Alternative Spliceosome Assembly Pathways

Revealed by Single Molecule Fluorescence Microscopy

Inna Shcherbakova, Aaron A. Hoskins, Larry J. Friedman, Victor Serebrov,

Ivan R. Corrêa Jr., Ming-Qun Xu, Jeff Gelles, and Melissa J. Moore

Supplemental Text

Supplemental Figure Legends

Figure S1. Binding of spliceosomal subcomplexes by surface-tethered pre-mRNAs, Related to Figure 2.

A. The distributions of the number of binding events for the U1, U2, U5 and NTC spliceosomal subcomplexes observed on individual pre-mRNA molecules over the first 20 min of the experiments are shown in Figure 2D. All data are from CoSMoS experiments in WCE in which a single spliceosomal subcomplex was labeled with Cy3-TMP. Only binding events observable in two or more consecutive frames (duration ≥ 1.5 sec) were included in the analysis. **B.** Distribution of the time intervals ($t_{U2} - t_{U1}$) between the binding of U1 and the binding of U2 for the formation of U1•U2•*RPS30A* pre-mRNA pre-spliceosomes in WCE CoSMoS experiments. U1 was labeled with red fluorophores on two different proteins and U2 was labeled with a single green fluorophore as described (Hoskins et al., 2011).

Figure S2. Control experiments to monitor binding of fluorescent U1, U2 and U5 snRNPs from triply-labeled WCE to surface-tethered pre-mRNAs, Related to Figure 3.

A. Subcomplex binding to *RPS30A* pre-mRNA in triply-labeled WCE. Plots indicate the number of U1, U2 and U5 spots in the same field of view during the first 1700 seconds of incubation of the surface-tethered pre-mRNA in WCE containing 2 mM ATP. **B.** Same as A, but using WCE depleted of ATP. **C.** Same as A, but in a control sample with no surface-tethered pre-mRNA, recorded during the first 900 seconds of incubation. **D.** Correlation between U5 binding and the preceding U1•U2•pre-mRNA complex formation event for complexes that formed by the U1→U2→U5 (red) and the U2→U1→U5 (green) pathways. Each point on the plot indicates the time after WCE addition for U5 snRNP binding to an individual *RPS30A* pre-mRNA molecule, and the time after WCE addition for formation on the same pre-mRNA molecule of the U1/U2/pre-mRNA complex that most closely preceded U5 binding. The experiment (see Figure 3) used triply-labeled WCE in the presence of 2 mM ATP. **E.** Profile of fluorescence of CUS1 labeled by CLIP-Surface^{TM} 547 from untreated (solid line) and micrococcal nuclease treated (dashed line) WCE after 10-30% glycerol gradient sedimentation.

Extended Experimental Procedures

Preparation of the pre-mRNAs

The sequences of gene segments used to transcribed pre-mRNAs are listed below, with exon sequences highlighted in gray, and splice site and branch point sequences shown in red (http://compbio.soe.ucsc.edu/yeast_introns.html)

RPS30A:

GGCTGCACATTAAATTTGCCACTGTAATAATCTTCCATATCCCCATACAAAAACTACGCAAA
TATG**GT**ACGTAATGCCTCATTAGGATTTTAAAATATGAAAAAAATGGACGTTTCAGCGAGGA
ATGAATAACAAGAAATAGCTAATTTGAATATATTACTGGTGTTTCATATCTATATATACCGCA
GAAAAAGAAATAAAACGCAATATATCTTTAAACTGGCTCCACCCTGAGAATATGATAGGAAA
CTTTTTAAATTTGAGTTCTGTTACATCAAAAACGTAATTGCTTGACACAGATTTTTCGAGGGC
AACAGCGTGGACTTAACATAATGTCGAGTTTAATAACAACATTTTTTGACGGAGAGTTAAAA
AGGCTTTGAGAAAATAGCAGAGTGAAGTTTATTTTTAGCATCCTTGCTGATACGCCAAATAA
TTCAAAAATTCACTA**TACTAAC**ATGGTCATTAATTTTTAATATTAATATTATTTATTT**AG**GCTA
AAGTTCATGGTTCTCTAGCTCGTGCTGGTAAAGTCAAGTCTCAAAC

RPS30B:

