Understanding Pre-mRNA Dynamics in Single Spliceosome Complexes

by

Ramya Krishnan

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemistry)
in the University of Michigan
2013

Doctoral Committee:

Professor Nils G. Walter, Chair
Professor Hashim M. Al-Hashimi
Professor Carol A. Fierke
Assistant Professor Aaron C. Goldstrohm

To *amma*, *appa*, Chikki, Girish and Abhi

Acknowledgements

First, I would like to thank my advisor, Dr. Nils Walter for taking me into his group and giving me the independence to pursue the science I wanted to. My days spent doing experiments have been fun alongside Mario Blanco and Matt Kahlscheuer, whom I have to thank for being great colleagues with a solid team spirit. I am honored to have worked in close collaboration with two great scientists in the field, Drs. John Abelson and Christine Guthrie. I would like to thank them for the insightful and encouraging discussions. For my time spent as a GSI, and for being a source of constant encouragement and support, I want to extend my wholehearted gratitude to Dr. Kathleen Nolta. Finally I would like to thank my friends in the Walter lab for maintaining a congenial environment, full of fun and life.

Of course, none of this would be possible without the unconditional love and support from mom, dad and my sister. My mom has always encouraged me to excel in whatever I did, and my dad has set an example for the confident person I am today. I know this will make them proud. I would also like to thank my in-laws for helping us when we needed it the most. Lastly to Girish and Abhinav, for making me laugh, loving me to pieces, taking me shopping and putting up with me, I owe it to you.

Table of Contents

Acknowledgements iii

List of Figures viii

List of Tables xi

List of appendices xii

Abstract xiii

CHAPTER 1:Introduction 1

1.1 RNA, the versatile biomolecule 1

1.2 Spliceosome, the cellular mystery machine 2

1.2.1 The interrupted gene 2

1.2.2 The Ribonucleoprotein complex 3

1.2.3 Finding splice sites in a cellular haystack 5

1.2.4 Types of RNA splicing 8

1.2.5 Helicases in splicing 8

1.2.6 An *in vitro* splicing system 10

1.3 Single molecule fluorescence microscopy- an ideal tool for studying dynamic systems 11

1.3.1 Single molecule FRET 11

1.3.2 Splice site juxtaposition using smFRET 12

1.3.3 SiMPull-FRET, a refined look at single molecules 14

CHAPTER 2: Biased Brownian ratcheting leads to pre-mRNA remodeling and capture prior to the first step of splicing 16

2.1 Introduction 16

2.2 Materials and Methods 18

2.2.1 Affinity purification of the B^{act} complex 18

2.2.2 Cloning, expression and purification of splicing factor proteins 19

2.2.3 Single molecule FRET of purified spliceosomal complexes 20

2.2.4 Fluorescent labeling of Cwc25 and distance estimation from FRET 21

2.2.5 Single Molecule Data Analysis 21

2.3 Results 23

2.3.1 Immunopurification of functional complex B^{act} with fluorophore labeled Ubc4 pre-mRNA 23

2.3.2 In the B^{act} complex, the pre-mRNA is in a stable low FRET state with distal 5'SS and BP 31

2.3.3 Prp2 mediates a large-amplitude, NTP-dependent conformational remodeling of the pre-mRNA 36

2.3.4 Cwc25 enhances the first step of splicing by increasing the residence time in the H state 41

2.3.5 Cwc25 dynamically interacts with the pre-mRNA close to the BP after Prp2-mediated pre-mRNA remodeling 50

2.4 Discussion 53

2.5 Acknowledgements 57

CHAPTER 3: Single molecule dissection of splice site juxtaposition along the pre-mRNA splicing cycle 58

3.1 Introduction 58

3.2 Materials and Methods 60

3.2.1 Synthesis of Ubc4 pre-mRNA variants 60

3.2.2 Yeast extracts and modifications 61

3.2.3 Single molecule measurements 61

3.2.4 Creating an HMM similarity matrix and Clustering Analysis 62

3.3 Results 63

3.3.1 Hierarchical clustering tools facilitate objective classification of complex smFRET data 63

3.3.2 Pre-mRNA in early splicing cycle shows dynamic variation in splice site proximity 79

3.3.3 Pre-mRNA predominantly explores close splice site proximity during late splicing cycle 84

3.3.4 Clustering analysis groups biologically similar data 88

3.3.5 The splice sites sample unique conformational paths towards splicing 91

3.4 Discussion 95

CHAPTER 4: An early ATP independent role for Prp28 in commitment complex formation 98

4.1 Introduction 98

4.2 Materials and Methods 100

4.2.1 Splicing extract preparation 100

4.2.2 Native gel analysis 100

4.2.3 Ubc4 pre-mRNA 100

4.2.4 Expression of recombinant Prp28 (rPrp28) protein 101

4.2.5 Single molecule FRET 101

4.3 Results 102

4.3.1 Absence of Prp28 protein inhibits ATP-independent CC2 formation 102

4.3.2 Prp28 allows the 5'SS- BP to explore conformations of close proximity 106

4.3.3 Titration of various concentrations of Prp28 shifts the equilibrium 113

4.4 Discussion 115

4.5 Acknowledgements 117

CHAPTER 5: Summary and Outlook 118

5.1 Pre-mRNA splice site dynamics using smFRET and SiMPull-FRET 118

5.2 Focusing on the first and second step splicing complexes 119

5.3 Labeling pre-mRNA and proteins for smFRET studies 123

5.4 A possible mechanism of action for spliceosomal helicases 125

APPENDICES 127

References 147

List of Figures

Figure 1.1 Canonical Splicing pathway 7
Figure 1.2 SmFRET setup for studying pre-mRNA dynamics 13
Figure 2.1 The SiMPull-FRET approach used here to interrogate active splicing complexes. 26
Figure 2.2 Verification of specificity of the affinity purification of single spliceosome complexes. 27
Figure 2.3 Affinity purified B^{act} complex is active and can go through both steps of splicing.....28
Figure 2.4 Protein expression and purification confirmed by SDS-PAGE analysis.....29
Figure 2.5 Affinity purified B^{act} complex formed with actin pre-mRNA shows very low levels of first step splicing when complemented with Prp2, Spp2 and ATP under splicing conditions. 30
Figure 2.6 In the Prp2-stalled B^{act} complex, the pre-mRNA is predominantly restricted to a stable low FRET state. 32
Figure 2.7 K-means clustering of HMM assigned FRET states. 33
Figure 2.8 Cross-correlation analysis showing the stable L2 state in the B^{act} complex with no significant cross correlation. 35
Figure 2.9 Upon the addition of ATP, Prp2 and Spp2, the pre-mRNA is able to explore splice site proximity. 37
Figure 2.10 The NTPase activity of Prp2 induces structural remodeling of the pre-mRNA. 40
Figure 2.11 Under C complex conditions, the pre-mRNA accesses dynamic and stable high-FRET states.....43
Figure 2.12 Cwc25 enhances the first step of splicing by stabilizing the H state. 44
Figure 2.13 Kinetic rate determination for B* and C complex transitions. 48
Figure 2.14 Prp2 mediated pre-mRNA remodeling creates a binding site for Cwc25 near the BP. 51
Figure 2.15 Model for the conformational mechanism of first-step splicing. 56
Figure 3.1 Stalling the progress of splicing at various steps 64
Figure 3.2 Mutant pre-mRNA substrates show expected stall in the splicing cycle 65

Figure 3.3 Depletion of U2 snRNA results in stalling of splicing cycle.....66
Figure 3.4 Depletion of U2 snRNA results in stalling of splicing cycle.....67
Figure 3.5 Heat-inactivated of Prp2 protein results in stalling of splicing cycle 68
Figure 3.6 FRET probability histograms comparing early and late steps of splicing.....70
Figure 3.7 FSM: A Hidden Markov Modeling derived similarity matrix.....72
Figure 3.8 Hierarchical tree constructed from clustering single molecules 74
Figure 3.9 Choosing a hierarchical tree pruning threshold 75
Figure 3.10 Vectorized FSM for each cluster.....77
Figure 3.11 Representative molecules showing dynamic behavior unique to each cluster 78
Figure 3.12 Conformational dynamics of early splicing stall at commitment complex formation 80
Figure 3.13 Conformational dynamics of early splicing stall at A complex formation.....83
Figure 3.14 Conformational dynamics of late splicing stall at B^{act} and C complex formation.....86
Figure 3.15 Conformational dynamics of pre-mRNA that can proceed through both steps of splicing 87
Figure 3.16 Hierarchical tree grouping the different experimental conditions 90
Figure 3.17 Consensus behavior of the identified hierarchical clusters.....93
Figure 3.18 Consensus behaviors aligned with splicing cycle 94
Figure 4.1 Depletion of Prp28 results in CC1 accumulation 104
Figure 4.2 Inhibition of CC2 formation in the absence of Prp28 in Ubc4 pre-mRNA.....107
Figure 4.3 SmFRET set-up and representative traces.....108
Figure 4.4 Addition of rPrp28 protein shifts the 5'SS-BP of Ubc4 pre-mRNA to explore conformations of close proximity 110
Figure 4.5 Post-synchronized histogram showing the increase in the high FRET transitions upon rPrp28 addition 111
Figure 4.6 BP mutant 5'SS-BP dynamics are not affected by Prp28 112
Figure 4.7 Pre-mRNA dynamics analyzed with varying concentrations of Prp28.....114

Figure A.1 Immunopurification of the Bact complex and reconstruction of the second step of splicing 131
Figure A.2 Recombinant expression of second-step splicing proteins Prp16, Prp18 and Slu7 ..132

Figure A.3 Model for the mechanism of second step catalysis via pre-mRNA remodeling.....133
Figure B.1 Fluorescent labeling of U5snRNA via 3'-periodate oxidation 139
Figure B.2 Reconstitution of snRNA depleted extracts with fluorescently labeled snRNA.....140
Figure B.3 Assembling Cy3-labeled U2 snRNA on doubly labeled (Cy3-Cy5) pre-mRNA. 141
Figure B.4 Single molecule binding experiment with Cy5-U2 snRNA and Ubc4 pre-mRNA...142

List of Tables

Table 1. 1 Examples of different types of RNA molecules found in a cell.....4

Table 2.1 Sequence information of oligonucleotides used in this study.....25
Table 2.2 K-means clustering parameters used on the HMM assigned FRET states. 34
Table 2.3 Comparison of average photobleaching times and number of molecules per condition. 46
Table 2.4 Classification of molecules from the observation of the same molecule chased from the B* to the C complex with the inclusion of a dark period during Cwc25 addition. 49

Table B.1 Sequence of DNA oligonucleotides for U2 and U6 snRNA depletions 145

List of appendices

Appendix A: SiMPull-FRET to study pre-mRNA dynamics at the second step of splicing 127

Appendix B: Utilizing fluorescently labeled snRNAs to characterize the spliceosome 136

Abstract

The spliceosome is a large RNA-protein complex that catalyzes pre-mRNA splicing by removing sequences called introns and joining of the remaining sequences, called exons, to produce a mature mRNA. Despite over a quarter century of research on splicing, little is known about the compositional and conformational rearrangements, timing, and coordination of this process. Particularly, the pre-mRNA, which provides the template for spliceosome assembly and the reactive sites for splicing chemistry, has been largely ignored. This thesis utilizes single molecule fluorescence resonance energy transfer (smFRET) approaches on a short budding yeast pre-mRNA to address this challenge. Specifically, we have placed fluorescent dyes near the conserved splice sites that are recognized by the spliceosome and have monitored dynamic changes in the distance between these sites, in real-time throughout the splicing cycle. We find that, contrary to conventional depictions, the splice sites are highly dynamic and explore reversible transitions. By stalling the progress of splicing at various steps, we have been able to associate unique dynamic information of the splice sites with specific steps of splicing using novel analysis methods that have broad applicability. We find that the splice sites explore reversible splice site proximity in a non-monotonic fashion throughout the process. Our results show that even at very early steps of splicing assembly, the splice sites are brought close together via a novel ATP-independent role of a helicase. Furthermore, employing a combination of smFRET and affinity purification (termed SiMPull-FRET), we have been able to describe the conformational dynamics of single isolated spliceosomes and find them to follow a biased Brownian ratcheting mechanism leading up to the first chemical step of splicing. Our results hint

at the possibility that, much like the ribosome, the spliceosome and its substrates often toggle between active and inactive conformations that are subsequently locked into the preferred state by a specific cofactor. The work presented in this thesis provides a structural and dynamic view of the pre-mRNA in the spliceosome, finds associated roles for protein factors, and pioneers single molecule techniques to answer focused questions about the mechanisms of RNA:protein complex assembly and catalysis in general.

CHAPTER 1 :Introduction

1.1 RNA, the versatile biomolecule

Nucleic acids are the biomolecules responsible for transmission of genetic information from one generation to another. The role of DNA as the genetic material has been well-known since the mid-1900s[1,2]. DNA has garnered much public attention due to its functional properties that make it the cell's master molecule. In an experimental tour-de-force, the cryptic code of DNA was cracked, and it became immediately obvious how the genetic information could be passed from one generation to the next[3-6]. RNA, the messenger molecule, allows to decode the message stored in DNA into proteins[7]. RNA is one of the few biomolecules that can perform two diametrically opposite functions; code for proteins, thereby functioning as a gene; and catalyze reactions, thereby functioning as an enzyme. Francis Crick already had the prescience to suggest that apart from its role as a messenger decoding DNA, RNA could have the ability to act as an enzyme, which was later proved by Tom Cech[8]. Even at this now seemingly mundane role, RNA exhibits structural and functional variety. Messenger RNA (mRNA) faithfully copies the genetic material, and serves as the template for ribosomes (composed of ribosomal RNA, rRNA) to synthesize proteins, which in turn is aided by transfer RNA (tRNA), the adaptor molecules bringing in the building blocks, amino acids.

With the discovery of reverse transcription[9,10], split genes[11,12], ribozymes[8], and more recently the completion of the human genome project, it has become increasingly clear that RNA plays diverse and unexpected roles (some well-known examples shown in Table 1. 1). Given the relatively small fraction of the human genome dedicated to protein coding, one might expect that only a small fraction of the genome is transcribed into RNA as well. This is far from the truth. Large portions of the genome in humans and other mammals are transcribed into RNA that is not classified as mRNA, tRNA or rRNA. These special-function RNAs, now known as non-coding RNAs, are being rapidly discovered and have roles ranging from structural scaffolding to regulation of gene expression. This versatility has led to much speculation about a possible role of RNA in the origin of life[13].

1.2 Spliceosome, the cellular mystery machine

1.2.1 The interrupted gene

After the discovery of the structure of DNA, it was assumed that a gene comprised of a continuous sequence of base pairs that in its entirety was used as a template for protein synthesis. In the simplest form, a gene is a length of DNA that directly corresponds to its protein product. Although largely correct for bacteria, seminal discoveries in the 1970s challenged this view for eukaryotes[11,12]. It was observed that the RNA transcribed from DNA in the nucleus was unusually long compared to the mRNA found in the cytosol. Resolution of this mystery led to the discovery of split genes and a process called RNA splicing. A simple yet profound comparison of the DNA and its corresponding RNA sequence shed light on the fact that some of the intervening sequences were absent from the mature cytosolic RNA product. It was thus

proved that genes can be composed of several separate segments. Drs. Philip A Sharp and Richard J. Roberts were awarded the Nobel Prize for Physiology or Medicine in 1993 for their independent discovery of split genes. This discovery has been of fundamental importance to today's basic research in biology, as well as more medically oriented research on disease development and research on how genes evolve in organisms.

1.2.2 The Ribonucleoprotein complex

In the cellular environment, RNA is often found complexed with proteins while playing both structural and functional roles in biomolecular machines. A well-known example is the ribosome, the cellular ribonucleoprotein (RNP) complex that facilitates protein biosynthesis. Another example involves the highly precise removal of interviewing RNA sequences called *introns* and the ligation of the flanking sequences called *exons* during the process of precursor-mRNA (pre-mRNA) splicing. This extraordinary task is carried out by a cellular RNP complex called the spliceosome[14]. Extensive genetic and biochemical studies in a variety of organisms have revealed that essential components of the spliceosome include five small RNAs – the U1, U2, U4, U5 and U6 small nuclear RNAs (snRNAs), each of which functions as an snRNP (small nuclear ribonucleoprotein). In addition to snRNPs, splicing requires many non-snRNP protein factors. More than two decades after its discovery, the spliceosome part list is slowly nearing its completion.

Category	Name	RNA Length (nts)	Functions
Coding RNA	Messenger RNA (mRNA)	variable	Protein coding
	Ribosomal RNA (rRNA)	~120-5,000	Translation
	Signal Recognition particle RNA	300	Translation
Housekeeping non-coding RNAs	Transfer RNA (tRNA)	70-100	Translation
	Small nuclear RNA (snRNA)	60-450	Splicing, Histone RNA processing
	Small nucleolar (snoRNA)	70-300	RNA editing
	Telomerase RNA	150-1300	Telomere length maintenance
Small non-coding RNAs (< 200 nt)	MicroRNA (miRNA)	20-24	Post-transcriptional gene silencing
	Small interfering RNA (siRNA)	20-24	Post-transcriptional gene silencing
	PIWI-interacting RNA (piRNA)	24-31	Germline gene silencing
Long non-coding RNAs (> 200 nt)	Long non-coding RNA (lncRNA)	>200	Epigenetic gene regulation
Ribozymes	Hepatitis Delta Virus (HDV) ribozyme	84 nt	viral replication

Table 1. 1 Examples of different types of RNA molecules found in a cell

The spliceosome, much like the ribosome, is an energy consuming machine, but neither requires NTP hydrolysis for the actual chemical reaction to take place[15]. It is intriguing to think that RNA may catalyze the basic chemical reaction of the spliceosome, leaving open what the roles of the various proteins are[16-18]. Here, once again, the similarities between the ribosome and spliceosome are instructive. The ribosome uses energy to power the various conformational rearrangements that occur in this RNP machine to ensure accuracy. Similarly, a widely accepted, yet untested hypothesis is that the spliceosome uses energy in terms of ATP hydrolysis to power the various RNA:RNA structural rearrangements needed to create the active catalytic core of the spliceosome[15,19-21].

1.2.3 Finding splice sites in a cellular haystack

The spliceosome assembles anew on each pre-mRNA conserved recognition sequence[22]. In metazoans, the splice sites are defined as conserved sequences at the intron-exon junctions, that is, the 5' splice site (5'SS) at the 5' end of an intron, the 3' splice site (3'SS) at the 3' end of the intron, and the upstream branchsite (BS) or branchpoint (BP) (Figure 1.1). The most conspicuous feature of the spliceosome is that it lacks a preformed catalytic core, which is created in a stepwise fashion, beginning with the assembly of the U1 and U2 snRNPs at the 5'SS and BP, respectively, to form the pre-spliceosome (A complex). The U4/U6.U5 tri-snRNP then binds to create the mature spliceosome (B complex). Notably, however, U1 and U4 snRNPs must be removed before catalysis, creating first the activated B^{act} complex, and after additional rearrangements, the catalytically active B* complex (Figure 1.1). The resulting post-first-step C complex then undergoes further remodeling required for the second step of splicing and the

formation of mature mRNA. The chemistry of splicing occurs in a two-step transesterification process. In the first step, the 2'OH of the branchpoint adenosine (BP) attacks the phosphodiester bond at the 5' splice site (5'SS), releasing the 5' exon and creating the branched, lariat structure; in the second step, the newly generated 3' hydroxyl of this exon attacks the phosphodiester bond at the 3' splice site (3'SS), releasing the lariat intron and creating the spliced mRNA product. This raises an interesting molecular question of how the spliceosome finds, recognizes and differentiates these sequences within the plethora of RNA sequences in the cell[23]. The sequences constituting the splice sites provide insufficient information to distinguish true splice sites from the greater number of false splice sites that populate transcripts. Even the simplest eukaryote, Saccharomyces *cerevisiae* commonly known as Baker's yeast, with its scarcity of introns that, when present, are short, this problem presents a potential roadblock. The answer to the mystery in part lies in the fact that the splice sites are only treated as global definitions of motifs while the spliceosome actively brings together the reactive units by manipulating the inherent structure of the intron in the midst of the variety surround it. In higher eukaryotes, this even more complex problem is simplified by the presence of auxiliary sequences that assist in highlighting the splice sites and aiding the spliceosome to act as a molecular bridge spanning the intron, thereby bringing the reactive exons together. The structure and conformations adopted by the intron play a critical role in regulating the progress through the splicing cycle. Although traditionally pre-mRNA is represented as a linear sequence until after the first chemical step, this is far from the truth. It has been shown in recent years that the pre-mRNA is modulated at the structural and

kinetic level throughout the splicing cycle with distinct conformational signatures (Chapters 2 and 3)[24].

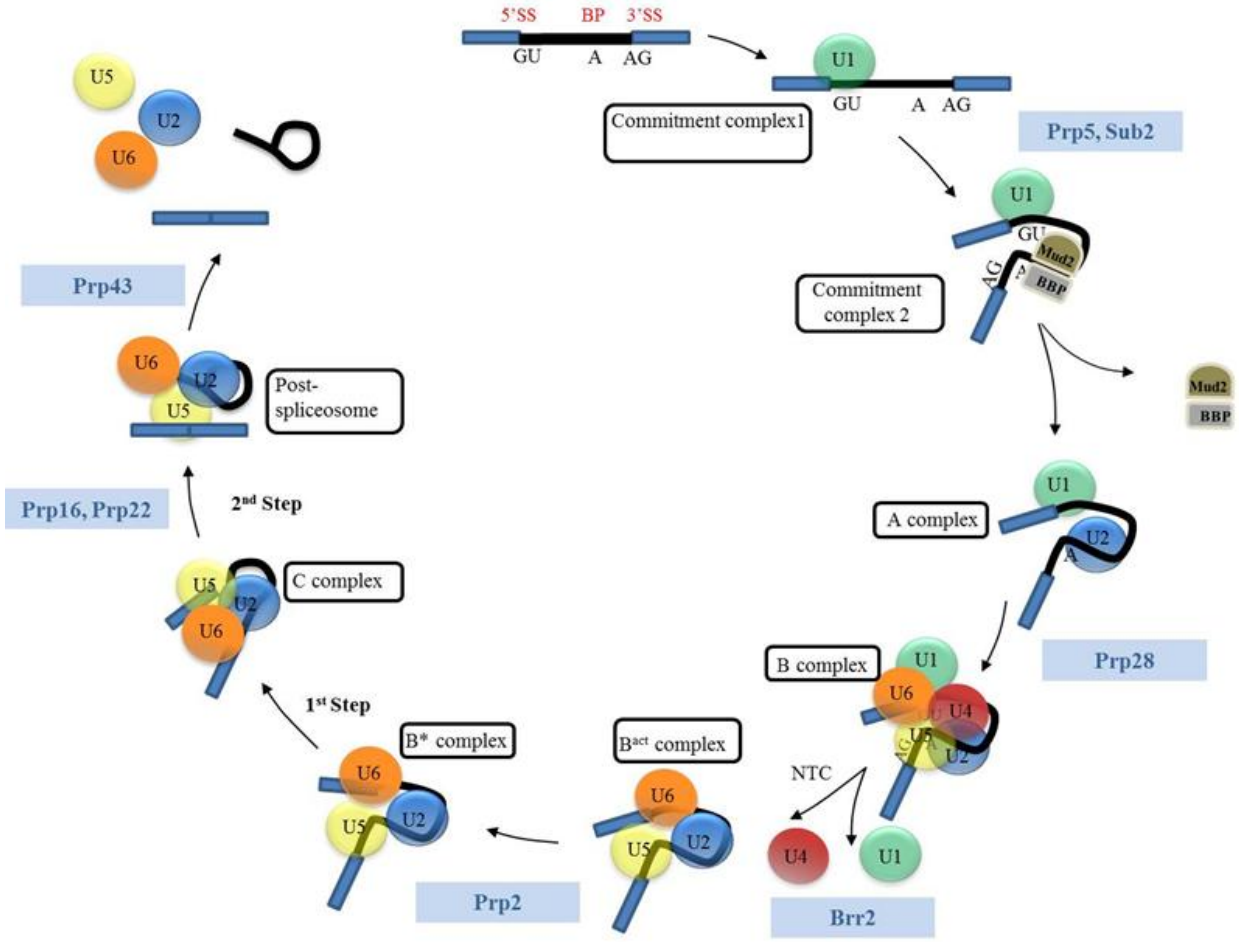


Figure 1.1 Canonical Splicing pathway

Schematic representation of the stepwise assembly of spliceosome on a pre-mRNA. The exons (blue) are spliced and the intron (black) is removed as a lariat. The conserved 5' splice site, 3' splice site and branchpoint are shown on the pre-mRNA. DExD/H box helicases involved are shown at each step (blue boxes).

1.2.4 Types of RNA splicing

There are four well-known types of splicing mechanism prevalent in eukaryotes. The first two called group 1 and group 2 introns share some key characteristics but differ in the details of their splicing mechanism. The third and largest type of splicing involving pre-mRNA takes place in the nucleus via two phosphoestertransfer reactions. Although the actual chemical reaction itself is ATP independent, energy in the form of ATP hydrolysis is required for the various conformational rearrangements leading to catalysis. This reaction is related in mechanism to the self-splicing mechanisms of group 1 and group 2 introns. Specifically, group 2 introns are thought to be spliced by a mechanism identical to that of the spliceosome, although there is no general requirement for either proteins or co-factors[25]. While the fourth and last type of splicing involves tRNA genes that are also interrupted by introns[26], it differs substantially in mechanism as it is mediated by protein enzymes with an intrinsic requirement for ATP hydrolysis.

1.2.5 Helicases in splicing

The spliceosome is a highly dynamic machine undergoing iterations of often mutually exclusive RNA-RNA unwinding events and rearrangements up until the active catalytic core if formed[19]. These rearrangements consume ATP and are facilitated by a large group of conserved RNA DExD/H box helicases. Among them, eight highly conserved helicases play crucial roles in key steps of splicing (Figure 1.1). These proteins are implicated in early events of spliceosome assembly and progression towards catalytic activation, actual catalysis and disassembly and have both ATP-dependent and independent roles. Among them, the ATPase activity of Sub2 is required in the formation of commitment complex by stabilizing the interaction of the Mud2 and

Branchpoint Binding Protein (BBP) proteins with the BP[27]. In an ATP independent manner, Sub2 also helps destabilize the Mud2-BBP heterodimer to allow for U2 snRNP binding at the BS. Prp5 is required for pre-spliceosome formation with its ability to bridge U1 and U2 snRNPs and uses ATP to cause rearrangements leading to the binding of U2 to the BP[28]. With the recruitment of the U4:U5:U6 tri-snRNP, the ATPase action of helicase Prp28 has been shown to be required for the removal of U1 from the 5'SS to allow for the mutually exclusive U6 binding to the 5'SS[15]. In Chapter 4, we have shown evidence for an early, ATP independent role for this protein in the splicing cycle. The U5 small nuclear ribonucleoprotein particle (snRNP) helicase Brr2 then disrupts the U4:U6 small nuclear RNA (snRNA) duplex and allows U6 snRNA to engage in an intricate RNA network at the active center of the spliceosome[29]. Brr2 is tightly regulated to ensure correct timing of spliceosome activation by its interactions with a protein termed Prp8, which is situated at the heart of the spliceosome[30]. Once the catalytic core is formed, spliceosome activation occurs wherein Prp2 is required for the first step and Prp16 and Prp22 are essential for the second step of splicing. The ATP dependent activity of Prp2 has been shown to cause a major remodeling of the spliceosome required for catalytic activation[31]. We found that this role is accompanied by a ratcheting motion of pre-mRNA, allowing it to explore close splice site proximity (Chapter 2). The next helicase Prp16 is required for the second step of catalysis (Appendix A) and for proofreading suboptimal substrates[30,32]. It is involved in the removal of Cwc25, a heat-stable protein that binds close to the BP adenosine and enhances the efficiency of first-step splicing (Chapter 2). Prp22 proofreads exon-ligation and mRNA production while Prp43 is involved in disassembling the spliceosome, allowing recycling of the individual components for yet another round of assembly and splicing[33,34]. Although these

important regulators have been extensively characterized using genetic and biochemical tools, a major goal in the future is to identify their molecular mechanisms of action, as well as the ligands and targets required to achieve accuracy in splicing.

1.2.6 An *in vitro* splicing system

The epitome of understanding any biological system is when the individual components can be assembled *in vitro* and the desired biological effect(s) be recapitulated. The secrets of the spliceosome began to rapidly unravel after the important achievement of the Abelson lab to identify and manipulate particles (which they termed 'spliceosome') that splice exons *in vitro*[14]. Simply put, incubation of a pre-mRNA with crude yeast extract, in the presence of ATP, crowding agents and essential ions, at room temperature, results in the production of mRNA. This technique has been widely exploited for over 25 years to unravel the splicing mechanism. Crude whole cell yeast extract has been the choice material for biomolecular scientists until it was shown that purified spliceosome intermediates could be isolated with the aid of yeast genetics as a tool for stalling the spliceosome at intermediate steps. These intermediary complexes can be coaxed with just a handful of proteins to go through the first and second catalytic steps of splicing[35]. This important discovery has enabled us to assign specific functions to a single protein (Chapter 2). Although, due to the sheer number of components involved, we are still far from able to set up an entirely recombinant *in vitro* splicing system by adding every single essential component, the available *in vitro* splicing systems together with yeast genetics have equipped us with an arsenal of tools to better understand the complex splicing process.

1.3 Single molecule fluorescence microscopy- an ideal tool for studying dynamic systems

1.3.1 Single molecule FRET

Major advances in understanding biological systems in recent years have been facilitated by the advent of single molecule (sm) techniques. These tools have opened a wide array of experimental avenues that were previously considered 'unanswerable'. Sm techniques have the unique and powerful advantage that they are able to probe molecular structure, dynamics and function unhindered by the averaging inherent in ensemble experiments. The study of single molecules can be broadly divided into two facets; one aimed at solving scientific problems and the second aimed at improving the associated sm analysis techniques themselves. One of the most widely used sm techniques is smFRET or single molecule Fluorescence Resonance Energy Transfer. SmFRET is powerful and has been used extensively to study intramolecular dynamics[36-39]. FRET is the non-radiative transfer of energy from a donor fluorophore to an acceptor fluorophore in a distance dependent manner. The amount of FRET is dependent on several factors, but primarily on the distance (R) between the acceptor and donor dyes and their relative dipole:dipole orientations. FRET efficiency decreases proportionally to R^6, where R is the donor:acceptor distance. The commonly used smFRET pairs Cy3 and Cy5 enable us to measure distances ranging from 20-100 Å, thus serving as a molecular ruler. As a result, smFRET techniques have found wide applications in fields ranging from quantum optics and photophysics to the assembly, dynamics and function of proteins and cellular RNP machines.

1.3.2 Splice site juxtaposition using smFRET

Each round of spliceosome assembly requires the step-by-step association and rearrangements of the splicing components on the pre-mRNA, the substrate on which the actual chemistry of splicing takes place. Thus, pre-mRNA conformations can serve as signatures for the different steps in the splicing cycle. To ensure proper fidelity, it is imperative that the nucleotides of the 5'SS, BP and 3'SS are aligned perfectly for chemistry. Although the recognition of the splice sites starts at a very early step in splicing, it is unclear as to when the pre-mRNA is funneled into the right conformation for catalysis to occur. Any mutation of the splice sites can affect this conformation and thereby hinder the assembly process[40]. In order to understand how the splice sites are positioned throughout the splicing cycle, we used smFRET as detected by Total Internal Reflection Microscopy (TIRF) to validate that single pre-mRNA could assemble into single spliceosome complexes (Figure 1.2). By judiciously choosing the location of the fluorophores –both exons labeled and exon/intron labeled – we were able to focus on the first and second step of splicing, respectively[37]. With this *in vitro* assay based on smFRET, we were able to monitor in real-time the dynamic conformational states of the pre-mRNA in spliceosome assembly[37]. Furthermore, we used the powerful tools of yeast genetics to stall the progression of splicing at critical points in the cycle to enrich for complexes at particular points (Chapter 3). Upon immobilizing the pre-mRNA on the microscope slide, we were able to use whole cell extracts to induce and monitor splicing and ATP-dependent changes in the splice site juxtaposition throughout the splicing cycle (Figure 1.2).

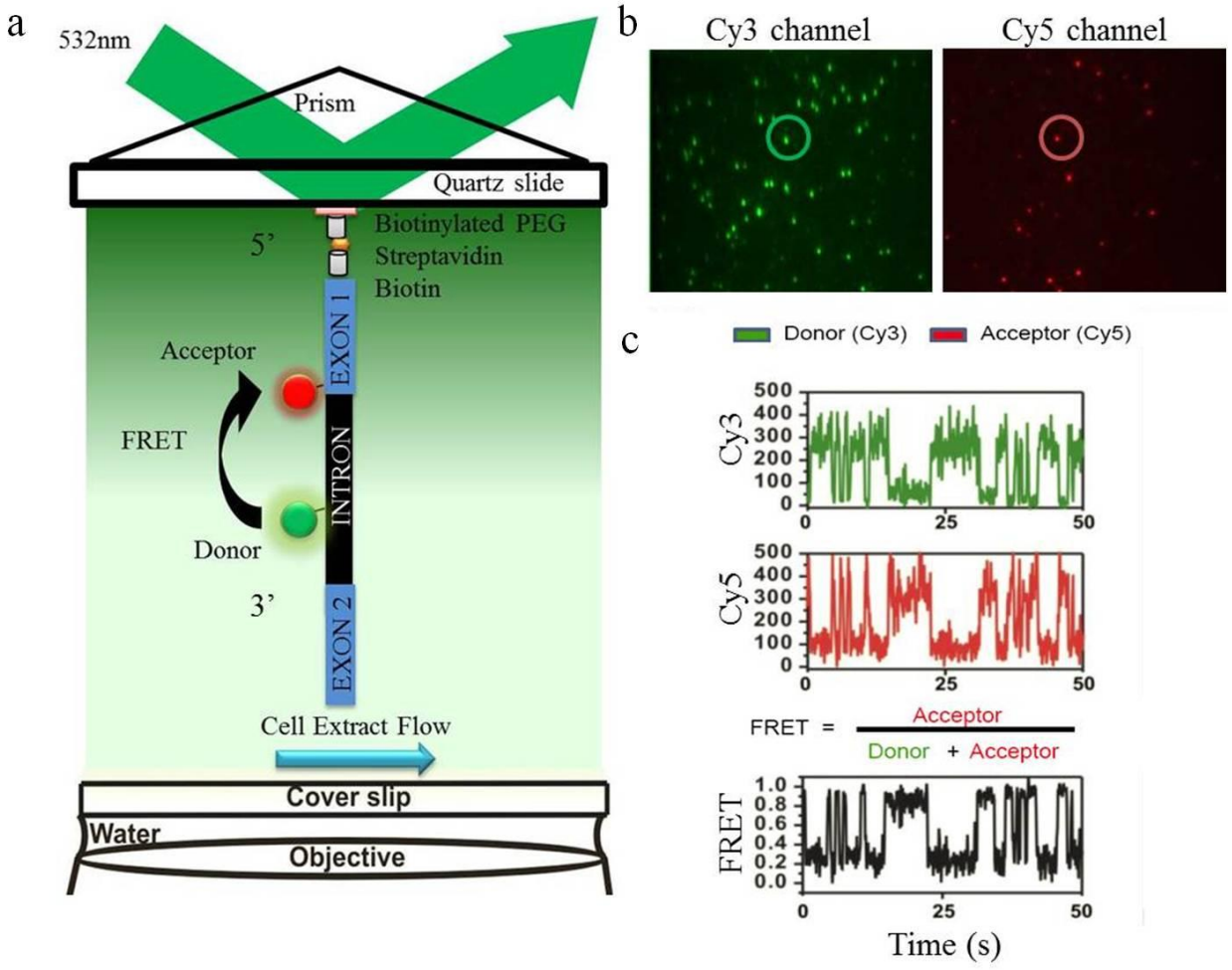


Figure 1.2 SmFRET setup for studying pre-mRNA dynamics

(a) Pre-mRNA with fluorophores at the splice sites is immobilized on a slide via a streptavidin-biotin linkage. Whole yeast cell extract is flowed into the channel. The Cy3 fluorophore (green) is excited using a 532 nm laser to collect smFRET data. (b) The Cy3 and Cy5 channels are mapped to one another to obtain information on single pre-mRNA molecules that are doubly labeled. (c) The anti-correlated signals from a single molecule's Cy3 (green) and Cy5 (red) dyes are tracked here for 50 s.

Due to the highly diverse and dynamic nature of the data obtained from these experiments, traditional sm-analysis tools were not sufficient to fully characterize them. We therefore used clustering algorithms to classify different behaviors (Chapter 3). This method has broad potential for teasing out complex dynamics in biological systems that cannot be efficiently extracted using traditional sm-analysis.

1.3.3 SiMPull-FRET, a refined look at single molecules

Traditionally, single molecule experiments on the spliceosome have been performed using exogenous pre-mRNAs in yeast whole cell extracts. Recently, an elegant approach to combine affinity purification and single molecule techniques was described[23]. We have used a similar Single Molecule Pull down (SiMPull) method on the spliceosome to measure the proximity of the 5'SS and BS during the first step of splicing and thereby give structural information on the spliceosome and its remodeling (Chapter 2). In addition to being able to stall the spliceosome at a defined step of splicing, this technique enabled us to purify a splicing complex of interest from the remaining extract components for TIRF microscopy. The complex thus assembled on the slide surface is free from most non-specific interactions caused by binding and dissociation of RNA binding proteins and extraneous protein-protein interactions not relevant to the particular step of splicing. A prime advantage of the technique used here is that reconstitution of the first step is achieved by the addition of single recombinant protein, which allows us to specifically assign shifts in the pre-mRNA conformation to the action of single components. The SiMPull-FRET experiment holds promise for future work as it opens up

opportunities to investigate other steps using purified components in unusually dynamic systems like the spliceosome.

The work presented in this thesis utilizes a combination of biochemical and single molecule FRET techniques and analysis as a way to characterize the pre-mRNA splice site dynamics at specific steps of splicing. We find that the 5'SS and BP come together in a non-monotonic fashion and are kinetically modulated to enable the chemistry of splicing. The FRET signature we identify at each step sheds light on the conformational remodeling of the pre-mRNA from one step to the other. Additionally, using hierarchical clustering tools, we have been able to identify the FRET dynamics at each step and re-create the entire conformational 'pathway' followed by a single pre-mRNA in order to be efficiently spliced. Although pre-mRNA is the prime focus, the experiments presented here also unravel unique roles for protein regulators in splicing. In particular, we have been able to assign novel roles to the protein Cwc25 in stabilizing a conformation favorable to the first step of splicing. Another focus of this thesis is DExD/H box helicase regulators, whose mechanism of action has been largely unknown. From our results on Prp2 (Chapter 2) and Prp28 (Chapter 4), we hypothesize that the role of helicases in general may be to unlock intrinsic thermal fluctuations of the pre-mRNA:spliceosome complex, allowing for a ratcheting motion that is rectified by Cwc25 as a cofactor pawl. Much like the ribosome, the spliceosome thus functions as a biased Brownian ratchet machine.

CHAPTER 2 : Biased Brownian ratcheting leads to pre-mRNA remodeling and capture prior to the first step of splicing[1]

2.1 Introduction

Introns are removed by the spliceosome, a large ribonucleoprotein complex, in a two-step transesterification process. In the first step, the 2'OH of the branchpoint adenosine (BP) attacks the phosphodiester bond at the 5' splice site (5'SS), releasing the 5' exon and creating the branched, lariat structure; in the second step, the 3' hydroxyl of this exon attacks the phosphodiester bond at the 3' splice site (3'SS), releasing the lariat intron and creating the spliced mRNA product[14]. The most conspicuous feature of this enzyme is that it lacks a preformed catalytic core, which is created in a stepwise fashion, beginning with the assembly of the U1 and U2 snRNPs at the 5'SS and BP, respectively, to form the pre-spliceosome (A complex)[41]. The U4/U6.U5 tri-snRNP then binds to create the mature spliceosome (B complex)[41]. Notably, however, U1 and U4 snRNPs must be removed before catalysis, creating first the activated B^{act} complex, and after additional rearrangements, the catalytically active B*

[1]Adapted from **Ramya Krishnan**[1], Mario Blanco[2], Matthew Kahlscheuer[1], John Abelson[3], Christine Guthrie[3], Nils G. Walter[1;] 'Biased Brownian ratcheting leads to pre-mRNA remodeling and capture prior to the first step of splicing' (under review at *Nat. Struct. Mol. Biol.*, 2013). R.K. performed the biochemical and single molecule experiments. M.L.K. cloned, expressed and labeled the proteins. R.K. and M.R.B. performed data analysis. R.K., J.A., C.G. and N.G.W. wrote the manuscript.

complex. The resulting post-first-step C complex then undergoes further remodeling required for the second step of splicing and the formation of mature mRNA[41].

