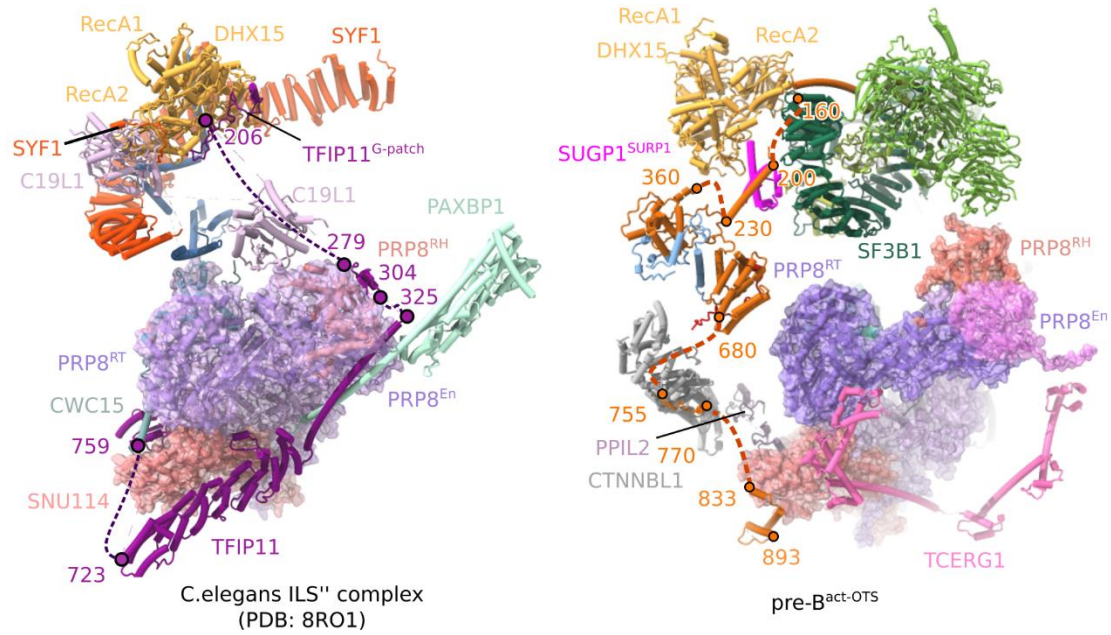
SR140 (aa 1–274) fragments were mixed and subjected to GST-affinity purification, indicating a direct interaction between the two protein fragments; lane 6, as a control, 20 μ g of GST protein and 7 μ g of SR140 (aa 1–274) fragments were mixed and subjected to GST-affinity purification, indicating that GST alone exhibits only background binding to SR140 (aa 1–274). **d**, SR140 interacts with SUGP1, CHERP and SPF45 through distinct, non-overlapping regions.



Supplementary Fig. 11 Recognition of peripheral surfaces of pre-B^{act}-OTS by SR140 and SR140's targeting of G-patch proteins to the spliceosome

a, Overview of the path of SR140 along the surface of the pre-B^{act}-OTS spliceosome. SR140 tethers DHX15 and SF3B1 through its N-terminal region. Its middle region interacts with the G-patch proteins SUGP1, CHERP and SPF45 (see also Supplementary Fig. 9, 10). The C-terminal region extends toward CTNNBL1 and finally reaches SNU114. Because the intron part of PM5-10 downstream of the BS/U2 helix is too short to extend beyond the inner circle of the HEAT domain, DHX15 cannot bind to intron RNA in PM5-10 pre-B^{act}-OTS. Frames, labelled b-f, indicate regions of