GGATTCTGAATTAGCACTCCTTCCTCAAGATACTAGATAAACAAATTATACAAAATG**GT**ACG
TTAATATTGTGTTAAAATGCTCCACGAAACGCTTTTTTTAGGAAACTTGTCCCCATACAGGA
AGAAGGTGTATTAAAAGGGAAAACTACGATCAAAGTAATATTTTGAATAGGAAAGATTGCCC
CATGATAATATTAATAATCCATCAAATGCCTGTAGATAATGACTGAAAAATGAAAAGAAAATA
AAATATTGTATCATTATTATGTTGAGATGTAAATGTACAAAAAAGGAAGAAAGACGAGAATTT
CAAAGGACGATTATTCTCGTTAATATTCACAACATGTTGAATATAGACCTTTCCTAAAAAATT
AGCGTCTGCGTTGCAGTAACTCCATAACACAAGTCTGAATATCTCAT**TACTAAC**ATTAAATG
AAAAATTTTTTTTTTTTCAAATTCTCCTT**AG**GCTAAAGTTCACGGTTCTCTAGCTCGTGCTGG
TAAAGTTAAGTCTCAAAC

ACT1:

GGTTTCTACTCAAACCAAGAAGAAAAAGAAAAGGTCAATCTTTGTTAAAGAATAGGATCTTC
TACTACATCAGCTTTTAGATTTTTCACGCTTACTGCTTTTTTCTTCCCAAGATCGAAAATTTA

CTGAATTAACAATGGATTCTG**GT**ATGTTCTAGCGCTTGCACCATCCCATTTAACTGTAAGAA
GAATTGCACGGTCCCAATTGCTCGAGAGATTTCTCTTTTACCTTTTTTTACTATTTTTCACTC
TCCCATAACCTCCTATATTGACTGATCTGTAATAACCACGATATTATTGGAATAAATAGGGG
CTTGAAATTTGGAAAAAAAAAAAAACTGAAATATTTTCGTGATAAGTGATAGTGATATTCTTC
TTTTATTTGCTACTGTTACTAAGTCTCATG**TACTAAC**ATCGATTGCTTCATTCTTTTTGTTGC
TATATTATATGTTT**AG**AGGTTGCTGCTTTGGTTATTGATAACGGTTCTGGTATGTGTAAAGC
CGGT

UBC4:

GGAAATCTATCCTGAAAAAAAATAAAGTAAAGAAGCCAGGAAAATCACTATCGCCACAAGTA
AATAAATTTCACTGACTATAGAGTACATACATAAACAAGCATCCAAAAAAACATGTCTTCTTC
TAAACGTATTGCTAAAGAACTAAGTGATCTAGAAAG**GT**ATGTCTAAAGTTATGGCCACGTTT
CAAATGCGTGCTTTTTTTTTAAAACTTATGCTCTTATT**TACTAAC**AAAATCAACATGCTATTG
AACT**AG**AGATCCACCTACTTCATGTTCAGCCGGTCCAGTCGGCGATGATCTATATC

TEF4:

GGACTGGTGTCCATCCTTAATACCAACTCAATTCTTTGATTTTGGCCCTCGATAGATTCAAG
ATGTCCCAAGGTACTTTATACATTAATAGATCTCCAAGAAACTACGCTTCTGAAGCTTTGAT
TTCTTACTTTAAACTAGATGTCAAGATCGTTGATCTAGAACAATCTAGCGAGTTTGCTTCTTT
GTTCCCATTGAAGCAAGCCCCTGCATTTTTGGGTCCAAAGGGCTTAAAGCTAACTGAAGCT
TTGGCTATCCAATTTTATT**GT**ATGTTCCATATTTTATTTTAACACTTCCACATACATTGTTTTG
CGCGTTTCCCGTTCGTTTATTTGGCACGTCATTTTTCTTCGAAACATAATGATGAAAAAAAAT
TTTATCAAACAGTTATCCCTGTCTGAATGGGTAATAATAGGTAACCTCTCATATGTTGATATT
TGTATTTCTGATATGTTTCTTAAAGAAAAATGAAAGTCAAAATAACAAAAAGGAAGTACAAAG
CGTGTCCTAATCCAGGAAAAAATATAAAGGATTGTTT**TACTAAC**ATTTCTTCAATTAATGGTT
GTATAACCATCGAGATGATGTATATTT**AG**TGGCTAATCAAGTTGCCGATGAAAAAGAAAGA
GCTCGCTTATTAGGTTCC

IMD4:

GGTTGGGCTTAAATTTTTTTACGAAAAAATGGTGGCTTTTACATTTTCTCTGGTCTTCTTTGC
TGCTTGAACCAAAGGATATTGGACCAATTCCATAGCTTTGAAGAAACCTAACAAACATTTTA

CGATGAGTGCTGCTCCATTGGATTACAAAAAGGCTTTAGAACATTTGAAGACATACAGTTCC
AAGGACGGTCTTTCTGTCCAGGAATTGATGGACTCCACAACTAGAGGTGGGTTGACCTACA
ATGATTTTTTGGTCTTACCAGGCTTGGTCAATTTCCCATCTTCTGCTGTCAGTTTGCAAACC
AAATTGACCAAGAAAATCACTTTGAACACTCCTTTTGTCTCTTCTCCTATGGACACAGTCAC
TGAAGCTGATATGGCTATTTATATGGCTTTATTGGGTGGTATTGGTTTCATCCATCACAACT
GTACTCCAAAGGAACAGGCTTCCATGGTCAAGAAAGTTAAAATGTTTGAAAACGGTTTCATC
AATTCTCCAATAGTAATTTCTCCAACCACCACTGTTGGTGAAGTTAAGGTTATGAAGAGAAA
GTTTGGTTTCTCTGGCTTCCCAGTTACTG**GT**ATGTTATAATAGACAAAATTTTTGGTATTATG
TTTCTTTCATTAGAATGCGCTTTTTAATCATCCATAAAGTTTTCCAAGGTGTGAAAAAGATCT
AAGATGATGATCAACTTTTTATATCAATAACTTTCGTTCTACTGACTGTGATCAAACGATCTT
GTAGAGAACTTTTACTCTGAATTAAATCATATTCGCACCTTTTTCGGAAAGCCCATTATGTGT
GATTAATTGACGTGCAGTGCACTTGAGGAACTTAATATATGTGATAAAAAAATATGTGGCTT
ATTT**TACTAAC**ATTGCTATGTAAATACTGTTATTTTTTTTCTTTTCTGGCTGGGCTGAAAATTT
CTTCAGTCTTCATAAAGTTGAGAAAATAAAAGGAAACAAGTACATCTTTTTGAACAC**AG**AAG
ACGGTAAGTGTCCAGGTAAGTTAGTTGGGTTAGTCACCTCTCGTGAT

RPS17A:

GGAAAGTCAAGATCTCGAGACTAGCAATAACAAAATG**GT**ATGTTAATATGGACTAAAGGAG
GCTTTTAAGGACACGTAATATTGAGTCGAC<u>ATGCGCAATAAAGTCATACAATAAATAAGTTA
AAGAAAACATATGACGATATTTTCAGACGATATGGAACTGCGATACCGAAAAAGTGGTGTG
ATTGAAATGATACTCATGTTTTGGCTATAGTTGCAGTGGCTTTCCAAAAAATTGGGAGTTAA
AATGCCTATAGCAGTATCACCGGAGATTCAGTCTCTTGTGGATCTTCTTTTTGCATAGTTGA
ACAAGGGATAATATGGCGCCATGAACTGCCTCTACACATTATAAT</u>GTGTTTTTGATATCAGT
A**TACTAAC**AAGTTGAATTGCATTTACAAACTTTTTATTTTGTATTGCTTTTCGTCATTTTAAT**A**
GGGTAGAGTTAGAACCAAGACCGTCAAGCGTGCTTCTAAGGCTTTGATTGA

The part of the *RPS17A* intron that is deleted in the *RP51A* pre-mRNA is underlined.

The intron-labeled *RPS30A* pre-mRNA was constructed from three transcripts (the exons, splice site and branch point sequences are shown as above) that were synthesized using T7 RNA polymerase:

The first transcript,

m^7G(5′)ppp(5′)GGCUGCACAUUAAAUUUGCCACUGUAAUAAUCUUCCAUAUCCCCA
UACAAAAACUACGCAAAUAUGGUACGUAAUGCCUCAUUAGGAUUUUAAAAUAUGA
AAAAAAUGGACGUUUCAGCGA,

was synthesized with a 5′ cap analog (New England Biolabs). The second transcript

