RNA modifications and Prp24 coordinate Lsm2–8 binding dynamics during *S. cerevisiae* U6 snRNP assembly

Ye Liu[1], Yuichiro Nomura[1], Samuel E. Butcher[1], and Aaron A. Hoskins[1,2,*]

From the [1]Department of Biochemistry, and [2]Department of Chemistry, University of Wisconsin-Madison, Madison, Wisconsin, USA

Reviewed by members of the JBC Editorial Board. Edited by Craig Cameron

In eukaryotes, the process of intron removal from nuclear pre-mRNA is performed by the spliceosome, a dynamic molecular machine composed of small nuclear ribonucleoproteins (snRNPs; U1, U2, U4, U5, and U6) and dozens of other protein splicing factors. The U6 snRNP contains the U6 small nuclear RNA (snRNA) and the proteins Prp24 and Lsm2–8 heteroheptamer. A key feature of the snRNP is a modified U6 snRNA 3′ end, which in *Saccharomyces cerevisiae* (yeast) contains a 3′ phosphate. U6 plays an essential role in splicing, and the U6 snRNP must be completely disassembled for splicing to occur. Once splicing is finished, the snRNP must then be reassembled to participate in a subsequent splicing reaction. While splicing efficiency depends on rapid U6 snRNP assembly, this process has not yet been kinetically characterized. Here, we use colocalization single-molecule spectroscopy to dissect the kinetic pathways of yeast U6 snRNA association with the Lsm2–8 complex and their dependence on the Prp24 protein and post-transcriptional snRNA modification. In the absence of 3′-end processing, Lsm2–8 association with the RNA is highly dependent on Prp24. However, processed RNAs can rapidly recruit Lsm2–8 in the absence of Prp24. Post-transcriptional processing facilitates Lsm2–8 association, whereas the presence of Prp24 promotes both recruitment and retention of the complex. This suggests that efficient U6 snRNP assembly could depend on kinetic selection of Lsm2–8 binding to 3′-end modified or Prp24-bound U6 snRNAs in order to discriminate against association with other RNAs.

In eukaryotes, the removal of introns from pre-mRNA is accomplished by a dynamic molecular machine known as the spliceosome. This machine is partly composed of five small nuclear ribonucleoproteins (snRNPs) named U1, U2, U4, U5, and U6, which each consist of small nuclear RNA (snRNA) molecules and associated protein factors. Throughout the splicing process, the spliceosome undergoes a precisely orchestrated series of assembly, activation, catalysis, dissociation, and recycling steps involving the snRNP particles (1–3). Among the snRNAs, U6 is the most highly conserved and the most dynamic (4, 5). For splicing to occur, U6 snRNPs containing the U6 snRNA, Prp24 protein, and Lsm2–8 protein heteroheptamer must first associate with the U4 snRNP and form base-pairing interactions between the U4 and U6 snRNAs (5). The U4–U6 di-snRNP must then associate with the U5 snRNP to form the U4/U6.U5 tri-snRNP, and Prp24 must dissociate at some point during this process (6). The tri-snRNP is subsequently integrated into the spliceosome and then drastically remodeled: the U4 snRNP and Lsm2–8 complexes are released to allow base pairing between the U2 and U6 snRNAs, between the U6 snRNA and the intron, and intramolecular base pairing within U6 itself. When splicing concludes, U6 snRNA must be released from the spliceosome by translocation of the Prp43 helicase along the RNA in the 3′ → 5′ direction (7–9). At some point, U6 snRNA then reassociates with Prp24 and Lsm2–8 to regenerate the U6 snRNP.

Initial biogenesis of the U6 snRNP follows a similar pathway involving Prp24 and Lsm2–8 (5). However, the nascent U6 snRNA must also be post-transcriptionally modified. U6 snRNA is generated by RNA polymerase (RNAP) III, which terminates transcription in response to a poly-U tract found at the 3′ end of U6 (10). After transcription, the 3′ terminal diol of U6 is initially bound by the protein chaperone Lhp1, providing protection against degradation (11). Following the association of Prp24, Lhp1 is released, and the 3′ end is exposed for processing by the Usb1 exonuclease (12). Usb1 removes a terminal uridine from U6 while generating a 3′ phosphate at the new terminus (12, 13). While not strictly essential for Lsm2–8 binding, this processing increases the equilibrium affinity of Lsm2–8 for U6 by threefold to sevenfold (12).

How Lsm2–8 correctly recognizes and binds to the U6 snRNA is an important biological problem since eukaryotic cells contain different Lsm complexes with distinct functions. The spliceosomal Lsm2–8 complex functions to stabilize the U6 snRNA and has been proposed to facilitate U2–U6 snRNA duplex formation (or promote duplex stability) during assembly of the spliceosome active site (14). The Lsm2–8 complex shares six of its seven subunits with the Lsm1–7 complex, differing only in exchange of Lsm8 for Lsm1. While Lsm2–8 promotes snRNA stability, Lsm1–7 binding to RNAs promotes their degradation (15). Both Lsm1–7 and 2–8 complexes can preferentially bind poly-U terminated RNAs, such as the U6 snRNA and other RNAPIII transcripts. In the case of the *Schizosaccharomyces pombe* proteins, 3′-end

* For correspondence: Aaron A. Hoskins, ahoskins@wisc.edu.

modification increases the affinity of Lsm2–8 $\sim$4-fold (from K_D = 100 to 26 nM), whereas decreasing the affinity of Lsm1−7 by more than 10-fold (from K_D = 25 to 329 nM) (16). This suggests that post-transcriptional modification plays an important role in targeting Lsm complexes to the correct RNAs and to reduce binding to the incorrect transcripts. Interestingly, it has also been shown in yeast that Lsm2–8 is important for nuclear degradation of unspliced or nuclear-retained mRNAs (17), reinforcing the notion that correct recognition of RNA 3′ ends by Lsm proteins can have consequences on the transcriptome.

The assembly of the *Saccharomyces cerevisiae* (*S. cerevisiae*, yeast) U6 snRNP has been reconstituted from purified components *in vitro* and is highly efficient (12, 18, 19). This represents a powerful system for dissecting how RNA and proteins dynamically interact with one another during molecular assembly and for connecting *in vivo* genetic results to quantifiable changes in structure, kinetic, or thermodynamic stabilities. While several *in vitro* studies have elucidated the equilibrium parameters that describe binding between the U6 snRNA and its associated proteins (*i.e.*, K_Ds) (12, 18, 20–22), there is little information about the kinetic details of U6 snRNP assembly starting either from a nascent transcript or the 3′-end-modified snRNA. In cells, it is possible that kinetic pathways dominate how U6 snRNPs are assembled since competing, nonreversible pathways of RNA export, degradation, and assembly into larger splicing complexes may prevent equilibrium conditions from being established. Indeed, it is well known that many steps in gene expression do not occur at equilibrium in cells (23).

In this study, we have analyzed the kinetic properties of yeast U6 snRNP assembly *in vitro* using purified components and colocalization single-molecule spectroscopy (CoSMoS). Single-molecule experiments like CoSMoS offer distinct advantages compared with ensemble assays by enabling the detection of transient intermediates, facilitating the identification of alternative assembly pathways, and visualizing complexes in real time as they form rather than with discontinuous experiments such as EMSAs (24, 25). By studying the kinetics of Lsm2–8 complex binding with U6 snRNA as a function of RNA 3′-end composition and the presence or absence of Prp24, we show that post-transcriptional processing of U6 significantly changes Lsm2–8 binding kinetics. Addition of Prp24 appears to optimize this interaction, further stabilizing the complex. In particular, the SNFFL-box motif in the C-terminal domain of Prp24 can modulate the accessibility of U6 for Lsm2–8 binding depending on the RNA's post-transcriptional modification status. Together, our results suggest that 3′-end modification of U6 may increase the efficiency of U6 snRNP assembly by providing access to an alternate pathway in which Lsm2–8 can rapidly bind prior to Prp24 association.

Results

Fluorophore-labeled Lsm2–8 binds 3′-end modified RNAs

For imaging single molecules of yeast Lsm2–8, we constructed an Lsm2–8 expression plasmid in which Lsm8 is genetically fused at the C terminus to the fast SNAP tag (Lsm2–8^{fSNAP}; Figure 1, A and B). We have previously shown that a similar C-terminal fSNAP fusion to Lsm8 is functional
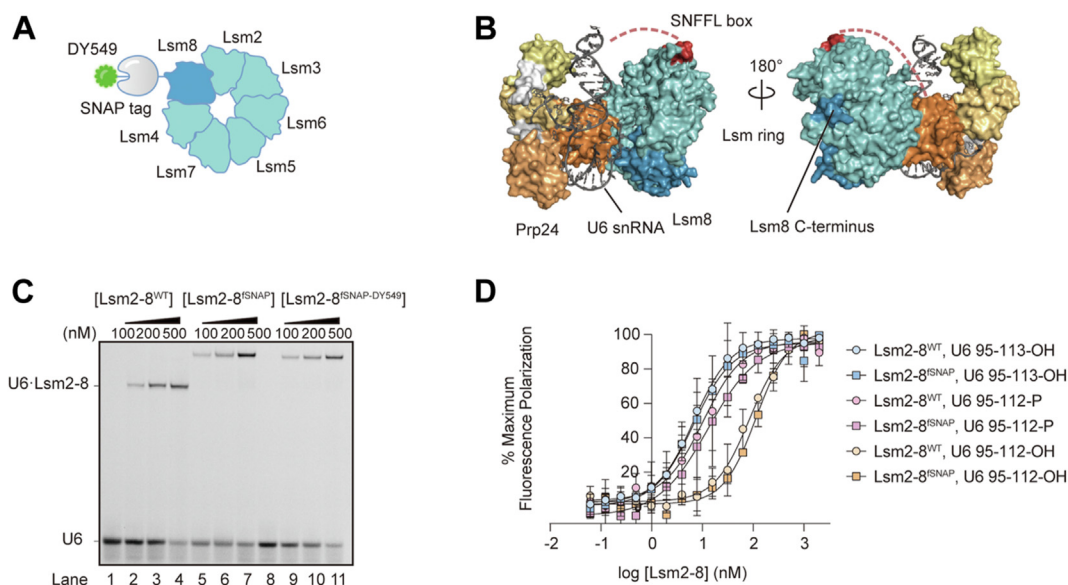

