1-1-2010

Single Molecule Studies Of Spliceosomal Snrnas U2-U6

Zhuojun Guo
Wayne State University

SINGLE MOLECULE STUDIES OF SPLICEOSOMAL snRNAs U2-U6

by

ZHUOJUN GUO

DISSERTATION

Submitted to the Graduate School

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

2010

MAJOR: CHEMISTRY

Approved by:

_____

Advisor Date

_____

_____

_____

DEDICATION

To my parents and husband

ACKNOWLEDGMENTS

First, I want to thank my advisor, Dr. David Rueda, who gave me the opportunity to join his lab. During the years in his lab, I had a great environment that is not only to learn science but also life. This is a very precious time and opportunity for me to become mature. He is very helpful and knowledgeable and gives me great guidance in my research. I learned so much in Rueda lab and everything will be great treasure in my future life.

I would like to thank my committee members, Dr. Andrew Feig, Dr. H. Bernhard Schlegel and Dr. Athar Ansari, for very helpful discussion and suggestions on my thesis. I also would like to thank Dr. Samuel Butcher, Dr. David Brow and Ashley Richie at the University of Wisconsin-Madison for their collaboration and providing protein Prp24. Dr. Neocles Leontis and Anton Petrov at the Bowling Green University for collaboration and suggestions on designing the base triple mutations and help with use their web site of a database of RNA 3D structures using FR3D program. I thank all members of Rueda lab for their help and spending such a fantastic time with me. This is a great group. For this I would like to thank Dr. Amanda solem, Dr. Elvin Alemán, Rui Zhao, Rajan Lamichhane, Krishanthi Karunatilaka, Sharla Wood, May Daher, Gayan Senavirathne, Chandani Wanarsooriya, Hansini Mundigala, Bishnu Paudel, Kyle Vritis. Especially Amanda for help me correct my thesis and lots of helpful discussion and suggestions of my research; Rui for technical assistance of single molecule; Rajan for lots of help in my experiments and

numerous useful suggestions for everything and Chandani for the fun time working with me.

Finally I would like to thank my friends in Wayne State for lots of help and memories in these years, as well as my friends and family in China for their support and encouragements.

TABLE OF CONTENTS

Dedication ii

Acknowledgements iii

List of Tables vi

List of Figures vii

List of Abbreviations xi

Chapter 1 – Introduction 1

1.1 The splicing reaction 1

1.2 The group II intron 3

1.3 The spliceosome and its assembly 5

1.4 The role of the spliceosomal proteins in splicing 6

1.5 The catalytic core of spliceosome: protein or RNA? 9

1.6 Roles of the spliceosomal snRNAs U2/U6 in splicing 10

1.7 Structural arrangement of the U2/U6 snRNAs 12

1.8 Base triples may exist in the U2/U6 complex 14

1.9 The single molecule FRET method and application in the studying of the spliceosome 15

1.10 References 20

Chapter 2 – Single-molecule analysis of protein-free U2–U6 snRNAs reveals key Mg2+ dependent folding dynamics 41

2.1 Introduction 41

2.2 Materials and methods 43

2.3 Results and discussion 48
2.3.1 Mg2+ induces a large amplitude conformational change 48
2.3.2 Single-molecule FRET reveals three dynamic conformations .. 49
2.3.3 Mg2+ ions influence the U2–U6 structural dynamics 50
2.3.4 Dwell time analysis determines the rates of folding 51
2.3.5 Helix IB forms in the low FRET state 52
2.3.6 The structural dynamics correlate with splicing activation 53
2.4 Conclusions 54
2.5 References 56
Chapter 3 – Single Molecule Study of protein free U2/U6 Reveals Existence of Base Triples 75
3.1 Introduction 75
3.2 Materials and methods 81
3.3 Results 84
3.3.1 A base triple forms between the metal binding U80 bulge and the G86-C61 base pair 84
3.3.2 Cytosine scanning reveals nucleotides possibly involved in base triple formation 88
3.3.3. Base U88-A59•A53 may not be essential in stabilizing the high FRET structure 93
3.3.4 Proposed base triple U87•G60•G52 is invariant 96
3.4 Conclusions 102

3.5 References 107

Chapter 4 – Single Molecule Study of binding of spliceosomal protein Prp24 to snRNA U2/U6……113

4.1 Introduction 113

4.2 Materials and methods..... 116

4.3 Results and discussion 118

4.3.1 Prp24 binding to U2/U6 induces conformational change. 118

4.3.2 Single molecule FRET reveals Prp24 stabilize a 0.2 state 120

4.3.3 Prp24 N1234 affects the structural dynamics of the U2/U6 complex 123

4.4 Conclusions 126

4.5 References..... 130

Chapter 5 – Conclusions and future directions 135

Abstract 151

Autobiographical Statement 153

LIST OF TABLES

Table 1. Sequences of RNA used in chapter 2 49

Table 2. Sequences of RNA used in chapter 3 82

Table 3. Summary of smFRET results of hypothetic base triples 100

LIST OF FIGURES

Figure 1.1 Principles of RNA processing and splicing 2

Figure 1.2 The self-splicing group II intron 6

Figure 1.3 Assembly pathway of the spliceosome 8

Figure 1.4 Spliceosomal protein Prp24 11

Figure 1.5 U2/U6 snRNA may form the catalytic core of the spliceosome..... 15

Figure 1.6 Principles of single molecule FRET 22

Figure 1.7 Single molecule FRET study of U2/U6..... 26

Figure 2.1 Secondary structure model of the spliceosomal snRNAs U2–U6 .... .42

Figure 2.2 Formation of the fluorophore-labeled U2/U6 complex 51

Figure 2.3 Single-molecule FRET reveals a three-state folding pathway 55

Figure 2.4 Mg2+ dependence of the folding rate constants for the U2–U6 Complex 59

Figure 2.5 Dwell time analysis for the rate constants..... 60

Figure 2.6 Mutations in the AGC triad (red) show that helix IB forms only in the low FRET state..... 62

Figure 2.7 Six fold flipped-triad mutations and A59C mutation 63

Figure 2.8 Single molecule data for the construct with an intact GNRUA loop .. 69

Figure 3.1 Secondary structure model of the spliceosomal snRNAs U2–U6 with base triple 77

Figure 3.2 The recent crystal structure of the self-splicing Group II intron..... 79

Figure 3.3 smFRET histograms for the U80 mutations 84

Figure 3.4 Structure models according to FR3D 85

Figure 3.5 Cytosine scanning 89

Figure 3.6 Structure models according to FR3D 90

Figure 3.7 smFRET histograms for the mutations of nucleotides involved in base triple U88-A59•A53 93

Figure 3.8 smFRET histograms for the mutations of nucleotides involved in base triple U87•G60•G52 96

Figure 3.9 Structures of base triples according to FR3D 97

Figure 3.10 Revised folding pathway of U2/U6 104

Figure 4.1 Prp24 and its binding site to U2/U6 113

Figure 4.2 Bulk titration of Prp24 118

Figure 4.3 Effect of prp24 on the conformation of U2U6 complex 121

Figure 4.4 Single molecule histograms indicate that the prp24 N1234 can induce significant conformation changes of U2U6 complex..... ..123

Figure 4.5 Single molecule histograms indicate that the full-length prp24 can induce the same conformation changes of U2U6 complex as Prp24 N1234.....….124

Figure 4.6 Proposed the mechanism how Prp24 modulate the structural dynamics of the U2/U6 127

LIST OF ABBREBEVIATIONS

DV	Domain 5 of the group II intron
EBS	Exon-binding sites
IBS	Intron-binding sites
J2/3	Linker between the DII and DIII
RRM	RNA recognition motifs
snRNA	Small nuclear RNA
snRNP	Small nuclear ribonucleoproteins
NMR	Nuclear magnetic resonance
FRET	Fluorescence resonance energy transfer
ISL	Intramolecular stem loop
TIRF	Total internal reflection fluorescence
EM	Electron microscopy
smFRET	Single molecule FRET

CHAPTER 1

Introduction

1.1 The splicing reaction

Twenty three years ago striking research showed that in eukaryotes, genes are interrupted by non-coding sequences that were named introns[1,2]. Although introns do not encode any proteins, they are transcribed with the rest of the gene to form precursor mRNAs (pre-mRNAs)[3]. The introns must be precisely removed and the protein encoding sections, exons, need to be ligated together for translation[3] (Figure 1.1A). This incredible discovery led to the Nobel Prize in Physiology or Medicine in 1993 for Phillip Allen Sharp and Richard J. Roberts.

This process, in which the introns are removed and exons are joined to form the mature mRNA, is called splicing. In eukaryotes, introns contain consensus regions at three locations: GUAUGU at the 5' splice, UACUAAC at branch site and YAG at the 3' splice site. The splicing reaction consists of two transeaterification steps. During the first step of the splicing reaction, the 2' hydroxyl group of a highly conserved adenosine in the branch site attacks the 5' splice site, resulting a free 5' exon and a lariat intermediate. In the second catalytic step, the 3' hydroxyl group of the free 5' exon attacks the 3' splice site to ligate the two exons and remove the intron in a lariat form[4] (Figure 1.1B).

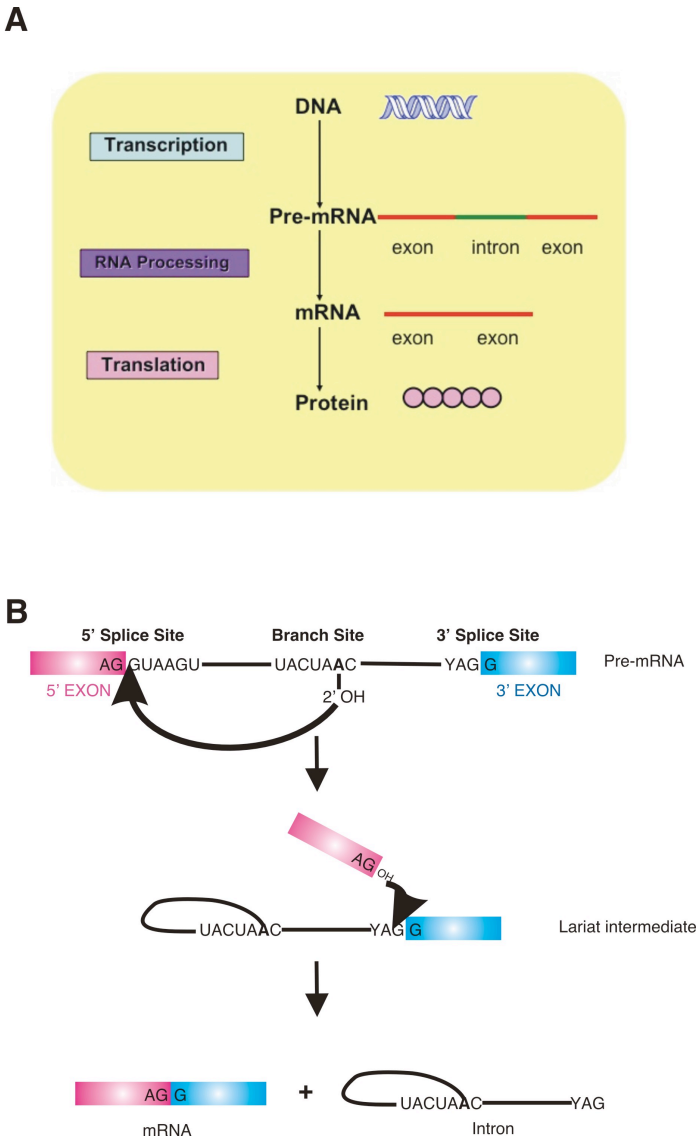


Figure 1.1 Principles of RNA processing and splicing (A) Central dogma of biology in eukaryotic cells. Genetic information stored in DNA flows to pre-mRNA by transcription. In eukaryotic cells the pre-mRNA contains both exons (Encoding protein) and introns (Not encoding protein). For protein translation introns must be removed and exons must be ligated together. This process is known as splicing. (**B**) 2-step splicing reaction[5,4]. During the first step of the splicing reaction, the 2' hydroxyl group of a highly conserved adenosine in the branch site attacks the 5' splice site, resulting a free 5' exon and a lariat intermediate. In the second catalytic step, the 3' hydroxyl group of the free 5' exon attacks the 3' splice site to ligate the two exons and remove the intron in a lariat form.

There are three major and well-defined types of introns: group I, group II and nuclear introns[6]. The group I and Group II are different in the mechanisms of their catalysis, but both of them are self-splicing, which means that the introns use the RNA component of the intron to catalyze their own splicing reaction and no protein enzyme is required. In contract, the splicing of the nuclear introns requires the spliceosome as the catalyst[7].

In cells, splicing is an essential step in the maturation of the pre-mRNA that is tightly regulated and maintained[4]. Splicing provides a mechanism for the cell to regulate gene expression by changing the level of the mature mRNA of a specific protein[8]. Furthermore, eukaryotic cells can achieve vast proteomic diversity via alternative splicing, in which a single pre-mRNA can be processed into multiple different mature mRNAs. Inaccurate or inefficient splicing leads to numerous human diseases, such as cancers and neurodegenerative disorders[9,10,11,12].

1.2 The group II intron

Among different types of introns, the group II intron is very interesting because it is self-splicing, and shares the same mechanism of reaction as nuclear intron splicing[13,7] (Figure 1.1A). The group II introns utilize RNA as the enzyme to catalyze their own splicing reaction. They are generally found in the bacteria, the mitochondrial or chloroplast genes of fungi, algae, yeast and plants. The splicing of the group II intron is very important for gene expression and regulation of variety of organisms[7, 14].

The secondary structure of the group II intron is comprised of six domains (DI-DVI) [14,15] (Figure 1.2A). The Domain 1 (DI) acts as the scaffold for the folding of other

parts of the intron. DI offers docking sites for other domains and determines the overall shape of the entire intron. It has two exon-binding sites (EBS1 and EBS2) that base-pair with the intron-binding sites (IBS1 and IBS2) in the 5'-exon respectively to position the exon and the catalytic center for the proper splicing reaction[16]. Domain 2 (DII) interacts with other regions of the intron via long-range interactions. Some of the nucleotides in the linker between the DII and DIII (J2/3) are highly conserved and important for the catalysis. Those nucleotides are components of the active site and play an important role in determining the specificity of the 3' splice-site[17,18]. The conserved loop in domain 3 (DIII) interacts with the intron core and catalytically crucial nucleotides in DI and DV[18a,19]. In the protein encoding group II introns, domain 4 (DIV) contains an the open reading frame (ORF) that encodes a protein which contributes to the intron mobility[7]. Domain 5 (DV) is the most phylogenetically conserved section of the intron, while the other domains are not highly conserved[20]. Deletion of DV is lethal for the function of this ribozyme, but catalysis can be rescued by adding DV in trans. These data suggest that this domain is vital for the catalysis of self-splicing and should be the catalytic core of the intron[21,22,23]. Domain V interacts with other parts of the intron intensively via tertiary interactions to position the essential components for catalysis properly. There are two conserved and crucial region in the domain 5: the DV bulge and the catalytic triad. The sequence of the DV bulge is usually AC. The catalytic triad is three consecutive nucleotides in the lower stem of DV, five base pairs away from the DV bulge (Figure 1.2B). This sequence is generally AGC, but is 5'-CGC for some species[24]. Domain V also binds functionally important Mg^{2+} ions in the bulge and in the catalytic triad since it is

highly possible that the self-splicing reaction utilize the two mental-ion mechanism[25,26]. Domain 6 (DVI) contains branch site adenosine residue that attacks the 5'-splice site in the first step of the spicing reaction. Domain 6 (DVI) is pivotal for branch site recognition and fidelity of the splicing reaction[27] (Figure1.1A and Figure 1.2A).

The recent high-resolution crystal structure of the group II intron from *Oceanobacillus ihyensis*, which is a eubacterium obtained from the seabed mud near the coast of Japan, shows that this big ribozyme folds into a very elaborate and compact three dimensional structure[28,29,30] (Figure 1.2C). The crystal structure agrees very well with previous biochemical and genetic studies, confirming most of the tertiary interactions predicted by other approaches[31].

Surprisingly, this giant ribozyme is modular, which means that some domains of the group II intron, for example DV, can be transcribed separately and reconstituted with other domains to perform the splicing reaction[32,33]. This domain modularity offers the opportunity for the group II intron to be developed into a variety of ribozymes[34,21, 35]. The group II intron is believed to be able to be divided into several functional fragments and reassembled *in trans* to perform catalysis. This indicates that it is highly possible that the group II intron is the molecular ancestor of the spliceosome, which catalyze nuclear pre-mRNA splicing[14].

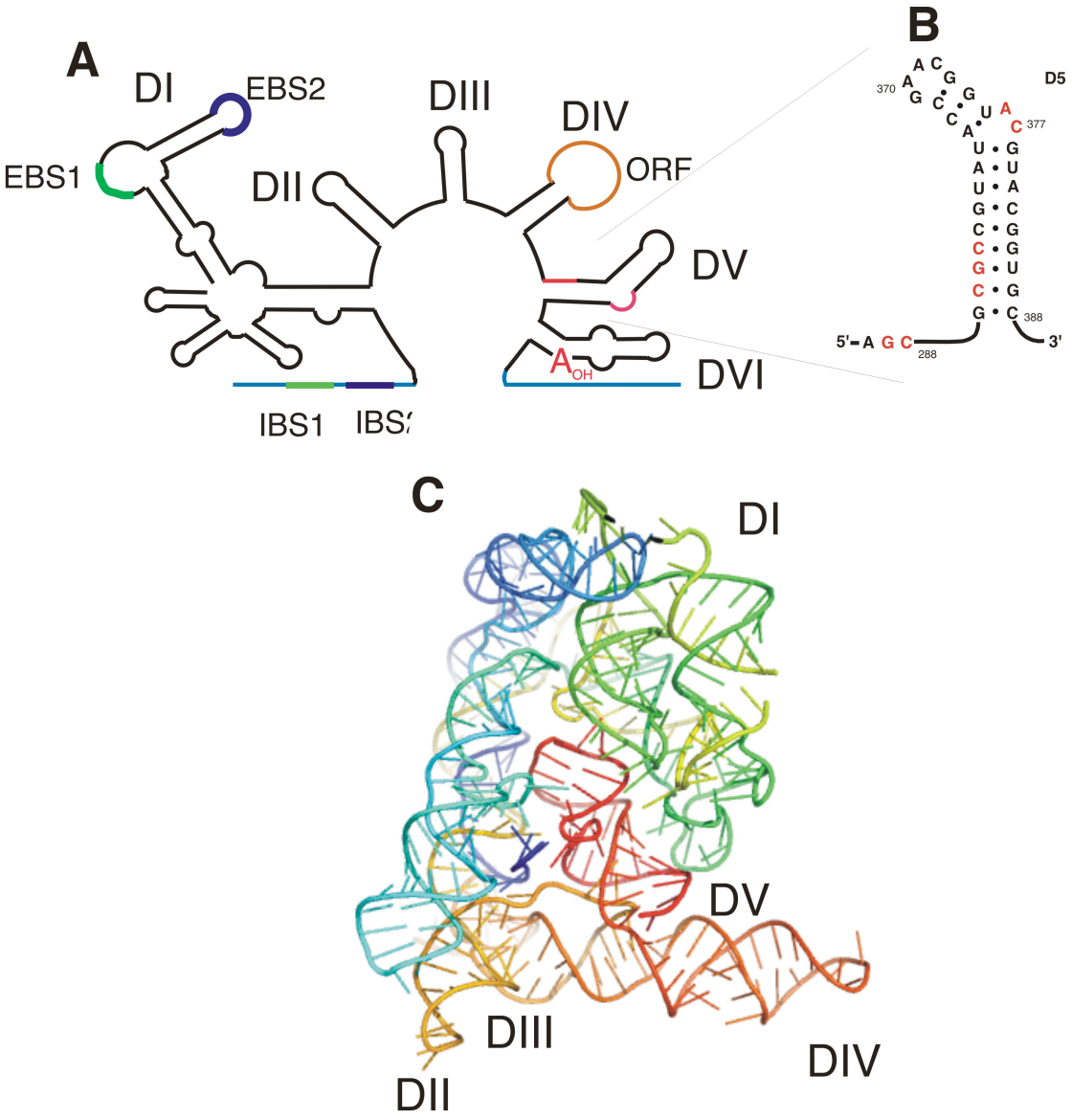


Figure 1.2 The self-splicing group II intron (A) Secondary structure of the group II intron[15,14]. The group II intron contains 6 domains (DI-DVI). It has two exon-binding sites (EBS1 and EBS2) that base pair with the intron-binding sites (IBS1 and IBS2) in the 5'-exon respectively to position the exon and the catalytic center for the proper splicing reaction. DVI contains the branch site. (**B**) Secondary structure of DV[15]. Highly conserved regions are highlighted in red. (**C**) Crystal structure of the group II intron with DI (Green, blue and purple), DII (dark yellow), DIII (Yellow), DIV (Orange) and DV (Red)[28,29].

1.3 The spliceosome and its assembly

Unlike the self-splicing group II intron, which is able to catalyze its own splicing reaction, the majority of introns are nuclear introns, which do not self-splice. The splicing reaction of nuclear pre-mRNAs must be catalyzed by the spliceosome. The spliceosome is a huge dynamic ribonucleoprotein (RNP) machine, consisting of five small nuclear RNAs (snRNAs, U1, U2, U4, U5 and U6) and about 170 proteins in human[5]. Each spliceosomal snRNA is packed into a small nuclear ribonucleoprotein (snRNP) by specific binding of numerous proteins. During each round of the splicing reaction, the spliceosome is built anew on the pre-mRNA substrate following a highly ordered and stepwise pathway during where the structure of the spliceosome is constantly changing due to the binding and release of snRNPs[36] (Figure 1.3). At first, U1 snRNP recognizes and binds the 5' splice site through base -pairing. Then, U2 associates with the branch site, yielding the prespliceosome (complex A). Subsequently, the pre-formed U4/U6•U5 triple snRNP binds to form the precatalytic spliceosome (complex B). During the activation of the spliceosome, the U1 and U4 snRNPs are destabilized from the complex and the resulting activated spliceosome (complex B*) is ready for the catalysis. During the two-step splicing reaction, only U2, U6 and U5 snRNPs remain in the spliceosome and comprise of the catalytic core of the machinery. The U1 and U4 leave the spliceosome during the activation process, so it is impossible for them to participate in the catalysis[37]. It was demonstrated that the U5 is functional dispensable for both step of splicing[38]. So it is highly possible that with the help of proteins, it is U2 and U6 that form the active site of the spliceosome (Figure 1.4B). The assembly of the spliceosome allows

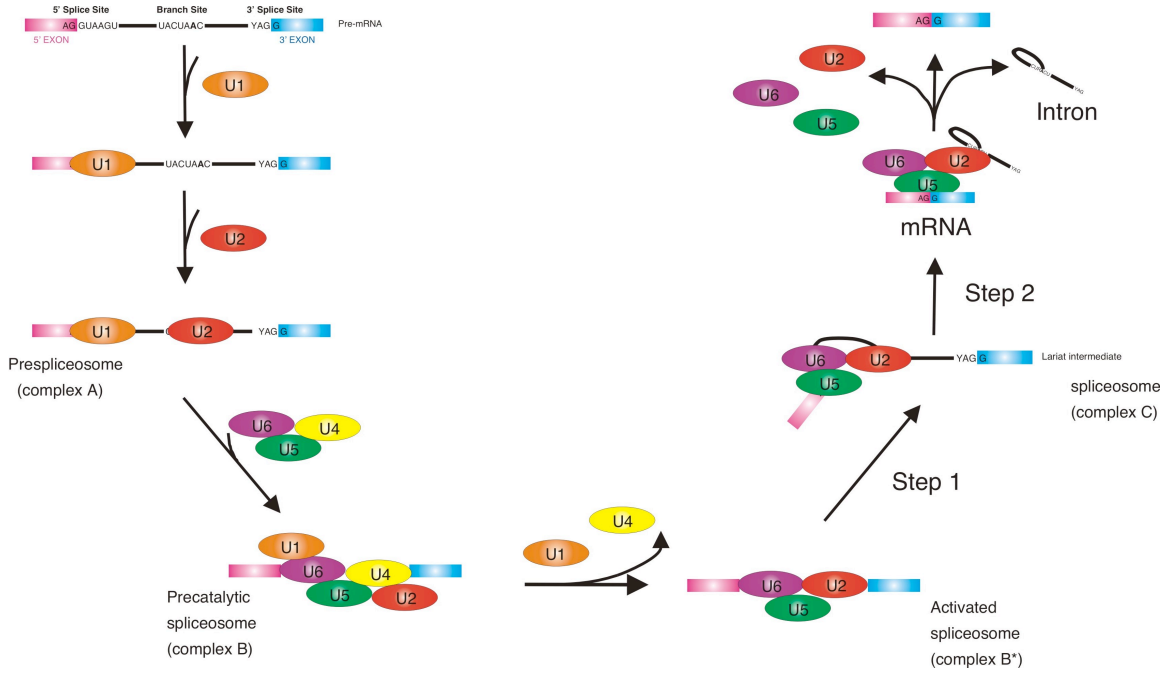


Figure 1.3 Assembly pathway of the spliceosome.[5] The spliceosome is a highly dynamic ribonucleoprotein complex, composed of 5 small nuclear RNAs (U1, U2, U4, U5 and U6 snRNAs) and about 80 proteins in yeast and 170 in human. Each snRNA binds a number of proteins and is packaged into small nuclear ribonucleoproteins (snRNPs). During the assembly, activation and catalysis of spliceosome, the snRNPs binds and disassociates following a highly ordered and regulated pathway. In the active spliceosome, only the U2, U6 and U5 snRNPs remain and the U5 is showed to be functionally dispensable for the catalysis *in vitro*, so it is highly possible that the U2 and U6 snRNAs form the catalytic core of the spliceosome and directly play a pivotal role in catalysis.

the multiple turnover reactions and the regulation of alternative splicing[5]. Strikingly, this assembly and activation pathway is conserved from yeast to humans[5].

1.4 The role of the spliceosomal proteins in splicing

1.4.1 Spliceosomal proteins play a significant role in splicing

The group II intron can catalyze the splicing reaction *in vitro* without any protein factors[14]. In contrast, protein free U2/U6 snRNAs can only catalyze a slow splicing-related reaction very inefficiently, indicating that spliceosomal proteins play an essential role in catalysis[39,40]. During the course of evolution, some functions of the catalytic RNA might have been transferred to proteins. In general, higher organisms have more proteins in the spliceosome[5]. The yeast spliceosome contains about 70 proteins, while in higher eukaryotes there are approximately 170 proteins[5]. Spliceosomal protein may participate in stabilizing the active structure, promoting RNA folding and remodeling, coordinating metal ion, binding substrates and substituting catalytically or structurally important metal ions[41]. Numerous studies showed that snRNAs form at least part of the active site of the spliceosome and play an important role in catalysis, but it is still not clear whether spliceosomal proteins are components of the active site as well. There are two proteins that are in or near the active site of the spliceosome. One is the U2-specific SF3b14a/p14 protein that interacts with the branch site adenosine[42] and the other one is the U5-specific Prp8 that interacts with the 5' splice site, the branch site and the 3' splice site[43,44,45]. These data lead to very interesting idea that both snRNAs and spliceosomal protein participate in the catalysis directly[46].

In spliceosomes of higher eukaryotes, the RNA:RNA interactions are weak, especially in humans[5,37], and proteins help to enhance the functional structures[5]. The recognition of splice sites and the branch site involves not only base paring interaction between snRNAs and the pre-mRNA but a combination of interaction involving both snRNA and proteins[47,48]. This offers a mechanism for the cell to regulate splicing and alternative splicing, during which multiple proteins can be produced using the same pre-mRNA[8]. There is a family of spliceosomal proteins named serine/arginine-rich proteins (SR proteins) in higher eukaryotes that perform alternative splicing regulation[5]. In contrast, in yeast, interactions between spliceosomal snRNAs themselves and the pre-mRNA are much stronger. The definition of the splice sites and the branch sites are achieved by perfect base pairing between them and corresponding snRNAs, so fewer proteins are required for this purpose[5]. Additionally, alternative splicing is very rare in the yeast, so there are no SR proteins required for the yeast spliceosome[5]. This may explain why there are fewer proteins in the yeast spliceosome than in that of humans. The rearrangements of the snRNAs require the assistance of DExD/H-type RNA-dependent ATPases/helicases and RNA chaperones that do not hydrolyze ATP. They help to unwind RNA helices, anneal duplexes and stabilize the active form of the RNA complexes[5,49].

1.4.2 Spliceosomal protein Prp24 promotes spliceosomal snRNA U6 rearrangements as an RNA chaperone

The spliceosomal protein Prp24 is an essential component of the U6 snRNP in yeast and helps U6 to remodel during assembly and catalysis of the spliceosome[50]. The details about the function and mechanism of Prp24 are still unclear. But it is well

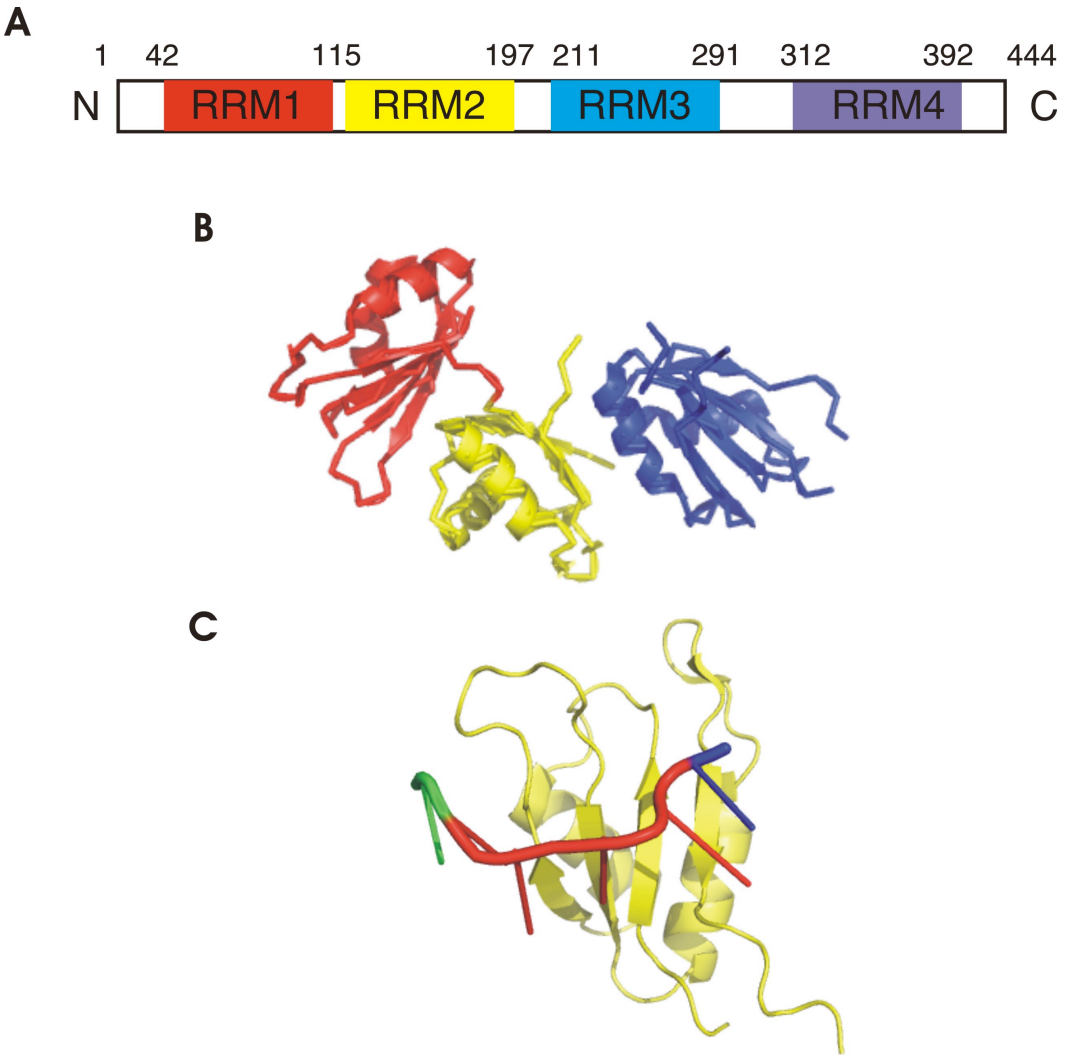


Figure 1.4 The spliceosomal protein Prp24 (A) Domain structure of Prp24 from S. cerevisiae. Prp24 contains 4 RNA recognition motifs (RRM)[51]. The four RRMs are shown in different colors. (**B**) Crystal structure of the first three RRMs of Prp24.RRM1 is in red, RRM2 in yellow and RRM3 in blue[51]. (**C**) NMR structure of RRM2 of Prp24 with a fragment of U6 (GAGA) bound to it. Prp24 RRM2 is in yellow and the GAGA sequence in red[52].

established that recombinant Prp24 promotes the annealing of the U4/U6 snRNA complex yielding U4/U6 bi-snRNP *in vitro*[53] (Figure 1.5C). However, recombinant Prp24 acts more efficiently in cell extract[54,53]. It may be that another member of the U6 snRNP, the Lsm protein complex, also helps the association of U4/U6 and works together with Prp24 to increase efficiency[55,56,57]. After the formation of the U4/U6 complex, Prp24 leaves the complex, but the Lsm proteins remain bound[58,59]. It has been proposed that Prp24 may return during the activation process of the spliceosome and help to unwind U4/U6 complex[54,60]. Prp24 does not hydrolyze ATP, so it is an RNA chaperone rather than an RNA heliase[53].

Prp24 form *Saccharomyces cerevisiae* has four RNA recognition motifs (RRMs)[51] (Figure 1.4A). Typically, RRMs preferentially bind single strand RNA[52]. Although a high-resolution structure of the whole Prp24 is not available now, the first three RRMs of Prp24 have been crystallized[51] (Figure 1.4B). Currently, the functions of each RRM are not well understood. Genetic studies indicate that RRMs 2 and 3 may stabilize the free U6 snRNA[60]. Gel shift binding assays demonstrated that RRMs 1 and 2 are important for the high-affinity binding of U6 snRNA and RRMs 3 and 4 may have other functions like controlling the stoichiometry of the Prp24 binding[50]. Biochemical studies with purified U6 snRNP and binding assays with recombined Prp24 and truncated U6 indicated that the binding site of Prp24 on U6 may be located near the conserved ACAGAGA loop (Nucleotides 40-58 and Figure 1.5B) [61,50, 58]. A more recent NMR study revealed that RRM2 binds the GAGA box in the conserved ACAGAGA loop of U6 in a sequence-specific fashion and RRM1 interacts with the phosphate backbone of 3’

downstream GAUCA sequence via electrostatic interactions, leaving the bases available for base pairing (Figure 1.4C). This can disrupt the internal helix of U6 and expose the bases for base paring[52], which is a very appealing model explaining the mechanism by which Prp24 promotes the annealing of the U4/U6 complex.