GGAAUGAAUAACAAGAAAUAGCUAAUUUGAAUAUAUUACUGGUGUUUCAUAUCU
AUAUAUACCGCAGAAAAAGAAAUAAAACGCAAUAUAUCUUUAAACUGGCUCCACC
CUGAGAAUAUGAUAGGAAACUUUUUAAAUUUGAGUUCUGUUACAUCAAAAACGU
AAUUGCUUGACACAGAUUUUUCGAGGGCAACAGCGUGGACUUAACAUAAUGUCG
AGUUUAAUAACAACAUUUUUUGACGGAGAGUUAAAAA,

was synthesized using 1:1 molar ratio of UTP:5-(3-aminoallyl)-UTP, then labeled with Alexa488 tetrafluorophenyl ester (ARES DNA labeling kit, Invitrogen, Molecular Probes). The third transcript,

GGCUUUGAGAAAAUAGCAGAGUGAAGUUUAUUUUUAGCAUCCUUGCUGAUACGC
CAAAUAAUUCAAAAAUUCACUA**UACUAAC**AUGGUCAUUAAUUUUUAAUAUUAAUA
UUAUUUAUUU**AG**GCUAAAGUUCAUGGUUCUCUAGCUCGUGCUGGUAAAGUCAA
GUCUCAAAC,

was post-transcriptionally biotinated using splint-directed extension by Klenow fragment (3'→5' exo$^-$) (New England Biolabs) and the splint oligonucleotide (5'-GTTTTTTATTATGTTTGAGACTTGACTTTACCAGddC-3') with biotin-dCTP in addition to dATP and dTTP nucleotides in the extension mix. The three transcripts were joined by splinted ligation using T4 RNA ligase 1 (Moore and Query, 2000; Stark et al., 2006) to produce the intron-labeled *RPS30A* pre-mRNA.

The average number of Alexa488 dyes per intron-labeled *RPS30A* was determined spectrophotometrically by first measuring the dye concentration using ε_{495} = 71,000 cm^{-1} M^{-1} for the Alexa488 absorbance at 495 nm. The RNA concentration was determined after correcting the absorbance at 260 nm as $A_{260} - 0.3\ A_{495}$, where 0.3 is the Alexa488 $\varepsilon_{260}/\varepsilon_{495}$ ratio. *The RPS30A* concentration was then calculated from the corrected RNA absorbance at 260 nm (one unit corrected absorbance at 260 nm = 40

µg/ml, M_r = 181,576 g/mol). For the intron-labeled RPS30A preparation this procedure measured eight Alexa488 dye moieties per RNA molecule.

Preparation of the yeast strain to monitor U1, U2 and U5 binding

The yeast strain used to simultaneously monitor U1, U2 and U5 binding was prepared from the yeast strain bearing C-terminal eDHFR tags on SNP1 and PRP40 of U1 snRNP (Hoskins et al., 2011) by sequential C-terminal tagging of BRR2 (U5 snRNP) with the $SNAP_f$ tag and of CUS1 (U2 snRNP) with the $CLIP_f$ tag through $(Gly\text{-}Ser\text{-}Gly)_2$ linkers as described in (Hoskins et al., 2011). The fast CLIP ($CLIP_f$) gene was amplified from the $pCLIP_f$ plasmid (New England Biolabs) by PCR, digested with the HindIII and BamHI restriction enzymes, and ligated into the pAG61 plasmid containing the *URA3* marker (Euroscarf) at the HindIII and BamHI sites. This created the pAG61-fCLIP plasmid (pAAH0026), and the $CLIP_f$ insert was fully sequenced. The fast SNAP ($SNAP_f$) gene was amplified from the $pSNAP_f$ plasmid (New England Biolabs) by PCR, digested with the HindIII and BamHI restriction enzymes, and ligated into the pAG25 plasmid containing the *NAT1* marker (Euroscarf) at the HindIII and BamHI sites. This created the pAG25-fSNAP plasmid (pAAH0034), and the $SNAP_f$ insert was fully sequenced.

Glycerol gradient fractionation of the triply labeled WCE

The triply labeled WCE was prepared and the $SNAP_f$ tag of BRR2 and $CLIP_f$ tag of CUS1 were labeled with SNAP-Surface® 488 and CLIP-Surface™ 547, respectively, as described in Experimental Procedures. For fractionation of snRNPs, 200 µL of 40% WCE in splicing buffer without PEG and ATP was loaded onto a 13 ml glycerol gradient (10 - 30% glycerol in the same buffer; gradient formed by a Gradient Master (Biocomp)). The WCE used for the fractionation was either untreated or treated with micrococcal nuclease as described in He et al., 2008. After the ultracentrifugation (29,000 rpm, Beckman SW40Ti rotor, 24 hours at 4 ºC) the 500 µL fractions were collected from the top of the gradient using a density gradient fractionator (Isco) by introducing a 50% glycerol solution to the bottom of the tube. Proteins from each fraction were precipitated

by trichloroacetic acid and analyzed by 8% SDS-PAGE. Fluorescent proteins bands were detected by a Typhoon scanner (GE) and band intensities were measured by ImageQuant software (GE).