Figure 1. Fluorophore-labeled yeast Lsm2–8 binds 3′-end modified RNAs. *A*, cartoon structure of Lsm2–8$^{fSNAP\_DY549}$. The SNAP tag is attached to the C terminus of Lsm8 *via* a flexible linker. *B*, crystal structure of the yeast U6 snRNP. U6 snRNA is shown in *dark gray*. Prp24 is depicted in a surface representation in *light gray* and *yellow-orange* colors. The SNFFL-box domain of Prp24 is shown in *red*, bound to the Lsm2–8 complex. The flexible linker connecting this domain to rest of Prp24 could not be modeled based on the crystallographic data and is represented as *red dashes*. The Lsm2–8 ring is shown in surface representation in *blue–green* colors with Lsm8 in *dark blue* (Protein Data Bank ID: 5VSU). *C*, EMSA analysis of Lsm2–8^{fSNAP} and Lsm2–8$^{fSNAP\_DY549}$ binding ability relative to WT protein to U6 RNAs. *D*, fluorescence polarization binding data comparing Lsm2–8^{WT} and Lsm2–8^{fSNAP} binding to U6 RNA fragments (nt 95–112 or −113) with either diol or 3′ phosphate-end modifications. Each point represents the average of three technical replicates, and bars represent ±SD. snRNA, small nuclear RNA; snRNP, small nuclear ribonucleoprotein.

in vivo and assembles into spliceosomes (14). We expressed and purified assembled Lsm2–8^{fSNAP} complexes from *Escherichia coli*. Lsm2–8^{fSNAP} was then labeled SNAP-Surface 549 (DY549), and the excess dye was removed by gel filtration chromatography (Fig. S1A). UV–visible spectroscopy was used to determine the extent of fluorophore labeling (~80%).

We tested the impact of fSNAP labeling on Lsm2–8 binding to U6 RNAs. EMSAs confirmed that the fluorophore-labeled Lsm2–8^{fSNAP} bound an RNA oligo representing U6 nucleotides (nt) 95 to 113 with a terminal 2′, 3′ diol (U6 95–113-OH) similarly to the untagged complex (Fig. 1C). We then used fluorescence polarization assays to study the affinities of unlabeled Lsm2–8^{fSNAP} complexes for RNAs with different modification states of U6 (Fig. 1D). Like Lsm2–8^{WT}, Lsm2–8^{fSNAP} showed the highest affinities for RNAs with either a five uridine 3′ end (U6 95–113-OH) or with a four uridine 3′ end containing a 3′ phosphate (U6 95–112-P; Table S1). RNAs containing four uridines and a 2′, 3′ diol at the 3′ end (U6 95–112-OH) bound both complexes with lower affinity. From these assays, we conclude that the labeled Lsm2–8^{fSNAP} complex interacts similarly with U6 RNA as does the untagged version.

Single-molecule assays of U6 RNA binding recapitulate ensemble observations

To image Lsm2–8^{fSNAP} proteins associated with U6 snRNAs, we used two-color CoSMoS assays (26). In these experiments, full-length, Cy5-labeled U6 RNAs (nts 1–113; 1–113-OH) were prepared by splinted ligation (Fig. S1B) and then immobilized to passivated glass slides with biotin–streptavidin. Interactions with DY549-labeled Lsm2–8^{fSNAP} could then be visualized by observing colocalization between the red (632 nm) laser– and green (532 nm) laser–excitable fluorophores on the RNA and Lsm2–8^{fSNAP}, respectively (Fig. 2A). Data were collected using a micromirror multiwavelength total internal reflection fluorescence microscope (27, 28).

In a specific field of view, we simultaneously tracked binding events occurring on hundreds of U6 1–113-OH RNA molecules (Fig. 2B). We frequently observed spots of fluorescence from Lsm2–8^{fSNAP} dynamically colocalizing with tethered RNA locations—indicating that Lsm2–8 association with the RNAs is reversible. As a control, if RNA was omitted from the surface, we observed significantly fewer (~15-fold lower) Lsm2–8^{fSNAP} binding events (Fig. S2). This suggests that colocalization events observed between Lsm2–8^{fSNAP} and the immobilized RNAs represent RNA-dependent interactions. To test if these interactions are also influenced by the free 3′ end of the U6 RNA, we first incubated the immobilized RNAs with the yeast Lhp1 protein, which tightly binds U6 RNAs containing a terminal 3′ diol. We predicted that this should sterically block the association of the fluorescent Lsm2–8^{fSNAP} complex with U6 1–113-OH. This prediction was correct, and addition of Lhp1 reduced the number of colocalized Lsm2–8^{fSNAP} binding events to background levels (Fig. S2). Finally, we used yeast Prp24 protein to release Lhp1 from the

immobilized U6 RNAs and restore Lsm2–8^{fSNAP} binding (Fig. S2). Together, these results indicate that Lsm2–8^{fSNAP} molecules colocalize with immobilized U6 RNAs depending on the accessibility of the RNA 3′ end and that Prp24 can release Lhp1 from immobilized U6 RNAs. We conclude that these single-molecule assays recapitulate key features of the U6 snRNP assembly pathway defined by bulk experiments (12).

3′-End modification of U6 changes Lsm2–8 occupancy and interaction kinetics

In yeast, U6 snRNAs are synthesized by RNAPIII, and transcription termination generates heterogeneously sized U6 tails ranging from 4 to 8 nt (5). Processing by Usb1 leads to the reduction of the U-tail length by one nucleotide and leaves behind a terminal 3′ phosphate group (12). In order to investigate the influence of different tail lengths and chemical groups on the binding behavior of Lsm2–8^{fSNAP}, we carried out CoSMoS assays with Lsm2–8^{fSNAP} on three different types of full-length, immobilized U6 RNAs: a mimic of the unprocessed form (U6 1–113-OH), a mimic of the product of yUsb1 processing (U6 1–112-P), and a final form lacking both one uridine in the 3′ tail and the 3′ phosphate group (U6 1–112-OH) (Fig. 2C).

Representative fluorescence trajectories of Lsm2–8^{fSNAP} fluorescence colocalized with individual U6 RNAs are shown in Figure 2D. These trajectories suggest that the modification status of the U6 RNA 3′ end can change Lsm2–8^{fSNAP} interaction dynamics: fewer and shorter events are observed with the U6 1–112-OH RNA relative to the other two. To provide overviews of binding events occurring on many different RNAs, we created rastergrams where the colocalization of Lsm2–8^{fSNAP} is represented by a purple segment and its absence by a white segment for 90 different RNA molecules. These rastergrams were further sorted by the timing of the first Lsm2–8^{fSNAP} binding event, with RNAs binding Lsm2–8^{fSNAP} earlier being assigned a lower number on the *y*-axis (Fig. 2, E–G).

From the rastergrams, it is evident that many more binding events are observed on the U6 1–112-P RNA relative to the others. In contrast, binding events to an RNA of the same length but lacking the 3′ phosphoryl modification (U6 1–112-OH) are both shorter and less frequent. Increasing the RNA tail length restores some long-lived binding events (U6 1–113 OH) but not the event density seen with the phosphoryl modification. Together, the rastergrams indicate that the U6 RNA mimicking modification by Usb1 has increased interaction with Lsm2–8^{fSNAP}.