1.5 The catalytic core of spliceosome: protein or RNA?

Similarities between the group II intron and the spliceosome support RNA based-catalysis of the spliceosome[14]. First, splicing in both systems fellows the same chemistry and two-step reaction[19] (Figure 1.1B). Both of them utilize a 2'-OH group of an adenosine in the branch site as the nucleophile to attack the 5' splice site, named the branch mechanism. Both the nuclear introns and the majority of the group II intron are removed in the form of a branched lariat structure. Second, DV is similar to the intramolecular stem loop of U6 (U6 ISL) both in the structure and function[15] (Figure 1.5A). In both systems, the most conserved residues include a metal ion binding bulge (AC in DV and U80 in yeast U6 ISL) and a catalytic triad five base pairs from the bulge that are thought to form the active site[62,63].

These striking similarities lead to the hypothesis that the group II intron is the evolutionary ancestor of the spliceosome and the spliceosome is a ribozyme[6,64]. It has been hypothesized that the domain modularity of the group II intron allows it to be divided into small fragments that evolutes into the 5 snRNAs of the spliceosome[14]. It has been proposed that the U2/U6 snRNA is responsible for the catalysis[49]. The ability of the RNA to perform catalysis in the spliceosome was confirmed by the finding that a protein-free minimal U2/U6 complex was able to catalyze reactions similar to both steps of

splicing[40]. Mg^{2+} ion concentration as well as mutations in the branch site, ACAGAGA loop and AGC triad affect the efficiency of the reaction[39]. This confirms the intrinsic splicing catalytic ability of the U2/U6 snRNAs. But the reactions are very slow and inefficient, indicating that spliceosomal proteins may also play a pivotal role in maintaining the correct structure, modulating the dynamics of the active site and even may be involved directly in the catalysis of the spliceosome[65]. This is consistent with the fact that *in vivo* the spliceosome requires essential protein factors for the function such as Prp8. It is possible that during the evolution, some functions of the RNA has been transferred into the spliceosomal proteins and the spliceosome may actually be a RNP enzyme[5,49].

1.6 Roles of the spliceosomal snRNAs U2/U6 in splicing

U6 is highly conserved from yeast to human. There are three invariant regions: the U6 ISL bulge, the ACAGAGA loop and the AGC triad[3]. (Figure 1.5B) U2/U6 complex form extensive base-pairing network and is the prime candidate for the catalytic core of the spliceosome.

The spliceosomal snRNAs U2/U6 have pivotal structural roles. The spliceosome utilizes snRNAs to recognize the substrates, pre-mRNAs. In the activated spliceosome, U6 replaces U1 and positions the 5' splice site via the conserved ACAGAGA loop[66,67]. The conserved loop in U5 binds and coordinates two exons through non-canonical base pairs[68]. It has been proposed that the U6 ISL and the conserved ACAGAGA loop must be

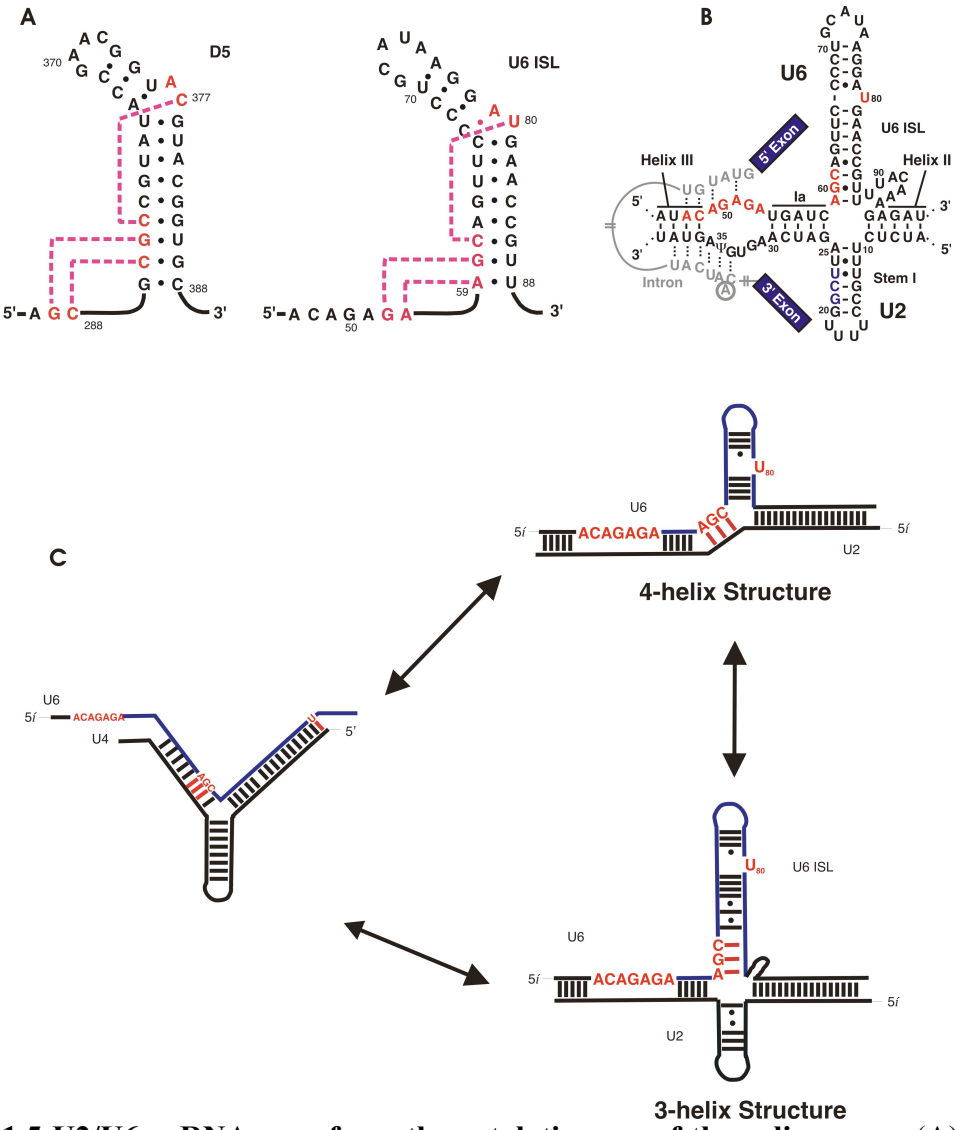


Figure 1.5 U2/U6 snRNA may form the catalytic core of the spliceosome (**A**) DV of the group II intron is very similar to U6 ISL of the spliceosome[15]. Highly conserved residues are highlighted in red. Red dash lines indicate base triple interactions. (**B**) Secondary structure model of the spliceosomal snRNAs U2–U6 from *Saccharomyces cerevisiae* with an intron bound.[69] Highly conserved residues in U6 (ACAGAGA loop, AGC triad and U80) are highlighted in red. (**C**) U6 Structural rearrangements. U6 first form the U4/U6 complex. After spliceosome activation, U4 RNA is unwound from U6 RNA. U6 then base pairs with U2 RNA to form the catalytically active form of the spliceosome. *The Figure **B** is adopted from paper: Guo, Z.; Karunatilaka, K. S.; Rueda, D., Single-molecule analysis of protein-free U2-U6 snRNAs. Nat Struct Mol Biol 2009, 16 (11), 1154-9.*

in close proximity for the first step of splicing to occur. This conformation brings the 5' splice site bound with ACAGAGA loop and the catalytically essential Mg^{2+} ion associated with U80 close to each other to facilitate the reaction[69,70]. Data from hydroxyl radical footprinting experiments support this hypothesis. After the first step of spicing, the catalytic core of the spliceosome may undergo rearrangements to position the 5' exon and the 3' splice site in the right orientation to enable the second step of splicing[71] (Figure 1.5B).

Furthermore, it has been proposed that the spliceosomal snRNAs U2/U6 also have pivotal catalytic roles. Mutational studies revealed that certain regions of the U6 snRNA, the ACAGAGA loop and the AGC triad, are highly conserved and any mutation or modification in these domains will lead to a defect in the first or second step of splicing[72,66]. This suggests that these sequences may be involved directly in the catalysis. These regions were shown to be able to form conserved secondary structure through base pairing[73] (i.e. Helix IB and U6 ISL), interact with specific protein factors[46] (i.e. Prp8 and Prp24), maintain three dimensional structure via short or long range tertiary interactions[15] (i.e. base triples), and bind and position catalytic important metal ions[74,75] (i.e. Mg^{2+}).

Like most of other ribozymes, catalysis of the spliceosome requires metal ions as cofactors[76,77]. It is highly possible that the spliceosome utilizes the two-metal ion mechanism in the catalysis[78]. Metal ions can facilitate the reaction by stabilizing the leaving group[79]. They may also play a significant structural role in stabilizing the tertiary interactions and co-helix stacking, neutralizing the negative charge of the RNA backbone to facilitate the folding of the U2/U6 complex. The metal ion bound in snRNAs and pre-

mRNA is required for the reaction[77]. Biochemical and structural studies of the group II intron reveal that the DV bulge and the catalytic triad bind catalytically important divalent ions[25,28]. Analogous regions in U6 were also implicated to bind Mg^{2+} ions involved in structure and catalysis[75,80]. By using phosphorothioate backbone substitutions that inhibit binding of Mg^{2+} but permit binding of Ca^{2+}, several phosphate groups in ACAGAGA loop, AGC triad and U80 bulge were shown to bind catalytically or structural important divalent ions[75,81,74]. The Mg^{2+} ion bound with U80 is required for the first step of splicing[75]. NMR studies also reveal the binding of Mg^{2+} ions in the U6 intramolecular stem loop (U6 ISL)[82].

1.7 Structural arrangement of the U2/U6 snRNAs

Because the spliceosome is a highly dynamic RNP machine, there are numerous RNA:RNA and RNA:protein rearrangements during the assembly and activation of the spliceosome[68] (Figure 1.3 and 1.4C). U6 is introduced to the spliceosome base paired with the U4 snRNA. Despite of this very stable interaction between U4 and U6 snRNAs in yeast, the base pairs are disrupted during the remodeling and activation of the spliceosome as U4 dissociates from the spliceosome before the first step of the splicing reaction[83]. In the activated spliceosome, U6 is base paired with U2 to form the active core[84]. During the two-step splicing reaction, the U2/U6 complex undergoes remodeling to enable different chemical reactions and the release of the products[85].

The active structure of U2/U6 has been a matter of debate over recent years. Genetics study in yeast revealed that the AGC triad forms base pairs with the GCU in U2 yielding the catalytically important helix IB and the U2/U6 complex adopts a three-way

junction structure[84]. Helix IB was shown to be very important for the second step of splicing, so this three-way junction structure may represent the active site of the second step of splicing[84,86]. However, according to a more recent NMR study, the U2/U6 adopts a four-way junction structure with an extended U6 ISL[69] (Figure 1.5B). In this structure, the ACG triad forms base pairs with CUU in U6, forming three additional base pairs to extend the U6 ISL. This structure is very similar to the human U2/U6 complex and DV in the group II intron[69,87,26]. The four-way junction structure provide a structure explanation for data from hydroxyl-radical cleavage showing that the ACAGAGA loop and the U6 ISL are close to each other during the first step of splicing in the activated spliceosome[70,73]. This structure may be able to bring the Mg^{2+} binding U80 and the 5' splice site close to each other, and it has been proposed that this structure is required for the first step of splicing[73,88,70]. The structure of the extended U6 ISL is mutually exclusive with the formation of helix IB, leading to the interesting hypothesis that each structure represents the conformation of one activation state of the spliceosome and intensive rearrangements occur between two steps of splicing[70,69]. The spliceosome may be in dynamic between different active structures and these dynamics may be very important for the stepwise assembly and catalysis of the spliceosome, as well as the regulation of the splicing[73,65].

Structural rearrangements of U6 require disruption of base pairs and secondary structures. The spliceosomal proteins, including RNA chaperones and DEAD-box proteins, may help to unwind and anneal spliceosomal snRNAs and stabilize active structures[49,4].

1.8 Base triples may exist in the U2/U6 complex

Compared to the well-characterized base pairs, there is relatively less understanding about how the tertiary interactions, such as base triples, maintain the native three-dimensional structure. A base triple forms when a single stranded nucleotide interacts with another nucleotide already involved in a base pair by hydrogen bonding[89]. Base triple interactions are common in RNA structures. For example, base triples were discovered by crystallography in yeast $tRNA^{Phe}$ and the group I intron by NMR[90,89]. In addition, base triple interactions in the telomerase RNA play a very important role in catalysis[91]. Recent crystal structure of the group II intron revealed that crucial base triples exist in the active core. It is highly possible that the triple helix locates in the active core of the spliceosome and is required for the reaction.

The similarities between the self-splicing group II intron and the spliceosome allow the recent atomic-resolution crystal structure of the group II intron to shed light on elucidating the structure of the spliceosomal snRNA U2/U6 complex[15]. The structure and behavior of DV is very similar to the U6. U80 is considered the analog of the DV bulge, the AGC catalytic triad is conserved from the group II intron to the spliceosome and the ACAGAGA loop in U6 is equivalent to J2/3 in the group II intron[7]. The most surprising feature of DV found in this crystal structure is the formation of three base triples involving DV[28,29,92]. According to the crystal structure of the group II intron, in DV, C360 and G383 form a Watson-Crick base pair and the sugar edge of C377 from DV bulge interacts with Hoogsteen edge of C360 yielding the G383-C360•C377 base triple. G359 and U384 form a Wobble base pair and the Hoogsteen edge of G288 from J2/3

pairs with G359 Watson edge to form the U384•G359•G288 base triple. Using the same set of interactions, C289 interacts with Watson-Crick base pair C358-G385, resulting in the C358-G385•C289 base triple (Figure1.4A). In the group II intron, the components participating in the base triples are far away from each other in primary sequence, but they are drawn together to form the active site (figure 1.2C). Based on the similarities, it has been hypothesized that similar base triples exist in the U6 snRNA[30,15] (Figure1.4A). In this model, U80 may interact with the C61-G86 base pair yielding a base triple, the G60•U87 and G52 in the ACAGAGA region form a triple in the middle of the triple helix and another one forms between the A59-U88 and an A53 within the ACAGAGA loop[15,93]. This model is very appealing because the catalytically crucial Mg^{2+} ion bound to U80 is brought close to the 5' splice site bound with ACAGAGA loop, branch site and the invariant AGC triad[73,70]. In this model, base triples bring the components necessary for the first step of splicing close together and should be pivotal for the function of spliceosome.

1.9 The single molecule FRET method and application in the studying of the spliceosome

1.9.1 Principles of FRET

FRET (fluorescence resonance energy transfer) is a very powerful tool to measure distance in the range of 10-100Å[94]. The energy of an exited donor is transferred to a nearby acceptor through dipole-dipole interactions (Figure 1.6A), causing decrease of donor emission and increase of acceptor emission[94]. The distance change of two dyes

(FRET donor and FRET acceptor) induces the FRET efficiency change[95]. By monitoring efficiency of FRET, the structural dynamics of biological molecules can be detected in real time[96,97]. The FRET efficiency is determined by

$$E = \frac{R_0^6}{R_0^6 + R^6} \qquad (1)$$

R is the distance between the two dyes and R_0 is the Forster distance. FRET is very sensitive to distance changes close to the R_0[94] (Figure 1.6B). R_0 is determined by the following equation

$$R_0 = 8.785 * 10^{-5} \frac{\kappa^2 \phi_D J}{n^4} \qquad (2)$$

In this equation, κ^2 is the orientation factor and, for free rotating dyes, this value is 2/3. ϕ_D is the quantum yield of the donor fluorescence. J is the overlap integral

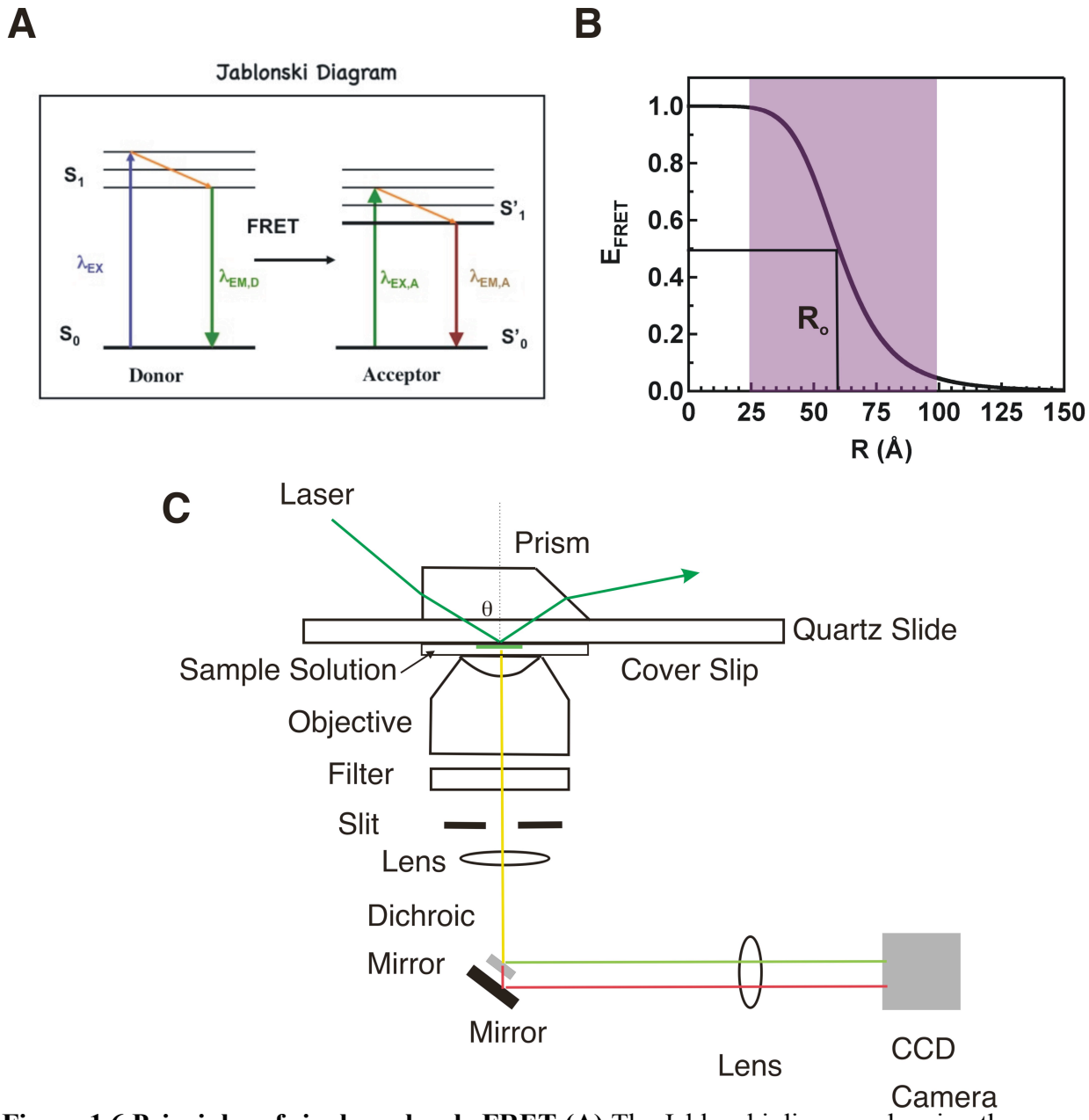


Figure 1.6 Principles of single molecule FRET (A) The Jablonski diagram showing the principles of the FRET experiments. The energy of an exited donor is transferred to a nearby acceptor through dipole-dipole interactions, decreasing donor emission and increasing acceptor emission. (**B**) The distance between two dyes (FRET donor and FRET acceptor) influences the FRET efficiency. R_0 is the forster distance that equal to the distance between the two dyes when FRET id 0.5. (**C**) Schematic diagram of TIRF single molecule[98,96]. Through a prism, the excitation laser hits the sample solution between the quartz slide and the cover slip at an angle (θ) slightly larger than the critical TIR angle and is totally reflected.

between the donor emission and the acceptor absorbance. The refractive index of medium n is 1.4 for aqueous solutions. According to equations 1 and 2, FRET efficiency depends on the distance between the FRET donor and acceptor, the relative orientation of the two dyes and the properties of the donor and acceptor[94]. R_0 of the FRET pair Cy3 (donor) and Cy5 (acceptor) is 55nM.

Experimentally, the FRET efficiency is determined with the following equation:

$$E = \frac{I_A}{I_A + I_D} \qquad (3)$$

I_A is the emission intensity from FRET acceptor and the I_D is the emission intensity from FRET donor. FRET has been widely used in studying the thermodynamics and kinetics of biological molecule conformational changes [94,99,100].

1.9.2 Principles of TIRF Single molecule techniques

The development of the single molecule techniques allowed the detection of FRET in the single molecule level[101]. To decrease the background and increase the signal/noise ratio, we use prism-based total internal reflection fluorescence microscopy (TIRFM) (Figure 1.6 C) Through a prism, the excitation laser hits the sample solution between the quartz slide and the cover slip at an angle slightly larger than the critical TIR angle (θ_c) and the laser is totally reflected. θ_c is determined by

$$\theta_c = \sin^{-1}\left(\frac{n_{sol}}{n_q}\right) \qquad (4)$$

In this equation, n_{sol} is the index of refraction of the sample solution and n_q is the index of refraction of the quartz slide. The totally reflected laser beam creates an evanescent wave at the interface between the slide and solution that penetrates about 100 nm into the sample solution, so only the molecules immobilized on the surface of quartz slide will be excited and the ones in the background solution will not. As a result, the excitation laser will not go deep into the solution and only the emission fluorescence from immobilized molecules will be detected (Figure 1.6 C). Therefore TIRFM is able to decrease background florescence dramatically. The florescence signals from samples are monitored and recorded by the objective and CCD camera[98].

TIRF single molecule FRET (smFRET) spectroscopy (Figure 1.6C) is a very powerful tool that enables us to uncover key structural and dynamic information otherwise hidden in bulk studies[97,101]. With the smFRET setup in our lab we can detect the structural dynamics of single molecules in real time (Figure 1.7C and 1.8B)[102,103]. This can provide detailed dynamic information about the conformational changes of biological molecules, capture transient intermediates which are difficult to detect in bulk solution experiments, and record conformational fluctuations at a time scale that is relevant for biological activity[104,96]. This technique also allows us to record movies of molecular motion and chemical activity of individual complexes[98]. Single molecule FRET has been widely used in studies on RNA and other biological molecules, such as folding of the group II intron, catalysis of the hairpin ribozyme and the movement of tRNAs in the ribosome [97,101,102,96, 105,106,107,108].

1.9.3 Single Molecule Experiments in studying spliceosomal snRNA U2/U6[109]

The U2 and U6 snRNAs play a very important role in the catalysis of spliceosome. However, the large size and dynamic nature of the spliceosome makes it very challenging to study by crystallography or NMR[5]. Some low-resolution structures were solved by electron microscopy (EM)[110], however, EM can only provide information about the overall structure and location of certain components of spliceosome[101,111]. So it is still impossible to access detailed structural information of the spliceosome. Previous genetic and NMR studies revealed possible structures of U2/U6, but the key structural dynamic information is still missing[84,69,87]. I used single molecule FRET (smFRET) to study the structural dynamics of U2/U6.

I have developed a smFRET assay studying the structural dynamics of U2/U6 with total internal reflection (TIR) Fluorescence spectroscopy[109] (Figure 1.6C and 1.7B). To investigate the dynamics of U2-U6, I have incorporated Cy5 (a FRET acceptor) at the U6 5′ end (nucleotides 45–70) and a biotin at the U6 3′ end (nucleotides 76–100) for surface mobilization (Figure 1.7A). To assess the dynamics between the ISL and the ACAGAGA loop, we have deleted the U6 pentaloop and labeled U70 with Cy3 (FRET donor) (Figure 1.7A). There are 3 strands of RNA forming the U2-U6 complex in our construct (Figure 1.7A): U6-1, U6-2 and U2. In my single molecule experiments, a sample containing 250 nM U6-1, 500nM U6-2 and 1 μM U2 was annealed in standard buffer (50 mM Tris-HCl, pH 7.5, 100 mM NaCl). The solution was heated to 90°C for 45 sec before cooling at room temperature over 20 min. The annealed biotinylated and fluorophore-labeled complex was then diluted to 25 - 50 pM and bound to a streptavidin-

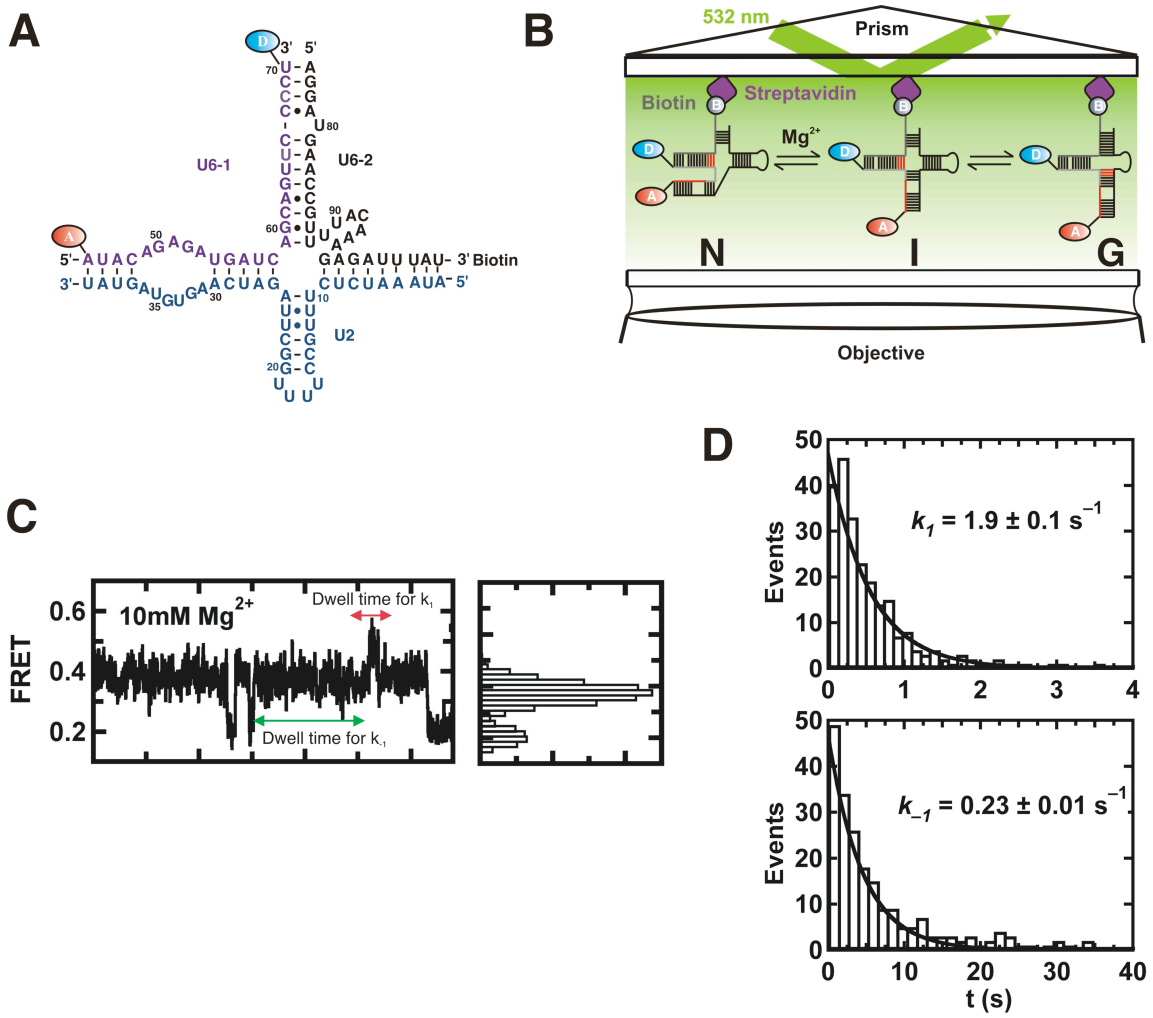


Figure 1.7 Single molecule FRET study of U2/U6 (**A**) Labeled U2/U6 construct. This complex consists of 3 strands, U6-1 (purple), U6-2 (black) and U2 (blue). The 5' of the U6-1 is labeled by a FRET acceptor Cy5 and the 3' is labeled by FRET donor Cy3. The 3' of U6-2 is labeled by biotin for surface immobilization. (**B**) Principle of single molecule spectroscopy with TIR excitation. Single U2-U6 molecules are immobilized via a Biotin-Streptavidin bridge. The donor fluorophore (D) is excited at 532 nm, and the donor and acceptor (A) fluorescence is collected through the objective. (**C**) Dwell time analysis of single molecule data. Dwell times of a single molecule staying in each state are calculated according to the single molecule trajectories. (**D**) Dwell time distributions in the N (top) and I (bottom) conformations in 40 mM Mg^{2+}. The distributions are fit to single exponential decays to yield the pseudo-first-order rates k_1 and k_{-1}, respectively. *The Figure is adopted from paper: Guo, Z.; Karunatilaka, K. S.; Rueda, D., Single-molecule analysis of protein-free U2-U6 snRNAs. Nat Struct Mol Biol 2009, 16 (11), 1154-9.*

coated quartz slide surface via biotin-streptavidin interaction to generate a surface density of ~0.1 molecules per μm^2. The donor (I_D) and acceptor (I_A) fluorescence signals of optically resolved single molecules (characterized by single-step photobleaching) were detected with a total internal reflection fluorescence microscope (Home-built) as described[98,106]. The donor and acceptor emission were separated using appropriate dichroic mirrors (610DCXR, Chroma, Rockingham, VT), and detected as two side-by-side images on a back-illuminated electron-multiplied CCD camera (Andor I-Xon$^+$, South Windsor, CT). The individual donor (I_D) and acceptor (I_A) intensities were measured by integration of their relative spot intensities. The donor (I_D) and acceptor (I_A) fluorescence signals of optically resolved single molecules were used to calculate the FRET ratio as $I_A/(I_A + I_D)$, and followed in real time for each molecule. The resulting FRET time trajectories are essentially recorded movies of the molecular motion of the U2/U6 complex, which enable us to study its conformational dynamics. Measurements were performed under variable magnesium concentrations (0 - 100 mM $MgCl_2$) at 25°C, with an oxygen scavenging system consisting of 10% (wt /vol) glucose, 1% (vol / vol) 2-mercaptoethanol, 750 µg/ml glucose oxidase, and 90 µg/ml catalase to reduce photobleaching. Dwell times of a single molecule staying in each state were calculated in Matlab with single molecule tine trajectories. Dwell time distributions are plotted and fit to single exponential decays to yield the pseudo-first-order rates constants (k_1, k_{-1}, k_2, k_{-2}). The smFRET histograms constructed with more than 100 molecules, (Figure 1.7C,D).

1.10 References

1. Berget, S. M.; Moore, C.; Sharp, P. A., Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proc Natl Acad Sci U S A* **1977,** *74* (8), 3171-5.

2. Chow, L. T.; Gelinas, R. E.; Broker, T. R.; Roberts, R. J., An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell* **1977,** *12* (1), 1-8.

3. Burge, C. B.; Tuschl, T. A.; Sharp, P. A., Splicing of precursors to mRNAs by the spliceosomes. *in The RNA World 2nd edition, R. F. Gesteland, T. R. Cech, J. F. Atkins (eds.), Cold Spring Harbor Laborotory Press, Cold Spring Harbor, NY, 1999, 525-560.* **1999**.

4. Staley, J. P.; Guthrie, C., Mechanical devices of the spliceosome: motors, clocks, springs, and things. *Cell* **1998,** *92* (3), 315-26.

5. Wahl, M. C.; Will, C. L.; Luhrmann, R., The spliceosome: design principles of a dynamic RNP machine. *Cell* **2009,** *136* (4), 701-18.

6. Cech, T. R., The generality of self-splicing RNA: relationship to nuclear mRNA splicing. *Cell* **1986,** *44* (2), 207-10.

7. Lambowitz, A. M.; Zimmerly, S., Mobile group II introns. *Annu Rev Genet* **2004,** *38*, 1-35.

8. Black, D. L., Mechanisms of alternative pre-messenger RNA splicing. *Annu Rev Biochem* **2003,** *72*, 291-336.

9. Wang, G. S.; Cooper, T. A., Splicing in disease: disruption of the splicing code and the decoding machinery. *Nat Rev Genet* **2007,** *8* (10), 749-61.

10. Kalnina, Z.; Zayakin, P.; Silina, K.; Line, A., Alterations of pre-mRNA splicing in cancer. *Genes Chromosomes Cancer* **2005,** *42* (4), 342-57.

11. Licatalosi, D. D.; Darnell, R. B., Splicing regulation in neurologic disease. *Neuron* **2006,** *52* (1), 93-101.

12. Cooper, T. A.; Wan, L.; Dreyfuss, G., RNA and disease. *Cell* **2009,** *136* (4), 777-93.

13. Lehmann, K.; Schmidt, U., Group II introns: structure and catalytic versatility of large natural ribozymes. *Crit Rev Biochem Mol Biol* **2003,** *38* (3), 249-303.

14. Pyle, A. M.; Lambowitz, A. M., Group II Introns: Ribozymes That Splice RNA and Invade DNA. *in The RNA World, 3rd edition (ed. R.F. Gesteland et al.), pp. 469-505. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.* **2006**.

15. Keating, K. S.; Toor, N.; Perlman, P. S.; Pyle, A. M., A structural analysis of the group II intron active site and implications for the spliceosome. *RNA* **2010,** *16* (1), 1-9.

16. Jacquier, A.; Michel, F., Multiple exon-binding sites in class II self-splicing introns. *Cell* **1987,** *50* (1), 17-29.

