Appendix file

Protein localisation by electron microscopy reveals the architecture of the yeast spliceosomal B complex

Norbert Rigo[†], Chengfu Sun[†], Patrizia Fabrizio, Berthold Kastner* and Reinhard Lührmann*

Department of Cellular Biochemistry, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany

* To whom correspondence should be addressed

Tel. +49-551-201-1405; Fax +49-551-201-1197; eMail reinhard.luehrmann@mpibpc.mpg.de

Tel. +49-551-201-1313; Fax +49-551-201-1197; eMail bkastne@mpibpc.mpg.de

[†] Authors contributed equally

Supplementary Methods

2 Appendix figures

1 Appendix table

Working title: Architecture of the yeast spliceosomal B complex

Appendix Supplementary Methods

Yeast strain growth tests

A pre-culture of each yeast strain was grown overnight and the OD_{600} was adjusted to 5.0. A dilution series from 10^0 to 10^{-5} was prepared in 96-well plates using 100 µl (0.5 OD) of the starting culture. These dilutions were then plated out onto YPD agar plates and incubated at four different temperatures (16 °C, 25 °C, 30 °C and 37 °C) for up to 6 days.

Yeast extract splicing tests

Uniformally [^{32}P]-labelled Actin pre-mRNA was transcribed *in vitro* with T7 RNA polymerase. The pre-mRNA then was incubated with yeast whole-cell extract prepared according to (Gottschalk *et al*, 1999) under splicing conditions (60 mM K_2HPO_4/KH_2PO_4 pH 7.4; 3% w/v PEG-8000; 2 mM ATP; 2 mM spermidine; 2.5 mM $MgCl_2$; 40% v/v yeast extract; 0.2 nM actin pre-mRNA). Aliquots were taken at 0 min and 30 min and digested with proteinase K from *Tritirachium album* (Sigma-Aldrich). The RNA was precipitated with ethanol, resuspended in sample buffer and separated by denaturing PAGE on an 8% (29:1) polyacrylamide gel containing 8.3 M urea. The gel was dried and the RNA was visualised by autoradiography.

Mass spectrometry

In order to analyse the proteome of the purified yeast B complexes that were used for locating spliceosomal proteins we performed LC-MSMS analysis. The spliceosomes were purified as described above with the following adjustments: No glutaraldehyde was present during density-gradient ultracentrifugation and the glycerol gradients were fractionated from the top. After fractionation the proteins in the gradient fractions were precipitated using 2.5 volumes of ethanol, resuspended in sample buffer and separated by SDS PAGE. The gel was stained with Coomassie Brilliant Blue and the entire lane was cut into slices. The proteins were subjected to in-gel digestion with trypsin according to (Shevchenko *et al*, 1996), the peptides were extracted and analysed in an LC-coupled ESI Q-ToF (Q-ToF Ultima, Waters) mass spectrometer. Using the search engine Mascot and the NCBI non-redundant database the proteins could be identified using the fragment spectra of the sequenced peptides (compare Supplementary Methods of (Fabrizio *et al*, 2009)).