To quantify these changes more accurately, we first determined the total fractions of experimental times each RNA was colocalized with Lsm2–8^{fSNAP} (Fig. 2H). As expected from the rastergrams, U6 1–112-P RNAs spent nearly twice as often colocalized with Lsm2–8^{fSNAP} proteins as did the other RNAs. We then analyzed the distributions of dwell times for binding events as well as the times in between binding events (Fig. 2, I and J). These distributions show that Lsm2–8 complexes tended to bind the U6 1–112-P RNA more quickly (lower
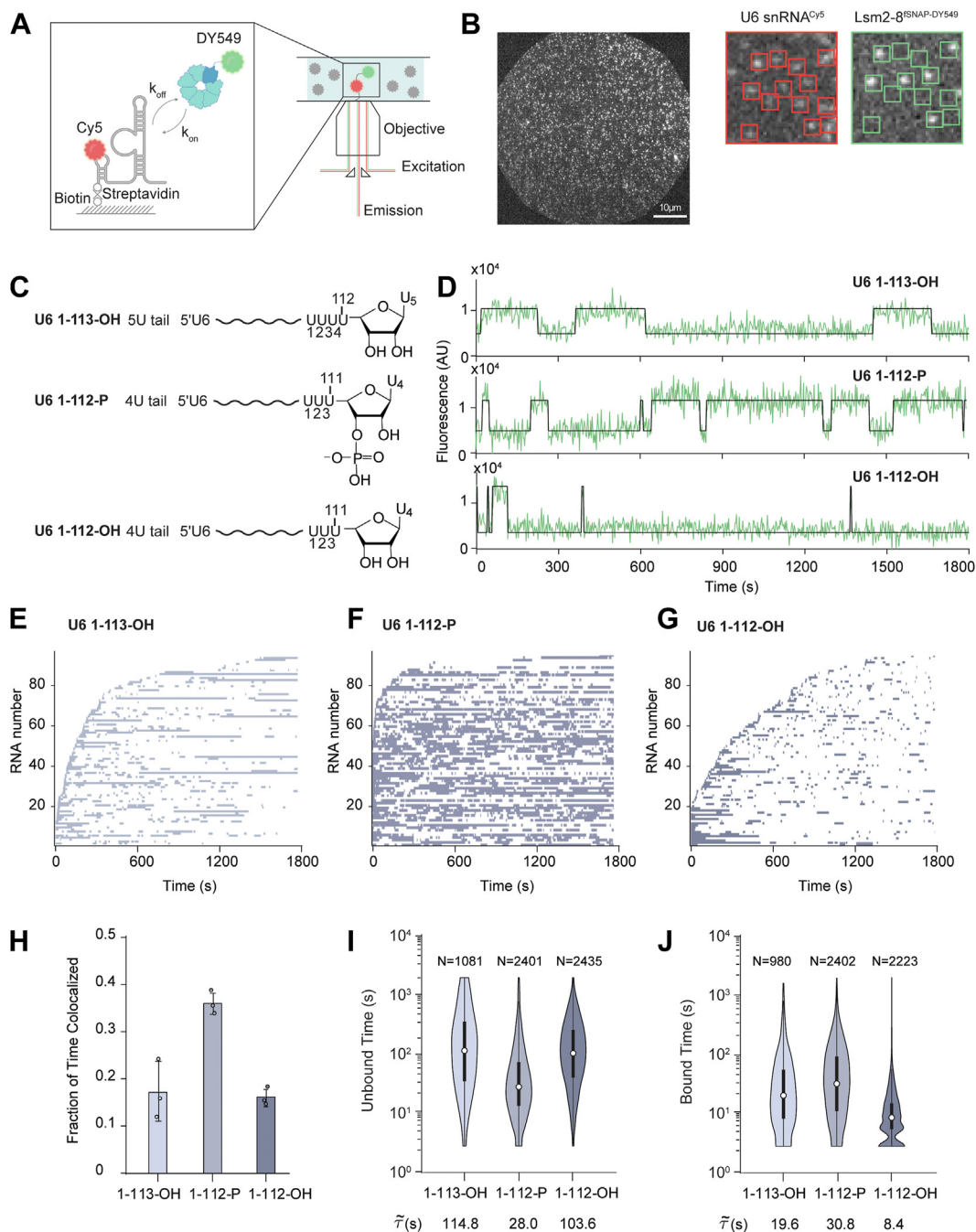
Kinetic analysis of Lsm2–8 binding to the U6 snRNA

Figure 2. U6 RNA 3′-end modification impacts Lsm2–8[fsnap] **binding kinetics.** *A*, schematic representation of the CoSMoS assay for monitoring Lsm2–8[fSNAP DY549] binding to immobilized, Cy5-labeled U6 RNAs. *B*, *left*, representative micrographs showing individual U6 snRNA molecules tethered to the slide surface in a field of view (FOV). *Right*, zoom-in of a portion of the FOV showing individual U6 RNA molecules (*red boxes* on the *left*) and colocalized Lsm2–8[fSNAP] molecules (*bright spots* inside the *green boxes* on the *right*). *C*, chemical structures of the RNAs used in the CoSMoS assays highlighting the structure of the RNA 3′ ends. *D*, representative fluorescence trajectories showing Lsm2–8[fSNAP DY549] binding events to the RNAs. *E–G*, rastergrams illustrating the binding of Lsm2–8 to 90 different molecules of U6 1–113-OH (*E*), U6 1–112-P (*F*), and U6 1–112-OH (*G*), sorted based on the time of first Lsm2–8[fSNAP DY549] binding event detection. *H*, average fraction of time U6 molecules spend colocalized with Lsm2–8[fSNAP DY549]. Bar heights represent the average from three technical replicates (*dots*), and bars represent ±SD. *I* and *J*, violin plots representing the dwell time distributions for unbound (*I*) and bound (*J*) states across three different RNAs. Each violin plot is overlaid with a box plot that indicates the median (*horizontal line*), interquartile range (IQR, *box*), and *whiskers* extending to 1.5 times the IQR. The numbers above each violin plot denote the number of lifetimes included in the respective distribution. The median dwell times are shown below each plot. CoSMoS, colocalization single-molecule spectroscopy.

unbound times; Fig. 2*I*) than either the 1–113-OH or 1–112-OH RNAs. Once bound, Lsm2–8 complexes tended to remain associated with the 1–113-OH and 1–112-P RNAs longer relative to the 1–112-OH RNA. To analyze the dwell times in more detail, we fit each distribution of dwell times to an equation containing two exponential terms to yield characteristic lifetimes for the bound (τ_{bound}) and unbound (τ_{unbound}) states (Tables S2 and S3). For the unbound times, we

observed a multiexponential distribution for each RNA with characteristic short- and long-$\tau_{unbound}$ parameters. In comparing the unmodified and modified U6 RNA mimics (U6 1–113-OH and U6 1–112-P, respectively), the most significant differences are found in a large increase in the amplitude of the short-$\tau_{unbound}$ parameter (increasing from 0.50 to 0.81) and a decrease in the value of $\tau_{unbound}$ from ~86 to 32 s because of the modification. One interpretation of these data is that U6-end modification increases both the rate of Lsm2–8 binding (as observed in the decrease $\tau_{unbound}$) and the likelihood of the RNA being able to bind Lsm2–8 rapidly (as observed by the increase in amplitude).

We also analyzed the distributions of bound dwell times (Table S3). In all cases, we observed a long-lived component of ~200 to 300 s. To determine if the lifetime of the longest-lived events could be shortened because of photobleaching, we analyzed the photobleaching rate of an immobilized fSNAP protein labeled with DY549 under the same conditions (Fig. S3). For the photobleaching control, the labeled SNAP protein had a characteristic lifetime of ~2682 s, at least 10-fold longer than the lifetime of the Lsm2–8^{fSNAP} proteins when associated with U6. Therefore, we assume that photobleaching had a minimal impact on the observed lifetime of the longest-lived Lsm2–8^{fSNAP} binding events.

In contrast with the analysis of unbound dwell times, the characteristic τ_{bound} parameters for the U6 1–113-OH and U6 1–112-P RNAs were quite similar to one another. In both cases, the fits were best described with two exponential terms. The amplitudes of the short- and long-lived parameters were only modestly higher (0.20 *versus* 0.31) for the long-lived parameter for U6 1–112-P. This suggests that while 3′-end modification can increase Lsm2–8 occupancy on U6, the origin of this effect stems mostly from facilitating protein association rather than just stabilization of the bound state. Interestingly, the U6 1–112-OH RNA showed a dramatic decrease in bound state lifetimes with an amplitude of the long-lived parameter of only 0.01. The vast majority of binding events observed on that RNA were short lived with a characteristic τ_{bound} of ~12 s. This RNA also showed an increase in the characteristic times for $\tau_{unbound}$ (Table S2). Lack of Lsm2–8^{fSNAP} occupancy on the U6 1–112-OH RNA originates both from difficulties in recruiting the protein complex to the RNA and stabilization of the complex once bound.

Kinetic mechanism for Lsm2–8 recruitment to modified U6 RNAs

We further analyzed the single-molecule data for Lsm2–8-^{fSNAP} binding the U6 1–112-P RNA to determine a kinetic mechanism that fit the observed dwell times. We carried out this analysis only with the U6 1–112-P RNA since it likely represents the predominant species for Lsm2–8 recruitment during U6 snRNP recycling from spliceosomes. We collected single-molecule data at various concentrations of Lsm2–8^{fSNAP} (1, 3, and 10 nM) and then globally fit the data to various kinetic schemes containing two, three, or four states using QuB software (29). We determined the best fit to be to a

four-state model (Fig. S4). Lsm2–8 binding predominantly occurs in a single step with a k_{on} of 1.94×10^7 M^{-1} s^{-1} and k_{off} = 0.045 s^{-1}. However, either or both U6 and Lsm2–8 can rarely transition to a state from which the U6/Lsm2–8 complex cannot form directly. In addition, the U6/Lsm2–8 complex itself can rarely transition to a state from which Lsm2–8 cannot directly dissociate. It is possible that these rare states represent alternate conformations of the protein or RNA that preclude Lsm2–8 association or dissociation.