17. Jacquier, A.; Jacquesson-Breuleux, N., Splice site selection and role of the lariat in a group II intron. *J Mol Biol* **1991,** *219* (3), 415-28.

18. (a) Fedorova, O.; Mitros, T.; Pyle, A. M., Domains 2 and 3 interact to form critical elements of the group II intron active site. *J Mol Biol* **2003,** *330* (2), 197-209; (b) Jacquier, A.; Michel, F., Base-pairing interactions involving the 5' and 3'-terminal nucleotides of group II self-splicing introns. *J Mol Biol* **1990,** *213* (3), 437-47.

19. Jacquier, A., Self-splicing group II and nuclear pre-mRNA introns: how similar are they? *Trends Biochem Sci* **1990,** *15* (9), 351-4.

20. Michel, F.; Umesono, K.; Ozeki, H., Comparative and functional anatomy of group II catalytic introns--a review. *Gene* **1989,** *82* (1), 5-30.

21. Jarrell, K. A.; Dietrich, R. C.; Perlman, P. S., Group II intron domain 5 facilitates a trans-splicing reaction. *Mol Cell Biol* **1988,** *8* (6), 2361-6.

22. Koch, J. L.; Boulanger, S. C.; Dib-Hajj, S. D.; Hebbar, S. K.; Perlman, P. S., Group II introns deleted for multiple substructures retain self-splicing activity. *Mol Cell Biol* **1992,** *12* (5), 1950-8.

23. Franzen, J. S.; Zhang, M.; Peebles, C. L., Kinetic analysis of the 5' splice junction hydrolysis of a group II intron promoted by domain 5. *Nucleic Acids Res* **1993,** *21* (3), 627-34.

24. Peebles, C. L.; Zhang, M.; Perlman, P. S.; Franzen, J. S., Catalytically critical nucleotide in domain 5 of a group II intron. *Proc Natl Acad Sci U S A* **1995,** *92* (10), 4422-6.

25. Gordon, P. M.; Piccirilli, J. A., Metal ion coordination by the AGC triad in domain 5 contributes to group II intron catalysis. *Nat Struct Biol* **2001,** *8* (10), 893-8.

26. Sigel, R. K.; Sashital, D. G.; Abramovitz, D. L.; Palmer, A. G.; Butcher, S. E.; Pyle, A. M., Solution structure of domain 5 of a group II intron ribozyme reveals a new RNA motif. *Nat Struct Mol Biol* **2004,** *11* (2), 187-92.

27. Zhang, L.; Doudna, J. A., Structural insights into group II intron catalysis and branch-site selection. *Science* **2002,** *295* (5562), 2084-8.

28. Toor, N.; Keating, K. S.; Taylor, S. D.; Pyle, A. M., Crystal structure of a self-spliced group II intron. *Science* **2008,** *320* (5872), 77-82.

29. Toor, N.; Rajashankar, K.; Keating, K. S.; Pyle, A. M., Structural basis for exon recognition by a group II intron. *Nat Struct Mol Biol* **2008,** *15* (11), 1221-2.

30. Toor, N.; Keating, K. S.; Pyle, A. M., Structural insights into RNA splicing. *Curr Opin Struct Biol* **2009,** *19* (3), 260-6.

31. Pyle, A. M.; Fedorova, O.; Waldsich, C., Folding of group II introns: a model system for large, multidomain RNAs? *Trends Biochem Sci* **2007,** *32* (3), 138-45.

32. Dib-Hajj, S. D.; Boulanger, S. C.; Hebbar, S. K.; Peebles, C. L.; Franzen, J. S.; Perlman, P. S., Domain 5 interacts with domain 6 and influences the second transesterification reaction of group II intron self-splicing. *Nucleic Acids Res* **1993,** *21* (8), 1797-804.

33. Jacquier, A.; Rosbash, M., Efficient trans-splicing of a yeast mitochondrial RNA group II intron implicates a strong 5' exon-intron interaction. *Science* **1986,** *234* (4780), 1099-104.

34. Michels, W. J., Jr.; Pyle, A. M., Conversion of a group II intron into a new multiple-turnover ribozyme that selectively cleaves oligonucleotides: elucidation of reaction mechanism and structure/function relationships. *Biochemistry* **1995,** *34* (9), 2965-77.

35. Podar, M.; Dib-Hajj, S.; Perlman, P. S., A UV-induced, Mg(2+)-dependent crosslink traps an active form of domain 3 of a self-splicing group II intron. *RNA* **1995,** *1* (8), 828-40.

36. Moore, M. J., Query, C.C., and Sharp, P.A., Splicing of precursors to mRNA by the spliceosome. *In TheRNA World, R. Gesteland and J. Atkins, eds. (New York: Cold Spring Harbor Laboratory Press)* **1993**, pp. 303–357.

37. Brow, D. A., Allosteric cascade of spliceosome activation. *Annu Rev Genet* **2002,** *36*, 333-60.

38. Segault, V.; Will, C. L.; Polycarpou-Schwarz, M.; Mattaj, I. W.; Branlant, C.; Luhrmann, R., Conserved loop I of U5 small nuclear RNA is dispensable for both catalytic steps of pre-mRNA splicing in HeLa nuclear extracts. *Mol Cell Biol* **1999,** *19* (4), 2782-90.

39. Valadkhan, S.; Mohammadi, A.; Jaladat, Y.; Geisler, S., Protein-free small nuclear RNAs catalyze a two-step splicing reaction. *Proc Natl Acad Sci U S A* **2009,** *106* (29), 11901-6.

40. Valadkhan, S.; Manley, J. L., Splicing-related catalysis by protein-free snRNAs. *Nature* **2001,** *413* (6857), 701-7.

41. Hoogstraten, C. G.; Sumita, M., Structure-function relationships in RNA and RNP enzymes: recent advances. *Biopolymers* **2007,** *87* (5-6), 317-28.

42. Will, C. L.; Schneider, C.; MacMillan, A. M.; Katopodis, N. F.; Neubauer, G.; Wilm, M.; Luhrmann, R.; Query, C. C., A novel U2 and U11/U12 snRNP protein that associates with the pre-mRNA branch site. *EMBO J* **2001,** *20* (16), 4536-46.

43. Yang, K.; Zhang, L.; Xu, T.; Heroux, A.; Zhao, R., Crystal structure of the beta-finger domain of Prp8 reveals analogy to ribosomal proteins. *Proc Natl Acad Sci U S A* **2008,** *105* (37), 13817-22.

44. Ritchie, D. B.; Schellenberg, M. J.; Gesner, E. M.; Raithatha, S. A.; Stuart, D. T.; Macmillan, A. M., Structural elucidation of a PRP8 core domain from the heart of the spliceosome. *Nat Struct Mol Biol* **2008,** *15* (11), 1199-205.

45. Pena, V.; Rozov, A.; Fabrizio, P.; Luhrmann, R.; Wahl, M. C., Structure and function of an RNase H domain at the heart of the spliceosome. *EMBO J* **2008,** *27* (21), 2929-40.

46. Abelson, J., Is the spliceosome a ribonucleoprotein enzyme? *Nat Struct Mol Biol* **2008,** *15* (12), 1235-7.

47. Cartegni, L.; Chew, S. L.; Krainer, A. R., Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat Rev Genet* **2002,** *3* (4), 285-98.

48. Singh, R.; Valcarcel, J., Building specificity with nonspecific RNA-binding proteins. *Nat Struct Mol Biol* **2005,** *12* (8), 645-53.

49. Collins, C. A.; Guthrie, C., The question remains: is the spliceosome a ribozyme? *Nat Struct Biol* **2000,** *7* (10), 850-4.

50. Kwan, S. S.; Brow, D. A., The N- and C-terminal RNA recognition motifs of splicing factor Prp24 have distinct functions in U6 RNA binding. *RNA* **2005,** *11* (5), 808-20.

51. Bae, E.; Reiter, N. J.; Bingman, C. A.; Kwan, S. S.; Lee, D.; Phillips, G. N., Jr.; Butcher, S. E.; Brow, D. A., Structure and interactions of the first three RNA recognition motifs of splicing factor prp24. *J Mol Biol* **2007,** *367* (5), 1447-58.

52. Martin-Tumasz, S.; Reiter, N. J.; Brow, D. A.; Butcher, S. E., Structure and functional implications of a complex containing a segment of U6 RNA bound by a domain of Prp24. *RNA* **2010,** *16* (4), 792-804.

53. Raghunathan, P. L.; Guthrie, C., A spliceosomal recycling factor that reanneals U4 and U6 small nuclear ribonucleoprotein particles. *Science* **1998,** *279* (5352), 857-60.

54. Ghetti, A.; Company, M.; Abelson, J., Specificity of Prp24 binding to RNA: a role for Prp24 in the dynamic interaction of U4 and U6 snRNAs. *RNA* **1995,** *1* (2), 132-45.

55. Achsel, T.; Brahms, H.; Kastner, B.; Bachi, A.; Wilm, M.; Luhrmann, R., A doughnut-shaped heteromer of human Sm-like proteins binds to the 3'-end of U6 snRNA, thereby facilitating U4/U6 duplex formation in vitro. *EMBO J* **1999,** *18* (20), 5789-802.

56. Verdone, L.; Galardi, S.; Page, D.; Beggs, J. D., Lsm proteins promote regeneration of pre-mRNA splicing activity. *Curr Biol* **2004,** *14* (16), 1487-91.

57. Ryan, D. E.; Stevens, S. W.; Abelson, J., The 5' and 3' domains of yeast U6 snRNA: Lsm proteins facilitate binding of Prp24 protein to the U6 telestem region. *RNA* **2002,** *8* (8), 1011-33.

58. Jandrositz, A.; Guthrie, C., Evidence for a Prp24 binding site in U6 snRNA and in a putative intermediate in the annealing of U6 and U4 snRNAs. *EMBO J* **1995,** *14* (4), 820-32.

59. Stevens, S. W.; Ryan, D. E.; Ge, H. Y.; Moore, R. E.; Young, M. K.; Lee, T. D.; Abelson, J., Composition and functional characterization of the yeast spliceosomal penta-snRNP. *Mol Cell* **2002,** *9* (1), 31-44.

60. Vidaver, R. M.; Fortner, D. M.; Loos-Austin, L. S.; Brow, D. A., Multiple functions of Saccharomyces cerevisiae splicing protein Prp24 in U6 RNA structural rearrangements. *Genetics* **1999,** *153* (3), 1205-18.

61. Karaduman, R.; Fabrizio, P.; Hartmuth, K.; Urlaub, H.; Luhrmann, R., RNA structure and RNA-protein interactions in purified yeast U6 snRNPs. *J Mol Biol* **2006,** *356* (5), 1248-62.

62. Konforti, B. B.; Abramovitz, D. L.; Duarte, C. M.; Karpeisky, A.; Beigelman, L.; Pyle, A. M., Ribozyme catalysis from the major groove of group II intron domain 5. *Mol Cell* **1998,** *1* (3), 433-41.

63. Sharp, P. A., On the origin of RNA splicing and introns. *Cell* **1985,** *42* (2), 397-400.

64. Cech, T. R.; Bass, B. L., Biological catalysis by RNA. *Annu Rev Biochem* **1986,** *55*, 599-629.

65. Valadkhan, S., snRNAs as the catalysts of pre-mRNA splicing. *Curr Opin Chem Biol* **2005,** *9* (6), 603-8.

66. Lesser, C. F.; Guthrie, C., Mutations in U6 snRNA that alter splice site specificity: implications for the active site. *Science* **1993,** *262* (5142), 1982-8.

67. Parker, R.; Siliciano, P. G.; Guthrie, C., Recognition of the TACTAAC box during mRNA splicing in yeast involves base pairing to the U2-like snRNA. *Cell* **1987,** *49* (2), 229-39.

68. Nilsen, T. W., RNA-RNA Interactions in Nuclear Pre-mRNA Splicing. *In RNA Structure and Function, R. Simons and M. Grunberg-Manago, eds. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press),* **1998**, pp. 279–307.

69. Sashital, D. G.; Cornilescu, G.; McManus, C. J.; Brow, D. A.; Butcher, S. E., U2-U6 RNA folding reveals a group II intron-like domain and a four-helix junction. *Nat Struct Mol Biol* **2004,** *11* (12), 1237-42.

70. Rhode, B. M.; Hartmuth, K.; Westhof, E.; Luhrmann, R., Proximity of conserved U6 and U2 snRNA elements to the 5' splice site region in activated spliceosomes. *EMBO J* **2006,** *25* (11), 2475-86.

71. Konarska, M. M.; Vilardell, J.; Query, C. C., Repositioning of the reaction intermediate within the catalytic center of the spliceosome. *Mol Cell* **2006,** *21* (4), 543-53.

72. Fabrizio, P.; Abelson, J., Two domains of yeast U6 small nuclear RNA required for both steps of nuclear precursor messenger RNA splicing. *Science* **1990,** *250* (4979), 404-9.

73. Butcher, S. E.; Brow, D. A., Towards understanding the catalytic core structure of the spliceosome. *Biochem Soc Trans* **2005,** *33* (Pt 3), 447-9.

74. Valadkhan, S.; Manley, J. L., Intrinsic metal binding by a spliceosomal RNA. *Nat Struct Biol* **2002,** *9* (7), 498-9.

75. Yean, S. L.; Wuenschell, G.; Termini, J.; Lin, R. J., Metal-ion coordination by U6 small nuclear RNA contributes to catalysis in the spliceosome. *Nature* **2000,** *408* (6814), 881-4.

76. Narlikar, G. J.; Herschlag, D., Mechanistic aspects of enzymatic catalysis: lessons from comparison of RNA and protein enzymes. *Annu Rev Biochem* **1997,** *66*, 19-59.

77. Sontheimer, E. J.; Sun, S.; Piccirilli, J. A., Metal ion catalysis during splicing of premessenger RNA. *Nature* **1997,** *388* (6644), 801-5.

78. Gordon, P. M.; Sontheimer, E. J.; Piccirilli, J. A., Metal ion catalysis during the exon-ligation step of nuclear pre-mRNA splicing: extending the parallels between the spliceosome and group II introns. *RNA* **2000,** *6* (2), 199-205.

79. Steitz, T. A.; Steitz, J. A., A general two-metal-ion mechanism for catalytic RNA. *Proc Natl Acad Sci U S A* **1993,** *90* (14), 6498-502.

80. Sontheimer, E. J.; Gordon, P. M.; Piccirilli, J. A., Metal ion catalysis during group II intron self-splicing: parallels with the spliceosome. *Genes Dev* **1999,** *13* (13), 1729-41.

81. Yu, Y. T.; Maroney, P. A.; Darzynkiwicz, E.; Nilsen, T. W., U6 snRNA function in nuclear pre-mRNA splicing: a phosphorothioate interference analysis of the U6 phosphate backbone. *RNA* **1995,** *1* (1), 46-54.

82. Huppler, A.; Nikstad, L. J.; Allmann, A. M.; Brow, D. A.; Butcher, S. E., Metal binding and base ionization in the U6 RNA intramolecular stem-loop structure. *Nat Struct Biol* **2002,** *9* (6), 431-5.

83. Yean, S. L.; Lin, R. J., U4 small nuclear RNA dissociates from a yeast spliceosome and does not participate in the subsequent splicing reaction. *Mol Cell Biol* **1991,** *11* (11), 5571-7.

84. Madhani, H. D.; Guthrie, C., A novel base-pairing interaction between U2 and U6 snRNAs suggests a mechanism for the catalytic activation of the spliceosome. *Cell* **1992,** *71* (5), 803-17.

85. Madhani, H. D.; Guthrie, C., Dynamic RNA-RNA interactions in the spliceosome. *Annu Rev Genet* **1994,** *28*, 1-26.

86. Hilliker, A. K.; Staley, J. P., Multiple functions for the invariant AGC triad of U6 snRNA. *RNA* **2004,** *10* (6), 921-8.

87. Sun, J. S.; Manley, J. L., A novel U2-U6 snRNA structure is necessary for mammalian mRNA splicing. *Genes Dev* **1995,** *9* (7), 843-54.

88. Ferre-D'amare A, R.; Rupert, P. B., The hairpin ribozyme: from crystal structure to function. *Biochem Soc Trans* **2002,** *30* (Pt 6), 1105-9.

89. Chastain, M.; Tinoco, I., Jr., A base-triple structural domain in RNA. *Biochemistry* **1992,** *31* (51), 12733-41.

90. Holbrook, S. R.; Sussman, J. L.; Warrant, R. W.; Kim, S. H., Crystal structure of yeast phenylalanine transfer RNA. II. Structural features and functional implications. *J Mol Biol* **1978,** *123* (4), 631-60.

91. Qiao, F.; Cech, T. R., Triple-helix structure in telomerase RNA contributes to catalysis. *Nat Struct Mol Biol* **2008,** *15* (6), 634-40.

92. Toor, N.; Keating, K. S.; Fedorova, O.; Rajashankar, K.; Wang, J.; Pyle, A. M., Tertiary architecture of the Oceanobacillus iheyensis group II intron. *RNA* **2010,** *16* (1), 57-69.

93. Michel, F.; Costa, M.; Westhof, E., The ribozyme core of group II introns: a structure in want of partners. *Trends Biochem Sci* **2009,** *34* (4), 189-99.

94. Klostermeier, D.; Millar, D. P., RNA conformation and folding studied with fluorescence resonance energy transfer. *Methods* **2001,** *23* (3), 240-54.

95. Selvin, P. R., Fluorescence resonance energy transfer. *Methods Enzymol* **1995,** *246*, 300-34.

96. Cornish, P. V.; Ha, T., A survey of single-molecule techniques in chemical biology. *ACS Chem Biol* **2007,** *2* (1), 53-61.

97. Bokinsky, G.; Zhuang, X., Single-molecule RNA folding. *Acc Chem Res* **2005,** *38* (7), 566-73.

98. Zhao, R.; Rueda, D., RNA folding dynamics by single-molecule fluorescence resonance energy transfer. *Methods* **2009,** *49* (2), 112-7.

99. Walter, N. G., Structural dynamics of catalytic RNA highlighted by fluorescence resonance energy transfer. *Methods* **2001,** *25* (1), 19-30.

100. Aleman, E. A.; Lamichhane, R.; Rueda, D., Exploring RNA folding one molecule at a time. *Curr Opin Chem Biol* **2008,** *12* (6), 647-54.

101. Zhuang, X., Single-molecule RNA science. *Annu Rev Biophys Biomol Struct* **2005,** *34*, 399-414.

102. Vale, R. D., Microscopes for fluorimeters: the era of single molecule measurements. *Cell* **2008,** *135* (5), 779-85.

103. Walter, N. G.; Huang, C. Y.; Manzo, A. J.; Sobhy, M. A., Do-it-yourself guide: how to use the modern single-molecule toolkit. *Nat Methods* **2008,** *5* (6), 475-89.

104. Roy, R.; Hohng, S.; Ha, T., A practical guide to single-molecule FRET. *Nat Methods* **2008,** *5* (6), 507-16.

105. Weiss, S., Measuring conformational dynamics of biomolecules by single molecule fluorescence spectroscopy. *Nat Struct Biol* **2000,** *7* (9), 724-9.

106. Steiner, M.; Karunatilaka, K. S.; Sigel, R. K.; Rueda, D., Single-molecule studies of group II intron ribozymes. *Proc Natl Acad Sci U S A* **2008,** *105* (37), 13853-8.

107. Rueda, D.; Bokinsky, G.; Rhodes, M. M.; Rust, M. J.; Zhuang, X.; Walter, N. G., Single-molecule enzymology of RNA: essential functional groups impact catalysis from a distance. *Proc Natl Acad Sci U S A* **2004,** *101* (27), 10066-71.

108. Blanchard, S. C.; Gonzalez, R. L.; Kim, H. D.; Chu, S.; Puglisi, J. D., tRNA selection and kinetic proofreading in translation. *Nat Struct Mol Biol* **2004,** *11* (10), 1008-14.

109. Guo, Z.; Karunatilaka, K. S.; Rueda, D., Single-molecule analysis of protein-free U2-U6 snRNAs. *Nat Struct Mol Biol* **2009,** *16* (11), 1154-9.

110. Stark, H.; Luhrmann, R., Cryo-electron microscopy of spliceosomal components. *Annu Rev Biophys Biomol Struct* **2006,** *35*, 435-57.

111. Luhrmann, R.; Stark, H., Structural mapping of spliceosomes by electron microscopy. *Curr Opin Struct Biol* **2009,** *19* (1), 96-102.

CHAPTER 2

Single-molecule analysis of protein-free U2–U6 snRNAs reveals key Mg^{2+} dependent folding dynamics

2.1 Introduction

Splicing is an essential step in eukaryotic precursor messenger RNA (pre-mRNA) maturation[1]. Single nucleotide errors can be lethal to the cell, and anomalous pre-mRNA splicing has been linked to many cancers and neurodegenerative disorders[2,3,4]. The spliceosome is a multi-megadalton protein–RNA complex consisting of five small nuclear RNAs (snRNAs U1, U2, U4, U5 and U6) and more than 150 proteins[5,6]. It catalyzes splicing and is a critical center for the regulation of alternative splicing[7]. During spliceosomal assembly and catalysis, snRNAs associate with proteins to form ribonucleoprotein complexes (snRNPs), which bind to and dissociate from the pre-mRNA substrate in a highly orchestrated process yielding catalytically competent spliceosomes[8]. Splicing consists of two sequential transesterification reactions that produce a free lariat intron and ligate two exons into a mature mRNA.

Only U2, U5 and U6 are present in active spliceosomes, but evidence suggests that U2 and U6 are directly involved in the first catalytic step. U2 and U6 form an extensive base pair network and directly bind the 5′ splice site and the branch point, positioning them for the first reaction (Fig. 2.1A)[9,10]. A highly conserved element of U5 is

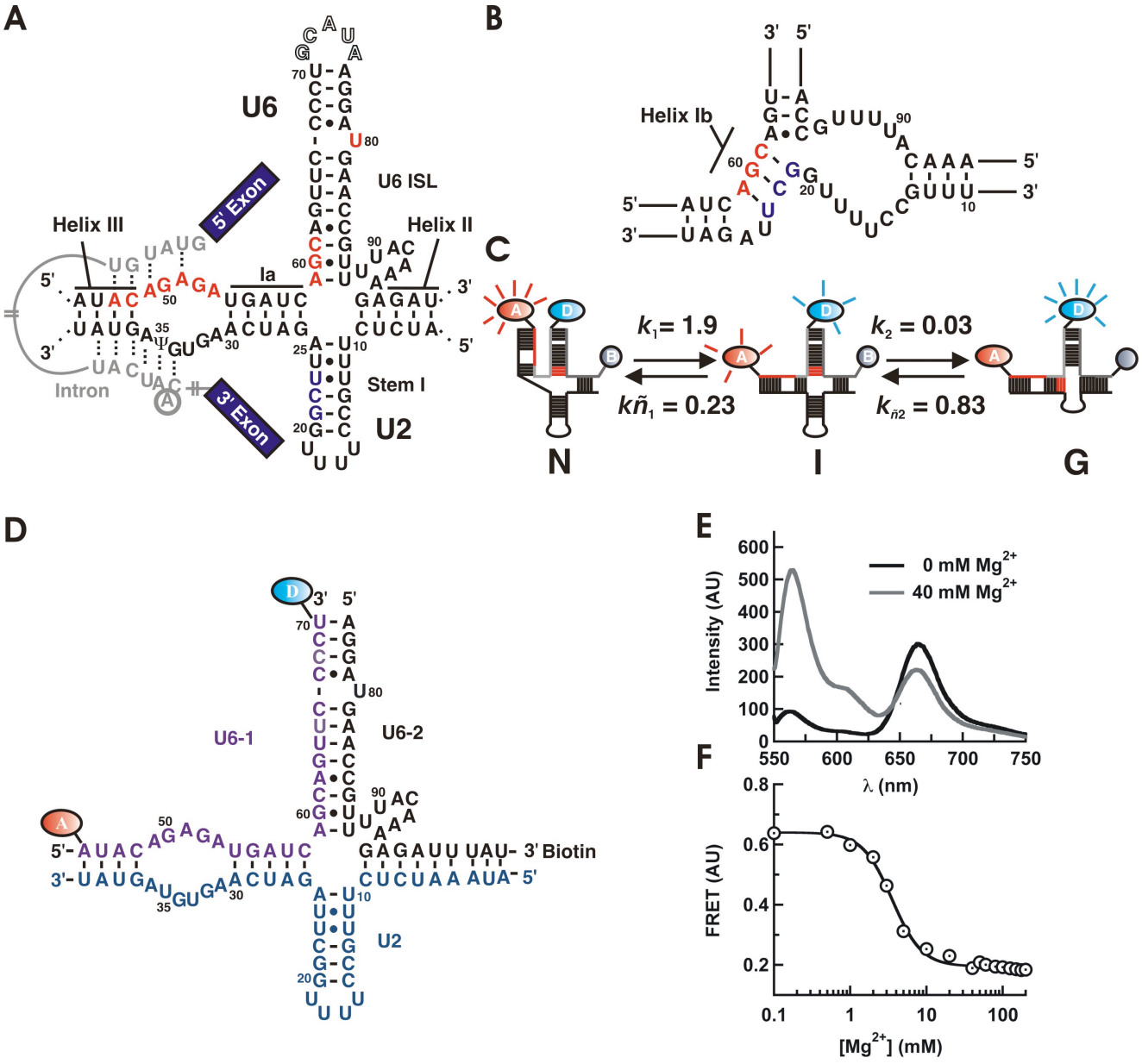


Figure 2.1 Secondary structure model of the spliceosomal snRNAs U2–U6 (**A**) Secondary structure model of the spliceosomal snRNAs U2–U6 from *Saccharomyces cerevisiae* with an intron bound[11]. Highly conserved residues in U6 (ACAGAGA loop, AGC triad and U80) are highlighted in red. U2 residues involved in helix IB formation are highlighted in blue. The U6 ISL pentaloop (outlined residues G71–A75) was deleted to facilitate construct labeling. (**B**) Proposed structure of helix IB. (**C**) Proposed folding reaction pathway for the U2–U6 complex (rates measured at 40 mM Mg^{2+}; units are s^{-1}). At least three distinct conformations are observed for the U2–U6 snRNA complex: a high FRET conformation that resembles the four-helix structure (N), a low FRET conformation that resembles the three-helix structure with helix IB (G), and a previously unobserved folding intermediate (I). (**D**) Labeled U2/U6 construct. This complex consists of 3 strands, U6-1 (purple), U6-2 (black) and U2 (blue). The 5' of the U6-1 is labeled by a FRET acceptor Cy5 and The 3' is labeled by FRET donor Cy3. The 3' of U6-2 is labeled by biotin for surface immobilization. (**E**) Mg^{2+}-induced conformational change in the U2–U6 spliceosomal complex. Fluorescence emission spectra of the fluorophore-labeled U2–U6 complex, in the absence (black) and presence (gray) of Mg^{2+} ions. The donor fluorophore (Cy3) emits at 565 nm and the acceptor (Cy5) at 665 nm. Mg^{2+} ions decrease the acceptor emission and increase the donor emission, revealing the presence of a conformational change. (**F**) Calculated FRET ratio as a function of [Mg^{2+}]. The line is a fit to a modified Hill equation (K_{Mg} = 3.0 ± 0.2 mM and n = 2.2 ± 0.2) *The Figure is adopted from paper: Guo, Z.; Karunatilaka, K. S.; Rueda, D., Single-molecule analysis of protein-free U2-U6 snRNAs. Nat Struct Mol Biol 2009, 16 (11), 1154-9.*

dispensable for the first splicing step *in vitro* and in mammalian cell extracts[12,13]. Splicing-related catalysis has been reported *in vitro* using a protein-free human U2–U6 complex[14,15]. Parallels between the structure and catalytic mechanism of the spliceosome and self-splicing group II introns support the idea of a common molecular ancestor and suggest that the spliceosome is a ribozyme[16,17,18]. However, recent crystal structures of a Prp8 fragment (a component of U5) suggest that both protein and RNA are involved in catalysis[19].

The U2–U6 structure has been a matter of debate in recent years. Early *in vivo* genetic studies support a three-helix structure in which the highly conserved AGC triad in U6 forms three base pairs with U2 (Fig. 2.1b)[20]. Believed to form during catalytic activation, this structure has been proposed to be a requirement for both splicing steps[21,22].

More recently, NMR studies have shown that in the absence of Mg^{2+} ions and proteins, the AGC triad forms intramolecular base pairs that extend the U6 internal stem loop (ISL, Fig. 2.1a), creating a four-helix structure[11,23]. To reconcile these two structures, it has been postulated that each corresponds to a different spliceosomal activation state[24], yet any conformational change has remained undefined. To address this issue, we have used fluorescence resonance energy transfer (FRET) and single-molecule fluorescence to characterize the structure and dynamics of a protein-free U2–U6 complex from yeast. Our results show that Mg^{2+} ions trigger a large amplitude conformational change that separates the ISL and the ACAGAGA loop (Fig. 2.1). This two-step conformational change contains a previously unobserved obligatory intermediate, where

only the first step is strongly Mg^{2+}-dependent. Mutations in the highly conserved AGC triad show that helix IB forms only in the lowest FRET conformation, and correlate the observed structural dynamics *in vitro* with previously published mutations linked to activation of the second step *in vivo*.

2.2 Materials and methods

2.2.1 RNA Preparations and Purification.

There are 3 strands of RNA forming the U2-U6 complex in our construct (Figure 2.1D): U6-1, U6-2 and U2. All the RNA samples were purchased from Keck Foundation Biotechnology Resource Laboratory at Yale University School of Medicine (New Haven, CT). These RNA samples have 2'OH protection groups that are removed by a 2'OH deprotection reaction according to the manufacture's protocol. The deprotected RNA was purified by denaturing gel electrophoresis (20% polyacrylamide and 8 M urea) and diffusion elution with elution buffer (0.5 M NH4OAc and 0.1 mM EDTA) at 4˚C overnight. This was followed by a chloroform extraction, ethanol precipitation, and C8 reverse-phase HPLC purification. The 5' end of U6-1 is labeled with Cy5 during the RNA synthesis and a C7 amino linker is attached to 3' end, enabling the labeling of a Cy3 to 3' end for our FRET studies. During the 3'-labeling reaction, the Cy5-labeled RNA was mixed with Cy3 dye (GE Healthcare) in labeling buffer (100 mM Sodium tetraborate, pH 8.5). This reaction mixture was kept overnight at room temperature. The doubly-labeled samples were purified by ethanol precipitation and reverse phase HPLC. RNA concentrations were measured by UV-Vis absorbance at 260 nm.

2.2.2 Steady-State FRET.

Steady-state FRET measurements of the double-labeled U2-U6 complex were carried out in a spectrofluorometer (Cary Eclipse, Varian Inc., Palo Alto, CA). Cy3 was excited at 550nm (10 nm bandwidth) and Cy3 and Cy5 emission intensities (I_{Cy3} and I_{Cy5}) were measured at 565 nm and 665 nm (5 nm bandwidth), respectively. Relative FRET efficiencies were calculated as FRET = $I_{Cy5}/(I_{Cy3} + I_{Cy5})$. A 130 µL solution with 250 nM U6-1, 500 nM U6-2 and 1 µM U2 in standard buffer (50 mM Tris-HCl, pH 7.5, 100mM NaCl and 25mM DTT) was first denatured by heating at 90 ˚C for 2 min and then slowly annealed at room temperature for 20 minutes. After this RNA-only solution was scanned, 10-µL aliquots of varying concentrations of $MgCl_2$ were manually added to the RNA. After equilibration, the final FRET value was measured. The Mg^{2+} dissociation constant (K_{Mg}) and cooperativity coefficients (n) were obtained by plotting FRET as a function of Mg^{2+} concentration and fitting to the modified Hill equation:

$$FRET = FRET_0 - \left(FRET_0 - FRET_\infty\right)\left([Mg^{2+}]^n \Big/ \left(K_M^n + [Mg^{2+}]^n\right)\right)$$

2.2.3 Single-molecule FRET.

The three RNA strands (Fig. 2.1D) were annealed at 1 µM concentration in standard buffer (50 mM Tris-HCl, pH 7.5, 100 mM NaCl) and variable [Mg^{2+}]. A 10 µL solution was heated to 90°C for 45 sec before cooling at room temperature over 20 min. The annealed biotinylated and fluorophore-labeled complex was diluted to 25 - 50 pM concentration and bound to a streptavidin-coated quartz slide surface via the biotin-streptavidin bridge to generate a surface density of ~0.1 molecules/µm^2. The donor

fluorophore was excited in a home-built total internal reflection microscopy by a laser (532 nm, 3 mW, Spectra-Physics Excelsior, Mountain View, CA). The donor and acceptor emission were separated using appropriate dichroic mirrors (610DCXR, Chroma, Rockingham, VT), and detected as two side-by-side images on a back-illuminated electron-multiplied CCD camera (Andor I-Xon, South Windsor, CT). The individual donor (I_D) and acceptor (I_A) intensities were measured by integration of their relative spot intensities. The donor (I_D) and acceptor (I_A) fluorescence signals of optically resolved single molecules (characterized by single-step photobleaching) were detected used to calculate the FRET ratio as $I_A/(I_A + I_D)$, and followed in real time for each molecule. The resulting FRET time trajectories are essentially recorded movies of the molecular motion of the U2/U6 complex, which enable us to study its conformational dynamics. Measurements were performed under variable [Mg^{2+}] (0 - 100 mM) at 22°C, with an oxygen scavenging system consisting of 10% (wt/vol) glucose, 2% (vol/vol) 2-mercaptoethanol, 750 µg/ml glucose oxidase, and 90 µg/ml catalase to reduce photobleaching, with trolox (Sigma). The dwell times of each folding event were calculated, histograms constructed, and the folding rate constants determined.

2.2.4 Fluoresce Based Gel Mobility Assay

15% non-denaturing polyacrylamide (29:1acrylamide: bisacrylamide ratio) gel electrophoresis was performed in 20 mM NaOAc, 20 mM Tris-acetic acid (pH6.5) and 40 mM $Mg(OAc)_2$ using low-fluorescence glass plates. 10 pmol doubly labeled RNA samples were heated at 90 °C for 45 seconds and cooled down to room temperature in standard buffer. Acrylamide gel was equilibrated for 15 minutes at 4 °C before loading.