The highly dynamic process of spliceosome assembly and catalysis is guided by a set of RNA-dependent ATPases of the DExD/H-box helicase family, which collectively function to insure the fidelity of splicing[15]. A major experimental challenge has been to understand the precise conformational rearrangements of RNA and protein that accompany each ATP-dependent step. DExD/H-box helicase Prp2 is required for the first chemical step of splicing and recent proteomic analyses of the B^{act}, B* and C complexes revealed that its action results in the destabilization of the U2 snRNP-associated proteins SF3a and SF3b[31,35,42,43]. An attractive hypothesis is that SF3b sequesters the BP adenosine[44] to prevent a premature attack on the 5'SS, and thus the ATP-dependent action of Prp2, together with its cofactor Spp2, would be required to initiate catalysis. In a biochemical tour-de-force, successful reconstitution of both steps of splicing with the addition of recombinantly expressed proteins to immunopurified splicing complexes has been demonstrated[35,45,46]. In particular, first-step chemistry could be achieved with the addition of ATP, Prp2, Spp2 and Cwc25, a small heat-stable factor splicing factor[35].

We have coupled this advance with single molecule fluorescence resonance energy transfer (FRET) to monitor the dynamics of the pre-mRNA as the BP and 5'SS are brought into close proximity for first-step chemistry. We used the resulting single molecule pull-down FRET (SiMPull[23]-FRET) technique to analyze a functional B^{act} complex assembled on a short (135-nucleotide) pre-mRNA with fluorophores near the scissile bonds. By assembling this complex in an extract with a temperature-sensitive allele of Prp2 (*prp2-1*)[47], the B^{act} spliceosome is stalled and can then be chased to B* with the addition of ATP, Prp2 and Spp2, and finally to C complex

with the addition of Cwc25. Using SiMPull-FRET, we show that ATP-dependent action of Prp2 along with its cofactor Spp2 unlocks reversible switching of the intron between splicing-active and –inactive conformations. Cwc25 then rectifies this thermal Brownian ratcheting by stabilizing the conformation in which the 5'SS and BP are in close proximity, driving the equilibrium toward catalysis.

2.2 Materials and Methods

2.2.1 Affinity purification of the B^{act} complex

Extracts were prepared from a *prp2-1 cef1-TAP* yeast strain, heated at 37 °C for 40 min to inactivate Prp2 and stall the spliceosome at the B^{act} complex. In a final volume of 135 µl, 40% (v/v) of this heat treated extract was incubated with ~50 pmoles FRET labeled Ubc4 pre-mRNA[37] in the presence of 2 mM ATP in splicing buffer (8 mM HEPES-KOH, pH 7.0, 2 mM $MgCl_2$, 0.08 mM EDTA, 60 mM $K_i(PO_4)$, 20 mM KCl, 8% (v/v) glycerol, 3% (w/v) PEG, 0.5 mM DTT) and incubated at 23 °C for 35 min. For biochemical experiments, streptavidin-coated magnetic beads (Dynabeads® MyOne™ pH 7.5, 50 mM NaCl). An equal volume of 0.5 mg/ml IgG-biotin (ZyMAX™ Rabbit Anti-Mouse IgG (H+L) - BT (ZyMAX™ Grade)) in T50 was added and incubated in a tube rotator at RT for 30 min. The beads were then pulled down using a magnet and the supernatant was discarded. To block any streptavidin not bound by biotin-IgG, the beads were incubated with excess free biotin at 1.5 mg/ml in T50 buffer in a tube rotator at RT for 20 min. After equilibration in splicing buffer, the independently assembled splicing reaction were added and incubated in a tube rotator for 30 min at RT to allow the protein A of the Cef1-TAP tag in the spliceosome complex to bind the biotin-IgG. Upon removal of the

supernatant, the beads were further washed three times with buffer A (20 mM HEPES-KOH, pH 7.9, 75 mM KCl, 0.01% NP40, 1.5 mM MgCl2, 5% (v/v) glycerol) and once with splicing buffer to further purify the B^{act} complex. The reactions are scaled up for reconstitution reactions pursued in parallel and split at this step. Prp2, Spp2 and Cwc25 were added at 90-120 nM final concentration in splicing buffer in the presence or absence of 2 mM ATP or AMPPNP or UTP and incubated in the tube rotator for 30-40 min for various levels of reconstitution. RNA was isolated and products of splicing were analyzed on a denaturing, 7 M urea, 15% polyacrylamide gel and scanned on a Typhoon variable mode imager (GE Healthcare).

2.2.2 Cloning, expression and purification of splicing factor proteins

The full-length PRP2 gene was PCR-amplified and ligated into plasmid pRSETA (Invitrogen) with a C-terminal hexahistidine tag. The N-terminally truncated form of SPP2 (coding for amino acids 37-185) containing a C-terminal hexahistidine tag, and the full-length Cwc25 gene were obtained from Reinhard Lührmann (Max Planck Institute for Biophysical Chemistry, Germany). Cwc25 was subcloned into a pRSETa plasmid containing a single cysteine residue and hexahistidine tag at the C-terminus. The constructs were then transformed into Escherichia coli strain Rossetta II (Novagen). Cultures were grown in 2-4 l of TB medium and induced with 125 μM IPTG. Cultures were then incubated at 20 ºC for 18 h. Cells were harvested, washed, and the pellets stored at -80 ºC. Purification of Cwc25-His and Spp2-His was performed as described[35]. Protein purity was confirmed by 16% SDS-PAGE and proteins were either first fluorescently labeled or directly aliquoted, flash frozen in liquid nitrogen, and stored at -80 ºC. Protein concentrations were determined by Bradford assay and measurement at A_{280}.

His-tagged Prp2 obtained from E.*coli* cell lysate was purified as described[48] Protein purity was confirmed by 10% SDS-PAGE and the final product aliquoted, flash frozen in liquid nitrogen, and stored at -80 ºC. Protein concentrations were determined by Bradford assay and measurement at A_{280}.

2.2.3 Single molecule FRET of purified spliceosomal complexes

For the single molecule FRET experiments on affinity-purified complexes, we prepared slides using previously published procedures[49]. In short, the surface of a quartz slide was amino functionalized, PEGylated and reacted with 0.2 mg/ml streptavidin in T50 buffer for 15 min at RT. 100 µl of 0.5 mg/ml IgG-biotin in T50 was flowed onto the slide and incubated for 20 min, followed by free biotin at 1.5 mg/ml in T50 buffer for 15 min. B^{act} spliceosomal complexes were assembled and stalled as described above by incubation of FRET labeled Ubc4 pre-mRNA with heat treated *prp2-1 cef1-TAP* yeast splicing extract in splicing buffer supplemented with 2 mM ATP and an oxygen scavenger system (OSS) composed of protocatechuate dioxygenase, protocatechuate and Trolox[37]. These complexes were then flowed onto the slide surface and incubated for 15-20 min to allow the Cef-1-TAP on the spliceosome to bind biotin-IgG. The slide surface was washed rigorously and reconstituted (in the presence of OSS) as described for the biochemical purification and incubated for 10-40 min before acquiring data. A home-built prism-based TIRF microscope was used to collect data as described[37,50,51]. To obtain FRET data, we directly excited the Cy3 donor near the BP adenosine with a 532-nm laser, and we recorded emission by Cy3 and Cy5 fluorophores at 100-ms time resolution using an intensified CCD camera (I-Pentamax, Princeton Instruments).

2.2.4 Fluorescent labeling of Cwc25 and distance estimation from FRET

The single-cysteine mutant of Cwc25 was labeled with Cy5-maleimide (GE Healthcare). Labeling was performed using 0.150 μmol of Cwc25 in storage buffer and 0.5 mg of dye containing 10 μM reducing agent Tris(2-carboxyethyl)phosphine (TCEP) (Sigma). Reactions were incubated at RT for 1 h followed by overnight at 4 ºC. Free dye was removed by re-purification of protein on a Ni^{2+} column and dialysis back into storage buffer. The degree of labeling was determined using GE Healthcare's protocol and was found to be 70%. Protein functionality was confirmed using an ensemble pull down assay as described above. The fluorophore distance, R, and the apparent FRET efficiency, E_{app}, were calculated as described[50,52,53] from the equations $E_{app} = c[1 + (R/R_0)^6]^{-1}$, where c = 0.69, R_0 = 54 Å, and $E_{app} = \frac{I_{Cy5}}{I_{Cy5} + I_{Cy3} \times \frac{(\phi_{Cy5} \times \eta_{Cy5})}{(\phi_{Cy3} \times \eta_{Cy3})}}$. φ and η signify the fluorophores quantum yields and detector channel efficiencies, respectively. The donor and acceptor intensities I_{Cy3} and I_{Cy5}, respectively, were corrected for leakage of donor photons into the acceptor channel.

2.2.5 Single Molecule Data Analysis

The donor and acceptor intensities I_{Cy3} and I_{Cy5}, respectively, were corrected for leakage of donor photons into the acceptor channel Cross-correlation analysis was carried out utilizing customized MATLAB scripts with built-in xcorr function. Time lags for the cross-correlation ranged from 0-50. Quality control for the raw smFRET trajectories obtained from the experimental conditions were performed as described[49]. Histograms for data sets measuring pre-mRNA dynamics were constructed by sampling 100 frames of data from each molecule.

Histograms for Cwc25-Cy5 and BP-Cy3 FRET experiments were constructed by sampling the entire length of data. Hidden Markov Modeling (HMM) was performed on trajectories utilizing the vbFRET software suite[49]. Each trajectory was individually fit with models ranging from 1-5 states with the optimal number of states determined by the vbFRET algorithm. The inherent experimental variations of the FRET signal between single molecules leads to a slightly different state assignment for similar states across different molecules. A K-means clustering approach was therefore performed in MATLAB to group similar states into larger macro states (L1, L2, M and H). A matrix cataloging the HMM assigned FRET state, raw FRET level, and difference in donor and acceptor intensities for each HMM derived event was utilized as input for the K-means algorithm. A total of four macro states were identified whose boundaries were used to re-assign the original HMM idealized FRET states. Transition Occupancy Density Plots (TODPs) were used to plot the fraction of molecules that contain any given HMM transition at least once[49]. Molecules that did not exhibit any transitions were plotted along the diagonal at their respective positions. For kinetic rate calculations, Transition Density Plots (TDPs) that are scaled by the number of times a transition occurs irrespective of how many molecules exhibit that transition were used as described[49]. A cumulative histogram scatter plot was then fit with a double-exponential association equation in MicroCal Origin. A weighted average ($k_{w,observed}$) of the two rate constants from the double-exponential fits was calculated based on the amplitude values of the exponential equation. To correct for bias introduced by the limited observation window used to measure dwell times, the measured $k_{w,observed}$ values were corrected by subtracting the photobleaching rate constant and the reciprocal of the observation window to yield $k_{w,actual}$ as described[54]. Equilibrium constants (K_{eq}) were calculated by taking a ratio of the

forward and backward rate constants for a set of state-to-state transitions. Post-synchronized histograms (PSHs) were constructed by synchronizing individual FRET events to the time where one of the macro-states (M, H) was achieved. The scale bar represents the fraction of FRET events which exhibit a certain FRET state at a given time.

2.3 Results

2.3.1 Immunopurification of functional complex B^{act} with fluorophore labeled Ubc4 pre-mRNA

It has been known for almost 25 years that a *prp2-1* yeast splicing extract can be heat-inactivated[47]. In this extract, the spliceosome is fully assembled but cannot carry out the first step of splicing. The prp2-1 immature spliceosome purified by centrifugation (B^{act}) was shown to be chased into first-step product by the addition of Prp2p and a heat stable factor(s)[45]. More recently, Lührmann and colleagues have repeated this experiment with purified Prp2, Spp2 and the (now identified) heat stable factor Cwc25[35]. Such a purified system is ideal for exploring substrate dynamics using single molecule FRET. To this end, we constructed a yeast strain containing the *prp2-1* mutation and a Tandem Affinity Purification (TAP) tag derivative of one of the NTC components, Cef1, known to be present in the spliceosome at this stage[31,35]. This allowed us to purify the stalled B^{act} complex via the TAP tag.

The Ubc4 pre-mRNA used in this study was synthesized chemically and labeled with the FRET donor Cy3 +6 nucleotides downstream from the BP adenosine and with the FRET acceptor Cy5 -7 nucleotides upstream from the 5'SS (Table 2.1). Heat inactivated whole cell extract from the aforementioned strain was then used to assemble splicing reactions containing the fluorophore labeled pre-mRNA in the presence of 2 mM ATP. The reaction was allowed to

proceed for 30 min to form the pre-catalytic B^{act} spliceosome, then bound either to streptavidin coated magnetic beads for biochemical analysis or to a PEG-passivated slide coated with streptavidin, biotin-IgG, and excess free biotin for single molecule analysis (Figure 2.1). Free biotin was added to block any biotin binding sites not associated with biotin-IgG and prevent direct binding of the 5' biotinylated pre-mRNA to the slide. Indeed, we observed at least 11-fold reduced surface binding of the pre-mRNA by itself compared to the Cef1-TAP-tagged B^{act} complex containing the pre-mRNA Figure 2.2). We detected similarly minimal non-specific binding of the tagged complex upon omission of IgG-biotin or when instead using a TAP-tag on protein Prp4, recently shown to be absent from the B^{act} complex[31] (Figure 2.2). To verify that our purification yielded functional B^{act} pre-spliceosome capable of catalyzing both steps of splicing, we added micrococcal nuclease (MN) treated (and thus RNA-free) whole cell extract to the stalled bead-purified B^{act} complex and incubated for 40 min at 23 ºC under splicing conditions. We found that both steps of splicing were reconstituted (Figure 2.3), demonstrating that only proteins are needed to chase B^{act} into splicing the FRET-labeled Ubc4 pre-mRNA. To verify that reconstitution of the first step of splicing could be achieved with the addition of recombinant proteins, we bead-purified the B^{act} complex and found it to require recombinant proteins Prp2 and Spp2 (Figure 2.4) to yield a low level of first-step splicing products (Figure 2.1), slightly above that observed for the previously characterized actin pre-mRNA[35] (Figure 2.5). When additionally Cwc25 was added, a 2- to 10-fold enhancement of first-step splicing was observed (Figure 2.1). Taken together, these experiments establish that FRET labeling the pre-mRNA substrate is compatible with the expected assembly and splicing activity of the immunopurified B^{act} complex, paving the way for SiMPull-FRET interrogation.

Ubc4 (20/20) 3' splice site mutant (3'ss)	5'-GAACUAAGUGAUC(5-N-U)AGAAAGGUAUGUCUAAAGUUAUGGCCACGUUUCAAAU *GCGUGCUUUUUUUUUAAAACUUAUGCUCUUAUUUACUA**A**CAAAA(5-N-U)CAACAUGCUAUUGAACUACA**C**AUCCA* CCUACUUCAUGUUT-3'
Ubc4 (20/20) Wildtype (WT)	5'-GAACUAAGUGAUC(5-N-U)AGAAAGGUAUGUCUAAAGUUAUGGCCACGUUUCAAAU *GCGUGCUUUUUUUUUAAAACUUAUGCUCUUAUUUACUA**A**CAAAA(5-N-U)CAACAUGCUAUUGAACUAGAGAUCCA* CCUACUUCAUGUUT-3'
DNA splint-dSplint	5' - GTTGATTTTGTTAGTA AATAAG(SP9)GTTTTAAAAAAAAAAGCACGC -3'

Table 2.1 Sequence information of oligonucleotides used in this study

Ubc4 intron is italicized, and the allyl-amine modified uridines are denoted as (5-N-U). The red and green colors represent positioning of the Cy3 and Cy5 fluorophores respectively. In the 3'SS mutant, the bold and underlined cytosine replaces the guanine in the wildtype. The bold and underlined A is the BP adenosine. dSplint is the DNA splint used for templated ligation to synthesize the pre-mRNA as described34 Sp9 denotes a 9-carbon linker.

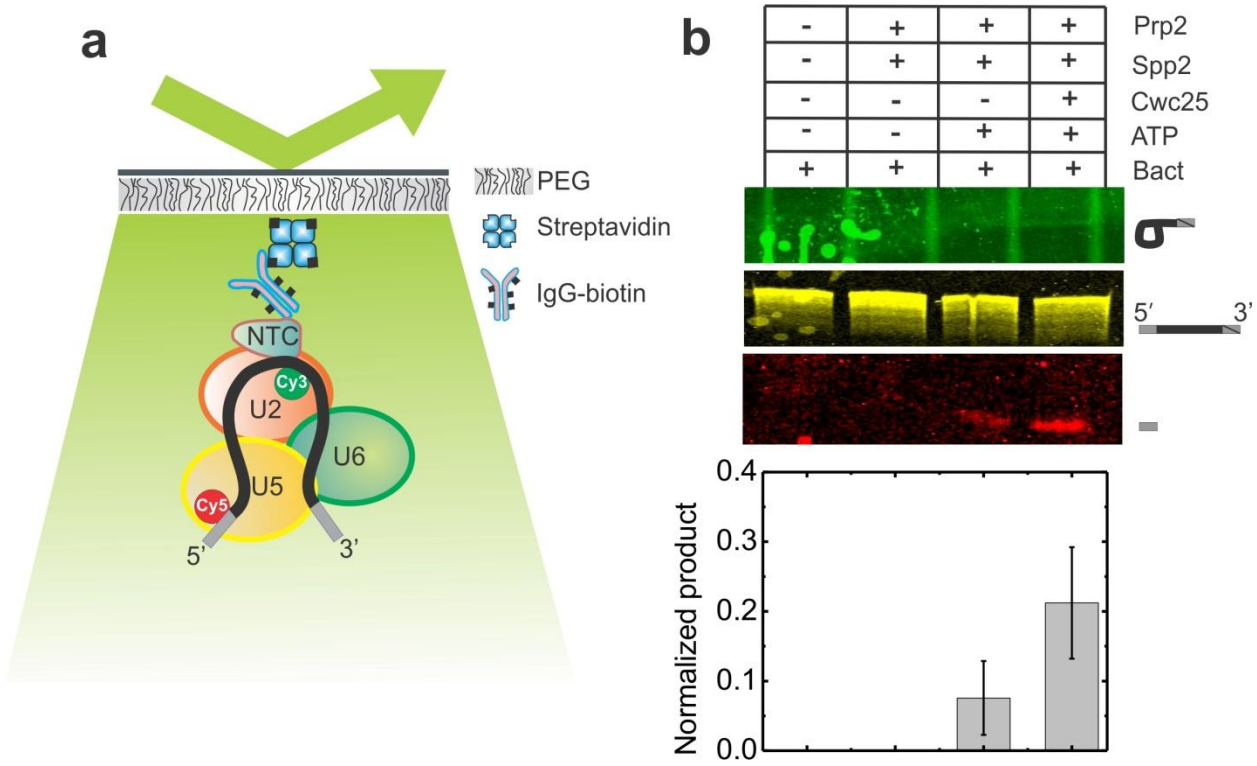


Figure 2.1 The SiMPull-FRET approach used here to interrogate active splicing complexes.

The SiMPull-FRET approach used here to interrogate active splicing complexes. (a) Schematic showing the affinity purified B^{act} complex immobilized to a streptavidin coated quartz slide via biotinylated-IgG. The green and red circles on the pre-mRNA represent Cy3 and Cy5 fluorophores, respectively, for FRET. (b) The affinity purified B^{act} complex is active and can be chased through the first step of chemistry with the addition of Prp2, Spp2, Cwc25 and ATP. A denaturing polyacrylamide gel scanned using a variable mode Typhoon imager shows pre-mRNA and first-step products (the top panel shows intron-lariat in a Cy3 scan, the mid panel shows pre-mRNA using an overlay of Cy3 and Cy5 scans, and the bottom panel shows the 5'-exon in a Cy5 scan). Histogram quantification shows enhancement of first-step splicing upon Cwc25 addition.

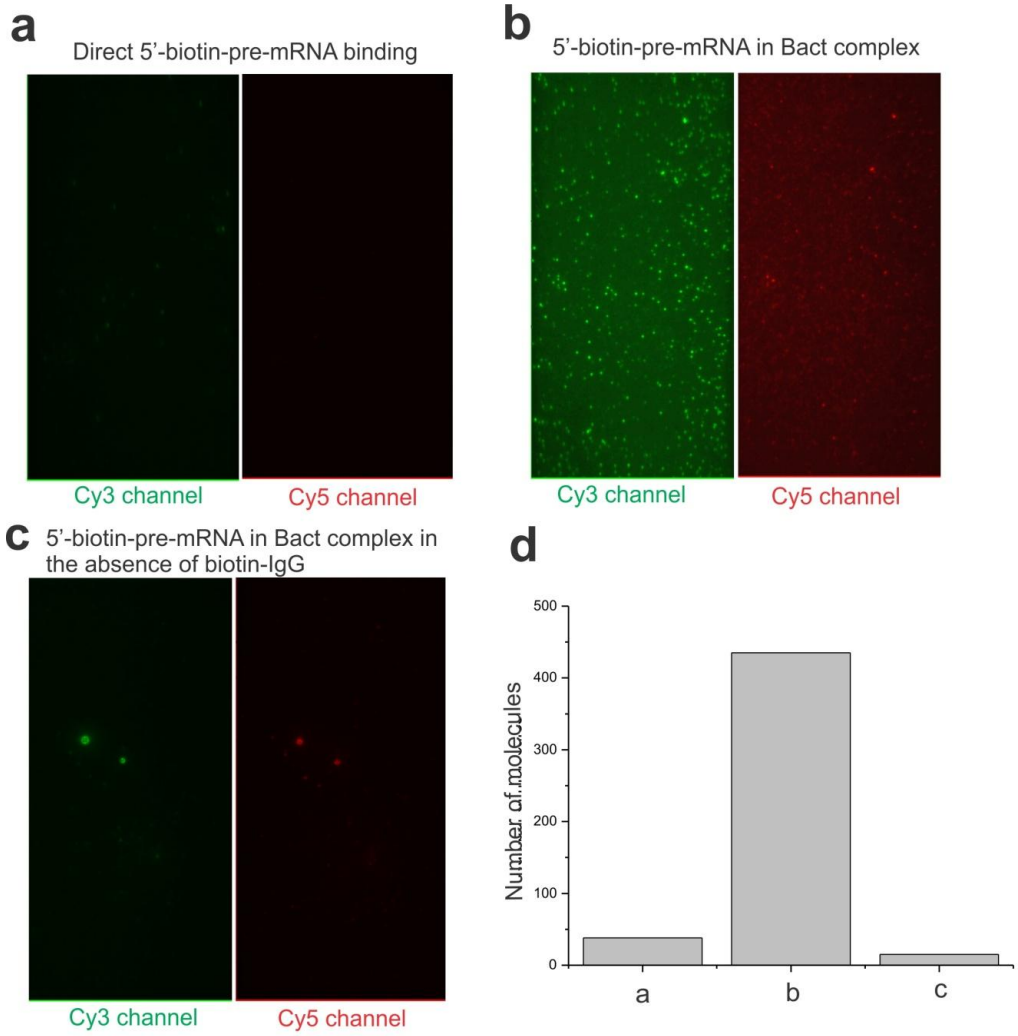


Figure 2.2 Verification of specificity of the affinity purification of single spliceosome complexes.

(a) Field of view showing direct binding of the 5' biotinylated pre-mRNA to the Streptavidin on the slide surface saturated with biotin-IgG and free biotin. (b) Field of view showing the binding of the immunopurified B^{act} spliceosome (with Cef1-TAP) to the streptavidin on the slide surface saturated with free biotin and biotin-IgG. (c) Field of view showing the binding of the immunopurified B^{act} spliceosome (with Cef1-TAP) to the streptavidin on the slide surface saturated with free biotin in the absence of IgG-biotin. Left and right panels are the Cy3 and Cy5 channels, respectively. A 532-nm laser was used for excitation under all conditions. (d) Quantification of number of molecules under conditions a-c.

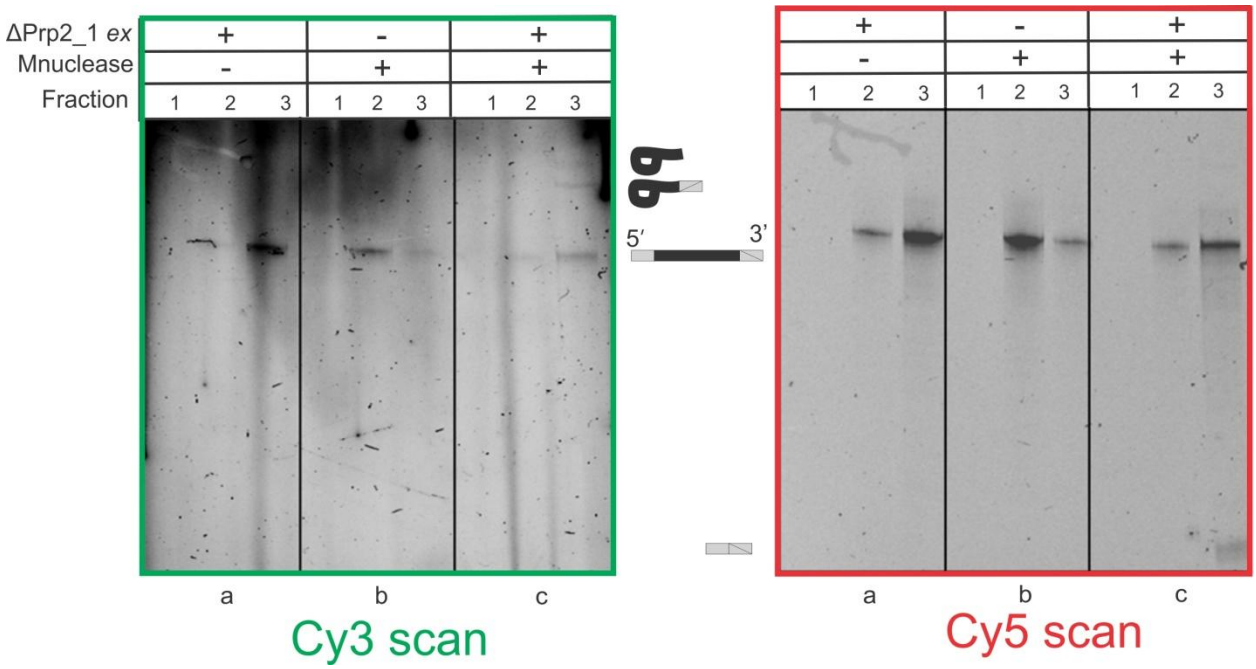


Figure 2.3 Affinity purified B^{act} complex is active and can go through both steps of splicing.

Lanes 1, 2 and 3 represent fractions wash, unbound and supernatant, respectively. Condition a is Ubc4 pre-mRNA assembled in ΔPrp2-1,Cef-I TAP extract under splicing conditions with ATP at 23 °C for 60 min. The B^{act} complex thus formed was affinity purified on streptavidin coated magnetic beads loaded with biotin-IgG. Unbound pre-mRNA is in lane 2 and the beads are further washed to remove non-specifically bound components (lane 1). The remaining fraction released from the beads and incubated with 2 mM ATP in splicing buffer is in lane 3. Condition b is wild-type pre-mRNA assembled in the absence of ΔPrp2-1,Cef-I TAP extract with 2 mM ATP in splicing buffer at 23 °C for 60 min. Most of the pre-mRNA is found in the unbound fraction (lane 2). Condition c is pre-mRNA assembled in ΔPrp2-1,Cef-I TAP extract to form the B^{act} complex as described in condition a. The fraction bound to the magnetic beads was then reconstituted with Micrococcal Nuclease (MNuclease) treated extract and analyzed in lane 3. The intron and intron-lariat products are observed in the Cy3 scan (left) and the mature mRNA product is visualized in the Cy5 scan (right).

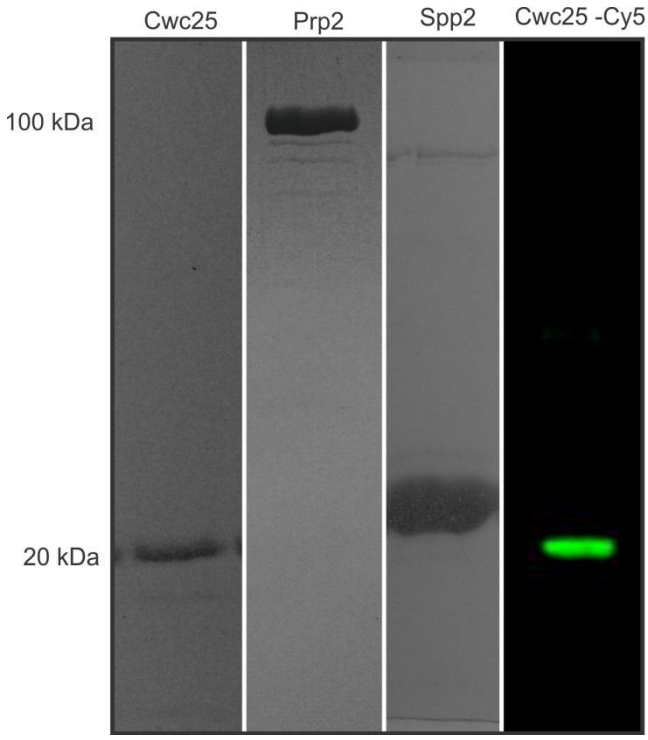


Figure 2.4 Protein expression and purification confirmed by SDS-PAGE analysis.

Histidine-tagged Prp2, Spp2 and Cwc25 are shown in lanes 1, 2 and 3, respectively. Cy5-fluorophore labeled single-cysteine Cwc25 is shown in lane 4. Prp2 was analyzed on a 10% SDS-PAGE, while the other proteins were analyzed on a 16% SDS-PAGE.

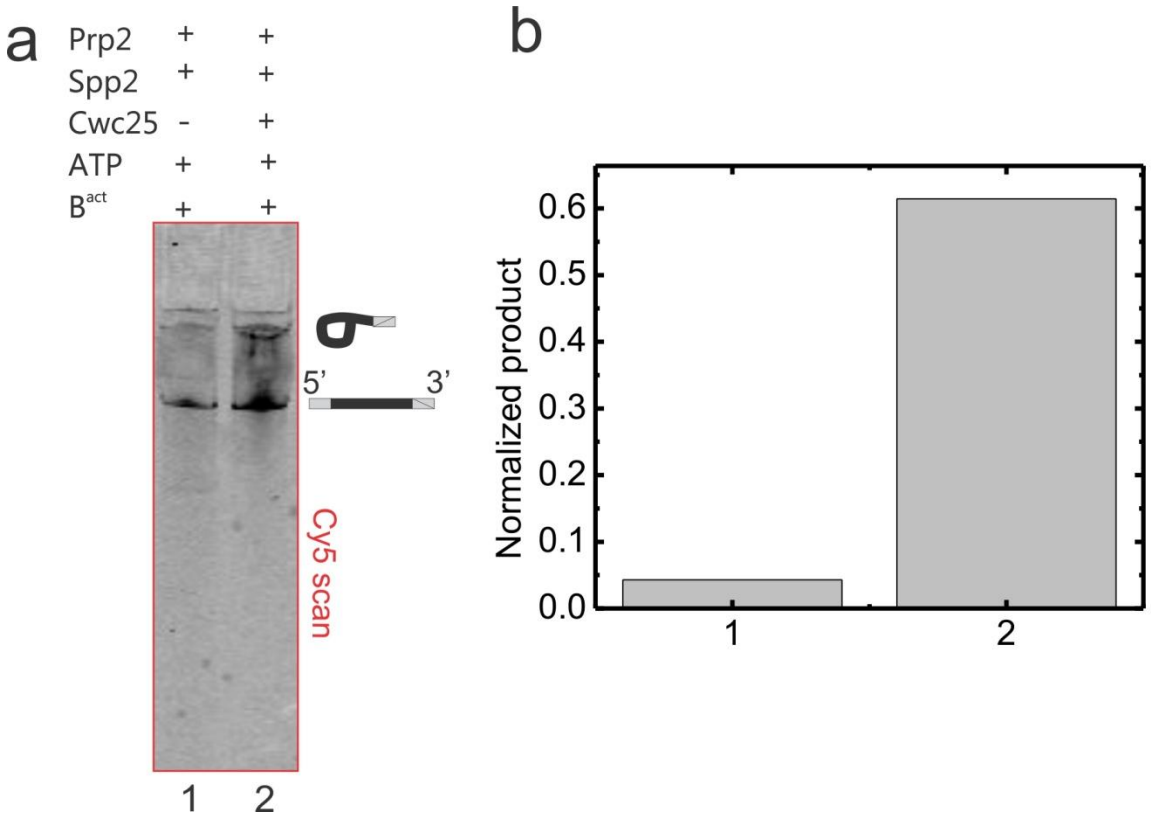


Figure 2.5 Affinity purified B^{act} complex formed with actin pre-mRNA shows very low levels of first step splicing when complemented with Prp2, Spp2 and ATP under splicing conditions.

(a) 6% Urea-polyacrylamide gel scanned using a variable mode Typhoon imager showing pre-mRNA and 1^{st} step products in the following conditions- affinity purified B^{act} complex supplemented with Prp2, Spp2 and 2 mM ATP (lane 1) and Prp2, Spp2, Cwc25 (lane 2). (b) Quantification of lanes 1 and 2 showing a 15-fold enhancement of 1^{st} step splicing products upon Cwc25 addition.

2.3.2 In the B^{act} complex, the pre-mRNA is in a stable low FRET state with distal 5'SS and BP

To obtain mechanistic insight into pre-mRNA splice site juxtaposition during the crucial Prp2-driven restructuring of the B^{act} complex into the catalytically activate B* complex, we carried out SiMPull-FRET on the slide-bound B^{act} complex (Figure 2.1) in standard splicing buffer. After initial quality control, FRET histograms plotted using the FRET values collected from 297 verified molecules over the first 100 frames (at 100 ms time resolution) indicated a single Gaussian distribution with an average FRET value of 0.3±0.15 (Figure 2.6). To objectively identify the underlying FRET states in the shot-noise dominated fluorescence signal, we employed Hidden Markov Modeling (HMM) as described[49]. Using a K-means approach, we then clustered the HMM assigned states using all the experimental conditions used in this study into four macro states with FRET values of 0.0-0.23 (state L1), 0.23-0.42 (L2), 0.42-0.60 (M), and 0.6-1.0 (H) (Figure 2.7 and Table 2.2). The dominant behavior in the B^{act} complex is a stable low-FRET state L2 (Figure 2.6). Transition Occupancy Density Plots (TODPs), which are scaled to emphasize the transitions found to be most common among a molecule population[49], indicate that the stable L2 state represents the only behavior in ~52% of all B^{act} molecules (Figure 2.6). In addition, molecules in this state have very few transitions. To test for dynamics too fast for detection by HMM, we performed cross-correlation analysis between the donor and acceptor trajectories of each molecule and in the resulting scatter around zero found no evidence for rapid transitions (Figure 2.6 and Figure 2.8). Although splice site recognition begins in the splicing cycle as early as the commitment complex[55] our results suggest that the 5'SS and BP in the B^{act} complex are kept stably apart, likely not yet close enough for splicing chemistry to occur.

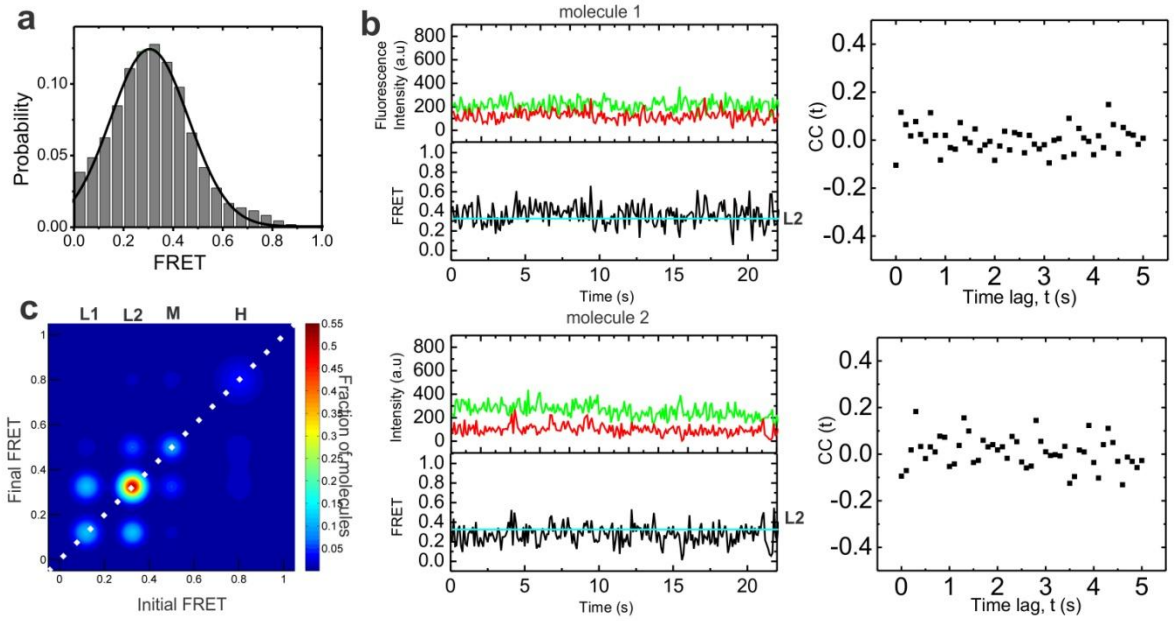


Figure 2.6 In the Prp2-stalled B^{act} complex, the pre-mRNA is predominantly restricted to a stable low FRET state.

(a) FRET histograms generated by binning the first 10 s of the raw single molecule FRET trajectories from the purified B^{act} complex. (b) Representative time traces of the B^{act} complex with raw donor (Cy3, green), acceptor (Cy5, red), and FRET (black) trajectories and their idealized HMM (cyan) showing a stable L2 FRET state. The right panel of each molecule is the corresponding cross-correlation of donor and acceptor intensities, demonstrating that the HMM did not miss any donor-acceptor anti-correlated events. (c) Transition Occupancy Density Plots (TODPs) showing that stable L2 FRET states, which lie on the diagonal (dotted white line), characterize more than 50% of all B^{act} complexes.

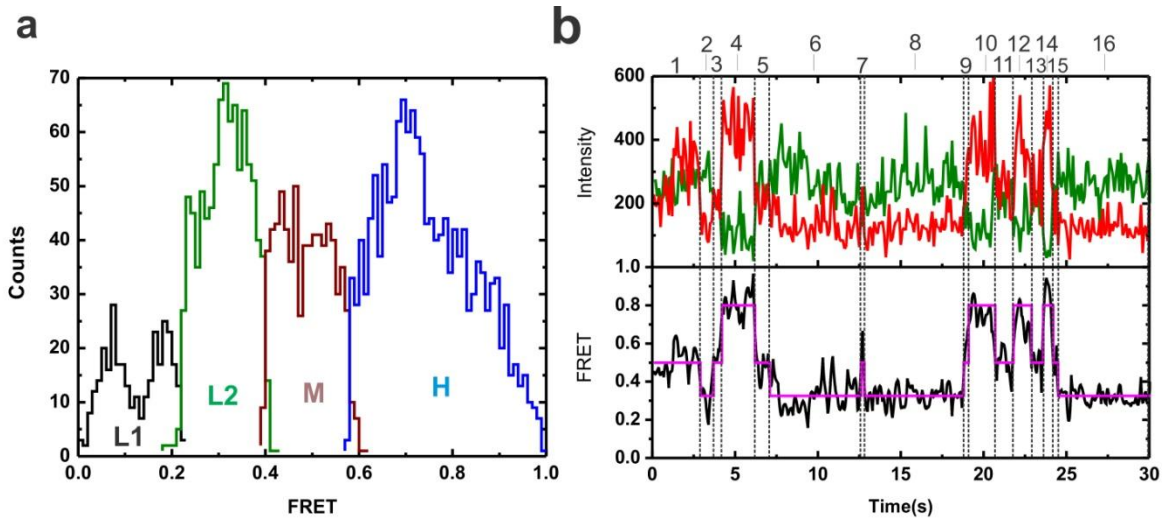


Figure 2.7 K-means clustering of HMM assigned FRET states.