Prp24 facilitates Lsm2–8 recruitment on unmodified RNAs

Since binding of Prp24 to displace Lhp1 is a prerequisite for subsequent Lsm2–8 binding on unmodified U6 RNAs (12, 30, 31), we next investigated the influence of Prp24 on Lsm2–8-^{fSNAP} interactions with U6 (Fig. 3A). In these experiments, we added a solution containing both Prp24 and Lsm2–8^{fSNAP} to slides containing immobilized U6 molecules and monitored Lsm2–8^{fSNAP} binding to the RNAs. One caveat of these experiments is that we did not monitor Prp24 binding directly; however, it was added at a concentration of 100 nM, ~5-fold higher than the reported K_D at 21 nM (18).

When we analyzed individual fluorescence trajectories, we noticed that while the majority RNAs bound only one Lsm2–8^{fSNAP} complex at a time, a small subset showed step-wise changes in fluorescence consistent with the presence of at least two Lsm2–8^{fSNAP} complexes associated with the same RNA (Fig. 3B). To simplify our analysis, we first analyzed only the subset of RNAs that bound a single Lsm2–8^{fSNAP} complex at a time. A rastergram analysis of these RNAs revealed that addition of Prp24 increases the association of Lsm2–8^{fSNAP} with U6 regardless of the 3′-end modification (Fig. 3, C–E). Lsm2–8 exhibited similar binding patterns on both U6 1–113-OH and 1–112-P, whereas binding on 1–112-OH appeared weaker but still stronger than was observed in the absence of Prp24 (Fig. 2F).

When we quantified the extent of Lsm2–8^{fSNAP} occupancy on RNAs in the presence of Prp24, we saw little change in occupancy with U6 1–112-P (Fig. 3F). However, Prp24 increased the occupancy fractions for both the 1–113-OH and 1–112-OH RNAs to levels more closely resembling that observed for 1–112-P. When comparing the distributions of dwell times (Fig. 3, G and H), it is apparent that Prp24 can perturb both unbound and bound times for the 1–113-OH and 1–112-OH RNAs. In contrast, the distributions of times observed on the 1–112-P RNA were similar regardless of the presence or absence of Prp24.

When comparing the characteristic $\tau_{unbound}$ and τ_{bound} times for the unmodified (1–113-OH) and modified (1–112-P) RNA mimics in the presence of Prp24, the effect of Prp24 is most pronounced in a change in the long-lived τ_{bound} parameter for U6 1–113-OH (Table S3). The amplitude of this parameter doubles from ~0.20 to 0.55 because of the presence of Prp24, and the associated characteristic lifetime increases from ~183 to 290 s. Similarly, parameters for the unbound state for U6 1–113-OH show a reduction in the times between binding events (Table S2). Overall, the parameters for the
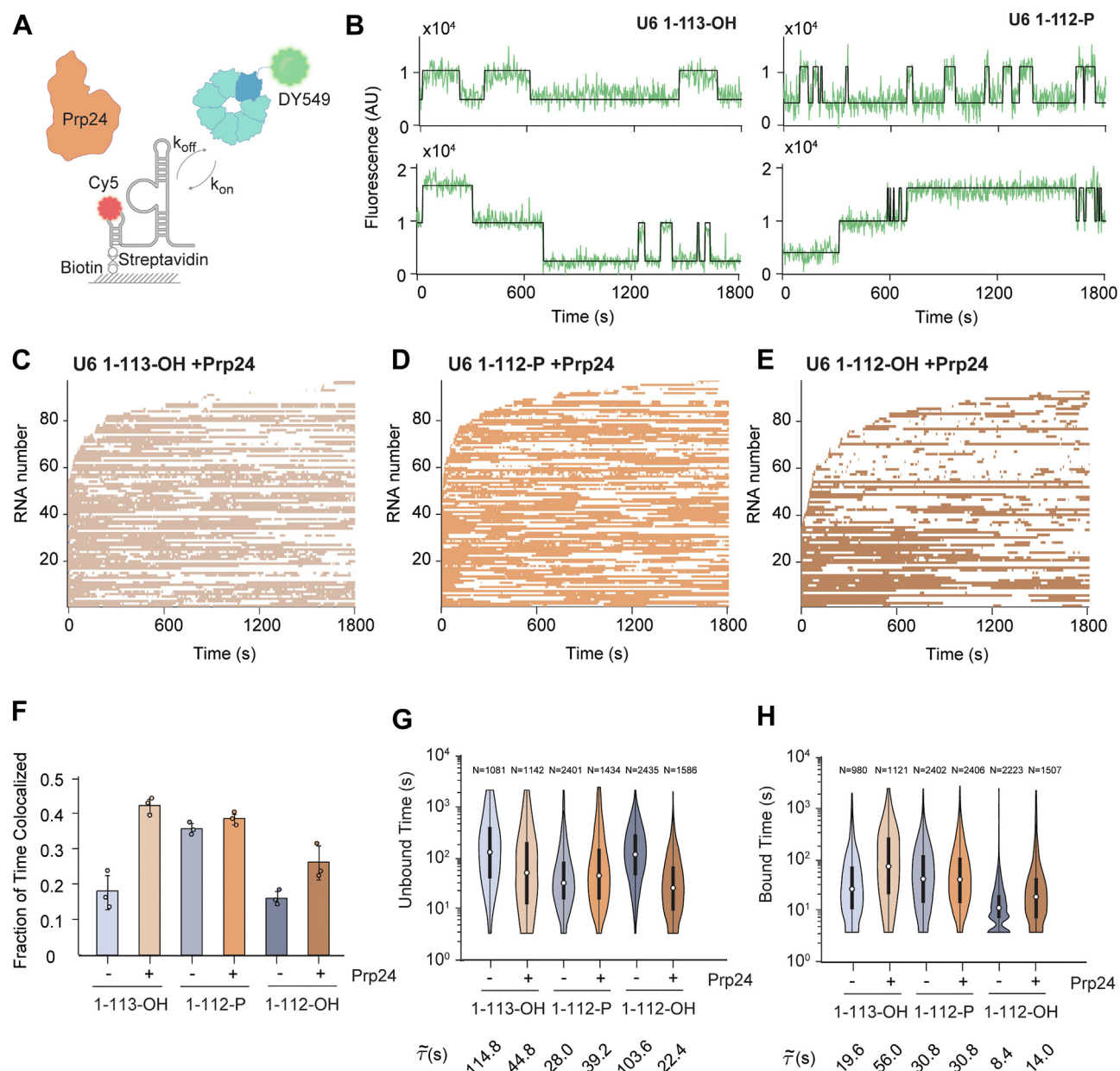
Figure 3. Prp24 enhances Lsm2–8^{fSNAP} binding kinetics to unmodified RNAs. *A*, schematic representation of the CoSMoS assay for monitoring Lsm2–8$^{fSNAP \ DY549}$ binding in the presence of unlabeled Prp24. *B*, representative fluorescence trajectories showing single Lsm2–8$^{fSNAP \ DY549}$ (*top traces*) or double (*bottom traces*) to RNAs in the presence of Prp24. *C–E*, rastergrams illustrating the binding of Lsm2–8 to 90 different molecules of U6 1–113-OH (*C*), U6 1–112-P (*D*), and U6 1–112-OH (*E*), sorted based on the time of first Lsm2–8$^{fSNAP \ DY549}$ binding event detection. *F*, average fraction of time U6 molecules spend colocalized with Lsm2–8$^{fSNAP \ DY549}$ in the presence of Prp24. Bar heights represent the average from three technical replicates (*dots*), and bars represent ±SD. Data for the absence of Prp24 are replicated from Figure 2H to facilitate comparison. *G* and *H*, violin plots representing the dwell time distributions for unbound (*G*) and bound (*H*) states across three different RNAs. Each violin plot is overlaid with a box plot that indicates the median (*horizontal line*), interquartile range (IQR, box), and *whiskers* extending to 1.5 times the IQR. The numbers above each violin plot denote the number of lifetimes included in the respective distribution. The median dwell times are shown below each plot. Data for the absence of Prp24 are replicated from Figure 2, *I* and *J* to facilitate comparison. For data shown in *C–H*, only RNAs containing single binding events were analyzed. CoSMoS, colocalization single-molecule spectroscopy.

Lsm2–8^{fSNAP} binding to U6 1–113-OH in the presence of Prp24 more closely resemble those for binding to U6 1–112-P than when Prp24 is absent. The U6 1–112-OH RNA is also impacted by Prp24 with increased binding because of changes in both the amplitude of the long-lived bound state (∼10-fold increase) and decreased times between binding events. In summary, Prp24 can facilitate the recruitment of Lsm2–8 to unmodified RNAs, and the longest-lived complexes between

Lsm2–8 and U6 have similar lifetimes independent of the 3′-end modification when Prp24 is present (300–400 s).

The SNFFL box is required for Prp24 to promote Lsm2–8 binding to unmodified U6 RNAs

Previous studies have shown that the SNFFL box of Prp24 interacts with Lsm2–8 (Fig. 1B). We wondered if a U6–Prp24

complex could present multiple surfaces for Lsm2–8 interaction, one of which could be dependent on the SNFFL box (19). These multiple surfaces could result in the double binding events we observed on some RNAs in the presence of Prp24 (Fig. 3*B*). To test this, we purified a Prp24 mutant lacking the SNFFL box. In comparison with WT Prp24, Prp24$^{\Delta SNFFL}$ is less efficient at forming U6–Prp24 and U6–Prp24–Lsm2–8 complexes by EMSA (Fig. 4*A*).

