

# Dynamics and evolutionary conservation of B complex protein recruitment during spliceosome activation

Xingyang Fu<sup>1,2</sup> and Aaron A. Hoskins<sup>1,3,\*</sup>

<sup>1</sup>Department of Chemistry, University of Wisconsin-Madison, Madison, WI 53706, United States

<sup>2</sup>Present address: Department of Neuroscience, Yale University, New Haven, CT 06520, United States

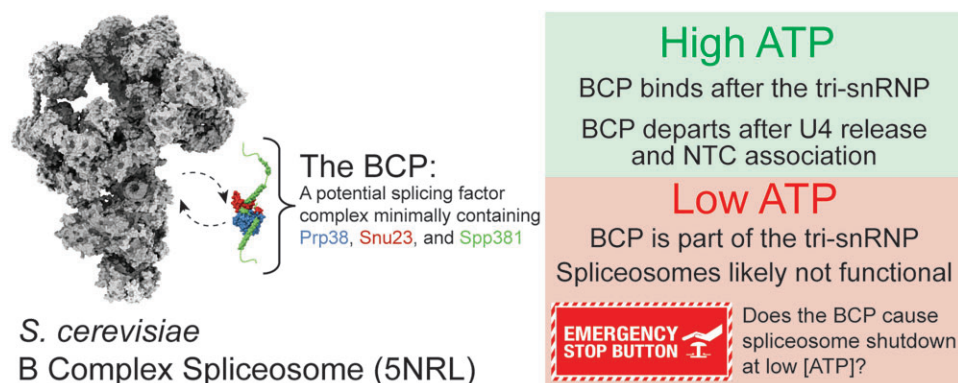
<sup>3</sup>Department of Biochemistry, University of Wisconsin-Madison, Madison, WI 53706, United States

\*To whom correspondence should be addressed. Email: [ahoskins@wisc.edu](mailto:ahoskins@wisc.edu)

## Abstract

Spliceosome assembly and catalytic activation involve dozens of protein and snRNA binding and unbinding events. The B-complex specific proteins (Prp38, Snu23, and Spp381) have critical roles in stabilizing the spliceosome during rearrangements essential for activation. While these proteins are conserved, different mechanisms have been proposed for their recruitment to spliceosomes. To visualize recruitment directly, we used Colocalization Single Molecule Spectroscopy (CoSMoS) to study the dynamics of Prp38, Snu23, and Spp381 during splicing in real time. These proteins can bind to and release from spliceosomes simultaneously and are likely associated with one another. We designate the assembly minimally containing Prp38, Snu23, and Spp381 as a potential B complex protein (BCP) subcomplex of the spliceosome. Under splicing conditions, BCP proteins associate with pre-mRNA after tri-snRNP binding. BCP protein release predominantly occurs after U4 snRNP dissociation and after NineTeen Complex (NTC) association. Under low concentrations of ATP, BCP proteins preassociate with the tri-snRNP resulting in their simultaneous binding to pre-mRNA. Together, our results reveal that BCP protein recruitment to the spliceosome is conserved between *Saccharomyces cerevisiae* and humans. The binding of BCP proteins to the tri-snRNP when ATP is limiting may result in the formation of unproductive complexes that could be used to regulate splicing.

## Graphical abstract



## Introduction

RNA splicing is the process of removing introns from precursor messenger RNAs (pre-mRNAs) concomitant with ligation of the flanking exons. The process is mediated by a megadalton ribonucleoprotein complex, the spliceosome. The components of spliceosomes can be broadly categorized as small nuclear ribonucleoprotein complexes (snRNPs; U1, U2, U4, U5, and U6 snRNPs), protein-only complexes (NineTeen Complex, NTC; NTC-related complexes, NTR), and transiently interacting protein splicing factors (e.g. several DExD/H-box ATPases required for splicing). While spliceosome assembly is

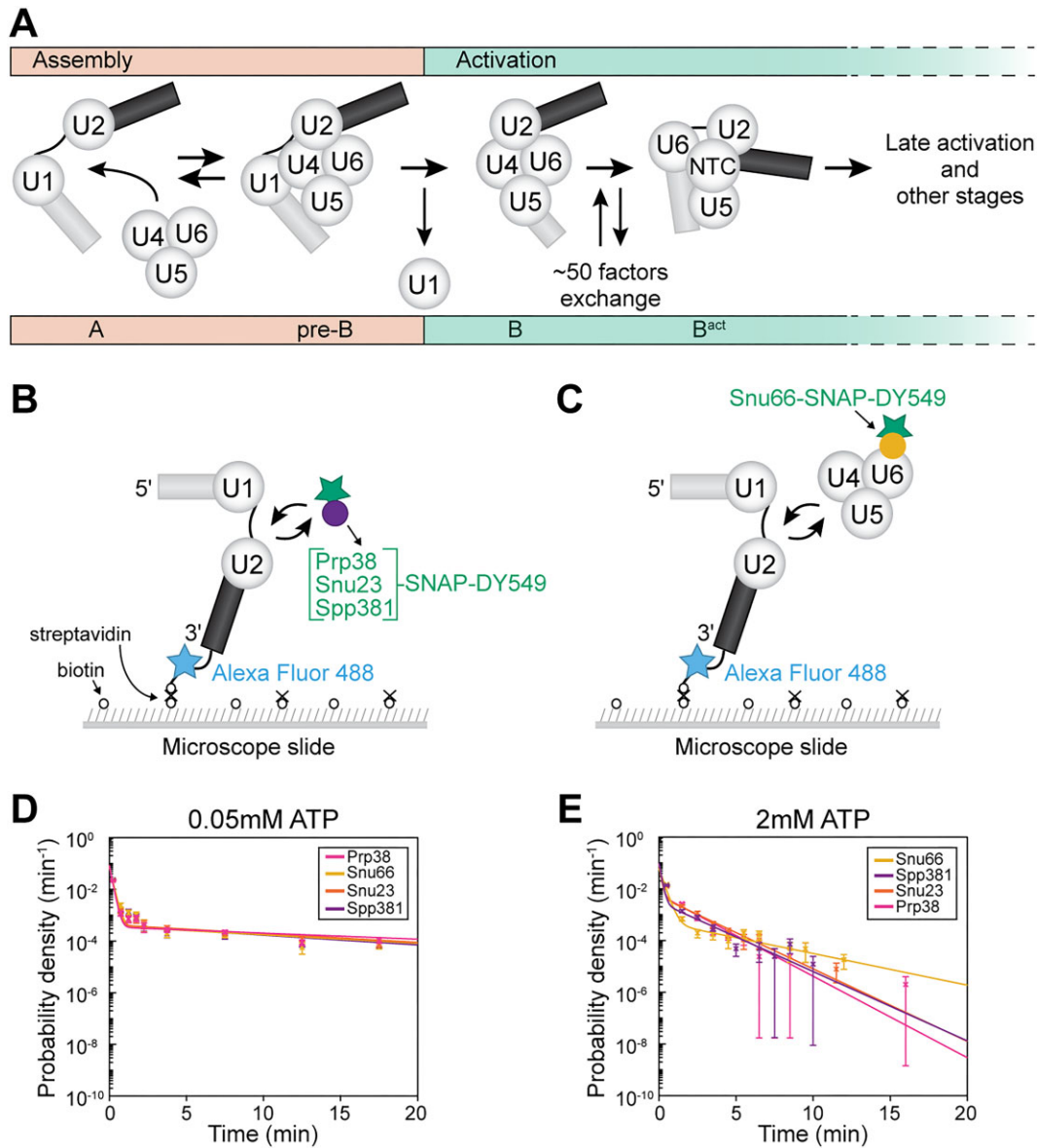
largely ordered, it is also a highly dynamic process (Fig. 1A) [1–3]. In brief, spliceosome assembly begins by recruitment of the U1, U2, and U4/U6.U5 tri-snRNPs to pre-mRNA to form the pre-B complex spliceosome. This is then followed by spliceosome activation during which the nascent active site is formed, the U1 and U4 snRNPs are released, and the NTC and NTR complexes are recruited. These steps involve the sequential formation of the B and B<sup>act</sup> intermediate complexes. Ultimately, large-scale conformational and compositional changes lead to the formation of a U2/U6 di-small nuclear RNA (snRNA) catalytic center associated with the intron

Received: August 8, 2024. Revised: January 31, 2025. Editorial Decision: February 4, 2025. Accepted: February 7, 2025

© The Author(s) 2025. Published by Oxford University Press on behalf of Nucleic Acids Research.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License

(<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact [reprints@oup.com](mailto:reprints@oup.com) for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact [journals.permissions@oup.com](mailto:journals.permissions@oup.com).



**Figure 1.** Overview of yeast spliceosome assembly and activation and CoSMoS assays. **(A)** Spliceosome activation begins after the tri-snRNP joins the spliceosome A complex to form the pre-B complex. Release of the U1 snRNP is then followed by rearrangements that include exchange of ~50 protein and snRNA factors to form the activated spliceosome (B<sup>act</sup>). This panel was adapted from Fu *et al.* [19]. **(B and C)** Schematics of two-color CoSMoS experiments for observing either Prp38, Snu23, or Spp381 binding dynamics panel (B) or Snu66-binding dynamics panel (C). **(D and E)** Probability density histograms for dwell times of Prp38, Snu66, Snu23, or Spp381 under conditions that either block (0.05 mM ATP, panel D) or permit spliceosome activation (2 mM ATP, panel E).

and capable of catalyzing splicing chemistry. Upon completion of exon ligation, the spliceosome disassembles, and its components are recycled for processing of other pre-mRNAs.

During the activation of human spliceosomes, B complex-specific factors (PRPF38, MFAP1, ZMAT2, FBP21, SMU1, RED, NPW38, and NPW38BP) are recruited during the transition from the pre-B to B complex spliceosome and released during the transition from the B to B<sup>act</sup> complex [4–7]. Cryo-EM structures of human spliceosomes have revealed where these factors are bound. ZMAT2 and PRPF38 are positioned to anchor the U6 snRNA/5' splice site (SS) duplex to the N-terminal domain of PRP8 whereas MFAP1 helps to stabilize the interaction between the 5' exon and U5 snRNA and interacts with the PRP8 switch loop [5]. The binding sites for these proteins are mutually exclusive with that of Prp28 (a DEAD-

box ATPase involved in the transfer of the 5' SS from the U1 to the U6 snRNA) indicating that Prp28 must be released before they can bind at that location [8]. Consequently, the presence of these proteins on the human spliceosome signifies a particular state in which activation has begun, the pre-B complex has been converted to the B complex, the U1 snRNP has dissociated, the 5' SS has been identified and transferred to U6, and Prp28 has departed.

Only PRPF38, ZMAT2, and MFAP1 are evolutionarily conserved between human and *Saccharomyces cerevisiae* (hereafter yeast; Prp38, Snu23, and Spp381, respectively) [9]. However, the three yeast proteins can be isolated as part of the yeast U4/U6.U5 tri-snRNP in the absence of splicing [9–13]. As a result, two different pathways for recruitment of B-specific factors have been proposed: ordered recruitment of the hu-

man proteins during spliceosome activation after the association of the tri-snRNP [5] or simultaneous recruitment of the yeast tri-snRNP and B-specific factors [14]. Given the evolutionary conservation of the splicing machinery, the origins and consequences for this variation remain elusive as do the implications for 5' SS identification and Prp28 function [9].

Prp38, Snu23, and Spp381 are all essential for yeast viability [15–17]. Though no crystal structures of a yeast Prp38–Snu23–Spp381 cocomplex have yet been solved, a crystal structure containing a ternary complex formed from fragments of *Chaetomium thermophilum* Prp38, Snu23, and MFAP1 has been obtained [18]. This structure and accompanying biochemical data suggest that they can potentially function as a heterotrimeric subcomplex during splicing and associate with one another in the absence of the spliceosome. Further, the yeast proteins can form an isolable complex by gel filtration experiments [9]. Despite evidence for interactions between the three proteins, it is not known if either the yeast or human B-specific factors are recruited to and release from tri-snRNPs or spliceosomes individually or as a single subunit.