(a) Histogram of HMM-idealized states for each of the K-means derived clusters L1, L2, M and H obtained by clustering single molecule trajectories from all the experimental conditions (B^{act}, B* and C). (b) Representative molecule showing the FRET states assigned by HMM upon K-means clustering.

Segment	μ_{FRET}	HMM state	μ_{ACP}-μ_{DNR}/ (μ_{ACP}+μ_{DNR})	Cluster
1	0.51	0.5	0.039	M
2	0.31	0.325	-0.37	L2
3	0.51	0.5	0.030	M
4	0.82	0.8	0.628	H
5	0.47	0.5	-0.075	M
6	0.32	0.325	-0.352	L2
7	0.56	0.5	0.123	M
8	0.32	0.325	-0.353	L2
9	0.55	0.5	0.111	M
10	0.76	0.8	0.525	H
11	0.50	0.5	0.012	M
12	0.72	0.8	0.454	H
13	0.52	0.5	0.061	M
14	0.87	0.8	0.747	H
15	0.45	0.5	-0.084	M
16	0.31	0.325	-0.366	L2

Table 2.2 K-means clustering parameters used on the HMM assigned FRET states.

Segment number (first column) the mean raw FRET value (second column), HMM state (third column), and normalized difference of mean acceptor and mean donor intensities for each segment (fourth column).

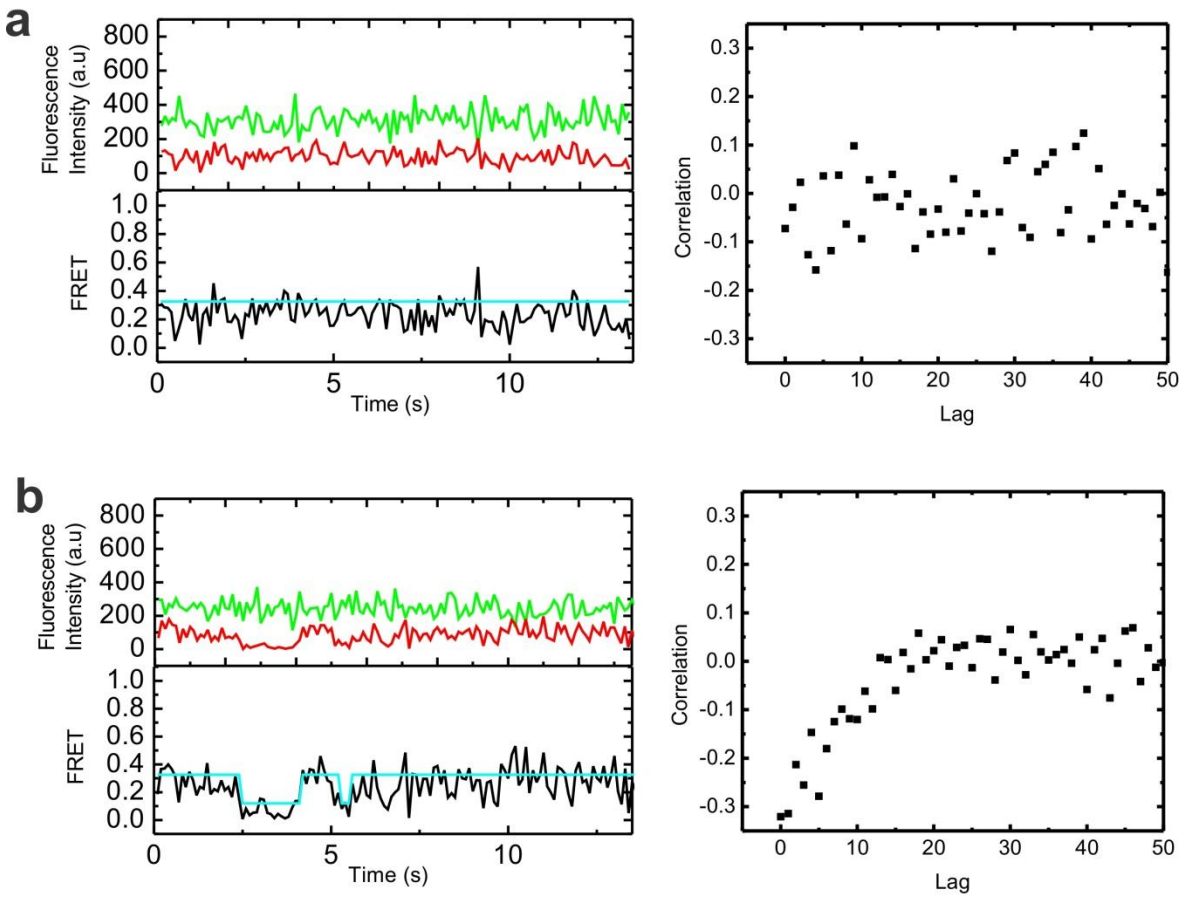


Figure 2.8 Cross-correlation analysis showing the stable L2 state in the B^{act} complex with no significant cross correlation.

(a) Sample trace from the B^{act} showing showing no transitions detected by HMM and no significant cross correlation indicating a true stable FRET state with the raw donor (Cy3, green), acceptor (Cy5, red), FRET (black) trajectories, idealized HMM models (cyan) and cross correlation analysis of donor and acceptor trajectories with time lags from 0-50. (b) Sample dynamic traces from the B^{act} showing anti correlated transitions with raw donor (Cy3, green), acceptor (Cy5, red), FRET (black) trajectories, idealized HMM models (cyan) and cross correlation analysis of donor and acceptor trajectories with time lags from 0-50.

2.3.3 Prp2 mediates a large-amplitude, NTP-dependent conformational remodeling of the pre-mRNA

The ATPase action of Prp2 has been shown to catalyze a large conformational change that activates the spliceosome for the first step of splicing[31,35,45]. Spliceosomal binding of Prp2 is dependent on its interaction with the G patch domain of its cofactor protein Spp2[56,57]. The addition of Prp2, Spp2 and ATP transforms the pre-catalytic B^{act} complex into the catalytically active, distinctly sedimenting B* complex and results in low levels of first-step splicing[35]. To investigate the role of Prp2 in pre-mRNA remodeling during this step, we incubated the B^{act} complex assembled on the slide surface with Prp2, Spp2 and 2 mM ATP (henceforth referred to as B* complex condition). B* complex conditions resulted in a significant shift in the FRET histogram toward a new ~45% population with a mean FRET value of 0.71±0.01, diminishing the lone 0.33±0.01 FRET distribution observed for the B^{act} complex (Figure 2.9). In contrast to the predominantly static L2 state of the B^{act} complex, molecules under B* conditions show dynamic excursions to high-FRET states, indicating that the 5'SS and BP can now reach the close proximity required for first-step chemistry. More specifically, the B* condition comprises the L2, M and H FRET states, where the H state is accessed from either the L2 or M states. TODP plots show that only ~11% of molecules retain the stable L2 state characteristic of the B^{act} complex, while ~39% of molecules exhibit at least one L2→H transition. Notably, transitions into the H state are short-lived in a majority of molecules. However, 12% of molecules show a stable high FRET state, as expected if they transitioned through the first step of splicing, in accord with the ~13% of first-step splicing we observe under B* conditions (Figure 2.1).

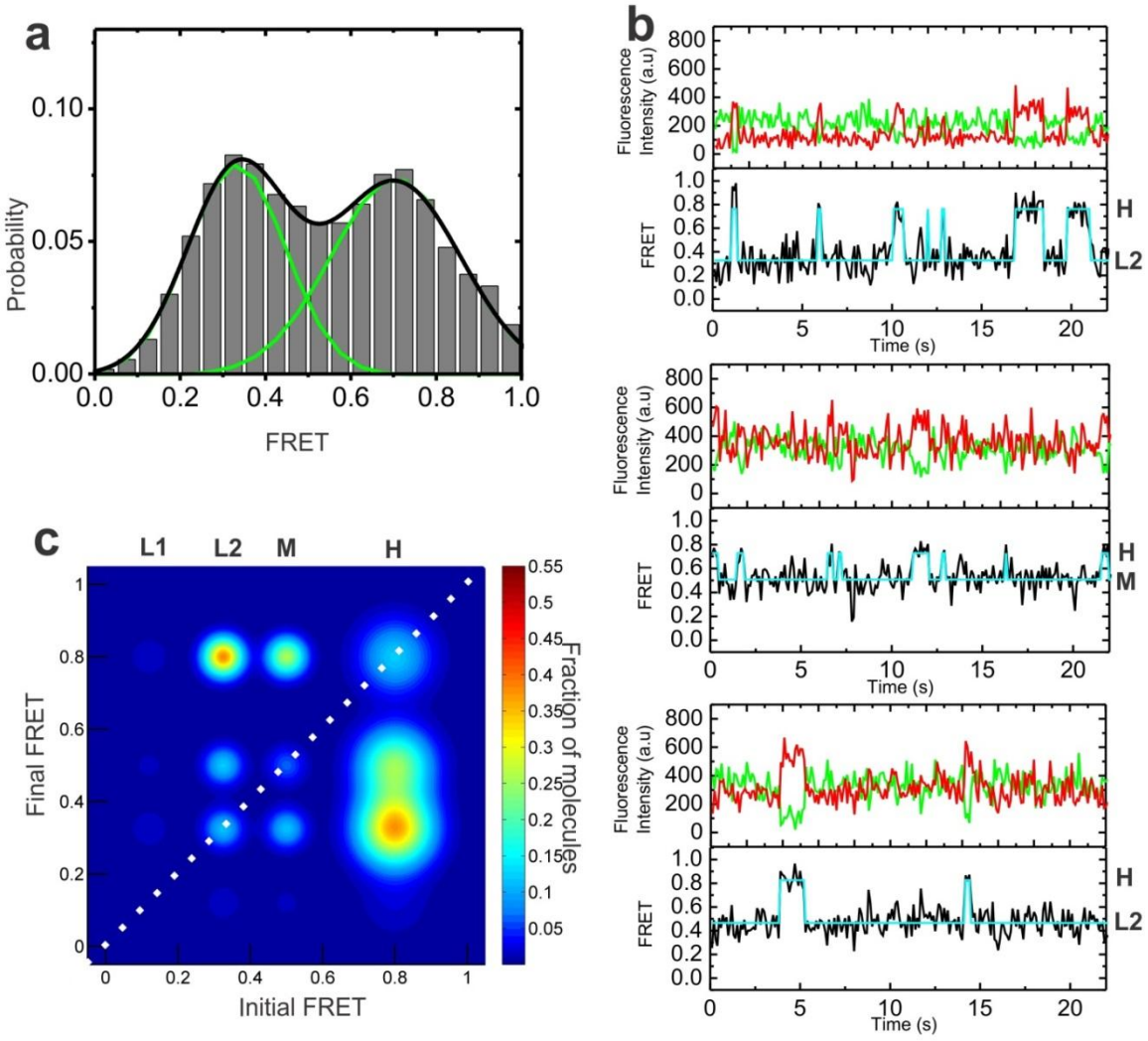


Figure 2.9 Upon the addition of ATP, Prp2 and Spp2, the pre-mRNA is able to explore splice site proximity.

(a) FRET histograms show higher FRET state sampling upon the addition of ATP and Prp2/Spp2 to the Bact complex, leading to formation of the B* complex. (b) Representative single molecule FRET time trajectories from this condition show that molecules that were in the stable L2 state prior to the addition of Prp2 now transition transiently to the H state from either the L2 state (top trajectory) or the M state (bottom two trajectories). (c) TODPs indicate that 11% of all molecules retain Bact-like behavior (stable L2 state), while a significant population explores higher FRET states. Approximately 12% of the molecules populate a stable H state, found along the diagonal.

Prp2 can directly bind a region in the pre-mRNA downstream of the BP adenosine, even in the absence of ATP[58,59]. To ask whether Prp2 alone can induce the pre-mRNA remodeling observed here by SiMPull-FRET, we omitted either Spp2 or Prp2 from our B* conditions and found the resulting FRET histograms to be indistinguishable from those of the starting B^{act} complex (Figure 2.10). Next, we studied the role of ATP in the remodeling of the B^{act} complex. For most spliceosomal DExD/H-box helicases, both ATP dependent and independent roles have been proposed[20,60-63]. Prp2 in particular has been shown to cause significant conformational remodeling of the spliceosome in the absence of ATP, whereas the displacement of SF3b is ATP dependent[31]. To ask whether the pre-mRNA remodeling observed here requires ATP, the B^{act} complex was incubated with Prp2 and Spp2 in the absence of ATP, leading again to no significant change in the FRET histogram (Figure 2.10). Similarly, when the non-hydrolysable ATP analog AMPPNP was used instead of ATP, the FRET histogram replicated that of the B^{act} complex, showing that ATP hydrolysis is required (Figure 2.10). Finally, the DExD/H box helicases involved in spliceosomal reorganization are either integral components of snRNPs or extrinsic components like Prp2. Mass spectrometry studies have shown that the stalled B^{act} complex contains stoichiometric amounts of the DExD/H-box helicase Brr2, the integral component of the U5 snRNP responsible for U4:U6 unwinding[64]. Direct interactions between Prp2 and the C-terminus of Brr2 have recently been discovered[59], suggesting a possible role for Brr2 in first-step catalytic activation. To rule out that Brr2 is responsible for the observed pre-mRNA remodeling, we exploited the fact that Brr2 is a strict ATPase[29] whereas Prp2 is a broad NTPase[65], and supplemented the B^{act} complex under B* conditions with UTP instead of ATP. The resulting FRET efficiency histogram is clearly distinct from that of B^{act} and overlays well

with that of the ATP-mediated B* condition with just a slightly less efficient shift towards the higher FRET population (Figure 2.10). This lower efficiency is consistent with the slightly (by ~2-fold) reduced activity of Prp2 in the presence of NTPs other than ATP[65]. Collectively, these results indicate that the NTP-driven helicase activity of Prp2 in complex with its activator Spp2 causes a large structural reorganization of the pre-mRNA that allows the distal 5'SS and BP of the B^{act} complex to reversibly access proximal conformations, which in turn enable first-step splicing.

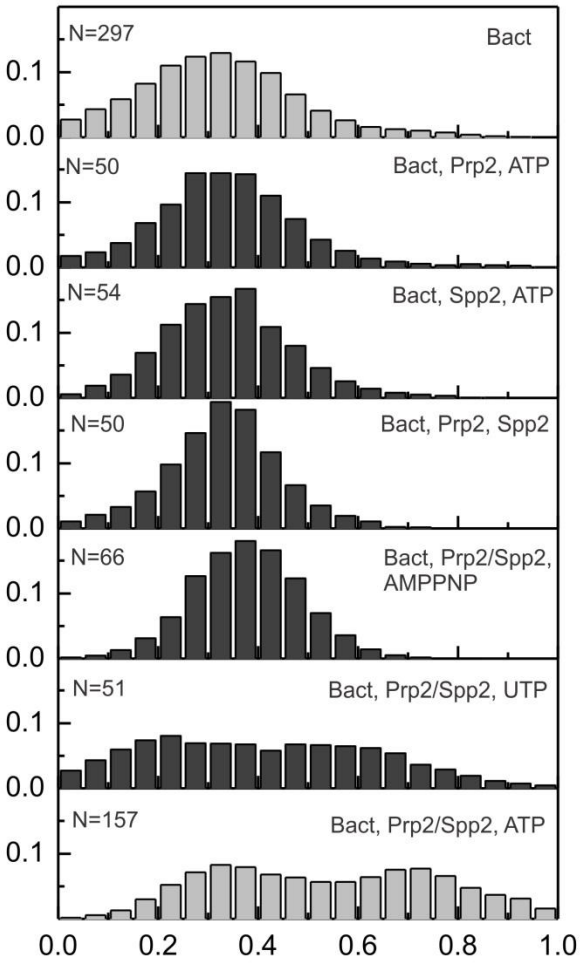


Figure 2.10 The NTPase activity of Prp2 induces structural remodeling of the pre-mRNA.

FRET probability densities show that the characteristic distribution of the B^{act} complex is retained when B^{act} is incubated with either Spp2, Prp2, or ATP missing, or when incubated with Prp2/Spp2 and non-hydrolysable ATP analog AMPPNP, as indicated. By contrast, the addition of UTP or ATP shifts the distribution notably towards a high-FRET population characteristic of the B* complex (bottom two panels).

2.3.4 Cwc25 enhances the first step of splicing by increasing the residence time in the H state

Although the pre-mRNA is remodeled by the NTPase action of Prp2/Spp2, it does not undergo very efficient first-step catalysis. To further enhance first-step splicing efficiency, addition of Cwc25 to the B^{act} complex incubated with Prp2, Spp2 and ATP is necessary. Cwc25 was identified as one of a group of proteins complexed with Cef1/Ntc85 of the NTC complex. It is a thermostable 20-kDa protein with short coiled coil domains[66]. To determine its role in remodeling of the pre-mRNA, we performed SiMPull-FRET on the purified B^{act} complex supplemented with Prp2, Spp2, ATP and Cwc25 (henceforth referred to as C complex conditions). This resulted in a FRET histogram with an enhanced ~73% population with a mean FRET value of 0.75±0.01 (Figure 2.11). We found the FRET states under C complex conditions to be the same as those under B* conditions, with the L2→H and M→H transitions prevalent; however, the occupancy in the H state is significantly enhanced under C conditions. TODP analysis revealed the fraction of molecules displaying at least one L2→H and M→H transition to be similar under B* and C conditions, whereas the stable H state occupancy is ~3-fold increased. This shift is similar in magnitude to the enhancement in first-step splicing that Cwc25 induces (Figure 2.1), consistent with the stable H state representing the C complex after the first step of splicing. Post-synchronized histograms (PSHs) created by aligning the HMM-fitted traces to start at the M state show that the molecules under C conditions both transition more frequently to the H state and exhibit a higher residence time once in the H state (Figure 2.12). A similar comparison of transitions starting at the H state further emphasizes the stabilization of this state by Cwc25 under C conditions (Figure 2.12). To rule out that a change in photostabilty of molecules in the C complex affects the relative prevalence of the H state, we analyzed the

average photobleaching time under B^{act}, B* and C conditions and found them to be comparable (Table 2.3). To quantitatively characterize the effects of Cwc25 on the conversion of the B* to the C complex, we plotted the cumulative dwell times for the forward and backward L2→H and M→H transitions under both conditions and fit them with double-exponential functions (Figure 2.13). A comparison of the weighted average rate constant for the L2→H transition showed similar forward and backward rate constants under both conditions, yielding equivalent equilibrium constants $K_{eq} = k_{forward}/k_{backward}$ of ~0.80. In contrast, the presence of Cwc25 accelerates the forward and reduces the backward rate constant of the observed M→H transition, leading to a K_{eq} that is ~3-fold more favorable for the H state under C compared to B* conditions. To show that the same pre-mRNA molecule can be chased from B* to C complex, we observed the same field of view before and after shifting from B* conditions (excluding Cwc25) to C conditions (including Cwc25) and incubating for 10 min in the dark. Before the dark period, molecules were dynamically shuttling between the L2 and H states. A subset of molecules were observable after the dark period and of those ~50% shifted to the stable H state (Figure 2.13 and Table 2.4). Taken together, our results suggest that Cwc25 acts kinetically to stabilize the catalytically favorable conformation, thereby effecting an enhancement of the first chemical step of splicing.

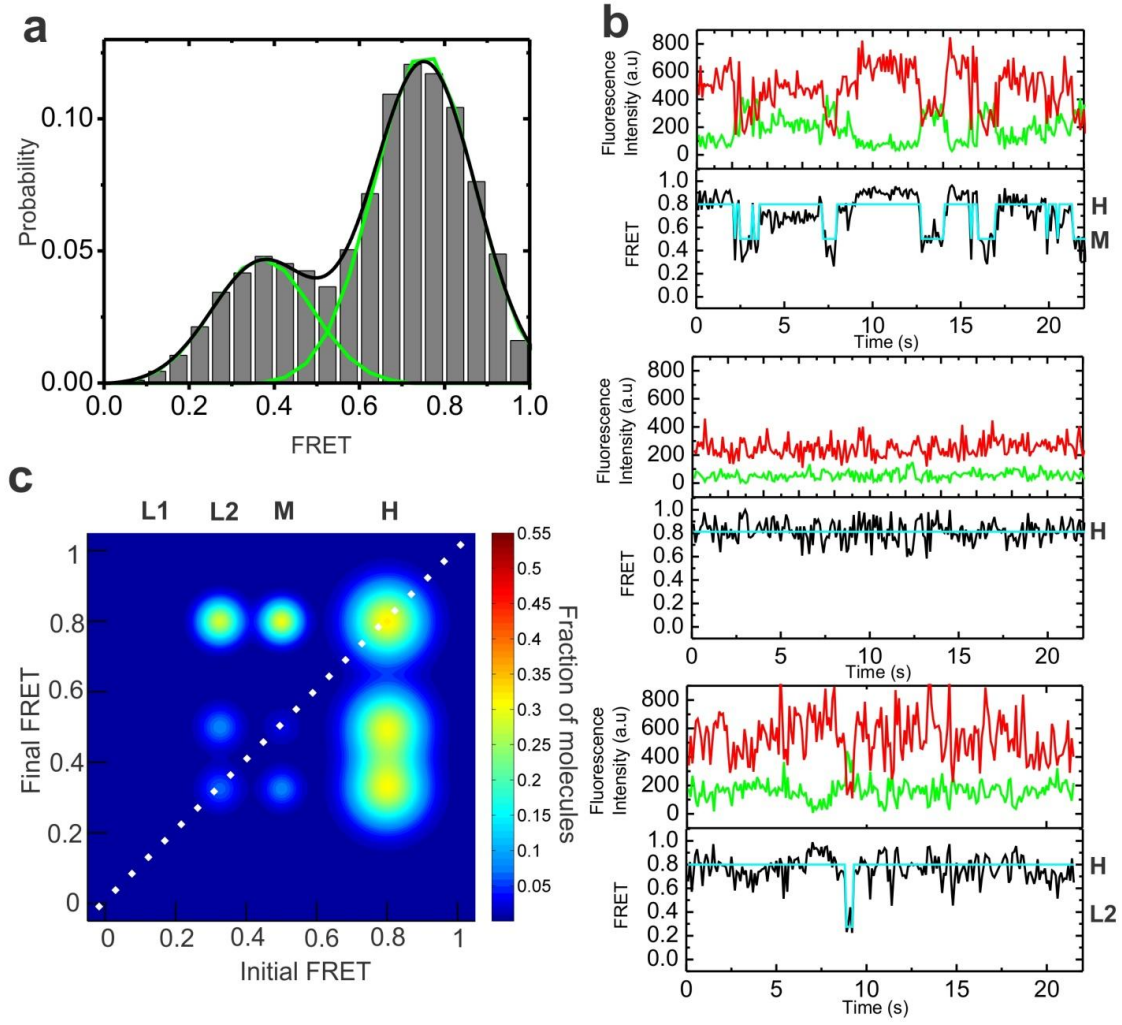


Figure 2.11 Under C complex conditions, the pre-mRNA accesses dynamic and stable high-FRET states.

(a) FRET probability densities plotted for molecules under C complex conditions where the addition of Cwc25 results in an enhanced shift to the high-FRET population. (b) Representative single molecule FRET trajectories from this condition show either M→H transitions (top panel), a stable H state (mid panel) or L2→H transitions (bottom panel), resulting in an overall increased residence time of the H state. (c) TODPs under this condition show that about 35% of the molecules exhibit a stable H state (so that they lie on the diagonal), which is 3-fold higher than under B* conditions, whereas the fraction of molecules showing L2→H and M→H transitions are similar.

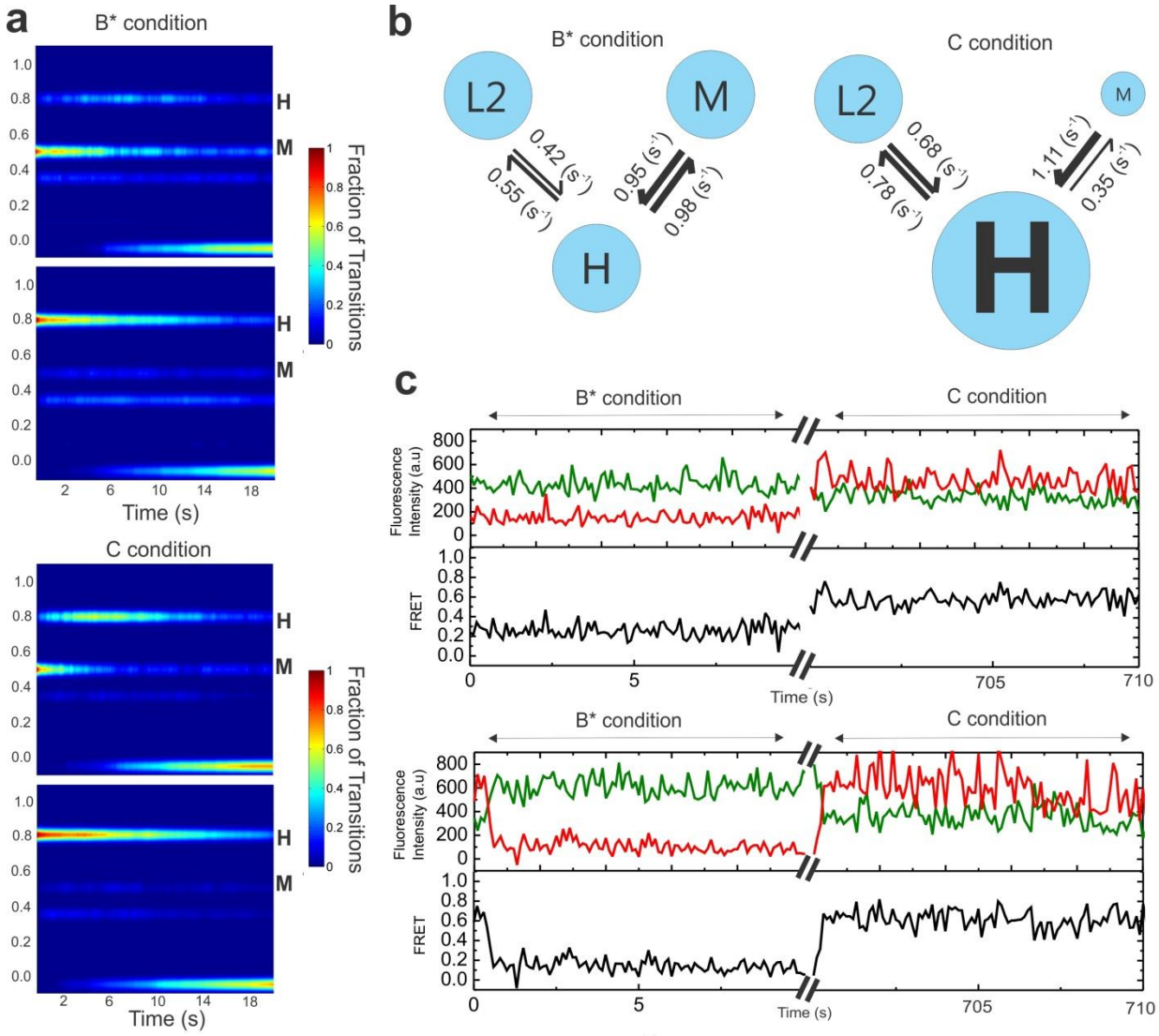


Figure 2.12 Cwc25 enhances the first step of splicing by stabilizing the H state.

(a) A comparison of the aggregate molecular behavior before (B*) and after Cwc25 addition (C) through post-synchronized histograms (PSH) with all trajectories synchronized to start from either the M (top) or H state (bottom) reveals a higher occupancy of the H state under C complex conditions. (b) (b) A comparison of the rate constants of the observable transitions under B* and C conditions reveals a ~3-fold increase in the stability of the dynamic H state for the C complex. The thickness of arrows corresponds to the relative rate constants. (c) Representative trajectories with transition dynamics from the same molecules imaged before (B*) and after Cwc25 (C) addition. Representative single molecule FRET trajectories with transition dynamics from the same molecules imaged before (B*) and after Cwc25 (C) addition. The axis breaks represent 10

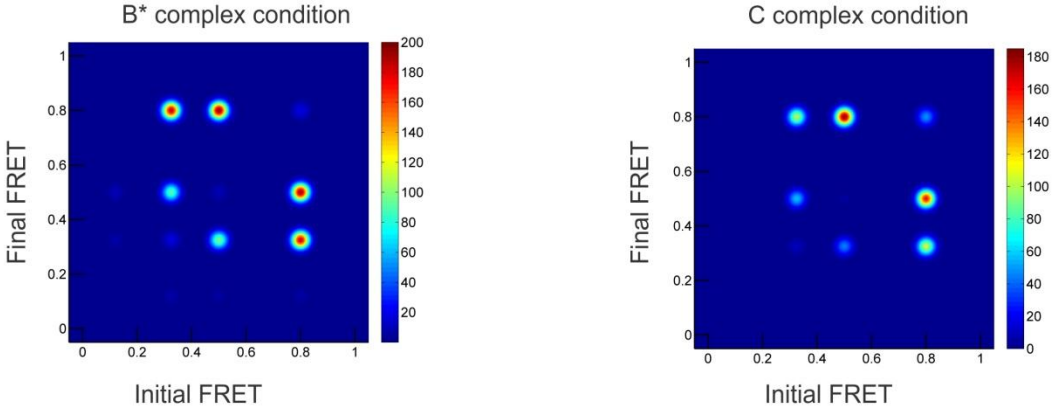
min of incubation after Cwc25 addition. The data confirm that individual molecules change their conformational dynamics upon Cwc25 addition.

Condition	Number of molecules	Average photobleaching time (Seconds)
B^{act} complex	297	11.8±7.0
B* complex	157	15.9±10.9
C complex	154	16.5±13.5

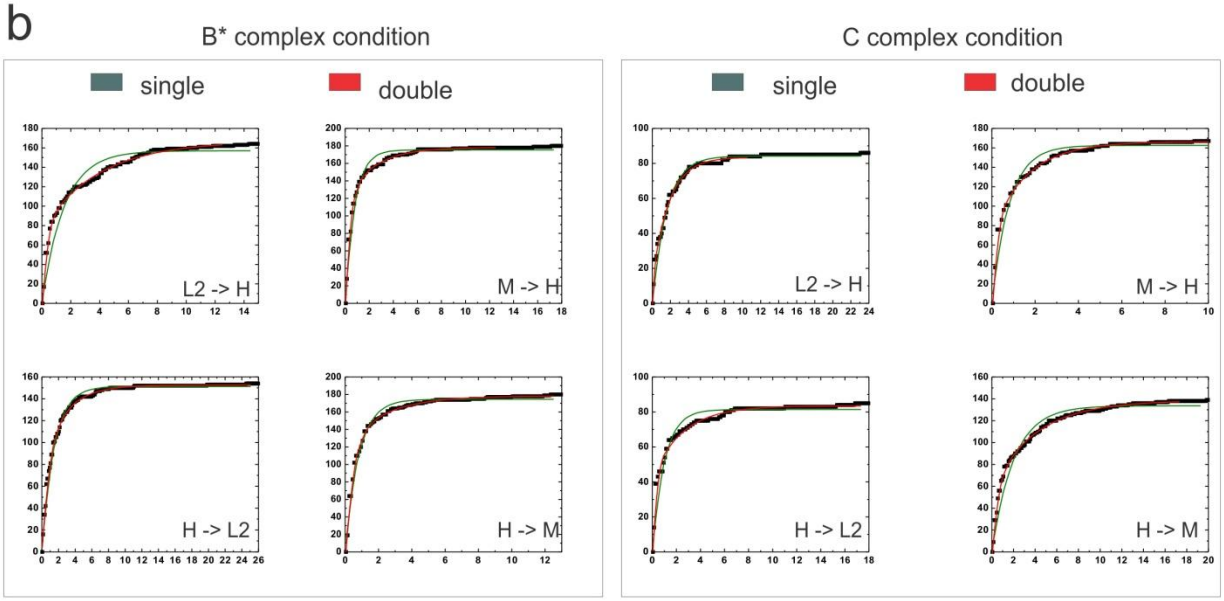
Table 2.3 Comparison of average photobleaching times and number of molecules per condition.

The average photobleaching times for the B* and C complex conditions are nearly identical. The increase in stable high FRET states in the C complex is thus not a result of a bias in the single molecule lifetimes under these conditions.

a



b



c

	L2 -> H		H -> L2		M -> H		H -> M	
Parameter	B*	C	B*	C	B*	C	B*	C
A1	90.50 +/- 1.32	70.53 +/- 1.10	65.60 +/- 4.30	40.58 +/- 1.23	61.00 +/- 3.03	91.65 +/- 2.08	54.90 +/- 4.14	80.33 +/- 1.18
t_1(s)	3.92 +/- 0.14	1.84 +/- 0.05	2.82 +/- 0.11	2.54 +/- 0.10	2.33 +/- 0.13	1.65 +/- 0.05	2.40 +/- 0.18	4.09 +/- 0.09
A2	94.40 +/- 2.75	27.76 +/- 5.61	89.70 +/- 4.01	55.51 +/- 4.01	135.84 +/- 3.24	105.75 +/- 3.40	140.41 +/- 4.10	68.46 +/- 1.59
t_2(s)	0.29 +/- 0.02	0.08 +/- 0.03	0.78 +/- 0.04	0.21 +/- 0.02	0.35 +/- 0.02	0.16 +/- 0.01	0.43 +/- 0.02	0.45 +/- 0.02
t_w(s)	2.07	1.34	1.64	1.19	0.96	0.85	0.98	2.42
$k_{w,observed}$ (s^{-1})	0.48	0.74	0.61	0.84	1.04	1.17	1.01	0.41
$k_{w,actual}$ (s^{-1})	0.42	0.68	0.55	0.78	0.98	1.11	0.95	0.35

Figure 2.13 Kinetic rate determination for B* and C complex transitions.

(a) Transition Density Plots (TDPs) for the B* and C complex conditions scaled to the number of transitions determined by HMM. The most prominent transitions in each condition are L2→H, H→L2, M→H, and H→M. (b) Dwells times were extracted for each corresponding transition, plotted as a cumulative count plot, and fit with either a single- or double-exponential rate equation. The double-exponential function best approximates the data and was chosen for further analysis. (c) Parameters for the double-exponential equations fitted to the dwell time data. To reduce the dimensionality of the data, a weighted average rate constant k_w was calculated by utilizing the amplitudes associated with each time constant as weighting factors. k_w was used for K_{eq} calculations and rate comparisons between B* and C complex conditions.

Fraction	Pre-Cwc25 addition	Post-Cwc25 addition	Classification
0.11	High FRET with long dwell	High FRET with long dwell	Catalysis of 1st step
0.39	Various behaviors (excluding high FRET with long dwell)	High FRET with long dwell	Cwc25 mediated enhancement
0.33	Fast Switching	Slow/No switch (no stable high FRET)	Prp2 unbound
0.08	Fast Switching	Fast Switching	Cwc25 not bound
0.09	N/A	Elevated Noise	N/A

Table 2.4 Classification of molecules from the observation of the same molecule chased from the B* to the C complex with the inclusion of a dark period during Cwc25 addition.

2.3.5 Cwc25 dynamically interacts with the pre-mRNA close to the BP after Prp2-mediated pre-mRNA remodeling

Previous studies have shown that Cwc25 binds stably to the spliceosome following Prp2 mediated SF3a/b destabilization[43]. It seems likely that Cwc25 enhances first-step chemistry by binding to the pre-mRNA since mutation at the BP abolishes this interaction[32]. Additionally, Cwc25 has been recently shown to crosslink near the BP of the pre-mRNA[32,67]. To directly observe the binding of Cwc25 to the pre-mRNA we labeled the protein's C-terminus with Cy5 fluorophore. The Cy5 near the 5'SS of the pre-mRNA was pre-bleached so that the pre-mRNA had a single fluorescent Cy3 label near the BP. We tested the activity of the Cy5-tagged Cwc25 using our bead pull-down assay and found it to be fully functional. SiMPull-FRET experiments were then carried out with Cwc25-Cy5 added to the B^{act} complex with and without the addition of Prp2, Spp2 and ATP (Figure 2.14). We observed repeated binding and dissociation of Cwc25-Cy5 and resulting FRET with the BP under B* condition. By contrast, there was little observable FRET between Cwc25 and the pre-mRNA BP in the absence of Prp2/Spp2 and ATP. From the FRET distribution of these binding events, centered around 0.37±0.03, we estimate that Cwc25 binds within FRET distance of the Cy3-Cy5 pair (<100 Å), roughly ~52 Å from the BP Adenosine. We conclude that Cwc25 activates the spliceosome for the first step by dynamically binding to the pre-mRNA near the branch point.

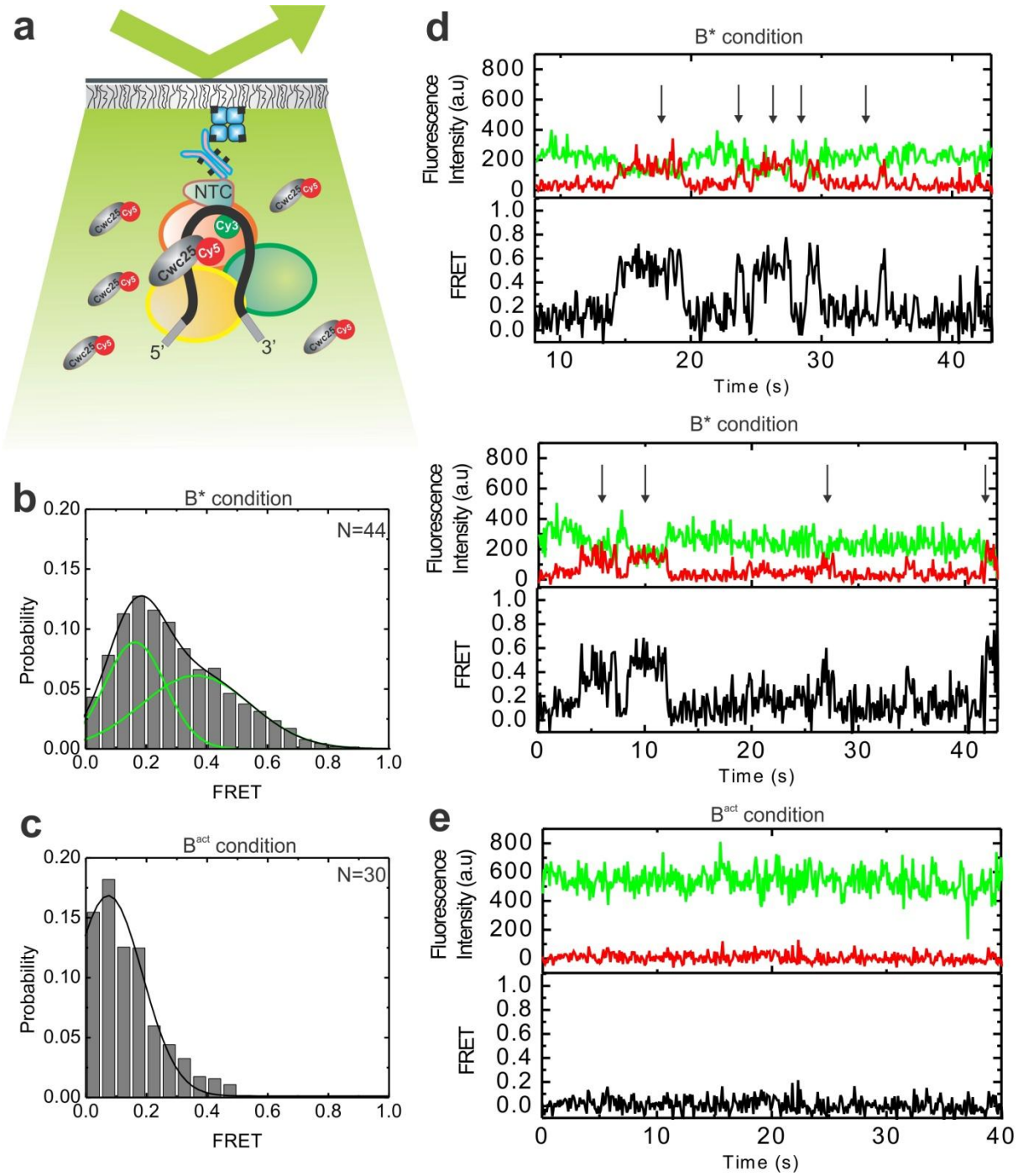


Figure 2.14 Prp2 mediated pre-mRNA remodeling creates a binding site for Cwc25 near the BP.