We then determined if use of the Prp24$^{\Delta SNFFL}$ protein would change the frequencies at which we observed double binding events of Lsm2–8^{fSNAP} on U6 RNAs in the CoSMoS assay. We tested this with the U6 1–112-P RNA since double binding events were much more rarely observed with the other RNAs (Fig. 4*B*). With U6 1–112-P, RNAs spent nearly 15% of the total experimental time associated with more than one Lsm2–8^{fSNAP} protein in the presence of Prp24 compared with ∼5% in its absence. However, when Prp24$^{\Delta SNFFL}$ was used, this increase was not observed and observations of double occupancy in the presence of Prp24$^{\Delta SNFFL}$ were similar in

frequency as to when Prp24 was omitted altogether. These data suggest that more than one Lsm2–8 complex can be recruited to a single U6 RNA at one time *via* the SNFFL domain of Prp24.

Finally, we tested if the SNFFL domain would have any impact on the ability of Prp24 to recruit Lsm2–8^{fSNAP} to unmodified U6 RNAs 1–112-OH or 1–113-OH. Much to our surprise, the addition of Prp24$^{\Delta SNFFL}$ prevented Lsm2–8^{fSNAP} binding to the U6 1–112-OH and 1–113-OH RNAs but not to U6 1–112-P (Fig. 4*D*). This was unexpected since we observed strong binding of Lsm2–8^{fSNAP} to both RNAs in the presence of full-length Prp24. In fact, Lsm2–8^{fSNAP} binding to the 1–113-OH RNA in the presence of Prp24$^{\Delta SNFFL}$ was less than if Prp24 was omitted entirely. This suggests Prp24$^{\Delta SNFFL}$ can inhibit Lsm2–8 from binding *cis*-diol-terminated U6 RNAs.

A possible mechanism to account for this could be sequestration of the 3′ end of U6 by Prp24$^{\Delta SNFFL}$ when it contains a *cis*-diol but not a terminal phosphate. This sequestration then prevents Lsm2–8 binding. To test this
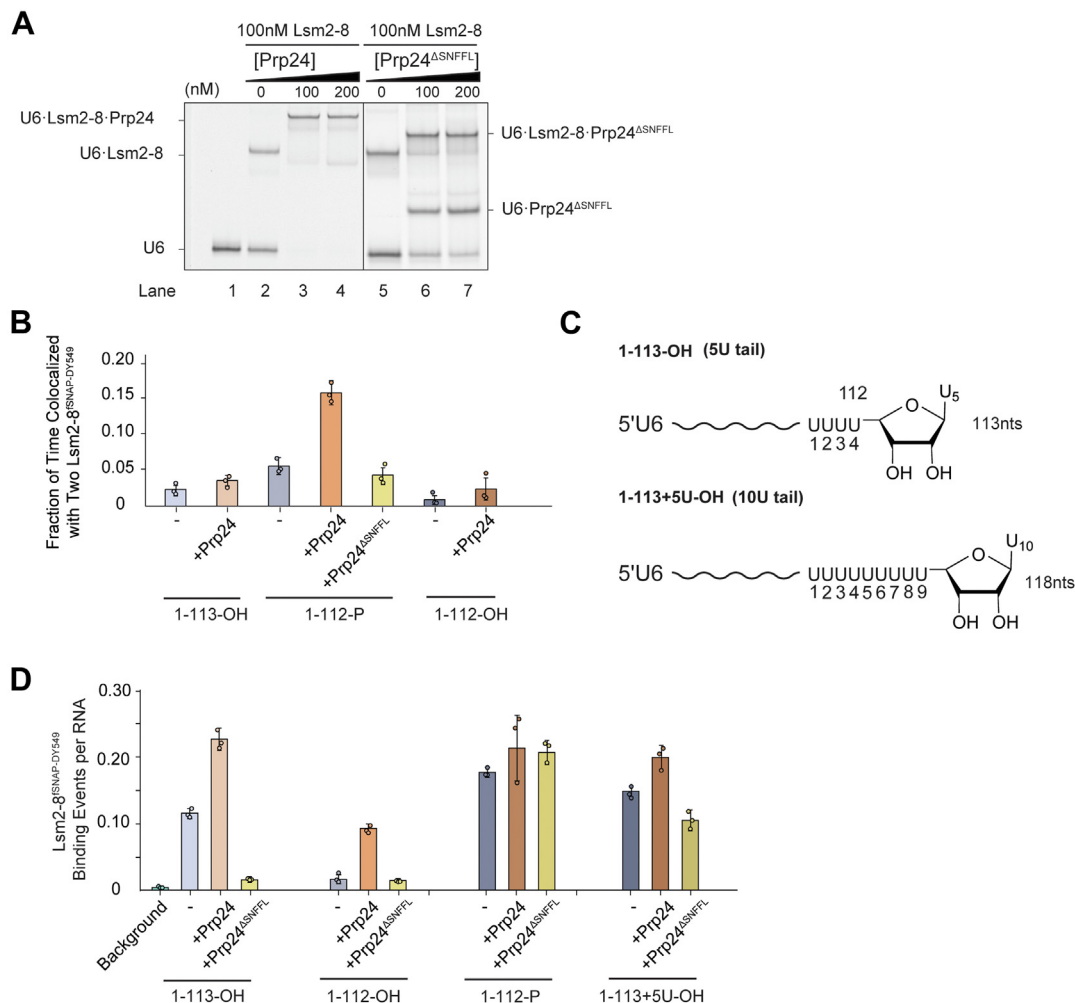

Figure 4. The Prp24 SNFFL box is required for recruitment of Lsm2–8^{fSNAP} to unmodified U6 RNAs. *A*, EMSA analysis of Lsm2–8–Prp24–U6 complex formation using EMSA using Prp24 containing (lanes 3 and 4) or lacking (lanes 6 and 7) the SNFFL-box domain. The concentrations of Cy5-labeled U6 oligos and Lsm2–8 were maintained at a constant 2 nM and 100 nM, respectively. *B*, average fractions of time U6 RNAs are simultaneously bound by two Lsm2–8$^{fSNAP\ DY549}$ complexes. *C*, chemical structures of the U6 1–113-OH and 1–113 + 5U-OH RNAs, highlighting the RNA 3′ ends. *D*, relative numbers of Lsm2–8$^{fSNAP\ DY549}$ binding events per U6 RNA in the presence or absence of Prp24 and Prp24$^{\Delta SNFFL}$. In *B* and *D*, bar heights represent the average from three technical replicates (*dots*), and bars represent ±SD.

hypothesis, we constructed a *cis*-diol-terminated U6 RNA with an extended 3′ end (1–113 + 5U-OH) in an attempt to attenuate the hindrance by Prp24$^{\Delta SNFFL}$ (Fig. 4C). In agreement of our hypothesis, Lsm2–8 exhibited increased binding to this longer RNA suggesting that the RNA 3′ end is now accessible (Fig. 4D). Together, our findings indicate that not only can the SNFFL box function to recruit Lsm2–8 complexes to U6 but that it is also essential for permitting Lsm2–8 recruitment to unmodified RNAs bound by Prp24.

Discussion

For efficient pre-mRNA splicing, it is important that the U6 snRNA be rapidly recycled from spliceosomes into U6 snRNPs and that the snRNA be protected from degradation (5). Previous work established a pathway by which yeast U6 snRNP can be post-transcriptionally modified on its 3′ end and assembled *in vitro* (12). In this work, we show that 3′-end modification can alter the kinetic pathways of assembly. In agreement with ensemble data, we observed a much higher degree of binding of Lsm2–8 to single U6 RNAs that mimic the processed form (U6 1–112-P). The origin of this effect likely stems from the presence of the 3′ phosphate since a shorter RNA containing a diol (U6 1–112-OH) did not stimulate binding. Dwell time analysis indicates that the increase in binding is primarily because of the RNA modification facilitating Lsm2–8 association (shorter $\tau_{unbound}$ times) rather than just stabilization of the bound state. Prp24 can moderate the impact of 3′-end modification and allow for efficient recruitment of Lsm2–8 on both processed and unprocessed RNAs. The increase in Lsm2–8 occupancy because of Prp24 can arise from both facilitating association and slowing dissociation. Interestingly, we observe that some U6 RNAs can recruit two Lsm2–8 complexes simultaneously in the presence of Prp24 and that this is dependent on the Prp24 SNFFL-box domain. This domain is also required for Prp24 to recruit Lsm2–8 to unmodified RNAs since Prp24 molecules lacking the SNFFL box can actually inhibit Lsm2–8 binding. We do not know if Lsm2–8 complexes bound to SNFFL box but not yet associated with U6 can be directly transferred to the RNA 3′ end.