Samples were loaded on the gel, and a constant electric field (6 V/cm) was immediately applied. After electrophoresis for 8 hrs at 4 °C, the gel was scanned in a Typhoon 9210 Variable Model Imager (GE Healthcare), and analyzed with ImageQuant software (Amershan Bioscience). For color calibration purposes, RNA labeled with only Cy3 and Cy5 was used. The Fluorsep software (Amershan Bioscience) was used to overlay the Cy3 and Cy5 gel images. Cy3 and Cy5 emissions appear as green and red color, respectively. In this gel, **Lane 1**: 10 nmol of fluorophore labeled U6-1 only. The red (high FRET) RNA migrates as a single band indicating its purity. **Lane 2**: 10 nmol of fluorophore labeled U6-1 and 20 nmol of biotinylated U6-2. The red (high FRET) RNA migrates as a single band indicating that these two RNAs alone do not from a stable U6-1/U6-2 complex. **Lane 3**: 10 nmol of fluorophore labeled U6-1 and 40 nmol of U2. The band migrates as in lane 1 indicating that these two RNAs alone do not from a stable U6-1/U2 complex. **Lane 4**: 10 nmol of fluorophore labeled U6-1, 20 nmol of biotinylated U6-2 and 40 nmol of U2. A yellow (low-mid FRET) slow migrating band clearly shows the formation of the ternary U6-1/U6-2/U2 complex. **Lane 5**: 10 nmol of fluorophore labeled triad mutant U6-1 and 20 nmol of biotinylated triad mutant U6-2. The red (high FRET) RNA migrates as a single band indicating that these two RNAs alone do not from a stable U6-1/U6-2 complex. **Lane 6**: 10 nmol of fluorophore labeled triad U6-1 and 40 nmol of U2. The band migrates as in lane 1 indicating that these two RNAs alone do not from a stable U6-1/U2 complex. **Lane 7**: 10 nmol of fluorophore labeled triad mutant U6-1, 20 nmol of biotinylated triad mutant U6-2 and 40 nmol of U2. An orange (mid FRET) slow migrating band clearly shows the formation of the ternary triad mutant U6-

1/U6-2/U2 complex. **Lane 8**: 10 nmol of fluorophore labeled single mutant U6-1 and 20 nmol of biotinylated U6-2. The red (high FRET) RNA migrates as a single band indicating that these two RNAs alone do not from a stable U6-1/U6-2 complex. **Lane 9**: 10 nmol of fluorophore labeled single mutant U6-1 and 40 nmol of U2. The band migrates as in lane 1 indicating that these two RNAs alone do not from a stable U6-1/U2 complex. **Lane 10**: 10 nmol of fluorophore labeled single mutant U6-1, 20 nmol of biotinylated U6-2 and 40 nmol of U2. A yellow-orange (low-mid FRET) slow migrating band clearly shows the formation of the ternary single mutant U6-1/U6-2/U2 complex. **Lane 11**: 10 nmol of fluorophore labeled single mutant U6-1, 20 nmol of biotinylated U6-2 and 40 nmol of double mutant U2. A yellow-orange (low-mid FRET) slow migrating band clearly shows the formation of the ternary double mutant U6-1/U6-2/U2 complex.

Table 1 Sequences of RNA used in chapter 2

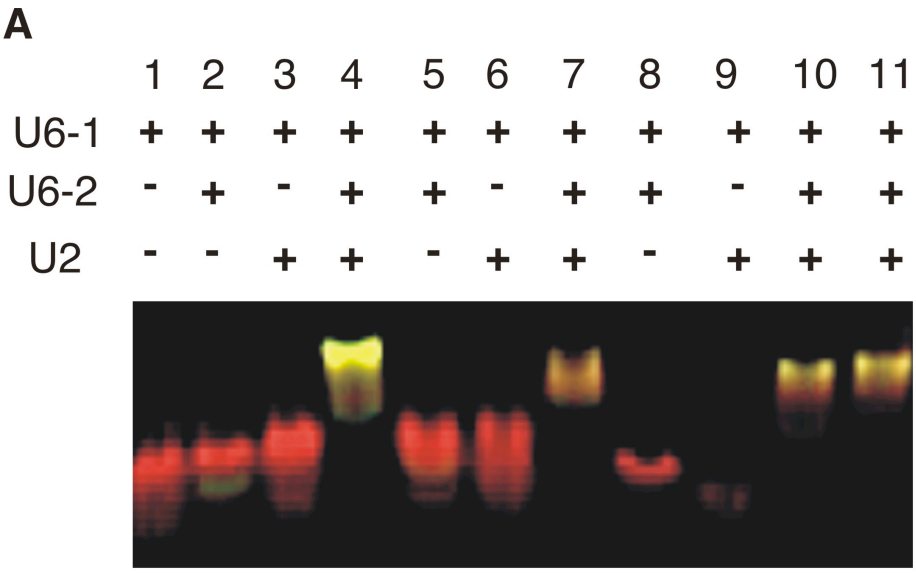
WT U6-1	5'-AUA CAG AGA UGA UCA GCA GUU CCC CU-3'
WT U6-2	5'-AGG AUG AAC CGU UUU ACA AAG AGA UUU AU-3'
WT U2	5'-AUC UCU UUG CCU UUU GGC UUA GAU CAA GUG UAG UAU-3'
Six-fold mutant U6-1	5'-AUA CAG AGA UGA UCU UGA GUU CCC CU-3'
Six-fold mutant U6-2	5'-AGG AUG AAC CCG AUU ACA AAG AGA UUU AU-3'
A59C U6-1	5'-AUA CAG AGA UGA UCC GCA GUU CCC CU-3'
U23G U2	5'-AUC UCU UUG CCU UUU GGC GUA GAU CAA GUG UAG UAU-3'
A91G U6-2	5'-AGG AUG AAC CGU UUG ACA AAG AGA UUU AU-3'
Full length control U6	5'-AUA CAG AGA UGA UCA GCA GUU CCC CdTG CAU AAG GAU GAA CCG UUU UAC AAA GAG AUU UAU-3'

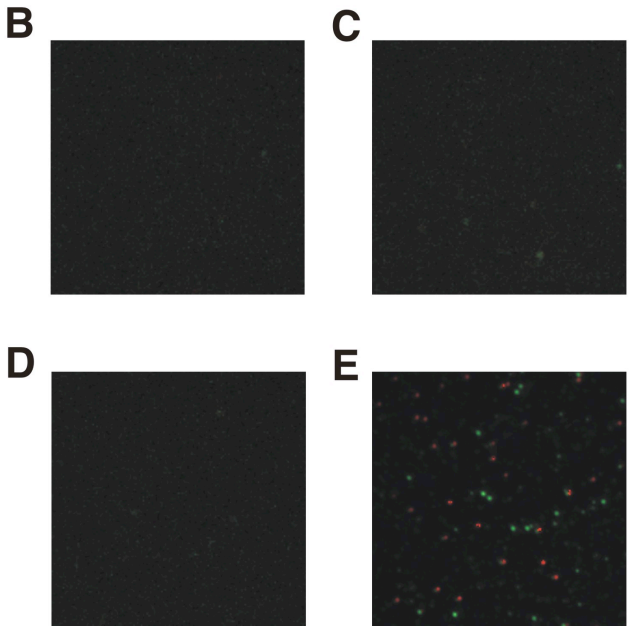
2.3 Results and discussion

2.3.1 Mg^{2+} induces a large amplitude conformational change

We have developed a labeling strategy that enables us to study the U2–U6 conformational dynamics by FRET and single-molecule fluorescence (Fig. 2.1C). We have incorporated Cy5 (a FRET acceptor) at the U6 5′ end (nucleotides 45–70) and a biotin at the U6 3′ end (nucleotides 76–100) for surface mobilization. To assess the dynamics between the ISL and the ACAGAGA loop, we have deleted the U6 pentaloop (Fig. 2.1A, outlined and 2.1D) and labeled U70 with Cy3 (FRET donor). To confirm three-strand complex formation, we used nondenaturing gel electrophoresis (Fig. 2.2A; single bands in lanes 4, 7, 10 and 11 are reliable indicators of homogeneity), fluorescence and single-molecule microscopy (Fig. 2.2B-E). The four- and three-helix structures are expected to yield high and low FRET ratios, respectively[11,20].

We first characterized the folding behavior of our fluorophore-labeled U2–U6 complex using bulk solution FRET (Fig. 2.1E,F). In the presence and absence of Mg^{2+} in standard buffer (50 mM Tris-HCl, pH 7.5, 100 mM NaCl). In the absence of Mg^{2+} (black line, Fig. 2.1E), the donor intensity is lower than the acceptor density (FRET = 0.64), indicating that U2–U6 adopts an initial conformation in which the ACAGAGA loop and the ISL are in close proximity, as suggested by the four-helix structure[11]. In 40 mM $MgCl_2$, the donor intensity increases and the acceptor intensity decreases (FRET = 0.19), indicating that Mg^{2+} ions induce a conformational change separating the ACAGAGA loop and the ISL, as expected for the three-helix struture[20]. Titrating the observed FRET





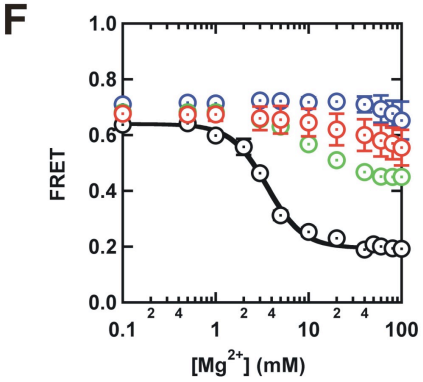


Figure 2.2. Formation of the fluorophore-labeled U2/U6 complex (**A**) monitored by non-denaturing gel electrophoresis with fluorescence detection. Each sample was prepared as in the FRET experiments. (**B, C, D, E**) smFRET control experiments testing how the heterogeneity affects the single molecule measurements. (**B**) Slide with only Cy3 and Cy5 labeled U6-1. Without biotin, U6-1 alone cannot be immobilized on the slide surface, and no molecules are observed. (**C**) Slide with U6-1 and U6-2. There are only very few molecules on the slide because the U6-1/U6-2 complex alone is not stable, consistent with the non-denaturing gel electrophoresis. The surface density of immobilized molecules is 20-fold lower than that of the 3-strand complex. (**D**) Slide with U6-1 and U2. No molecule is immobilized on the slide surface because there is no biotin present. (**E**) Slide with U6-1, U6-2 and U2. The 3-strand complex can form stably, and it can be immobilized on the slide surface efficiently. The buffer conditions for all these experiments are 50 mM Tris-HCl, pH 7.5, 10 mM MgCl2, 100 mM NaCl. (**F**) Ensemble-averaged magnesium titration of the fluorophore labeled three-strand complex (black), U6-1 alone (blue), U6-1/U2 (red), and U6-1/U6-2 (green). Only the three-strand complex exhibits a significant Mg^{2+} dependent conformational change. These data, together with B-E show that the observed conformational dynamics reflect only the behavior of the three-strand complex.
The Figure is adopted from paper: Guo, Z.; Karunatilaka, K. S.; Rueda, D., Single-molecule analysis of protein-free U2-U6 snRNAs. Nat Struct Mol Biol 2009, 16 (11), 1154-9.

efficiency as a function of [Mg^{2+}] (Fig. 2.1F) reveals a cooperative decrease in FRET efficiency with increasing [Mg^{2+}]. This conformational change occurs with a dissociation constant K_{Mg} = 3.3 ± 0.7 mM, near the physiological range (~1 mM), and a cooperativity coefficient n = 2.2 ± 0.2. These data indicate that U2–U6 acts as a Mg^{2+}-dependent conformational switch. The high cooperativity coefficient probably reflects the presence of multiple Mg^{2+} binding sites, in agreement with previous studies that have linked several coordination sites near the highly conserved base U80 in U6, the ACG triad and the ACAGAGA loop (Fig. 2.1A) with spliceosomal activation[25-26].

2.3.2 Single-molecule FRET reveals three dynamic conformations

Single-molecule FRET (smFRET) uncovers key structural and dynamic information otherwise hidden in ensemble-averaged experiments[27]. Here, we have used smFRET to elucidate the folding reaction of the U2–U6 complex (Fig. 2.3A). Characteristic FRET time trajectories of surface-immobilized U2–U6 in 10 and 100 mM Mg^{2+} are shown in Figure 2.3B,C, respectively. The observed FRET efficiency jumps randomly between three different values (~0.6, ~0.4 and ~0.2), revealing the presence of at least three distinct conformational states in dynamic equilibrium. On the basis of their FRET ratios, we initially assigned the high FRET state to the four-helix structure (N)[11], the low FRET state to the three-helix structure (G) and the mid-FRET state to a previously unobserved folding intermediate (I, Fig. 2.1C).

FRET histograms from the trajectories (Fig. 2.3B,C, right panels) reveal that at high [Mg^{2+}] (>10 mM), I and G are the predominant states, and N is only transiently

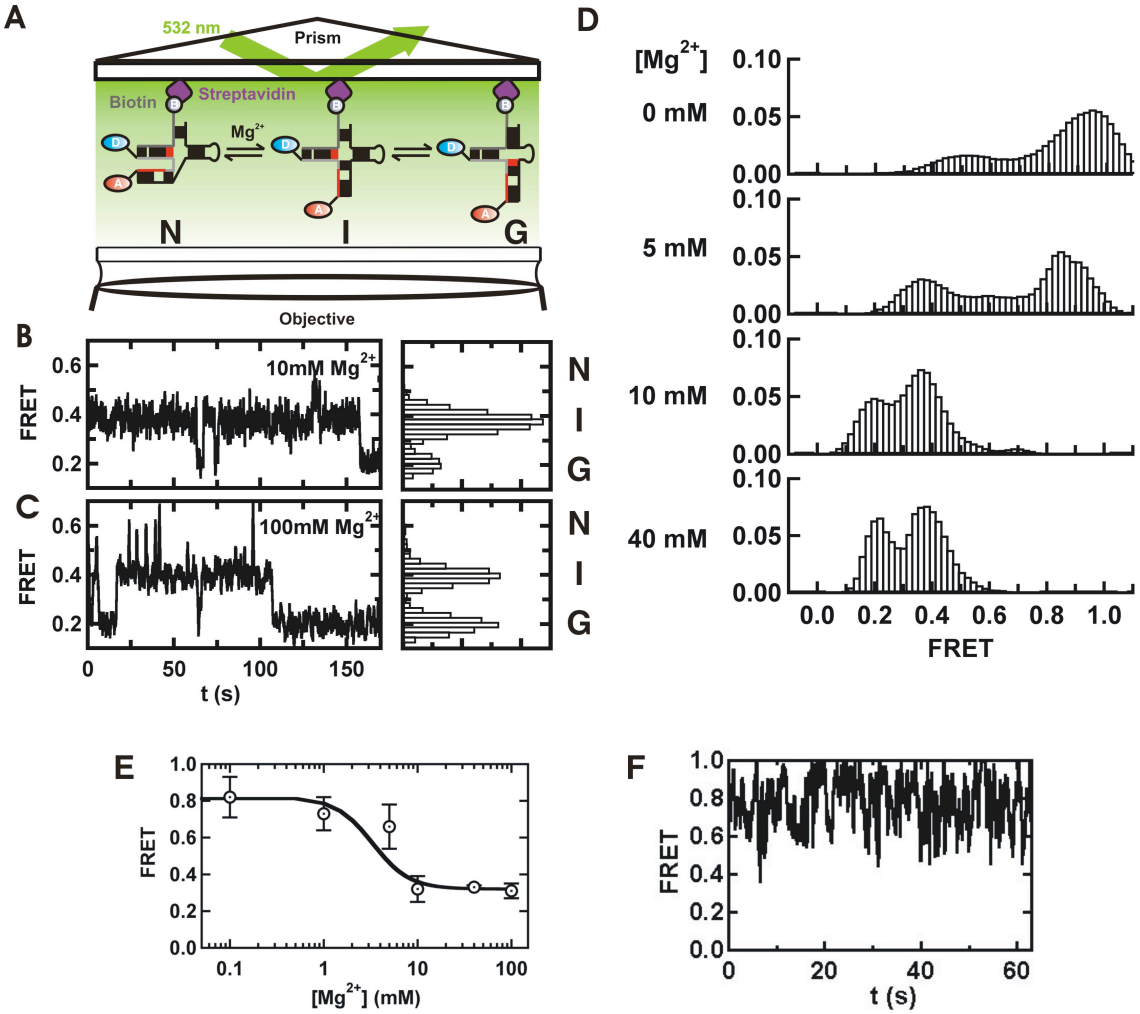


Figure 2.3 Single-molecule FRET reveals a three-state folding pathway (**A**) Schematic diagram of the single-molecule experiments. The RNA complex is surface immobilized via a biotin-streptavidin bridge. The fluorophores are excited in a prism-based total internal reflection microscope. Fluorescence is collected by the objective and monitored with a CCD camera. (**B** and **C**). Single-molecule FRET time trajectories in 10 and 100 mM Mg^{2+}, respectively. The different conformations (N, I and G) can be identified by their FRET ratios (~0.6, ~0.4 and ~0.2, respectively) in the corresponding FRET histograms. (**D**) Mg^{2+} ions modulate the stability of the U2–U6 conformations. Each panel corresponds to an average smFRET histogram from >100 single-molecule trajectories as a function of [Mg^{2+}], as indicated. In the absence of Mg^{2+}, a high FRET state predominates the histogram. As [Mg^{2+}] increases, the 0.4 and 0.2 FRET states become more populated. (**E**) Comparison of the Mg^{2+} titration in ensemble-averaged and singlemolecule experiments. The calculated average FRET values from the single-molecule histograms in Figure 2.3D (circles) overlay within error on the fit of the ensemble averaged measurements from Figure 2.1F (line). The error bars stem from the standard deviation of the single-molecule FRET distributions. The initial and final FRET values of the fit had to be adjusted because the filters used in the two experiments are different yielding somewhat different initial and final FRET values, but the K_{Mg} and the Hill coefficient n were held constant to their values in Figure 2.1F. This results shows that the surface immobilization scheme does not significantly affect the U2/U6 structural dynamics. The buffer conditions for these smFRET experiments are 50 mM Tris-HCl, pH 7.5, 100 mM NaCl, and for bulk titration 50 mM Tris-HCl, pH 7.5, 100mM NaCl and 25mM DTT. (**F**) Single-molecule time trajectory in the absence of Mg^{2+}. The trace is smoothed with a running 5-point average. The FRET value oscillates randomly between 1.0 and 0.5 without the presence of well-defined conformational states. The observed oscillations are much wider than our experimental noise indicating the presence of dynamics faster than our 33 ms time resolution. These dynamics result in the broad distributions in Figure 2.3D. A possible explanation is that the 5'end of U6 does not base pair stably with U2 in the absence of Mg^{2+}. The buffer condition for this experiment is 50 mM Tris-HCl, pH 7.5, 0 mM MgCl2, 100 mM NaCl. *The Figure is adopted from paper: Guo, Z.; Karunatilaka, K. S.; Rueda, D., Single-molecule analysis of protein-free U2-U6 snRNAs. Nat Struct Mol Biol 2009, 16 (11), 1154-9.*

populated. The high FRET state becomes more apparent at [Mg^{2+}] < 10 mM (Fig. 2.3D). Of the >3,200 transitions observed, only 40 (1.2%) go directly from N to G. In these instances, the dwell time in the I state may be shorter than our 33-ms time resolution, which is direct evidence that I is an obligatory intermediate. Based on our single-molecule measurements, we propose a two-step folding pathway for the U2–U6 complex (Fig. 2.1C).

2.3.3 Mg^{2+} ions influence the U2–U6 structural dynamics

To assess the effect of Mg^{2+} ions on the U2–U6 folding dynamics, we measured smFRET histograms between 0 and 40 mM Mg^{2+} (Fig. 2.3D). These histograms, each built from more than 100 trajectories, show how Mg^{2+} ions modulate the folding dynamics of the U2–U6 complex. In the absence of Mg^{2+} (top panel), most molecules exist in the high FRET state, and only a small fraction adopts the intermediate FRET conformation. Increasing the [Mg^{2+}] to 5 mM (second panel) shifts some of the high FRET population into the intermediate FRET population. In 10 mM Mg^{2+} or higher (Fig. 2.3D, bottom panels), the high FRET state becomes more transiently populated and almost disappears. Only brief excursions to the high FRET state are observed at these concentrations, and the U2–U6 molecules spend most of their time in the intermediate and low FRET states. The average FRET values of the single-molecule results overlay the ensemble-averaged results within the calculated s.d. (Fig. 2.3E), indicating that surface immobilization does not appreciably affect the U2–U6 folding dynamics.

Notably, the high FRET peak gradually shifts from ~0.95 at 0 mM Mg^{2+} to ~0.6 at 40 mM Mg^{2+}. Such smooth changes in FRET distributions are usually an indication

that multiple dynamic states are present in solution but their dynamics are faster than the available time resolution[28]. This idea is further corroborated by the broad distributions with flat shoulders at 0 and 5 mM Mg^{2+} (Fig. 2.3D). A similar behavior has been reported for the Diels-Alderase ribozyme[29]. It is possible that in low [Mg^{2+}], the AU-rich helix III is partially unwound (Fig. 2.1A and Fig. 2.3F), yielding the 0.95 FRET value. As [Mg^{2+}] increases, helix III forms and effectively distances the U6 5′ end from the ISL, yielding the 0.6 FRET value. The highly dynamic time trajectories observed below 5 mM Mg^{2+} (Fig. 2.3F), as well as previously reported Mg2+ binding sites in this region[25,26], support this hypothesis.

2.3.4 Dwell time analysis determines the rates of folding

We have determined the folding rate constants k_1, k_{-1}, k_2 and k_{-2} (Fig. 2.1C) for [Mg^{2+}] = 5–100 mM. Below this range, dwell times were unidentifiable owing to the highly dynamic time trajectories (see above). At 40 mM Mg^{2+}, the dwell time distributions can be readily fitted with single exponential decays to yield pseudo-first-order rate constants (Fig. 2.4A and Fig. 2.5). Transitions out of the I state were divided into two categories depending on their final state: those that returned to N were used to determine k_{-1}, whereas those that moved on to G were used to determine k_2[17]. The resulting rate constants and their Mg^{2+} dependence are shown in Figure 2.4B.

These data explain how Mg^{2+} ions affect the U2–U6 folding dynamics and how to interpret them in terms of a folding potential energy surface comprising three minima (N,

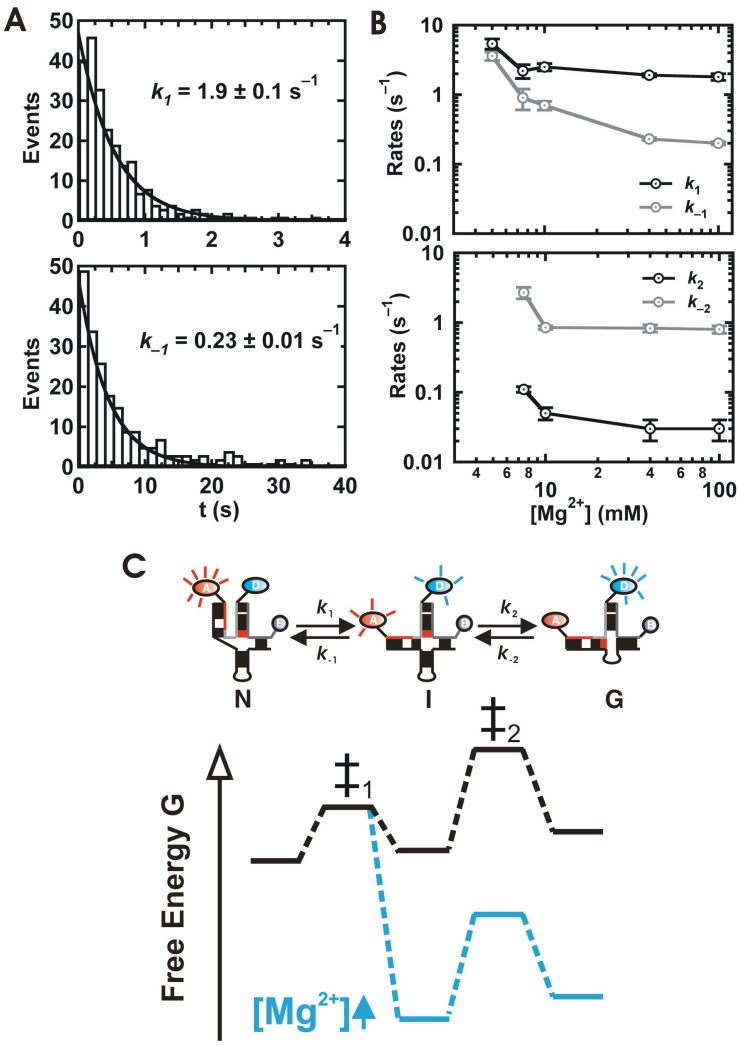


Figure 2.4 Mg2+ dependence of the folding rate constants for the U2–U6 complex (**A**) Dwell time distributions in the N (top) and I (bottom) conformations in 40 mM Mg^{2+}. The distributions are fit to single exponential decays to yield the pseudo-first-order rates k_1 and k_{-1}, respectively. (**B**) Mg^{2+} dependence of k_1 and k_{-1} (top) and k_2 and k_{-2} (bottom). (**C**) Schematic diagram of the folding potential energy surface for the protein-free U2–U6 spliceosomal complex from yeast. In the absence of Mg^{2+} (black), the high FRET state (N) is the most stable. On addition of Mg^{2+} ions (green), the previously unobserved mid-FRET conformation (I) becomes more stable, whereas N and the first transition state (‡1) remain approximately constant. This is shown by the large decrease in k_{-1} but not in k_1. Meanwhile, the low FRET state (G) and the second transition state (‡2) are also similarly stabilized, as shown by the mild Mg^{2+} dependence of k_2 and k_{-2}. *The Figure is adopted from paper: Guo, Z.; Karunatilaka, K. S.; Rueda, D., Single-molecule analysis of protein-free U2-U6 snRNAs. Nat Struct Mol Biol 2009, 16 (11), 1154-9.*

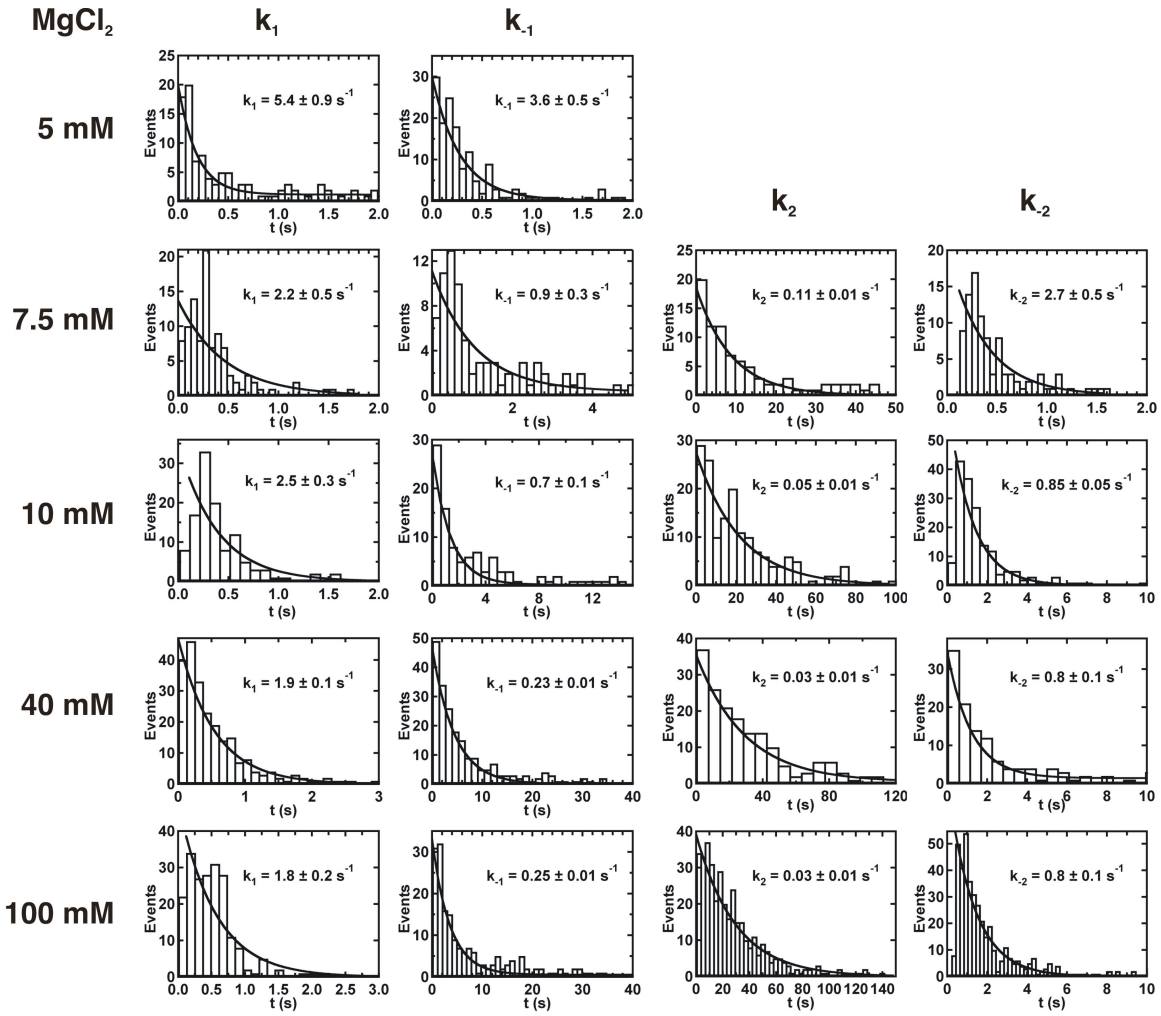


Figure 2.5 Dwell time analysis for the rate constants k_1 (column1), k_{-1} (column2), k_2 (column3) and k_{-2} (column4) as a function of [Mg^{2+}]. The pseudo-first order rates were determined by fitting the dwell time distributions to single exponential decays (black lines). The resulting values were used in Fig 2.4B. The buffer conditions for these experiments are 50 mM Tris-HCl, pH 7.5, 100 mM NaCl and variable [Mg^{2+}]. *The Figure is adopted from paper: Guo, Z.; Karunatilaka, K. S.; Rueda, D., Single-molecule analysis of protein-free U2-U6 snRNAs. Nat Struct Mol Biol 2009, 16 (11), 1154-9.*

I and G) separated by two transition states (Fig. 2.4C). Between 5 and 100 mM Mg^{2+}, k_1 decreases 3-fold whereas k_{-1} decreases 18-fold. The large decrease in k_{-1} indicates that Mg^{2+} ions preferentially stabilize I relative to the first transition state, leaving N approximately unchanged. Between 5 and 100 mM Mg^{2+}, k_2 and k_{-2} decrease only four- and three-fold, respectively, indicating that I, G and the second transition state are similarly stabilized by Mg^{2+} ions.

2.3.5 Helix IB forms in the low FRET state

To test for the presence of helix IB in the low FRET conformation (G), we designed a six-fold mutant that prevents helix IB formation while maintaining the stability of the extended ISL (Fig. 2.6A), because the identity of junction-closing base pairs can affect the junction stability and dynamics[30]. An ensemble-averaged Mg^{2+} titration shows that these mutations do not affect K_{Mg}, but the observed FRET ratio at high [Mg^{2+}] is higher than that of the wild type (Fig. 2.7A). The absence of the 0.2 FRET state in the single-molecule trajectories and corresponding FRET histogram (Fig. 2.6A) support our initial assignments. Dwell time analysis of this mutant shows that k_1 and k_{-1} are within two-fold of the wild-type values (Fig. 2.7B), showing that this mutation does not affect the first step of folding.