Since the transition from B to B<sup>act</sup> spliceosome involves the exchange of ~50 factors (including the release of B-specific factors and the U4 snRNP and recruitment of the NTC and NTR), it is unsurprising that this transition likely involves the formation of several intermediates [14, 19]. Two recent human pre-B<sup>act</sup> spliceosome structures stalled during this transition using a small molecule splicing inhibitor revealed at least two substeps [5]. In pre-B<sup>act-1</sup> spliceosome (representing the first substep), U4 snRNP has been released, PRPF38, ZMAT2, and MFAP1 are present, and a subset of NTC proteins have associated and can be resolved in the structure. PRPF38, ZMAT2, and MFAP1 are also retained in pre-B<sup>act-2</sup> where they function to impede docking of the U6/5' SS duplex into the nascent active site and Prp8 rearrangement. Together, these data suggest that PRPF38, ZMAT2, and MFAP1 are released after U4 snRNP dissociation and after NTC recruitment by an unknown mechanism. Similar activation intermediates have not yet been structurally characterized for the yeast spliceosome.

The highly compositional and conformational dynamic activation process renders the underlying mechanism intrinsically complicated and difficult to study. Even though cryo-EM structures have provided a few snapshots along the activation path, the number of kinetic steps and characteristic spliceosome complexes involved in each transition are challenging to predict. Previously, our lab has employed Colocalization Single Molecule Spectroscopy (CoSMoS) to study the dynamics of major factors (U4 snRNP proteins, NTC proteins, and Lsm2–8 proteins) involved in the activation pathway [19, 20]. In this work, we analyze the dynamics of yeast Prp38, Snu23, and Spp381 during activation and compare these to other splicing factors to order their arrival to and departure from single spliceosomes. Our results show that a human-like activation pathway is conserved for the yeast splicing machinery. Prp38, Snu23, and Spp381 can bind to and dissociate from spliceosomes likely as a subcomplex, which we refer to as the putative B complex protein (BCP) subcomplex. Under splicing conditions, spliceosomes can assemble by ordered, sequential addition of the tri-snRNP and BCP to pre-mRNAs. Surprisingly, the association of BCP proteins with the tri-snRNP is ATP-dependent. Under low ATP conditions that prevent splicing, BCP proteins and the tri-snRNP instead associate with pre-mRNA simultaneously. This observation likely explains the presence of the yeast BCP in tri-snRNPs and tri-snRNP-

containing complexes purified under nonsplicing conditions. Together, our results provide insight into how tri-snRNP composition may be regulated by ATP and demonstrate evolutionary conservation of a fundamental step in spliceosome activation.

## Materials and methods

### Yeast strains

Yeast strains (Supplementary Table S1) were derived from the protease-deficient strain BJ2168. Splicing factors were C-terminally tagged by integrating fast SNAP or dihydrofolate reductase (DHFR) tags at the appropriate genomic locations by homologous recombination as described [1, 20].

### Yeast growth assay

Selected yeast strains were grown in YPD media to saturation in a 30°C shaking incubator. Concentrated yeast cultures were then diluted to OD<sub>600</sub> = 0.03 with YPD media and placed in a 48-well plate (Corning Costar 48-well Flat Bottom Cell Culture Plate). The plate was covered with a Breathe-Easy plate sealing membrane to limit evaporation during incubation. A Tecan Infinite 200Pro plate reader was then used to monitor yeast growth by measuring the OD<sub>600</sub> for 30 h at 30°C with shaking.

### Preparation of yeast whole cell extract

Yeast whole cell extract (WCE) for splicing was prepared according to published protocols using the ball mill method [21]. SNAP-tagged proteins were fluorophore labeled as previously described [22]. Briefly, SNAP-Surface® 549 dye (S9112S, New England BioLabs; abbreviated elsewhere as SNAP-DY549) in dimethylsulfoxide (DMSO) was added to 1.2 ml of yWCE to a final concentration of 1 μM. The reaction tube was rotated in the dark for 30 min at room temperature. The reaction was then loaded onto a pre-equilibrated G-25 Sephadex column (Kontes Flex Column) in SEC buffer [25 mM HEPES-KOH, pH 7.9, 50 mM KCl, 1 mM dithiothreitol (DTT), and 10% (v/v) glycerol] at 4°C to remove excess dye. A peristaltic pump was used for pre-equilibrating the column and eluting the labeled extract at a flow rate of ~0.25 ml/min with the SEC buffer. Fluorophore labeling of the proteins was confirmed by sodium dodecylsulfate–polyacrylamide gel electrophoresis (SDS–PAGE) and fluorescence using a Typhoon FLA 9000 scanner (Cytiva) at 532nm. Results were analyzed with ImageQuantTL (Cytiva) software.

### In vitro splicing assays

[α-<sup>32</sup>P] uridine 5'-triphosphate (UTP) radiolabeled (PerkinElmer) and m<sup>7</sup>G(5')ppp(5')G capped (New England Biolabs) RP51A pre-mRNA substrates were made by *in vitro* transcription of a linear DNA template with T7 RNA polymerase (Agilent or purified in the laboratory). The DNA template was produced from a polymerase chain reaction (PCR) reaction of pBS117 plasmid [23] using Taq DNA polymerase (M7122, Promega), followed by gel purification of the products with a Wizard SV Gel and PCR Clean-Up System kit (Promega). Transcription products were separated on a 6% denaturing polyacrylamide gel (AccuGel 19:1, National Diagnostics), followed by ethanol precipitation of the extracted RP51A transcripts. Transcripts were resuspended in nuclease-free water (Ambion, Fischer Scientific) and quanti-



fied with a liquid scintillation counter (LSC, tri-carb 2900TR, Packard).

A typical *in vitro* splicing reaction included 40% (v/v) WCE and 0.2 nM RP51A substrate in a splicing buffer [final concentrations: 100 mM KPi, pH 7.3, 3% (w/v) PEG-8000, 1 mM DTT, 2.5 mM MgCl<sub>2</sub>, and 0.2 U/μl RNasin Plus (Promega)]. The reaction was incubated at room temperature for 45 min. The reaction was quenched in a splicing dilution buffer [100 mM Tris base, pH 7.5, 10 mM EDTA, pH 8.0, 1% (w/v) SDS, 150 mM NaCl, and 300 mM NaOAc, pH 5.3]. RNAs from the reaction were extracted using phenol–chloroform, ethanol precipitated, resuspended in deionized formamide, and separated on a 12% denaturing polyacrylamide gel. Gels were then dried and exposed to a phosphor screen overnight. The screen was imaged with a Typhoon FLA 9000 scanner (Cytiva), and results were analyzed with ImageQuantTL (Cytiva) software. The intensities of RP51A pre-mRNAs and splicing products were determined by integrating the signals within same-sized rectangles around the band. The background corrected intensities for bands were then used for calculating splicing efficiencies. The first step efficiency was calculated as the ratio of the sum of the intensities of bands corresponding to products having completed 5' SS cleavage (the mRNA and intron lariat-3' exon) over the sum of those intensities plus that for the pre-mRNA band. The second step efficiency was calculated as the ratio of the intensity of the mRNA band to the sum of the intensities of the pre-mRNA, intron lariat-3' exon, and mRNA bands.

### Preparation of fluorescently labeled RP51A pre-mRNAs

Fluorescent, biotinylated RP51A pre-mRNAs were prepared as described by splinted ligation of a fluorescent biotin handle oligonucleotide to a capped, RP51A transcript prepared by *in vitro* transcription using T4 RNA Ligase 2 [19, 24].

### CoSMoS assays

Microscope slides (100490-396, VWR) and cover glasses (12-553-455, Fischer) were cleaned and passivated using a mixture of mPEG-SVA (MPEG-SVA-5K, Laysan Bio) and mPEG-biotin-SVA (BIO-PEG-SVA-5K, Laysan Bio) at a ratio of 1:100 (w/w) in 100 mM NaHCO<sub>3</sub> (pH 8) buffer as previously described [19].

For each CoSMoS assay, the mPEG mixture was washed off the slide with 1× phosphate-buffered saline (PBS), and then streptavidin (0.2 mg/ml, SA10, Kelowna International Scientific) was added and allowed to bind the surface. The excess streptavidin was washed off the surface using PBS and the biotinylated, fluorescent RNAs were then bound. The slide was washed a final time with PBS before the addition of splicing assay buffer containing 40% (v/v) WCE, 20 nM Cy5-TMP, oxygen scavenger [5 mM protocatechuic acid (PCA) and 0.96 U/ml protocatechuic dioxygenase (PCD)], 1 mM Trolox, and triplet quenchers (0.5 mM propyl gallate, 1 mM cyclooctatetraene, and 1 mM 4-nitrobenzyl alcohol) [1]. The triplet quenchers were added as a mixture from a 100× stock in DMSO. Additional details of these procedures have been recently described elsewhere [19].

### CoSMoS data acquisition

Single molecule data were collected on a custom-built objective-type TIRF microscope using GLIMPSE software (<https://github.com/gelles-brandeis/Glimpse>) as previously

described [19]. For the laser illumination, powers were set to the ranges of 1.2–1.5 mW, 500–600 μW, 400–440 μW, and 2.5 mW for the 488 nm (blue), 532 nm (green), 633 nm (red), and 785 nm (infrared) lasers, respectively. The 488, 532, and 633 nm lasers were used to image fluorophores while the 785 nm laser was used for focusing.

Cycles of time-lapse imaging were used according to the following excitation scheme with each frame lasting 1 s: the sample was illuminated with the 785 nm for focusing, the 488 nm blue laser was then turned on to collect two consecutive frames to image the immobilized RNAs, and then the 532 and 633 nm lasers were turned on to simultaneously collect 14 frames with a 3 s delay between adjacent frames. The total cycle time was ~1 min, and this cycle was repeated 50× to collect videos lasting for ~50 min. To avoid photo bleaching of DY-549 and Cy5 fluorophores by the 488 nm laser, the path of the laser was physically blocked after the collection of 10 frames of blue laser-illuminated images. Data were collected following a typical dead time of ~5 min during which the samples were loaded and the images were aligned and focused.

### Photobleaching control

To estimate the influence of photobleaching of the SNAP tag fluorophore (DY549) on the experiments, we utilized a purified, biotinylated, and fluorescently labeled SNAP protein. The protein was imaged under the same conditions used to follow splicing reactions, and its average fluorophore lifetime ( $\tau = 995$  s) was determined by measuring the loss in fluorescence signal overtime as previously described [19].

### CoSMoS data analysis

Single molecule data were analyzed as described previously [1, 19, 25] using a custom program imscroll ([https://github.com/gelles-brandeis/CoSMoS\\_Analysis](https://github.com/gelles-brandeis/CoSMoS_Analysis)) written in MATLAB (The Mathworks). Binding events were identified as signals centered within the area-of-interest (AOI) that appeared with intensities  $>3.6\times$  the standard deviation above the mean of the background noise. Loss of signals were identified as points in time at which the signal fell below  $1\times$  standard deviation above the mean background.