Image adjustments for optimum visibility of the protein localisation results

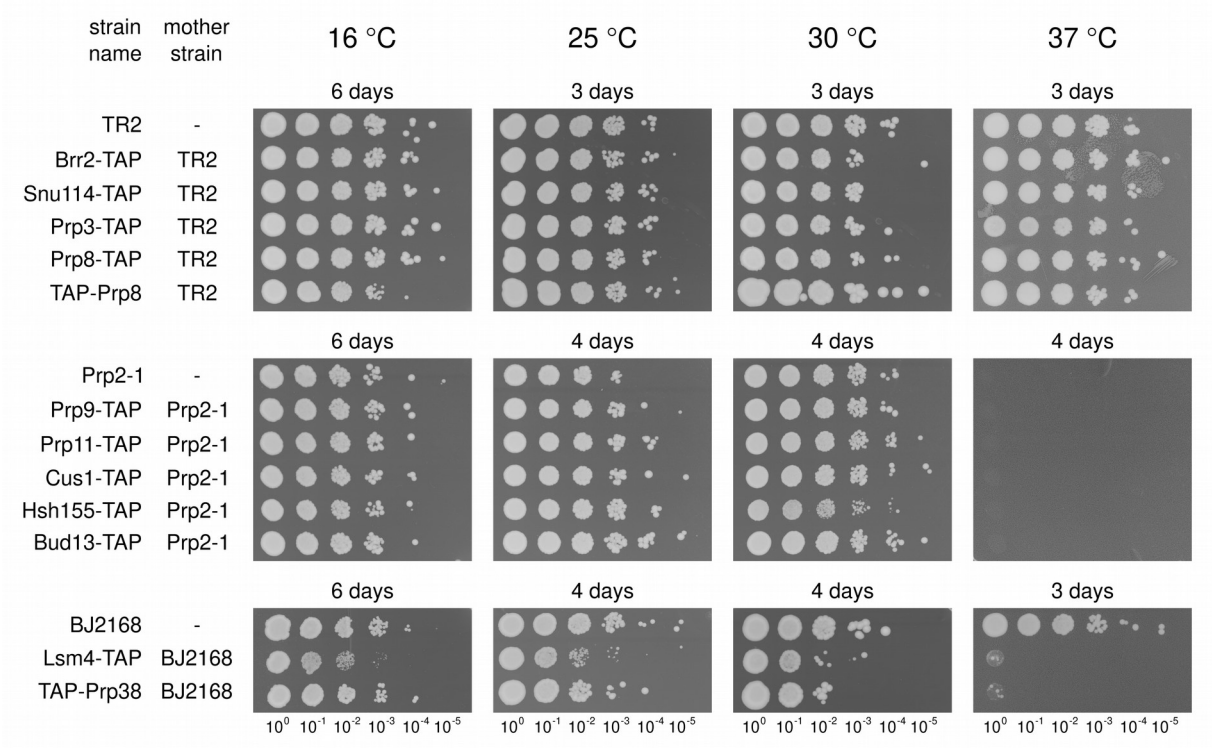
Overlaying of all 100 potential TAP tag location areas (see Methods section) was performed using Inkscape (https://inkscape.org). The schematic drawing of the B complex main view (together with the area inside a 14 nm radius of the second particle outline) for all 100 dimers was aligned according to the main-view particle. Then the areas were coloured black (R/G/B 0/0/0) with an alpha channel value of 4 to render the areas transparent. This resulted in the best possible display for the fraction where all areas overlapped on a PC screen and the analysis could be performed. However, printing this image resulted in poor quality. To obtain a result suitable for printing, the alpha channel was set to a value of 1. The image was then exported to a file in the format 'portable network graphics' (png) and all further processing steps were performed using the GNU Image Manipulation Program (GIMP; http://www.gimp.org/). To increase the brightness difference, and thus the visibility of the borders between the segments where n areas and n-1 areas overlap, the image was duplicated to create 4 layers overlying each other. The three top layers were set to an opacity of 33% and the layer mode 'Burn' was used (program documentation: "Burn mode inverts the pixel value of the lower layer, multiplies it by 256, divides that by one plus the pixel value of the upper layer, then inverts the result. … somewhat similar to "Multiply" mode."). After that, the colour curve of the image was adjusted to give optimum contrast, i.e. to attain the best possible brightness difference between the grey value resulting from 100 overlapping areas and the grey values resulting from fewer areas overlapping.

References

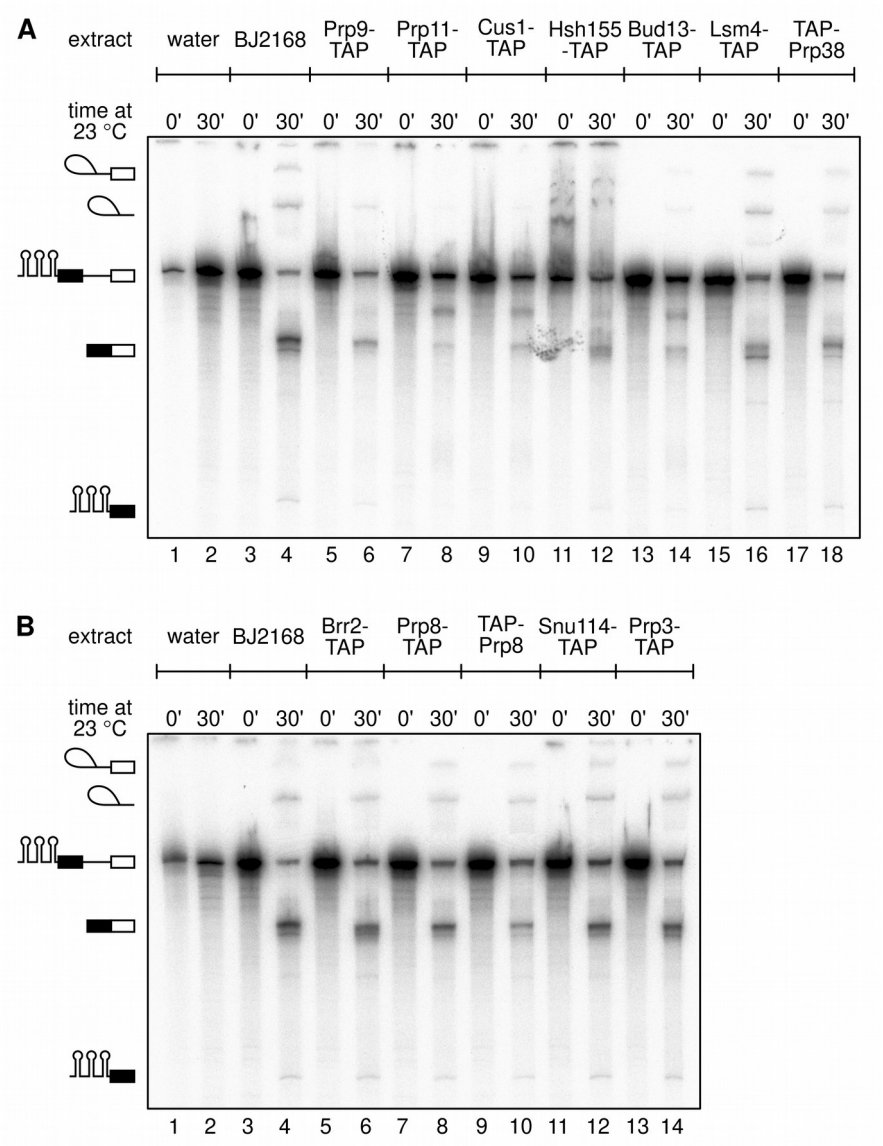
Fabrizio P, Dannenberg J, Dube P, Kastner B, Stark H, Urlaub H & Lührmann R (2009) The Evolutionarily Conserved Core Design of the Catalytic Activation Step of the Yeast Spliceosome. Mol. Cell 36: 593–608