Together, our results indicate that post-transcriptional modification of U6 facilitates snRNP assembly by opening an alternate route of Lsm2–8 acquisition (Fig. 5). In the absence of modification, Lsm2–8 binding would be highly dependent on the presence of Prp24. For this pathway, we assume that Prp24 binds U6 first and subsequently recruits Lsm2–8. However, we did not simultaneously monitor Prp24 and Lsm2–8 recruitment, and it is possible that a preformed Prp24–Lsm2–8 complex may also be involved in snRNP assembly. U6 snRNA modification opens a second pathway in which efficient and stable Lsm2–8 binding occurs in the absence of Prp24. Both Prp24-dependent and -independent pathways may be functionally relevant for splicing *in vivo*. In the case of *de novo* U6 snRNP assembly from an unprocessed transcript, Prp24 dependence likely assures that the Lsm2–8 complex assembles specifically on U6 snRNAs rather than on any transcript harboring a 3′ polyU sequence. Notably, such
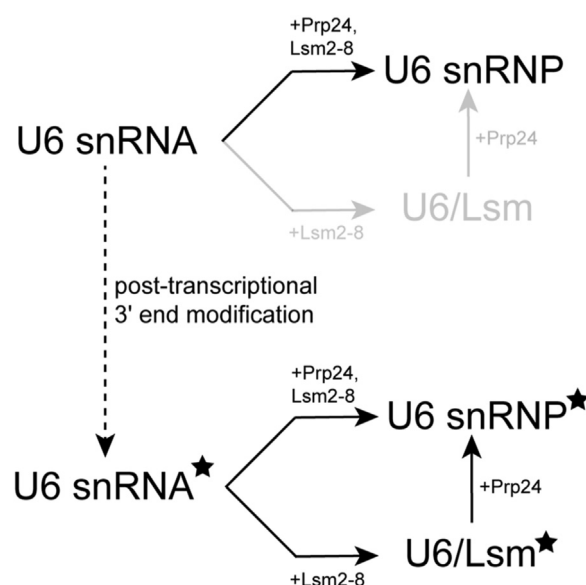

Figure 5. U6 RNA 3′-end modification permits an alternate assembly route for the U6 snRNP. The unmodified U6 snRNA depends on interaction with Prp24 for Lsm2–8 recruitment, whereas U6 modification (3′ phosphate modification is denoted by the *black star*) permits Prp24-dependent and -independent Lsm2–8 recruitment. snRNP, small nuclear ribonucleoprotein.

transcripts are abundant in the cell since all RNAPIII transcripts terminate with polyU including tRNAs and 5S rRNAs. In terms of recycling U6 snRNA from spliceosomes at the conclusion of splicing, direct recruitment of Lsm2–8 to modified U6 snRNA may facilitate rapid reassembly of the snRNP while the spliceosome is being dismantled and before Prp24 has the opportunity to bind. In addition to making reassembly more efficient, this may also be important for protecting U6 snRNA from nucleases by ensuring fast sequestration of the RNA 3′ end (30, 32).

Our work also unexpectedly uncovered a novel function for the Prp24 SNFFL-box domain in facilitating Lsm2–8 recruitment to 3′-end unmodified U6 RNAs. In the absence of this domain, Lsm2–8 complexes failed to bind stably to the U6 RNA unless a 3′ phosphate was present. We speculate that this is due to sequestration of the U6 3′ end by Prp24$^{\Delta SNFFL}$ but do neither yet have a structural explanation for this nor have we located a potential binding site for diol-terminated 3′ ends on the protein. Nonetheless, our observations are consistent with *in vivo* phenotypes reported in *S. cerevisiae* for Prp24$^{\Delta SNFFL}$ mutants (22). Deletion of the C-terminal 10 amino acids of Prp24 (including the SNFFL box; Prp24$^{\Delta 10}$) results in a cold-sensitive growth phenotype, reduced levels of U4/U6 di-snRNAs, reduced yeast two-hybrid interactions between Prp24 and Lsm proteins, and changes in the kinetic parameters (K_m for U4 and/or U6 snRNAs) for U4/U6 di-snRNA formation *in vitro* using snRNAs isolated from yeast. Since Prp24 and Lsm proteins influence multiple steps in U6 and U4/U6 di-snRNP assembly (5, 12) and the relative abundances of 3′-end-modified and -unmodified U6 snRNAs were not analyzed in experiments involving Prp24$^{\Delta 10}$, it is difficult to define the precise origin(s) of the *in vivo* phenotypes. However, our work indicates that inhibition of Lsm2–8 binding to

3′-end-unmodified U6 snRNAs by Prp24$^{\Delta 10}$ could contribute to lower levels of U4/U6 di-snRNA formation because of some U6 snRNAs lacking the Lsm2–8 as well as weakened two-hybrid interactions between Prp24 and Lsm proteins if such interactions were mediated in part by endogenous U6 snRNAs.

Finally, it is interesting to note that the related Lsm1–7 complex from *S. pombe* can tightly bind polyU tracts on RNAs lacking modifications and in the absence of protein cofactors (Pat1) (16). Even though these complexes share six of seven subunits, their functions are very different: Lsm1–7 association marks mRNAs for degradation, but Lsm2–8 association is needed for U6 snRNP assembly. While we have not carried out a detailed kinetic analysis of Lsm1–7/RNA interactions, this observation suggests that it may be equally important for the cell to prevent Lsm2–8 (or Lsm1–7) complex assembly on the wrong RNAs as it is to facilitate assembly on U6. Dependence of the Lsm2–8 binding kinetics on both 3′-end modification and the presence of Prp24 could provide two independent layers of proofreading *via* kinetic selection to facilitate preferential recruitment to the U6 snRNA.

Experimental procedures

Cloning and mutagenesis

The fSNAP variant of yeast Lsm2–8 was created by first removing the Lsm8 sequence from the pQLink-Lsm2–8 plasmid (a gift from Yigong Shi; (33)) by restriction enzyme digestion with SacI and NheI. A double-stranded DNA fragment coding for a C-terminal fusion of the fSNAP tag to Lsm8 was synthesized (ThermoFisher, GeneArt Synthetic Gene) and ligated into those same restriction sites. To ensure the accuracy of recombinant plasmids and deletion mutants, DNA sequencing was performed for the whole plasmid.

Protein expression and purification

Prp24 was purified as described previously (12). Briefly, proteins were expressed in *E. coli* BL21 (DE3) pLysS cells in LB medium after induction by addition of IPTG to a final concentration of 1 mM for 20 h at 16 °C. Proteins were then purified by nickel–nitrilotriacetic acid (Ni–NTA) agarose chromatography followed by removal of the His-tag with tobacco etch virus (TEV) protease and cation-exchange chromatography as described except that a HiTrap SP FF (GE Healthcare) column was used.

Lhp1 was also purified as previously described (12). Lhp1 protein was expressed in *E. coli* BL21 (DE3) pLysS cells in LB medium after induction by addition of IPTG (1 mM) for 20 h at 16 °C. Lhp1 was then purified by Ni–NTA agarose chromatography followed by removal of the His-tag with TEV protease and cation-exchange chromatography as described except that a HiTrap S (GE Healthcare) column was used.

Lsm2–8 and Lsm2–8^{fSNAP} were purified using a previously described protocol for Lsm2–8 (12). Lsm proteins were expressed in *E. coli* BL21 (DE3) pLysS cells in LB medium after induction by addition of IPTG (1 mM) for 20 h at 37 °C. Lsm protein complexes were then purified by Ni–NTA agarose and glutathione-*S*-transferase-affinity chromatography followed by removal of the His-tag with TEV protease and anion-exchange chromatography as described.

For all proteins, the final concentrations were determined by using calculated extinction coefficients (ProtParam (34)) and measuring the absorbance at 280 nm.

Fluorophore labeling of Lsm2–8^{fSNAP}

SNAP-Surface 549 dye (S9112S; New England BioLabs) in dimethyl sulfoxide was added to a protein solution (500 nM; 500 µl) at a dye ratio 1.25 dye:1 Lsm2–8^{fSNAP} complex in size-exclusion chromatography buffer (25 mM Hepes–KOH [pH 7.9], 50 mM KCl, 1 mM Tris(2-carboxyethyl)phosphine hydrochloride [TCEP], and 10% v/v glycerol). The reaction tube was then incubated in darkness at room temperature for 1 h. Excess dye was removed by gel filtration chromatography using G-25 Sephadex (Sigma) equilibrated in size-exclusion chromatography buffer. For the labeled proteins, the labeling efficiency was determined by the concentrations of both dye and protein in the final product measured by the absorbance at 280 and 550 nm.

RNA preparation

RNA oligos were purchased from Integrated DNA Technologies (IDT). Sequences and modifications are described in Table S4.

Full-length U6 RNAs (112 or 113 nt in length) containing a 3′ *cis*-diol were prepared as previously described by splinted RNA ligation using T4 RNA ligase 2 (12). For CoSMoS assays, the 5′ piece of the U6 RNA encoding nt 1 to 12 also contained a 5′ biotin and internal Cy5 fluorophore. The RNA molecules were then purified using PAGE with 8 M urea in a 20% (w/v) acrylamide: *bis*-acrylamide gel, followed by extraction into buffer (1 mM EDTA, 10% v/v acid phenol, 300 mM sodium acetate, pH 5.0). RNAs were then isolated by ethanol precipitation and dissolved in nuclease-free water (Ambion). RNA products were then quantified by UV–visible spectroscopy.

3′ Phosphate-modified U6 molecules for CoSMoS assays were prepared using a double-splinted ligation reaction with T4 RNA ligase 2 and 3 RNA fragments corresponding to U6 nt 1 to 12 (synthetic and containing a 5′ biotin and internal Cy5, purchased from IDT), nt 13 to 94 (*in vitro* transcribed), and nt 95 to 112 (synthetic, purchased from IDT), and two DNA splints. In the ligation reaction, the three pieces were used at a ratio of 1 (nt 1–12):2 (nt 13–94):2 (nt 95–112):1.5 (DNA splints). Ligation products were gel purified, extracted, isolated, and quantified as aforementioned.