Analysis of the measured dwell times and fits to kinetic equations were carried out using MATLAB and AGATHA software (<https://github.com/hoskinslab/AGATHA>) using maximum likelihood methods and fitting to equations containing one (Eq. 1), two (Eq. 2), or three (Eq. 3) exponential terms as described [25].

$$\frac{1}{\left(e^{-tm_{\tau}} - e^{-tx_{\tau}}\right)} * \left(\frac{1}{\tau} e^{-\frac{intervals}{\tau}}\right) \quad (1)$$

$$\frac{1}{a\left(e^{-tm_{\tau}1} - e^{-tx_{\tau}1}\right) + (1-a)\left(e^{-tm_{\tau}2} - e^{-tx_{\tau}2}\right)} * \left(\frac{a}{\tau_1} e^{-\frac{intervals}{\tau_1}} + \frac{1-a}{\tau_2} e^{-\frac{intervals}{\tau_2}}\right) \quad (2)$$

$$\frac{1}{a_1\left(e^{-tm_{\tau}1} - e^{-tx_{\tau}1}\right) + a_2\left(e^{-tm_{\tau}2} - e^{-tx_{\tau}2}\right) + (1-a_1-a_2)\left(e^{-tm_{\tau}3} - e^{-tx_{\tau}3}\right)} * \left(\frac{a_1}{\tau_1} e^{-\frac{intervals}{\tau_1}} + \frac{a_2}{\tau_2} e^{-\frac{intervals}{\tau_2}} + \frac{1-a_1-a_2}{\tau_3} e^{-\frac{intervals}{\tau_3}}\right) \quad (3)$$

Bootstrapping was used to calculate standard errors for all fitted parameters. Fit parameters are included in Supplementary Tables S2 and S3. Histograms for the distribution of events were generated in MATLAB with empty bins

removed. The error (standard deviation) for each bin was calculated using Eq. (4), with the assumption that the number of events within a bin follows a binomial distribution.

$$\sigma = \sqrt{NP(1-P)} \quad (4)$$

## Results

### Prp38, Snu23, and Spp38 have similar pre-mRNA binding kinetics

To visualize splicing factor dynamics directly, we genetically encoded C-terminal fast SNAP tags on Prp38, Snu23, and Spp381 in three separate yeast strains (Supplementary Table S1, strains #5–7). As a control, we also labeled the tri-snRNP protein Snu66 in a separate strain (Supplementary Table S1, strain #4). Snu66 is released during the B to B<sup>act</sup> transition but occupies a different binding site on the spliceosome than the other three factors [12]. We confirmed that these SNAP-tagged proteins can be readily labeled with benzylguanine-fluorophores (SNAP-DY549) in yeast WCEs and that the labeled extracts efficiently spliced RP51A substrate pre-mRNAs (Supplementary Figs S1 and S2). We do note, however, that strains containing tagged Prp38 grew more slowly and produced WCE with lower splicing activity when Snu23 and Spp381 or NTC components were also DHFR tagged (Supplementary Fig. S2) relative to others. It is possible that tagging of Prp38 disrupts expression of its neighboring gene, SMD1 (a component of the U snRNP Sm ring), and that this contributes to the observed differences in growth and splicing activity [26].

In order to observe splicing factor dynamics, we flowed a pre-mixed solution consisting of a fluorescently labeled WCE and splicing buffer into a reaction chamber containing surface-immobilized, Alexa Fluor 488-labeled RP51A pre-mRNAs (Fig. 1B and C). We carried out the experiments at two ATP concentrations that either permit spliceosome activation and splicing (2 mM ATP) or inhibit activation and splicing but permit assembly (0.05 mM ATP) [27]. Using a custom-built CoSMoS microscope [28], we could observe binding and unbinding dynamics of fluorescently labeled proteins on the fixed fluorescent pre-mRNA molecules in real time.

All of the proteins showed ATP-dependent dynamics. At 0.05 mM ATP, the dwell times of each protein bound to single pre-mRNA molecules often lasted for tens of minutes (Supplementary Fig. S3). In comparison, the binding events at 2 mM ATP were much shorter and lasted from a few seconds to a few minutes. In both cases, we observed some binding events at the start of data collection. These were likely due to some amount of spliceosome assembly occurring during the dead time of the experiment (~5 min) during sample loading. In addition, at least some of these proteins (Snu23 and Prp38) directly interact with RNA in the spliceosome [14] and may have some intrinsic affinity for nucleic acid. These results resemble single molecule data obtained from other proteins involved in spliceosome activation, such as Prp3 and Prp4 from U4 snRNP complexes [20] and Lsm8 from the Lsm2–8 ring complex [19]. It is likely that the long dwell times observed at 0.05 mM ATP originated from stalled spliceosome complexes while the shorter times seen at 2 mM ATP originated from transient binding due to release of these factors during spliceosome activation.

We further analyzed the binding dynamics by fitting the collected, unbinned dwell times to exponential-based func-

tions (Supplementary Table S2) and generating probability density histograms. At low ATP, histograms and fits for Prp38, Spp381, Snu23, and Snu66 are very similar to one another (Fig. 1D and Supplementary Table S2). In contrast, under conditions that permit activation (2 mM ATP), the distributions and fits for Prp38, Spp381, and Snu23 remained similar to one another but were distinct from those of Snu66 (Fig. 1E and Supplementary Table S2). Together, these results indicate that Prp38, Spp381, Snu23 interact with the splicing machinery with similar kinetic features. These features are shared with Snu66 only under low ATP conditions that inhibit spliceosome activation and splicing.

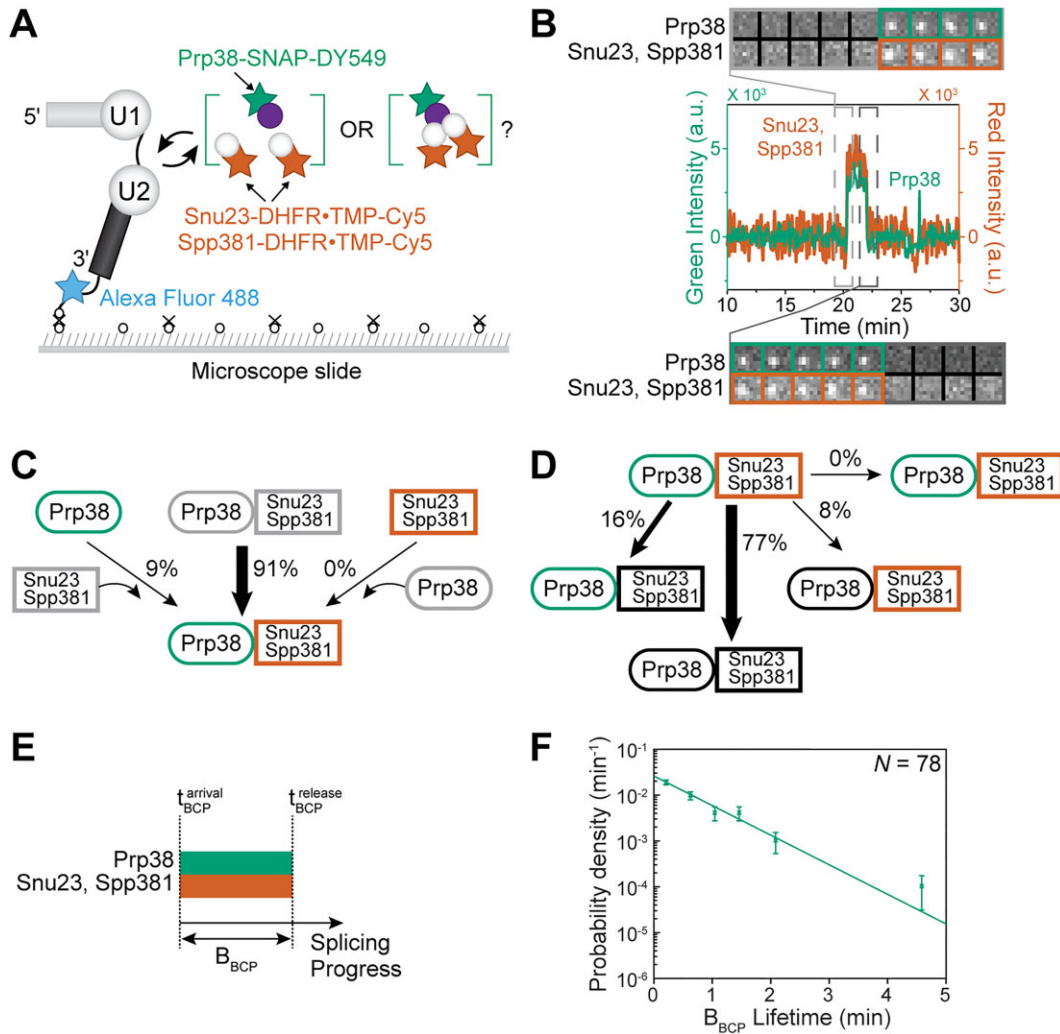
### Prp38, Spp381, and Snu23 can be recruited to and released from spliceosomes simultaneously

To see if the similarity in Prp38, Spp381, and Snu23 kinetics is coincidental or results from their presence together in a single spliceosomal subcomplex, we carried out experiments to visualize the recruitment and release of the proteins relative to one another during splicing. We carried out three-color CoSMoS assays using WCE containing Prp38–SNAP labeled with SNAP–DY549 fluorophores and DHFR tags on Spp381 and Snu23 labeled with Cy5–TMP fluorophores (Fig. 2A and Supplementary Table S3).

In these experiments, we noted that not all events due to Prp38 binding also contained an overlapped binding event from Snu23 and Spp381 and vice versa. Approximately 24% (112/460) of Prp38-binding events colocalized with Snu23/Spp381 binding events, and ~44% (112/256) of Snu23/Spp381 binding events colocalized with Prp38. While we cannot exclude the presence of partially labeled complexes, the data suggest the presence of ternary complexes of Prp38/Snu23/Spp381 alongside single proteins and partial complexes.

We then identified the subset events in which DY549 signals from Prp38 and Cy5 signals from Spp381 and Snu23 were colocalized on the immobilized RNAs since during spliceosome activation Prp38, Snu23, and Spp381 should all be simultaneously present (Fig. 2B and Supplementary Fig. S4). For these events, the binding patterns were then classified based on the relative recruitment and release orders for Prp38 and Snu23/Spp381 (Supplementary Fig. S5). For example, we determined if the DY549 signal appeared or disappeared before, after, or simultaneously with the Cy5 signal for each colocalized pair. We found that most of the event pairs (91%; 102/112) showed simultaneous arrival of Prp38–SNAP and Snu23/Spp381–DHFR signals (Fig. 2C). Among this subset, 77% showed simultaneous release of all three proteins (Fig. 2D). While it is possible that the proteins individually bind and dissociate too quickly one after the other for our measurements to detect, our data are most consistent with structural and gel filtration data [4, 5, 14, 18] supporting formation of a Prp38/Snu23/Spp381 ternary complex in the absence of a spliceosome.