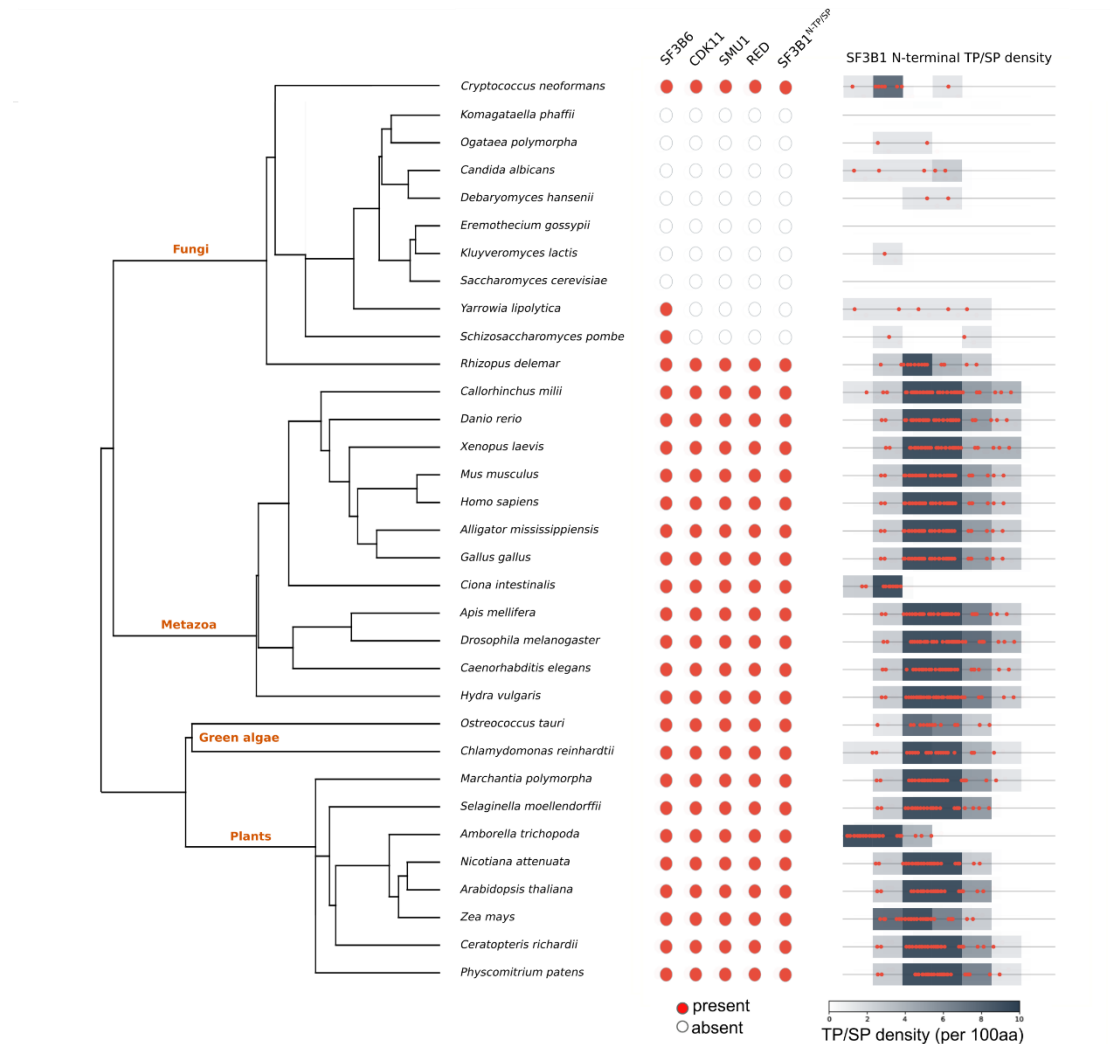
PM5-10 pre-B^{act-OTS}, the structures of which are shown in more detail in b–f. **b**, AF3-predicted molecular interactions between an N-terminal α helix of SR140 (aa 100–127) and HR5 of SF3B1^{HEAT}. Dashed blue lines, hydrogen bonds. **c**, AF3-predicted molecular interactions between an N-terminal region of SR140 (aa 133–163) and DHX15. **d**, A flexible density bridge, potentially containing the globular domains of the complex between SR140 and the G-patch proteins CHERP, SPF45 and SUGP1 (as mapped in Supplementary Fig. 9, 10), connects the SF3B1/DHX15 complex and the region comprising CTNNBL1, SKIP and CDC5L. The map is low-pass filtered to 30 Å. **e**, AF3-predicted molecular interactions between the HEAT domain of CTNNBL1 and the region of SR140 comprising aa 755–770; consistent with our crosslinks (see Fig 7c). **f**, AF3-predicted molecular interactions between a C-terminal region of SR140 (aa 838–859) and domain IV of SNU114 in pre-B^{act-OTS}. In B^{act} complexes, a region of SRRM2 (aa 51–100) binds at the domain IV of SNU114 through very similar interactions as SR140 in pre-B^{act-OTS}.