(a) Schematic of our SiMPull experiment to test for binding of Cy5-tagged Cwc25 to the spliceosome assembled on pre-mRNA with a single active Cy3 fluorophore. (b) FRET

histogram under B* conditions (in the presence of Prp2-mediated remodeling) showing the average FRET value obtained from the experiment. The lower FRET peak centered at 0.15 is the background associated with the system while the broader peak centered at 0.37 is the mean FRET between the fluorophores. (c) FRET histogram under B^{act} conditions (in the absence of Prp2-mediated remodeling) showing the low levels of background centered at 0.08. (d) Representative single molecule FRET trajectories showing the binding and associated FRET between the Cy5 on Cwc25 and the Cy3 near the pre-mRNA BP under B^{*} conditions. Arrows indicate binding events. (e) Representative smFRET trajectory showing the absence of FRET between the Cy5 on Cwc25 and the Cy3 near the pre-mRNA BP under B^{act} conditions.

2.4 Discussion

Here we have combined single molecule FRET between fluorophores attached near the 5'SS and BP of the pre-mRNA substrate with affinity purification into a technique we term SiMPull[68]-FRET to study the spliceosomal B^{act} complex stalled by heat inactivation of Prp2. Stepwise addition of ATP and the recombinant proteins Prp2, Spp2 and Cwc25, which together are required for efficient first-step splicing and conversion of the B^{act} into first the B* then the C complex, revealed the role of each factor in pre-mRNA remodeling as summarized in Figure 2.15. We find that the pre-mRNA remains in the stable low-FRET L2 state of the B^{act} complex (which for clarity we term $L2^{act}$) with distal 5'SS and BP until the activation by Prp2/Spp2 in the presence of ATP produces the B* complex. Prp2-mediated hydrolysis of ATP (or UTP) in this step weakens the binding of some seven proteins, including SF3a and SF3b[31,35,42,43] that bind the pre-mRNA upstream and downstream of the BP adenosine, presumably preventing its premature nucleophilic attack on the 5'SS. Accordingly, the B*-associated low-FRET state (L2*) allows the pre-mRNA to transiently and reversibly visit both mid- (M*) and high-FRET (H*) states with more proximal 5'SS and BP. First-step splicing now proceeds with low efficiency, leading to post-catalytic C complex formation signified by a stable high-FRET state (H^{C}). This finding indicates that the increased proximity of the reactive sites is sufficient for catalysis. However, reaction chemistry is significantly enhanced by the addition of Cwc25, which binds the pre-mRNA substrate near the BP and slows particularly the rate constant of the high- to mid-FRET transition, leading to a longer dwell time in the pre-catalytic transient-FRET state $H^{C\text{-}pre}$. In turn,

this event leads to enhanced occurrence of the stable high-FRET state associated with the post-catalytic C complex (H^C) (Figure 2.15).

Our data show that before Prp2/Spp2/ATP action, the spliceosome keeps the reactive sites of the pre-mRNA strictly apart. This observation is consistent with and refines a recent report that suggested that stable splice site juxtaposition occurs some time after the NTC assembles on the pre-mRNA[24]. Furthermore, it has previously been speculated that the catalytically activated B* complex, formed after Prp2/Spp2/ATP but before Cwc25 action, may shift back and forth between inactive and active conformations[35]. We here have presented direct evidence for this notion by showing that only in the B* state dynamic excursions between low- and high-FRET states become observable, and only the high-FRET state places the reactive 5'SS and BP in close enough proximity for subsequent catalysis, correlated with the appearance of the stable high-FRET state H^C. The same authors also proposed that Cwc25 binding may shift the equilibrium between inactive and active conformations towards the latter, which we directly observe and assign to a marked increase of the dwell time in the active conformation $H^{C\text{-}pre}$ with proximal 5'SS and BP. This behavior strikingly resembles that of a classical biased Brownian ratchet machine that draws path directionality from the stochastic thermal fluctuations to which it is constantly subjected, through a form of "rectification" or "biasing"[69,70]. In fact, the ribosome has been described as a biased Brownian ratchet machine[71-73] and our previous single molecule FRET probing of pre-mRNA dynamics in whole yeast cell extract suggested that the spliceosome, like the ribosome, works close to thermal equilibrium[37]. We therefore propose that the ATP-driven helicase activity of Prp2/Spp2 acts to remove SF3a/SF3b as an impediment to the intrinsic thermal fluctuations of the spliceosome-substrate complex, while Cwc25 provides

directionality to the reaction pathway by then acting as a "pawl" to stabilize the catalytically competent conformation.

DExD/H-box helicases such as Prp2 are widespread enzymes that participate in many aspects of RNA processing[21,74]. In general, they are thought to use ATP hydrolysis to remodel RNA and RNP complexes by binding, unwinding and releasing the RNA. Unlike all previous single molecule approaches[24,37,75-77], our SiMPull-FRET approach allowed us to directly interrogate a specific intermediate of spliceosomal assembly and catalysis and discover the specific mechanism by which the helicase activity of Prp2/Spp2 is augmented by Cwc25. Such biased Brownian ratcheting may be widespread among helicase-driven RNPs, and SiMPull-FRET will allow us to test this hypothesis further.

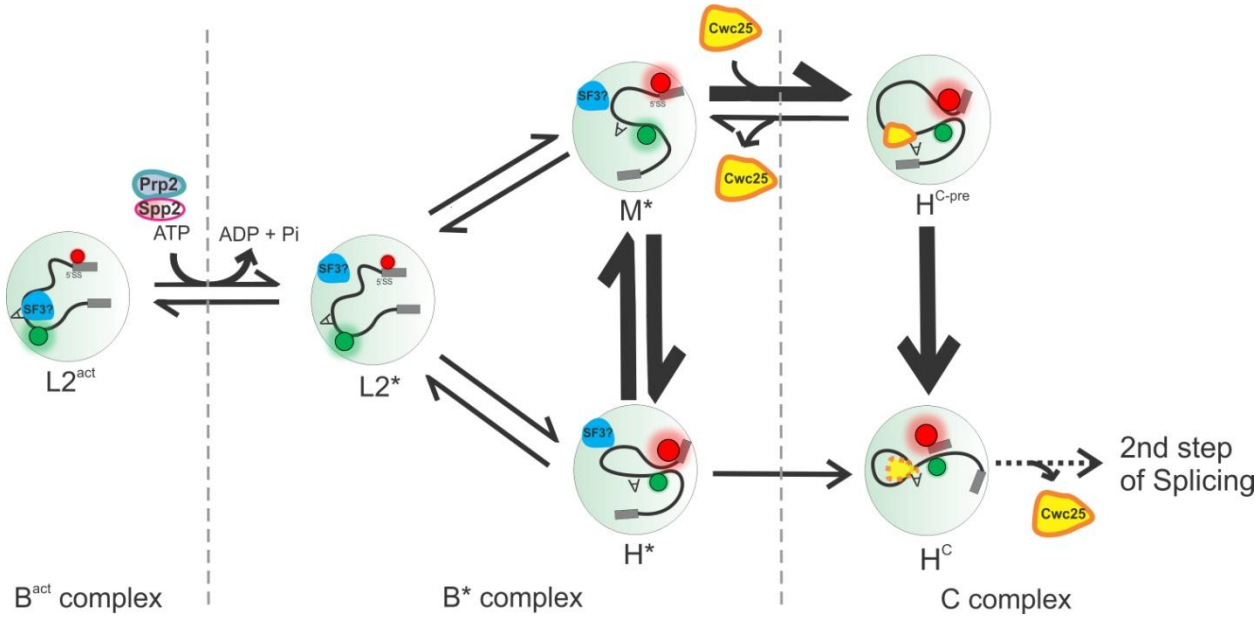


Figure 2.15 Model for the conformational mechanism of first-step splicing.

The 5'SS and BP of the pre-mRNA in the B^{act} complex are predominantly in the stable distal $L2^{act}$ conformation with low FRET (i.e., high Cy3 and low Cy5 fluorescence, indicated by the green and red circles, respectively). Upon ATP hydrolysis and conversion into the B* complex, Prp2 along with its cofactor Spp2 unlocks the B*-associated low-FRET state L2* to reversibly sample the mid- and high-FRET (spatially proximal) conformations M* and H* (and $H^{C\text{-}pre}$ under C complex conditions). Cwc25 binds near the BP of the pre-mRNA, thus reducing the rate constant of the high- to mid-FRET transition and favoring first-step chemistry, upon which the pre-mRNA adopts the stable high-FRET state H^{C}.

2.5 Acknowledgements

The authors wish to thank R. Lührmann (Max Planck Institute for Biophysical Chemistry, Göttingen, Germany) and R. J. Lin (Fujian Medical University, Fuzhou, China) for generously supplying expression plasmids for Spp2, Cwc25 and Prp2; H. Hadjivassiliou and A. Price (University of California, San Francisco) for providing Cy5-body labeled actin pre-mRNA substrate.

CHAPTER 3 : Single molecule dissection of splice site juxtaposition along the pre-mRNA splicing cycle [2]

3.1 Introduction

Excision of introns from pre-mRNA transcripts is carried out at single nucleotide resolution by a large and elaborate macromolecular RNP machine called spliceosome. The spliceosome lacks a preformed catalytic core and is assembled in a stepwise fashion from the U1, U2, U4, U5 and U6 small nuclear ribonucleoprotein particles (snRNPs). Each of the snRNPs consists of a specific small nuclear RNA (snRNA) and associated proteins. Assembly is carefully orchestrated in a stepwise manner on each pre-mRNA template, resulting in the formation of a mature mRNA in a two-step transesterification process. In the first catalytic step of splicing, the 2'OH of the branchpoint adenosine (BP) attacks the phosphodiester bond at the 5' splice site (5'SS), creating a lariat structure and a free 5' exon. In the second step, the 3' hydroxyl of this exon attacks the phosphodiester bond at the 3' splice site (3'SS), releasing the lariat intron and mRNA product.

[2]Adapted from: Ramya Krishnan*, Blanco, Mario*, Joshua Martin, Matthew Kahlscheuer, John Abelson, Christine Guthrie, Alain Laederach and Nils Walter; "Single molecule dissection of splice site juxtaposition along the pre-mRNA splicing cycle" (manuscript in preparation). Single molecule experiments and biochemical characterizations of the various blocks were carried out by Ramya Krishnan. Single molecule data pre-filtering was jointly done by Ramya Krishnan and Matthew Kahlscheuer. Data clustering was jointly done by Mario Blanco and Joshua Martin. *Contributed equally to this work.

The conserved sequences and the secondary structure of the pre-mRNA serve as template on which the components of the spliceosome assemble in a dynamic and highly regulated manner. While the focus in the field has largely been on the protein components, advances in the last decade in single molecule fluorescence microscopy techniques have allowed us to track the pre-mRNA structural rearrangements with accuracy. Notably, these studies have identified that components of the spliceosome are in a state of constant flux, characterized by steady-state dynamics and reversibility. Our previous work showed that the pre-mRNA reversibly samples a variety of conformations during the ATP-dependent and -independent assembly steps in splicing extracts. The multitude of different conformations sampled stochastically in the absence of molecular synchronization did, however, limit our ability to assign dynamics to specific parts of the assembly cycle.

In an effort to reliably characterize pre-mRNA splice site dynamics throughout the splicing cycle and integrate it with the mechanistic information available, we here have employed a combination of genetics to stall the spliceosome at distinct stages, single molecule microscopy and hierarchical clustering tools to detail the 5'SS-BP conformational dynamics explored by the pre-mRNA throughout splicing. Using a hierarchical clustering approach to extract the consensus behaviors of pre-mRNAs throughout the splicing cycle, we find that the splice sites are dynamic and explore dynamic conformational proximity with the early assembly restricted to predominantly low- to mid-FRET (high- to mid-distance) states. Using these snippets of FRET dynamics of each splicing stall, we have recreated the entire pre-mRNA conformational consensus pathway in the splicing cycle. The approach presented here is widely applicable to macromolecular systems with complex sets of dynamics.

3.2 Materials and Methods

3.2.1 Synthesis of Ubc4 pre-mRNA variants

The Ubc4 pre-mRNA variants used in this study were synthesized chemically and labeled with the FRET donor Cy3 +6 nucleotides downstream from the BP adenosine and with the FRET acceptor Cy5 -7 nucleotides upstream from the 5'SS via allyl-amine modified uridines. A DNA splint (dSplint) was used for templated ligation to synthesize the pre-mRNA as described[3,4]. The 5'- end of the pre-mRNA is biotinylated and used for direct immobilization on the microscope surface. Sequences for the mutants are given here (mutated sites are bolded and highlighted):

WildtypeUbc4:

5'GAACUAAGUGAUCUAGAAAGGUAUGUCUAAAGUUAUGGCCACGUUUCAAAUGC
GUGCUUUUUUUUUAAAACUUAUGCUCUUAUUUACUAACAAAAUCAACAUGCUAU
UGAACUAGAGAUCCACCUACUUCAUGUUT3'

5'Hyperstable mutant Ubc4:

5'GAACUAAGUGAUCAGAAAGGUA**A**GU**A**U**U**A**U**GUUAUGGCCACGUUUCAAAUGCG
UGCUUUUUUUUUAAAACUUAUGCUCUUAUUUACUAACAAAAUUCAACAUGCUAU
UGAACUAGAGAUCCACCUACUUCAUGUUT3'

Branch point mutant Ubc4:

5'GAACUAAGUGAUCUAGAAAGGUAUGUCUAAAGUUAUGGCCACGUUUCAAAUGC
GUGCUUUUUUUUUAAAACUUAUGCUCUUAUUUACUA**C**CAAAAUCAACAUGCUAU
UGAACUAGAGAUCCACCUACUUCAUGUUT3'

3’Splicesite mutant Ubc4:

5’GAACUAAGUGAUCUAGAAAGGUAUGUCUAAAGUUAUGGCCACGUUUCAAAUGC

GUGCUUUUUUUUUAAAACUUAUGCUCUUAUUUACUAACAAAAUCAACAUGCUAU

UGAACUA**C**A**C**AUCCACCUACUUCAUGUUT3’

3.2.2 Yeast extracts and modifications

Wildtype yeast splicing extracts (Snu66-TAP background) were prepared as previously described[37]. U2 and U6 snRNAs were ablated as described with the following modifications[78,79]. DNA oligo SRU2 was used at a concentration of 400nM to achieve efficient depletion of U2 snRNA. In vitro transcribed U2 snRNA at 50 nM concentration was then incubated with the depleted extract to achieve U2 reconstitution and efficient splicing. DNA oligo D1 complementary to U6 snRNA was used at 300 mM concentration to achieve efficient depletion of U6 snRNA from the extract. In vitro transcribed U6 snRNA at 10 nM concentration was sufficient to reconstitute active splicing. Prp2-1 heat sensitive extract was incubated for 40 min at 37 ºC to inactivate the Prp2 protein. Each of the splicing stalls and reconstitution of activity were verified with ensemble splicing assays after each single molecule experiment.

3.2.3 Single molecule measurements

Single molecule measurements were carried out as described[37]. The Ubc4 pre-mRNA used in this study were biotinylated at the 5’-end and attached directly via streptavidin-biotin linkage onto a quartz slide. Single molecules were selected for further analysis after verification that they had each a single donor (Cy3) and acceptor (Cy5) as evident from photobleaching analysis or direct excitation of the Cy5 dye with a 635 nm laser. Histograms were constructed by

sampling the first 10 seconds of each trajectory. Hidden Markov Model (HMM) analysis was carried out by applying a global model to the entire data set of molecules utilizing the vbFRET79 program. The number of states allowed ranged from 1-10 and vbFRET was allowed to choose the final number of states, found to be 7 states. The global fit ensures that the same idealized FRET values are sought in all data sets, allowing for a direct comparison of state abundance and state-to-state transitions.

3.2.4 Creating an HMM similarity matrix and Clustering Analysis

The FRET data for all conditions examined were fit with a global HMM to determine and assign idealized states. The HMM analysis was performed with vbFRET, allowing for the number of states to range from 1-10. The VbFRET algorithm determined that a total of 7 states exist in the system. Transition probability (TP) matrices were created for each individual molecule by using the path of idealized states determined by vbFRET and the hmmgenerate function in MATLAB. When a specific molecule did not sample a particular state that state's position on the TP matrix was assigned a zero to make sure that all TP matrices have the same size (7 states x 7 states). The FRET Similarity Matrix (FSM) combined the TP matrix with the fraction occupancy of a molecule at each state (1 x 7 matrix); making the FSM an overall 7 x 8 matrix. The FSM was then clustered using the hierarchical clustering toolbox in MATLAB. Before clustering, molecules sampling only one (static) state were separated and grouped independently because their low complexity resulted in the identification of 5 distinct static clusters. The pairwise distance of each molecule's FSM was calculated using the Euclidean metric. These distances were then joined using the linkage function that created a hierarchical

tree joined at nodes based on the median Euclidean distance between the trajectories. The hierarchical tree can be clustered by determining a height cutoff between pairs of nodes. We determined the appropriate cutoff by measuring the intra and inter cluster distances at several different thresholds. A threshold was chosen that resulted in 14 different clusters. These clusters were then joined with the 5 static clusters previously identified. Once these 19 clusters were identified, we could quantify the number of molecules present in each cluster under each experimental condition. To determine consensus behaviors for each cluster we calculated the mean TP matrix for each cluster and then determined the trajectory with the shortest Euclidean distance to that mean.

3.3 Results

3.3.1 Hierarchical clustering tools facilitate objective classification of complex smFRET data

Previous smFRET studies on splice site dynamics in Ubc4 substrate have shown that the dynamics are time- and ATP-dependent and largely reversible[37]. More recently, work on 5'SS-BS dynamics in RP51 pre-mRNA has shown that these sites are held farther apart from each other during assembly until the binding of the NTC complex, after which the splice sites explore close proximity[24]. We have used established genetic and biochemical tools to stall the progress of the splicing cycle at various steps (Figure 3.1) and observed 5'SS-BP dynamics in immobilized Ubc4 pre-mRNA in whole yeast cell extract (YE). These stalls were characterized biochemically via standard splicing assays by incubating either the wildtype/mutant pre-mRNA with wildtype/mutant YE, with and without ATP for 30-40 min at 23 °C to confirm their ability to reliably stall the cycle progress at the specific steps (Figure 3.2, Figure 3.3, Figure 3.4 and

Figure 3.5). The Ubc4 pre-mRNA used was 5'-biotinylated and labeled with a donor fluorophore (Cy3) at position +6 from the branchsite adenosine while maintaining the acceptor fluorophore (Cy5) at position -7 from the 5' splice site.

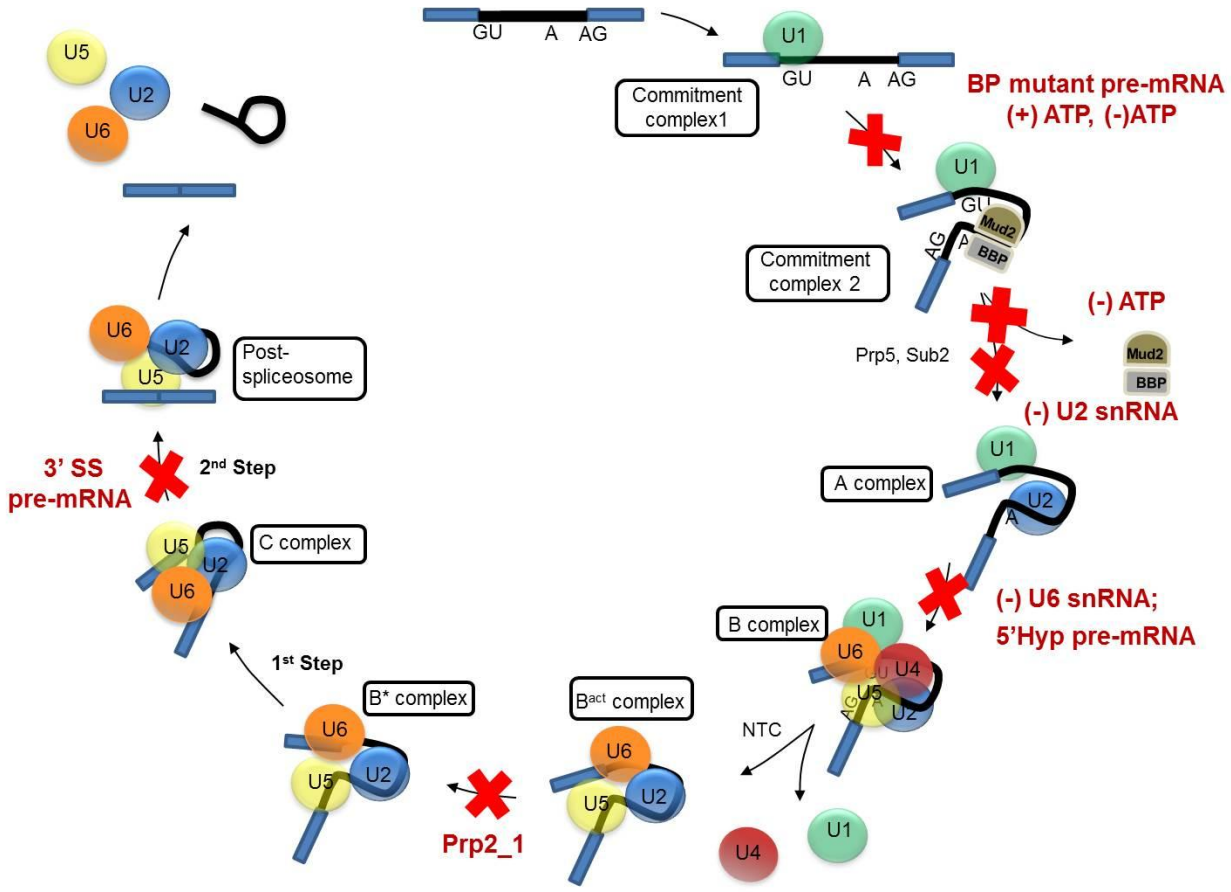


Figure 3.1 Stalling the progress of splicing at various steps

Schematic showing the assembly, activation, catalysis and disassembly of the spliceosome during the process of pre-mRNA splicing. Shown in red cross marks are the independent stalls employed in the study. The mutant pre-mRNAs (BP mutant, 5'Hyp mutant and 3'SS mutant) are used with wildtype extract while the other stalls have extract modifications used as described in methods.

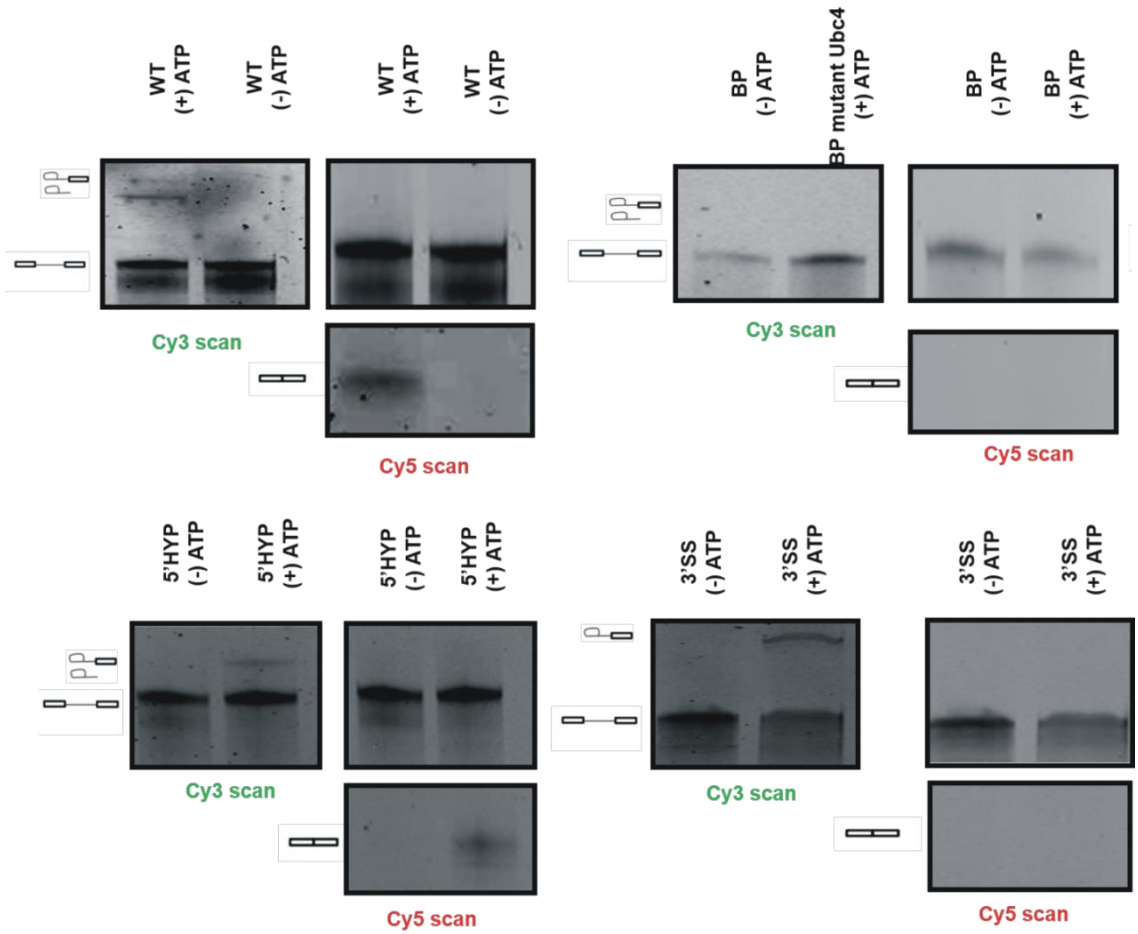


Figure 3.2 Mutant pre-mRNA substrates show expected stall in the splicing cycle

Wildtype (WT) and mutant (BP, 5'Hyp, 3'SS) Ubc4 pre-mRNA assembled in wildtype extract under splicing conditions without and with ATP ((-) and (+), respectively) at 23 °C for 60 min. RNA was extracted and the products were analyzed on a denaturing, 7 M urea, 15% polyacrylamide gel. A variable mode Typhoon scanner was used to obtain the Cy3 scan (left-green) and Cy5-scan (right-red). The intron and intron-lariat products are observed in the Cy3 scan (left) and the mature mRNA product is visualized in the Cy5 scan (right).

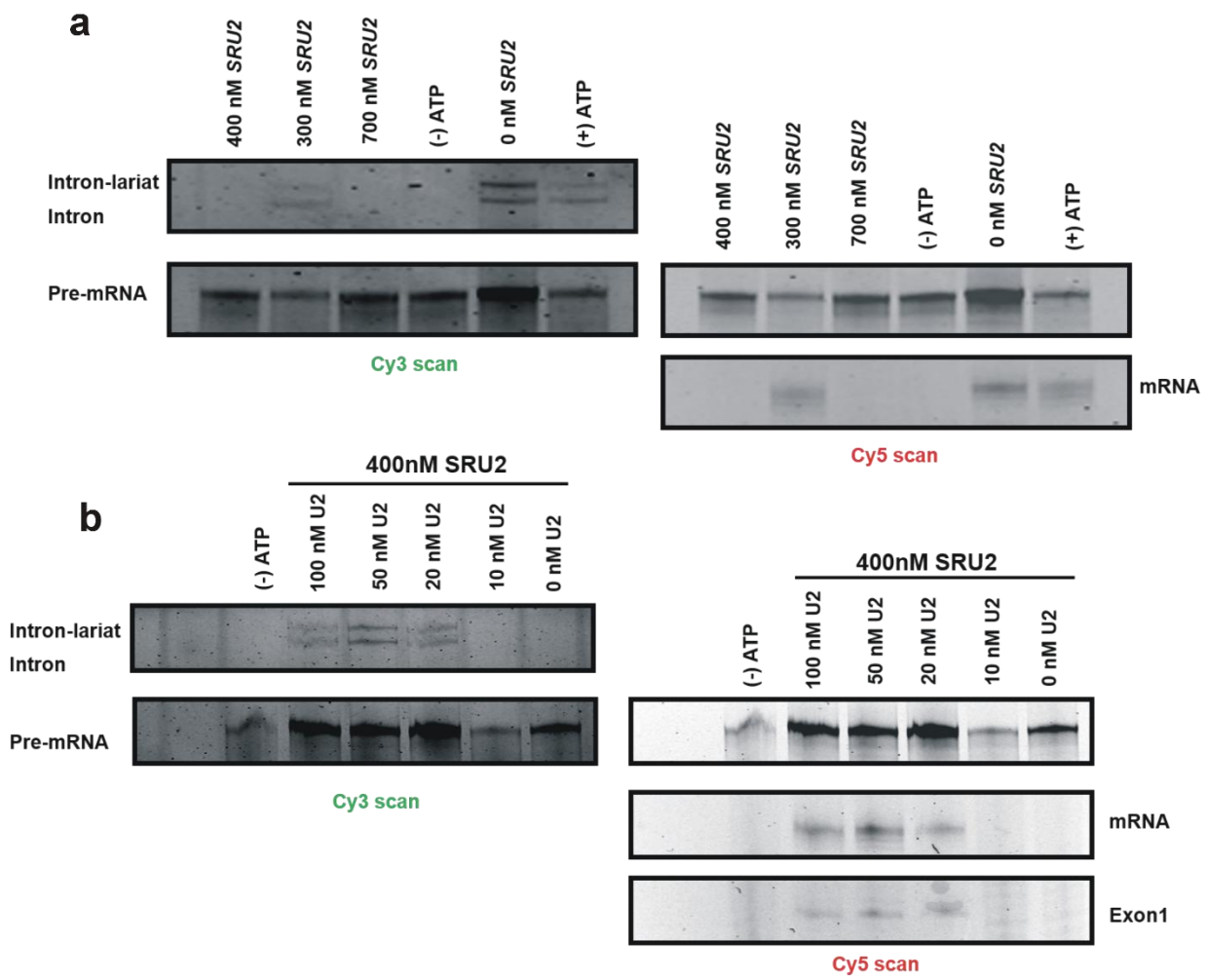


Figure 3.3 Depletion of U2 snRNA results in stalling of splicing cycle

DNA oligonucleotide SRU2 is complementary to the U2 snRNA and when incubated at a sufficiently high concentration can deplete WT extract of U2 snRNA. (a) SRU2 at 400 nM was sufficient to abolish all splicing activity of WT UBC4 pre-mRNA. (b) *In vitro* transcribed U2snRNA at 50 nM was able to efficiently reconstitute activity when supplemented to extract depleted of U2 (with 400 nM SRU2).

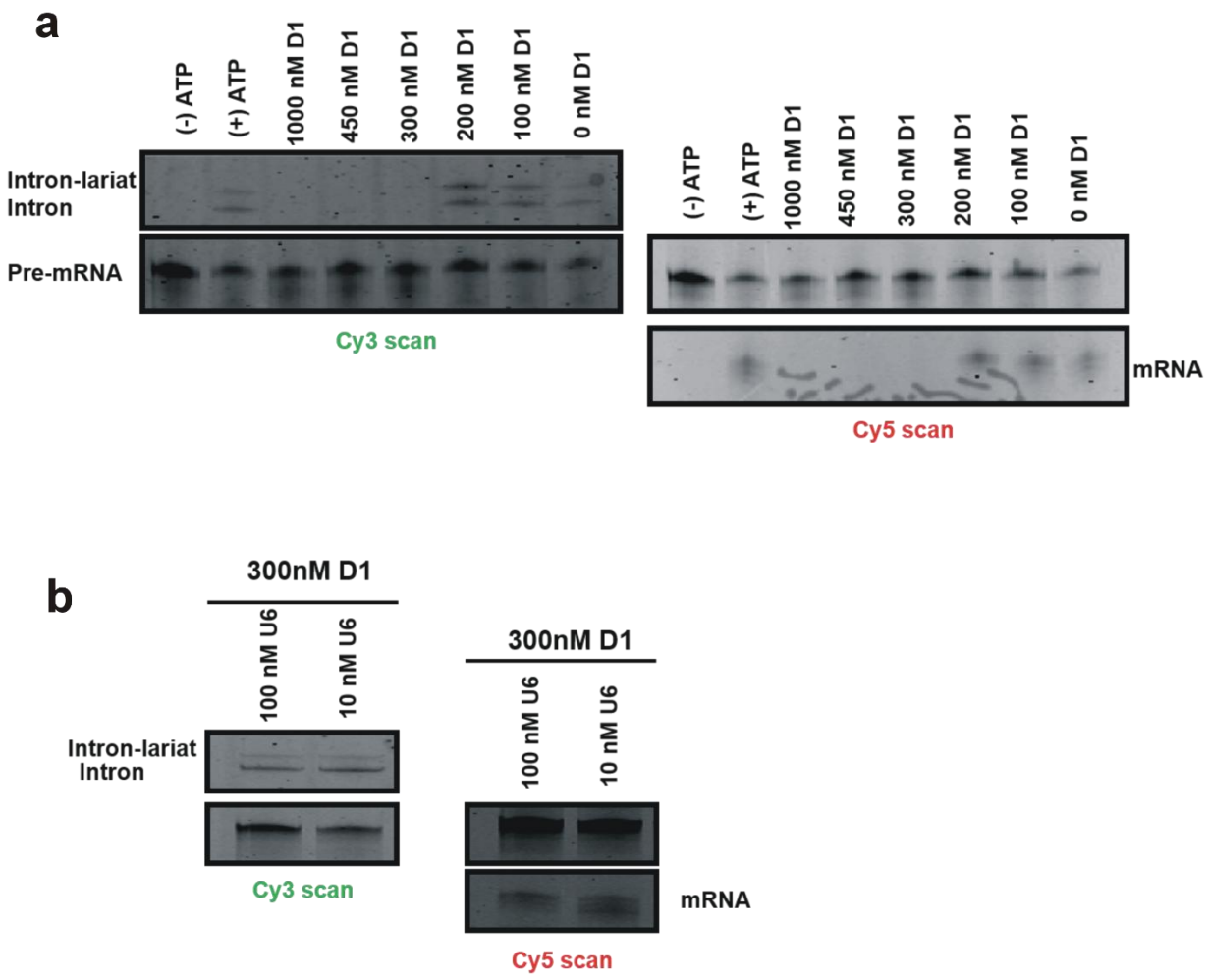


Figure 3.4 Depletion of U2 snRNA results in stalling of splicing cycle

DNA oligonucleotide D1 is complementary to the U6 snRNA and when incubated at a sufficiently high concentration can deplete WT extract of U6 snRNA. (a) D1 at 300 nM was sufficient to abolish all splicing activity of WT UBC4 pre-mRNA. (b) In vitro transcribed U6 snRNA at 10 nM was able to efficiently reconstitute activity when supplemented to extract depleted of U6 (with 300 nM D1).

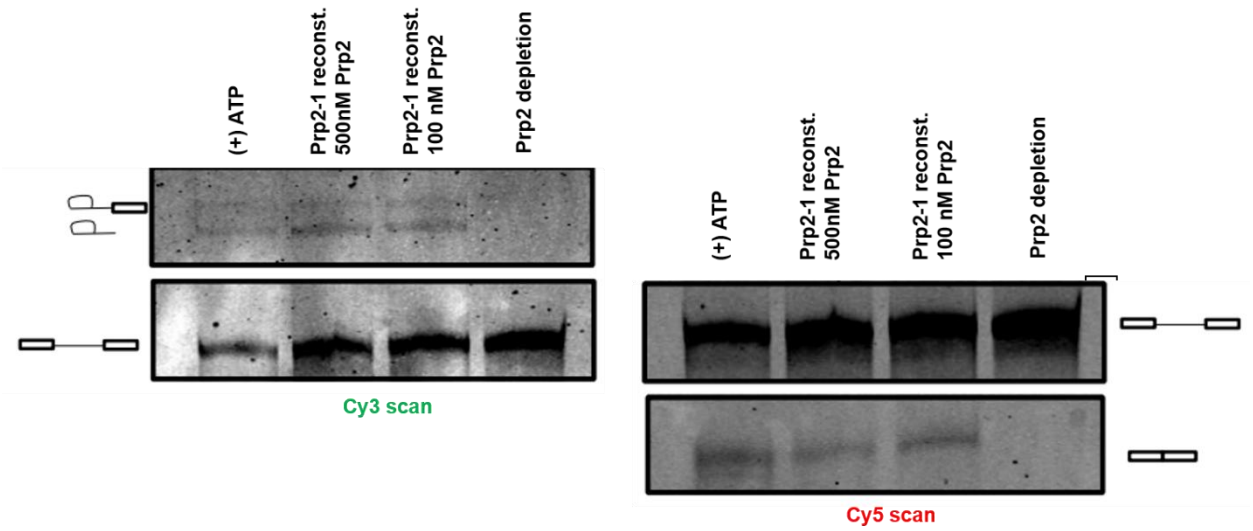


Figure 3.5 Heat-inactivated of Prp2 protein results in stalling of splicing cycle

Prp2-1 extract carries a temperature sensitive mutation in the RNA dependent ATPase Prp2 that can be heat inactivated at 37 ºC for 40 min. This blocks any catalysis but allows assembly to proceed until just before the first step of catalysis. This block can be relieved by supplementing heat inactivated extract with 100 nM recombinantly expressed Prp2 protein.

To obtain insights into pre-mRNA splice site juxtaposition up to each step where splicing was stalled, we performed smFRET studies on either WT or mutant pre-mRNA immobilized on quartz slides via streptavidin-biotin interaction. To this pre-mRNA, either the WT or modified yeast extract (YE) was flowed and the RNA imaged. To relieve a particular stall, we added back mutant extract reconstituted with the specific component and imaged the pre-mRNA. After initial quality control, FRET histograms were plotted using the FRET values collected from all verified molecules over the first 100 video frames (at 100 ms time resolution). As expected, we observed ATP-dependent changes in the WT pre-mRNA when incubated in YE (Figure 3.6 b and c). WT pre-mRNA incubated in standard splicing buffer exhibited a wide range of conformational dynamics with prevalence of high FRET states which are disfavored upon extract addition (Figure 3.6 a and b). The earliest stall in the splicing cycle was achieved by using a BP mutant pre-mRNA that showed a highly populated low FRET state irrespective of ATP (Figure 3.6 d). We similarly plotted probability density histograms for all the splicing stalls used in this study to achieve enrichment at either the CC1 complex, CC2 complex, A complex, B^{act} complex or C complex stages. Our previous work had established that the conformational behaviors of single pre-mRNAs undergoing splicing are complex with multiple interconverting states and a variety of kinetic behaviors[37]. The data obtained with the 5'SS-BP labeled substrate similarly exhibited intricate behaviors and therefore required sophisticated techniques to dissect.

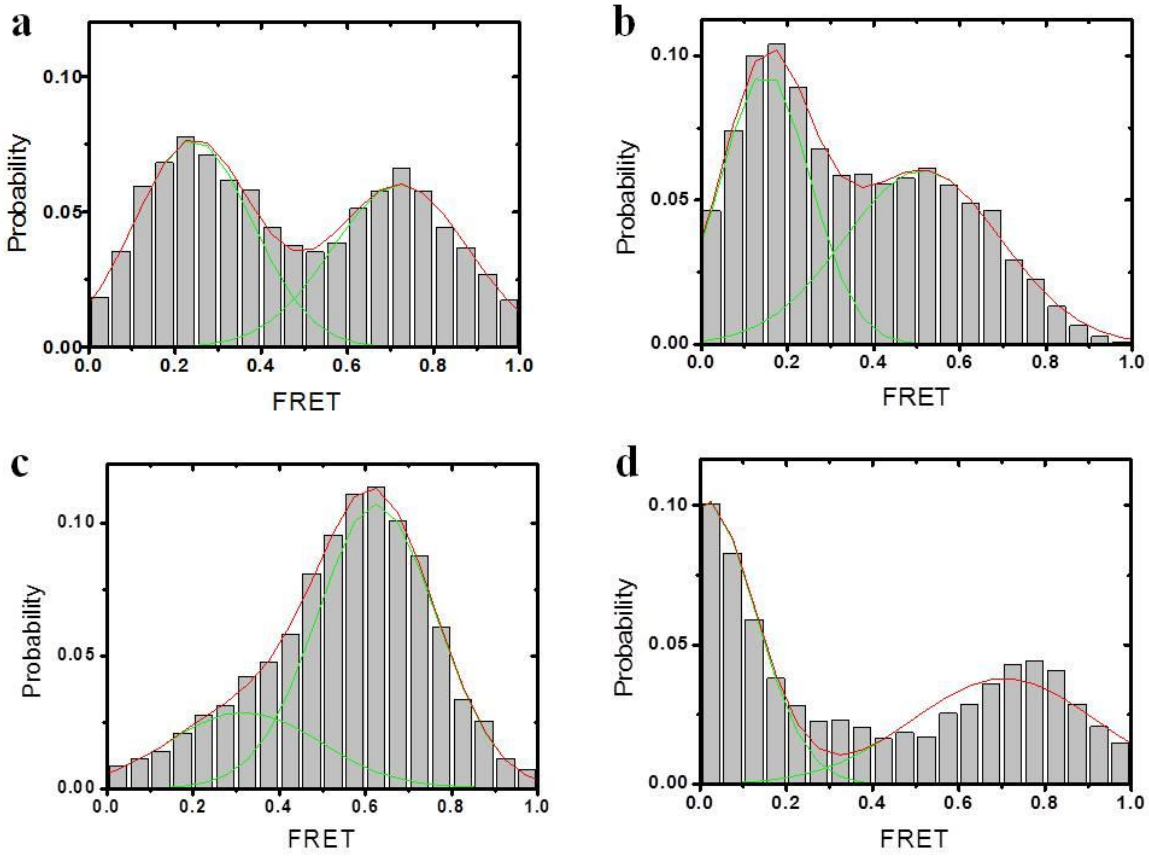


Figure 3.6 FRET probability histograms comparing early and late steps of splicing

FRET histograms were plotted using the FRET values collected from all the verified molecules over the first 100 frames showing WT Ubc4 pre-mRNA in (a) splicing buffer, (b) wildtype YE depleted of ATP and (c) wildtype YE with 2 mM ATP. (d) FRET histogram showing Branchpoint (BP) mutant pre-mRNA in YE in the absence of ATP.