We additionally used the measured dwell times for the colocalized event pairs to determine the lifetime of the BCP-containing spliceosome (the B<sub>BCP</sub> complex, Fig. 2E). The distribution of lifetime measurements could be described by a function containing a single exponential term, resulting in a characteristic lifetime of ~41 s for B<sub>BCP</sub> (Fig. 2F and Supplementary Table S4). Our determined lifetime of the B<sub>BCP</sub> species is unlikely to be limited by photobleaching since simul-



**Figure 2.** Prp38, Snu23, and Spp381 can simultaneously bind to and release from spliceosomes. **(A)** Schematic of a three-color CoSMoS assay in which Spp381/Snu23 are labeled with Cy5-TMP, Prp38 is labeled with a DY549, and the surface-tethered pre-mRNA is labeled with Alexa Fluor 488. **(B)** Representative time record segment showing peaks in fluorescence intensity corresponding to colocalization of Snu23 and Spp381 (red) with Prp38 (green) on the same pre-mRNA molecule. The dashed rectangles mark examples of the simultaneous appearance and disappearance of Snu23, Spp381, and Prp38 spots; galleries show consecutive images taken from the indicated part of the recording showing that both spot appearance and disappearance are simultaneous. **(C)** Routes for recruitment of either the Prp38 or Spp381/Snu23 fluorescent spots at 2 mM ATP for  $N = 112$  pairs of overlapping events. Green and red shapes represent observation of fluorescence from the corresponding DY549 or Cy5 fluorophores on Prp38 or Snu23, Spp381, respectively. Gray colors represent the absence of fluorescence. Percentages represent the fraction of complexes in which fluorescence appeared by the indicated pathway; more prevalent pathways are emphasized with thicker arrows. **(D)** Routes for loss of either the Prp38 or Snu23, Spp381 fluorescent spots at 2 mM ATP for  $N = 102$  pairs of overlapping events that showed simultaneous recruitment. Black shapes represent the disappearance of fluorescence. **(E)** Schematic depicting how the lifetime of the  $B_{BCP}$  complex is defined relative to the recruitment and release times of BCP-constituent proteins. **(F)** Probability density histograms of  $B_{BCP}$  complex lifetimes obtained from the subset of events ( $N = 78$ ) showing both simultaneous arrival and departure of Prp38, Snu23, and Spp381. The green line represents a fit of the distribution to an equation containing a single exponential term that yielded the parameters reported in [Supplementary Table S4](#) ( $\tau_1 \sim 41$  s). Error bars ( $\pm$ SD) were calculated for each point as described in the "Materials and methods" section.

taneous bleaching of the DY549 and Cy5 signals should be rare and we were readily able to measure lifetimes  $> 10$ -fold larger for the SNAP-labeled proteins when ATP was limiting (Fig. 1D and [Supplementary Table S2](#)). Since splicing *in vitro* occurs over tens of minutes, dissolution of the  $B_{BCP}$  complex occurs rapidly enough to support splicing and likely involves just a single, rate-limiting kinetic step. Notably, the distribution of the subset of BCP binding events was nonidentical to the distributions of all the measured Prp38 or Spp381/Snu23 events ([Supplementary Tables S3 and S4](#)). The former was single exponential with a characteristic lifetime of  $\sim 41$  s, while the latter were best fit using three exponential terms with characteristic lifetimes of  $\sim 8$ ,  $\sim 40$ , and  $> 300$  s and amplitudes of

$\sim 0.4$ ,  $0.5$ , and  $\sim 0.05$ , respectively. This suggests that a ternary complex of Prp38, Snu23, and Spp381 exhibits distinct binding kinetics in comparison to the individual proteins or smaller complexes. Incomplete complexes may interact more briefly with the RNA/spliceosome (dwell times of  $\sim 8$  s) or rarely associate very tightly (dwell times  $> 300$  s).

#### BCP proteins join the spliceosome after tri-snRNP recruitment under splicing conditions

We then studied when the BCP subcomplex is recruited to the spliceosome and when it is released during activation relative to the U4 snRNP. The U4 snRNP is a core component



of the tri-snRNP and thus reports on tri-snRNP binding to the pre-mRNA. Release of U4 snRNP is a key event during the B-to-B<sup>act</sup> complex transition since it permits formation of the U2/U6 di-snRNA catalytic site. We carried out three-color CoSMoS assays using extracts containing SNAP-tagged Snu23 labeled with SNAP-DY549 fluorophore and DHFR-tagged U4 snRNP proteins (Prp3 and Prp4) labeled with Cy5-TMP fluorophores along with surface-immobilized Alexa Fluor 488-labeled pre-mRNAs (Fig. 3A and [Supplementary Fig. S6A](#)). We first identified all paired binding events in which fluorescent signals from the U4 snRNP proteins were simultaneously present on the pre-mRNA with signals from Snu23. We found that most of the event pairs (89%; 31/35 events) showed sequential arrival. The U4 snRNP protein signals appeared first followed by appearance of Snu23, suggesting that BCP proteins are not recruited as part of the tri-snRNP (Fig. 3B and C, and [Supplementary Fig. S7A](#)). Consistently, the same recruitment pattern was also found when extracts containing SNAP-tagged Spp381 or Prp38 were used along with DHFR-tagged U4 snRNP proteins ([Supplementary Figs S6B and C, and S7B–E](#)). Together these results indicate that the BCP is recruited to the spliceosome after the tri-snRNP associates.

We further analyzed the order of BCP and U4 snRNP release for pairs of colocalized events. In the case of the BCP protein Snu23, the majority of event pairs (86%) showed loss of signals from the U4 snRNP proteins occurring before loss of signals from Snu23 (Fig. 3D and [Supplementary Fig. S7A](#)). This was also the predominant pattern for signal loss from Spp381 and Prp38 when WCEs containing those labeled proteins were used ([Supplementary Fig. S7B, C, F, and G](#)). This result is consistent with ordered release of the U4 snRNP and then the BCP from the spliceosome as well as a previously determined structure of the human pre-B<sup>act-1</sup> activation intermediate that contains the human BCP but not the U4 snRNP [5]. Thus, the yeast spliceosome appears to share with humans conserved pathways for both recruitment and release of BCP proteins during activation.

The single molecule data also indicate the existence of at least three types of spliceosome B complexes: B<sub>U4</sub> in which the tri-snRNP has associated with the pre-mRNA but the BCP has not yet joined, B<sub>U4+BCP</sub> in which the BCP has bound but U4 has not been released, and B<sub>BCP</sub><sup>ΔU4</sup> in which the U4 snRNP has been released and the BCP has not yet dissociated (Fig 3E). In this nomenclature, we use the subscript to identify components of the B complex spliceosome and the superscript to signify loss of a particular subunit.

We were able to determine the characteristic lifetimes of each complex to be ~57, ~13, and ~101 s for the B<sub>U4</sub>, B<sub>U4+BCP</sub>, and B<sub>BCP</sub><sup>ΔU4</sup> complexes, respectively ([Supplementary Fig. S8 and Supplementary Table S4](#)). These data indicate that under splicing conditions, the BCP and U4 snRNP only transiently coexist in the same spliceosome complex before U4 snRNP release is completed.

### Snu66 is released with the U4 snRNP and before BCP proteins

We next tested if another protein released during the B-to-B<sup>act</sup> transition but not part of the BCP also dissociates subsequent to U4 snRNP release. We carried out three-color CoSMoS assays using WCEs containing Snu66–SNAP labeled with SNAP-DY549 fluorophore and doubly DHFR-tagged U4

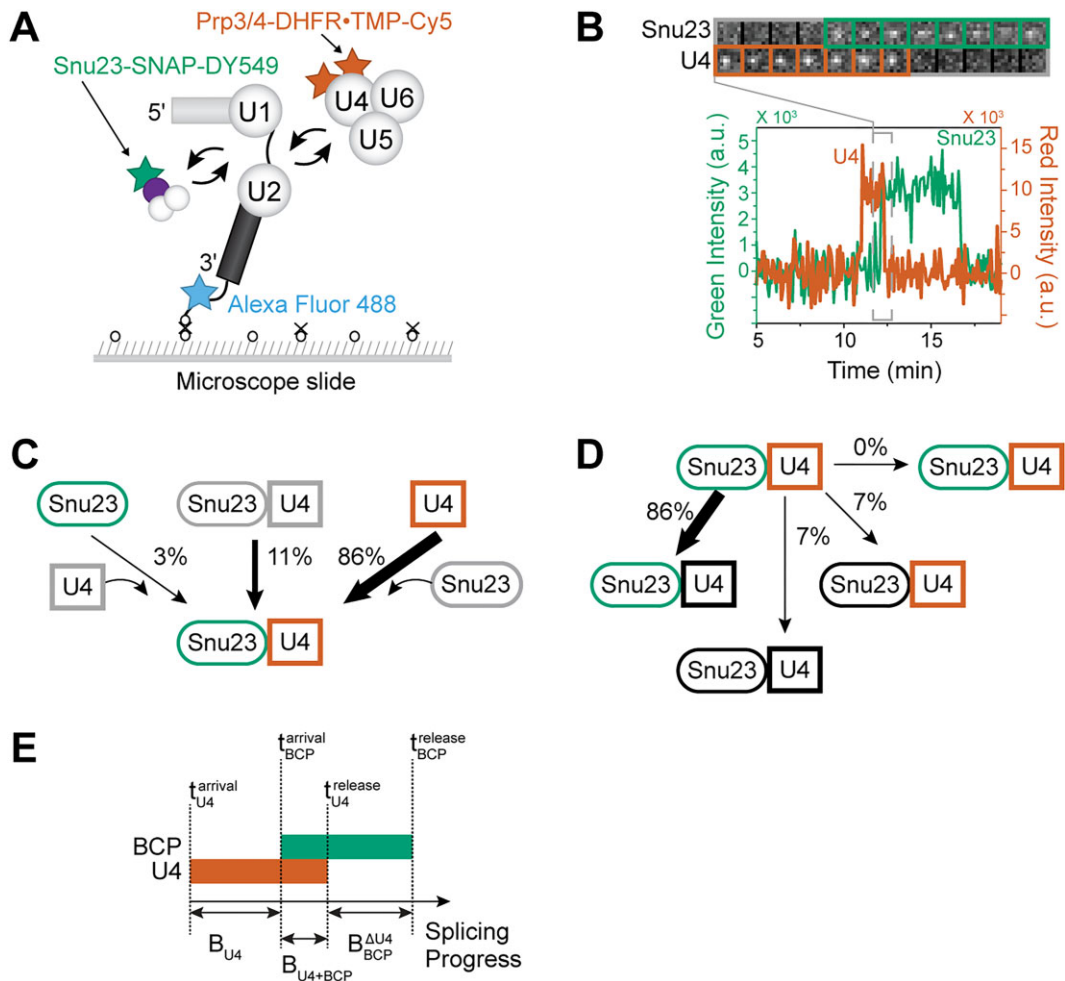
snRNP proteins (Prp3 and Prp4) labeled with Cy5-TMP fluorophores (Fig. 4A). As before, we identified colocalized pairs of events and then studied the orders in which the signals appeared or were lost. At 2 mM ATP, the predominant pathway for Snu66 association with the pre-mRNA is coincident with U4 snRNP proteins (80%; 77/96 event pairs; Fig. 4B and C, and [Supplementary Fig. S9](#)). Among the subset in which Snu66 and U4 snRNP protein signals appeared simultaneously, the majority also showed simultaneous loss of the signals (72%, 55/77 event pairs; Fig. 4B and D). While we cannot exclude the possibility that Snu66 binding or release occurs very swiftly (<4 s) relative to tri-snRNP binding and U4 dissociation, our data are most consistent with Snu66 being recruited to spliceosomes as part of the tri-snRNP and being released along with the U4 snRNP. The release of Snu66 along with U4 snRNP proteins could be due to direct interaction between Snu66 and Prp3, as has been observed in human tri-snRNPs [29].

The above data show that not all proteins released during the B-to-B<sup>act</sup> transition dissociate simultaneously. When combined with the data from Fig. 3, we were able to predict that Snu66 and the BCP should be released sequentially. Specifically, since BCP proteins are released after the U4 snRNP dissociates and Snu66 dissociates along with the U4 snRNP, then BCP proteins should also be released after Snu66. To verify this prediction, we carried out a three-color CoSMoS assay using extracts containing Snu66–SNAP labeled with SNAP-DY549 fluorophore and doubly DHFR-tagged B complex proteins (Spp381 and Snu23) labeled with Cy5-TMP fluorophores ([Supplementary Fig. S10A](#)). As expected, 84% (66/79) of event pairs showed that Snu66 associates with the pre-mRNA prior to BCP protein binding ([Supplementary Fig. S10B and C](#)). Among this subset of event pairs, 67% (44/66) showed that BCP proteins are released after Snu66 ([Supplementary Fig. S10D](#)), consistent with our prediction.