Helix IB formation can be favored by the U6 mutation A91G, which extends helix II by one base pair and stabilizes it relative to stem I in U2 (Figs. 2.1A and 2.6B and J. Halsig, D.G. Sashital and S.E. Butcher, University of Wisconsin, Madison, personal communication). The trajectories clearly show that excursions to the low FRET state last

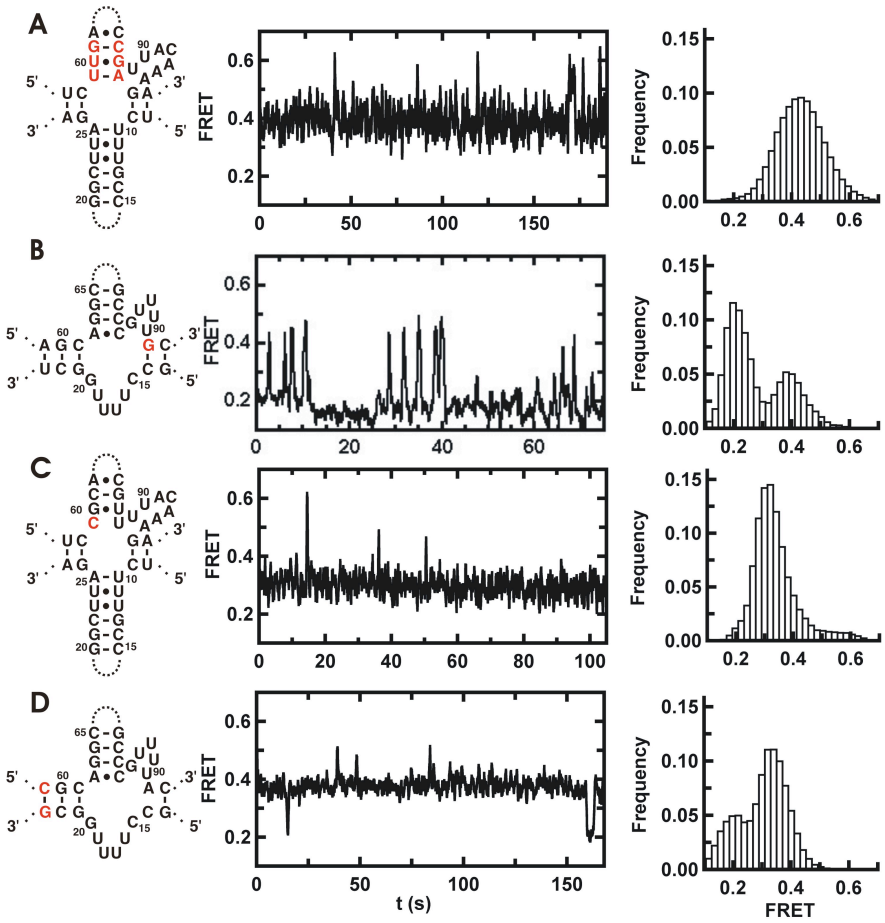


Figure 2.6 Mutations in the AGC triad (red) show that helix IB forms only in the low FRET state (**A**) Single-molecule trajectory of a six-fold mutant in the U6 snRNA with a base-flipped AGC triad that prevents formation of helix IB. The low FRET state is never observed in the trajectories or in the corresponding FRET histogram (right). (**B**) Single-molecule trajectory of an A91G mutant in the U6 snRNA previously shown to stabilize the formation of helix IB. Excursions to the low FRET state last longer, and the corresponding peak in the FRET histogram (right) is larger. (**C**) Single-molecule trajectory of an A59C mutant in the U6 snRNA that blocks the second step of splicing *in vivo*. Neither the 0.4 nor 0.2 FRET states are observed in the corresponding trajectories. The FRET histogram (right) reveals the presence of a new state with intermediate FRET (~0.3). (**D**) A U23G mutation in the U2 snRNA rescues formation of helix IB in the presence of the U6 A59C mutation. The single-molecule trajectories and corresponding histograms show the recovery of the 0.4–0.2 FRET dynamics *in vitro*. All measured in 50 mM Tris-HCl, pH 7.5, 100 mM NaCl and 40 mM $MgCl_2$. The structures on the left are intended only to show the location of the mutations. *The Figure is adopted from paper: Guo, Z.; Karunatilaka, K. S.; Rueda, D., Single-molecule analysis of protein-free U2-U6 snRNAs. Nat Struct Mol Biol 2009, 16 (11), 1154-9.*

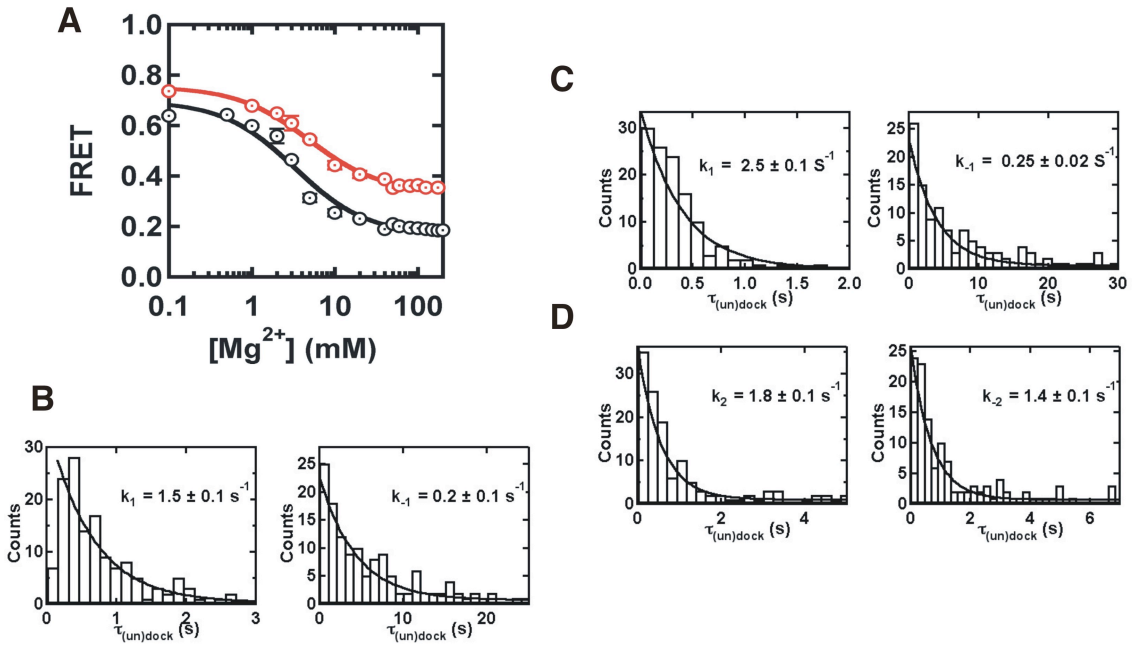


Figure 2.7 Six fold flipped-triad mutations and A59C mutation (**A**) Ensemble-averaged magnesium titration of the fluorophore labeled U2/U6 complex with the six fold flipped-triad mutations (red). A fit to the modified Hill equation yields a dissociation constant K_{Mg} = 4.8 ± 0.5 mM, similar to the wild type (black). However, the observed FRET ratio at high [Mg^{2+}] is significantly higher than the wild type (black), in agreement with the single-molecule data for this mutant. This data supports that Helix IB only forms in the low FRET state (G). (**B, C, D**) Dwell time analysis for the rate constants of triad mutant, single mutant and A91G mutant. (**B**) Dwell time analysis for the rate constants k_1, k_{-1} of triad mutant. The values are comparable to that of the wild-type U2/U6 complex, suggesting that the six-fold mutation does not affect the transition between N and I. (**C**) Dwell time analysis for the rate constants k_1, k_{-1} of A59C single mutant. The values are comparable to that of the wild-type U2/U6 complex, suggesting that A59C mutation does not affect the transition between the N and I. (**D**) Dwell time analysis for the rate constants k_2, k_{-2} of A91G mutant. The value of k_{-2} is comparable to that of wild type, while the k_2 is 60-fold larger than that of wild type. This indicates that this mutation favors the formation of G. *The Figure is adopted from paper: Guo, Z.; Karunatilaka, K. S.; Rueda, D., Single-molecule analysis of protein-free U2-U6 snRNAs. Nat Struct Mol Biol 2009, 16 (11), 1154-9.*

longer in the mutant than in the wild type, and the corresponding smFRET histogram reveals a larger low FRET peak for the mutant than for the wild type. Dwell-time analysis shows that k_1, k_{-1} and k_{-2} are very similar to the folding rate constants for the wild type, but k_2 is 60-fold faster, in agreement with these observations (Fig. 2.7C,D).

Taken together, these results show that the second folding step corresponds to a spontaneous intramolecular junction migration from I to G. The viability of the six-fold mutant has never been tested *in vivo*, but mutations that block helix IB formation can be lethal for the cell[21,20], thus raising the interesting possibility that the observed dynamics *in vitro* are linked to spliceosomal activation *in vivo*.

2.3.6 The structural dynamics correlate with splicing activation

To test the role of the observed dynamics in spliceosomal activation, we used previously published mutations. The U6 mutation A59C (Fig. 2.6C) prevents helix IB formation and blocks the spliceosome between the two splicing steps *in vivo*[21]. The U2 compensatory mutation U23G (Fig. 2.6D) restores base pairing in helix IB and rescues the second step of splicing.

We have tested the effect of these mutations on the structural dynamics of U2–U6 in 40 mM Mg^{2+}. The trajectories for A59C show an effect on both the junction structure and its dynamics (Fig. 2.6C). The low FRET states (I and G) disappear, and a new state (~0.30 FRET) replaces them. Two possible scenarios explain these results. The junction mutation may disrupt the junction structure, trapping it in an intermediate conformation with the U6 5′ end pointing in an intermediate direction, explaining the observed FRET ratio. Alternatively, the mutation may accelerate the junction dynamics faster than our

time resolution; making the two FRET states appear as a single state with an average FRET value between 0.2 and 0.4.

The trajectories for the compensatory mutation clearly show the rescue of the low FRET state and the structural dynamics (Fig. 2.6D), in agreement with the *in vivo* experiments. Notably, the FRET histogram shows that the intermediate-state FRET ratio is 0.32, indicating that the disrupted junction structure, rather than the averaged fast dynamics, is the most probable explanation for the A59C mutant intermediate FRET ratio.

These results link the observed structural dynamics of the U2–U6 complex to a conformational rearrangement before the second splicing step, suggesting that junction dynamics may contribute to successful splicing *in vivo.*

2.4 Conclusions

We have recently elucidated the folding pathway of spliceosomal snRNAs U2/U6, which has structural and catalytic similarities to the a self-splicing group II intron ribozyme[31,17]. Its pathway also involves obligatory intermediates and is dominated by a Mg^{2+} capture step that activates junction dynamics for catalysis. The similarities between the U2–U6 structural dynamics and those of the group II intron ribozyme now expand the parallels between these two enzymes and suggest the existence of evolutionarily conserved structural dynamics.

The U2–U6 complex lies at the heart of the catalytic core of the eukaryotic spliceosome, but its structure has been highly debated. Here, we have used smFRET to show that this important RNA complex acts as a Mg^{2+}-dependent conformational switch

that can adopt at least three distinct conformations. This supports the hypothesis that U2–U6 adopts multiple conformations at various splicing stages[24], which may or may not reflect the presence of unique active sites for each step[32].

In the high FRET conformation (N), the AGC triad extends the ISL as predicted for the four-helix structure[11], and the ACAGAGA loop and U80 in the ISL are brought into close proximity by a stabilizing tertiary contact. Because of the expected relationship between these three regions during the first splicing step, it is possible that this conformation resembles the active conformation. The recent group II intron ribozyme crystal structure provides a framework for this scenario[18]. In this structure, C358G359C360 (proposed equivalents of the AGC triad) are clearly base-paired with G383U384G385, extending domain 5 (proposed equivalent of the U6 ISL).

In the mid-FRET conformation, the AGC triad still forms the extended ISL (four-helix structure)[11], but the tertiary contact between the ACAGAGA loop and U80 is no longer formed, as its stability does not depend on the U80 mutations (Chapter 3). Our current data, however, do not allow us to conclude whether this conformation has a more direct role in splicing.

In the low FRET conformation, a junction migration takes place, resulting in the formation of helix IB (three-helix junction)[20]. Helix I is important for both steps of splicing, and a conformational rearrangement must precede each step[22]. This result is in agreement with our mutational data that link the formation of the low FRET conformation to an activating step between the first and second splicing steps. It is tempting to propose that the observed structural dynamics also have an activating role in

the first splicing step, but the mutations tested here do not provide enough support for this conclusion.

The role of Mg^{2+} ions at different stages of spliceosomal activation and catalysis has been established by experiments using phosphorothioate-substituted RNAs[33,34,35,36]. A key sensitive nucleotide is the phylogenetically conserved U80, which suggests that this base either has a direct role in catalysis or is at least involved in a closely related step[34]. Phosphorothioate substitutions in the 5′ and 3′ splice sites also involve Mg^{2+} ions in both splicing steps by activating the nucleophilic attack and stabilizing the leaving groups[33,35], and this suggests the presence of a Mg^{2+}-dependent conformational change specific to the second step of splicing[35,36]. We propose that this Mg^{2+}-induced conformational change corresponds to the one observed here. On the basis of this model, however, one might expect an inverse effect, whereby Mg^{2+} ions would preferentially stabilize the high FRET structure. In the presence of an essentially constant supply of Mg^{2+} ions *in vivo*, it is possible that the role of at least some of the spliceosomal proteins is to adjust the relative stability of these conformations to time the U2–U6 structural dynamics for accurate and efficient splicing.

We removed the highly conserved U6 ISL loop, but in accordance with a previous genetic study in which the ISL was extended by one base pair, we do not expect this loop to significantly affect the structure of U2–U6 *in vitro*[37]. Nonetheless, we have tested the effect of this deletion on the U2–U6 structural dynamics using a two-strand construct (Fig 2.8). Our results show that the two constructs behave almost identically. The three-

strand construct, however, offers greater efficiency of synthesis, labeling and purification, and will facilitate future mutational studies.

The whole chapter 2 is adopted from paper: Guo, Z.; Karunatilaka, K. S.; Rueda, D., Single-molecule analysis of protein-free U2-U6 snRNAs. Nat Struct Mol Biol 2009, 16 (11), 1154-9.

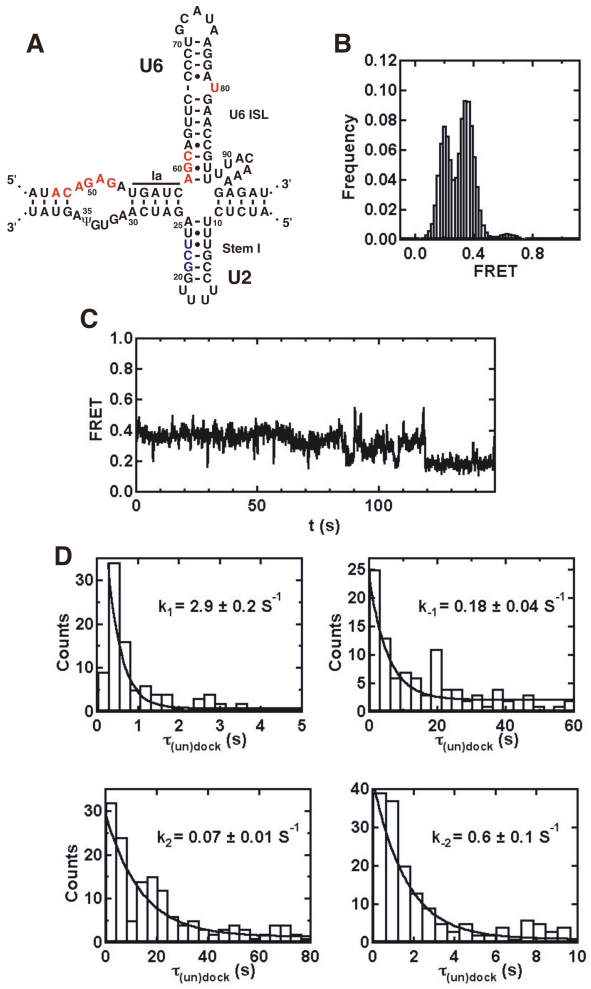


Figure 2.8 Single molecule data for the construct with an intact GNRUA loop (**A**) The construct with GNRUA loop used in smFRET experiments. The donor fluorophore is at the 5' end of U6, the acceptor is on the U6 base U70 and the biotin is on the 3' end of U6, equivalent to the three-strand construct. (**B**) FRET histogram of this 2-strand construct in 40 mM $MgCl_2$ There are 2 major peaks at FRET 0.2 and 0.4, which are comparable to that of the 3-strand construct. (**C**) Single-molecule trajectory of the 2-strand construct. The molecule shows dynamic transitions between the 0.2, 0.4 and 0.6 states, which are similar to the wild type. (**D**) Dwell time analysis for the rate constants of the 2-strand construct. The resulting folding rate constants for the two-strand construct are all within two-fold of the three-strand construct folding rates. These smFRET data of the 2-strand U2/U6 complex validate the use of the three-strand construct. The buffer conditions for these experiments are 50 mM Tris-HCl, pH 7.5, 40 mM $MgCl_2$, 100 mM NaCl. *The Figure is adopted from paper: Guo, Z.; Karunatilaka, K. S.; Rueda, D., Single-molecule analysis of protein-free U2-U6 snRNAs. Nat Struct Mol Biol 2009, 16 (11), 1154-9.*

2.5 References

1. Burge, C. B.; Tuschl, T. A.; Sharp, P. A., Splicing of precursors to mRNAs by the spliceosomes. *in The RNA World 2nd edition, R. F. Gesteland, T. R. Cech, J. F. Atkins (eds.), Cold Spring Harbor Laborotory Press, Cold Spring Harbor, NY, 1999, 525-560.* **1999**.

2. Kalnina, Z.; Zayakin, P.; Silina, K.; Line, A., Alterations of pre-mRNA splicing in cancer. *Genes Chromosomes Cancer* **2005,** *42* (4), 342-57.

3. Wang, G. S.; Cooper, T. A., Splicing in disease: disruption of the splicing code and the decoding machinery. *Nat Rev Genet* **2007,** *8* (10), 749-61.

4. Licatalosi, D. D.; Darnell, R. B., Splicing regulation in neurologic disease. *Neuron* **2006,** *52* (1), 93-101.

5. Wahl, M. C.; Will, C. L.; Luhrmann, R., The spliceosome: design principles of a dynamic RNP machine. *Cell* **2009,** *136* (4), 701-18.

6. Staley, J. P.; Guthrie, C., Mechanical devices of the spliceosome: motors, clocks, springs, and things. *Cell* **1998,** *92* (3), 315-26.

7. Black, D. L., Mechanisms of alternative pre-messenger RNA splicing. *Annu Rev Biochem* **2003,** *72*, 291-336.

8. Brow, D. A., Allosteric cascade of spliceosome activation. *Annu Rev Genet* **2002,** *36*, 333-60.

9. Parker, R.; Siliciano, P. G.; Guthrie, C., Recognition of the TACTAAC box during mRNA splicing in yeast involves base pairing to the U2-like snRNA. *Cell* **1987,** *49* (2), 229-39.

10. Lesser, C. F.; Guthrie, C., Mutations in U6 snRNA that alter splice site specificity: implications for the active site. *Science* **1993,** *262* (5142), 1982-8.

11. Sashital, D. G.; Cornilescu, G.; McManus, C. J.; Brow, D. A.; Butcher, S. E., U2-U6 RNA folding reveals a group II intron-like domain and a four-helix junction. *Nat Struct Mol Biol* **2004,** *11* (12), 1237-42.

12. O'Keefe, R. T.; Norman, C.; Newman, A. J., The invariant U5 snRNA loop 1 sequence is dispensable for the first catalytic step of pre-mRNA splicing in yeast. *Cell* **1996,** *86* (4), 679-89.

13. Segault, V.; Will, C. L.; Polycarpou-Schwarz, M.; Mattaj, I. W.; Branlant, C.; Luhrmann, R., Conserved loop I of U5 small nuclear RNA is dispensable for both catalytic steps of pre-mRNA splicing in HeLa nuclear extracts. *Mol Cell Biol* **1999,** *19* (4), 2782-90.

14. Valadkhan, S.; Manley, J. L., Splicing-related catalysis by protein-free snRNAs. *Nature* **2001,** *413* (6857), 701-7.

15. Valadkhan, S.; Mohammadi, A.; Jaladat, Y.; Geisler, S., Protein-free small nuclear RNAs catalyze a two-step splicing reaction. *Proc Natl Acad Sci U S A* **2009,** *106* (29), 11901-6.

16. Sontheimer, E. J.; Gordon, P. M.; Piccirilli, J. A., Metal ion catalysis during group II intron self-splicing: parallels with the spliceosome. *Genes Dev* **1999,** *13* (13), 1729-41.

17. Steiner, M.; Karunatilaka, K. S.; Sigel, R. K.; Rueda, D., Single-molecule studies of group II intron ribozymes. *Proc Natl Acad Sci U S A* **2008,** *105* (37), 13853-8.

18. Toor, N.; Keating, K. S.; Taylor, S. D.; Pyle, A. M., Crystal structure of a self-spliced group II intron. *Science* **2008,** *320* (5872), 77-82.

19. Abelson, J., Is the spliceosome a ribonucleoprotein enzyme? *Nat Struct Mol Biol* **2008,** *15* (12), 1235-7.

20. Madhani, H. D.; Guthrie, C., A novel base-pairing interaction between U2 and U6 snRNAs suggests a mechanism for the catalytic activation of the spliceosome. *Cell* **1992,** *71* (5), 803-17.

21. Hilliker, A. K.; Staley, J. P., Multiple functions for the invariant AGC triad of U6 snRNA. *RNA* **2004,** *10* (6), 921-8.

22. Mefford, M. A.; Staley, J. P., Evidence that U2/U6 helix I promotes both catalytic steps of pre-mRNA splicing and rearranges in between these steps. *RNA* **2009,** *15* (7), 1386-97.

23. Sun, J. S.; Manley, J. L., A novel U2-U6 snRNA structure is necessary for mammalian mRNA splicing. *Genes Dev* **1995,** *9* (7), 843-54.

24. Rhode, B. M.; Hartmuth, K.; Westhof, E.; Luhrmann, R., Proximity of conserved U6 and U2 snRNA elements to the 5' splice site region in activated spliceosomes. *EMBO J* **2006,** *25* (11), 2475-86.

25. Fabrizio, P.; Abelson, J., Thiophosphates in yeast U6 snRNA specifically affect pre-mRNA splicing in vitro. *Nucleic Acids Res* **1992,** *20* (14), 3659-64.

26. Yuan, F.; Griffin, L.; Phelps, L.; Buschmann, V.; Weston, K.; Greenbaum, N. L., Use of a novel Forster resonance energy transfer method to identify locations of site-

bound metal ions in the U2-U6 snRNA complex. *Nucleic Acids Res* **2007,** *35* (9), 2833-45.

27. (a) Zhao, R.; Rueda, D., RNA folding dynamics by single-molecule fluorescence resonance energy transfer. *Methods* **2009,** *49* (2), 112-7; (b) Aleman, E. A.; Lamichhane, R.; Rueda, D., Exploring RNA folding one molecule at a time. *Curr Opin Chem Biol* **2008,** *12* (6), 647-54.

28. Nir, E.; Michalet, X.; Hamadani, K. M.; Laurence, T. A.; Neuhauser, D.; Kovchegov, Y.; Weiss, S., Shot-noise limited single-molecule FRET histograms: comparison between theory and experiments. *J Phys Chem B* **2006,** *110* (44), 22103-24.

29. Kobitski, A. Y.; Nierth, A.; Helm, M.; Jaschke, A.; Nienhaus, G. U., Mg2+-dependent folding of a Diels-Alderase ribozyme probed by single-molecule FRET analysis. *Nucleic Acids Res* **2007,** *35* (6), 2047-59.

30. McKinney, S. A.; Freeman, A. D.; Lilley, D. M.; Ha, T., Observing spontaneous branch migration of Holliday junctions one step at a time. *Proc Natl Acad Sci U S A* **2005,** *102* (16), 5715-20.

31. Pyle, A. M.; Lambowitz, A. M., Group II Introns: Ribozymes That Splice RNA and Invade DNA. *in The RNA World, 3rd edition (ed. R.F. Gesteland et al.), pp. 469-505. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.* **2006**.

32. Moore, M. J.; Sharp, P. A., Evidence for two active sites in the spliceosome provided by stereochemistry of pre-mRNA splicing. *Nature* **1993,** *365* (6444), 364-8.

33. Sontheimer, E. J.; Sun, S.; Piccirilli, J. A., Metal ion catalysis during splicing of premessenger RNA. *Nature* **1997,** *388* (6644), 801-5.

34. Yean, S. L.; Wuenschell, G.; Termini, J.; Lin, R. J., Metal-ion coordination by U6 small nuclear RNA contributes to catalysis in the spliceosome. *Nature* **2000,** *408* (6814), 881-4.

35. Gordon, P. M.; Sontheimer, E. J.; Piccirilli, J. A., Metal ion catalysis during the exon-ligation step of nuclear pre-mRNA splicing: extending the parallels between the spliceosome and group II introns. *RNA* **2000,** *6* (2), 199-205.

36. Sontheimer, E. J., The spliceosome shows its metal. *Nat Struct Biol* **2001,** *8* (1), 11-3.

37. Sun, J. S.; Manley, J. L., The human U6 snRNA intramolecular helix: structural constraints and lack of sequence specificity. *RNA* **1997,** *3* (5), 514-26.

CHAPTER 3

Single Molecule Study of protein free U2/U6 Reveals Existence of Base Triples

3.1 Introduction

In eukaryotic cells, the immature precursor messenger RNA (pre-mRNA) must go through several RNA processing steps to be converted into the mature messenger RNA for translation[1,2]. Splicing is one of the essential steps in the maturation of the pre-mRNA, during which the non-coding sequences (introns) are removed and the sequences encoding proteins (exons) are ligated together[3,4].

In contrast to the autocatalytic self-splicing group II intron that is a ribozyme, the nuclear mRNA splicing is catalyzed by the spliceosome to accomplish, though both of them share exactly the same chemistry. The spliceosome is a highly dynamic ribonucleoprotein (RNP) complex, composed of 5 small nuclear RNAs (U1, U2, U4, U5 and U6 snRNAs) and more than 150 proteins[2,5]. Each snRNA binds a variety of spliceosomal proteins yielding small nuclear ribonucleoproteins (snRNPs). During the assembly, activation and catalysis of the spliceosome, the snRNPs bind and disassociate following a highly ordered and regulated pathway. During this process, the conformation of each snRNA undergoes dramatic rearrangements[6].

The group II intron and the nuclear spliceosome are similar in numerous ways including the splicing chemistry, catalytic mechanism, structure of conserved domains and metal ion binding[7]. It was proposed based on those similarities that the group II intron is the evolutionary ancestor of the spliceosome and the spliceosome may be a ribozyme[8].

In an active spliceosome, only U2, U6 and U5 snRNPs remain and U5 has been shown to be functionally dispensable for catalysis *in vitro*[9,10], so it is highly likely that the U2 and U6 snRNAs form the catalytic core of the spliceosome and play a pivotal, directly role in catalysis (Figure 3.1A)[11]. Protein free U2/U6 was shown to catalyze slow reactions similar to both steps of splicing, supporting the hypothesis that RNA performs catalysis in the spliceosome and the significance of the U2/U6 complex[12,13]. Lots of effort has been devoted to study the structure and function of the U2/U6 complex. U6 is phylogenetically conserved, the most conserved spliceosomal snRNA, from yeast to humans. The conserved residues were demonstrated to be essential for function by mutational studies[14-15] (Figure 3.1A). The U2 and U6 snRNAs form a stable complex through an extensive base-pairing network, recognizes pre-mRNA substrates, and bind catalytically essential metal ions[16]. Different conformations of the U2/U6 complex were discovered with a variety of approaches and it has been proposed that U2/U6 goes through conformational changes between the two steps of splicing to form multiple active sites[11,17-18].

However, in the absence of a high-resolution three-dimensional structure of the spliceosome, it is still impossible to access the detailed structural information of the

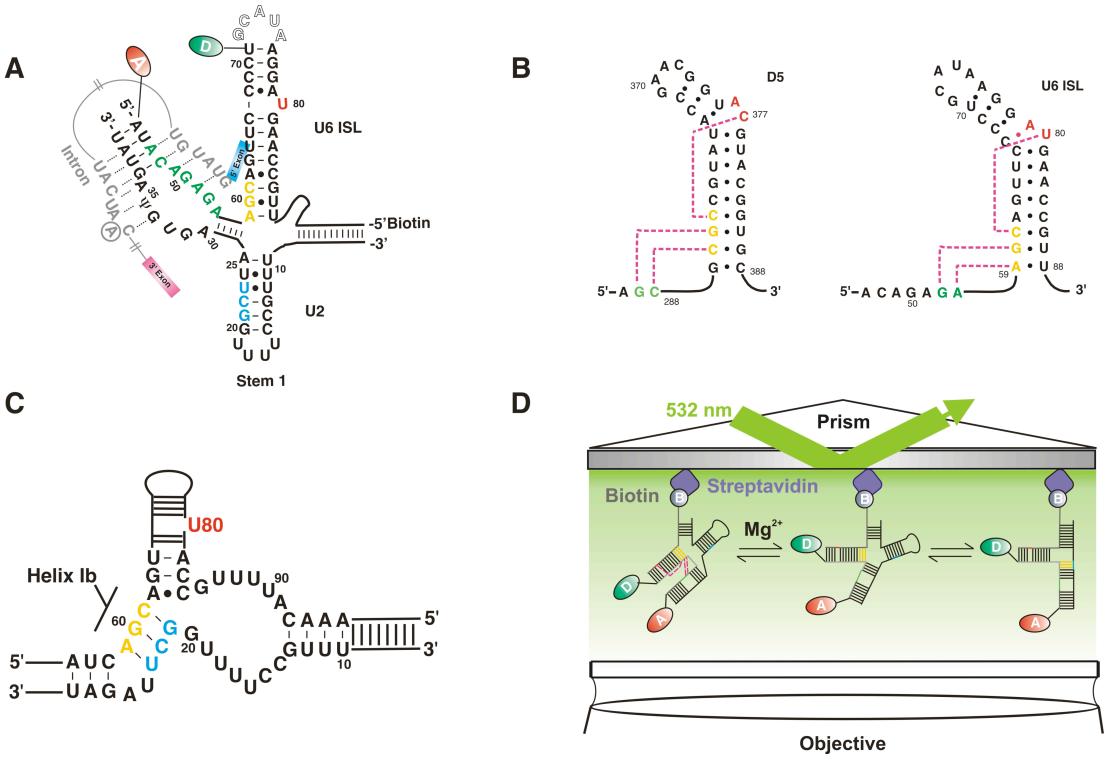


Figure 3.1 Secondary structure model of the spliceosomal snRNAs U2–U6 with base triple (**A**) Proposed secondary structure of spliceosomal snRNA U2/U6 from *Saccharomyces cerevisiae*[17]. Highly conserved regions are highlighted in colors (U80 in red, AGC triad in yellow and ACAGAGA loop in green). The region in U2 that pairs with AGC to form Helix IB is in blue. U2/U6 binds intron (grey) and bring 5' splice site and branch site in close proximity. The 5' end of U6 is labeled by FRET donor (Cy3) and we cut the GCAUA loop (outline) to facility the labeling of U70 by FRET acceptor (Cy5) (**B**) Secondary structure of DV of the group II intron (left) and U6 ISL of the spliceosome (right)[19]. Highly conserved regions are highlighted in the same color code as in (A). Red dash lines indicate base triple interactions discovered from recent crystal structure of the group II intron (left) and proposed base triples in the spliceosome (right). (**C**) Proposed secondary structure of Helix IB[11]. (**D**) Schematic diagram of single molecule spectroscopy with TIR excitation. Single U2-U6 molecules are immobilized via a Biotin-Streptavidin bridge. The donor fluorophore (D) is excited at 532 nm, and the donor and acceptor (A) fluorescence is collected through the objective and recorded by CCD camera.

spliceosomal snRNPs, specifically the functionally essential U2/U6 complex. Notably, the structure of group II intron contains five domains and among them, domain V (DV) is highly conserved both in term of sequences and secondary structure and is considered the active site of the group II intron[7,20] (Figure 3.1B). DV contains a catalytically important bulge (AC) and three consecutive nucleotides named the catalytic triad (usually AGC, CGC for some species)[21,22,23]. The structure and behavior of DV is very similar to that of the U6 intramolecular stem loop (U6 ISL) (Figure 3.1B). U80 is considered the analogue of the DV bulge and the AGC catalytic triad is conserved from the group II intron to the spliceosome. The structure of DV from a recent atomic-resolution crystal structure of group II intron could provide valuable information about the structure of the spliceosomal snRNAs U2/U6 complex[19-24]. The most surprising feature of DV found in this crystal structure is the formation of base triples between DV and conserved nucleotides in the linker between DII and DIII (J2/3) (Figure 3.1B and 3.2). The components participating in the base triples are far away from each other in primary sequence, but they are drawn together to form the active site. Previous studies and the similarities allow us to make structural parallels between DV and the U6. It has been hypothesized that the existence of similar base triples in the U2/U6 complex (Figure 3.1B)[19,25,26]. U80 has been proposed to interact with the C61-G86 base pair resulting in a base triple. In the center of this triple helix, the G60•U87 and a guanosine in the ACAGAGA region that is the analogue of J2/3 in U6 have been proposed to form a conventional base triple. Another base triple has been proposed between A59-U88 and an adenosine within the ACAGAGA loop[19,25]. This mode is very appealing because in this model the catalytically crucial Mg^{2+} ion bound by

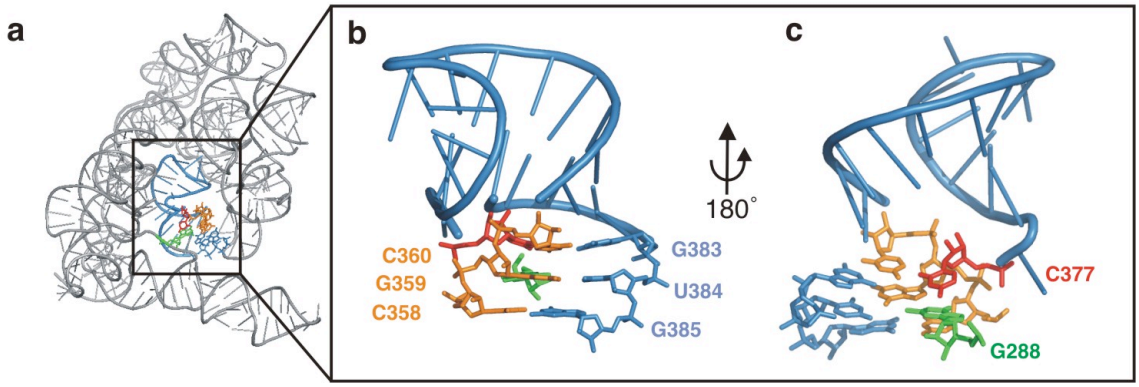


Figure 3.2 The recent crystal structure of the self-splicing Group II intron ribozyme supports the hypothesis that the high FRET conformation resembles the active conformation adopted during splicing given the similarity between the U6 ISL and the domain 5 of Group II intron. (**a**) The crystal structure of the whole self-splicing Group II intron. The boxed region is the catalytic active domain 5. (**b**) and (**c**) Higher magnifications of the boxed region (domain 5). In this structure, C358G359C360 (equivalents of the AGC triad, orange) are clearly base paired with G383U384G385 (blue), extending domain 5 (equivalent of the U6 ISL). In addition, the C360-G383 base pair forms a base triple with C377 (equivalent of U80, red), which in turn forms a base stacking interaction with G288 from J2/3 (equivalent of the ACAGAGA loop, green). Finally, G288 forms a base triple with the G359-U384 base pair as well. All these interactions could readily take place in the high FRET structure observed in our experiments. *The Figure is adopted from paper: Guo, Z.; Karunatilaka, K. S.; Rueda, D., Single-molecule analysis of protein-free U2-U6 snRNAs. Nat Struct Mol Biol 2009, 16 (11), 1154-9.*

U80 is brought in close proximity of the 5' splice site bound to the ACAGAGA loop, branch site and the invariant AGC triad[27,28]. The base triples bring the components necessary for the first step of splicing close together and should be pivotal for the function of spliceosome (Figure 3.1A).

Two key questions remain: Do the base triple interactions discovered in the group II intron also exist in the U2/U6 complex, and how is the catalytically important high FRET structure stabilized by the base triple interactions. Here we report single molecule FRET data that provide further insight into these issues (Figure 3.1D). We elucidated the evidence showing the existence of base triples in the U2/U6 complex that in a similar fashion as in the group II intron. Our data agree well with the structure predictions based on the U2/U6 folding and a database of 3D structures using FR3D program[29]. The triple helix may play significant role in position substrate and metal ions required for first step splicing, as well as maintain active structure of the active core.