Supplementary Fig. 12 Comparison of the extended, but significantly different, paths of SR140 at the human pre-B^{act-OTS} complex with that of TFIP11/PAXBP1 at the *C. elegans* ILS'' complex.

In the *C. elegans* ILS'' complex, TFIP11 and PAXBP1 bind to the peripheral of the spliceosome, and the G-patch region of TFIP11 activates DHX15 for spliceosome disassembly¹². The RecA2 domain of DHX15 packs against a C-terminal region of SYF1, which together with TFIP11, orients DHX15 to engage its U6 snRNA target, priming it to dissociate the U6/U2 catalytic core RNP¹². In the pre-B^{act-OTS} complex, SR140 also binds to the outer surface of the spliceosome, but takes a significantly different path. Unlike TFIP11, SR140 does not contain a G-patch region, but it is associated with several G-patch-containing proteins, which are probably recruited as a pre-formed complex^{13,14}, consistent with our crosslinks between those G-patch proteins and SR140 in pre-B^{act-OTS} (see Supplementary Figs. 9, 10). Similarly to SYF1 in *C. elegans* ILS''¹², the N-terminal region of SR140 packs against DHX15's RecA2 domain, likely ensuring its correct orientation. However, in pre-B^{act-OTS}, DHX15 is positioned adjacent to the intron exit site near SF3B1 HRs 5–7 (see also Fig. 6f). From this location, it is probably poised to bind the intron near the 3' splice site/exon junction, dismantling a remodeled aberrant spliceosome. In contrast, in the ILS complex, DHX15 has accommodated the 3' end of U6 snRNA in its RNA-binding channel¹², poised to

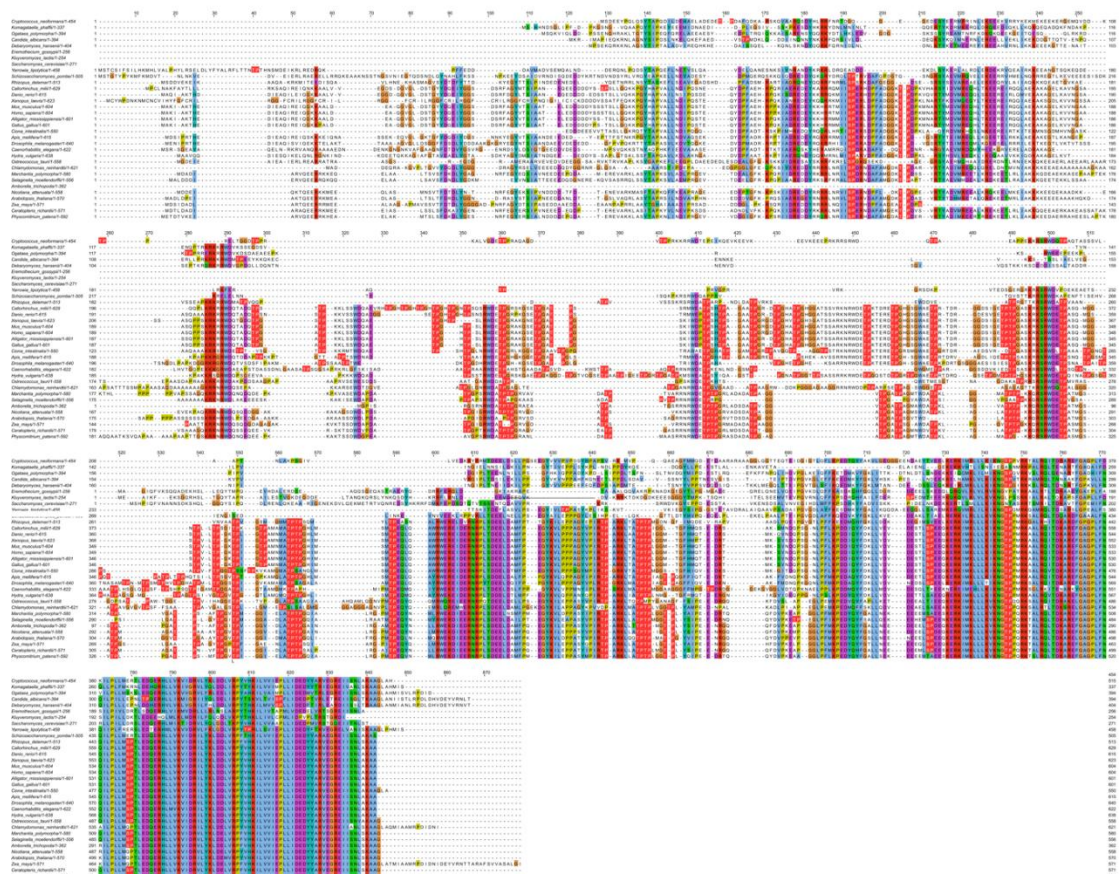
dissociate the U6/U2 catalytic core RNP. For clarity, only the structural elements of the *C. elegans* ILS'' and pre-B^{act-OTS} complexes are shown, which are relevant to the discussion here.