### BCP proteins are tri-snRNP components at low ATP

The observation that BCP proteins are associating with the pre-mRNA after tri-snRNP binding was unexpected since they have previously been identified as tri-snRNP components by mass spectrometry [10, 11, 15, 30] and were present in samples used to determine the cryo-EM structure of the tri-snRNP, although not modeled [31]. One difference between those experiments and ours is that we carried out our single-molecule assays under conditions that permit spliceosome activation and splicing (2 mM ATP). Under these conditions, the yeast tri-snRNP is unstable. Consequently, structural and analytical experiments were instead carried out in the absence of added ATP [12, 31]. We wondered if the presence or absence of ATP changes BCP association with the tri-snRNP. To test this, we repeated our three-color CoSMoS assays using labeled BCP and U4 snRNP but under conditions that permit spliceosome assembly but inhibit activation (0.05 mM ATP) (Fig. 5A).

At 0.05 mM ATP, Snu23 no longer predominantly associated with the pre-mRNA separately from the tri-snRNP. Instead, the most common pathway was simultaneous binding of Snu23 and U4 snRNP proteins (73%; 68/93 event pairs, Fig. 5B and C, and [Supplementary Fig. S11A](#)). Simultaneous binding with U4 snRNP proteins was also observed for Spp381, Prp38, and Snu66 when these proteins were labeled instead of Snu23 ([Supplementary Fig. S11B–G](#)). This obser-



**Figure 3.** BCP association occurs after tri-snRNP binding and persists after U4 release. **(A)** Schematic of a three-color CoSMoS assay in which U4 snRNP proteins (Prp3 and Prp4) are labeled with Cy5-TMP, the BCP is labeled with a DY549 (here, Snu23), and the surface-tethered pre-mRNA is labeled with Alexa Fluor 488. **(B)** Representative time record segment showing peaks in fluorescence intensity corresponding to colocalization of the U4 snRNP (red) with Snu23 (green) on the same pre-mRNA molecule. The dashed rectangle marks an example of ordered recruitment of the tri-snRNP and BCP, followed by U4 release. The image galleries show consecutive images taken from the indicated part of the recording showing that both spot appearances and disappearances are not simultaneous. **(C)** Routes for the appearance of Snu23 and U4 fluorescent spots at 2 mM ATP for  $N = 35$  pairs of overlapping events. Red and green shapes represent observation of fluorescence from the corresponding DY549 or Cy5 fluorophores on Snu23 or U4, respectively. Percentages represent the fraction of complexes in which fluorescence appeared by the indicated pathway; more prevalent pathways are emphasized with thicker arrows. **(D)** Routes for loss of either Snu23 or U4 fluorescent spots at 2 mM ATP for  $N = 30$  pairs of overlapping events in which the U4 spot appearance preceded arrival of Snu23. **(E)** Schematic depicting how the lifetimes of the  $B_{U4}$ ,  $B_{U4+BCP}$ , and  $B_{BCP}^{\Delta U4}$  complexes are defined relative to the recruitment and release times of constituent factors. Corresponding data for Prp38 and Spp381 are included in [Supplementary Figs S6–S10](#).

vation likely explains why BCP proteins and Snu66 showed similar binding kinetics in two-color CoSMoS assays at 0.05 mM but not 2 mM ATP (Fig. 1D and E). These results indicate that under conditions that stall spliceosome activation, BCP proteins are recruited to the pre-mRNA as stable components of the tri-snRNP and remain stably bound to these spliceosome complexes.

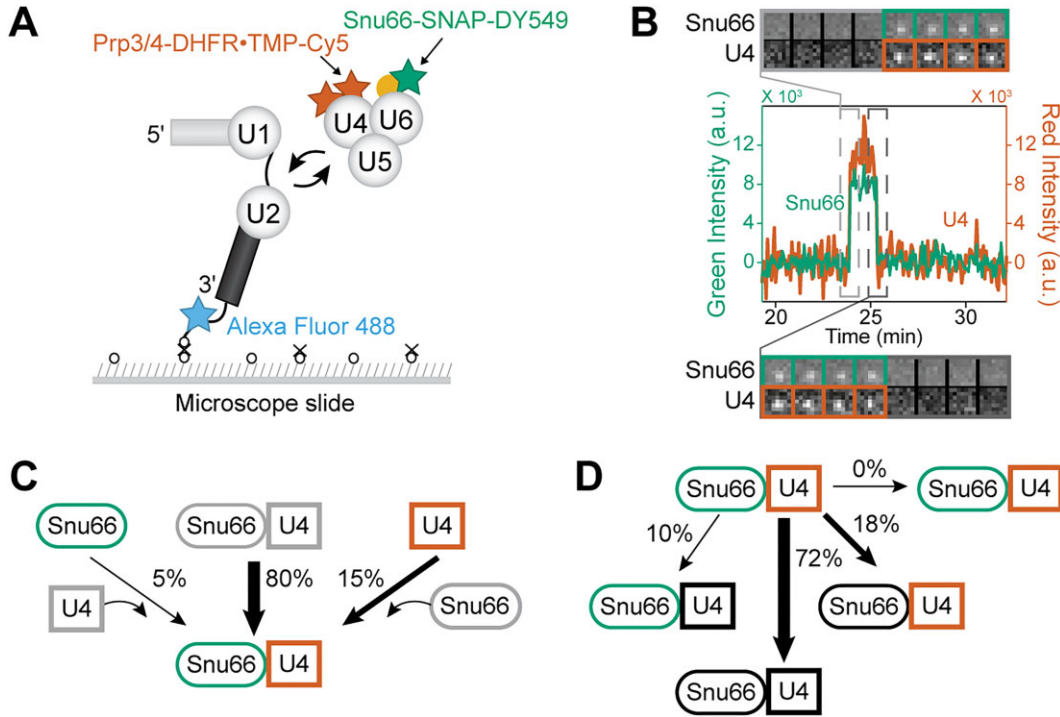
### BCP proteins are released after NTC recruitment

Finally, we wished to order release of the BCP relative to recruitment of the NTC. We previously determined that the NTC is predominantly recruited after U4 snRNP release [20]. Since BCP proteins are also released after U4 departure, we wondered if this occurs coincident with, before, or after NTC recruitment. To answer this question, we carried out three-color CoSMoS assays using WCE containing SNAP-tagged BCP (via Snu23 labeling with SNAP-DY549 fluorophore) and

two DHFR tags on NTC components (Syf1 and Cef1) labeled with Cy5-TMP fluorophores (Fig. 6A).

We identified pairs of BCP (Snu23) and NTC binding events closest in time to one another with a requirement that binding events for both components should be at least two frames in duration ( $\geq 8$  s) to eliminate sampling interactions of the NTC [19]. For each pair, we then subtracted the time of NTC binding ( $t_{\text{NTC}}^{\text{recruitment}}$ ) from the time of BCP signal loss ( $t_{\text{BCP}}^{\text{release}}$ ) to yield a distribution of  $t_{\text{BCP}}^{\text{release}} - t_{\text{NTC}}^{\text{recruitment}}$  values. In this distribution, positive values would indicate that BCP is released after NTC recruitment while negative values would indicate that BCP release occurs before NTC recruitment. The distribution of  $t_{\text{BCP}}^{\text{release}} - t_{\text{NTC}}^{\text{recruitment}}$  values for the selected pairs of events showed that 85% had positive values (97/114 events), suggesting that the NTC is recruited while BCP proteins are still present and that the BCP is released after the NTC binds (Fig. 6B–E and [Supplementary Fig. S12](#)). In comparison, the same analysis carried out for extracts containing SNAP-tagged





**Figure 4.** Snu66 is released along with the U4 snRNP during activation. **(A)** Schematic of a three-color CoSMoS assay in which U4 snRNP proteins (Prp3 and Prp4) are labeled with Cy5-TMP, Snu66 is labeled with a DY549, and the surface-tethered pre-mRNA is labeled with Alexa Fluor 488. **(B)** Representative time record segment showing peaks in fluorescence intensity corresponding to colocalization of the U4 snRNP (red) with Snu66 (green) on the same pre-mRNA molecule. The dashed rectangles mark examples of the simultaneous appearance and disappearance of U4 and Snu66 spots; galleries show consecutive images taken from the indicated part of the recording. **(C)** Routes for the appearance of Snu66 and U4 fluorescent spots at 2 mM ATP for  $N = 96$  pairs of overlapping events. Red and green shapes represent observation of fluorescence from the corresponding DY549 or Cy5 fluorophores on Snu66 or U4, respectively. Percentages represent the fraction of complexes in which fluorescence appeared by the indicated pathway; more prevalent pathways are emphasized with thicker arrows. **(D)** Routes for loss of either Snu66 or U4 fluorescent spots at 2 mM ATP for  $N = 77$  pairs of overlapping events in which Snu66 and U4 bound the pre-mRNA simultaneously.

Snu66 and doubly DHFR-tagged NTC proteins showed predominantly negative values, consistent with loss of Snu66 preceding NTC recruitment (Supplementary Fig. S13).

The single molecule data also indicate the existence of several intermediate complexes:  $B_{BCP}$  in which BCP proteins (Snu23) have associated with the pre-mRNA but the NTC has not yet joined,  $B_{BCP+NTC}$  in which the NTC and BCP are both simultaneously present, and  $B_{NTC}^{\Delta BCP}$  and later complexes in which the BCP has released but the NTC remains (Fig. 6F). The characteristic lifetimes of the  $B_{BCP}$  and  $B_{BCP+NTC}$  complexes are relatively short ( $\sim 1$  min; Fig. 6G and H, and Supplementary Table S4). The total lifetime of the Snu23-SNAP-labeled BCP in these assays was 9-fold shorter than that for the fluorophore measured under the same conditions but attached to an immobilized SNAP protein ( $\sim 110$  s versus  $\sim 995$  s) [19]. This supports the conclusion that the measured lifetimes are indicative of binding and release of the BCP from splicing complexes rather than being primarily due to photobleaching. The  $B_{BCP+NTC}$  intermediate has not yet been described for yeast spliceosomes and likely possesses a similar composition and structure to the human pre- $B^{act-1}$  complex [5].

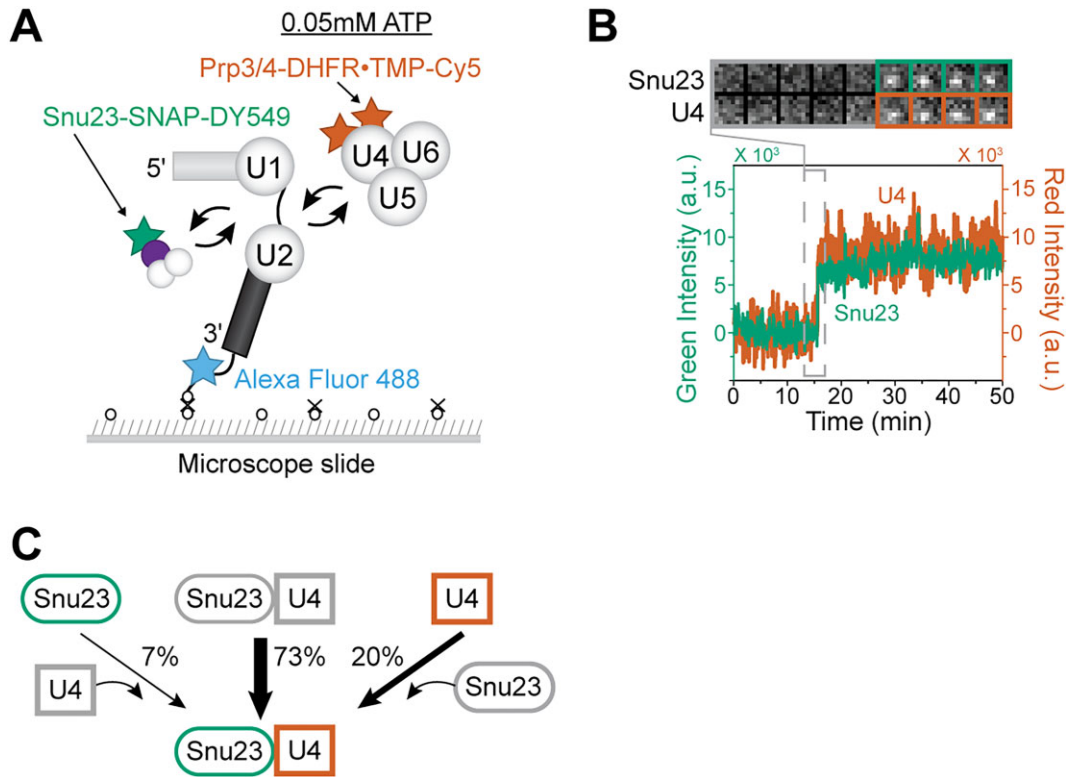
## Discussion

Integrating previous work from our group [19, 20] with single molecule data described here, we can derive a kinetic scheme for spliceosome activation *in vitro* (Fig. 7A). In our scheme,

the tri-snRNP, lacking BCP proteins, joins the spliceosome A complex containing U1 and U2 to form the pre-B complex. After  $\sim 57$  s, a putative BCP complex containing Prp38, Snu23, and Spp381 can join in a single step to form the  $B_{U4+Lsm+BCP}$  spliceosome. We note that a splicing-competent BCP complex has not yet been biochemically purified from yeast and characterized by mass spectrometry. Assemblies of Prp38, Snu23, and Spp381 could contain additional factors, and several proteins have reported physical interactions with all of three components (see Supplementary Table S5).