The large amount (>10,000 single molecules) of complex data collected from these experiments required the use of novel analysis methods to objectively classify the similarities and differences between them. We created a FRET similarity matrix (FSM) by combining the traditional Transition Probability matrix (TP) and a histogram that describes the overall occupancy of states (Figure 3.7). A global Hidden Markov Model fit to all single molecule data collected yielded a total of seven discrete FRET states (7 x 7 TP matrix). We then created a FSM for each single molecule trajectory. A vector form FSM can now be used as input into clustering algorithms. The data sets collected for all the blocks were mixed and then a pairwise distance between all possible trajectories was calculated using the Euclidean distance metric. Static molecules (no transitions with occupancy in only one state) were then separated since their features are simple and can be treated separately. The Euclidean distances between FSMs for dynamic molecules were calculated and used to build a hierarchical cluster tree by utilizing the linkage function in Matlab (Figure 3.8). Molecules with short distances (high similarity) were grouped together and connected to other groups of molecules based on their average inter-cluster distance. A tree-pruning threshold was then selected based on the inter-cluster distance that minimized the intra-cluster distance (differences within a cluster) and maximized inter-cluster distance (differences between clusters) (Figure 3.9). The threshold yielded a total of 19 clusters,h 14 of which were dynamic and 5 static. Importantly, the data were not segregated *a priori* by the experimental conditions, thereby allowing molecules with similar dynamics across conditions to be grouped objectively. The uniqueness of the 19 clusters created from the hierarchical clustering

analysis were assessed by visually verifying that similar molecules were grouped together to create clusters of high homogeneity.

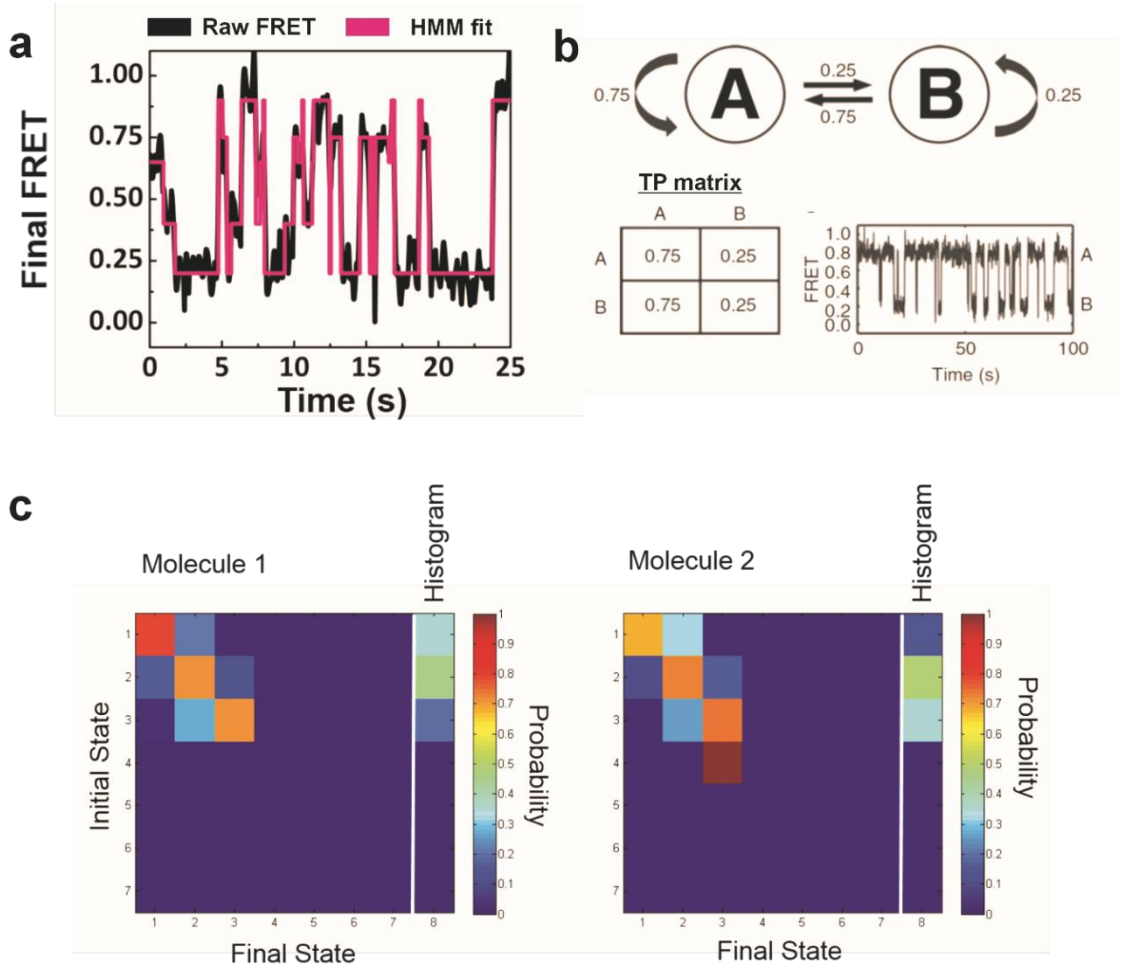


Figure 3.7 FSM: A Hidden Markov Modeling derived similarity matrix

(a) Hidden Markov Models (HMM) can objectively and accurately analyze single molecule FRET signals with complex behaviors. A complex single molecule trajectory with several underlying states is fit with an HMM algorithm that can identify change points within the trajectory. (b) A simplified two-state HMM with defined transition probabilities. Transition probabilities describe the likelihood of any given state either transiting to another state or remaining in the same state. The Transition Probability (TP) matrix

summarizes all the pairwise probabilities and can be used to describe the kinetic behavior of a single molecule. (c) A FRET similarity matrix (FSM) based on a TP matrix can be constructed for individual molecules and used as a clustering tool. The similarity matrix shown here is a 7 x 7 matrix (7 idealized states) based on the TP matrix calculated for each molecule and a final column (column 8) which is the distribution of data points at each state, i.e., a histogram. The FSM of two similar molecules, depicted as heat maps, are shown. A vectorized form of the FSM was used for clustering analysis.

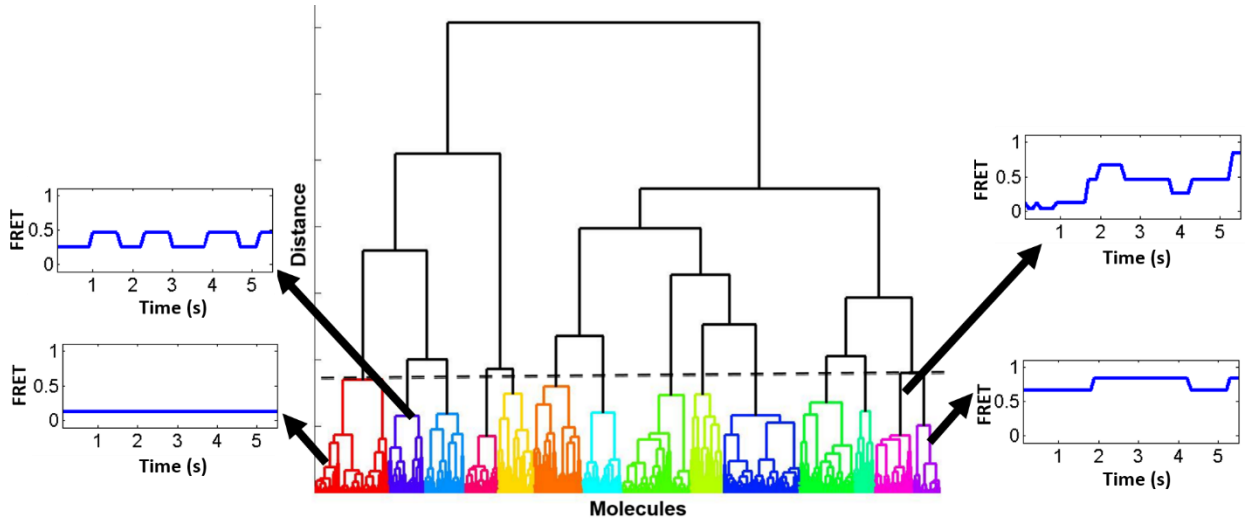


Figure 3.8 Hierarchical tree constructed from clustering single molecules

Single-molecule FRET data were collected for each substrate and extract and globally fit with a 7-state HMM. An HMM similarity matrix was then constructed for each molecule and clustered. Molecules with only 1 state (static) were clustered separately and later incorporated as separate clusters. Hierarchical clustering was performed utilizing Matlab routines by calculating the pairwise Euclidean distance of every single molecule and building an agglomerative hierarchical cluster tree utilizing the unweighted average distance to compute distances between clusters. This tree was pruned by using a distance criterion that resulted in the formation of 14 well separated clusters.

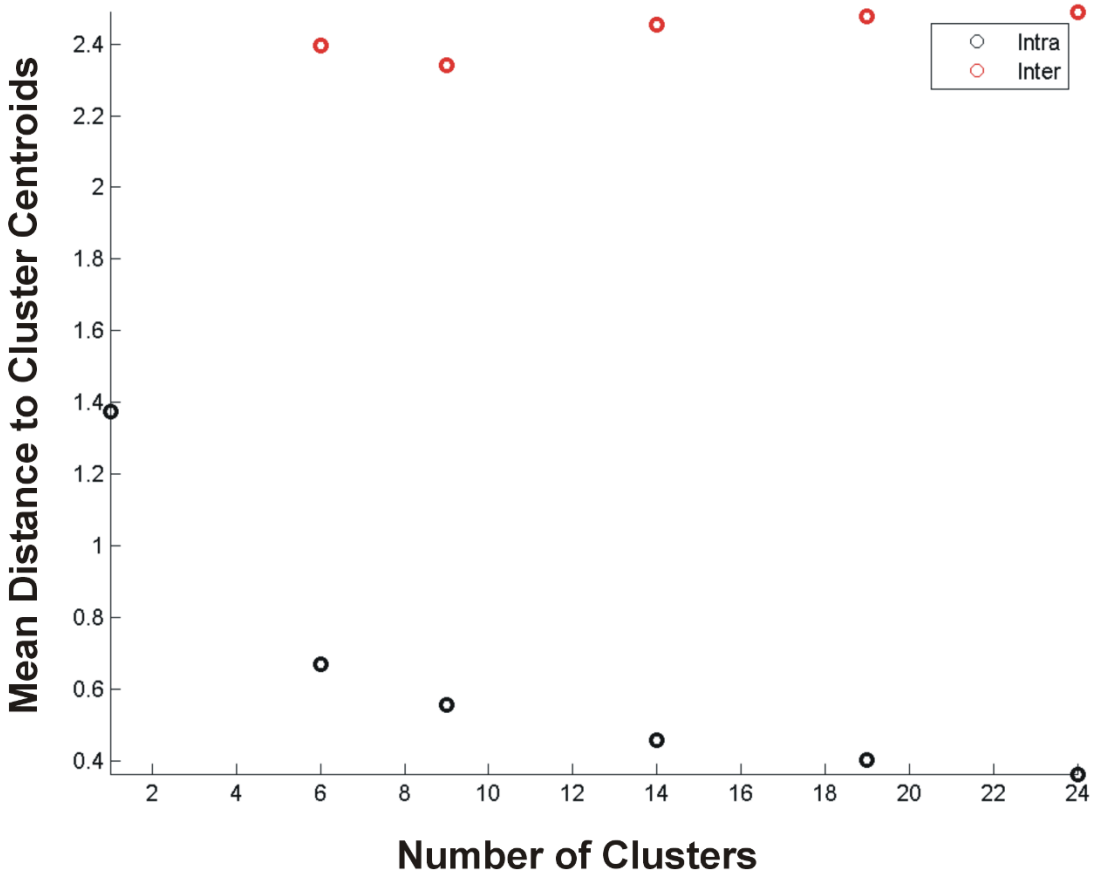


Figure 3.9 Choosing a hierarchical tree pruning threshold

The hierarchical tree was cut several times varying the node height criterion to yield a range of clusters. Intra- and inter-cluster distances were computed by measuring the mean Euclidean distance of the trajectories to the cluster centroids formed at the various cuts. 14 clusters were selected because there was minimal improvement in intra cluster distance beyond this point.

To perform this cluster analysis with minimal user introduced bias, the linearized form of the FSM (i.e., the 7 x 8 matrix converted to a 1 x 56 matrix) of each molecule was plotted in its assigned clusters as a heat map of the values at each position in the matrix. These heat maps were organized such that each row represents a single molecule, while the columns are the different values for each point in the FSM (Figure 3.10). We found that each cluster thus projected has a distinct set of values for the FSM but high homogeneity within the clusters. The HMM fit for 20 single molecules selected at random for each cluster was concatenated into a single plot with vertical bars separating individual molecules. The randomly selected molecules within clusters have similar dynamics and illustrate the behaviors that represent each cluster (Figure 3.11).

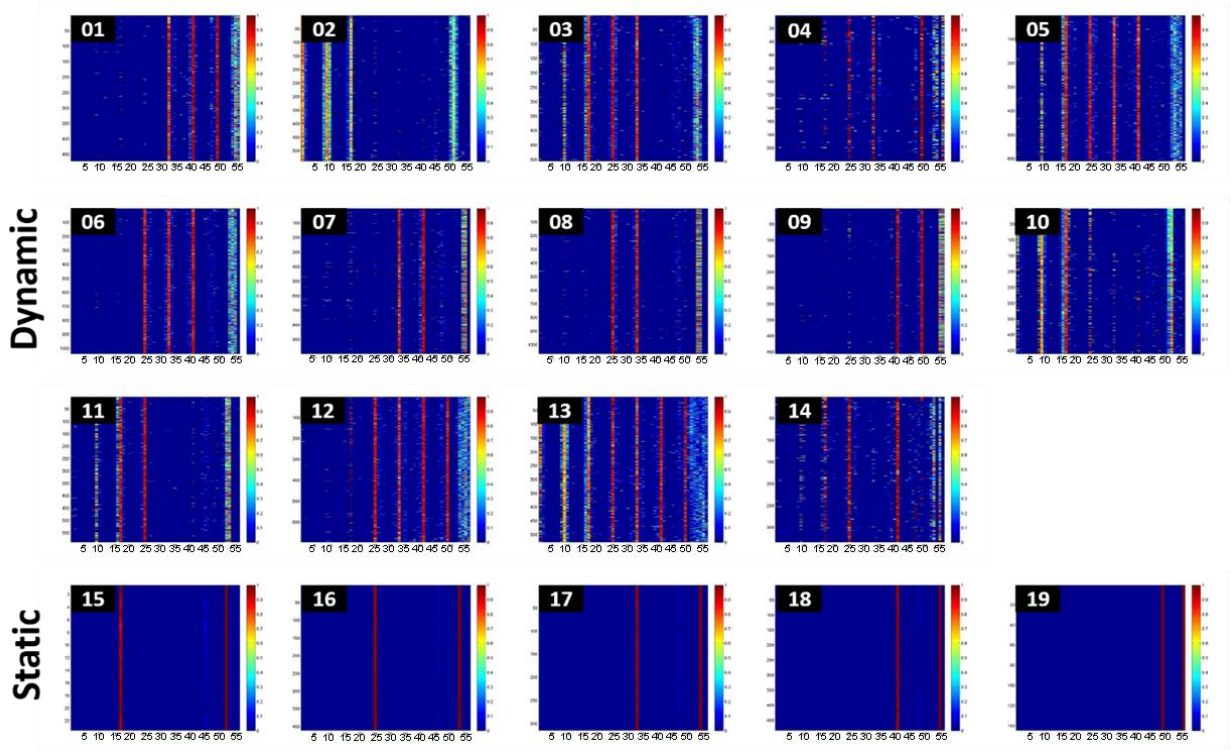


Figure 3.10 Vectorized FSM for each cluster

A total of 19 clusters were derived from the clustering of single molecule splicing data. 14-dynamic clusters, and 5-static clusters. Heat maps of the vectorized HMM similarity matrices for the molecules assigned to each cluster are represented here. These vectorized matrices (columns) represent the features that were used for clustering the individual molecules (rows) by the hierarchical clustering routine.

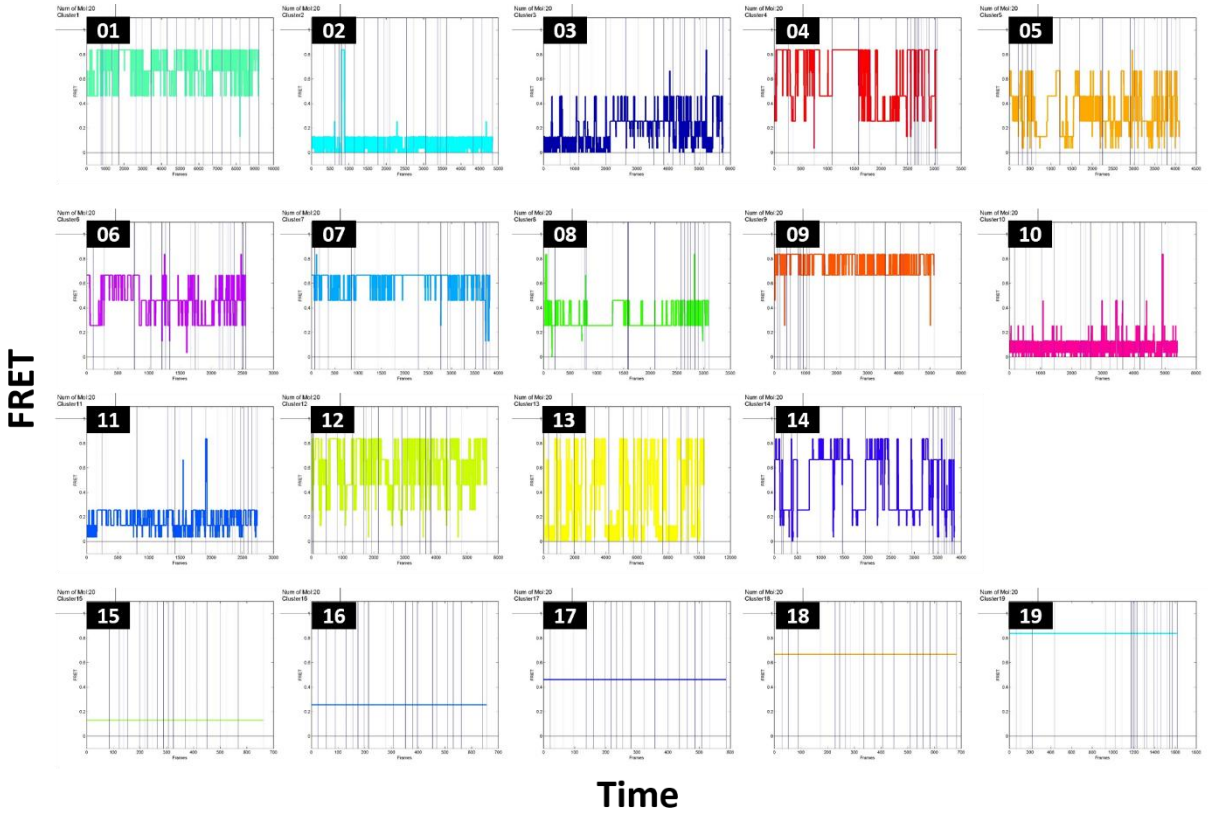


Figure 3.11 Representative molecules showing dynamic behavior unique to each cluster

Single molecules were randomly selected from each of the 19 clusters and concatenated into a single plot. Vertical bars serve to separate the individual molecules. The randomly selected molecules exemplify the typical kinetic behavior of molecules with the assigned cluster. As expected, the behaviors within the clusters are homogeneous and distinct from those in other clusters. Cluster assignment given in the inset.

3.3.2 Pre-mRNA in early splicing cycle shows dynamic variation in splice site proximity

The BP mutant substrate has an intact 5'SS that allows the U1 snRNP to bind. However, it blocks further assembly steps by preventing the binding of the Mud2/BBP complex at the branch site[44]. The FRET histogram of the BP mutant in splicing buffer shows that it can adopt high-FRET conformations in the absence of any splicing component. Upon addition of YE depleted of ATP, the substrate becomes constrained and mainly adopts a low-FRET distribution (Figure 3.6). The addition of ATP does not have a dramatic effect on the FRET distribution with the most noticeable change manifesting as a slight increase in the abundance of high-FRET states. Using the data obtained from the FSM clustering, we analyzed the abundance of clusters (organized by ascending mean FRET value from left to right) to identify the dynamics that lead to the differences between conditions. The most abundant dynamics in the BP buffer are rapid inter-conversions within the high FRET states (Figure 3.12). This shifts to dramatically lower FRET states populated by 55% of WT pre-mRNA molecules in YE depleted of ATP. The slight increase in the high-FRET population with the addition of ATP is due to a redistribution of molecules in cluster 2 to cluster 13, which exhibits the same low-FRET state but with a reversible excursion to the high FRET state (Figure 3.12). This systematic study allows us to link the extensive genetic knowledge available in the field to dynamic changes in the splice site proximity. Although the BP mutant pre-mRNA is recognized independent of ATP, genetic interactions show that the ATP-dependent DExD/H helicase Sub2p interacts early in spliceosome assembly[80]. Accordingly, our results imply that even at an early step, the pre-mRNA dynamics are influenced by ATP.

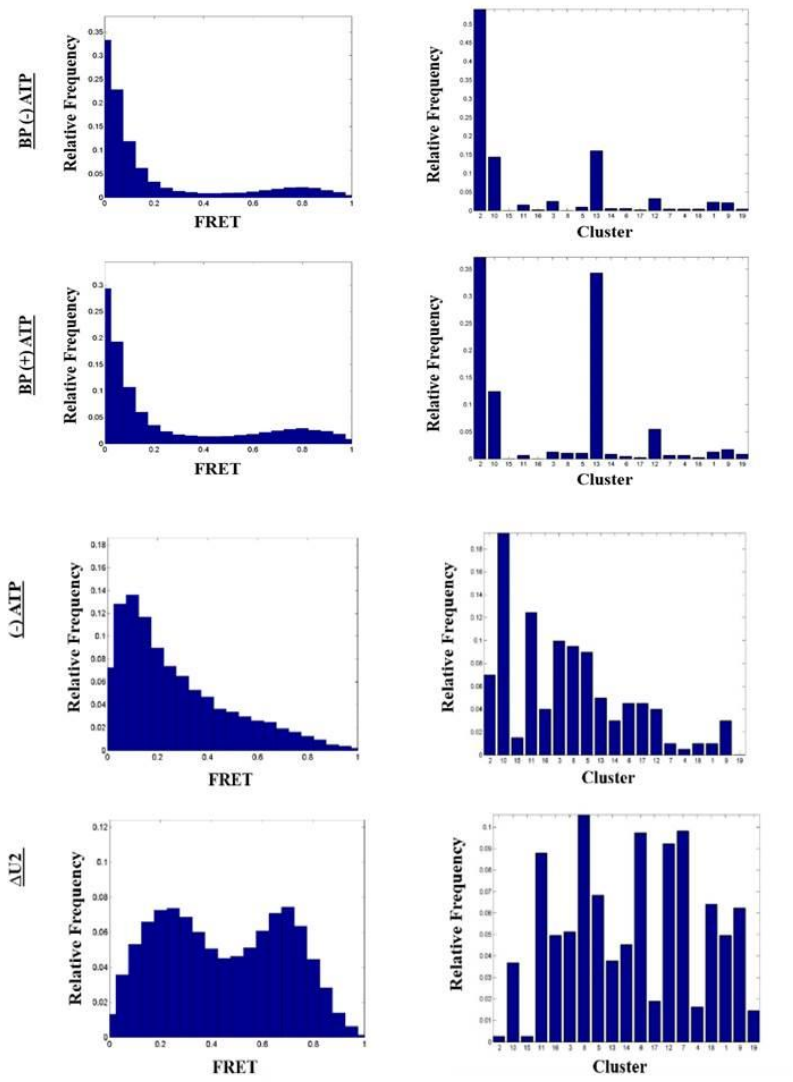


Figure 3.12 Conformational dynamics of early splicing stall at commitment complex formation

FRET histograms and hierarchical clusters compared for the early splicing assembly stalls at commitment complex 1 and commitment complex 2. Clusters are reorganized based on mean FRET value within the cluster so the plot reads from lowest mean FRET (left) to highest mean FRET (right).

According to the current understanding in the field, commitment complex 2 can assemble in the absence of ATP on a pre-mRNA with intact splice sites. Binding of Mud2 and BBP in addition to U1 snRNP now allows the pre-mRNA splice sites to sample excursions to mid-FRET states. However, U2 snRNA depletion, which stalls the spliceosome at CC2 similar to the omission of ATP from WT pre-mRNA, results in a broad sampling of several FRET states not restricted to low-FRET states. Since ATP is required to deplete U2 snRNA in these experiments, we speculate that this observation may be due to the removal of BBP-Mud2 complex by the ATP-dependent action of Sub2. The binding of U4:U5:U6 to the spliceosome results in the formation of the B complex. Depletion of U6 snRNA prevents this complex from forming and stalls the spliceosome at the A complex with U1 bound to the 5'SS and U2 bound to the BP (Figure 3.1). The most abundant dynamics (cluster 11) under these condition shows surprisingly low 5'SS-BP proximity, which may serve the purpose to facilitate the extensive remodeling that the pre-mRNA and snRNAs undergo at this step (Figure 3.13). Additionally, the 5'hyperstable pre-mRNA, which stabilized the interaction between the U1 snRNA and the 5'splice site and results in the inability of the spliceosome to exchange U1 snRNA for U6 snRNA, also leads to a stall at the A complex. We find that the 5'Hyp substrate in buffer has a FRET distribution that is lower in FRET than the BP mutant, suggesting that the 5'Hyp mutation leads to a change in the conformational space sampled by the naked RNA (data not shown). The addition of extract depleted of ATP diminishes the bulk of the high-FRET dynamics, leading to a lower FRET distribution (Figure 3.13). The addition of ATP has a minimal effect on the FRET distribution, which could be a result of the inefficient assembly of this substrate. We find that irrespective of ATP, the 5'Hyp pre-mRNA is dominated by the relatively slow 0.25-0.50 FRET transitions of

cluster 8. The inability of this substrate to assemble efficiently spliceosomes even in the presence of ATP results in early assembly complex behaviors.

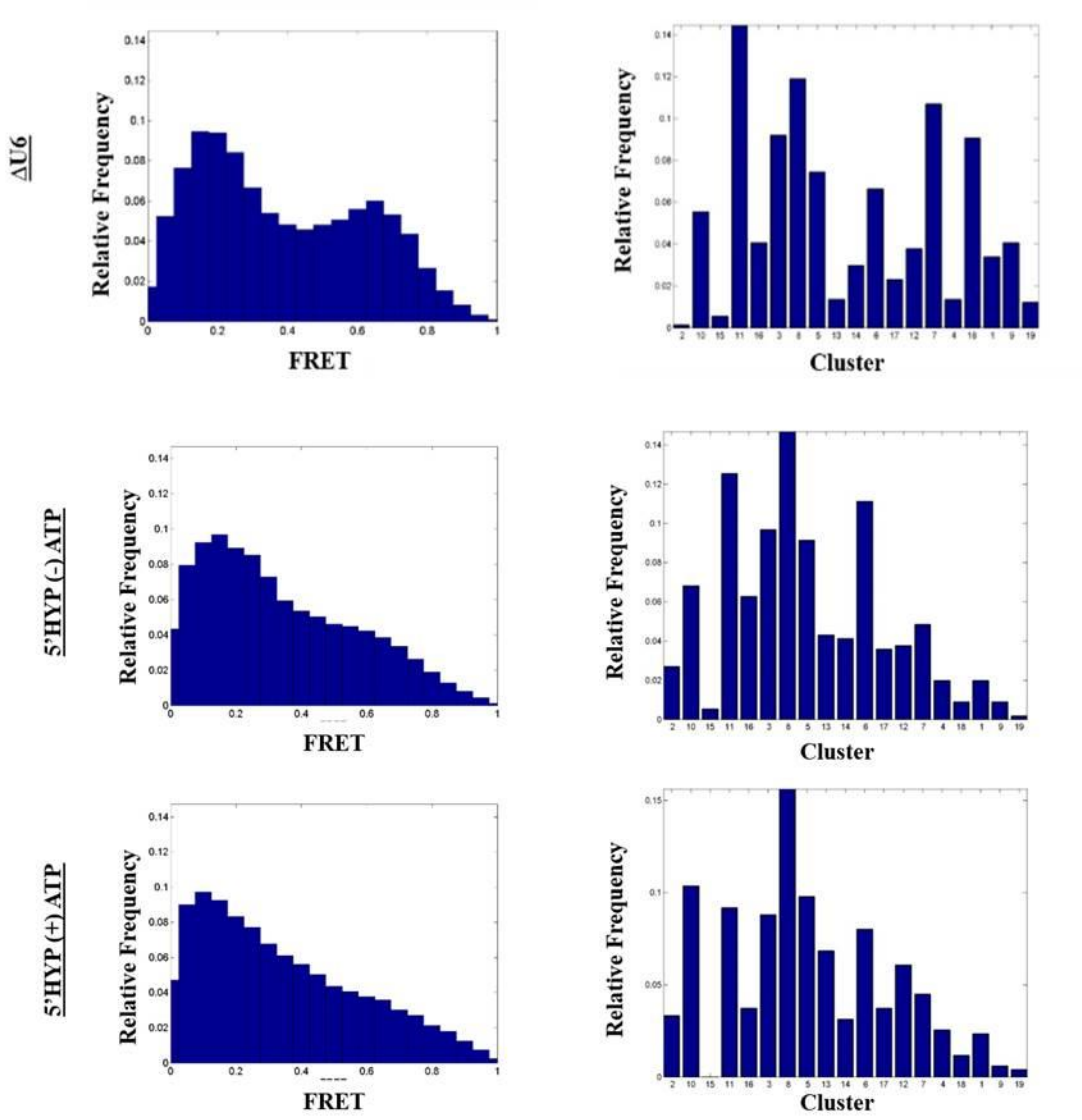


Figure 3.13 Conformational dynamics of early splicing stall at A complex formation

FRET histograms and hierarchical clusters compared for the early splicing assembly stalled at A complex formation. Clusters are reorganized based on their mean FRET value within the cluster so the plot reads from lowest mean FRET (left) to highest mean FRET (right).

3.3.3 Pre-mRNA predominantly explores close splice site proximity during late splicing cycle

With the spliceosome assembled on the pre-mRNA, several splicing regulators modulate the structure and conformation of the splice sites to facilitate the occurrence of the first chemical step of splicing. The exact timing and mechanism of how the splice sites are brought sufficiently close for chemistry is largely unknown. We performed smFRET experiments on two late splicing stalls at the B^{act} complex (via the Prp2_1 stall) and the C complex (via the 3'SS mutant pre-mRNA stall). The Prp2_1 stall allows for the spliceosome to progress until the B^{act} complex which has the U2:U5:U6 tri-snRNP assembled on it. The DExD/H box helicase Prp2 is then required for the spliceosome to undergo extensive remodeling to facilitate the first step of chemistry[35]. The most abundant cluster under this condition is cluster 5 which samples low-, mid-, and high-FRET states. Although we find that high-FRET states are sampled by the pre-mRNA at various steps of splicing, albeit at lower frequency, it is at this stage that we find this behavior to predominate. Results presented in Chapter III of this thesis show that the splice sites are brought into close proximity by the action of Prp2 protein, whereas prior to that the pre-mRNA shows a stable ~0.4 FRET state in the B^{act} complex. The sampling of high-FRET states under this condition could represent the pre-mRNA being extensively remodeled in the B complex to achieve the final tri-snRNP composition found in the B^{act} complex conformation. We find that relieving this stall by the addition of Prp2 protein allows for the spliceosome to go through both steps of splicing.

The 3'SS substrate blocks exon ligation through a mutation in the AG acceptor site near the 3'exon. When examined under single molecule conditions, the 3'SS mutant undergoes a similar shift towards lower FRET behaviors from higher FRET after the addition of ATP depleted extract (Figure 3.14). This seems to be a general phenomenon as all three substrate mutants adopt lower FRET behaviors during early assembly steps. In the case of the 3'SS mutant, the increase in low-FRET population originates from molecules now occupying cluster 8 as opposed to cluster 6, the predominant cluster in buffer (data not shown). Cluster 8 has dynamics similar to the cluster 6 found in buffer except that the transition to a higher FRET state is disallowed, suggesting the early assembly events are preventing the substrate from reaching the high FRET state, therefore leading to the lower FRET distribution. The addition of ATP leads to a slight shift towards high FRET that can be attributed to an increase in cluster 1 occupancy, which exhibits rapid transitions in the high-FRET regime and a decrease in the abundance of molecules in clusters with low-FRET state consensus sequences.

To determine how similar or different the predominant behavior of the pre-mRNA that were relieved of a splicing stall to achieve splicing were when compared to an unblocked pre-mRNA splicing. WT pre-mRNA was found to undergo a shift towards lower FRET states from the high FRET states observed in buffer upon the addition of (-)ATP extract (Figure 3.12). Upon addition of extract with ATP, we find that the predominant behavior shifts to high-FRET conformations with infrequent sampling of lower FRET states indicating proximity between the 5'SS and BP for the first step of splicing. A comparison of the most abundant cluster for the WT(+)ATP data with the reconstitutions that remove the stalls introduced at CC2, A and B^{act} complex via U2, U6 and Prp2 mutations shows that the predominant behaviors sampled under

these conditions are similar to those of the WT pre-mRNA (+)ATP as they populate cluster 7 (Figure 3.15).

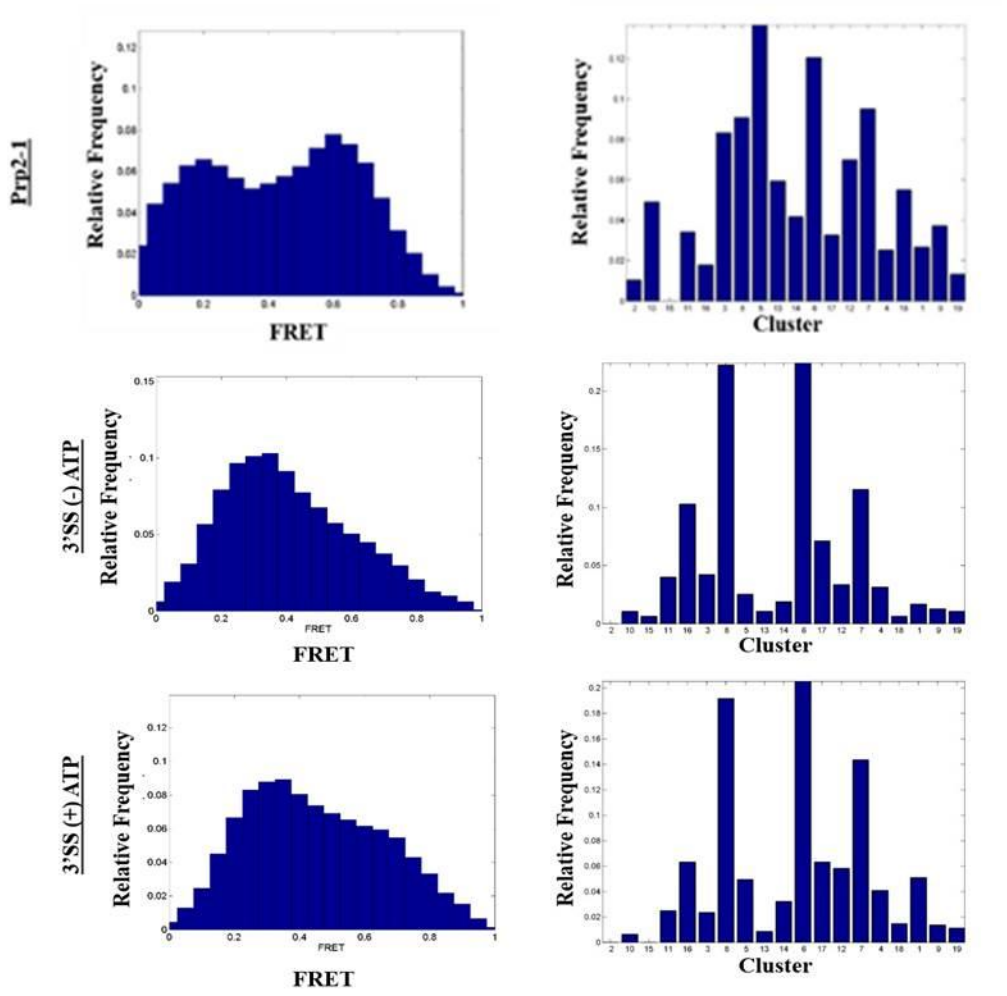


Figure 3.14 Conformational dynamics of late splicing stall at B^{act} and C complex formation

FRET histograms and hierarchical clusters compared for the early splicing assembly stalls at A complex formation. Clusters are reorganized based on mean FRET value within the cluster so the plot reads from lowest mean FRET (left) to highest mean FRET (right).

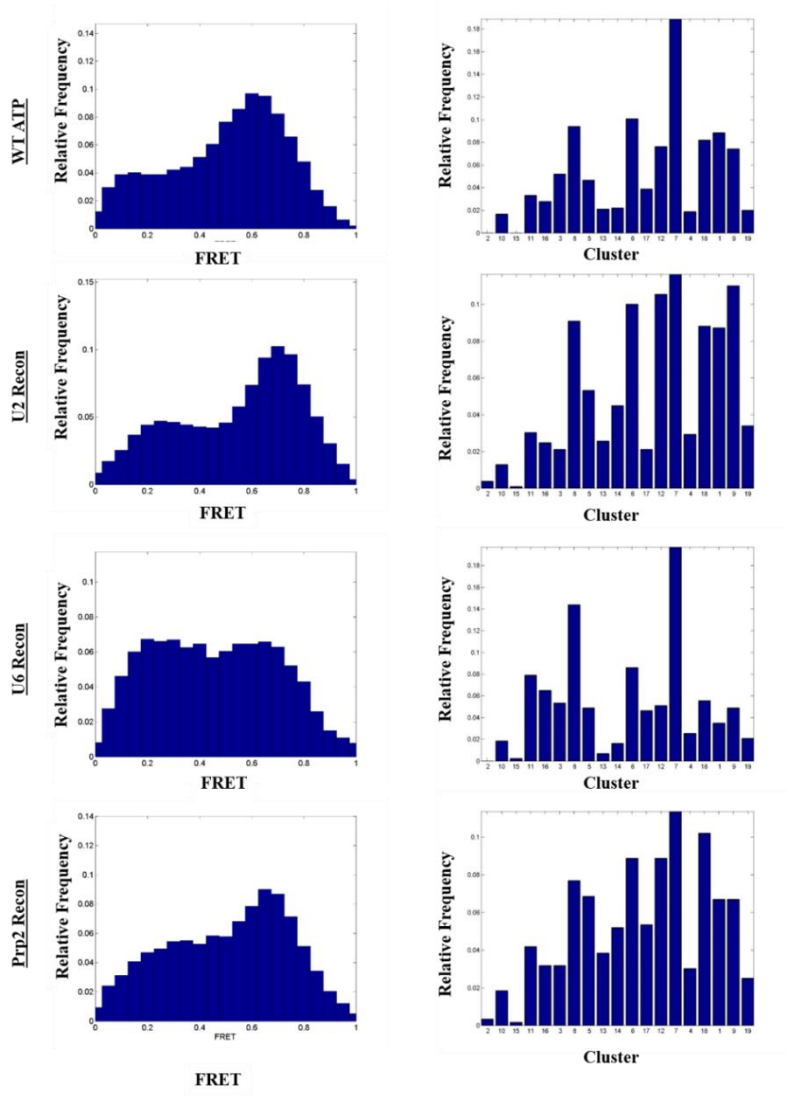


Figure 3.15 Conformational dynamics of pre-mRNA that can proceed through both steps of splicing

FRET histograms and hierarchical clusters compared for the early splicing assembly stalls at A complex formation. Clusters are reorganized based on mean FRET value within the cluster so the plot reads from lowest mean FRET (left) to highest mean FRET (right).