Gottschalk A, Neubauer G, Banroques J, Mann M, Lührmann R & Fabrizio P (1999) Identification by mass spectrometry and functional analysis of novel proteins of the yeast [U4/U6.U5] tri-snRNP. EMBO J. 18: 4535–4548

Shevchenko A, Wilm M, Vorm O & Mann M (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. Anal. Chem. 68: 850–858



Appendix Fig S1: Growth tests of *S. cerevisiae* strains expressing TAP-tagged spliceosomal proteins.
The strains expressing the given TAP-tagged proteins were plated on YPD agar plates at the dilutions given at the bottom (the dilution 10^0 corresponds to 0.5 OD_{600}) and incubated at the respective temperature for the given time. For construction of the TAP-tagged strains three different strains were used: TR2, BJ2168 and Prp2-1; the latter showed a temperature-sensitive phenotype. To determine a potential growth defect, all strains expressing a TAP-tagged protein were plated out and compared with the corresponding untagged strain. For splicing-active extract preparation the yeast cells of the strains having TR2 and BJ2168 background were grown at 30 °C and the strains having a Prp2-1 background at 25 °C.
Addition of the TAP tag to the N terminus of Prp38 resulted in a temperature-sensitive phenotype at 37 °C. Adding the TAP tag to the C terminus of Lsm4 resulted in a temperature-sensitive phenotype and a generally slower growth at lower temperatures. Nevertheless the yeast cells expressing Lsm4-TAP grow comparatively well at 30 °C. All other strains show no significant difference in their growth when compared with their respective mother strains.



Appendix Fig S2: Splicing tests of whole-cell extracts of *S.-cerevisiae* expressing TAP-tagged spliceosomal proteins.

Yeast extracts prepared from yeast strains expressing the respective TAP-tagged protein were incubated with [^{32}P]-labelled Actin pre-mRNA carrying three MS2 aptamers at the 5' end under splicing conditions. Samples were withdrawn at 0 min and 30 min and they were analysed by denaturing PAGE and autoradiography. The identity of the bands is given on the left as schematic drawings. The MS2 aptamers are shown as three stem-loops, the intron depicted as a line, the 5' exon as a black box and the 3' exon as a white box. The tagged protein present in the extract used is named above each lane and the lanes of each gel are numbered in the bottom.

A) Splicing tests of the yeast extracts from strains that were constructed by using either BJ2168 (TAP-Prp38 and Lsm4-TAP) or Prp2-1 (Prp9-TAP, Prp11-TAP, Cus1-TAP and Hsh155-TAP).

B) Splicing test of the yeast extract from strains that were constructed using TR2.