3.2 Materials and methods

3.2.1 RNA Preparation and Purification

There are 3 strands of RNA forming the U2-U6 complex in our construct (Figure 2.1D): U6-1, U6-2 and U2. The sequences of RNA samples used in this chapter are listed in Table 1. All the RNA samples were purchased from Keck Foundation Biotechnology Resource Laboratory at Yale University School of Medicine (New Haven, CT). These RNA samples have 2'OH protection groups that are removed by a 2'OH deprotection reaction according to the manufacture's protocol. The deprotected RNA was purified as

described in Chapter 2. The 5' end of U6-1 is labeled with Cy5 during the RNA synthesis and a C7 amino linker is attached to 3' end, enabling the labeling of a Cy3 to the 3' end for our FRET studies. The RNAs were labeled and purified as described in Chapter 2. RNA concentrations were measured by UV-Vis absorbance at 260 nm.

3.2.2 Single-molecule FRET

The three RNA strands were heated to 90°C for 45s and annealed at room temperature over 20 min at 1 μM concentration in a 10 μL solution with the standard buffer (50 mM Tris-HCl, pH 7.5, 100 mM NaCl) and variable concentrations of Mg^{2+} ions. The single molecule experiments were conducted as described in Chapter 2. The histograms constructed with Matlab.

3.2.3 Base triple predictions by FR3D

All of the base triple predictions were obtained from the website of a database of 3D structures using FR3D program (http://rna.bgsu.edu/Triples/triples.php). There are two ways to search for a given base triple by base triple family or by combination. Base triples are divided into families according to their glycosidic bonds orientations and the interacting edges (for examples: Watson-Crick/Watson-Crick, Hoogsteen/Sugar Edge, Watson-Crick/Hoogsteen). All the structures showed in this chapter are constructed according to this website with ChemDraw. Comparisons of geometries are based on the PDB files of base triple structure downloaded from this website and constructed in Pymol.

Table 2 Sequences of RNA used in chapter 3

WT U6-1	5'-AUA CAG AGA UGA UCA GCA GUU CCC CU-3'
WT U6-2	5'-AGG AUG AAC CGU UUU ACA AAG AGA UUU AU-3'
WT U2	5'-AUC UCU UUG CCU UUU GGC UUA GAU CAA GUG UAG UAU-3'
G50C U6-1	5'- AUA CAC AGA UGA UCA GCA GUU CCC CU- 3'
A51C U6-1	5'- AUA CAG CGA UGA UCA GCA GUU CCC CU- 3'
G52C U6-1	5'- AUA CAG ACA UGA UCA GCA GUU CCC CU- 3'
A53C U6-1	5'- AUA CAG AGC UGA UCA GCA GUU CCC CU- 3'
U80A U6-2	5'-AGG AAG AAC CGU UUU ACA AAG AGA UUU AU - 3'
U80C U6-2	5'-AGG ACG AAC CGU UUU ACA AAG AGA UUU AU - 3'
U80G U6-2	5'-AGG AGG AAC CGU UUU ACA AAG AGA UUU AU - 3'
U80Δ U6-2	5'-AGG AG AAC CGU UUU ACA AAG AGA UUU AU - 3'
G52CG60A U6-1	5'-AUA CAG ACA UGA UCA ACA GUU CCC CU- 3'
U88C U6-2	5'-AGG AUG AAC CGU CUU ACA AAG AGA UUU AU- 3'
U87C U6-2	5'-AGG AUG AAC CGC UUU ACA AAG AGA UUU AU- 3'
U88A U6-2	5'-AGG AUG AAC CGU AUU ACA AAG AGA UUU AU- 3'
A59U U6-1	5'-AUA CAG AGA UGA UCU GCA GUU CCC CU- 3'
C61G U6-1	5'-AUA CAG AGA UGA UCA GGA GUU CCC CU- 3'
G86C U6-2	5'-AGG AUG AAC CCU UUU ACA AAG AGA UUU AU- 3'
G86CU80A U6-2	5'-AGG AAG AAC CCU UUU ACA AAG AGA UUU AU- 3'
G86CU80c U6-2	5'-AGG ACG AAC CCU UUU ACA AAG AGA UUU AU- 3'

3.3 Results

3.3.1 A base triple forms between the metal binding U80 bulge and the G86-C61 base pair

According to the crystal structure of the group II intron, in DV, C360 and G383 form a Watson-Crick base pair and the sugar edge of C377 from DV bulge interacts with Hoogsteen edge of C360 yielding G383-C360•C377 base triple. The hypothetical base triple G86-C61•U80 in U2/U6 is proposed to form in the same family adopting the same conformation (Figure 3.1B and 3.4A,B).

To confirm the existence of G86-C61•U80 base triple, we tested mutations within this base triple. The results are reported in table 3. As predicted by the model, if there are base triple interactions in the U2/U6 complex, the regions where the two fluorophores attached should be in close proximity and the high FRET should be observed, if we disrupt the base triple, the high FRET state should be destabilized.

The U6 base U80 is highly conserved and has been involved in binding a metal ion important for catalysis[30]. To assess its role in base triple formation, we deleted and mutated it to A, G or C (Figure 3.3) and with single molecule FRET, which can uncover key structural and dynamic information otherwise hidden in bulk studies and allows to record movies of molecular motion and chemical activity of individual complexes[31,32,33,34]. The U80G mutation and deletion are lethal in yeast, whereas U80A and U80C are viable[35]. The corresponding 5 mM Mg^{2+} smFRET histograms (Figure 3.3)

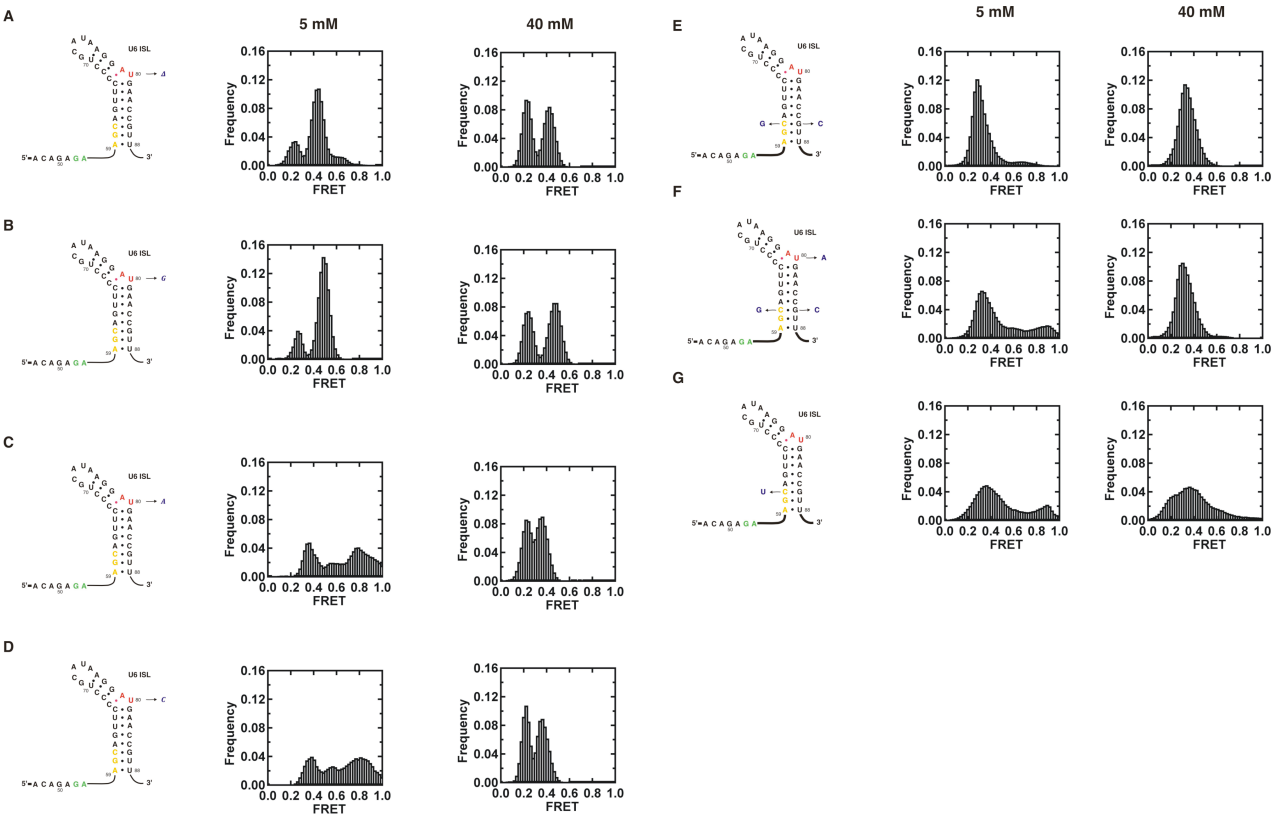


Figure 3.3 smFRET histograms for the U80 mutations (U80, U80G, U80A and U80C) and nucleotides involved in base triple G86-C61•U80 in 5 and 40 mM Mg^{2+}, as indicated. (**A, B**) The U80 deletion and G mutation result in the disappearance of the high FRET conformation in the low-[Mg^{2+}] histogram, but do not affect the intermediate and low FRET conformations in the high-[Mg^{2+}] histogram. (**C, D**) The smFRET histograms for U80A and U80C are similar to that of the wild type. (**E**) C61G/G86C mutant disrupt high FRET structure (**F**) C61G/G86C/U80A mutations can restore the high FRET, consistent with the predictions. (**G**) C61U can retain the high FRET state, which is agree well with the model.

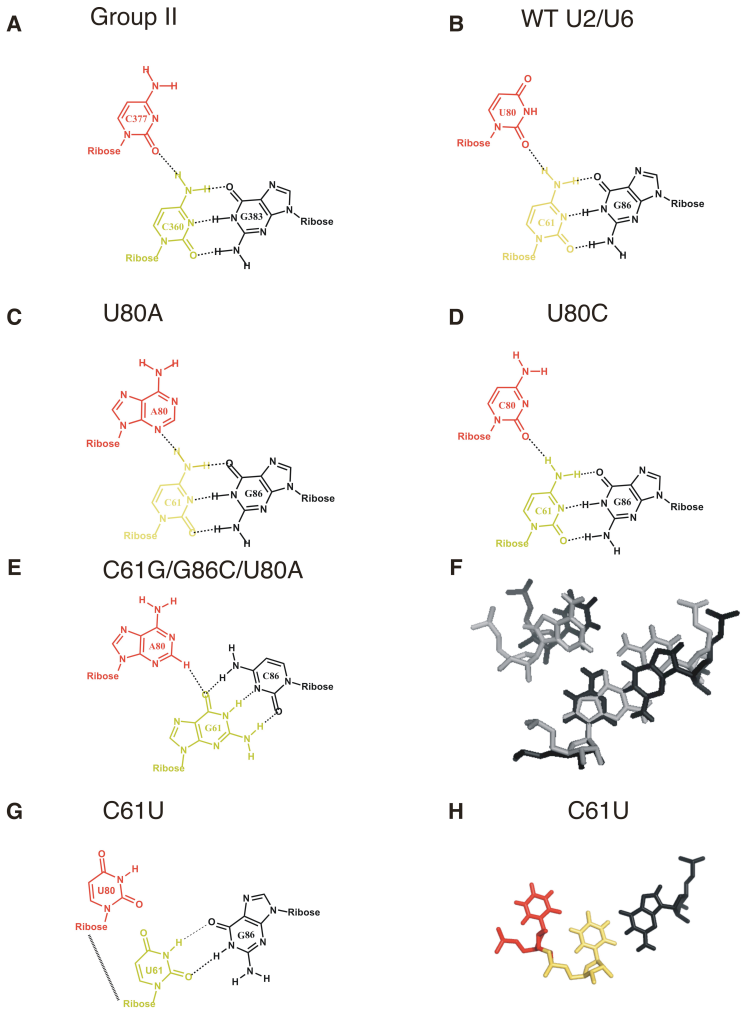


Figure 3.4 Structure models according to FR3D[29] (A-E) Structures of base triples according to FR3D[29]. (**A**) G383-C360•C377 base triple discovered in the group II intron. (**B**) G86-C61•U80 base triple proposed in U2/U6. (**C**) G86-C61•A80 base triple proposed in U2/U6. (**D**) G86-C61•C80 base triple proposed in U2/U6 (**E**) C86-G61•A80 base triple proposed in U2/U6. (**F**) Comparison of geometry of wild type G86-C61•U80 base triple (Black) and mutant C86-G61•A80 base triple (Grey) according to FR3D[29].(**G**) G86-U61•U80 base triple proposed in U2/U6.The interaction between U61 and U80 is through the backbone. (**H**) Conformation of G86-U61•U80 base triple according to crystal structure from FR3D[29].

show that the broad distribution characteristic of the high FRET conformation disappears for the U80G mutation and deletion, but it remains present in both the U80A and U80C mutants. In 40 mM Mg^{2+}, however, the histograms are similar for all mutants (Figure 3.3). The centers of the intermediate FRET distributions for the U80 deletion and U80G mutant shift to 0.43 and 0.47, respectively, indicating that the ISL bends differently in these mutants. A comparison of the NMR structures of the wild-type ISL and the U80G mutant shows a ~12Å displacement of U70 (location of Cy3), supporting this idea[36,37,34].

U80 deletion eliminates the high FRET conformation, possibly due to disruption of the G86-C61•U80 base triple since one base in this base triple is completely missing. U80G also destabilizes the high FRET state and could inhibit the formation of base triple G86-C61•U80 by base pairing with C67 resulting in the C67–G80 base pair, thereby changing the structure of the U6 ISL. U80A and U80C could still form base triple and the high FRET state, with the premise that the high FRET structure is stabilized by base triple interactions (Figure 3.1A)[34]. A database of 3D structures using FR3D program shows that resulting base triples, by U80A and U80C respectively, GCA and GCC are possible and stable (Figure 3.4C,D)[29].

To disrupt the G86-C61•U80 base triple but maintain the U6 ISL stem, the C61-G86 base pair was swapped, altered to G61-C86. Most of the high FRET state is disrupted at 5mM $MgCl_2$, under which condition we usually see dominant high FRET with wild type U2/U6. FR3D predicts the resulting C-G•U to be unstable, and the data support the hypothesis that this triple is important in forming the high FRET state. Furthermore, at 40mM Mg^{2+}, there is only one major peak with FRET 0.3. We reported

in Chapter 2 that transitions to the 0.2 FRET state is corresponding to the formation of helix IB (Figure 3.1C) and mutations that destabilize Helix IB eliminate the 0.2 FRET state[34]. Similar to other mutations that disrupt the formation of helix IB, C61G also destabilizes the 0.2 state, conforming our earlier conclusion (Figure 3.3E and see below).

To recue the formation of the base triple from the swap mutant G61-C80, according to prediction of FR3D, the U80 was mutated to A. Single molecule data at 5 mM Mg^{2+} of this mutant show inefficient recovery of the high FRET state (Figure 3.3F). This may be because this base triple serves as the stacking terminal of other two triples and the CGA base triple is not quite isosteric; changing U with a larger purine may alter the geometry and interfere the folding (Figure 3.4E,F).

Another mutation, C61U, is expected to retain the formation of base triple (Figure 3.4G,H). We observe high FRET in 5 mM $MgCl_2$ with this mutation, confirming our model (Figure 3.3G). The different effect of C61G/G86C and C61U on high FRET structure may explain why C61G/G86C is lethal while spliceosomes with a C61U mutation still retain 50%-80% splicing activity *in vitro*[35,38].

In summary, the single molecule FRET data provide evidence supporting the hypothesis that a base triple forms between the U80 bulge of the U6 ISL and the C61-G86 base pair in a similar fashion as in DV of the group II intron (Figure 3.4A,B). This base triple may play an important role in bringing a proven catalytically essential metal ion into the active site and stabilizing the active structure of spliceosome.

3.3.2 Cytosine scanning reveals nucleotides possibly involved in base triple formation

In DV of the group II intron, G359 and U384 form a Wobble base pair and the Hoogsteen edge of G288 from J2/3 pairs with G359 Watson edge to form U384•G359•G288 base triple. Using the same set of interactions, C289 interact with Watson-Crick base pair C358-G385, resulting in C358-G385•C289 base triple (Figure 3.1B and 3.6A,B). We propose that in U2/U6, two nucleotides in the ACAGAGA loop that is the counterpart of J2/3 of the group II intron form base triple with U87•G60 and U88-A59 respectively adopting same conformations (Figure 3.1B and 3.6C,D).

We mutated the U2 bases 31–36 to replace the ACAGAGA loop with a double-stranded helix (Figure 3.5A)[34]. The corresponding single-molecule experiments in 5 mM Mg^{2+} also show the disappearance of the high FRET conformation (Figure 3.5B). These results show that the high FRET state is stabilized by a tertiary interaction involving ACAGAGA loop. The presence or absence of the high FRET conformation in all of these mutants correlates well with viable or lethal mutations in yeast, supporting the idea that the *in vitro* dynamics may be important for splicing *in vivo*[35,34]. To further demonstrate which nucleotides are involved in the base triple formation, we mutated the region covering G50-A53 to cytosine individually (Figure 3.5C-F).

To form the hypothetical base triple U88-A59•A, U88-A59 interacts with an adenosine in the ACAGAGA loop. Changing this adenosine to a cytosine would allow the formation of a U88-A59•C base triple (Figure 3.6E). Furthermore, according to the crystal structure of the group II intron, cytosine in this position has an additional interaction with a phosphate group of DV backbone (Figure 3.6F). Similarly, this A to C

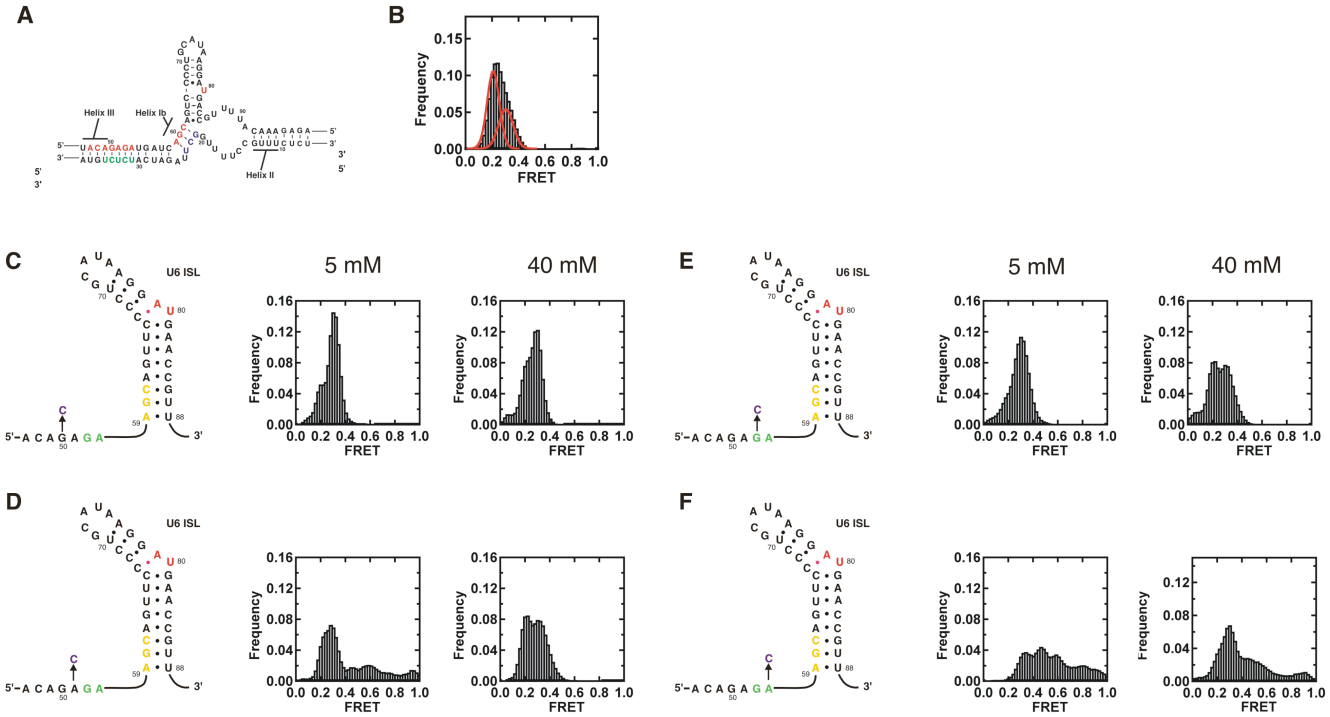


Figure 3.5 Cytosine scanning (**A**) The construct with ACAGAGA loop deletion. The U2 bases in green were mutated to base pair with the ACAGAGA sequence in U6. (**B**) The smFRET histogram of ACAGAGA loop deletion mutant in 5 mM $MgCl_2$. There are 2 major peaks at ~0.2 and ~0.3 FRET. (**C-F**) smFRET histograms for the cytosine scanning in the ACAGAGA loop (in 5 and 40 mM Mg^{2+}, as indicated). (**C, E**) The G50C and G52C mutation result in the disappearance of the high FRET conformation in the low-[Mg^{2+}] histogram. (**D**) The smFRET histograms for A51C are similar to that of the wild type. (**F**) A53C stabilizes high FRET both with low and high Mg^{2+} concentrations. These results indicate that G50 or G52 and A53 are involved in a functionally important base triple tertiary contact that stabilizes the high FRET conformation. *The Figure A is adopted from paper: Guo, Z.; Karunatilaka, K. S.; Rueda, D., Single-molecule analysis of protein-free U2-U6 snRNAs. Nat Struct Mol Biol 2009, 16 (11), 1154-9.*

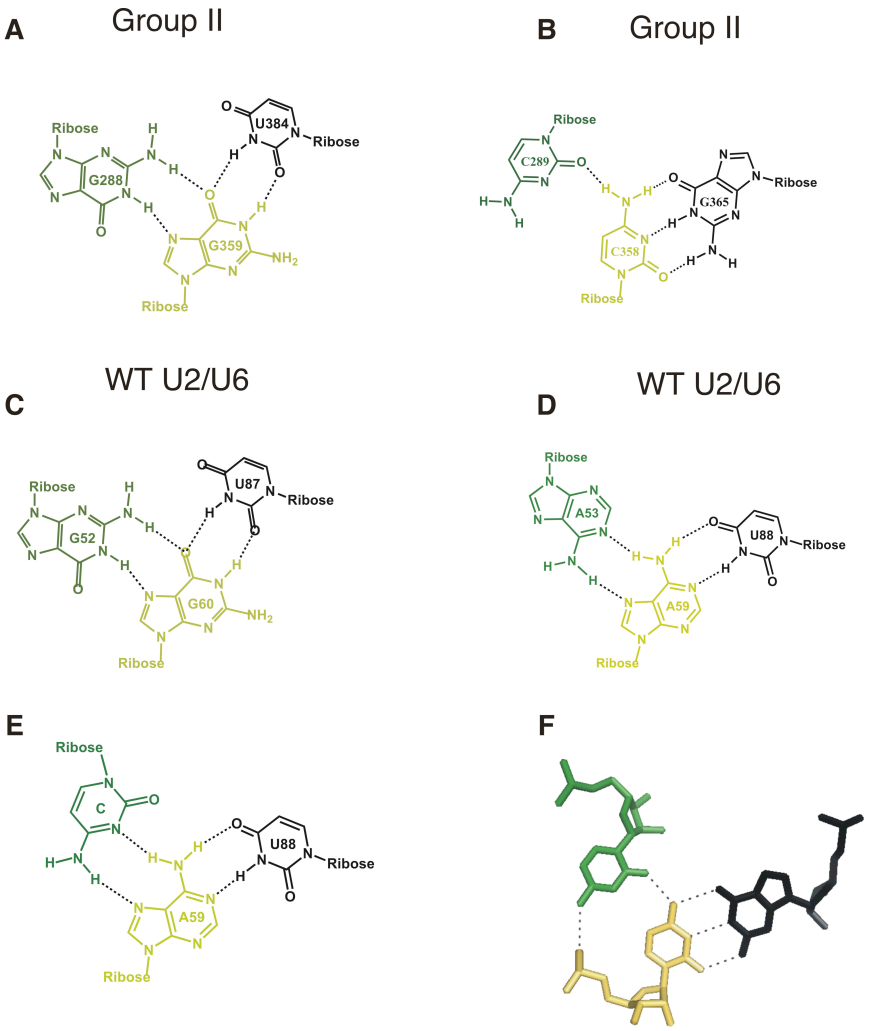


Figure 3.6 Structure models according to FR3D[29] (A-E) Structures of base triples according to FR3D. (**A**) U384• G359•G288 base triple discovered in the group II intron. (**B**) C358-G385•C289 base triple discovered in the group II intron. (**C**) U87•G60•G52 base triple proposed in U2/U6. (**D**) U88-A59•A53 base triple proposed in U2/U6. (**E**) U88-A59•C53 base triple proposed in U2/U6. (**F**) The structure of C358-G385•C289 base triple according crystal structure of the group II intron. There is an additional interaction between the C289 and the phosphate group in the backbone.

mutation should stabilize the high FRET more. According to the single molecule data, A51C shows near WT behavior with both 5mM and 40 mM Mg^{2+}, while the A53C mutant stabilizes the high FRET state at both 5mM Mg^{2+} and 40mM (Figure 3.5D, F). The A51 in the ACAGAGA loop has been shown to interact with the 5' splice site and plays an important role in the recognition and positioning of the 5' splice site[16]. The near WT result obtained with the A51C mutant further suggests that this nucleotide does not participate in the tertiary base triple interaction. A53C stabilizes the high FRET as we predicted, suggesting that the A53 is part of the U88-A59•A base triple.

According to our model, U87•G60 interacts with a guanosine in the ACAGAGA loop, yielding a base triple. If we mutate this guanosine to a cytosine, the base triple U87•G60•G is predicted to be disrupted due to steric clashes. We found that both G50C and G52C can destabilize the high FRET structure with 5mM Mg^{2+} (Figure 3.5C, E). We hypothesize that G52 is involved in the base triple rather than G50. Because in group II intron, two consecutive nucleotides in J2/3 (G288C289) form base triples with DV, we think it is more likely that it is two similar consecutive nucleotides G52A53 that participate in the triplex in U2/U6. G50C may cause the formation of other alternative structures with low FRET. However, we cannot totally rule out the possibility that G50 is part of the base triple (Also see below). All the data are summarized in table 3.

In summary, according to our data, nucleotides G52A53 most probably participates in the base triple interactions.

3.3.3. Base U88-A59•A53 may not be essential in stabilizing the high FRET structure

According to our cytosine scanning results, we proposed that A53 in the ACAGAGA loop interacts with the A59-U88 base pair to form a base triple (Figure 3.1B). To further test the existence and the influence of this proposed base triple to the overall structure, a serial of mutants were tested. All the data are summarized in table 3.

With an A59C mutation, both the U6 ISL and the base triple are predicted to be disrupted. Single molecule data show that the high FRET structure is destabilized greatly but not completely at 5mM $MgCl_2$ (Figure 3.7A). However, either destabilization of the junction of the U6 ISL or the base triple may reduce the high FRET state. So we designed another mutant in which the A59-U88 base pair was swapped to U59-A88. This mutation retains the stability of the U6 ISL, and only the base triple is destabilized according to FR3D[29]. We find that the high FRET state was not disrupted significantly by this mutation at 5 mM $MgCl_2$, suggesting that destabilization of this base triple is not sufficient to disrupt the high FRET structure with the other two base triples in place and the reduction of high FRET observed with A59C may be due to the disruption of the U6 ISL that is important for numerous tertiary interactions (Figure 3.7B). This is consistent with the observation that the analog of this base triple in DV of the group II intron (G385-C358•C289) only has a supporting role in stabilizing the structure and the same flipped mutation in the group II intron dose not effect splicing *in vivo*[39]. Furthermore, with this swap mutant, at 40mM $MgCl_2$, the majority of the molecules adopt the 0.3 state, and a small fraction stay in the high FERT state (Figure 3.7B). It is obvious that A59U can disrupt helix IB and it was demonstrated to inhibit the second step of splicing like

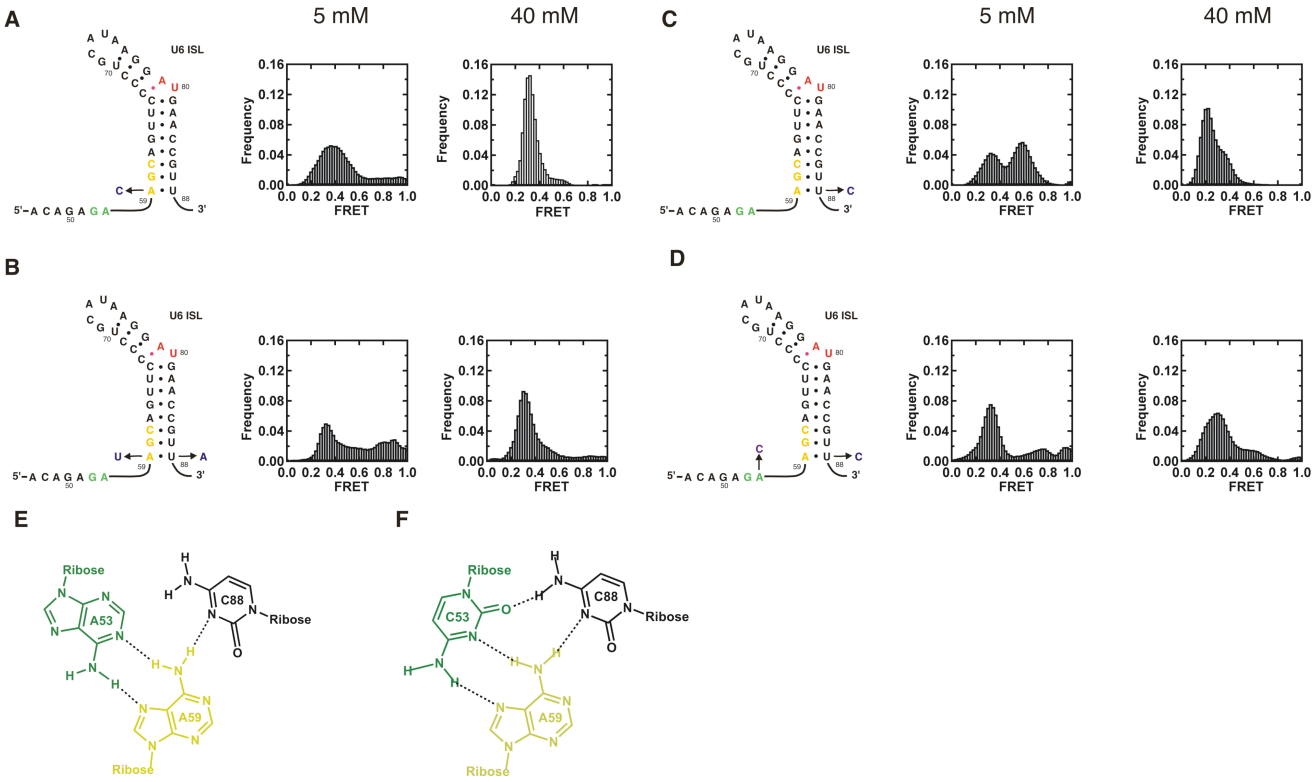


Figure 3.7 smFRET histograms for the mutations of nucleotides involved in base triple U88-A59•A53 in 5 and 40 mM Mg^{2+}, as indicated. (**A, C**) The A59C and U88C mutations result in the disappearance of the high FRET conformation in the low-[Mg2+] histograms. (**B**) The smFRET histogram A59U/U88A with low ion concentration is similar to that of the wild type. (**D**) U88C/A53C double mutant can recues the high FRET state from U88C alone. (**E, F**) Structures of base triples according to FR3D[29]. (**E**) C88-A59•A53 base triple proposed in U2/U6. (**F**) C88-A59•C53 base triple proposed in U2/U6. This base triple is more stable than the C88-A59•A53 base triple because of additional interactions.

A59C, but to an even greater extent[40,34,35]. These data confirm our pervious conclusion that the 0.2 state corresponds to the formation of Helix IB which plays a important role in the second step of the splicing[34,35] (See above).

The data suggest at this position of the triple helix it is the destabilization of the U6 ISL junction other than the base triple that effect the high FRET structure One can predict that if we only destabilize the junction not the base triple, the high FRET state should still be disrupted. We made a U88C mutant, which is predicted to form a stable CAA base triple and Helix IB, but the U6 ISL is destabilized by introducing the C•A mismatch (Figure 3.1A and 3.7 E). With 5 mM $MgCl_2$ most of the high FRET structures disappear from the histogram, leaving a 0.6 state peak, consistent with our prediction (Figure 3.7C). In addition, with 40mM $MgCl_2$, the majority of molecules adopt the 0.2 FRET structure and only small fraction is in the 0.3 state (Figure 3.7C). It is may be because the disruption of the U6 ISL allows the complex to form helix IB (0.2 state) more easily. The double mutant A53C/U88C destabilizes the U6 ISL as well, but compared to U88C alone, there are two additional interactions contributing to hold the base triple together; one is the hydrogen bond between the $-NH_3$ in C4 of C88 and the O2 of C53, and the other one is between the $-NH_3$ in C4 of C53 and the oxygen of phosphate group in backbone, the same as found in the crystal structure of the group II intron (Figure 3.7F). So the double mutant A53C/U88C is expected to populate the high FRET state (~0.8) while the U88C dose not. In our single molecule FRET assay, we observed high FRET with 5mM $MgCl_2$, which is less than the A53C because of the disruption of the U6 ISL. (Figure 3.7D and 3.5F)

In summary, similar to its analog in the group II intron, the hypothetical base triple U88A59A53 may only have a supporting role in stabilizing structure necessary for catalysis. Destabilizing only this base triple but keeping the other two will not significantly destroy the high FRET structure and functions associated with this structure. This conclusion is in good agreement with the *in vitro* splicing assay data, in which any mutation in the 53rd position is well tolerated, even though A53U disrupts the base triple. Neither A59C nor A59U/U88A is predicted to form a base triple but both retain high FRET structure to some extent. This is consistent with the *in vitro* splicing assay that the first step of splicing is not significantly affect by those mutations. But both of them destabilize the helix IB and inhibit the second step of splicing, explaining their lethality in yeast[38,35].