Based on cryo-EM structures of yeast and human spliceosomes, the BCP associates near the U6 snRNA/5' SS duplex at the same site vacated by Prp28 [5, 32, 33]. Thus, the  $\sim 57$ s time interval reports on both BCP association and transfer of the 5' SS from the U1 snRNA to the U6 snRNA by Prp28. We have not yet analyzed U1 snRNP dynamics during activation, but it is likely that it also dissociates during this time.

Importantly, we cannot exclude the presence of alternate pathways due to the presence of partial BCP complexes and due to the time resolution of our experiments. It is possible that spliceosomes could also be assembled by sequential binding of various assemblies of Prp38, Snu23, and Spp381. A limitation of our method is that identifying such sequential pathways may require four or more color CoSMoS experiments and/or more targeted analysis of factor dynamics (such as on-rates). This could be the subject of future investigations. Our kinetic scheme also assumes that when we measured lifetimes of Prp38, Snu23, and Spp381 relative to other factors (e.g. U4



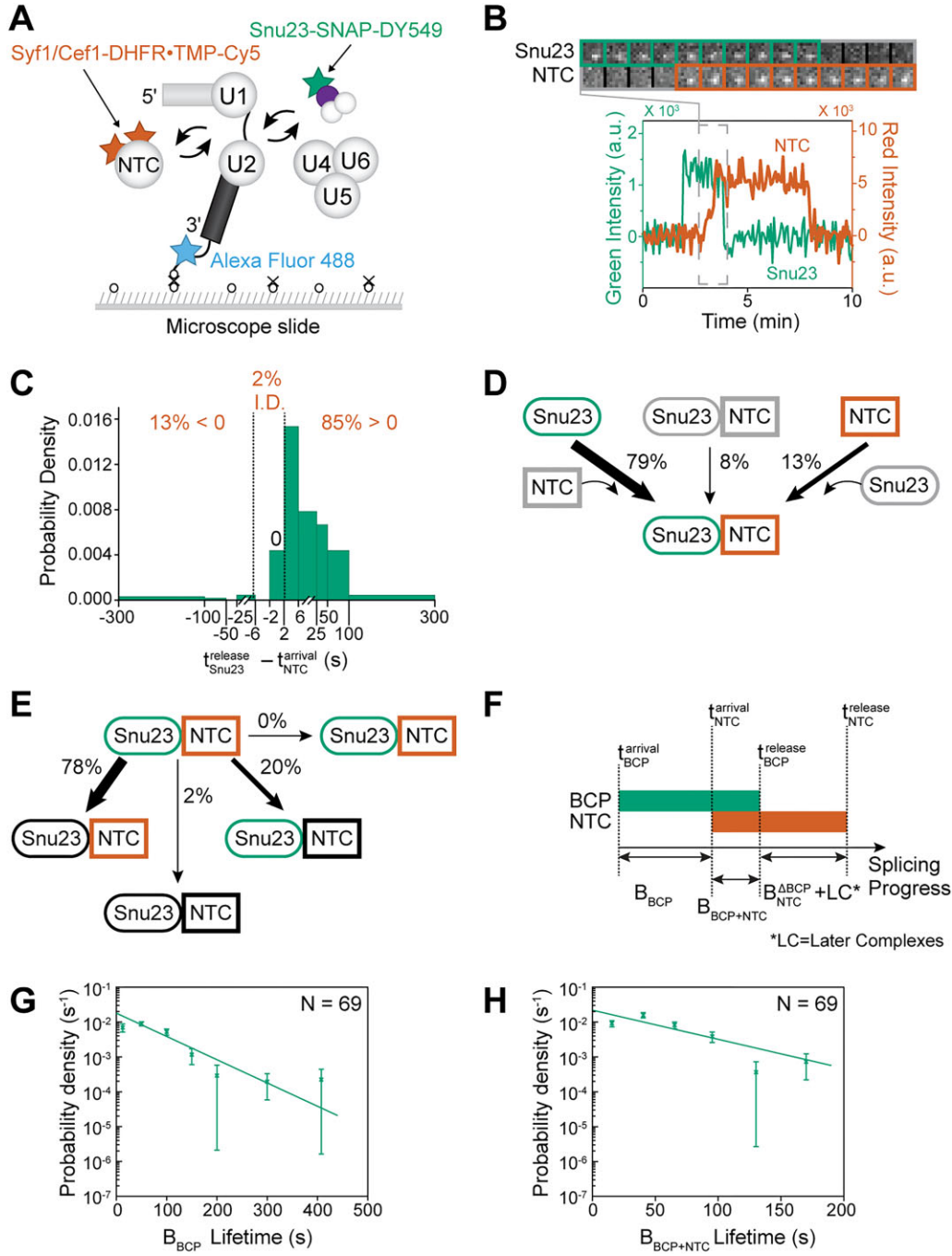
**Figure 5.** BCP association with the tri-snRNP is ATP-dependent. **(A)** Schematic of a three-color CoSMoS assay in which U4 snRNP proteins (Prp3 and Prp4) are labeled with Cy5-TMP, the BCP is labeled with a DY549 via Snu23, and the surface-tethered pre-mRNA is labeled with Alexa Fluor 488. A low concentration of ATP (0.05 mM) is used to inhibit spliceosome activation. **(B)** Segment of a representative time record showing peaks in fluorescence intensity corresponding to colocalization of U4 (red) and Snu23 (green) with the same pre-mRNA molecule. The dashed rectangle marks an example of the simultaneous appearance of U4 and Snu23 spots; galleries show consecutive images taken from the indicated part of the recording showing. **(C)** Routes for the appearance of Snu23 and U4 fluorescent spots at 0.05 mM ATP for  $N = 93$  pairs of overlapping events. Red and green shapes represent observation of fluorescence from the corresponding DY549 or Cy5 fluorophores on Snu23 or U4, respectively. Percentages represent the fraction of complexes in which fluorescence appeared by the indicated pathway; more prevalent pathways are emphasized with thicker arrows. Corresponding data for the BCP labeled on either Spp381 and Snu23 are included in [Supplementary Fig. S11](#).

snRNP proteins), we are measuring the lifetime of an intact BCP. This assumption is based on three observations. First, no structural data yet exist for recruitment of a partial BCP to a spliceosome during activation. Second, the lifetimes of colocalized Prp38/Snu23/Spp381 complexes (Fig. 2) are  $\sim 40$  s, and most closely resemble those of total Prp38, Snu23, or Spp381 lifetimes measured when colocalized with other factors ( $\sim 100$  s) (Figs 3, 4, and 6, and [Supplementary Table S4](#)). Third, we carried out many of the three-color experiments in an orthogonal manner with each BCP component individually labeled ([Supplementary Tables S3 and S4](#)). These data are remarkably consistent no matter if the labeled protein is Prp38, Snu23, or Spp381.

With the above caveats and assumptions in mind, we predict that the lifetime of  $B_{U4+Lsm+BCP}$  is very short,  $\sim 13$  s, before the U4 snRNP is lost along with at least a subset of tri-snRNP specific factors including Snu66. This forms a spliceosome complex lacking U4 and Snu66 but containing the Lsm2-8 and BCP complexes ( $B_{Lsm+BCP}^{\Delta U4}$ ). This complex then persists for 49–53 s, before the NTC associates to form the  $B_{Lsm+BCP+NTC}^{\Delta U4}$  complex.  $B_{Lsm+BCP+NTC}^{\Delta U4}$  in turn persists for another  $\sim 52$  s before dissociation of the BCP and Lsm2–8 complexes to form the  $B_{NTC}^{\Delta U4}$  complex, which likely has a similar composition and structure to the  $B^{act}$  spliceosome [34, 35]. We do not yet know the order of BCP and Lsm2–8 complex departure, as we have

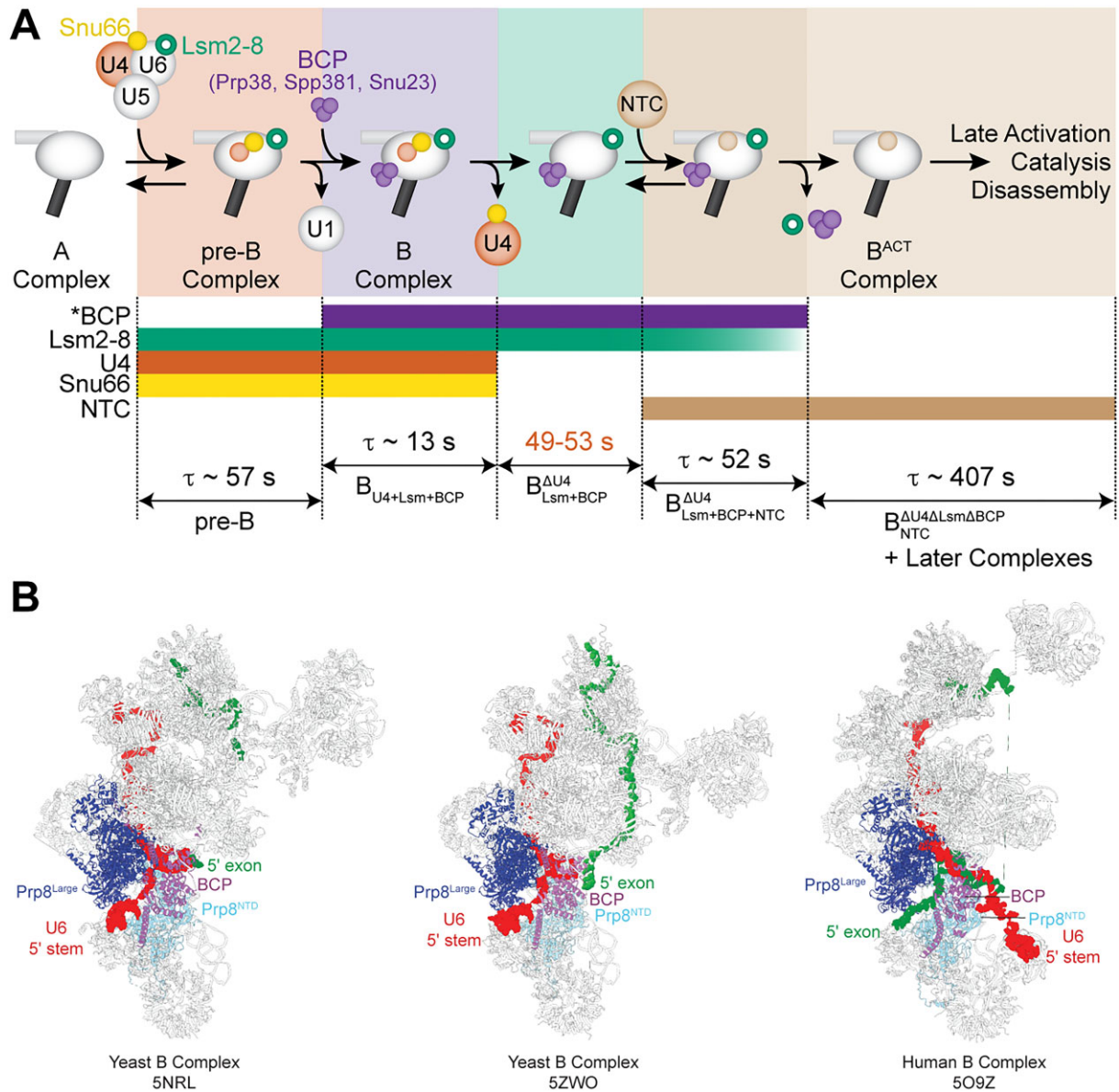
not yet been able to assay this particular step (not shown). Overall, this process takes  $\sim 3$  min from tri-snRNP binding through Lsm and BCP release. It is likely activation occurs more quickly *in vivo* since the entire splicing reaction has been estimated to be complete in a similar time frame [36]. One possibility is that high local concentrations of factors such as the BCP and NTC near the spliceosome may accelerate activation due to increasing effective on-rates. Once  $B^{act}$  is formed, spliceosomes usually persist for  $\sim 407$  s before loss of the NTC due to either successful splicing or disassembly of stalled complexes.