Supplementary Fig. 13. Co-evolution of the SF3B6-CDK11-SF3B1^{N-TP/SP} axis with SMU1/RED

Phylogenetic tree showing the evolutionary divergence of selected species across the major eukaryotic kingdoms. Distribution of SF3B6, CDK11, a TP/SP dipeptide-rich SF3B1 N-terminal region (SF3B1^{N-TP/SP}), and the B complex proteins SMU1/RED are shown in the central column. Red circles indicate detected homologs, and empty circles indicate their absence. The data reveal a tight evolutionary coupling of these factors across metazoan and plant lineages. Although a putative CDK11 homolog can be identified in *S. pombe*, it lacks the critical N-terminal region that is conserved in metazoans and plants, and CDK11's function is reportedly unrelated to splicing¹⁵. Notably, *C. neoformans*, which is intron-rich and, compared with *S. cerevisiae*, contains more degenerate splicing signatures¹⁶, stands out as a unique fungal exception

possessing SF3B6, a putative CDK11 homolog, an SF3B1 N-terminal region with numerous TP/SP dipeptides and SMU1/RED proteins. TP/SP dipeptide repeat density in the SF3B1 N-terminus is quantified, and the density of TP/SP motifs per 100 amino acids (aa) is indicated by the gray-scale gradient. Red data points represent individual TP/SP motifs that occurs within the SF3B1 N-terminal region. Note that species lacking the SF3B6-CDK11-SMU1-RED axis generally exhibit significantly reduced or absent TP/SP clusters within the SF3B1 N-terminus.



Supplementary Fig. 14. Evolutionary conservation of the SF3B1 N-terminal region.

Multiple sequence alignment of the N-terminal region of the SF3B1 protein across diverse species. Residues are colored according to chemical properties. TP/SP sites are highlighted by light-red background.

Cryo-EM data collection, refinement and validation statistics

	pre-B ^{act} -OTS (EMDB-53843) (PDB 9R8V)	pre-B ^{act} -OTS core (EMDB-53554) (PDB 9R3D)
Data collection and processing		
Magnification	103,700	103, 700
Voltage (kV)	300	300
Electron exposure (e-/Å ²)	39	39
Defocus range (μm)	1-2.5	1-2.5
Pixel size (Å)	1.35	1.35
Symmetry imposed	C1	C1
Initial particle images (no.)	1.5 million	1.5 million
Final particle images (no.)	133,575	35,015
Map resolution (Å)	5.9	3.12
FSC threshold	0.143	0.143
Map resolution range (Å)	5.4-30	2.9-4.5
Refinement		
Initial model used (PDB code)	7ABI	7ABI
Model resolution (Å)	--	3.4
FSC threshold		0.5
Model resolution range (Å)	--	3.2-4
Map sharpening <i>B</i> factor (Å ²)	--	-60
Model composition	--	
Non-hydrogen atoms		41,826
Protein residues		5,145
Ligands		0
<i>B</i> factors (Å ²)	--	
Protein		85.11
Ligand		184.48
R.m.s. deviations	--	
Bond lengths (Å)		0.002
Bond angles (°)		0.486
Validation	--	
MolProbity score		1.61
Clashscore		9.30
Poor rotamers (%)		0
Ramachandran plot	--	
Favored (%)		97.43
Allowed (%)		2.57
Disallowed (%)		0

Supplementary Table 1. Cryo-EM data collection, refinement and validation statistics