3.3.4 Proposed base triple U87•G60•G52 is invariant

G60 is invariant in U6 and its corresponding analog in the group II intron, G359, is invariant as well[22,7] (Figure 3.1B). From cytosine scanning, we find that mutating the third G (Either G50 or G52) of this base triple to C, with which the base triple is predicted to be impossible, disrupts the high FRET conformation. According to the FR3D[29], mutating the G60 to A or U87 to C should rescue the base triple formation from the G52C mutation and restore the high FRET state (Figure 3.9A,C). However, we found that neither of these mutants rescues the high FRET state (Figure 3.8A,B). With the double mutant G52C/G60A (Figure 3.8A), compared to the G52C only, only small fragment of the high FRET state was restored, suggesting the rescue is not quite efficient. It is may be because the G60A mutation changes the relative orientation and geometry of

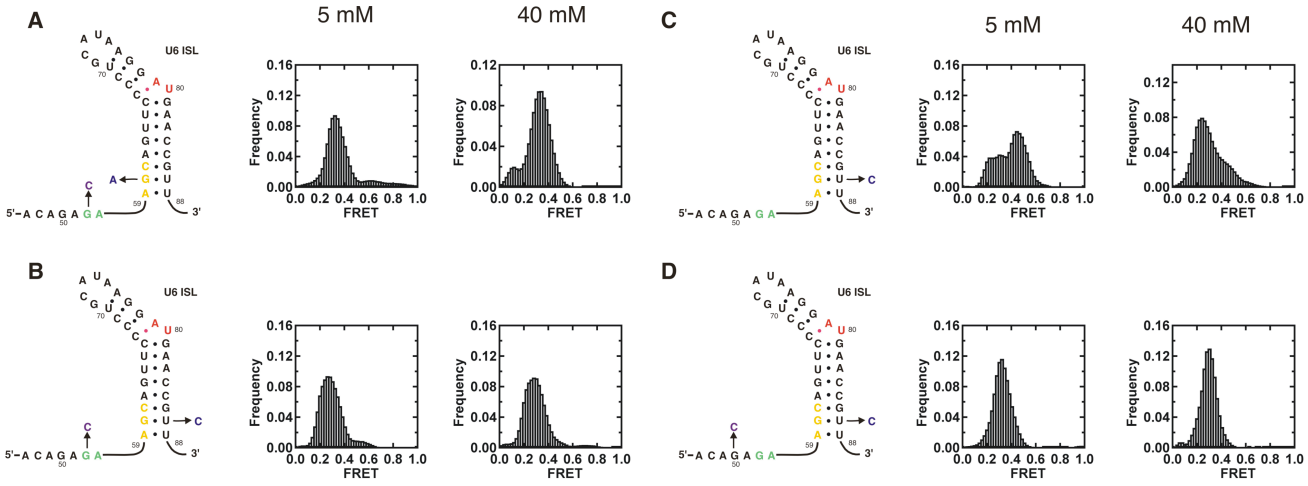


Figure 3.8 smFRET histograms for the mutations of nucleotides involved in base triple U87•G60•G52 in 5 and 40 mM Mg^{2+}, as indicated. (**A, B**) Neither G60A nor U87C can restore the high FRET. (**C**) U87C itself can decrease the population of high FRET. (**D**) U87C cannot rescue high FRET deficiency caused by G50C mutation. These data indicates that this base triple is invariant and may play an important role in stacking interactions stabilizing the active structure. The analog of this base triple in the group II intron is invariant as well[19].

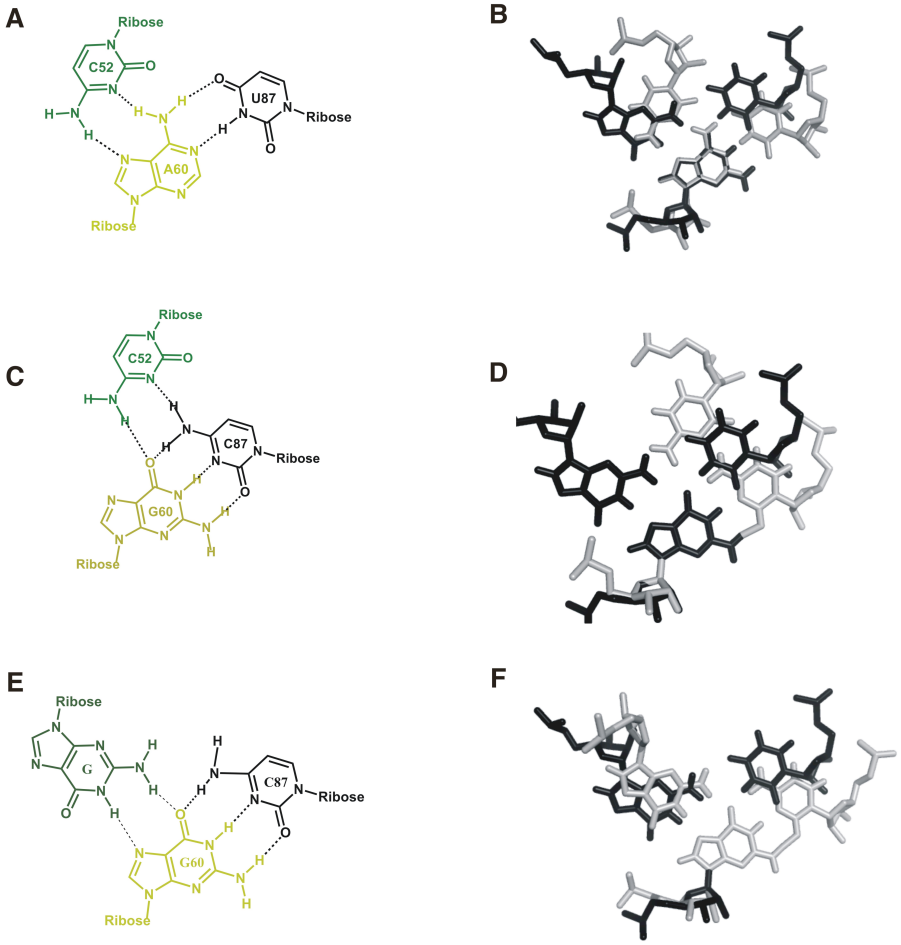


Figure 3.9 (A, C, E) Structures of base triples according to FR3D[29]. (**B, D, F**) comparisons of geometry of wild type (Black) and mutant base (Grey) triples according to FR3D[29]. (A) U87•A60•C52 base triple proposed in U2/U6. (B) Comparison of geometry of wild type U87•G60•G52 (Black) and mutant U87•A60•C52 base triple (Grey). (C) C87•G60•C52 base triple proposed in U2/U6. (D) Comparison of geometry of wild type U87•G60•G52 (Black) and mutant C87•G60•C52 base triple (Grey). (E) C87•G60•G52 base triple proposed in U2/U6. (F) Comparison of geometry of wild type U87•G60•G52 (Black) and mutant C87•G60•G52 base triple (Grey).

the base triple, breaking the stacking interactions (Figure 3.9B). This U87•G60•G52 base triple may be the most crucial for the stacking interactions holding all of three base triples together since it is in the middle of this base triple sandwich. U87C cannot suppress the high FRET deficiency caused by G52C mutation either (Figure 3.8B). U87C also twists the relative orientation of the base triple as G52C/G60A and probably breaks the stacking too (Figure 3.9D). It is quite surprising to find that a single U87C mutation is be able to change the histogram profile dramatically compared to the wild type U2/U6. With U87C, at 5mM $MgCl_2$, there are two major peaks in the histogram, one is a very broad peak between 0.2 and 0.4 and the other is a FRET 0.5 peak, in contrast to a major 0.9 and a minor 0.4 peak of the wild type U2/U6 (Figure 3.8C). The stable base triple from U87C mutation disrupt the high FRET structure either by distorting of the geometry as aforementioned (Figure 3.9F) or stabilizing the U6 ISL and constraining the dynamics essential for base triple formation or proper folding.

Alternatively, U87•G60 may form a base triple with G50 instead of G52. So we test the G50C/U87C. Even though the base triple C87-G60•C50 is predicted by FR3D to be stable, the high FRET deficiency was failed to be recued (Figure 3.8D). As a result, whether it is G50 or G52 that truly participate in the base triple may be difficult to indentify.

In conclusion, base triple U87•G60•G52 may be invariant, the same as its counterpart in the group II intron (base triple U384•G359•G288). Any mutation made in this triple changes the geometry and breaks the stacking. Compared to the other two base triples, the type of this base triple is more common and conventional[19]. It is highly

conserved indicating that it may play a very important role in stacking interactions and catalysis of spliceosome. All the data are summarized in table 3.

Table 3. Summary of smFRET results of hypothetic base triples

Mutant	Phenotype	BT Prediction	Predicted	Observed
U80 deletion	Lethal	No	Low FRET	Low FRET
U80A	80%-100% WT	Yes	High FRET	High FRET
U80C	80%-100% WT	Yes	High FRET	High FRET
U80G	Lethal	No	Low FRET	Low FRET
C61GG86C	Lethal	No	Low FRET	Low FRET +few high FRET
C61GG86CU80A	Not tested	Yes	High FRET	Medium fraction of high FRET
C61U	50%-80% WT	Yes	High FRET	High FRET
G52C	Lethal	No	Low FRET	Low FRET
U87C	Viable	Yes	High FRET	Low FRET
G60A/G52C	Not tested	Yes	High FRET	Low FRET+ few High FRET
U87C/G52C	Not tested	Yes	High FRET	Low FRET
U87C/G50C	Not tested	Yes	High FRET	Low FRET
A53C	80%-100% WT	Yes	High FRET	High FRET
A59C	Lethal	No	Low FRET	Low FRET +few high FRET
U88C	Viable	Yes	High FRET	Low FRET
A59U/U88A	Lethal	No	Low FRET	High FRET

3.4 Conclusions

Here we report the first direct structural evidence that supports the existence of the base triples in the spliceosomal snRNA U2/U6. These interactions were proposed according to a corresponding set of base triple interactions discovered in the recently published crystal structure of the self-splicing group II intron[19-24, 41-19, 42]. The data support the hypothesis that U2/U6 and the group II intron contain analogous base triples that are important for function.

The AGC triad in U6 is highly conserved and most mutations within it lead to a severe deficiency of splicing[11,40,35]. In contrast, mutations of their base pairing partners in U2 or in U6 ISL do not[11,15,35]. These data indicate that the AGC triad may have multiple roles besides forming base pairs in extended U6 ISL or helix IB. The recent crystal structure of group II intron suggest that the AGC triad may participate in base triple interactions that are important in maintaining the compact and delicate structure of the active core[19].

The putative base triples in U2/U6 may have a very significant role in the function of the spliceosome. They are able to draw essential parts for the first step of splicing in close proximity, which are otherwise far away from each other in primary sequence. The base triple G86-C61•U80 brings the catalytic important Mg^{2+} close to the proposed active site for the first step of splicing, containing the 5'splice site bound with ACAGAGA loop, branch site and the catalytic triad AGC. Both U88-A59•A53 and U87•G60•G52 bring the 5'splice site and the branch site to the active core (Figure 3.1A). But in contrast to the other two that play a central role in maintaining the overall structure and catalysis, the

U88-A59•A53 base triple may only have a supporting role in stabilizing the active structure, similar to its counterpart in the group II intron. This may be because it is on the bottom of this base triple sandwich and not important for the stacking. The G60 is invariant in U6 and so is the analogous G359 in the group II intron. This may be able to explain the failure of rescue of this base triple. Any change made in this base triple disrupts the overall structure, which might be able to be attributed to the breaking of the stacking interactions among three base triples.

The base triples may also play a significant role in position catalytically important Mg^{2+} ions for the two-metal ion catalysis proposed for the spliceosome. Most ribozymes require divalent ions such as Mg^{2+} as cofactors for catalysis[43]. Numerous biochemical and structural studies of the group II intron suggest a two-metal ion mechanism for the splicing reaction[7,22]. And the spliceosome may also utilize the same machanism[44]. It was revealed that catalytically important divalent ions bind in the DV bulge and the catalytic triad[23-24, 41]. The crystal structure of the group II intron demonstrated that the triple helix provides a platform for metal ion binding with a 3.9 Å distance between two metal ions, which agrees very well with the two-metal ion model (Figure 3.2)[24, 41,44]. Analogous regions in U6 has been shown to bind Mg^{2+} ions involved in structure and catalysis[40,30,45]. U80 binds a Mg^{2+} ion essential for first step of splicing reaction[30]. Several positions in the ACAGAGA loop and AGC triad may have a role in metal ion binding[40]. The hypothetic triple helix in U2/U6 that brings these regions together may also be able to position two metal ions in the active core of the spliceosome for the catalysis in a similar way as in the group II intron.

Intriguingly, mutational studies revealed that A59 plays a more important role in the formation of helix IB than the other two nucleotides in the AGC triad[11]. According to our data, A59 plays a less essential role in stabilizing the base triple interaction than the other two. It was proposed there are multiple active sites for different steps of splicing reaction but they may overlap with each other to some degree[11]. Our data agree with this interesting idea that the A59 is involved more in the second step active structure while G60C61 are more important for the one of first step.

Our data support a model that extended U6 ISL with the AGC base pair with GUU of U6 that participates in the triple helix formation[17]. If the extended U6 ISL is disrupted, the high FRET state is also decreased. In Chapter 2, we demonstrated that under low $MgCl_2$ condition (5mM), most of U2/U6 complexes prefer the high FRET four-way junction structure with an extended U6 ISL, and with high concentration of $MgCl_2$ (40mM-100mM), the intermediate state with a medium FRET is stabilized drastically and enable the formation of low FRET three-way helix structure with helix IB[34] (Figure 3.10). This is surprising and people expected the opposite. The base triple model proposed in this report offers an explanation for these data (Figure 3.10). According to a NMR study of U6 ISL, with low concentration of $MgCl_2$, U80 prefer to flip out of the helix of U6 ISL[46], which enables the base triple formation between U80 and C61-G86. This is also observed for DV bulge in the crystal structure of the group II intron[41]. Higher concentration of $MgCl_2$ stabilizes the conformation with U80 stacked in U6 ISL[46] and disrupts the base triples (Figure 3.10). This conformation is highly possible to be the intermediate state in our folding pathway of U2/U6[34]. This explanation is

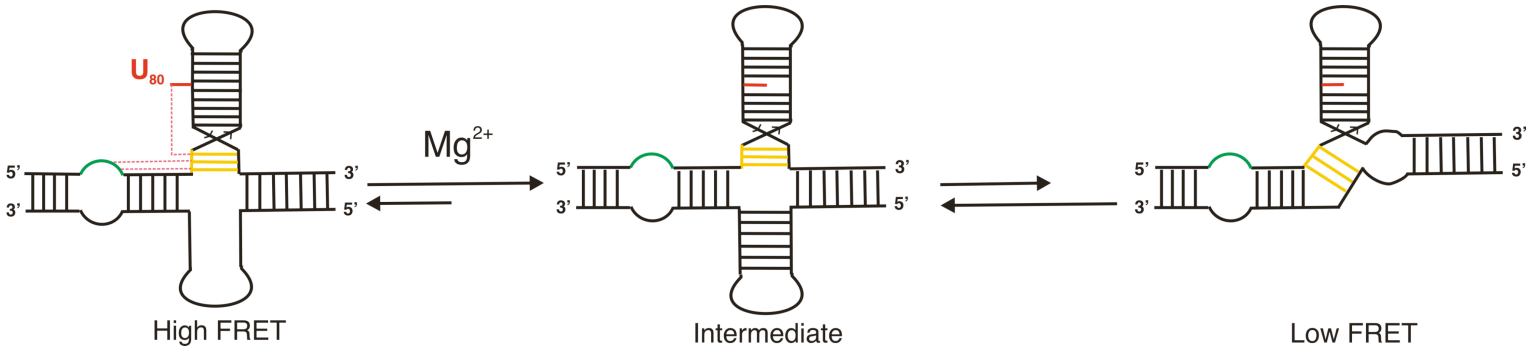


Figure 3.10 Revised folding pathway of U2/U6. With low concentration of $MgCl_2$, U80 prefer to flip out of the helix of U6 ISL, which enables the base triple formation between U80 and C61-G86. Higher concentration of $MgCl_2$ stabilizes the conformation with U80 stacked in U6 ISL and disrupts the base triples. This conformation is highly possible to be the intermediate state. Mg^{2+} ions stabilize the intermediate state that enable the formation of low FRET state with helix IB. This model explains why Mg^{2+} ions stabilize the intermediate state reported in Chapter 2.

consistent with the fact that Mg^{2+} ions inhibit the rescue of first step of splicing by Mn^{2+} in a thiophosphate substitution study of U80[30]. This may be because with high concentration of Mg^{2+}, U80 is stacked in the helix and not available for base triples that important for metal ion coordination for first step of splicing. In addition, according to the folding nature of the four-way junction structure, it may be able to adopt a conformation in which the ACAGAGA loop and the U80 are in close proximity (Figure 3.1D)[27,17,47-48]. However, because the triple interactions occur mainly between the AGC triad and a third strand (U80, G52 and A53) and the base-paring partner of AGC triad (G86-U88 in U6 or G21-U23 in U2) does not directly participate in the interactions. So we cannot totally rule out the possibility that the AGC forms a triplex in the form of helix IB. It is possible the base triples form during both steps of splicing and play pivotal structural and functional role in catalysis, similar to proposed role of base triples of group II intron[19,49]. Interestingly, both in the group II intron and spliceosomal snRNA U2/U6, the catalytic triad AGC is highly sensitive to mutations, while their base pairing partners only lead to minor defects in splicing[39,50].

It is highly possible that spliceosomal proteins are able to stabilize the U6 ISL and base triples *in vivo*. This may be able to explain why some mutations, like U88C, that disrupt the U6 ISL are not lethal *in vivo*[35]. For example, prp8 has been shown to interact extensively with snRNAs in the active core and is required for catalysis. This leads to the hypothesis that both RNA and protein components of spliceosome are involved directly in catalysis[51]. Proteins that bind U6 and are essential for the first step of splicing, for example Cwc 25, may stabilize the triple helix *in vivo*[52].

Part of Chapter 3 is adopted from paper: Guo, Z.; Karunatilaka, K. S.; Rueda, D., Single-molecule analysis of protein-free U2-U6 snRNAs. Nat Struct Mol Biol 2009, 16 (11), 1154-9.

3.5 References

1. Moore, M. J., Query, C.C., and Sharp, P.A., Splicing of precursors to mRNA by the spliceosome. *In TheRNA World, R. Gesteland and J. Atkins, eds. (New York: Cold Spring Harbor Laboratory Press)* **1993**, pp. 303–357.

2. Staley, J. P.; Guthrie, C., Mechanical devices of the spliceosome: motors, clocks, springs, and things. *Cell* **1998,** *92* (3), 315-26.

3. Nilsen, T. W., RNA-RNA Interactions in Nuclear Pre-mRNA Splicing. *In RNA Structure and Function, R. Simons and M. Grunberg-Manago, eds. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press),* **1998**, pp. 279–307.

4. Burge, C. B.; Tuschl, T. A.; Sharp, P. A., Splicing of precursors to mRNAs by the spliceosomes. *in The RNA World 2nd edition, R. F. Gesteland, T. R. Cech, J. F. Atkins (eds.), Cold Spring Harbor Laborotory Press, Cold Spring Harbor, NY, 1999, 525-560.* **1999**.

5. Wahl, M. C.; Will, C. L.; Luhrmann, R., The spliceosome: design principles of a dynamic RNP machine. *Cell* **2009,** *136* (4), 701-18.

6. Brow, D. A., Allosteric cascade of spliceosome activation. *Annu Rev Genet* **2002,** *36*, 333-60.

7. Pyle, A. M.; Lambowitz, A. M., Group II Introns: Ribozymes That Splice RNA and Invade DNA. *in The RNA World, 3rd edition (ed. R.F. Gesteland et al.), pp. 469-505. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.* **2006**.

8. Valadkhan, S., The spliceosome: a ribozyme at heart? *Biol Chem* **2007,** *388* (7), 693-7.

9. O'Keefe, R. T.; Norman, C.; Newman, A. J., The invariant U5 snRNA loop 1 sequence is dispensable for the first catalytic step of pre-mRNA splicing in yeast. *Cell* **1996,** *86* (4), 679-89.

10. Segault, V.; Will, C. L.; Polycarpou-Schwarz, M.; Mattaj, I. W.; Branlant, C.; Luhrmann, R., Conserved loop I of U5 small nuclear RNA is dispensable for both catalytic steps of pre-mRNA splicing in HeLa nuclear extracts. *Mol Cell Biol* **1999,** *19* (4), 2782-90.

11. Madhani, H. D.; Guthrie, C., A novel base-pairing interaction between U2 and U6 snRNAs suggests a mechanism for the catalytic activation of the spliceosome. *Cell* **1992,** *71* (5), 803-17.

12. Valadkhan, S.; Manley, J. L., Splicing-related catalysis by protein-free snRNAs. *Nature* **2001,** *413* (6857), 701-7.

13. Valadkhan, S.; Mohammadi, A.; Jaladat, Y.; Geisler, S., Protein-free small nuclear RNAs catalyze a two-step splicing reaction. *Proc Natl Acad Sci U S A* **2009,** *106* (29), 11901-6.

14. Madhani, H. D.; Bordonne, R.; Guthrie, C., Multiple roles for U6 snRNA in the splicing pathway. *Genes Dev* **1990,** *4* (12B), 2264-77.

15. Ryan, D. E.; Abelson, J., The conserved central domain of yeast U6 snRNA: importance of U2-U6 helix Ia in spliceosome assembly. *RNA* **2002,** *8* (8), 997-1010.

16. Collins, C. A.; Guthrie, C., The question remains: is the spliceosome a ribozyme? *Nat Struct Biol* **2000,** *7* (10), 850-4.

17. Sashital, D. G.; Cornilescu, G.; McManus, C. J.; Brow, D. A.; Butcher, S. E., U2-U6 RNA folding reveals a group II intron-like domain and a four-helix junction. *Nat Struct Mol Biol* **2004,** *11* (12), 1237-42.

18. Sun, J. S.; Manley, J. L., A novel U2-U6 snRNA structure is necessary for mammalian mRNA splicing. *Genes Dev* **1995,** *9* (7), 843-54.

19. Keating, K. S.; Toor, N.; Perlman, P. S.; Pyle, A. M., A structural analysis of the group II intron active site and implications for the spliceosome. *RNA* **2010,** *16* (1), 1-9.

20. Michel, F.; Umesono, K.; Ozeki, H., Comparative and functional anatomy of group II catalytic introns--a review. *Gene* **1989,** *82* (1), 5-30.

21. Lehmann, K.; Schmidt, U., Group II introns: structure and catalytic versatility of large natural ribozymes. *Crit Rev Biochem Mol Biol* **2003,** *38* (3), 249-303.

22. Lambowitz, A. M.; Zimmerly, S., Mobile group II introns. *Annu Rev Genet* **2004,** *38*, 1-35.

23. Gordon, P. M.; Piccirilli, J. A., Metal ion coordination by the AGC triad in domain 5 contributes to group II intron catalysis. *Nat Struct Biol* **2001,** *8* (10), 893-8.

24. Toor, N.; Keating, K. S.; Fedorova, O.; Rajashankar, K.; Wang, J.; Pyle, A. M., Tertiary architecture of the Oceanobacillus iheyensis group II intron. *RNA* **2010,** *16* (1), 57-69.

25. Michel, F.; Costa, M.; Westhof, E., The ribozyme core of group II introns: a structure in want of partners. *Trends Biochem Sci* **2009,** *34* (4), 189-99.

26. Dayie, K. T.; Padgett, R. A., A glimpse into the active site of a group II intron and maybe the spliceosome, too. *RNA* **2008,** *14* (9), 1697-703.

27. Butcher, S. E.; Brow, D. A., Towards understanding the catalytic core structure of the spliceosome. *Biochem Soc Trans* **2005,** *33* (Pt 3), 447-9.

28. Rhode, B. M.; Hartmuth, K.; Westhof, E.; Luhrmann, R., Proximity of conserved U6 and U2 snRNA elements to the 5' splice site region in activated spliceosomes. *EMBO J* **2006,** *25* (11), 2475-86.

29. Sarver, M.; Zirbel, C. L.; Stombaugh, J.; Mokdad, A.; Leontis, N. B., FR3D: finding local and composite recurrent structural motifs in RNA 3D structures. *J Math Biol* **2008,** *56* (1-2), 215-52.

30. Yean, S. L.; Wuenschell, G.; Termini, J.; Lin, R. J., Metal-ion coordination by U6 small nuclear RNA contributes to catalysis in the spliceosome. *Nature* **2000,** *408* (6814), 881-4.

31. Zhao, R.; Rueda, D., RNA folding dynamics by single-molecule fluorescence resonance energy transfer. *Methods* **2009,** *49* (2), 112-7.

32. Bokinsky, G.; Zhuang, X., Single-molecule RNA folding. *Acc Chem Res* **2005,** *38* (7), 566-73.

33. Cornish, P. V.; Ha, T., A survey of single-molecule techniques in chemical biology. *ACS Chem Biol* **2007,** *2* (1), 53-61.

34. Guo, Z.; Karunatilaka, K. S.; Rueda, D., Single-molecule analysis of protein-free U2-U6 snRNAs. *Nat Struct Mol Biol* **2009,** *16* (11), 1154-9.

35. Hilliker, A. K.; Staley, J. P., Multiple functions for the invariant AGC triad of U6 snRNA. *RNA* **2004,** *10* (6), 921-8.

36. Sashital, D. G.; Allmann, A. M.; Van Doren, S. R.; Butcher, S. E., Structural basis for a lethal mutation in U6 RNA. *Biochemistry* **2003,** *42* (6), 1470-7.

37. Reiter, N. J.; Nikstad, L. J.; Allmann, A. M.; Johnson, R. J.; Butcher, S. E., Structure of the U6 RNA intramolecular stem-loop harboring an S(P)-phosphorothioate modification. *RNA* **2003,** *9* (5), 533-42.

38. Fabrizio, P.; Abelson, J., Two domains of yeast U6 small nuclear RNA required for both steps of nuclear precursor messenger RNA splicing. *Science* **1990,** *250* (4979), 404-9.

39. Boulanger, S. C.; Belcher, S. M.; Schmidt, U.; Dib-Hajj, S. D.; Schmidt, T.; Perlman, P. S., Studies of point mutants define three essential paired nucleotides in the domain 5 substructure of a group II intron. *Mol Cell Biol* **1995,** *15* (8), 4479-88.

40. Fabrizio, P.; Abelson, J., Thiophosphates in yeast U6 snRNA specifically affect pre-mRNA splicing in vitro. *Nucleic Acids Res* **1992,** *20* (14), 3659-64.

41. Toor, N.; Keating, K. S.; Taylor, S. D.; Pyle, A. M., Crystal structure of a self-spliced group II intron. *Science* **2008,** *320* (5872), 77-82.

42. Toor, N.; Rajashankar, K.; Keating, K. S.; Pyle, A. M., Structural basis for exon recognition by a group II intron. *Nat Struct Mol Biol* **2008,** *15* (11), 1221-2.

43. Cech, T. R.; Bass, B. L., Biological catalysis by RNA. *Annu Rev Biochem* **1986,** *55*, 599-629.

44. Steitz, T. A.; Steitz, J. A., A general two-metal-ion mechanism for catalytic RNA. *Proc Natl Acad Sci U S A* **1993,** *90* (14), 6498-502.

45. Yu, Y. T.; Maroney, P. A.; Darzynkiwicz, E.; Nilsen, T. W., U6 snRNA function in nuclear pre-mRNA splicing: a phosphorothioate interference analysis of the U6 phosphate backbone. *RNA* **1995,** *1* (1), 46-54.

46. Huppler, A.; Nikstad, L. J.; Allmann, A. M.; Brow, D. A.; Butcher, S. E., Metal binding and base ionization in the U6 RNA intramolecular stem-loop structure. *Nat Struct Biol* **2002,** *9* (6), 431-5.

47. Valadkhan, S.; Manley, J. L., Intrinsic metal binding by a spliceosomal RNA. *Nat Struct Biol* **2002,** *9* (7), 498-9.

48. Murchie, A. I.; Thomson, J. B.; Walter, F.; Lilley, D. M., Folding of the hairpin ribozyme in its natural conformation achieves close physical proximity of the loops. *Mol Cell* **1998,** *1* (6), 873-81.

49. de Lencastre, A.; Hamill, S.; Pyle, A. M., A single active-site region for a group II intron. *Nat Struct Mol Biol* **2005,** *12* (7), 626-7.

50. Peebles, C. L.; Zhang, M.; Perlman, P. S.; Franzen, J. S., Catalytically critical nucleotide in domain 5 of a group II intron. *Proc Natl Acad Sci U S A* **1995,** *92* (10), 4422-6.

51. Abelson, J., Is the spliceosome a ribonucleoprotein enzyme? *Nat Struct Mol Biol* **2008,** *15* (12), 1235-7.

52. Warkocki, Z.; Odenwalder, P.; Schmitzova, J.; Platzmann, F.; Stark, H.; Urlaub, H.; Ficner, R.; Fabrizio, P.; Luhrmann, R., Reconstitution of both steps of Saccharomyces cerevisiae splicing with purified spliceosomal components. *Nat Struct Mol Biol* **2009,** *16* (12), 1237-43.

CHAPTER 4

Single Molecule Study of binding of spliceosomal protein Prp24 to snRNA U2/U6

4.1 Introduction

In eukaryotes, pre-mRNA contains both protein encoding sections, exons, and intervening sequences, introns. For proper protein translation, introns must be precisely removed and the exons need to be ligated together for translation[1,2]. This process is called splicing[3]. Nuclear pre-mRNA splicing is catalyzed by the spliceosome, a huge and dynamic complex containing five small nuclear RNAs (snRNAs U1, U2, U4, U5 and U6) and about 170 proteins in humans[2,4]. Accumulating evidence indicates that the spliceosome is at least in part a ribozyme and snRNAs U2 and U6 are responsible for catalysis of splicing[5,6,7]. Conserved regions in U6 (U80, AGC triad and ACAGAGA loop) have been proposed to participate in catalysis directly (Figure 4.1A)[8,9].However, recent structural studies on Prp8 lead to the hypothesis that in the catalytic heart of the spliceosome, both RNA and protein components are involved in catalysis[10,11,12,].

During each round of the splicing reaction, the spliceosome is built anew on the pre-mRNA substrate following a highly ordered and stepwise pathway in which the structure of the spliceosome is in constant changing[2]. During the assembly and catalysis of the spliceosome, the snRNAs undergo dramatic structural rearrangements[13,14,15] (Figure 1.5C). The rearrangements of the snRNAs require the assistance of DExD/H-type

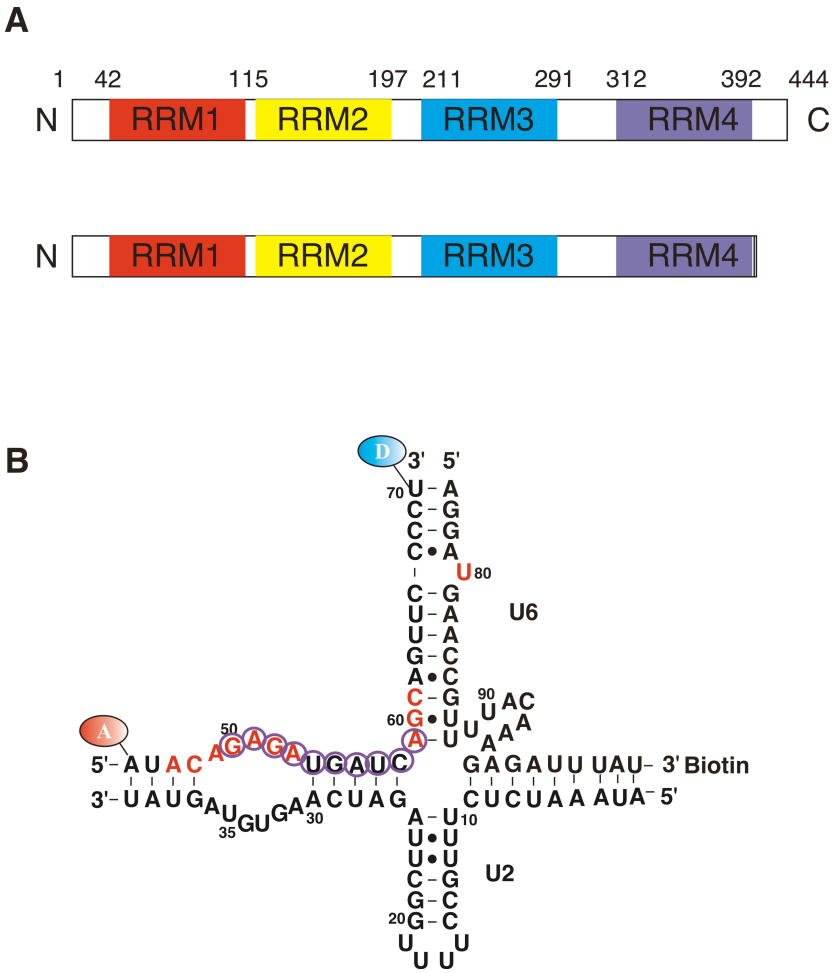


Figure 4.1 Prp24 and its binding site to U2/U6 (A) Domain structure of Prp24 from S. cerevisiae. Prp24 contains 4 RNA recognition motifs (RRM)[16]. The four RRMs are shown in different colors. (**B**) Labeled U2/U6 construct. The 5' of the U6-1 is labeled by a FRET acceptor Cy5 and the 3' is labeled by FRET donor Cy3. The 3' of U6-2 is labeled by biotin for surface immobilization. Highly conserved regions are highlighted in red. Prp24 binding sites are in purple circles[17,14].

RNA-dependent ATPases/helicases and RNA chaperones that do not require ATP. They help to unwind RNA helices, anneal duplexes and stabilize the active form of the RNA complexes[18,6].

Prp24 has been proposed to be the RNA chaperone of U6[19,20]. The spliceosomal protein Prp24 is an essential component of the U6 snRNP in yeast and helps U6 to remodel during assembly and catalysis of the spliceosome[20,16]. The details about the function and mechanism of Prp24 are still unclear. However, it is well established that recombinant Prp24 promotes the annealing of the U4/U6 snRNA complex yielding U4/U6 bi-snRNP *in vitro*[19]. However, recombinant Prp24 acts more efficiently in cell extract[21,19]. It may be that another member of the U6 snRNP, the Lsm protein complex, also helps the association of U4/U6 and works together with Prp24 to increase efficiency[22,23,24]. After the formation of the U4/U6 complex, Prp24 leaves the complex, but the Lsm proteins remain bound[25,26]. It has been proposed that Prp24 may return during the activation process of the spliceosome to help unwind the U4/U6 complex[21,27]. Prp24 does not hydrolyze ATP and may be an RNA chaperone[19].