Currently, no structural data is available for yeast  $B_{Lsm+BCP}^{\Delta U4+NTC}$  or  $B_{Lsm+BCP+NTC}^{\Delta U4}$  complexes. However, it is likely that they may be structurally analogous to the human pre- $B^{act-1}$  and pre- $B^{act-2}$  complexes, respectively, with the exception that yeast likely recruit the NTC as a single subunit while the Lsm complex remains bound [5, 19]. Importantly, our model is consistent with evolutionary conservation of BCP recruitment and release between yeast and humans: BCP proteins arrive independent of the tri-snRNP to B complex and are released after U4 snRNP dissociation and arrival of at least a partial complement of NTC proteins [5]. Since our work indicates that yeast  $B_{Lsm+BCP}^{\Delta U4+NTC}$  or  $B_{Lsm+BCP+NTC}^{\Delta U4}$  complexes exist only transiently, obtaining cryo-EM structural data for these may benefit from either synchronizing spliceosome activation followed by rapid



**Figure 6.** Three-color CoSMoS observation of NTC and Snu23 binding dynamics during activation. **(A)** Schematic of a three-color CoSMoS assay in which NTC proteins (Syf1 and Cef1) are labeled with Cy5-TMP, the BCP is labeled with a DY549 via Snu23, and the surface-tethered pre-mRNA is labeled with Alexa Fluor 488. **(B)** Segment of a representative time record showing peaks in fluorescence intensity corresponding to colocalization of NTC (red) and Snu23 (green) with the same pre-mRNA molecule. The light gray dashed rectangle marks an example of the ordered recruitment of the NTC followed by the release of Snu23; galleries show consecutive images taken from that part of the recording. **(C)** Probability density histogram of measured delays between the times of NTC arrivals and Snu23 release events. Most often (85% of  $N = 117$  total events), the NTC arrived before release of the Snu23 ( $t_{\text{release Snu23}} - t_{\text{arrival NTC}} > 0$ ). In 2% of cases, the exact order of events was indeterminate (I.D.) due to simultaneous loss of both signals or both signals remaining at the conclusion of the experiment. **(D)** Routes for the appearance of Snu23 and NTC fluorescent spots at 2 mM ATP for  $N = 116$  pairs of the subset of overlapping events. Red and green shapes represent observation of fluorescence from the corresponding DY549 or Cy5 fluorophores on Snu23 or the NTC, respectively; gray shapes represent the absence of fluorescence. Percentages represent the fraction of complexes in which fluorescence appeared by the indicated pathway; more prevalent pathways are emphasized with thicker arrows. **(E)** Routes for loss of either Snu23 or NTC fluorescent spots at 2 mM ATP for  $N = 92$  pairs of overlapping events in which the Snu23 spot appearance preceded arrival of the NTC. **(F)** Schematic depicting how the lifetimes of the  $B_{\text{BCP}}$ ,  $B_{\text{BCP+NTC}}$ , and  $B_{\text{BCP+NTC}}^{\text{ABCP}}$  complexes are defined relative to the recruitment and release times of constituent factors. **(G and H)** Probability density histograms of  $B_{\text{BCP}}$  (panel G,  $t_{\text{release Snu23}} - t_{\text{arrival NTC}}$ ) and  $B_{\text{BCP+NTC}}$  (panel H,  $t_{\text{release Snu23}} - t_{\text{arrival NTC}}$ ) obtained from the subset of events ( $N$ ) showing ordered arrival of Snu23 and then NTC spots followed by ordered loss of the Snu23 and then NTC signals. Lines represent fits of the lifetime distributions with equations containing single exponential terms that yielded the parameters reported in [Supplementary Table S4](#) ( $\tau_1 \sim 66$  and  $\sim 53$  s for  $B_{\text{BCP}}$  and  $B_{\text{BCP+NTC}}$ , respectively). Error bars ( $\pm$ SD) were calculated for each point as described in the "Materials and methods" section.





**Figure 7.** A kinetic scheme for spliceosome activation. **(A)** In this kinetic scheme, a spliceosome complex containing an intact tri-snRNP (as evidenced by presence of the U4 snRNP and Lsm2-8 complexes) and the BCP ( $B_{U4+Lsm+BCP}$ ) forms very transiently ( $\tau \sim 13$  s) before the U4 snRNP is released to form the  $B_{Lsm+BCP}^{\Delta U4}$  complex. After  $\sim 50$  s, the NTC joins to form the  $B_{Lsm+BCP+NTC}^{\Delta U4}$  spliceosome. The Lsm2-8 and BCP complexes are then released in an undetermined order after another  $\sim 50$  s. The lifetime of the  $B_{Lsm+BCP}^{\Delta U4}$  complex (red) was not directly measured in these experiments but can be deduced from our studies in combination with previously reported experiments that measured the delay between U4 release and NTC recruitment [19, 20]. Note that we cannot exclude the possibility that other pathways may exist including those that involve step-wise recruitment of individual BCP proteins or subcomplexes. **(B)** Structural alignment of yeast (left, middle) and human (right) B complex spliceosomes highlights differences in positioning of the U6 snRNA 5' stem region (red) and 5' exon (green) of the pre-mRNA substrates in yeast and human spliceosomes. Spliceosomes were aligned based on the Prp8 N-terminal domain (Prp8<sup>NTD</sup>) using PyMol and visualized using PyMol and ChimeraX software.

freezing of the samples at short time intervals or use of small molecules or splicing factor mutations that result in accumulation of these states.

#### BCP Association with the tri-snRNP is ATP dependent

It has previously been proposed that BCP components are part of the tri-snRNP and recruited to the spliceosome as part of this assembly. This was largely based on mass spectrometry data and biochemical assays of complexes isolated under no or limiting ATP conditions [10–13, 15, 30]. Conversely, data

from the human splicing machinery indicate that the corresponding human proteins are recruited during activation after tri-snRNP binding [33, 37]. Our CoSMoS data resolve these observations. Under conditions that permit activation and splicing, the BCP is recruited to the spliceosome after incorporation of the tri-snRNP, in agreement with a model proposed by the Shi laboratory based on cryo-EM structures of endogenous yeast pre-B and B complexes [32]. This mechanism simplifies models for yeast spliceosome activation by eliminating the need to reconcile BCP occupancy near the 5' SS with Prp28 activity. Stepwise recruitment of the tri-snRNP and BCP to spliceosomes would allow Prp28 to transfer the 5'

SS from U1 to U6 and dissociate from the spliceosome prior to BCP binding. The BCP can then be positioned to stabilize interactions and conformational changes needed for subsequent steps in activation [5, 7].

However, limiting ATP results in the BCP being recruited to the yeast spliceosome as part of the tri-snRNP. Thus, the BCP is an ATP-dependent component of the yeast tri-snRNP, and this explains its presence in complexes isolated under low ATP conditions. Whether or not the BCP could similarly be recruited to human tri-snRNPs in the absence of spliceosome formation is uncertain. Human tri-snRNPs appear to be stabilized in an unactivated conformation with Brr2 located far from its U4 snRNP substrate and with stably bound Prp28 [37]. These factors may preclude premature BCP association.

### Premature association of the BCP may result in unproductive spliceosomes

We believe that yeast tri-snRNPs formed at low ATP and containing the BCP are unlikely to be competent for splicing. Previous single molecule data showed that the majority (68%) of spliceosomes formed at 0.05 mM ATP were disassembled at 2 mM ATP [20]. In addition, tri-snRNPs purified in the absence of ATP and containing the BCP disassemble when ATP is added [31]. This would support the hypothesis that many of the BCP-containing tri-snRNP complexes that bind to the pre-mRNA are not able to carry out subsequent steps and represent nonfunctional species. It should be noted; however, that we have not directly assayed the splicing competency of the BCP-associated tri-snRNP.

Formation of nonfunctional yeast spliceosome complexes at low ATP is also supported by analysis of cryo-EM data for yeast and human B complexes. Importantly, the yeast B complex spliceosomes were assembled at low, 0.05 mM ATP (5NRL) or purified from yeast in the absence of ATP (5ZWO) whereas the human B complex was isolated at high ATP (2 mM) using a low  $Mg^{2+}$  concentration (5O9Z) [4, 14, 32]. When comparing these structures, it is obvious that the 5' stem of the U6 snRNA and 5' exon of the pre-mRNA have exchanged locations (Fig. 7B). In catalytic spliceosomes from both yeast and humans (B\* or C complexes that form after the B complex), the 5' exon is accommodated in an “exon channel” formed by Prp8. In the human B spliceosome, the 5' exon is already located in the exon channel and the 5' SS is base paired with the U6 snRNA ACAGAGA sequence proximal to the channel (Fig. 7B, right). These interactions are stabilized by the human BCP. However, in the published yeast B spliceosome complexes the 5' SS has not yet base paired with the U6 ACAGAGA sequence and is instead pairing with an adjacent upstream element (the ACAGAGA stem). Additionally, the 5' exon is located outside the channel while the U6 5' stem is located inside, stabilized by the BCP (Fig. 7B, left and middle). Previously, the differences in 5' SS/U6 base pairing and around the exon channel were interpreted as divergent activation mechanisms taken by yeast and human splicing machinery with yeast requiring additional steps to correctly pair U6 and position the 5' exon [38].

We propose that yeast B complex structures formed at limiting ATP represent off-pathway or nonfunctional intermediates due to pre-mature association of the BCP with the tri-snRNP. The BCP may structurally lock the U6 5' stem in the exon channel of the B complex as well as block correct transfer of the 5' SS to U6 by preventing Prp28 binding. This results

in formation of spliceosomes not competent for carrying out the splicing reaction and that are substrates for disassembly at higher concentrations of ATP. It is interesting to note that insertion of the U6 snRNA 5' stem into the exon channel was also observed in structures of endogenous spliceosomes purified from lysates (Fig. 7B, middle) [32]. This suggests that formation of this structure may be biologically relevant and not specific to just one pre-mRNA substrate.