Electrophoretic mobility shift assays

RNAs were first diluted into RNA dilution buffer (100 mM KCl, 20% w/v sucrose, 20 mM Hepes [pH 7], 1 mM EDTA, 1 mM TCEP, 0.01% v/v Triton X-100, 0.2 mg/ml yeast tRNA, and 0.2 mg/ml sodium heparin), and proteins were diluted into protein dilution buffer (100 mM KCl, 20% w/v sucrose, 20 mM Hepes [pH 7], 1 mM EDTA, 1 mM TCEP, 0.01% v/v Triton

Kinetic analysis of Lsm2–8 binding to the U6 snRNA

X-100, and 0.2 mg/ml bovine serum albumin [BSA]) as described (12). Binding was initiated by mixing 5 μl of RNA with 5 μl of protein, and samples were then incubated for 30 min at room temperature. Mixtures were then loaded onto a prerun native 6% (w/v) PAGE gel in 1x Tris–borate–EDTA and separated for 2 h at 4°C at 5 W. Fluorescence analysis was performed with a Typhoon FLA 9000 scanner (Cytiva) with excitation at 635 nm, and results were analyzed using ImageJ software (version 1.52v) (35).

Fluorescence polarization assays

Fluorescence polarization experiments were performed according to previous methods (12). In summary, a 2× RNA solution (2 nM, 100 μl) in buffer (100 mM NaCl, 20 mM Bis–Tris [pH 7.0], 10 mM HCl, 1 mM TCEP, 5% v/v glycerol, and 0.02 mg yeast tRNA) was combined with 100 μl of protein at various concentrations (0.01–1000 nM) in buffer that also contained 0.2 mg/ml BSA. The mixture was then added to black, nontransparent, flat-bottomed 96-well microplates. Fluorescence polarization data were then collected in triplicate using a Tecan Infinite M1000Pro plate reader with an excitation wavelength of 470 nm and an emission wavelength of 519 nm. The data were normalized to the values obtained with 0 nM protein (the smallest value) and to the highest value and then averaged across three technical replicates. Binding curves were fit using nonlinear regression in GraphPad Prism 4 with a four-parameter logistic equation where FP_{min} and FP_{max} represent the normalized minimum and maximum % bound, K_D stands for the binding dissociation constant, and H represents the Hill coefficient (Equation 1).

$$\% \; Bound = FP_{min} + \frac{(FP_{max} - FP_{min})}{1 + 10^{(log \; K_D - log[Protein]) \times H}} \qquad (1)$$

Microscope slide preparation

Microscope slides (100490-396; VWR) and cover slips (12-553-455; Fisher) were cleaned as described (14) using 2% v/v micro-90, 100% ethanol, and 1 M KOH with rinsing in between each solution with MilliQ water. After the final rinse, the slides were dried with high purity nitrogen (NI UHP300; Airgas) and aminosilanized with VECTABOND (NC9280699; Fisher Scientific). Slides were then passivated by incubation overnight with a mixture of mPEG-SVA (MPEG-SVA-5K; Laysan Bio) and mPEG-biotin-SVA (BIO-PEG-SVA-5K; Laysan Bio) in a 1:100 w/w ratio in 100 mM NaHCO$_3$ buffer.

Single-molecule microscopy

Single-molecule data were obtained using a custom-built objective-type total internal fluorescence microscope system (36). The microscope setup includes a 60 × 1.49 Numerical Aperture PlanApo objective (Olympus) and 532 nm (CrystalLaser), 633 nm (Power Technology, Inc), and 785 nm (Power Technology, Inc) lasers for total internal reflection fluorescence excitation. Typically, fluorophores were imaged with the 532 nm laser set at a power of 800 to 900 μW and the 633 nm laser set to 600 to 700 μW. The 785 nm laser was used for autofocus and was set at 2.5 mW. The emission light was split into <635 nm and >635 nm channels to produce two images that were then imaged on separate sCMOS detectors (Hamamatsu ORCA-Flash4.0 V3) with 2 × 2-pixel binning. The microscope was controlled using Micro-Manager 2.03 (37).

Prior to data collection, streptavidin-labeled fluorescent beads (T10711; Invitrogen) in PBS were introduced into a channel on the passivated slide to serve as fiduciary markers for alignment. The lane was then washed 0.2 mg/ml streptavidin (SA10-10; Agilent; 50 μl) for 2 min, followed by washing with PBS to remove any unbound beads and streptavidin. Biotin_Cy5_U6 snRNA was diluted to 10 pM in imaging buffer (20 mM Hepes [pH 7.0], 100 mM KCl, 1 mM EDTA, 20% w/v sucrose, 10 mg/ml BSA, 5 mM protocatechuic acid, 1 U/ml protocatechuate dioxygenase, 1 mM trolox, and 2% v/v dimethyl sulfoxide) and incubated in the lane for 1 min, followed by another wash with imaging buffer. A 100 μl solution containing varying concentrations of Prp24, Lsm2–8, and Lhp1 in imaging buffer was added, and alternating laser excitation imaging was performed based on the following excitation scheme: The 532 and 633 nm lasers were alternatively turned on (1 s/fr) to capture images with an ∼0.5 s lag between each image for ∼1800 s.

Microscopy data analysis

The single-molecule data were analyzed using custom MATLAB software (smtoolbox, available at https://github.com/David-Scott-White/cosmos-toolbox). The analysis procedure involved several steps: (1) a mapping file that correlated the positions in the <635 nm channel to the >635 nm channel was created utilizing the fluorescence beads as fiducial markers, (2) lateral drift was corrected by computing a non-reflective similarity transformation between temporally separated images, (3) potential areas of interest (AOIs) containing single RNA molecules were identified in the >635 nm channel by averaging the first five frames and applying a generalized likelihood ratio test to the resulting image. These detected AOIs were then fit to a two-dimensional Gaussian within a 5 × 5 pixel area, (4) the accepted AOIs were subsequently mapped to the <635 nm channel using the mapping transformations, and (5) the fluorescence trajectories in the <635 nm channel (Lsm2–8 proteins) for each AOI were processed using the divisive segmentation and clustering algorithm (DISC, (38)) to interpret each trajectory as a binary signal indicating bound or unbound states over time.

The durations of bound and unbound events (dwells) were treated as single or double exponential distributions, and the underlying parameters of each distribution were estimated using maximum likelihood methods (39, 40). A comparison was made between models described by single exponential terms, the sum of multiple exponential terms, or gamma

distributions. The log-likelihood ratio test was employed for model selection. For exponential distributions, the likelihood of observing a dwell time of a specific duration is expressed as (Equation 2):

$$P(x) = \sum_{k=1}^{K} A_k \tau_k e^{-\tau_k x} \qquad (2)$$

where the variable k denotes the number of exponentials being estimated, A represents the amplitude associated with a particular exponential term, and τ signifies the estimated time constant corresponding to that term.

To control photobleaching, a biotinylated SNAP protein was derivatized with SNAP-Surface 549, immobilized on a slide, and its fluorescence lifetime measured under identical conditions as previously described (14).

Kinetic analysis using QuB

To analyze the binding kinetics of Lsm2–8 to U6 RNA, we performed Hidden Markov Model fitting in QuB using single-molecule trajectories collected at 1, 3, and 10 nM Lsm2–8 concentrations (29). Transition rates were optimized using Maximum Idealized Point likelihood rate estimation, which globally optimizes state transition probabilities across all observed molecules (41). A range of user-defined models with different numbers of states (two to four) and transition pathways were tested to identify the best kinetic description of Lsm2–8 binding. Model selection was based on the Bayesian Information Criterion (42):

$$BIC = k \log(N) - 2 \log(L(\theta))$$

where k is the number of free parameters, N is the number of frames analyzed, and $L(\theta)$ is the log likelihood of the dataset given the model. After evaluating multiple models, we found that a model with six free parameters provided the best fit, as indicated by the lowest Bayesian Information Criterion value. The final kinetic scheme, presented in Fig. S4, represents the key states and transition pathways identified from the optimal model.

Data availability

Raw microscopy data and gel images have been uploaded to Zenodo (DOI: 10.5281/zenodo.14790398).

Supporting information—This article contains supporting information, including supporting figures illustrating protein and RNA purification results, spot accumulation curves, photobleaching controls, and kinetic models. It also includes data tables showing fitted binding affinity parameters, bound and unbound time distributions, and oligonucleotide sequences. These materials can be accessed online alongside the full text of the article.

Author contributions—Y. L. and A. A. H. conceptualization; Y. L. methodology; Y. L. validation; Y. L. formal analysis; Y. L. and Y. N. investigation; S. E. B. resources; Y. L. data curation; Y. L. and A. A. H. writing–original draft; Y. L., S. E. B., and A. A. H. writing–review & editing; Y. L. and A. A. H. visualization; S. E. B. and A. A. H. supervision; S. E. B. and A. A. H. project administration; S. E. B. and A. A. H. funding acquisition.

Conflict of interest—A. A. H. is a member of the scientific advisory board and is carrying out sponsored research for Remix Therapeutics. All the other authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: AOI, area of interest; BSA, bovine serum albumin; CoSMoS, colocalization single-molecule spectroscopy; IDT, Integrated DNA Technologies; Ni–NTA, nickel–nitrilotriacetic acid; RNAP, RNA polymerase; snRNA, small nuclear RNA; snRNP, small nuclear ribonucleoprotein; TCEP, Tris(2-carboxyethyl)phosphine hydrochloride; TEV, tobacco etch virus.