Appendix Table S1: Mass-spectrometric analysis of purified monomeric and dimeric yeast spliceosomal B complexes. Spliceosomal B complexes harbouring one TAP-tagged protein were assembled, anti-TAP antibodies were added and the complexes were purified. The protein content of the gradient fractions of the 70S region (B complex dimers) and the 40S region (B complex monomers) were separated by SDS PAGE and subjected to LC-MSMS analysis. The numbers given for the spliceosomal proteins represent the absolute number of peptides sequenced. In addition the molecular weight (MW) for the *S. cerevisiae* proteins are given. The protein carrying the TAP tag in the respective B complex dimer is indicated at the top of each column. Untagged spliceosomes did not dimerise and only B complex monomers were analysed as a control. No step 2 factors and almost no splicing factors known to be present in stages later than the B complex were found.

protein name	MW	number of peptides sequenced: monomer untagged	dimer Brr2-TAP	dimer Prp3-TAP
Sm proteins				
SMB1	22 kDa	7	5	4
SMD1	16 kDa	2	2	4
SMD2	13 kDa	4	3	4
SMD3	11 kDa	2	2	2
SME	10 kDa	3	1	1
SMF	10 kDa	0	0	1
SMG	9 kDa	1	2	1
U1 snRNP proteins				
PRP39	75 kDa	12	9	9
SNU71	71 kDa	10	10	9
PRP40	69 kDa	9	6	6
PRP42	65 kDa	6	4	3
NAM8	57 kDa	4	2	2
SNU56	57 kDa	7	6	5
SNP1	34 kDa	5	4	4
MUD1	34 kDa	7	5	2
LUC7	30 kDa	4	6	2
YHC1	27 kDa	2	0	2
U2 snRNP proteins				
RSE1	154 kDa	47	36	35
HSH155	110 kDa	27	20	16
PRP9	63 kDa	14	9	9
CUS1	50 kDa	20	16	15
PRP21	33 kDa	8	5	7
PRP11	30 kDa	6	7	6
LEA1	27 kDa	9	8	9
HSH49	25 kDa	8	6	4
MSL1	13 kDa	5	3	3
RDS3	12 kDa	4	2	2
YSF3	10 kDa	1	1	3
U5 snRNP proteins				
PRP8	280 kDa	73	64	55
BRR2	246 kDa	84	61	63
SNU114	114 kDa	33	26	31
PRP6	104 kDa	42	31	31
PRP28	67 kDa	0	1	1
LIN1	40 kDa	0	0	0
DIB1	17 kDa	6	5	5
U4/U6 snRNP proteins				
PRP31	56 kDa	18	13	11
PRP3	56 kDa	10	9	7
PRP4	52 kDa	16	10	10
SNU13	14 kDa	1	2	1
U5·U4/U6 snRNP proteins				
SNU66	66 kDa	24	19	20
SAD1	52 kDa	0	0	0
SPP381	34 kDa	5	6	7
PRP38	28 kDa	6	2	4
SNU23	23 kDa	2	4	4
Disassembly proteins				
PRP43	88 kDa	13	17	13
SPP382	83 kDa	0	0	0
NTR2	37 kDa	0	0	0

protein name	MW	number of peptides sequenced: monomer untagged	dimer Brr2-TAP	dimer Prp3-TAP
Lsm proteins				
LSM4	21 kDa	2	1	2
LSM7	13 kDa	1	2	2
LSM8	12 kDa	2	1	1
LSM2	11 kDa	6	5	4
LSM5	10 kDa	0	0	0
LSM3	10 kDa	0	0	0
LSM6	9 kDa	1	1	0
RES complex				
BUD13	30 kDa	8	6	5
PML1	24 kDa	5	3	2
IST3	17 kDa	4	3	2
NTC proteins				
SYF1	100 kDa	10	9	8
CLF1	82 kDa	10	7	10
CEF1	68 kDa	9	5	7
PRP19	57 kDa	10	8	7
ISY1	28 kDa	5	3	2
SYF2	25 kDa	5	5	4
SNT309	21 kDa	3	2	0
NTC20	16 kDa	0	2	0
NTC-related proteins				
PRP46	51 kDa	10	9	4
PRP45	42 kDa	8	9	4
ECM2	41 kDa	1	1	0
CWC2	38 kDa	0	0	0
CWC15	20 kDa	0	1	0
BUD31	18 kDa	3	1	1
Early splicing factors				
MUD2	60 kDa	2	3	2
PRP5	96 kDa	5	0	0
URN1	54 kDa	11	10	6
Known splicing factors				
PRP2	100 kDa	0	0	0
SPP2	21 kDa	0	0	0
YJU2	32 kDa	2	2	0
CWC21	16 kDa	4	5	5
CWC22	67 kDa	0	0	0
CWC24	38 kDa	0	0	0
CWC27	35 kDa	0	0	0
CWC23	33 kDa	0	0	0
CWC25	20 kDa	0	0	0
Step 2 proteins				
PRP17	52 kDa	0	0	0
PRP22	130 kDa	0	0	0
PRP16	122 kDa	0	0	0
SLU7	45 kDa	0	0	0
PRP18	28 kDa	0	0	0
CBP proteins				
STO1	100 kDa	4	13	7
CBC2	24 kDa	1	1	2