References:

- 1 Townsend, C. *et al.* Mechanism of protein-guided folding of the active site U2/U6 RNA during spliceosome activation. *Science* **370** (2020). <https://doi.org/10.1126/science.abc3753>
- 2 Loyer, P. *et al.* Characterization of cyclin L1 and L2 interactions with CDK11 and splicing factors: influence of cyclin L isoforms on splice site selection. *J Biol Chem* **283**, 7721-7732 (2008). <https://doi.org/10.1074/jbc.M708188200>
- 3 Roscigno, R. F. & Garcia-Blanco, M. A. SR proteins escort the U4/U6.U5 tri-snRNP to the spliceosome. *RNA* **1**, 692-706 (1995).
- 4 Wang, C. *et al.* CDK11 requires a critical activator SAP30BP to regulate pre-mRNA splicing. *EMBO J* **42**, e114051 (2023). <https://doi.org/10.15252/embj.2023114051>
- 5 Zhang, Z. *et al.* Cryo-EM analyses of dimerized spliceosomes provide new insights into the functions of B complex proteins. *EMBO J* **43**, 1065-1088 (2024). <https://doi.org/10.1038/s44318-024-00052-1>
- 6 Loerch, S. *et al.* The pre-mRNA splicing and transcription factor Tat-SF1 is a functional partner of the spliceosome SF3b1 subunit via a U2AF homology motif interface. *J Biol Chem* **294**, 2892-2902 (2019). <https://doi.org/10.1074/jbc.RA118.006764>
- 7 Haselbach, D. *et al.* Structure and Conformational Dynamics of the Human Spliceosomal B(act) Complex. *Cell* **172**, 454-464.e411 (2018). <https://doi.org/10.1016/j.cell.2018.01.010>
- 8 Aravind, L. & Koonin, E. V. G-patch: a new conserved domain in eukaryotic RNA-processing proteins and type D retroviral polyproteins. *Trends Biochem Sci* **24**, 342-344 (1999). [https://doi.org/10.1016/s0968-0004\(99\)01437-1](https://doi.org/10.1016/s0968-0004(99)01437-1)
- 9 Bohnsack, K. E., Ficner, R., Bohnsack, M. T. & Jonas, S. Regulation of DEAH-box RNA helicases by G-patch proteins. *Biol Chem* **402**, 561-579 (2021). <https://doi.org/10.1515/hsz-2020-0338>
- 10 Studer, M. K., Ivanović, L., Weber, M. E., Marti, S. & Jonas, S. Structural basis for DEAH-helicase activation by G-patch proteins. *Proc Natl Acad Sci U S A* **117**, 7159-7170 (2020). <https://doi.org/10.1073/pnas.1913880117>
- 11 Nameki, N. *et al.* Structural basis for the interaction between the first SURP domain of the SF3A1 subunit in U2 snRNP and the human splicing factor SF1. *Protein Sci* **31**, e4437 (2022). <https://doi.org/10.1002/pro.4437>
- 12 Vorländer, M. K. *et al.* Mechanism for the initiation of spliceosome disassembly. *Nature* **632**, 443-450 (2024). <https://doi.org/10.1038/s41586-024-07741-1>
- 13 Martín, E., Vivori, C., Rogalska, M., Herrero-Vicente, J. & Valcárcel, J. Alternative splicing regulation of cell-cycle genes by SPF45/SR140/CHERP complex controls cell proliferation. *RNA* **27**, 1557-1576 (2021). <https://doi.org/10.1261/rna.078935.121>
- 14 De Maio, A. *et al.* RBM17 Interacts with U2SURP and CHERP to Regulate Expression and Splicing of RNA-Processing Proteins. *Cell Rep* **25**, 726-736.e727 (2018). <https://doi.org/10.1016/j.celrep.2018.09.041>
- 15 Drogat, J. *et al.* Cdk11-cyclinL controls the assembly of the RNA polymerase II mediator complex. *Cell Rep* **2**, 1068-1076 (2012). <https://doi.org/10.1016/j.celrep.2012.09.027>
- 16 Sales-Lee, J. *et al.* Coupling of spliceosome complexity to intron diversity. *Curr Biol* **31**, 4898-4910.e4894 (2021). <https://doi.org/10.1016/j.cub.2021.09.004>