References

1. Wahl, M. C., Will, C. L., and Luhrmann, R. (2009) The spliceosome: design principles of a dynamic RNP machine. *Cell* **136**, 701–718
2. Nilsen, T. W. (2003) The spliceosome: the most complex macromolecular machine in the cell? *Bioessays* **25**, 1147–1149
3. Staley, J. P., and Guthrie, C. (1998) Mechanical devices of the spliceosome: motors, clocks, springs, and things. *Cell* **92**, 315–326
4. Brow, D. A., and Guthrie, C. (1988) Spliceosomal Rna U6 Is Remarkably Conserved from Yeast to Mammals. *Nature* **334**, 213–218
5. Didychuk, A. L., Butcher, S. E., and Brow, D. A. (2018) The life of U6 small nuclear RNA, from cradle to grave. *RNA* **24**, 437–460
6. Makarov, E. M., Makarova, O. V., Urlaub, H., Gentzel, M., Will, C. L., Wilm, M., *et al.* (2002) Small nuclear ribonucleoprotein remodeling during catalytic activation of the spliceosome. *Science* **298**, 2205–2208
7. De Bortoli, F., Espinosa, S., and Zhao, R. (2021) DEAH-box RNA helicases in pre-mRNA splicing. *Trends Biochem. Sci.* **46**, 225–238
8. Toroney, R., Nielsen, K. H., and Staley, J. P. (2019) Termination of pre-mRNA splicing requires that the ATPase and RNA unwindase Prp43p acts on the catalytic snRNA U6. *Genes Dev.* **33**, 1555–1574
9. Vorlander, M. K., Rothe, P., Kleifeld, J., Cormack, E. D., Veleti, L., Riabov-Bassat, D., *et al.* (2024) Mechanism for the initiation of spliceosome disassembly. *Nature* **632**, 443–450
10. Booth, B. L., Jr., and Pugh, B. F. (1997) Identification and characterization of a nuclease specific for the 3' end of the U6 small nuclear RNA. *J. Biol. Chem.* **272**, 984–991
11. Pannone, B. K., Xue, D., and Wolin, S. L. (1998) A role for the yeast La protein in U6 snRNP assembly: evidence that the La protein is a molecular chaperone for RNA polymerase III transcripts. *EMBO J.* **17**, 7442–7453
12. Didychuk, A. L., Montemayor, E. J., Carrocci, T. J., DeLaitsch, A. T., Lucarelli, S. E., Westler, W. M., *et al.* (2017) Usb1 controls U6 snRNP assembly through evolutionarily divergent cyclic phosphodiesterase activities. *Nat. Commun.* **8**, 497
13. Shchepachev, V., Wischnewski, H., Missiaglia, E., Soneson, C., and Azzalin, C. M. (2012) Mpn1, mutated in poikiloderma with neutropenia

protein 1, is a conserved 3'-to-5' RNA exonuclease processing U6 small nuclear RNA. *Cell Rep.* **2**, 855–865

14. Fu, X., Kaur, H., Rodgers, M. L., Montemayor, E. J., Butcher, S. E., and Hoskins, A. A. (2022) Identification of transient intermediates during spliceosome activation by single molecule fluorescence microscopy. *Proc. Natl. Acad. Sci. U. S. A.* **119**, e2206815119

15. Tharun, S. (2009) Roles of eukaryotic Lsm proteins in the regulation of mRNA function. *Int. Rev. Cell Mol. Biol.* **272**, 149–189

16. Montemayor, E. J., Virta, J. M., Hayes, S. M., Nomura, Y., Brow, D. A., and Butcher, S. E. (2020) Molecular basis for the distinct cellular functions of the Lsm1-7 and Lsm2-8 complexes. *RNA* **26**, 1400–1413

17. Kufel, J., Bousquet-Antonelli, C., Beggs, J. D., and Tollervey, D. (2004) Nuclear pre-mRNA decapping and 5' degradation in yeast require the Lsm2-8p complex. *Mol. Cell Biol.* **24**, 9646–9657

18. Didychuk, A. L., Montemayor, E. J., Brow, D. A., and Butcher, S. E. (2016) Structural requirements for protein-catalyzed annealing of U4 and U6 RNAs during di-snRNP assembly. *Nucleic Acids Res.* **44**, 1398–1410

19. Montemayor, E. J., Didychuk, A. L., Yake, A. D., Sidhu, G. K., Brow, D. A., and Butcher, S. E. (2018) Architecture of the U6 snRNP reveals specific recognition of 3'-end processed U6 snRNA. *Nat. Commun.* **9**, 1749

20. Karaduman, R., Fabrizio, P., Hartmuth, K., Urlaub, H., and Luhrmann, R. (2006) RNA structure and RNA-protein interactions in purified yeast U6 snRNPs. *J. Mol. Biol.* **356**, 1248–1262

21. Kwan, S. S., and Brow, D. A. (2005) The N- and C-terminal RNA recognition motifs of splicing factor Prp24 have distinct functions in U6 RNA binding. *RNA* **11**, 808–820

22. Rader, S. D., and Guthrie, C. (2002) A conserved Lsm-interaction motif in Prp24 required for efficient U4/U6 di-snRNP formation. *RNA* **8**, 1378–1392

23. Pincus, D. (2015) Keeping up with the 'omics: non-equilibrium models of gene regulation. *BMC Biol.* **13**, 9

24. Larson, J. D., Rodgers, M. L., and Hoskins, A. A. (2014) Visualizing cellular machines with colocalization single molecule microscopy. *Chem. Soc. Rev.* **43**, 1189–1200

25. Walter, N. G., Huang, C. Y., Manzo, A. J., and Sobhy, M. A. (2008) Do-it-yourself guide: how to use the modern single-molecule toolkit. *Nat. Methods* **5**, 475–489

26. Hoskins, A. A., Rodgers, M. L., Friedman, L. J., Gelles, J., and Moore, M. J. (2016) Single molecule analysis reveals reversible and irreversible steps during spliceosome activation. *Elife* **5**, e14166

27. Larson, J., Kirk, M., Drier, E. A., O'Brien, W., MacKay, J. F., Friedman, L. J., *et al.* (2014) Design and construction of a multiwavelength, micromirror total internal reflectance fluorescence microscope. *Nat. Protoc.* **9**, 2317–2328

28. Friedman, L. J., Chung, J., and Gelles, J. (2006) Viewing dynamic assembly of molecular complexes by multi-wavelength single-molecule fluorescence. *Biophys. J.* **91**, 1023–1031

29. Nicolai, C. S. (2013) Solving ion channel kinetics with the qub software. *Biophys. Rev. Lett.* https://doi.org/10.1142/S1793048013300053

30. Spiller, M. P., Boon, K. L., Reijns, M. A., and Beggs, J. D. (2007) The Lsm2-8 complex determines nuclear localization of the spliceosomal U6 snRNA. *Nucleic Acids Res.* **35**, 923–929

31. Mroczek, S., and Dziembowski, A. (2013) U6 RNA biogenesis and disease association. *Wiley Inter. Rev. RNA* **4**, 581–592

32. Ryan, D. E., Stevens, S. W., and Abelson, J. (2002) The 5' and 3' domains of yeast U6 snRNA: Lsm proteins facilitate binding of Prp24 protein to the U6 telestem region. *RNA* **8**, 1011–1033

33. Zhou, L., Hang, J., Zhou, Y., Wan, R., Lu, G., Yin, P., *et al.* (2014) Crystal structures of the Lsm complex bound to the 3' end sequence of U6 small nuclear RNA. *Nature* **506**, 116–120

34. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., and Bairoch, A. (2005) *Protein Identification and Analysis Tools on the Expasy Server*, Humana Press, Totowa, NJ

35. Rasband, W. S. (1997-2015) *ImageJ*. U.S. National Institutes of Health, Bethesda, MD

36. DeHaven, A. C., Norden, I. S., and Hoskins, A. A. (2016) Lights, camera, action! Capturing the spliceosome and pre-mRNA splicing with single-molecule fluorescence microscopy. *Wiley Inter. Rev. RNA* **7**, 683–701

37. Edelstein, A. D., Tsuchida, M. A., Amodaj, N., Pinkard, H., Vale, R. D., and Stuurman, N. (2014) Advanced methods of microscope control using muManager software. *J. Biol. Methods* **1**, e10

38. White, D. S., Goldschen-Ohm, M. P., Goldsmith, R. H., and Chanda, B. (2020) Top-down machine learning approach for high-throughput single-molecule analysis. *Elife* **9**, e53357

39. Kaur, H., Jamalidinan, F., Condon, S. G. F., Senes, A., and Hoskins, A. A. (2019) Analysis of spliceosome dynamics by maximum likelihood fitting of dwell time distributions. *Methods* **153**, 13–21

40. Floyd, D. L., Harrison, S. C., and van Oijen, A. M. (2010) Analysis of kinetic intermediates in single-particle dwell-time distributions. *Biophys. J.* **99**, 360–366

41. Milescu, L. S., Yildiz, A., Selvin, P. R., and Sachs, F. (2006) Extracting dwell time sequences from processive molecular motor data. *Biophys. J.* **91**, 3135–3150

42. Schwarz, G. (1978) Estimating dimension of a model. *Ann. Stat.* **6**, 461–464

ASBMB