3.3.4 Clustering analysis groups biologically similar data

Clustering analysis allows us to objectively assess data sets by comparing the abundance of each cluster for each condition. We performed a similar linkage analysis on the different data sets collected. The abundance profile (fraction of molecules in a given cluster) for each condition was used as a similarity matrix for the analysis. The result of this analysis is represented by a hierarchical cluster tree whose leaf order is based on the similarity among conditions (Figure 3.16). An immediate observation from this plot is that the clustering routine was able to group conditions that are expected to be similar to each other. In addition, our results show that both the BP and 3'SS mutant pre-mRNA remarkably segregate into distinct clusters of their own separate from the other experimental conditions. This finding alludes to a possibility that the BP and the 3'SS mutations are recognized very early on in the splicing cycle and that they follow a kinetically and conformationally distinct path from the WT pre-mRNA. WT, BP and 5'Hyp pre-mRNA substrates in splicing buffer similarly fall into a separate branch, indicating a similarity in the secondary structures of these substrates, a finding that was also observed for the doubly exon labeled Ubc4 initially characterized by smFRET[37]. On the contrary, the 3'SS pre-mRNA, which can go through the first step of splicing, might need to be strictly regulated to prevent the formation of an aberrant mRNA. The pre-mRNA molecules in splicing buffer are in general less restricted than the BP mutant pre-mRNA in cell extract and often exhibit FRET dynamics that mimic dynamics during spliceosome assembly. Additionally, we find that the stalls employed at CC2, A and B^{act} complexes via U2 snRNA, U6 snRNA and Prp2 depletions respectively fall

close together in the tree away from both early ATP independent assembly events and late splicing events. Finally, the wildtype pre-mRNA in extract with ATP and conditions where extract depletions were relieved via reconstitution allowing them to undergo splicing again, are grouped most closely to each other (Fig 3.15). These results highlight the sensitivity of single molecule measurements and the clustering analysis to differentiate inherent structural differences in the RNA without any bias created by user input.

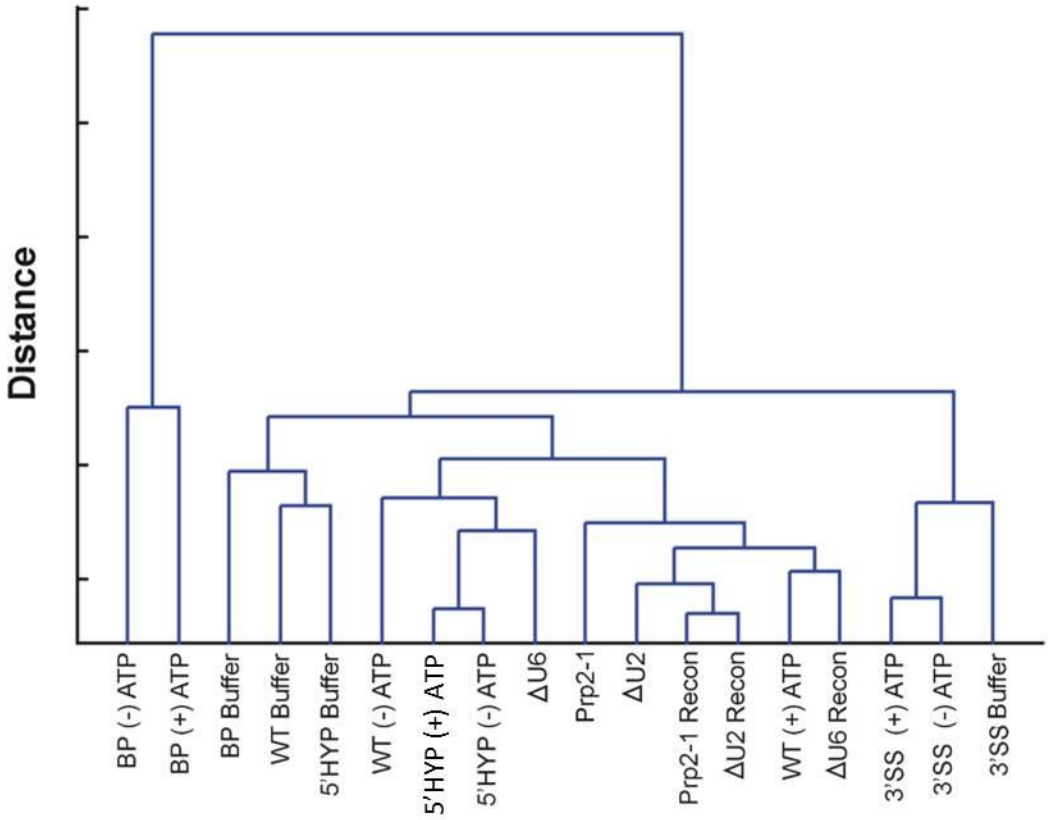


Figure 3.16 Hierarchical tree grouping the different experimental conditions

The abundance of each cluster for each condition was used as a metric for hierarchical clustering. Conditions are ordered based on similarity.

3.3.5 The splice sites sample unique conformational paths towards splicing

The pre-mRNA dynamics that lead to chemistry must be regulated in a strict manner to prevent mis-splicing and achieve precision. It has been observed that close proximity of the splice sites, even in the absence of proteins can lead to chemistry. A WT pre-mRNA that the spliceosome recognizes by binding to conserved sequences must adopt a set of defined conformations as it progresses along the pathway. Recently, it has been proposed that the spliceosome exists in two distinct conformations before the first and second steps of chemistry[81]. In Chapter III, we describe how the pre-mRNA conformations are kinetically modulated to favor the first step of splicing. In order to identify the predominant conformational behavior of the pre-mRNA at each slicing step, the consensus behavior within a cluster was determined by calculating the mean behavior of all the molecules within a cluster and then determining the single molecule with the shortest distance to the mean. These "consensus" molecules represent the average behavior of each of the 19 clusters (Figure 3.17). Comparison of the consensus behavior of the different splicing stalls as the pre-mRNA progresses along the splicing cycle is shown in Figure 3.18 a. We find that during early stages of spliceosome assembly, the intrinsic structure of the pre-mRNA is disrupted, leading to the 5'SS and BP predominantly in an extended conformation. However, in the presence of ATP, we find that pre-mRNA in CC2 shows some mid-FRET excursions. Additionally, we find that upon binding of U2snRNP, the pre-mRNA splice sites are removed farther from each other possibly to facilitate the extensive snRNA:pre-mRNA rearrangements to accommodate tri-snRNP binding. The 5'SS and BP are then brought closer together through the addition of ATP once the spliceosome has been activated by the loss of U1 and U4 snRNPs. Our results indicate that during the later stages of

assembly, the splice sites are predominantly limited to proximal FRET conformations that facilitate first step of catalysis. Relieving the splicing stalls using reconstitutions show behaviors similar to the pre-mRNA in extract with ATP (Figure 3.18 b) indicating that even if the splicing cycle is stalled at certain steps, relieving the stalls allow the pre-mRNA to go through a conformational pathway similar to what is adopted by a wildtype pre-mRNA undergoing splicing.

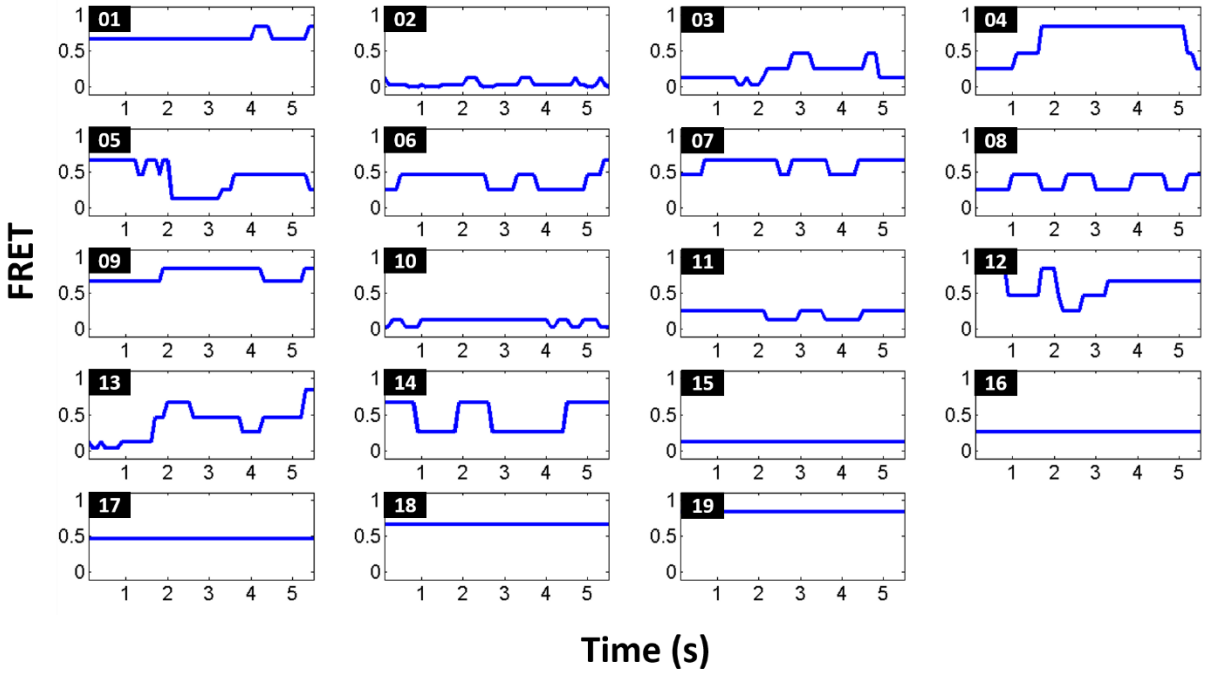


Figure 3.17 Consensus behavior of the identified hierarchical clusters

The pair wise distance from each molecule to the centroid of its assigned cluster was determined and the molecule with the smallest distance (highest similarity) to the centroid was plotted. These molecules represent the average or 'consensus' behavior of the molecules within the given cluster (inset).

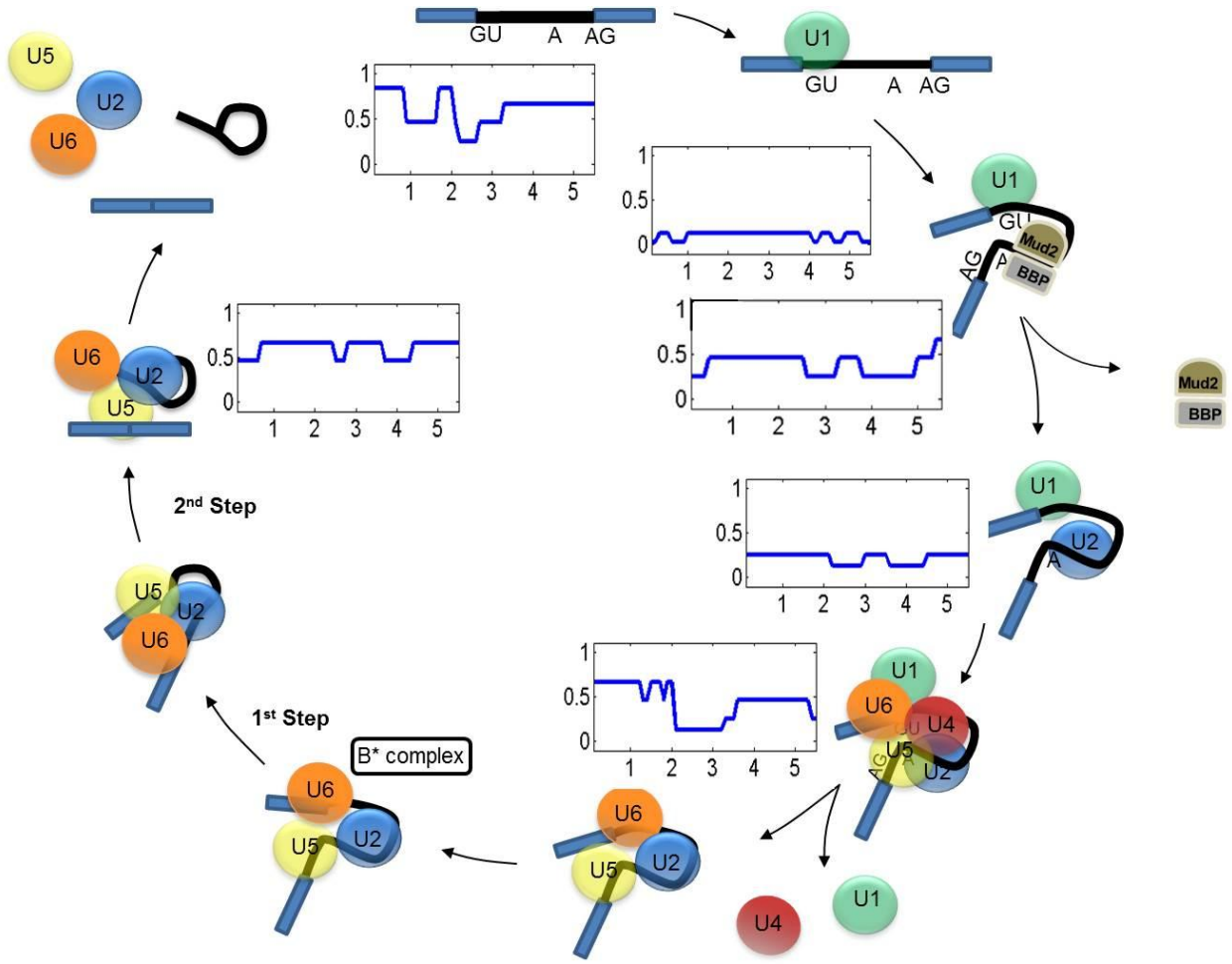


Figure 3.18 Consensus behaviors aligned with splicing cycle

(a) The consensus behaviors of the most abundant cluster for each block is paired with the relevant complex enriched in the splicing cycle. (b) The predominant consensus behavior in the conditions where the stall was relieved using a reconstitution was similar to a wildtype pre-mRNA that can go through the entire splicing process (WT (+) ATP).

3.4 Discussion

In this study, we have collected single molecule FRET data from Ubc4 pre-mRNAs labeled at the 5'SS and BP that can reach various steps of splicing assembly. To achieve enrichment up to specific points along the splicing cycle, we introduced substrate mutations and depleted specific splicing factors. Stalling the progress of splicing at these specific steps allowed us to delineate the dynamic proximity of 5'SS and BP during the early and late steps of splicing. We have employed a clustering routine to tease out the multi-state dynamic behavior of the single molecules observed under these experimental conditions. By utilizing the transition probability matrix that results from Hidden Markov Modeling, we created a FRET Similarity Matrix (FSM) that can be utilized in hierarchical clustering algorithms to group molecules of similar state and kinetic behavior. This technique has enabled identification of consensus behaviors of molecules at various stages of splicing. We find that the WT substrate as well as the mutant BP, 3'SS, 5'Hyp substrates all undergo a shift to low FRET during the ATP independent assembly events. The BP mutant, which is stalled during the earliest stages of assembly, can only bind the U1 snRNP in a stable fashion and adopts the lowest FRET conformation. The WT substrate can undergo all the necessary assembly steps to achieve both steps of catalysis. We controlled the assembly of the WT substrate by depleting the extract of ATP or ablating the activity of the U2 snRNP, U6 snRNP, and Prp2 protein. Although the U2 snRNP ablation (ΔU2) is expected to enrich the same complexes as the (-) ATP condition, we detected differences in the most abundant conformations present during the two blocks. The low FRET dynamics present in

WT(-)ATP mirror those of the BP mutant, while the most abundant dynamics in the ΔU2 can now sample mid-FRET states from a low-FRET state, suggesting that the presence of ATP but absence of U2 snRNP has effects on the assembly of components on the pre-mRNA. Genetic and biochemical evidence previously established a role for the ATP dependent helicases Sub2p and Prp5 near the site of U2 snRNP assembly[27,60]. We attribute the differences seen here to the ATPase action of Sub2 located near the branchsite. However, the absence of U2snRNP prevents the spliceosome assembly to progress beyond this stage. The ΔU6 block allows for the assembly of the U2 snRNP near the branchsite enriching the A complex. Although cross-linking studies have shown that U2 snRNP promotes cross-intron interaction between the BP and 5'SS through U1-U2 snRNP interactions[28], the results presented here show that at the A complex, Ubc4 pre-mRNA adopts a predominantly low FRET behavior with 5'SS and BP kept apart from each other. This complex, however, does explore transient high-FRET excursions, possibly contributing to the results observed in the cross-linking studies. The stall at the B^{act} complex mediated by the *prp2-1* mutation allows the substrate to assemble up to the point just prior to catalytic activation. At this stage, the most common pre-mRNA dynamics are molecules that sample low-, mid-, and high-FRET states. Relieving the splicing stalls with the reconstitution of splicing components allows the splicing cycle to proceed similar to the wildtype pre-mRNA in extract with ATP. The most common dynamics sampled are restricted primarily to mid- and high-FRET transitions with kinetics that favor the high FRET state. This suggests that the spliceosome components facilitate the stabilization of conformations where the 5'SS and BP are in close proximity during the latter stages of splicing cycle to promote catalysis. These results collectively indicate that the splice sites in the pre-mRNA approach proximal high-FRET states

in a non-monotonic or dynamic manner with frequent excursions sampling high- and low-FRET states. A progressive sequence of dynamics across the spliceosome assembly landscape has provided us with the ability to determine the effects the spliceosomal components on the juxtaposition of the 5'SS and branchsite as shown in Figure 3.18.

The novel combination of hierarchical clustering approach to understand complex single molecule data presented here has allowed us to group large data sets without any user input bias into biologically meaningful results. The analysis routine is generalizable to any set of smFRET data and we believe that it will become a commonly used tool as the behavior of increasingly complex systems is investigated by smFRET.

CHAPTER 4 : An early ATP independent role for Prp28 in commitment complex formation[3]

4.1 Introduction

Pre-mRNA is processed via the removal of intervening introns and joining of the exons in a two-step transesterification process called splicing. This process is catalyzed by a large and dynamic RNP complex termed the spliceosome that in *Saccharomyces cerevisiae* is composed of five small nuclear RNAs (snRNAs) and about 80 proteins. Exquisite precision in this process is imperative and derives from controlled rearrangements catalyzed by a group of RNA-dependent proteins called the DExD/H-box helicases. In *Saccharomyces cerevisiae*, eight DExD/H-box helicases facilitate extensive rearrangements in the spliceosome using energy from ATP hydrolysis, thereby ensuring splicing accuracy. By contrast, the actual chemical steps of pre-mRNA splicing are isoenergetic and do not require ATP hydrolysis.

The stepwise assembly of the spliceosome components on pre-mRNA starts with the binding of the U1 snRNP at the 5' splice site (5'SS) to form commitment complex 1 (CC1, Figure 4.1). In the next step, two proteins, BBP and Mud2 recognize and bind the branch site and 3' end of the intron forming commitment complex 2 (CC2). These steps do not require ATP

[3]Ramya Krishnan performed the single molecule experiments and analysis. Argenta M. Price at University of California, San Francisco, carried out the biochemical experiments characterizing the commitment complexes. Matthew Kahlscheuer expressed and purified Prp28 protein and helped with the titration experiments.

hydrolysis. The first ATP-dependent rearrangement occurs with Sub2, followed by Prp5, and U2snRNP joining to form the A complex. This event is followed by the joining of the tri-snRNP to form the B complex. The action of the DExD/H-box protein Prp28 then facilitates the exchange of U1 for U6 from the 5'SS in an ATP-dependent manner. This is followed by Prp8 and Brr2 mediated RNA:RNA and RNA:protein rearrangements to form the catalytic core of the spliceosome. The actual steps of splicing are promoted by the helicases Prp2, Prp16 and Prp22, and finally the spliceosome is disassembled by Prp43[82].

The ATP-dependent role for Prp28 in displacing U1 snRNA from the branch site has been extensively studied[83-86]. Both ATP-dependent and –independent roles have been proposed for most of the spliceosome DExD/H box helicases. Using a cold-sensitive Prp28 strain, *prp28-1*, it was found that Prp28 not only plays a role in spliceosome activation but also one of the earliest steps of assembly, formation of the ATP-independent Commitment Complex 2 (CC2) (unpublished results, A. Price et. al., 2013). In order to precisely characterize the role played by Prp28 in early pre-spliceosome formation, we assembled early spliceosome in the absence of ATP, in Prp28 depleted extracts on Ubc4 pre-mRNA labeled with fluorophores on the branch site and 5'SS. Our biochemical and single molecule FRET results suggest that Prp28 plays a novel role in promoting/stabilizing the ATP-independent formation of CC2.

4.2 Materials and Methods

4.2.1 Splicing extract preparation

The Prp28-depletion (ΔPrp28) strain was created by deleting the genomic copy of Prp28 and complementing with a plasmid containing Prp28 under control of a galactose inducible promoter. The galactose-induced copy of Prp28 had a 3xHA tag inserted 3aa into its N-terminus, which partially destabilized the protein (Thesis of Pratima Raghunathan, UCSF, 1997). The yeast strain was grown in YEP complemented with 2% (w/v) galactose and 0.04% (w/v) sucrose, until reaching an OD_{600} of 0.5. After filtering, the cells are washed with YEP, and resuspended in either YEP with 2% (w/v) galactose and 0.04% (w/v) sucrose (for Wt) or in YEP with 2% (w/v) glucose (for ΔPrp28). The cells were grown for 4.5 h, until the yeast were at OD ~1.8 and extract was prepared as described (Umen and Guthrie 1995). Extracts were checked by Western blot with anti-HA antibody to confirm that Prp28 was no longer detectable.

4.2.2 Native gel analysis

Commitment complex gels were performed essentially as described[87], except that reactions were incubated at 16 °C or 23°C for 20 min in the absence of ATP. The reaction was run on a 3% (w/v) polyacrylamide, 0.5% (w/v) agarose, 5% (v/v) glycerol gel electrophoresis in 0.5x TBE, run at ~5 V/cm for ~22 h in the cold.

4.2.3 Ubc4 pre-mRNA

Wildtype (WT) and branch point (BP) mutant Ubc4 pre-mRNA sequences were used and labeled with fluorophores near the branch site and 5'SS as described in Chapters 2 and 3.

4.2.4 Expression of recombinant Prp28 (rPrp28) protein

A 25 mL overnight culture of Rosetta II cells containing Prp28 in pRSetA vector plasmid was used to inoculate 3-4 L of LB broth. Bacterial cells were grown to an OD_{600} of 0.5-0.7. The cells were then cooled to 20 °C, induced with 125 μM IPTG and incubated for 12-18 h at 18-22 °C. The cells were then harvested and lysed in lysis buffer (50 mM Tris-HCl, pH 8.0, 300 mM NaCl, 10% (w/v) imidazole, 10% (v/v) glycerol). Any insoluble material was removed via centrifugation and the supernatant applied to Ni^{2+}-NTA resin equilibrated in lysis buffer. The resin was then washed with lysis buffer and eluted in elution buffer (50 mM Tris-HCl, pH 8.0, 300 mM NaCl, 150 mM (w/v) imidazole, 10% (v/v) glycerol). The eluted rPrp28 protein was dialyzed against dialysis buffer (50 mM Tris-HCl, pH 7.5, 300 mM NaCl, 1 mM EDTA, 10% (v/v) glycerol), snap cooled and stored at -80 °C.

4.2.5 Single molecule FRET

For smFRET experiments, the 5' biotinylated pre-mRNA was attached to a streptavidin coated quartz slide as described in Chapter 3. Yeast ΔPrp28 extract was then depleted of any remaining ATP by adding Glucose at 1 mM concentration along with Oxygen Scavenging System as described in Chapter 3 and added to either immobilized WT or BP mutant pre-mRNA on a slide. The slide was then imaged as the (-ATP, Prp28Δ) condition. For Prp28 reconstitution, expressed recombinant Prp28 protein (rPrp28) was added to the specified concentration to Prp28-depleted extract and incubated for ~10 min. This reconstituted extract was then flowed onto a slide with Ubc4 pre-mRNA bound to flow out Prp28 depleted extract that was present. The obtained single molecule trajectories were prefiltered by visually inspecting for the presence

of any substantial anticorrelation, presence of two fluorophores, and single-step photobleaching. We screened transition events found within the prefiltered FRET index trajectories for localized anticorrelation in the donor and acceptor traces. The postfiltering system required the use of the HMM tool vbFRET. Post-synchronized histograms (PSHs) were constructed by synchronizing individual FRET events. For the titration experiments, rPrp28 at various concentrations was added to reconstitute the extract. Starting from the lowest concentration, this extract was flowed onto the slide surface, incubated for 2-5 min and imaged.

4.3 Results

4.3.1 Absence of Prp28 protein inhibits ATP-independent CC2 formation

In order to investigate the role played by Prp28 in early splicing commitment complex formation, splicing reactions with radiolabeled RP51A pre-mRNA were assembled with either Wildtype (WT), Prp28-depleted (ΔPrp28), csPrp28 (*prp28-1* at 16 °C) or U2snRNA-depleted (ΔU2) extract, in the presence and absence of ATP (+/-ATP).). The resulting complexes were resolved on native gels (Figure 4.1, Price et al. 2013 in preparation). As expected, WT splicing reactions depleted of ATP are stalled right after CC2 formation and hence accumulate CC2, which migrates slower than CC1. Addition of ATP converts most of the CC2 into co-migrating pre-spliceosome/spliceosome (P/SP). In contrast, both ΔPrp28 and *prp28-1* splicing reactions in the absence of ATP accumulate the majority of pre-mRNA in CC1. Upon addition of ATP to the *prp28-1* splicing reaction, only about half as much pre-mRNA was converted to P/SP than was with WT extract, and most of the pre-mRNA that remained in commitment complexes remained

in CC1. The *prp28-1* does not completely block splicing at 16 ˚C, while in the ΔPrp28 strain the expression of Prp28 can be turned off by growth in media with glucose as the sugar source.

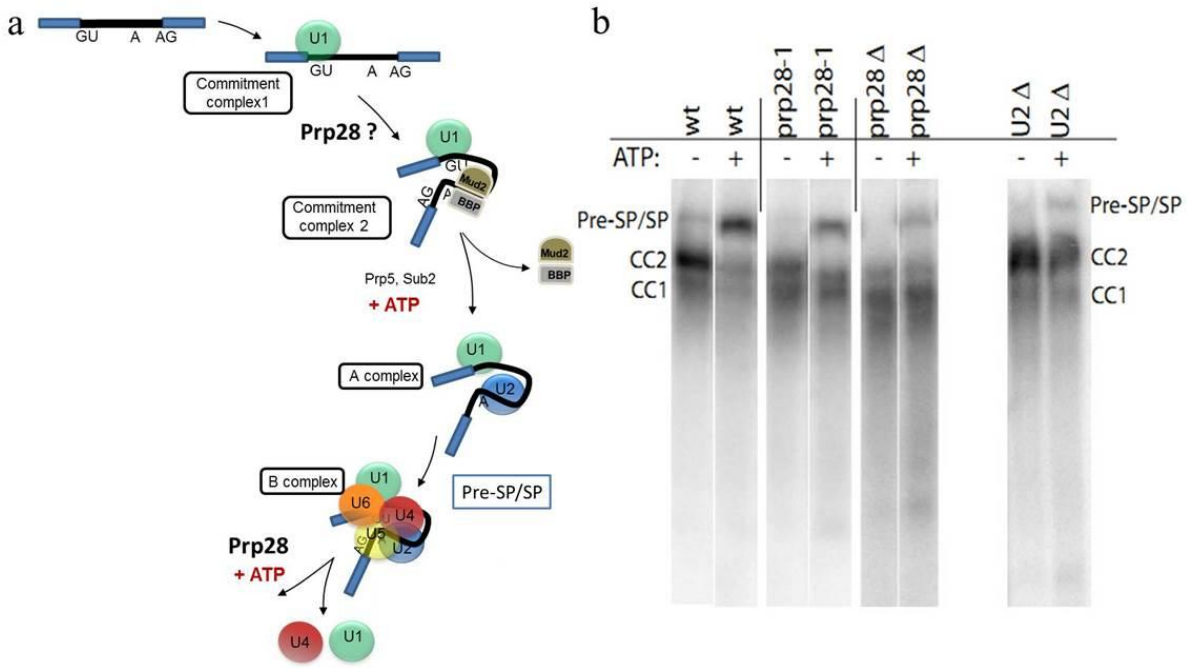


Figure 4.1 Depletion of Prp28 results in CC1 accumulation

(a) Schematic of a section of the splicing cycle depicting spliceosome assembly on pre-mRNA and the ATP-dependent role played by Prp28 in exchanging U1 for U6. (b) Native polyacrylamide gel showing complex formation using either WT, ΔPrp28, cold sensitive *prp28*-1 or U2 depleted (ΔU2) extract in the absence (-) or presence (+) of ATP on RP51 pre-mRNA (Price et al. 2013, in preperation). In the absence of Prp28 and ATP, the conversion to CC2 is inefficient.

After 3.5-5 h of growth in the presence of glucose, splicing extract made from these cells no longer contained HA-Prp28 detectable by Western blot for HA and failed to splice at 16 ˚C. The accumulation of pre-mRNA in CC1 in the absence of ATP is more pronounced in the ΔPrp28 extract. By contrast, extracts that were U2snRNA depleted, which should prevent complex A formation and result in the accumulation of CC2 complex, exhibits the expected behavior both in the presence and absence of ATP. It seems likely that the absence of Prp28 affects the association of BBP and Mud2 with CC1. Our results indicate that Prp28 plays an early, ATP-independent role for Prp28 in the conversion of CC1 to CC2 and thereby commits the pre-mRNA to splicing. To confirm that the Prp28-mediated effect seen is not limited to the RP51 substrate, we performed a similar set of experiments using the Ubc4 substrate. In this case, depletion of Prp28 combined with the absence of ATP resulted in an accumulation of CC1 complex as with RP51 substrate (Figure 4.2). To confirm that this was due to the depletion of Prp28, we added back rPRp28 to the extract. This however, only resulted in a marginal increase of CC2 band (adding back rPrp28 in both RP51 and Ubc4 has been unsuccessful in our hands to date). A wildtype control depleted of ATP, as expected resulted in the accumulation of CC2 band. However, when Prp28 was depleted and ATP present in the extract, the resulting effect caused an accumulation of CC1 instead of the P/Sp band, consistent with our hypothesis. If Prp28 only has a late, ATP-dependent role in splicing, depletion of Prp28 should have resulted in the pre-spliceosome accumulation as with a wildtype control extract in the presence of ATP (Figure 4.2). Although this result adds strength to the hypothesis, addition of rPrp28 needs to be optimized.

4.3.2 Prp28 allows the 5'SS- BP to explore conformations of close proximity

To investigate the role of Prp28 in CC2 formation and the effect on pre-mRNA conformation, we chose Ubc4, our well-studied pre-mRNA with fluorophores near the 5'SS and BP (as described in Chapters 2 and 3). The WT pre-mRNA was biotinylated at the 5'end and immobilized on a streptavidin coated quartz slide. Splicing reactions were assembled in the ΔPrp28 extract in the absence of ATP (-ATP) and flowed onto the slide (Figure 4.3). After initial imaging, the ΔPrp28 extract was reconstituted with recombinantly expressed rPrp28 protein at 80 nM concentration and flowed onto the slide.

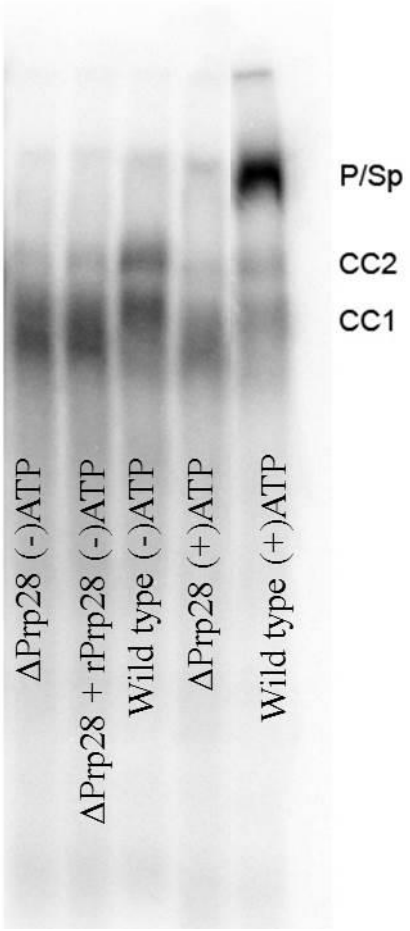


Figure 4.2 Inhibition of CC2 formation in the absence of Prp28 in Ubc4 pre-mRNA

Yeast extract assembled on radiolabeled Ubc4 pre-mRNA in either Prp28-depleted extract (ΔPrp28) or wildtype (WT) extract analyzed on a 3% acrylamide, 0.5% agarose, 5% glycerol gel.

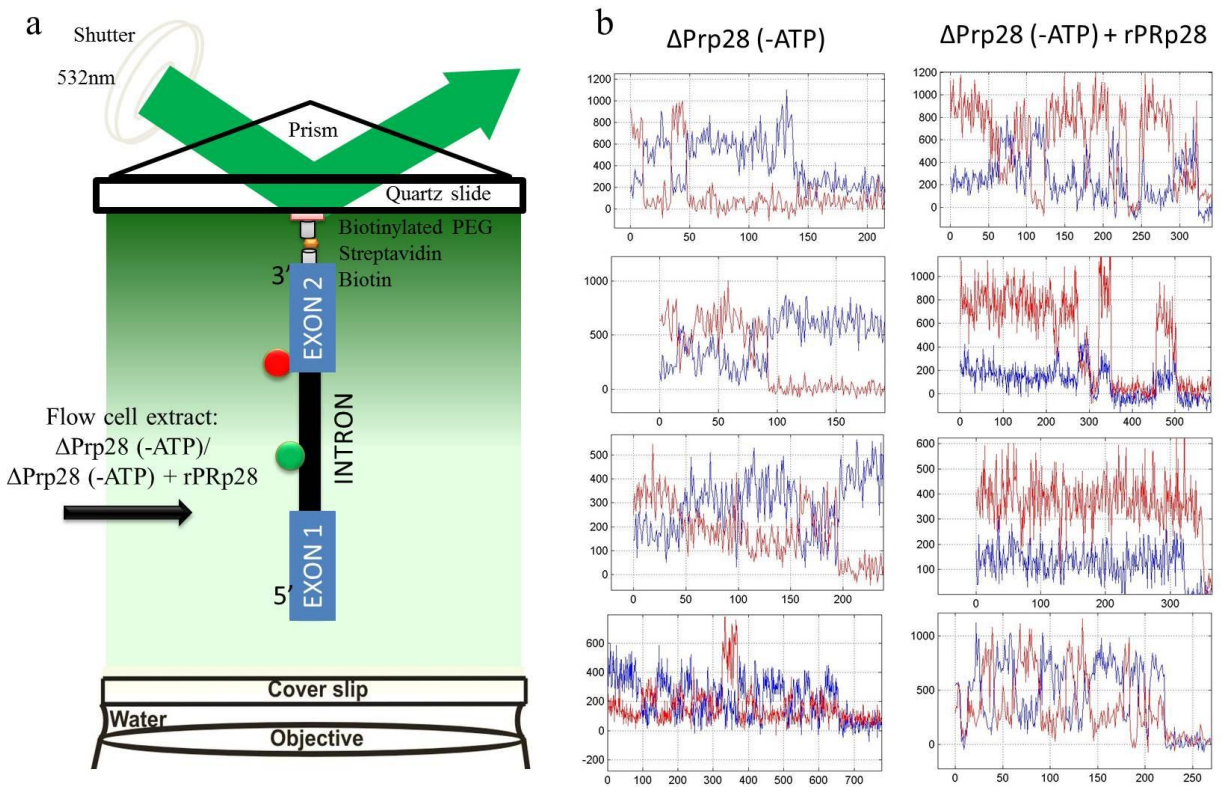


Figure 4.3 SmFRET set-up and representative traces

(a) Schematic of the smFRET experiment showing Ubc4 pre-mRNA immobilized via the 5' biotin onto a streptavidin coated quartz slide. ΔPrp28 extract in the absence of ATP (-ATP) or the same extract reconstituted with rPrp28 is flowed onto the slide and imaged. (b) Representative single molecule traces from ΔPrp28 (-ATP) and ΔPrp28 (-ATP) reconstituted with 80 nM recombinantly expressed Prp28 protein. Red trace represents Cy5 and blue trace represents Cy3.

Representative single molecule traces from these conditions show anti-correlated behavior under both conditions (Figure 4.3). FRET probability histograms plotted by sampling the first 100 frames of each single molecule trajectory show that under Prp28-depleted condition, the pre-mRNA exhibits a bimodal distribution with centers at ~0.21 and ~0.7 (Figure 4.4). Addition of rPrp28 induces the pre-mRNA molecules to explore higher FRET states (Figure 4.4). Post-synchronized histograms (PSHs) also show that high FRET transitions are favored in the presence of rPrp28 (Figure 4.5).

We next assembled early spliceosomes in the absence of ATP in the branchpoint A-to-C mutant (BP) pre-mRNA substrate and imaged it under single molecule conditions. Substrates with mutant or missing branchpoint regions are known to only form the CC1 complex[88], since the association of the branchpoint binding protein (BBP) and Mud2 are dependent on the presence of an intact BP sequence[89,90]. If the high FRET transitions observed in the WT pre-mRNA to be induced by rPrp28 were representative of 5'SS-BP proximity in CC2, then these transitions should be absent from the BP mutant substrate since the lack of a correct BP sequence prevents Mud2-BBP complex from binding it. FRET probability histograms show that, in the absence of ATP, ΔPrp28 extract assembled on BP mutant pre-mRNA shows a bimodal distribution that does not change significantly with the addition of rPrp28 protein (Figure 4.5). This result supports the conclusion that Prp28 is indeed responsible for driving the proximity between 5'SS and BP in WT pre-mRNA. Pre-mRNA in CC1 may be in equilibrium between high- and low-FRET states, and rPrp28 drives this equilibrium towards CC2 signature wherein the 5'SS and BP are more stable in the proximal high-FRET state.