Whether or not BCP-dependent formation of non-functional tri-snRNP or spliceosome complexes at low ATP is used to regulate splicing *in vivo* is unknown. It is possible that switching of the tri-snRNP between catalytically competent and incompetent conformations depending on the presence or absence of the BCP could be used to globally regulate cellular splicing in response to ATP levels. In *S. cerevisiae*, cellular ATP concentration is usually maintained at a constant level of ~4 mM [39]. However, under certain conditions, such as deletion of the Bas1 transcription factor that regulates purine biosynthesis or in the presence of glycolysis inhibitors such as 2-deoxyglucose, intracellular levels can fall dramatically, including to levels that we predict would cause BCP association with the tri-snRNP ( $\leq 0.05$  mM) [40]. Under such conditions, “spliceosome shutdown” caused by formation of BCP-associated tri-snRNPs and nonfunctional spliceosomes could lead to reduced ribosome synthesis (since the majority of yeast introns are found within ribosomal protein genes [41]) and a redirection of cellular ATP towards other metabolic or cellular needs.

### Acknowledgements

We thank Laura Vanderploeg (Department of Biochemistry, UW-Madison) for help in creating the figures and illustrations. We also thank Dr. Kathy Senn and Ethan Aubuchon for critical reading of the manuscript. We thank Cheng Ern Yap with molecular biology assistance.

**Author contributions:** X.F. and A.A.H. conceptualized the project. X.F. collected and analyzed the data. X.F. and A.A.H. wrote the manuscript.

### Supplementary data

Supplementary data is available at NAR online.

### Conflict of interest

A.A.H. is a member of the scientific advisory board and carrying out sponsored research for Remix Therapeutics.

### Funding

This work was supported by grants from the National Institutes of Health (R35 GM136261 to A.A.H.). Funding to pay the Open Access publication charges for this article was provided by the NIH grant.

### Data availability

Single molecule data (microscope video recordings) are deposited in Zenodo at DOI: 10.5281/zenodo.14766507.

## References

- Hoskins AA, Friedman LJ, Gallagher SS *et al.* Ordered and dynamic assembly of single spliceosomes. *Science* 2011;331:1289–95. <https://doi.org/10.1126/science.1198830>
- Wilkinson ME, Charenton C, Nagai K. RNA splicing by the spliceosome. *Annu Rev Biochem.* 2020;89:359–88. <https://doi.org/10.1146/annurev-biochem-091719-064225>
- Shcherbakova I, Hoskins AA, Friedman LJ *et al.* Alternative spliceosome assembly pathways revealed by single-molecule fluorescence microscopy. *Cell Rep* 2013;5:151–65. <https://doi.org/10.1016/j.celrep.2013.08.026>
- Bertram K, Agafonov DE, Dybkov O *et al.* Cryo-EM structure of a pre-catalytic human spliceosome primed for activation. *Cell* 2017;170:701–13. <https://doi.org/10.1016/j.cell.2017.07.011>
- Townsend C, Leelaram MN, Agafonov DE *et al.* Mechanism of protein-guided folding of the active site U2/U6 RNA during spliceosome activation. *Science* 2020;370:eabc3753. <https://doi.org/10.1126/science.abc3753>
- Zhang X, Yan C, Zhan X *et al.* Structure of the human activated spliceosome in three conformational states. *Cell Res* 2018;28:307–22. <https://doi.org/10.1038/cr.2018.14>
- Zhang Z, Kumar V, Dybkov O *et al.* Structural insights into the cross-exon to cross-intron spliceosome switch. *Nature* 2024;630:1012–19. <https://doi.org/10.1038/s41586-024-07458-1>
- Kastner B, Will CL, Stark H *et al.* Structural insights into nuclear pre-mRNA splicing in higher eukaryotes. *Cold Spring Harb Perspect Biol* 2019;11:a032417. <https://doi.org/10.1101/cshperspect.a032417>
- Ulrich AK, Wahl MC. Human MFAP1 is a cryptic ortholog of the *Saccharomyces cerevisiae* Spp381 splicing factor. *BMC Evol Biol* 2017;17:91. <https://doi.org/10.1186/s12862-017-0923-1>
- Gottschalk A, Neubauer G, Banroques J *et al.* Identification by mass spectrometry and functional analysis of novel proteins of the yeast [U4/U6.U5] tri-snRNP. *EMBO J* 1999;18:4535–48. <https://doi.org/10.1093/emboj/18.16.4535>
- Stevens SW, Abelson J. Purification of the yeast U4/U6.U5 small nuclear ribonucleoprotein particle and identification of its proteins. *Proc Natl Acad Sci USA* 1999;96:7226–31. <https://doi.org/10.1073/pnas.96.13.7226>
- Nguyen THD, Galej WP, Bai XC *et al.* Cryo-EM structure of the yeast U4/U6.U5 tri-snRNP at 3.7 Å resolution. *Nature* 2016;530:298–302. <https://doi.org/10.1038/nature16940>
- Wan R, Yan C, Bai R *et al.* The 3.8 Å structure of the U4/U6.U5 tri-snRNP: insights into spliceosome assembly and catalysis. *Science* 2016;351:466–75. <https://doi.org/10.1126/science.aad6466>
- Plaschka C, Lin PC, Nagai K. Structure of a pre-catalytic spliceosome. *Nature* 2017;546:617–21. <https://doi.org/10.1038/nature22799>
- Lybarger S, Beickman K, Brown V *et al.* Elevated levels of a U4/U6.U5 snRNP-associated protein, Spp381p, rescue a mutant defective in spliceosome maturation. *Mol Cell Biol* 1999;19:577–84. <https://doi.org/10.1128/MCB.19.1.577>
- Blanton S, Srinivasan A, Rymond BC. PRP38 encodes a yeast protein required for pre-mRNA splicing and maintenance of stable U6 small nuclear RNA levels. *Mol Cell Biol* 1992;12:3939–47.
- Stevens SW, Barta I, Ge HY *et al.* Biochemical and genetic analyses of the U5, U6, and U4/U6 x U5 small nuclear ribonucleoproteins from *Saccharomyces cerevisiae*. *RNA* 2001;7:1543–53.
- Ulrich AKC, Seeger M, Schütze T *et al.* Scaffolding in the spliceosome via single  $\alpha$  helices. *Structure* 2016;24:1972–83. <https://doi.org/10.1016/j.str.2016.09.007>
- Fu X, Kaur H, Rodgers ML *et al.* Identification of transient intermediates during spliceosome activation by single molecule fluorescence microscopy. *Proc Natl Acad Sci USA* 2022;119:e2206815119. <https://doi.org/10.1073/pnas.2206815119>
- Hoskins AA, Rodgers ML, Friedman LJ *et al.* Single molecule analysis reveals reversible and irreversible steps during spliceosome activation. *eLife* 2016;5:e14166. <https://doi.org/10.7554/eLife.14166>
- Mayas RM, Maita H, Staley JP. Exon ligation is proofread by the DExD/H-box ATPase Prp22p. *Nat Struct Mol Biol* 2006;13:482–90. <https://doi.org/10.1038/nsmb1093>
- Anderson EG, Hoskins AA. Single molecule approaches for studying spliceosome assembly and catalysis. *Methods Mol Biol* 2014;1126:217–41. [https://doi.org/10.1007/978-1-62703-980-2\\_17](https://doi.org/10.1007/978-1-62703-980-2_17)
- S raphin B, Rosbash M. The yeast branchpoint sequence is not required for the formation of a stable U1 snRNA-pre-mRNA complex and is recognized in the absence of U2 snRNA. *EMBO J* 1991;10:1209–16. <https://doi.org/10.1002/j.1460-2075.1991.tb08062.x>
- Crawford DJ, Hoskins AA, Friedman LJ *et al.* Visualizing the splicing of single pre-mRNA molecules in whole cell extract. *RNA* 2008;14:170–9. <https://doi.org/10.1261/rna.794808>
- Kaur H, Jamalidinan F, Condon SGF *et al.* Analysis of spliceosome dynamics by maximum likelihood fitting of dwell time distributions. *Methods* 2019;153:13–21. <https://doi.org/10.1016/j.ymeth.2018.11.014>
- Rymond BC. Convergent transcripts of the yeast PRP38–SMD1 locus encode two essential splicing factors, including the D1 core polypeptide of small nuclear ribonucleoprotein particles. *Proc Natl Acad Sci USA* 1993;90:848–52. <https://doi.org/10.1073/pnas.90.3.848>
- Tarn WY, Lee KR, Cheng SC. Yeast precursor mRNA processing protein PRP19 associates with the spliceosome concomitant with or just after dissociation of U4 small nuclear RNA. *Proc Natl Acad Sci USA* 1993;90:10821–5. <https://doi.org/10.1073/pnas.90.22.10821>
- Larson J, Kirk M, Drier EA *et al.* Design and construction of a multiwavelength, micromirror total internal reflectance fluorescence microscope. *Nat Protoc* 2014;9:2317–28. <https://doi.org/10.1038/nprot.2014.155>
- Liu S, Rauhut R, Vornlocher HP *et al.* The network of protein–protein interactions within the human U4/U6.U5 tri-snRNP. *RNA* 2006;12:1418–30. <https://doi.org/10.1261/rna.55406>
- Xie J, Beickman K, Otte E *et al.* Progression through the spliceosome cycle requires Prp38p function for U4/U6 snRNA dissociation. *EMBO J* 1998;17:2938–46. <https://doi.org/10.1093/emboj/17.10.2938>
- Nguyen TH, Galej WP, Bai XC *et al.* The architecture of the spliceosomal U4/U6.U5 tri-snRNP. *Nature* 2015;523:47–52. <https://doi.org/10.1038/nature14548>
- Bai R, Wan R, Yan C *et al.* Structures of the fully assembled. *Science* 2018;360:1423–9. <https://doi.org/10.1126/science.aau0325>
- Charenton C, Wilkinson ME, Nagai K. Mechanism of 5' splice site transfer for human spliceosome activation. *Science* 2019;364:362–7. <https://doi.org/10.1126/science.aax3289>
- Rauhut R, Fabrizio P, Dybkov O *et al.* Molecular architecture of the *saccharomyces cerevisiae* activated spliceosome. *Science* 2016;353:1399–405. <https://doi.org/10.1126/science.aag1906>
- Fabrizio P, Dannenberg J, Dube P *et al.* The evolutionarily conserved core design of the catalytic activation step of the yeast spliceosome. *Mol Cell* 2009;36:593–608. <https://doi.org/10.1016/j.molcel.2009.09.040>
- Carrocci T, Neugebauer K. Emerging and re-emerging themes in co-transcriptional pre-mRNA splicing. *Mol Cell* 2024;84:3656–66. <https://doi.org/10.1016/j.molcel.2024.08.036>
- Agafonov DE, Kastner B, Dybkov O *et al.* Molecular architecture of the human U4/U6.U5 tri-snRNP. *Science* 2016;351:1416–20. <https://doi.org/10.1126/science.aad2085>
- Plaschka C, Newman AJ, Nagai K. Structural basis of nuclear pre-mRNA splicing: lessons from Yeast. *Cold Spring Harb Perspect Biol* 2019;11:a032391. <https://doi.org/10.1101/cshperspect.a032391>



39. Takaine M, Ueno M, Kitamura K *et al.* Reliable imaging of ATP in living budding and fission yeast. *J Cell Sci* 2019;132:jcs.230649. <https://doi.org/10.1242/jcs.230649>
40. Takaine M, Imamura H, Yoshida S. High and stable ATP levels prevent aberrant intracellular protein aggregation in yeast. *eLife* 2022;11:e67659. <https://doi.org/10.7554/eLife.67659>
41. Ares M Jr, Grate L, Pauling MH. A handful of intron-containing genes produces the lion's share of yeast mRNA. *RNA* 1999;5:1138–9. <https://doi.org/10.1017/S1355838299991379>