Supplemental References

He, F., Amrani, N., Johansson, M.J.O., Jacobson, A. (2008) Methods In Enzymology *449*, 127-147.

http://compbio.soe.ucsc.edu/yeast_introns.html

Moore, M.J., and Query, C.C. (2000). Joining of RNAs by splinted ligation. Methods In Enzymology *317*, 109-123.

Stark, M.R., Pleiss, J.A., Deras, M., Scaringe, S.A., and Rader, S.D. (2006). An RNA ligase-mediated method for the efficient creation of large, synthetic RNAs. RNA *12*, 2014-2019.

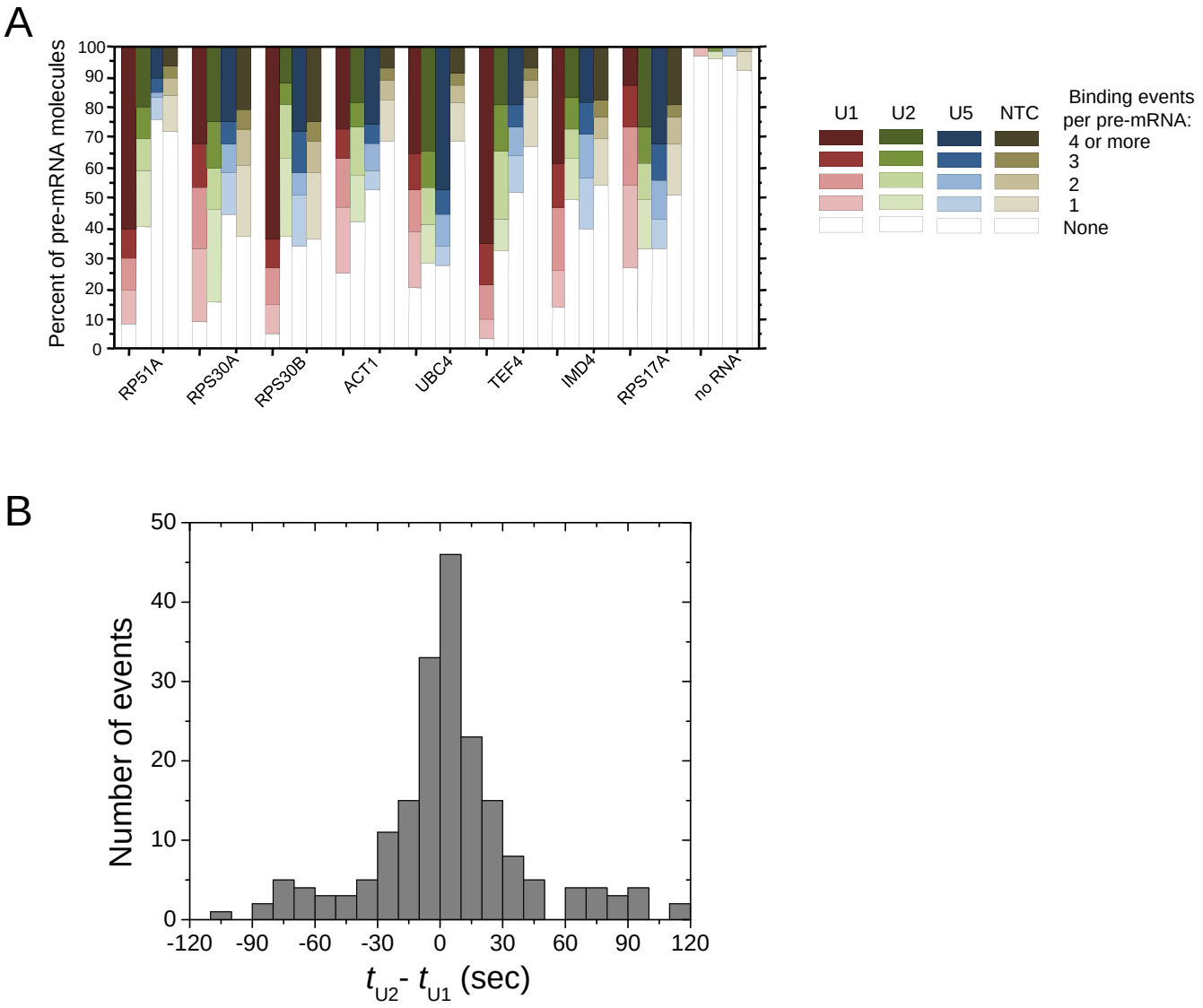

Figure S1. Related to Figure 2.