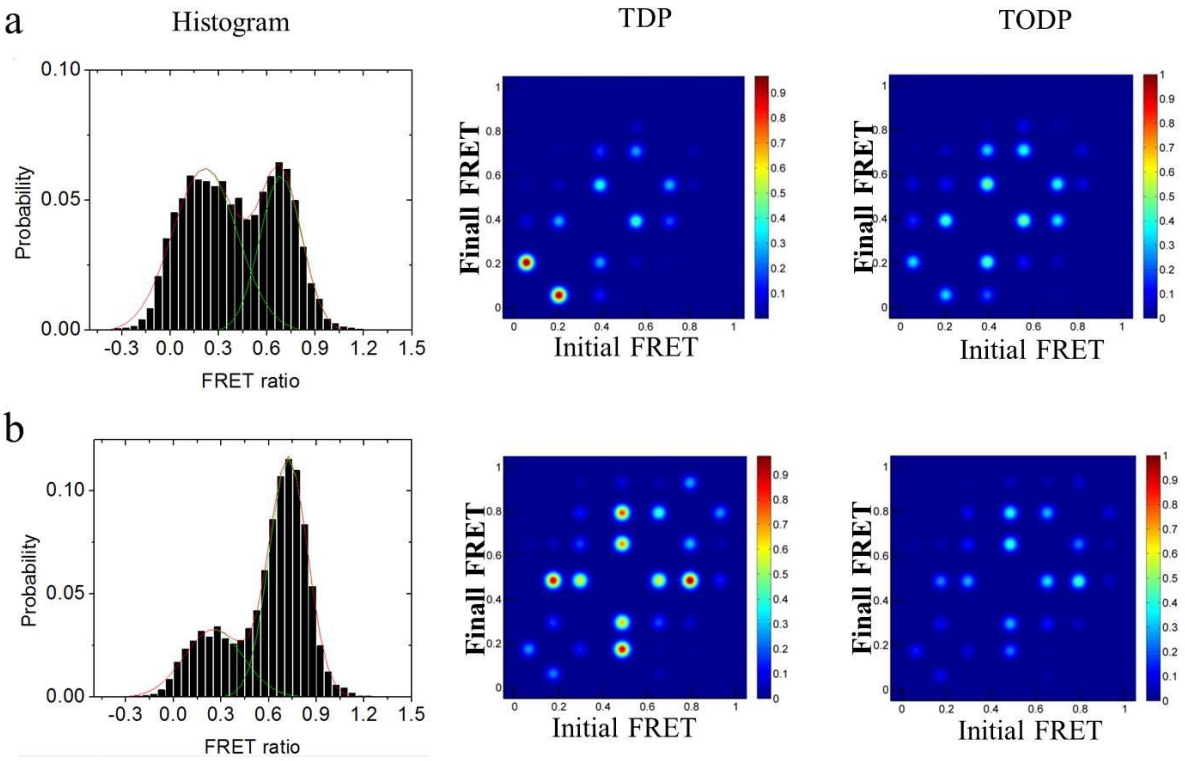


Figure 4.4 Addition of rPrp28 protein shifts the 5'SS-BP of Ubc4 pre-mRNA to explore conformations of close proximity

FRET histograms generated by binning the first 10 s of the raw single molecule FRET trajectories from doubly labeled Ubc4 pre-mRNA imaged under (a) ΔPrp8, (-) ATP extract and (b) ΔPrp8, (-) ATP extract incubated with 80 nM rPrp28 protein. Transition Density Plots (TDPs) show that the predominant transition under condition a is restricted to low FRET ranges while condition b shows more higher FRET transitions. Transition Occupancy Density Plots (TODPs) depicting all the transitions encountered by the molecules under both conditions.

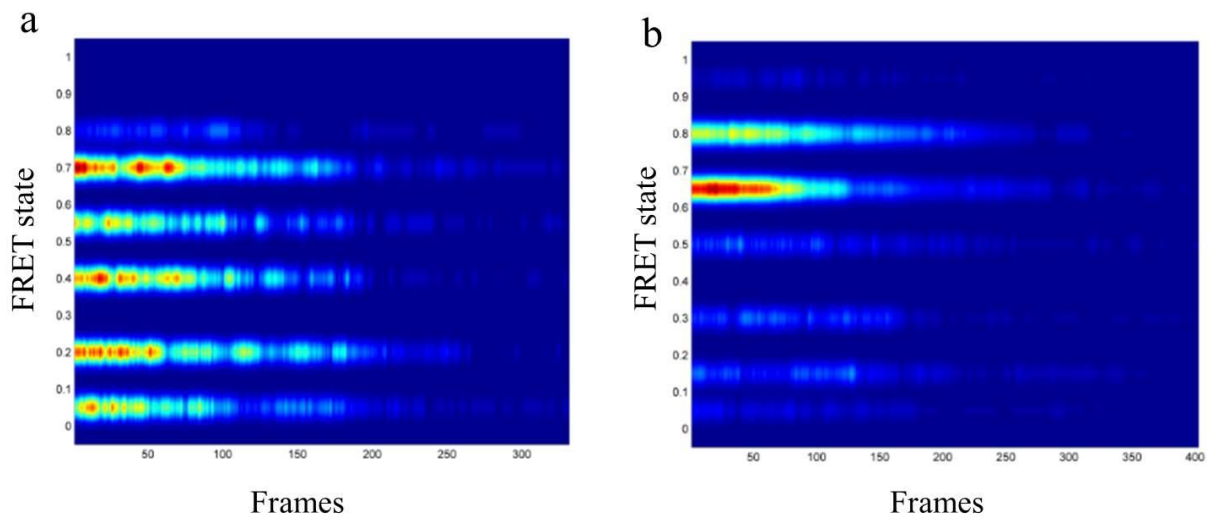


Figure 4.5 Post-synchronized histogram showing the increase in the high FRET transitions upon rPrp28 addition

Post-synchronized histograms of pre-mRNA dynamics in the absence of ATP in either (a) ΔPrp28 extract or (b) ΔPrp28 extract supplemented with recombinant Prp28.

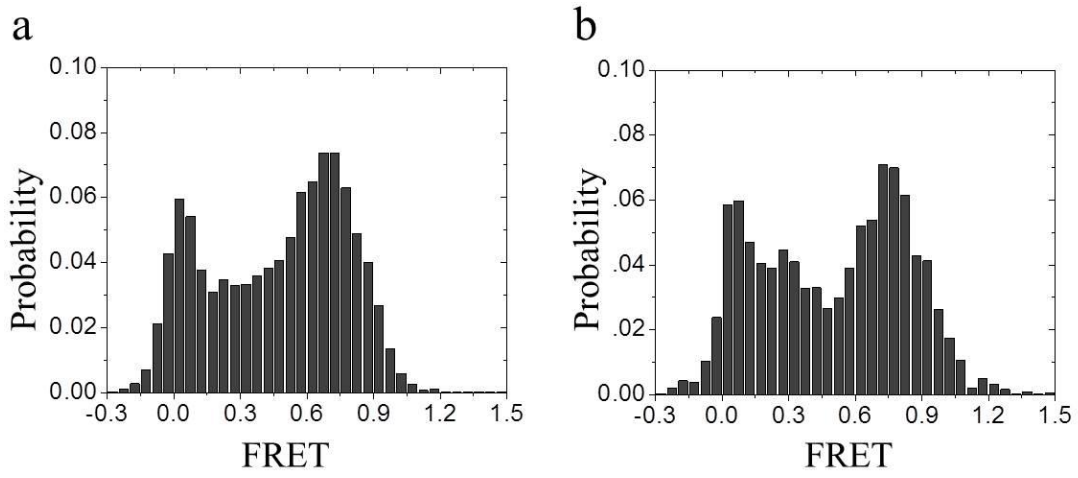


Figure 4.6 BP mutant 5'SS-BP dynamics are not affected by Prp28

FRET histograms generated by binning the first 10 s of the raw single molecule FRET trajectories from doubly labeled BP mutant Ubc4 pre-mRNA imaged under conditions of (a) ΔPrp8, (-) ATP extract and (b) ΔPrp8, (-) ATP extract incubated with 80 nM rPrp28 protein.

4.3.3 Titration of various concentrations of Prp28 shifts the equilibrium

To assess if the addition of rPrp28 indeed shifts the equilibrium between the high- and low-FRET states observed in the pre-mRNA traces, we added increasing concentrations of recombinant rPrp28 to Ubc4 pre-mRNA incubated with ΔPrp28 extract in the absence of ATP. In the absence of rPrp28, the probability histogram of 5'SS-BP dynamics shows bimodal sampling of both high- and low-FRET states. Addition of 20, 40 and 80 nM rPrp28 biases this sampling increasingly towards the high-FRET state (Figure 4.7). The 40 nM and 80 nM rPrp28 concentrations do not significantly differ in the equilibrium between these states. It appears that very low concentrations of rPrp28 are sufficient to drive the equilibrium towards high-FRET states under these conditions.

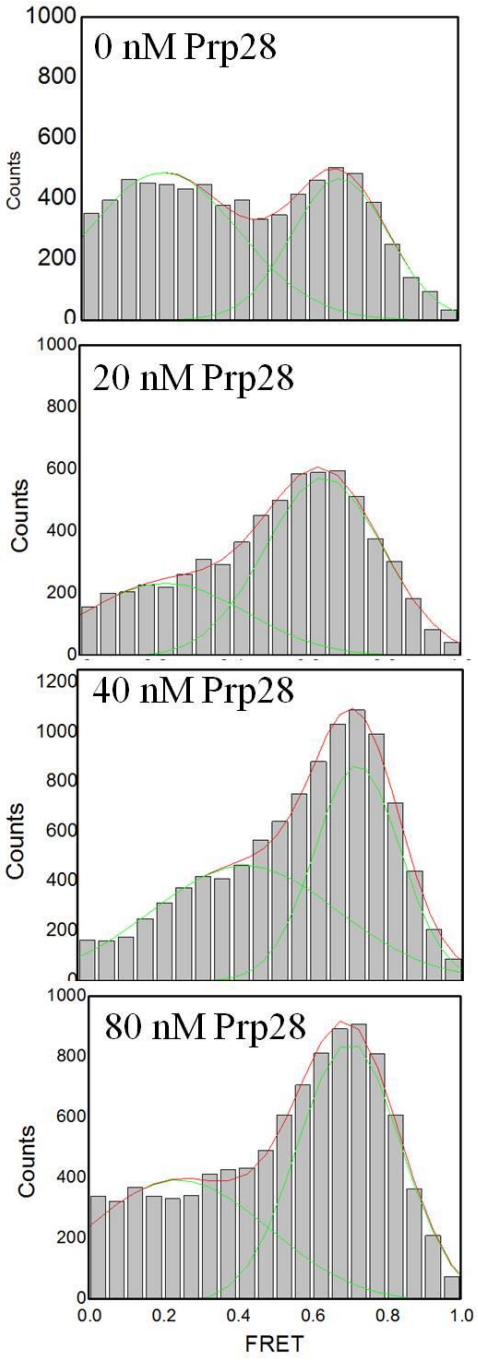


Figure 4.7 Pre-mRNA dynamics analyzed with varying concentrations of Prp28

FRET probability histograms showing 5'SS-BP FRET distributions at varying concentrations of rPrp28.

4.4 Discussion

DExD/H box helicases have been studied quite extensively in recent years. It has become clear that we still have a long way to go before fully understanding them. Questions ranging from what is their mechanism of action to what is their substrate remain unanswered. The spliceosomal DEXD/H-box proteins have been shown to function at different points in the splicing cycle in ATP-dependent and –independent roles. These roles are well known for Prp5[28,91,92], Prp2[31,35] Prp22 and Prp16[32]. In the work presented here, we have shown that Prp28 plays an early, ATP-independent role in spliceosome assembly in addition to its well-known ATP-dependent role in substituting U1 for U6 snRNP[85,93]. We show that Prp28 promotes the formation of CC2 complex. The 5'SS and BP of pre-mRNA assembled in CC1 toggle between low-FRET (distal) and high-FRET (proximal) states. Binding of the Mud2-BBP complex and conversion to CC2 pushes the conformation to closer proximity (high-FRET state). We show that rPrp28 induces this change possibly by stabilizing BBP and Mud2 binding on the branchsite of the pre-mRNA during CC2 formation, since a BP mutant pre-mRNA does not show a Prp28-dependent high-FRET state. We believe that this principle of toggling might be a common theme that the spliceosome exploits to achieve precision[94] (Chapter 2).

Another approach that can be used to tease out the hypothesis is to use extract from a U1-TAP yeast strain. Using a pre-mRNA assembled in this extract in the absence of ATP, we can affinity purify any complex that is bound by U1snRNP in the absence of ATP, the CC1 and CC2 complex. Using a similar set up, in the presence of the BP mutant pre-mRNA, we can purify the CC1 complex. After purification, one half of this can be run on a native gel after TEV protease mediated cleavage of the complex from the IgG-biotin tag and the other half on the single

molecule microscope for side by side comparison of signatures of CC1 and CC2 in Ubc4 substrate. Once definitive signatures of CC1 and CC2 are obtained, a variety of routes can be taken to confirm if CC2 formation is enhanced by Prp28. We could deplete the pulled down complex in the absence of ATP of any Prp28 using a specific antibody and then analyze the complexes using smFRET or make the ΔPrp28 strain in the U1-TAP background.

The results presented in Chapter 2 hint towards the possibility that the 5'SS and BP in Ubc4 pre-mRNA only come into close proximity after remodeling by Prp2. The experiments presented there with Ubc4 substrate in wildtype extract (Snu66-TAP background) under conditions of ATP depletion and U2snRNA depletion should essentially be signatures of CC2 complex. However, we find that under these conditions, 5'SS and BP of Ubc4 are still farther apart from each other. If according to the results presented here, Prp28 drives the spliceosome to form and possibly stabilize CC2 complex, this should result in low FRET states population. However, we find this not to be the case. There could be several reasons for this; 1) the splicing extracts used under these two experimental conditions might have slightly altered spliceosome to pre-mRNA stoichiometry. To resolve this, a wildtype extract with the same background as the ΔPrp28 strain needs to be tested in the absence of ATP. 2) Another possibility is that these experiments were carried out at 80 nM Prp28. If Prp28 stabilizes CC2 formation in a concentration dependent manner, the conditions used in these experiments could be overly saturating. If this is the case, Prp28 reconstitution experiments should be done at a lower concentration of Prp28.

Additionally, our preliminary results from using a Prp28 mutant with a single cysteine labeled with a fluorophore shows that even in the absence of ATP, Prp28 binds the 5'SS of the pre-mRNA transiently. Further experiments to characterize if the binding and dissociation rates of

labeled Prp28 varies in the presence and absence of ATP are required to kinetically assess the mechanism of Prp28-mediated early complex formation.

4.5 Acknowledgements

I would like to thank Chanrith Siv for assisting with the collection of rPrp28 titration data.

CHAPTER 5 : Summary and Outlook

5.1 Pre-mRNA splice site dynamics using smFRET and SiMPull-FRET

Structural information about the spliceosome has largely been derived from electron microscopy images of stable complexes[95]. Recently crystallization of spliceosome components like Prp8, known to be at the heart of the spliceosome are providing clues into how this enigmatic machine functions[30,81,96]. However, the pre-mRNA so far is largely excluded from these studies. Although some efforts have been made towards mapping the position of the splice sites in these complexes[97], these at their best, capture only static images of the system. The dynamics of any biological system hold the key to how it actually functions. The results presented herein on pre-mRNA dynamics clearly reveal that splice sites transiently and reversibly sample close splice site proximity many times throughout the splicing cycle (Chapters 2, 3 and 4). Some of the early biochemical experiments meant to probe splice site proximity pointed towards a possibility where the splice sites were brought close together very early in the splicing cycle[98,99]. Our results show that the splice sites adopt close proximity only after the action of Prp2 (Chapter 2). Careful analysis of FRET dynamics provides vital insights into the coordinated interplay between the snRNPs, pre-mRNA and DExD/H box helicases to achieve efficient splicing.

In this thesis, we have described the use of two different techniques to understand pre-mRNA dynamics - one, a direct attachment of 5' biotinylated pre-mRNA on a streptavidin coated quartz slide and the other, more recently developed, SiMPull-FRET, where the immunopurification of a complex is followed by an antibody mediated immobilization on a quartz slide. Direct immobilization of pre-mRNA has its advantages in that it is simple and allows the user to change the experimental conditions simply by flowing in and out different variations of extracts and mutants. While it does not allow for the same flexibility in varying the conditions, SiMPull-FRET on the other hand offers a cleaner system to address focused questions to associate specific FRET changes with single components. To the best of our knowledge, this is the first time smFRET dynamics has been studied in the spliceosome using a purified system. Both techniques have provided the field with significant advancement as tools for investigating mechanistic details of various stages of splicing and are fully extendable to other biomolecular systems.

5.2 Focusing on the first and second step splicing complexes

In order to gain insights into the structure of the spliceosome and the mechanism by which splicing occurs, significant efforts in recent years have been directed at the catalytic steps themselves[31,32,35,43,59,100,101]. It is known that conformational change in the spliceosome precedes both catalytic steps of splicing. In Chapter 2, we have focused on the pre-mRNA dynamics at the first step of splicing by assigning specific roles to the proteins known to be involved in this. Specifically, we have shown that Prp2 acts to remodel the pre-mRNA bringing about a large conformational change positioning the 5'SS and BP in close proximity. This opens up the obvious question of

what the target of Prp2 could be. One speculation is that Prp2 remodels the spliceosome by weakening the protein interactions near the BP. One of the critical snRNA-pre-mRNA interactions in the spliceosome involves a short duplex between the U2 snRNA and the BP region where the 5′ end of the U2 snRNA contains a sequence complementary to the BP, with the exception of a bulged branchpoint adenosine[44,102]. It has been proposed that in order to avoid a premature attack, the adenosine residue is thought to be physically sequestered by the SF3b subcomplex, one of the two multimeric protein complexes required for the addition of U2 snRNP during pre-spliceosome formation[31]. Structural reorganization of the spliceosome by Prp2 possibly displaces SF3b protein allowing for the BP to become accessible for the first step chemistry to occur. This hypothesis could be tested by smFRET between a labeled SF3b protein and pre-mRNA substrate with a fluorophore label near the BP in the presence and absence of Prp2. In human spliceosomes, structural analysis have shown that SF3b complex crosslinks to the BP adenosine[103]. If according to the proposed model the BP is sequestered by SF3 sub-complex, we would be able to look for direct interactions between BP and the labeled protein which should subsequently get altered with the addition of Prp2.

Another interesting facet to explore is the mechanism by which Prp2 as a helicase mediates the displacement of SF3b. One model is that Prp2 slides along the BP of the pre-mRNA fulfilling its role as a helicase much like Prp22[104], weakening the U2-BP base pairing and thereby destabilizing SF3b subcomplex. To test this, we can utilize the SiMPull-FRET set-up assembled using the heat inactivatable Prp2-1 extract that has been reconstituted with 5'-end labeled U2 snRNA as described in Chapter 2. Ubc4 pre-mRNA with fluorophore near the BP assembled in this set-up can essentially serve as a single molecule helicase assaying system to test if the addition of Prp2 (possibly ATP-dependent) weakens the interaction between U2 snRNA and pre-mRNA. On the contrary, our findings may already hold the answer to how this occurs. Prp2 might affect the local structure of the

pre-mRNA (as shown in Chapter 2) to destabilize SF3b. This testable hypothesis will provide vital clues to what the target of this important first step protein might be.

In Chapter 2 we have shown that Cwc25, a protein whose mechanistic roles functions were unknown, plays a prime role in kinetically modulating the spliceosome towards enhanced first step catalysis. Our findings indicate that Cwc25 does this by binding to a site within FRET distance to the BP of the pre-mRNA (Figure 3.14). However, binding a dissociation of Cwc25 evidences the transition between first and second step catalytic conformational switches. With the singly labeled Cwc25, we can address kinetically when and how long Cwc25 is bound near the BP prior to the first step. Which factors are responsible for the removal of Cwc25 bound near the BP in Ubc4 pre-mRNA, where the distance between the BP and 3'SS lies somewhat in between the longer and shorter version of actin substrates previously used[100]. Additionally, we can also explore if the displacement of Cwc25 is mediated by the ATPase action of the helicase Prp16 by using an ATPase deficient Prp2 mutant.

With the reconstitution of second step of splicing, it can become clear that the spliceosome exploits structural reorganizations to regulate chemistry. Our findings detailed in Chapter 2 show that prior to the first step of splicing, Prp2, in an ATP-dependent manner mediates a large conformational change in the pre-mRNA itself. Similarly, does the second step helicase Prp16 mediate a conformational change in the pre-mRNA, possibly bringing the 3'SS and 5'SS close together? If so, is the role of other second step factors (Prp18-Slu7) to stabilize the second step conformation? SiMPull-FRET offers easy means to answer these questions by simply immobilizing the B^{act} complex as described earlier with the sequential addition of step 1 and step 2 factors while imaging the pre-mRNA. In order to image single molecules throughout this process, a traditional continuous exposure is almost impossible with the current dyes. A simple B^{act} to C complex chase had proven to be

extremely challenging in our previous experiments (Figure 3.12 c). To circumvent this, we will use time-lapse illumination and increase our observation window with short exposures where we capture dynamics and longer dark periods with no illumination that can be used to image pre-mRNA molecules to an hour in buffer conditions. Alternatively, we could use dyes with enhanced photostability that have a proximal protective agents linked to it instead of the traditional cyanine dyes as described[105]. It would be interesting to observe how Prp16 affects the molecules with a stable high FRET state (post-first step chemistry) versus those molecules that are still toggling between high and mid-FRET states (Prp2-induced) even after Cwc25 addition, providing clues towards how Prp16 mediates proofreading.

The catalytic center of the spliceosome holds several secrets. With the recent crystal structure of Prp8 protein, a highly conserved factor known to be at the active site of the spliceosome[81], it has been increasingly evident that the RNaseH-like domain on this protein could be important either as the docking site for catalytic events or directly take part in catalysis. It has been proposed that Prp8 has two alternative states that facilitate either the first or the second step[106,107]. In a screen for the identification of suppressors of splicing defects caused by pre-mRNA mutations, Prp8 suppressors were identified, that can rescue or suppress first step and second step splicing defects[106,108]. First step alleles favor the first step, accumulate intermediates and produce very little mRNA. Second step alleles favor the second step, result in very little intermediates but make second step products. Some of the suppressors of 5'SS, 3'SS and BP mutations map to the b-finger of the RNaseH-like domain[81]. This domain has been shown to assume several different conformations which in theory could be relevant for the transition between first and second steps of splicing. It is plausible that the conformations adopted by the RNaseH domain influences and informs the structure of the pre-

mRNA towards catalysis. Our SiMpull-FRET set up offers a way to test if in the presence of extract with the first step alleles, the pre-mRNA adopts conformations favorable for the first step as opposed to the second step and conversely, in the presence of extract with second step alleles, the pre-mRNA adopts conformations favorable for the second step. If this is true, it could mean that either these suppressor mutations of Prp8 are directly involved in catalytic steps of splicing or that these mutant residues encourage the rearrangements in the RNA structure necessary for catalysis.

5.3 Labeling pre-mRNA and proteins for smFRET studies

The ability to site-specifically attach fluorescent dyes to proteins and nucleic acids is crucial for understanding their structure and function. SmFRET techniques are limited by our ability to conjugate specific fluorophores to either RNA or protein, yet retain their normal functionality, to study RNA-RNA, RNA-protein and protein-protein interactions. A handful of techniques have been developed to conjugate dyes to RNA molecules[109], but these are predominantly for end-labeling RNA molecules. In the case of RNA, we are primarily restricted to shorter lengths because of the ease of chemical synthesis and site-specific conjugation with dyes. Actin pre-mRNA, which has a 308-nts long intron, has been the choice for traditional yeast *in vitro* splicing experiments. In an effort to find a shorter and more easily synthesizable pre-mRNA, Ubc4 was identified as suitable for single molecule studies[37]. Ubc4 pre-mRNA has been predominantly used in the experiments described in this thesis. This pre-mRNA has the advantage of being short (90 nts intron) and yet splicing efficient. This allowed us to synthesize the pre-mRNA in two pieces and ligate them to obtain the full-length product. The individual pieces were conjugated with fluorophores at the desired splice site locations, upon which they were purified on 6% denaturing polyacrylamide gels (7 M urea). The fluorophore conjugated

pieces were then ligated using T4 RNA ligase 1 with a DNA splint. The ligated product was purified away from the remainder of unligated pieces by electrophoresis on a 6% polyacrylamide gel (7 M urea). The pre-mRNA band was then carefully excised from the gel and electroeluted on a home-built setup to finally extract it. We are still limited in terms of overall yield because of the cumulative loss in the multiple steps involved in the process, but can easily produce sufficient quantities for single molecule detection. Nevertheless, there is high need for new and more efficient labeling methods.

Fluorescent proteins have swept through world of biology. With the introduction of GFP[110] there has been a leap in the use of fusion protein to understand cellular processes. Many methods have been developed to introduce site-specific, covalently attached labels into proteins; however, each method has its own limitation[109]. With over 80 proteins making up the spliceosome in yeast, there is no dearth of potential protein labeling targets. The experiments presented in Chapter 2 on a labeled single-cysteine mutant of the Cwc25 protein is an example of the questions that can be addressed using labeled proteins. To the best of our knowledge, Cwc25-pre-mRNA FRET experiments are the first of their kind since there has been no other RNA-protein FRET study in the splicing field. The results point to the idea that Cwc25 binds close to the BP, modulating the kinetics required for favorable first-step splicing. There has been a recent advance that eliminates the requirement to recombinantly express and purify proteins for labeling[77]. The desired advantage is that crude whole cell yeast extracts can be incubated with labeling reagent(s) and yet be able to achieve specific labeling of a single protein. Similar techniques based on inteins[111] are currently being developed in our lab. In order to meet the

burgeoning demand in fluorescent microscopy, it is imperative that further advances are made in the development of new, easy and efficient labeling strategies.

5.4 A possible mechanism of action for spliceosomal helicases

The work presented in this thesis spans three important DExD/H box helicases, Prp2 (Chapter 2), Prp28 (Chapter 4) and Prp16 (Appendix A), which bring about significant conformational rearrangements for catalytic activation of the spliceosome. RNA helicases in general are characterized by an NTP binding domain and usually use the energy derived from hydrolysis to unwind dsRNA. However *in vitro* strand displacement activity is yet to be demonstrated for most of the spliceosome helicases whose direct targets are largely unknown. Our results indicate that Prp2 with NTP and protein cofactor Spp2 rearranges the spliceosome-substrate complex to reversibly explore conformations with proximal 5'SS and BP that accommodate chemistry. Similarly, experiments on Prp28 show that the pre-mRNA at this stage toggles between two FRET conformations and that Prp28 may drive the equilibrium towards a higher FRET state. We thus draw analogy to the ribosome and find the spliceosome to be a biased Brownian ratchet machine. In the Brownian motor mechanism, stochastic thermal fluctuations in a system allow for a factor to bias the directedness of the process. We believe that such a toggling mechanism may be prevalent throughout the spliceosome[94]. In the case of Prp2, a stable FRET state seen in the B^{act} pre-mRNA is unlocked by the helicase Prp2, while in the case of Prp28, the equilibrium between two FRET states in the CC1 complex is biased towards the higher FRET state corresponding to CC2. Thus, the spliceosome in principle can toggle between active and inactive, large-scale conformations at critical steps in splicing. Although we have no

direct evidence yet that this may be the case with other helicases, we believe that a similar mechanism is likely operating to facilitate the second step of splicing and that this theme might be recurrent in the function of DExD/H box helicases. Interestingly, much like proofreading in the ribosome, the toggling seen in here might be the mechanism by which proofreading is accomplished in the spliceosome, where each conformational selection by toggling is separated by an irreversible step leading to increased commitment as we progress into the splicing cycle. This scenario presents interdependent opportunities and challenges for structural and dynamics studies in the spliceosome and other biomolecular machines.

Cell biology is at its most exciting time now. With the development of new and improved microscopy techniques and instrumentation, we are able to push the limits of resolution and 'see' things that were invisible previously. In all likelihood, single molecule methods will continue to expand and provide valuable answers to important biological questions in the future.

APPENDIX A: SiMPull-FRET to study pre-mRNA dynamics at the second step of splicing

A.1 Reconstitution of the second step of splicing using SiMPull-FRET

The second step of pre-mRNA splicing involves a transesterification reaction where the 3' hydroxyl group of the 5'SS attacks the 3'SS to splice the exons together and to remove the intron as a lariat intermediate. In Chapter 2, the roles played by Prp2, Spp2 and Cwc25 in facilitating the first step of splicing were shown to involve the remodeling of the B^{act} complex by ATPase action of Prp2 and the stabilization of the necessary first-step conformation by Cwc25. With the *in vitro* reconstitution of the second step of splicing, it was demonstrated that the essential proteins for this step are Prp16, Prp18 and Slu7[35,100]. The helicase Prp16 is involved in the removal of Cwc25 and Yju2 from the BP and remodeling of the C complex[32,100,101,112]. For the second step of splicing, the 3'SS needs to be positioned at the catalytic site. This is achieved by the ATP-independent action of the Prp18-Slu7 heterodimeric complex[113]. The primary role for the helicase Prp22 lies in the ATP-driven disassembly of the spliceosome, with an additional ATP-independent role in assisting step 2 catalysis[114].

The nucleotide distance between the BP and 3'SS of the pre-mRNA has been shown to play a crucial role in which second step factors are required[100]. For a truncated actin pre-mRNA substrate, where this distance is as short as 7 nucleotides (nts), only Prp16 and ATP are required,

while when the distance is 38 nts, Prp16, ATP, Prp18 and Slu7 are all required for efficient product formation[100]. In order to investigate the role played by these factor(s) in restructuring the pre-mRNA to bring the 3'SS and 5'SS close together, we decided to use our preferred smFRET substrate, Ubc4 pre-mRNA, with fluorophores near the 3'SS and 5'SS. As discussed in Chapter 2, SiMPull-FRET can be effectively utilized to assign specific pre-mRNA dynamics to specific proteins. Using Ubc4 pre-mRNA and heat-inactivatable Prp2 extract (ΔPrp2, Cef-1-TAP), we developed a system to reconstitute both steps of splicing starting from the stalled B^{act} complex. Immunopurification of the B^{act} complex was achieved as described in Chapter 2, by flowing the assembled splicing reaction onto streptavidin coated magnetic beads (Invitrogen) that had been pre-incubated with biotinylated IgG. The TAP-tag on the Cef-1 protein binds the protein A domain on the immobilized IgG. The complex is then washed under stringent salt conditions to obtain a pure complex devoid of any non-specifically bound proteins. The pre-mRNA in the B^{act} complex alone does not undergo splicing, as expected (Figure A.1, lane 1). Addition of first-step factors Prp2, Spp2 and Cwc25 along with ATP leads to efficient first-step products as seen in Figure A.1, lane 2. Addition of Prp16 to this complex yielded second-step products albeit at a lower efficiency (Figure A.1, lane 3). Furthermore, addition of Prp18 and Slu7 increased the second step products by a small factor (Figure A.2, lane 4). Quantification of the free 5' exon released shows that's there is a slight reduction with the addition of Prp18, Slu7 with respect to the pre-mRNA, possibly due to its incorporation into the mature mRNA. Due to high background signal, an accurate value for this enhancement could not be reported here.

Prp22 has been shown to be not required for second-step reconstitution with actin pre-mRNA. This seems to be the case with Ubc4 pre-mRNA here also. The lower splicing efficiency

with the second step may be in part due to the stepwise addition of second step factors. Our results show that for Ubc4, where the distance between the BP and 3'SS is 23 nts, Prp16 seems largely sufficient to drive second-step catalysis, while Prp18 and Slu7 only offer a slight enhancement, less than what is seen with Cwc25 in the first step. This immunopurified and reconstituted system can be used in future SiMPull-FRET experiments to address several pertinent questions in the field of splicing. We believe that, like the first step of splicing, the second step may also exhibit thermal ratchet-like motion that is locked by a factor in the active second-step conformation. To test this, the B^{act} complex needs to be chased as described in Chapter 2 to go through the first step of splicing, after which second-step factors can be added one at a time to observe what conformational changes each induces in Ubc4 pre-mRNA with fluorescent labels near the 3'SS and 5'SS. Additionally, the role played by Prp16 in remodeling the spliceosome such that Cwc25 is now displaced from the BP can be explored with fluorescently labeled Cwc25 described in Chapter 2. A hypothetical model by which the pre-mRNA explores two alternative routes to get through step 2 catalysis with the remove Cwc25 to is shown in Figure A.3. Our results in Chapter 2 show that in the C complex the 5'SS and BP explore stable high FRET states. This could mean that the 3'SS is yet to get coordinated into the catalytic center, positioning the 5' and 3' exon distal from each other (low FRET state in H^C complex as shown in Figure A.3. In pre-mRNA with a short BP-3'SS distance, Prp16 mediated remodeling weakens the binding of Cwc25 to the BP, giving the intron loop the flexibility to move out of the catalytic center. This leads to some second-step product formation in substrates with short BP-3'SS distance (Step 2* complex). In pre-mRNA with longer BP-3'SS distance, removal of Cwc25 is enhanced by Prp18-$Slu7^{100}$, which when bound also stabilize the second-

step conformation and product formation (Step 2 complex). Binding of Cwc25 does not necessarily mean reversal of second step chemistry, since such a role for Cwc25 has not been shown. Removal of Cwc25 is necessary for the spliceosome to adopt second step conformation and subsequently catalysis. This model can be tested in future SiMPull-FRET experiments.

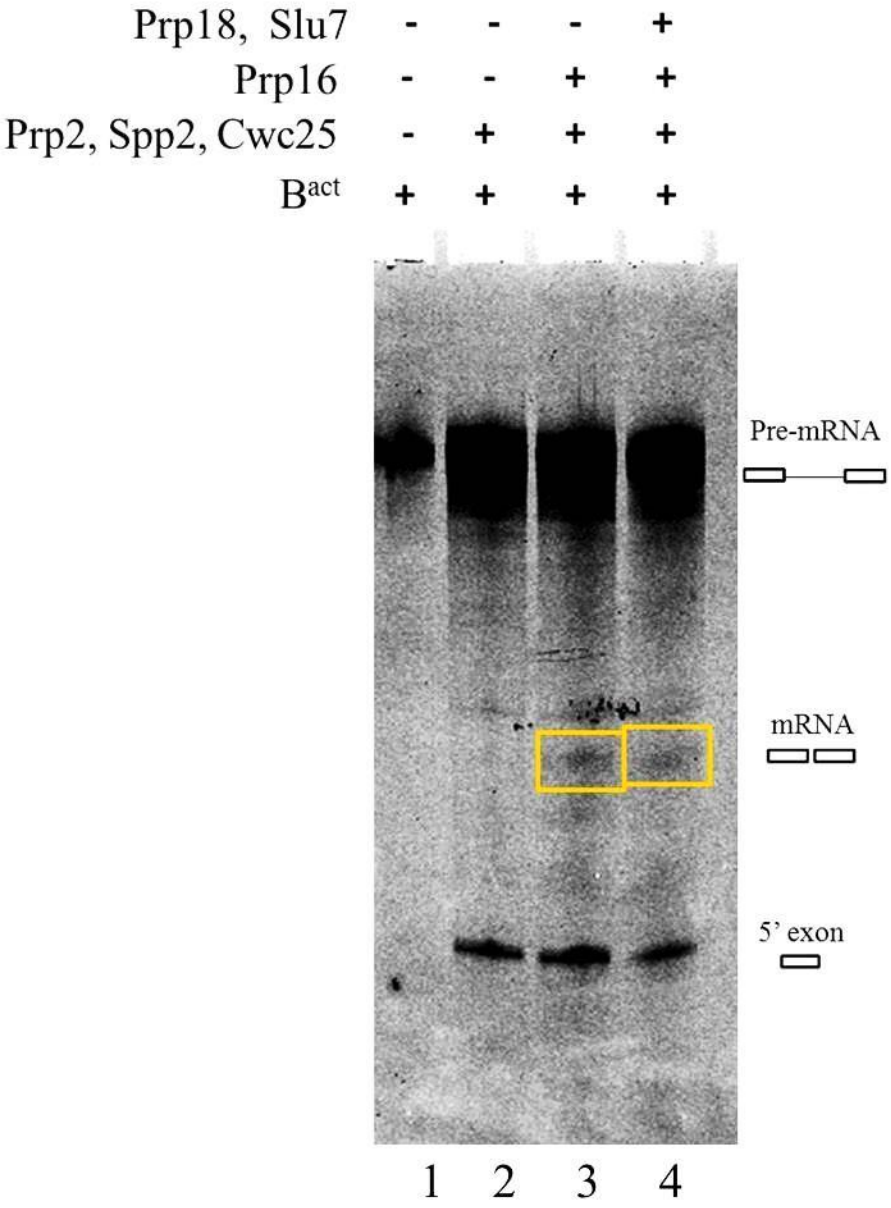


Figure A.1 Immunopurification of the Bact complex and reconstruction of the second step of splicing

Reconstitution of second step of splicing with immunopurified B^{act} complexes analyzed on a 15% polyacrylamide, 7 M urea gel. Immunopurified B^{act} complex (lane 1) is incubated with first-step factors Prp2, Spp2, Cwc25 and ATP (lane 2) alone, or additionally with Prp 16 and ATP (lane 3), or additionally with Prp16, Prp18 and Slu7 and ATP (lane 4).

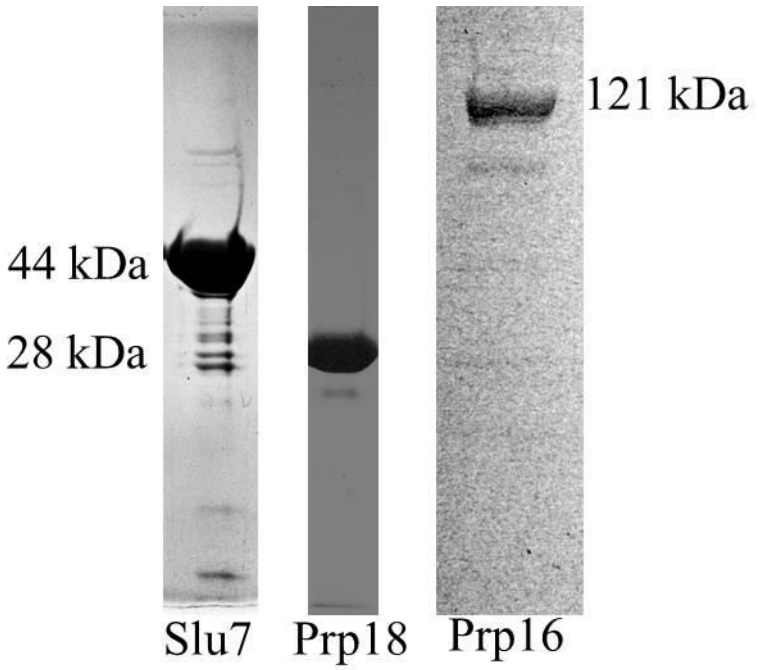


Figure A.2 Recombinant expression of second-step splicing proteins Prp16, Prp18 and Slu7

The second-step proteins Prp16, Prp18 and Slu 7 were expressed, purified and analyzed on 10% SDS PAGE gels.

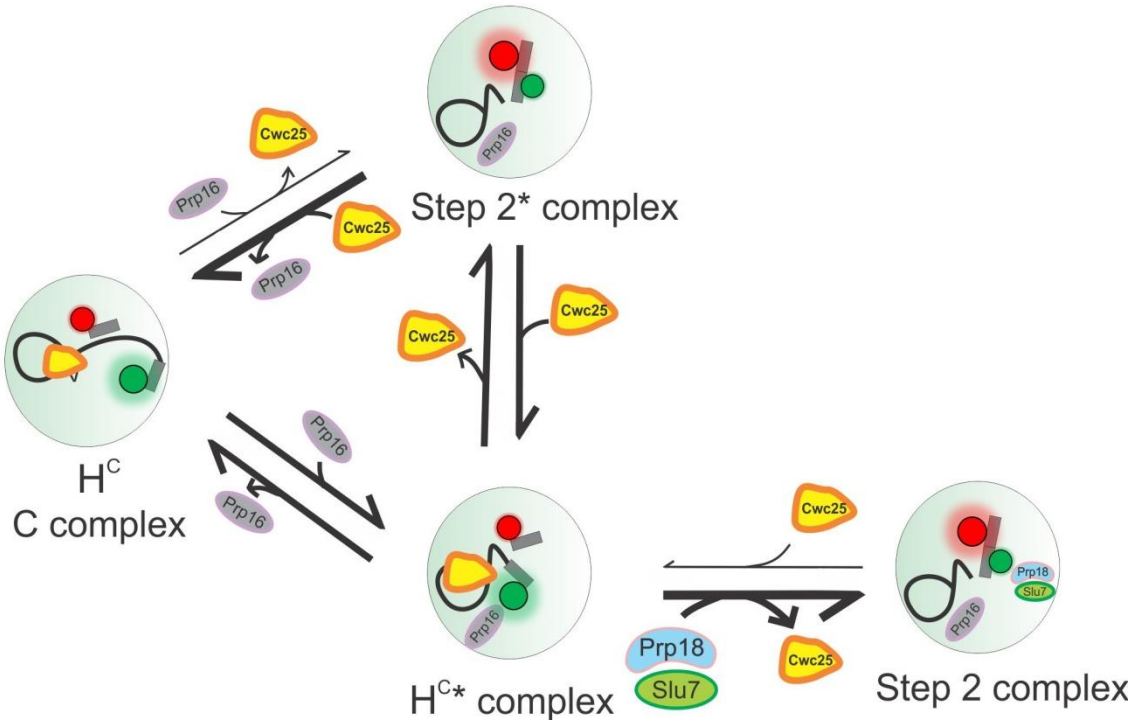


Figure A.3 Model for the mechanism of second step catalysis via pre-mRNA remodeling

The 3' and 5' exons of the pre-mRNA at the C complex are positioned farther apart from each other in a stable low-FRET state. Much like Prp2, the helicase Prp16 unlocks this state, remodeling the pre-mRNA, thereby allowing the exons to now sample close conformational proximity and second step product formation (Step 2* complex). For the efficient removal of Cwc25, Prp18-Slu7 heterodimer binds this complex, helping the displacement of Cwc25 thereby stabilizing the step 2 conformation and enhancing second step chemistry (Step 2 complex).

A.2 Materials and Methods

Expression and purification of Prp16, Prp18 and Slu7

Prp16 was cloned into a pET22b plasmid with a C-terminal hexa-histidine tag, Slu7 was subcloned from pGEX6P1 into pRsetA plasmid to carry a C-terminal hexa-histidine tag, and Prp18 was cloned into a pETM11 plasmid with an N-terminal hexa-histidine tag. The constructs were then transformed into Escherichia coli strain Rossetta II (Novagen). Cultures were grown in 2-4 l of TB medium and induced with 125 μM IPTG. Cultures were then incubated at 20 ºC for 18 h. Cells were harvested, washed, and the pellets stored at -80 ºC. Purification was performed as described[35]. Protein purity was confirmed by 10% SDS-PAGE (Figure A.2) and proteins were aliquoted, flash frozen in liquid nitrogen, and stored at -80 ºC. Protein concentrations were determined by Bradford assay and measurement at A_{280}.

Immunopurification and second-step reconstitution

Extracts prepared from *prp2-1 cef-1-TAP* yeast strain were heated at 37 °C for 40 min to inactivate Prp2 and stall the spliceosome at the B^{act} complex. In a final volume of 135 μl, 40% (v/v) of this heat treated extract was incubated with ~50 pmoles 3'SS and 5'SS labeled Ubc4 pre-mRNA in the presence of 2 mM ATP in splicing buffer (8 mM HEPES-KOH, pH 7.0, 2 mM $MgCl_2$, 0.08 mM EDTA, 60 mM $K_i(PO_4)$, 20 mM KCl, 8% (v/v) glycerol, 3% (w/v) PEG, 0.5 mM DTT) and incubated at 23 °C for 35 min. Streptavidin-coated magnetic beads were incubated with an equal volume of 0.5 mg/ml IgG-biotin in T50 in a tube rotator at RT for 30 min. The beads were then pulled down using a magnet and the supernatant was discarded. Free

biotin was added to block any streptavidin molecules not yet bound to biotin as described in Chapter 2. After equilibration in splicing buffer, the independently assembled splicing reaction was added and incubated in a tube rotator for 30 min at RT, followed by washes with buffer A and splicing buffer to further purify the B^{act} complex (see Materials and Methods of Chapter 2). The reactions are scaled up for reconstitution reactions pursued in parallel and split at this step. One part of this B^{act} complex was analyzed on a 15% polyacrylamide gel (Figure A.1, lane 1). To the other part, Prp2, Spp2 and Cwc25 were added at 90-120 nM final concentration in splicing buffer in the presence or absence of 2 mM ATP and incubated in the tube rotator for 30-40 min. At this step, the reaction was split into two aliquots, one of which was loaded into lane 2, whereas to the other Prp16 at 80 nM concentration was added in splicing buffer with 2 mM ATP and incubated for 30 min. The reaction was once again split into two aliquots, one was loaded into lane 3, whereas to the other Prp18 and Slu7 were added at each ~100 nM final concentration in splicing buffer and incubated in the tube rotor for 30 min (lane 4). To visualize the products of splicing, RNA was isolated and analyzed on a denaturing, 7 M urea, 15% polyacrylamide gel and scanned on a Typhoon variable mode imager (GE Healthcare). Shown in Figure A.1 is the Cy5 scan.

APPENDIX B: Utilizing fluorescently labeled snRNAs to characterize the spliceosome

B.1 An *in vitro* splicing system with fluorescently labeled snRNAs

The spliceosome is made up of five subunits, each consisting of a single uridine-rich small nuclear RNA (snRNA). Assembly occurs in a stepwise manner on the pre-mRNA substrate[77]. Various rearrangements and conformational changes take place to finally form the catalytically active core of the spliceosome, which consists of the U2, U5 and U6 snRNA:protein (snRNP) complexes. During the course of this assembly, pre-mRNA:snRNA interactions are disrupted and new ones are created. Despite extensive research in the field, very little is known about the timing of snRNA assembly and the pre-mRNA conformational changes accompanying it. By establishing a splicing system where the endogenous snRNAs are replaced with fluorescent snRNAs, we can in real-time monitor several critical assembly and catalytic steps of splicing. Towards this goal, U2, U5 and U6 snRNA were transcribed and fluorescently labeled at the 3' end by coupling with hydrazide derivative of Cy3, Cy5, Alexa488 or TAMRA (Figure B.1). Each of the snRNAs could be depleted independently from the wildtype yeast extract using previously published complementary DNA oligonucleotides (Chapter 3 and Table B.1). The snRNA-depleted samples were then tested for *in vitro* splicing by incubating the depleted extract with radiolabeled pre-mRNA for 30 min and visualizing the product on a 15% PAGE. Since

snRNAs are essential in the extract for pre-mRNA splicing, upon depletions using complementary DNA oligonucleotides pre-mRNA splicing was inhibited, as expected (Figure B.2). To these inhibited reactions *in vitro* transcribed and fluorophore labeled snRNAs were added; we were able to restore splicing for U2 snRNA (Figure B.2). Once the conditions for depletion and reconstitution for U2, U5 and U6 snRNAs were established, we performed smFRET experiments on the system to address specific questions.

To characterize splice site proximity and the role played by specific snRNAs with respect to it, we performed what we term a coincidence analysis to observe pre-mRNA dynamics and correlate it with snRNA binding to pre-mRNA. Binding of U2 snRNA is one of the first ATP dependent steps in the spliceosome assembly process. We first imaged Ubc4 wildtype pre-mRNA with Cy3 near the intron, and Cy5 near the BP (Chapters 2 and 3) in the presence of yeast whole cell extract depleted of U2 snRNA (to monitor pre-snRNA assembly dynamics). We then flowed on whole cell yeast extract reconstituted with U2 snRNA labeled with Cy3 fluorophore at its 3'end (to monitor post-snRNA assembly dynamics) (Figure B.3). U2 snRNA was labeled at the 3'end because the truncated Δ107 U2 snRNA construct used in these experiments increases the proximity between the 3' end of U2 snRNA and the BP of the pre-mRNA, while it lacks the extensive base pairing with U6 snRNA that a full-length 5' end has. In a converse experiment, we were able to observe binding of Cy5-labeled U2 snRNA on the pre-mRNA with a single Cy3 label near the BP when excited with both red and green lasers (Figure B.4).

The system presented here has wide applicability. Kinetic information on snRNA assembly on the pre-mRNA can be obtained using a similar experimental strategy. Interactions between U2:U6 snRNAs are restructured several times during activation and catalysis. Using a

fluorescently labeled U2 and U6 snRNA assembled on 5'-biotinylated pre-mRNA in the depletion/reconstitution system described, the structural rearrangements between these snRNAs to form the catalytic core can be investigated. Labeling U2 and U6 snRNAs, components of the only two snRNPs thought to be involved in catalysis, can be used to gain insights into the much-debated penta- versus tri-snRNP assembly pathway[115],[116] by visualizing the time-dependent assembly of U2 and U6 snRNA on the pre-mRNA substrate. The experiments described here merely serve as examples of the wide range of questions that can be addressed.

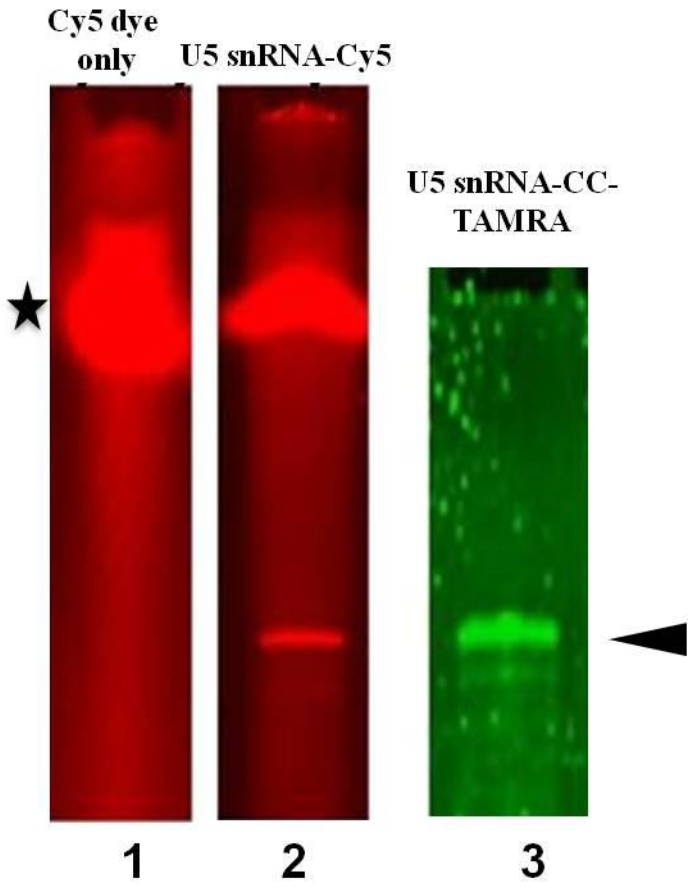


Figure B.1 Fluorescent labeling of U5snRNA via 3'-periodate oxidation

15% polyacrylamide, 7 M urea gel showing fluorescently labeled U5 snRNA. From left to right, lane 1 is free Cy5 dye (marked with a black star), lane 2 shows Cy5 labeled U5snRNA and lane 3 shows U5 snRNA labeled with TAMRA fluorophore (black arrow head indicates snRNA band).

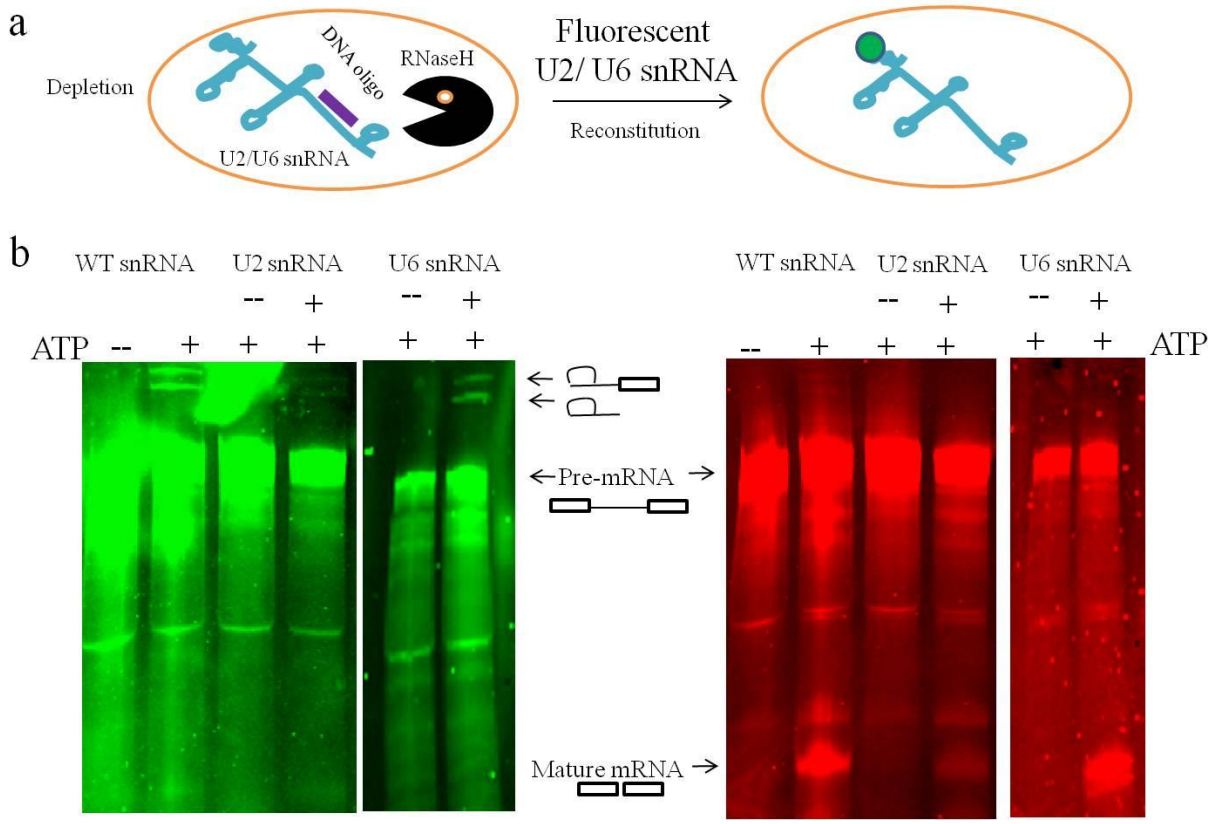


Figure B.2 Reconstitution of snRNA depleted extracts with fluorescently labeled snRNA.

(a) Schematic showing RNaseH mediated degradation of snRNA-DNA hybrid and reconstitution with Cy3 labeled snRNA (green circle). (b) 15% polyacrylamide, 7 M urea gel showing wildtype snRNA (WT snRNA, unmodified), U2 snRNA -, depleted and +, reconstituted with Cy3-labeled U2 snRNA, U6snRNA -, depleted and +, reconstituted with Cy3-labeled U6snRNA in the presence or absence of ATP (+/-). Efficient splicing is achieved upon reconstitution with labeled snRNA as indicated by the splicing products mRNA, lariat intron and lariat.

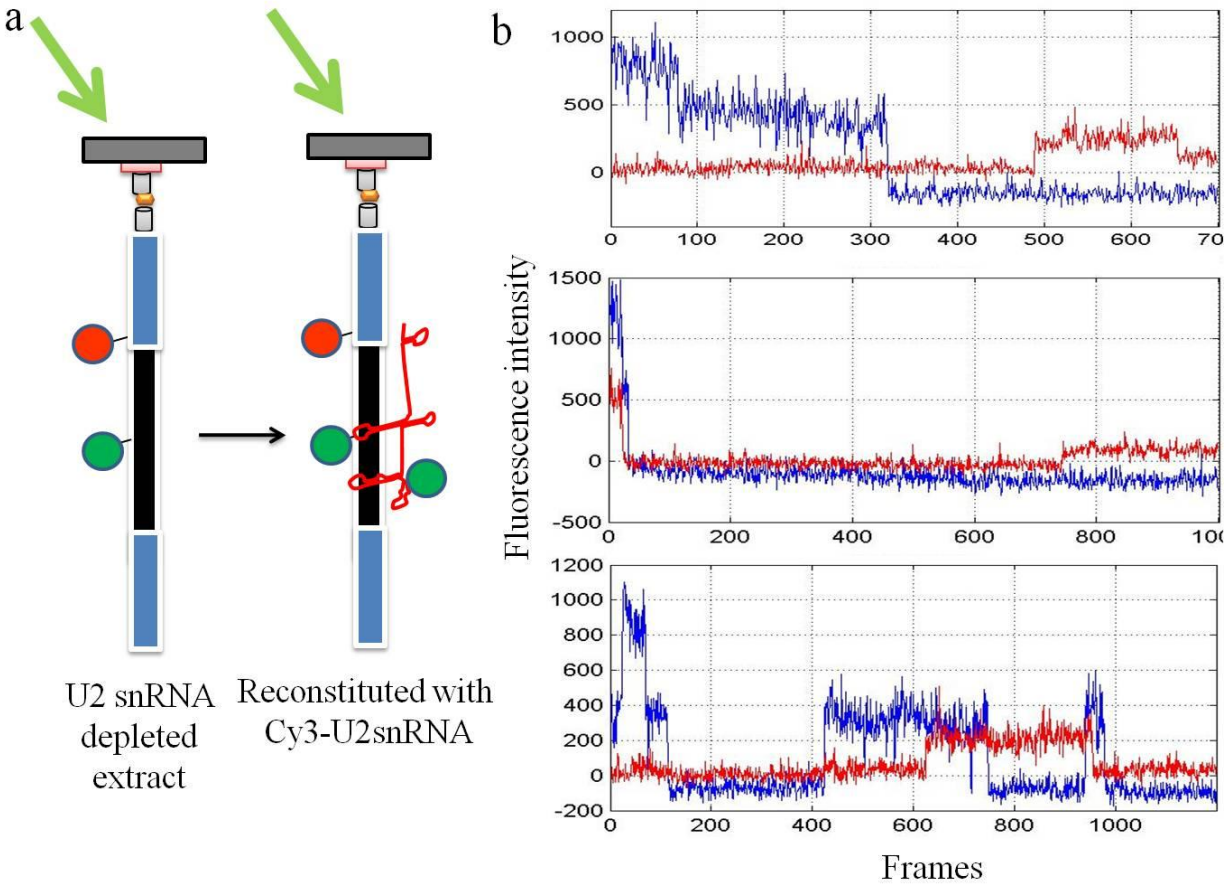


Figure B.3 Assembling Cy3-labeled U2 snRNA on doubly labeled (Cy3-Cy5) pre-mRNA.

(a) Schematic of the experimental setup showing immobilized pre-mRNA (with Cy3 near the BP and Cy5 near the 5'SS) on a slide surface in extract depleted of U2 snRNA and extract reconstituted with Cy3-U2 snRNA. (b) Representative smFRET trajectories from the experiment showing the assembly of Cy3-U2snRNA on pre-mRNA (blue, Cy3 signal; red, Cy5 signal). The presence of two Cy3 fluorophores (one on the U2 snRNA and the other on the BP of the pre-mRNA) is confirmed by the two-step photobleaching observed.

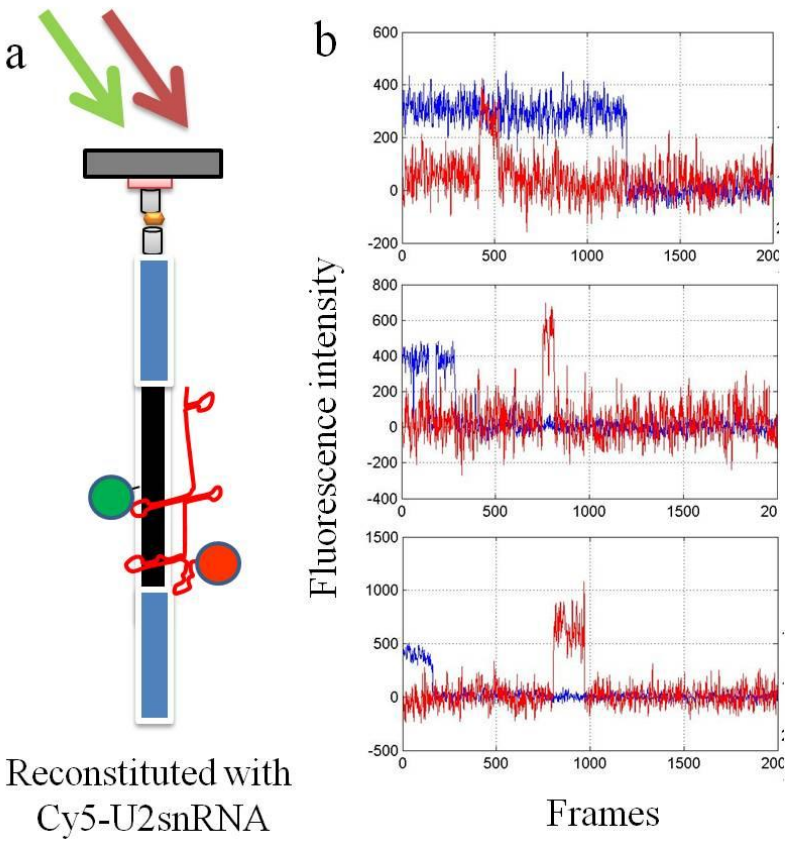


Figure B.4 Single molecule binding experiment with Cy5-U2 snRNA and Ubc4 pre-mRNA.

(a) Schematic of a binding experiment showing Cy5-labeled U2 snRNA binding to Ubc4 pre-mRNA labeled with a single Cy3 fluorophore. (b) Representative traces from this experiment indicating binding of Cy5-U2snRNA (red) on single labeled Cy3-Ubc4 (blue).

B.2 Materials and Methods

Plasmids used

Plasmids for U2, U5 and U6 snRNA (cloned in pUC18) were obtained from the Guthrie lab at the University of California, San Francisco. U2 and U6 snRNA plasmids have a T7 promoter upstream of the snRNA sequence. U5 snRNA plasmid was cloned into the pUC18 without the T7 promoter. The length of U2, U5 and U6 are 307, 220 and 120 nts, respectively.

PCR amplification and *in vitro* transcription

U5 snRNA was amplified from the plasmid using a forward primer with a T7 promoter and a reverse primer. The PCR product was analyzed on a 1% agarose gel and subsequently gel purified. This DNA template was then used directly for standard *in vitro* transcription reactions. U2 and U6 snRNA were linearized using restriction enzymes Afl1 and Dra1, respectively, with cleavage sites located at the 3' end of the snRNA sequences. The product was analyzed on a 1% agarose gel and subsequently gel eluted. Transcription reactions were then carried out on this template using home-made T7 polymerase and buffer. The transcription reaction were carried out for 8 h at 37 ºC and treated with DNase enzyme to digest the DNA template. The products of the transcription were run out on a 15%, 7 M urea PAGE, gel eluted using 1 mM EDTA and precipitated using 300 mM sodium acetate, pH 5.3, and 2.5 volumes of 100% ethanol.

Periodate oxidation and Cy3/Cy5/Alexa488-hydrazide coupling of snRNA

To fluorescently label the snRNAs, we used periodate oxidation of the 3' terminal ribose followed to couple the RNA with Cy3, Cy5 or Alexa488-hydrazide (GE life sciences) as described[109]. 50% (v/v) glycerol was added to the reaction mixture and the product was analyzed

on a 15%, 7 M urea PAGE. The bands were eluted using standard protocol (see above) and ethanol precipitated.

DNA oligonucleotides for U2 and U6 snRNA depletion

U2 and U6 snRNA were depleted using DNA oligonucleotides complementary to specific accessible regions. DNA oligonucleotide SRU2 was hybridized to U2 snRNA and DNA oligonucleotide D1 was hybridized to U6 snRNA. The sequence of SRU2 and D1 are shown in Table B.1. Wildtype yeast cell extracts were incubated with 450 nM of SRU2 and 300 nM of D1 independently for 30 min at 33 ºC in the presence of 3.3 mM ATP. Depletions were tested using the *in vitro* splicing reaction conditions described above and the products analyzed on a 15%, 7 M urea PAGE for splicing.

Reconstitution of splicing

Yeast cell extract depleted of endogenous U2 and U6 snRNA was reconstituted with 15 nM and 10 nM concentrations of *in vitro* transcribed and 3' fluorophore labeled U2 and U6 snRNA, respectively. The ability of the reconstituted system to direct pre-mRNA splicing was tested using the *in vitro* splicing reaction conditions described above and the products analyzed on a 15%, 7 M urea PAGE.

SRU2 (anti U2 snRNA)	5’-CAGATACTACACTTG-3’
D1 (anti-U6 snRNA)	5’-CATCTCTGTATTGTTTCAAATTGACCAA-3’

Table B.1 Sequence of DNA oligonucleotides for U2 and U6 snRNA depletions

PCR amplification, *in vitro* transcription and fluorophore coupling

U2 and U6 snRNA plasmids were linearized and *in vitro* transcriptions were performed using T7 polymerase to generate the respective snRNAs (data not shown). U5 snRNA plasmid was PCR amplified using a T7 promoter added to the forward primer. The product of the PCR reaction was used to obtain U5 snRNA via *in vitro* transcription. Periodate oxidation and coupling with the hydrazide derivative of the desired fluorophore yielded the labeled snRNAs shown in Figure B.3.

U2 and U6 snRNA depletion and reconstitution reactions

Splicing cell extracts were depleted of their endogenous U2 and U6 snRNAs using complementary DNA oligonucleotides as described (Chapter 3) and reconstituted with *in vitro* transcribed and fluorescently labeled U2 and U6 snRNA to reconstitute splicing as visualized by a standard *in vitro* splicing assay.

References

(1) Hershey, A. D.; Chase, M. *J Gen Physiol* **1952**, *36*, 39.
(2) Avery, O. T.; Macleod, C. M.; McCarty, M. *J Exp Med* **1944**, *79*, 137.
(3) Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 737.
(4) Wilkins, M. H. F.; Stokes, A. R.; Wilson, H. R. *Nature* **1953**, *171*, 738.
(5) Franklin, R. E.; Gosling, R. G. *Nature* **1953**, *171*, 740.
(6) Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 964.
(7) Brenner, S.; Jacob, F.; Meselson, M. *Nature* **1961**, *190*, 576.
(8) Cech, T. R. *Proc Natl Acad Sci U S A* **1986**, *83*, 4360.
(9) Baltimore, D. *Nature* **1970**, *226*, 1209.
(10) Temin, H. M.; Mizutani, S. *Nature* **1970**, *226*, 1211.
(11) Berget, S. M.; Moore, C.; Sharp, P. A. *Proc Natl Acad Sci U S A* **1977**, *74*, 3171.
(12) Gelinas, R. E.; Roberts, R. J. *Cell* **1977**, *11*, 533.
(13) Ganem, B. *Nature* **1987**, *328*, 676.
(14) Brody, E.; Abelson, J. *Science* **1985**, *228*, 963.
(15) Staley, J. P.; Guthrie, C. *Cell* **1998**, *92*, 315.
(16) Valadkhan, S.; Mohammadi, A.; Jaladat, Y.; Geisler, S. *Proc Natl Acad Sci U S A* **2009**, *106*, 11901.
(17) Lee, C.; Jaladat, Y.; Mohammadi, A.; Sharifi, A.; Geisler, S.; Valadkhan, S. *RNA* **2010**, *16*, 2226.
(18) Valadkhan, S. *RNA Biol* **2010**, *7*, 345.
(19) Horowitz, D. S. *RNA* **2011**, *17*, 551.
(20) Perriman, R.; Barta, I.; Voeltz, G. K.; Abelson, J.; Ares, M. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 13857.
(21) Tanner, N. K.; Linder, P. *Mol. Cell* **2001**, *8*, 251.
(22) Cheng, S. C.; Abelson, J. *Genes Dev* **1987**, *1*, 1014.
(23) Jain, A.; Liu, R.; Xiang, Y. K.; Ha, T. *Nat Protoc* **2012**, *7*, 445.
(24) Crawford, D. J.; Hoskins, A. A.; Friedman, L. J.; Gelles, J.; Moore, M. J. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 6783.
(25) Pyle, A. M. *Crit Rev Biochem Mol Biol* **2010**, *45*, 215.
(26) Abelson, J.; Trotta, C. R.; Li, H. *J Biol Chem* **1998**, *273*, 12685.
(27) Kistler, A. L.; Guthrie, C. *Gene Dev* **2001**, *15*, 42.
(28) Xu, Y. Z.; Newnham, C. M.; Kameoka, S.; Huang, T.; Konarska, M. M.; Query, C. C. *Embo Journal* **2004**, *23*, 376.
(29) Raghunathan, P. L.; Guthrie, C. *Curr Biol* **1998**, *8*, 847.
(30) Nguyen, T. H.; Li, J.; Galej, W. P.; Oshikane, H.; Newman, A. J.; Nagai, K. *Structure* **2013**.

(31) Lardelli, R. M.; Thompson, J. X.; Yates, J. R., 3rd; Stevens, S. W. *RNA* **2010**, *16*, 516.

(32) Tseng, C. K.; Liu, H. L.; Cheng, S. C. *Rna* **2011**, *17*, 145.

(33) Mayas, R. M.; Maita, H.; Staley, J. P. *Nature Structural & Molecular Biology* **2006**, *13*, 482.

(34) Koodathingal, P.; Novak, T.; Piccirilli, J. A.; Staley, J. P. *Mol Cell* **2010**, *39*, 385.

(35) Warkocki, Z.; Odenwalder, P.; Schmitzova, J.; Platzmann, F.; Stark, H.; Urlaub, H.; Ficner, R.; Fabrizio, P.; Luhrmann, R. *Nat. Struct. Mol. Biol.* **2009**, *16*, 1237.

(36) Cornish, P. V.; Ermolenko, D. N.; Noller, H. F.; Ha, T. *Mol Cell* **2008**, *30*, 578.

(37) Abelson, J.; Blanco, M.; Ditzler, M. A.; Fuller, F.; Aravamudhan, P.; Wood, M.; Villa, T.; Ryan, D. E.; Pleiss, J. A.; Maeder, C.; Guthrie, C.; Walter, N. G. *Nat. Struct. Mol. Biol.* **2010**, *17*, 504.

(38) Lee, T. H.; Blanchard, S. C.; Kim, H. D.; Puglisi, J. D.; Chu, S. *Proc Natl Acad Sci U S A* **2007**, *104*, 13661.

(39) Chen, J.; Tsai, A.; O'Leary, S. E.; Petrov, A.; Puglisi, J. D. *Curr Opin Struct Biol* **2012**, *22*, 804.

(40) Vijayraghavan, U.; Company, M.; Abelson, J. *Genes Dev* **1989**, *3*, 1206.

(41) Wahl, M. C.; Will, C. L.; Luhrmann, R. *Cell* **2009**, *136*, 701.

(42) Chen, J. H.; Lin, R. J. *Nucleic Acids Res.* **1990**, *18*, 6447.

(43) Ohrt, T.; Prior, M.; Dannenberg, J.; Odenwalder, P.; Dybkov, O.; Rasche, N.; Schmitzova, J.; Gregor, I.; Fabrizio, P.; Enderlein, J.; Luhrmann, R. *RNA* **2012**, *18*, 1244.

(44) Berglund, J. A.; Rosbash, M.; Schultz, S. C. *RNA* **2001**, *7*, 682.

(45) Kim, S. H.; Lin, R. J. *Mol. Cell. Biol.* **1996**, *16*, 6810.

(46) Chiu, Y. F.; Liu, Y. C.; Chiang, T. W.; Yeh, T. C.; Tseng, C. K.; Wu, N. Y.; Cheng, S. C. *Mol. Cell. Biol.* **2009**, *29*, 5671.

(47) Vijayraghavan, U.; Company, M.; Abelson, J. *Genes Dev.* **1989**, *3*, 1206.

(48) Edwalds-Gilbert, G.; Kim, D. H.; Kim, S. H.; Tseng, Y. H.; Yu, Y.; Lin, R. J. *RNA* **2000**, *6*, 1106.

(49) Blanco, M.; Walter, N. G. *Methods Enzymol.* **2010**, *472*, 153.

(50) Pereira, M. J. B.; Nikolova, E. N.; Hiley, S. L.; Jaikaran, D.; Collins, R. A.; Walter, N. G. *J. Mol. Biol.* **2008**, *382*, 496.

(51) Ditzler, M. A.; Rueda, D.; Mo, J. J.; Hakansson, K.; Walter, N. G. *Nucleic Acids Res.* **2008**, *36*, 7088.

(52) Sabanayagam, C. R.; Eid, J. S.; Meller, A. *J Chem Phys* **2005**, *122*, 061103.

(53) Cosa, G.; Harbron, E. J.; Zeng, Y.; Liu, H. W.; O'Connor, D. B.; Eta-Hosokawa, C.; Musier-Forsyth, K.; Barbara, P. F. *Biophys. J.* **2004**, *87*, 2759.

(54) Rueda, D.; Bokinsky, G.; Rhodes, M. M.; Rust, M. J.; Zhuang, X.; Walter, N. G. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 10066.

(55) Black, D. L. *RNA* **1995**, *1*, 763.

(56) Silverman, E. J.; Maeda, A.; Wei, J.; Smith, P.; Beggs, J. D.; Lin, R. J. *Mol. Cell. Biol.* **2004**, *24*, 10101.

(57) Roy, J.; Kim, K.; Maddock, J. R.; Anthony, J. G.; Woolford, J. L. *RNA* **1995**, *1*, 375.

(58) Teigelkamp, S.; McGarvey, M.; Plumpton, M.; Beggs, J. D. *EMBO J.* **1994**, *13*, 888.
(59) Liu, H. L.; Cheng, S. C. *Mol. Cell. Biol.* **2012**, *32*, 5056.
(60) Newnham, C. M.; Query, C. C. *RNA* **2001**, *7*, 1298.
(61) Schwer, B.; Gross, C. H. *EMBO J.* **1998**, *17*, 2086.
(62) Company, M.; Arenas, J.; Abelson, J. *Nature* **1991**, *349*, 487.
(63) Schwer, B. *Mol. Cell* **2008**, *30*, 743.
(64) Bartels, C.; Klatt, C.; Luhrmann, R.; Fabrizio, P. *Embo Rep* **2002**, *3*, 875.
(65) Kim, S. H. *EMBO J.* **1992**, *11*, 2319.
(66) Ohi, M. D.; Link, A. J.; Ren, L. P.; Jennings, J. L.; McDonald, W. H.; Gould, K. L. *Mol. Cell. Biol.* **2002**, *22*, 2011.
(67) Chen, H. C.; Tseng, C. K.; Tsai, R. T.; Chung, C. S.; Cheng, S. C. *Mol. Cell. Biol.* **2013**, *33*, 514.
(68) Jain, A.; Liu, R. J.; Ramani, B.; Arauz, E.; Ishitsuka, Y.; Ragunathan, K.; Park, J.; Chen, J.; Xiang, Y. K.; Ha, T. *Nature* **2011**, *473*, 484.
(69) Cordova, N. J.; Ermentrout, B.; Oster, G. F. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 339.
(70) Astumian, R. D. *Science* **1997**, *276*, 917.
(71) Spirin, A. S. *Biochemistry* **2009**, *48*, 10688.
(72) Frank, J.; Gonzalez, R. L., Jr. *Annu. Rev. Biochem.* **2010**, *79*, 381.
(73) Rodnina, M. V.; Wintermeyer, W. *Biochem. Soc. Trans.* **2011**, *39*, 658.
(74) Jankowsky, E. *Trends Biochem. Sci.* **2011**, *36*, 19.
(75) Robertson, K. L.; Yu, L.; Armitage, B. A.; Lopez, A. J.; Peteanu, L. A. *Biochemistry* **2006**, *45*, 6066.
(76) Crawford, D. J.; Hoskins, A. A.; Friedman, L. J.; Gelles, J.; Moore, M. J. *RNA* **2008**, *14*, 170.
(77) Hoskins, A. A.; Friedman, L. J.; Gallagher, S. S.; Crawford, D. J.; Anderson, E. G.; Wombacher, R.; Ramirez, N.; Cornish, V. W.; Gelles, J.; Moore, M. J. *Science* **2011**, *331*, 1289.
(78) McPheeters, D. S.; Fabrizio, P.; Abelson, J. *Genes Dev* **1989**, *3*, 2124.
(79) Fabrizio, P.; McPheeters, D. S.; Abelson, J. *Genes Dev* **1989**, *3*, 2137.
(80) Wang, Q.; Zhang, L.; Lynn, B.; Rymond, B. C. *Nucleic Acids Res* **2008**, *36*, 2787.
(81) Galej, W. P.; Oubridge, C.; Newman, A. J.; Nagai, K. *Nature* **2013**, *493*, 638.
(82) Cordin, O.; Beggs, J. D. *RNA Biol* **2013**, *10*, 83.
(83) Strauss, E. J.; Guthrie, C. *Nucleic Acids Res* **1994**, *22*, 3187.
(84) Ismaili, N.; Sha, M.; Gustafson, E. H.; Konarska, M. M. *RNA* **2001**, *7*, 182.
(85) Staley, J. P.; Guthrie, C. *Mol Cell* **1999**, *3*, 55.
(86) Yang, F.; Wang, X. Y.; Zhang, Z. M.; Pu, J.; Fan, Y. J.; Zhou, J. H.; Query, C. C.; Xu, Y. Z. *Nucleic Acids Research* **2013**, *41*, 4660.
(87) Seraphin, B.; Rosbash, M. *Cell* **1989**, *59*, 349.
(88) Seraphin, B.; Rosbash, M. *EMBO J* **1991**, *10*, 1209.
(89) Abovich, N.; Liao, X. C.; Rosbash, M. *Genes Dev* **1994**, *8*, 843.
(90) Rutz, B.; Seraphin, B. *RNA* **1999**, *5*, 819.

(91) Perriman, R.; Barta, I.; Voeltz, G. K.; Abelson, J.; Ares, M. *P Natl Acad Sci USA* **2003**, *100*, 13857.
(92) Kosowski, T. R.; Keys, H. R.; Quan, T. K.; Ruby, S. W. *Rna* **2009**, *15*, 1345.
(93) Ismaili, N.; Sha, M.; Gustafson, E. H.; Konarska, M. M. *Rna* **2001**, *7*, 182.
(94) Abelson, J. *Nat. Struct. Mol. Biol.* **2013**, *20*, 645.
(95) Stark, H.; Luhrmann, R. *Annu Rev Biophys Biomol Struct* **2006**, *35*, 435.
(96) Krummel, D. A. P.; Oubridge, C.; Leung, A. K. W.; Li, J.; Nagai, K. *Nature* **2009**, *458*, 475.
(97) Wolf, E.; Kastner, B.; Deckert, J.; Merz, C.; Stark, H.; Luhrmann, R. *EMBO J* **2009**, *28*, 2283.
(98) Kent, O. A.; MacMillan, A. M. *Nat Struct Biol* **2002**, *9*, 576.
(99) Donmez, G.; Hartmuth, K.; Kastner, B.; Will, C. L.; Luhrmann, R. *Molecular Cell* **2007**, *25*, 399.
(100) Ohrt, T.; Odenwalder, P.; Dannenberg, J.; Prior, M.; Warkocki, Z.; Schmitzova, J.; Karaduman, R.; Gregor, I.; Enderlein, J.; Fabrizio, P.; Luhrmann, R. *RNA* **2013**, *19*, 902.
(101) Tseng, C. K.; Liu, H. L.; Cheng, S. C. *RNA* **2011**, *17*, 145.
(102) Lin, Y.; Kielkopf, C. L. *Biochemistry* **2008**, *47*, 5503.
(103) MacMillan, A. M.; Query, C. C.; Allerson, C. R.; Chen, S.; Verdine, G. L.; Sharp, P. A. *Genes Dev* **1994**, *8*, 3008.
(104) Schwer, B. *Mol Cell* **2008**, *30*, 743.
(105) Altman, R. B.; Zheng, Q.; Zhou, Z.; Terry, D. S.; Warren, J. D.; Blanchard, S. C. *Nat Methods* **2012**, *9*, 428.
(106) Query, C. C.; Konarska, M. M. *Mol Cell* **2004**, *14*, 343.
(107) Ritchie, D. B.; Schellenberg, M. J.; Gesner, E. M.; Raithatha, S. A.; Stuart, D. T.; Macmillan, A. M. *Nat. Struct. Mol. Biol.* **2008**, *15*, 1199.
(108) Umen, J. G.; Guthrie, C. *Genetics* **1996**, *143*, 723.
(109) Qin, P. Z.; Pyle, A. M. *Methods* **1999**, *18*, 60.
(110) Prasher, D. C.; Eckenrode, V. K.; Ward, W. W.; Prendergast, F. G.; Cormier, M. J. *Gene* **1992**, *111*, 229.
(111) Ghosh, I.; Considine, N.; Maunus, E.; Sun, L.; Zhang, A.; Buswell, J.; Evans, T. C., Jr.; Xu, M. Q. *Methods Mol Biol* **2011**, *705*, 87.
(112) Schwer, B.; Guthrie, C. *EMBO J* **1992**, *11*, 5033.
(113) Aronova, A.; Bacikova, D.; Crotti, L. B.; Horowitz, D. S.; Schwer, B. *Rna* **2007**, *13*, 1437.
(114) Schwer, B. *Molecular Cell* **2008**, *30*, 743.
(115) Stevens, S. W.; Ryan, D. E.; Ge, H. Y.; Moore, R. E.; Young, M. K.; Lee, T. D.; Abelson, J. *Mol Cell* **2002**, *9*, 31.
(116) Johnston, N. *Scientist* **2002**, *16*, 30.