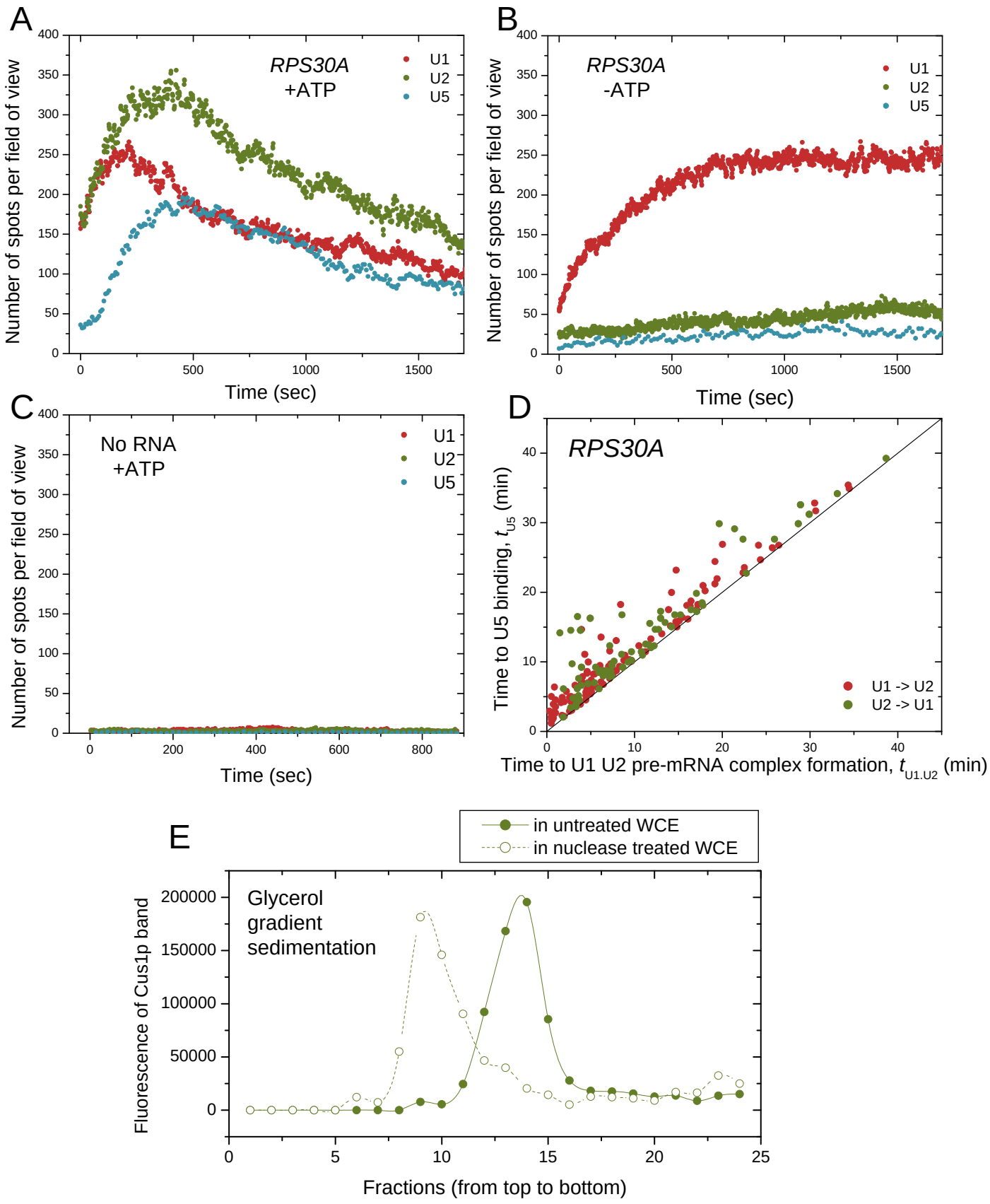

Figure S2. Related to Figure 3.