Prp24 form *Saccharomyces cerevisiae* has four RNA recognition motifs (RRMs)[16] (Figure 4.1A). Typically, RRMs preferentially bind single strand RNA[17]. Although a high-resolution structure of the full-length Prp24 is not yet available, the first three RRMs of Prp24 have been crystallized[16]. Currently, the functions of each RRM are not well understood. Genetic studies indicate that RRMs 2 and 3 may stabilize the free U6 snRNA[27]. Gel shift binding assays demonstrated that RRMs 1 and 2 are important for high-affinity binding of U6 snRNA and RRMs 3 and 4 may have other functions like controlling the stoichiometry of the Prp24 binding[20]. Biochemical studies with purified

U6 snRNP and binding assays with recombined Prp24 and truncated U6 indicat that the binding site of Prp24 on U6 may be located near the conserved ACAGAGA loop (Nucleotides 40-58 and Figure 4.1B) [28,20, 25]. A more recent NMR study revealed that RRM2 binds the GAGA box in the conserved ACAGAGA loop of U6 in a sequence-specific fashion and RRM1 interacts with the phosphate backbone of 3' downstream GAUCA sequence via electrostatic interactions, leaving the bases available for base pairing (Figure 3.1B). This can disrupt the internal helix of U6 and expose the bases for base paring[17], which is a very appealing model explaining the mechanism by which Prp24 promotes the annealing of the U4/U6 complex.

Pervious structural studies about Prp24 were conducted with U6 snRNA or snRNP only. The genetic studies only considered the effect of Prp24 on the U4/U6 complex. The question remains: does Prp24 change the structural dynamics of the U2/U6 complex that is considered the active heart of the spliceosome? To address this question, we use single molecule FRET to characterize the effect of Prp24 on the structure and dynamics of U2/U6 complex. We observed that Prp24 can induce conformational changes of U2/U6 and stabilize a low FRET structure.

4.2 Materials and methods

4.2.1 Sample Preparation and Purification.

There are 3 strands of RNA forming the U2-U6 complex in our construct (Figure 1.7A): U6-1, U6-2 and U2. All the RNA samples were purchased from Keck Foundation Biotechnology Resource Laboratory at Yale University School of Medicine (New Haven, CT). RNA purification and labeling was preformed as described in Chapter 2.

The protein Prp 24 was obtained form the Butcher Lab in university of Wisconsin-Madison. Protein samples came in the storage buffer containing 50% glycerol, 25mM Tris, 125 mM NaCl, 1.25mM DTT and 0.2 mM EDTA. Protein samples were dialyzed before use with dialysis buffer containing 50mM Tris, 100mM KCl and 2mM DTT.

4.2.2 Steady-State FRET

Steady-state FRET measurements of the binding of Prp24 to double-labeled U2-U6 complex were carried out in a spectrofluorometer (Cary Eclipse, Varian Inc., Palo Alto, CA). Cy3 was excited at 550nm (10 nm bandwidth) and Cy3 and Cy5 emission intensities (I_{Cy3} and I_{Cy5}) were measured at 565 nm and 665 nm (5 nm bandwidth), respectively. Relative FRET efficiencies were calculated as FRET = $I_{Cy5}/(I_{Cy3} + I_{Cy5})$. A 130 μL solution with 25 nM U6-1, 50 nM U6-2 and 100 nM U2 in standard buffer (50 mM Tris-HCl, pH 7.5, 100mM NaCl and 25mM DTT) was first denatured by heating at 90 ˚C for 2 min and then slowly annealed at room temperature for 20 minutes. After this RNA-only solution was scanned, various amount of protein were manually added to the RNA solution to reach final protein concentrations. After equilibration, the final FRET value was measured. The protein dissociation constant (K_D) and cooperativity coefficients (n) were obtained by plotting FRET as a function of protein concentration and fitting to the modified Hill equation:

$$FRET = FRET_0 - \left(FRET_0 - FRET_\infty\right)\left([\mathrm{Pr}\ p24\ ]^n \Big/ \left(K_M^n + [\mathrm{Pr}\ p24]^n\right)\right)$$

4.2.3 Single-molecule FRET.

The three RNA strands (Fig. 1.7A) were annealed in standard buffer (50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 1% (vol/vol) 2-mercaptoethanol) and variable [Mg^{2+}] from 1-10mM. A 10 μL solution was heated to 90°C for 45 sec before cooling at room temperature over 20 min. The annealed biotinylated and fluorophore-labeled complex was diluted to 25-50 pM and desired concentration of Prp24 was added in the diluted solution. RNA-protein complexes are immunized on PEGlated surface of quartz slides. This method was reviewed in detail in Ref 29. Single molecule measurements were conducted as described in Chapter 2.

4.3 Results and discussion

4.3.1 Prp24 binding to U2/U6 induces conformational change.

We have developed a minimal fluorophore labeled U2/U6 construct that enables us to study the U2/U6 structural dynamics with FRET and single molecule spectroscopy[14]. We used this construct to characterize the Prp24 and U2/U6 interactions in this study. We first characterized the Prp24 and U2/U6 interactions using bulk FRET with various concentration of $MgCl_2$ (Figure 4.2). With a fixed $MgCl_2$ concentration, C-terminal truncated Prp24 (Prp24 N1234) was added to the RNA solution. The bulk titrations were repeated with different concentrations of Mg^{2+}, which are 0.1mM, 3mM, 7mM and 20mM, respectively. For all the Mg^{2+} concentrations, with the increase of the Prp24 N1234 concentration, the FRET decreased (Figure 4.2A), indicating that the Prp24 N1234 induces a conformational change of U2/U6 that separates ACAGAGA loop and the U6 intramolecular stem loop (U6 ISL) apart form each other. The apparent

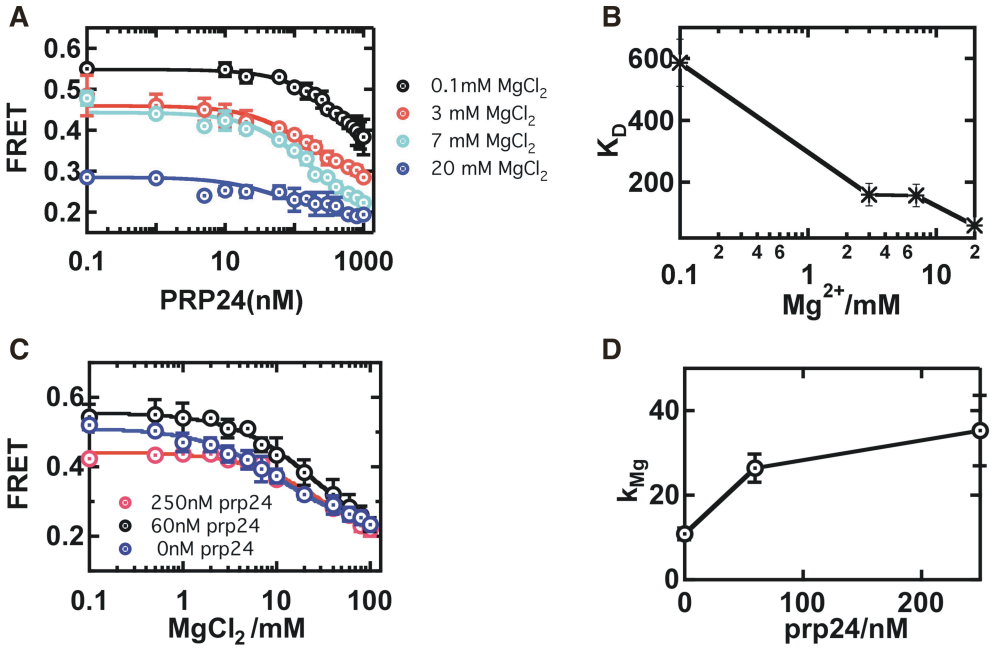


Figure 4.2 Bulk titration of Prp24 (A,B) Bulk fluorescence measurements of the labeled U2U6 complex in the dependence of prp24 concentration. **(A)** The bulk titrations are conducted with different concentrations of Mg^{2+}, which are 0.1mM (in black), 3mM (in red), 7mM(in cyan) and 20mM (in blue), respectively. For all the Mg^{2+} concentrations, with the increase of the prp24 concentration, the FRET decreased. **(B)** K_D obtained from the titrations decreased with the increase of the Mg^{2+} concentration. **(C,D)** Bulk fluorescence measurements of the labeled U2U6 complex in the dependence of Mg^{2+} concentration. (**C**) The bulk titrations are conducted with different concentrations of prp24, which are 0mM (in blue), 60mM (in black), and 250mM (in red), respectively. For all the prp24 concentrations, with the increase of the Mg^{2+} concentration, the FRET decreased. (**D**) The K_{Mg} increased while the prp24 concentration was increasing.

dissociation constants (K_D) of Prp24 N1234 with different $MgCl_2$ concentration were obtained from the titrations that are K_D (0.1) = 585.7 ± 75.8 nM, K_D (3.0) = 159.2 ± 36.2 nM, K_D (7.0) = 156.5 ± 35.9 nM and K_D (20) = 59.8 ± 17.3 nM, respectively. K_D decreased with the increase of the Mg^{2+} concentration (Figure 4.2B), suggesting that Mg^{2+} ions help the binding of Prp24 N1234 to U2/U6. We also titrated change of FRET efficiency as a function of $MgCl_2$ concentration with a fixed concentration of Prp24 N1234. The bulk titrations are conducted with 0mM, 60mM, and 250mM of prp24 N1234. For all the Prp24 N1234 concentrations, when the concentration of the Mg^{2+} increased, the FRET ratio decreased. K_{Mg} was obtained from titrations (Figure 4.2C), which are K_{Mg} (0.0) = 10.7 ± 1.3 mM, K_{Mg} (60) = 26.3 ± 3.3 mM, and K_{Mg} (250) = 35.3 ± 8.3 mM, respectively. The K_{Mg} increased with increasing concentration of Prp24 N1234 (Figure 4.2D), indicating that the binding of Prp24 N1234 may compete with the Mg^{2+} binding. This is consistent with the fact that binding site of Prp24 N1234 overlap with the proposed Mg^{2+} binding site in U6 (ACAGAGA loop and AGC triad)[30,31,32].

In conclusion, Prp24 N1234 can induce a conformational change of U2/U6 in our bulk FRET experiments and the binding of Prp24 N1234 is Mg^{2+} dependent. Surprisingly, interaction between Prp24 N1234 and U2/U6 may inhibit the binding of Mg^{2+}, indicating that Prp24 N1234 may bind and bury one or more the Mg^{2+} binding sites in U6.

4.3.2 Single molecule FRET reveals Prp24 stabilize a 0.2 FRET state

Form bulk FRET data, we can only obtain information about the averaged behavior of many molecules, while single molecule FRET (smFRET) provides detailed dynamic information about the conformational changes of biological molecules, capture

transient intermediates which are difficult to detect in bulk solution experiments, and record conformational fluctuations at a time scale that is relevant for biological activity[33,34]. Characteristic single molecule trajectories of single U2/U6 with or without Prp24 N1234 that reveal the effect of Prp24 N1234 on the conformation of U2U6 complex, are shown in Figure 4.3 A,B. Without Prp24 N1234, in 10mM $MgCl_2$ U2/U6 is dynamic, jumping among three FRET states (0.6, 0.4 and 0.2) randomly. After adding Prp24 N1234, it becomes static in the 0.2 state stably. Furthermore, we built smFRET histograms with more than 100 single molecule trajectories in the presence or absence of Prp24 N1234. According to the smFRET histograms, without Prp24 N1234, the majority of molecules stay in the intermediate state (FRET=0.4) and low FRET state (FRET =0.2). In contrast, with 30 μM Prp24 N1234, most of molecules stay in a low FRET state (FRET=0.2) (Figure 4.3 A,B).

We also performed a flow experiment in which with 3mM $MgCl_2$ U2/U6 complexes were immobilized on the surface of the quartz slide and Prp24 N1234 solution was injected while monitoring and recording with a CCD camera. This enabled monitoring of prp24 N1234-induced conformational changes in the real time. Before the addition of Prp24 N1234, the molecule is dynamic and stays mostly at high FRET state. After the addition of Prp24 N1234 (33 s), the molecule soon changed to the 0.2 state and became static, staying in the 0.2 state for the rest of time before photobleaching.

In summary, our single molecule data reveal that Prp24 N1234 stabilizes a low FRET (0.2) state and restrict the dynamics of the U2/U6.

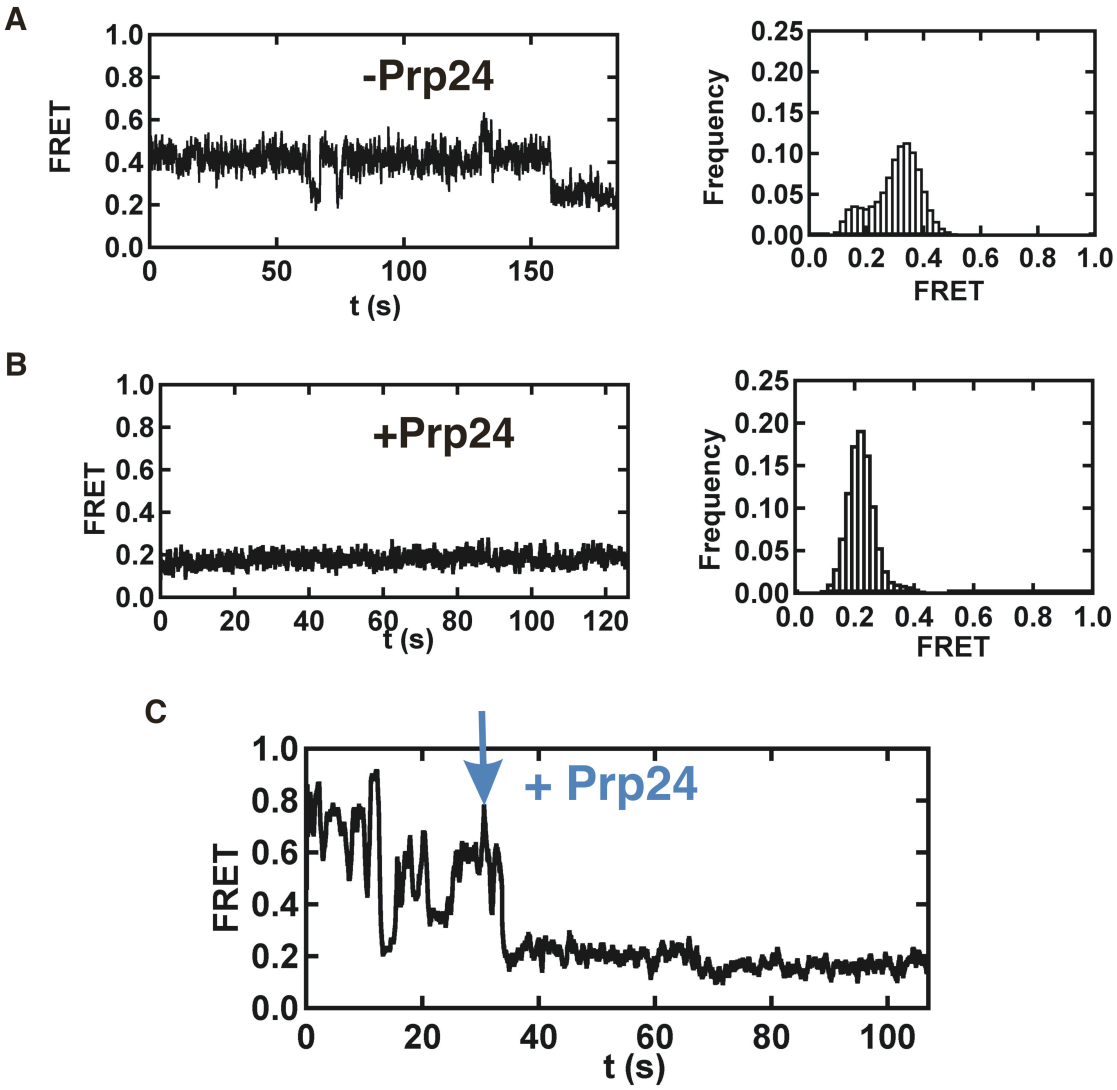


Figure 4.3 Effect of prp24 on the conformation of U2U6 complex (A,B) Single molecule traces reveal the effect of prp24 on the conformation of U2U6 complex and single molecule histograms indicate that the prp24 can induce significant conformation changes of U2U6 complex. (**A**) Without prp24 and with 10mM $MgCl_2$ the molecule stay at the 0.4 state mostly and change to 0.2 and 0.6 states sometimes. Single molecule histogram has both 0.2 and 0.4 peaks. (**B**) After adding of prp24, it stays at the 0.2 state stably. Single molecule histogram has only 0.2 peak. (**C**) The monitor of prp24 induced conformational change in the real time. This experiment is carried out with 3mM $MgCl_2$. Before the adding of prp24, the molecule stays at high FRET state but very unstable. After the adding of prp24 at the 33rd second, the molecule changes to the 0.2 state shortly and stay there for the rest of time before photobleaching.

4.3.3 Prp24 N1234 affects the structural dynamics of the U2/U6 complex

To access the influence of spliceosomal protein Prp24 N1234 on U2/U6 folding dynamics, smFRET distributions were plotted with or without Prp24 N1234 between 1mM and 10mM Mg^{2+}. Each histogram is constructed with about 100 single molecule trajectories (Figure 4.4). With 1mM Mg^{2+} and no Prp24 N1234, most molecules stay in the high FRET state. However, after adding 30 nM Prp24 N1234, most of them adopt a 0.2 FRET structure. At 3 and 5mM Mg^{2+} without Prp24 N1234, the majority of molecules stay in the high FRET state and only a small fraction of them stay in the intermediate state as described in the chapter 2[14]. Addition of Prp24 N1234 promotes most of U2/U6 to stay in a 0.2 FRET state. With 10mM $MgCl_2$, after adding Prp24, FRET change from two major peaks at 0.4 and 0.2 to an only peak at 0.2. Single molecule data indicate that the Prp24 can induce significant conformation changes of U2U6 complex (Figure 4.4).

We also conducted single molecule experiments with various concentration of full length Prp24 in 10mM $MgCl_2$. With 10mM $MgCl_2$ and 45 nM full length Prp24, most of molecules stay in a 0.2 state that is very similar to that with Prp24 N1234 (Figure 4.5A), consistent with the data from gel shift binding experiments showing that there is no significant difference in binding affinity with or without C-terminus of Prp24[20]. If the Prp24 concentration is decreased to 4.5 nM, a small fraction of molecules stay in a 0.35 state, while majority of them adopting the 0.2 state (Figure 4.5B). With 45 pM Prp24, there are two major peaks in the corresponding histogram: a broad peak at 0.35 and a broad 0.2 peak (Figure 4.5C), indicating that full length Prp24 induces conformational

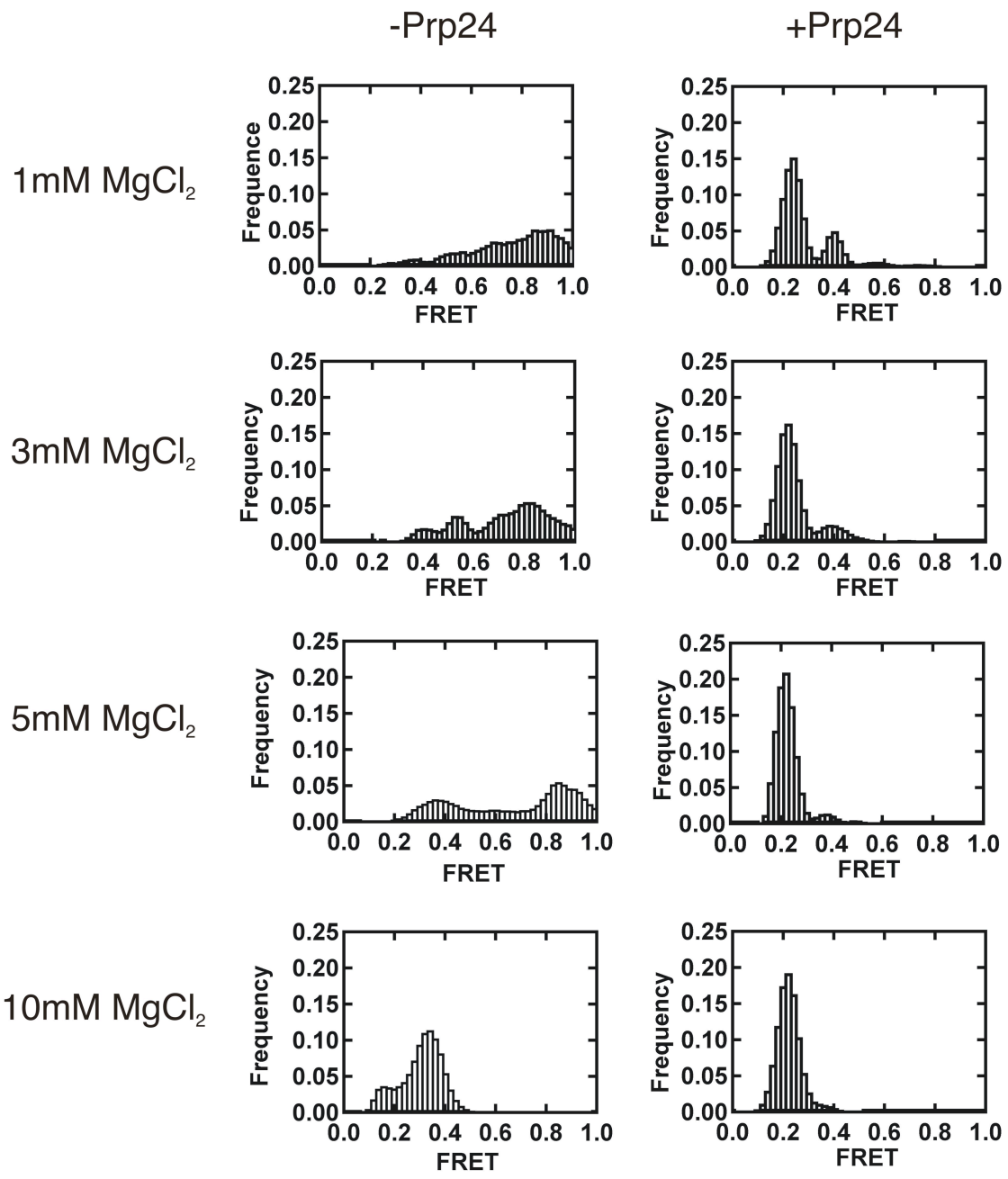


Figure 4.4 Single molecule histograms indicate that the prp24 N1234 can induce significant conformation changes of U2U6 complex. With 1mM $MgCl_2$, after the adding of prp24 the FRTE change from a major peak around 0.8-1.0 to a big peak at 0.2 and a smaller peak at 0.4. With 3mM $MgCl_2$, after adding of prp24 the FRTE change from two major peaks at 0.8 and 0.6 to a big peak at 0.2 and a very small peak at 0.4. With 5mM $MgCl_2$, after adding of prp24 the FRTE change from two major peaks at 0.8 and 0.4 to a big peak at 0.2 and a very tiny peak at 0.4. With 10mM $MgCl_2$, after adding of prp24 the FRTE change from two major peaks at 0.4 and 0.2 to an only peak at 0.2.

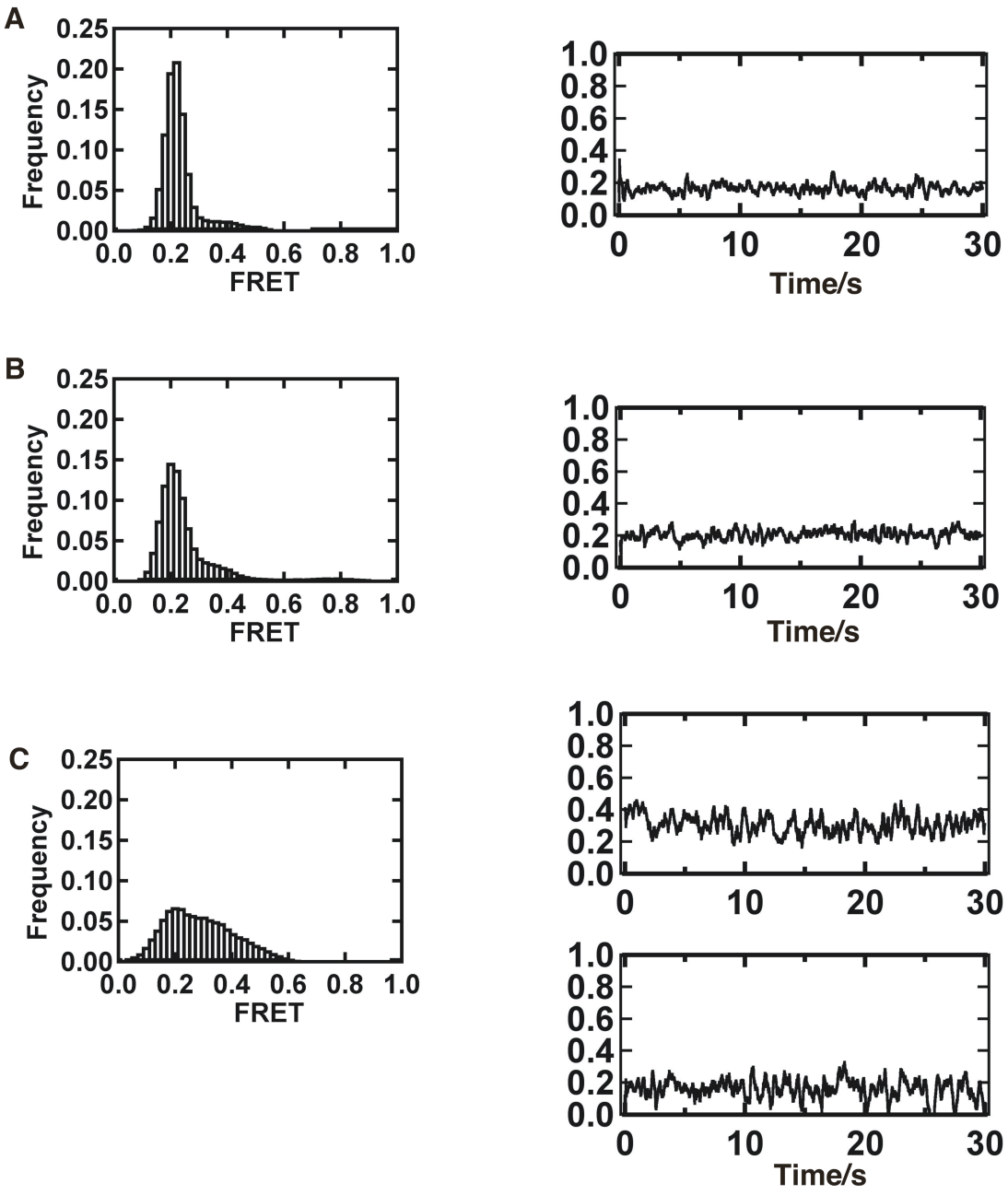


Figure 4.5 Single molecule histograms indicate that the full-length prp24 can induce the same conformation changes of U2U6 complex as Prp24 N1234. (A) With 10mM $MgCl_2$ and 45nm full length Prp24, most of molecules stay in a 0.2 state that is very similar to with Prp24 N1234 (**B**) With 4.5 nM and 10 mM $MgCl_2$, a small fraction of molecules stay in a 0.35 state, while majority of then in 0.2 state. (**C**) With 45 pM Prp24 and 10 mM $MgCl_2$, there are two major peaks in the corresponding histogram: a broad peak in 0.35 and a broad 0.2 peak.

change of U2/U6 to a 0.2 state, but with low concentration of Prp24, the intermediate and low FRET states become less defined or in fast dynamics.

In conclusion, Prp24 interaction with U2/U6 is dependent on Mg^{2+} and Prp24 concentrations.

4.4 Conclusions

There are a large number of proteins in the spliceosome that play vital structural and catalytic roles[18,4]. RNA Chaperones, like Prp24, help the structural rearrangements of spliceosomal snRNAs. Here we report preliminary single molecule FRET data showing that spliceosomal protein Prp24 can induce a conformational change in U2/U6 complex. This RNA:protein interaction is Mg^{2+} and protein concentration dependent.

Mg^{2+} ions help the binding of Prp24 to U2/U6, but in contrast, Prp24 inhibits the binding Mg^{2+}. The Mg^{2+} binding sites have been proposed in the highly conserved regions, the ACAGAGA loop, AGC triad and U80[32,35,31,30], overlapping with binding sites of Prp24[17,28]. Prp24 may bind these regions and inhibit binding of incoming Mg^{2+} ions (Figure 4.1B).

Prp24 has been shown to be able to anneal U4 and U6 to form a complex[27,19]. A recent NMR study with a segment of U6 and RRM2 of Prp24 revealed possible binding sites of Prp24 in U6[17]. In the resulting model, RRM2 of Prp24 binds the GAGA box in the ACAGAG loop in a sequence-specific fashion and RRM1 interacts with U6 via charge-charge interactions. Binding of Prp24 may be able to unwind part of the internal helix of U6 snRNA and expose bases for annealing with U4[17]. According to that model and our single molecule data, we proposed a mechanism by which Prp24 modulates the

structural dynamics of U2/U6 (Figure 4.6). As described in Chapter 3, with low concentration of Mg^{2+}, there are base triples that stabilize the high FRET state. Prp24 RRM1 and RRM2 bind the nucleotides that participate in the base triples (G52, A53 and AGC triad) and the binding of Prp24 may protect these nucleotides and disrupt the base triples and the high FRET state. At the same time, helix Ia and part of the U6 ISL may be unwound and this partial open conformation of U2/U6 may correspond to the intermediate FRET state observed. The opening of the U6 ISL facilitates the formation of helix IB and Prp24 may stay bound after the formation of three-helix structure with helix IB, stabilizing it form going back to the intermediate state and high FRET state. This may explain why with low concentrations of Prp24, the intermediate state becomes less defined. If the protein dissociates form U2/U6 before annealing, the partial open complex may adopt random conformations. In addition, low FRET state is less populated and also become less defined with low concentration of Prp24, suggesting that the binding of Prp24 stabilizes this conformation.

It has been shown that the first RRMs of Prp24, RRM1 and RRM2, are important for high affinity binding of Prp24 to U6 snRNA and the function of RRM3 and RRM4 is unclear. It is possible that RRM3 and RRM4 are able to bind U2 and stabilize the complex formed between U2 and U6. The reason why Prp24 prefers to stabilize the three-way helix structure of U2/U6 may be that binding of Prp24 disrupts the base-triple interactions that stabilize the four–way helix structure. Furthermore, we showed evidence that destabilizing of the base triples may favor the formation of intermediate state and enable the transitions to the three-way helix conformation in Chapter 3.

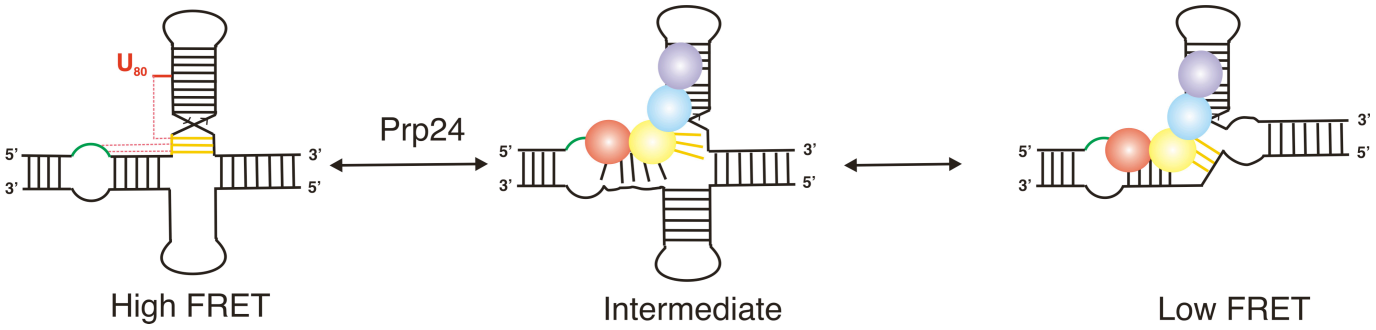


Figure 4.6 Proposed the mechanism how Prp24 modulate the structural dynamics of the U2/U6. As described in Chapter 3, under low concentration of Mg^{2+}, there are base triples that stabilize the high FRET state. Prp24 RRM1 and RRM2 binds the nucleotides that participate in the base triples (G52, A53 and AGC triad) The binding of Prp24 may protect these nucleotides disrupting the base triples and the high FRET state and unwind helix Ia and part of the U6 ISL. This partial open conformation of U2/U6 may be corresponding to the intermediate FRET state observed. The opening of the U6 ISL may facilitate the formation of helix IB and Prp24 may stay bound after the formation of three-helix structure with helix IB, stabilizing it form going back to the intermediate state and high FRET state.

Many functions of Prp24 have been suggested and proposed, including annealing and unwinding of U4/U6[17,19,27], modulating the structural dynamics of the U2/U6. To conduct these multiple functions that are sometimes conflicting with each other, there must be some regulation mechanism for Prp24 function. Lsm proteins interact with Prp24 and enhance the annealing of U4/U6[23,24,36]. In addition, other spliceosomal protein factors, especially regulation factors, may modulate the function of Prp24 at different stages of splicing.

According to genetic studies, Prp24 promotes the annealing of U4/U6[19,27]. Structural studies with full length and truncated U6 and Prp24 revealed the possible binding sites of Prp24 to U6[20,28]. More recently, NMR studies revealed the sequence-specific binding of RRM2 of Prp24 and proposed that Prp24 can open partially the U6 ISL to facilitate the annealing of U4/U6. Because U2/U6 is believed to be the catalytic core of the spliceosome and Prp24 is an essential component of U6 snRNP[16,15], we hypothesize that Prp24 also play an important role in modulating the structural dynamics of U2/U6. Our bulk and single molecule data is consistent with this idea. Prp24 may bind U6 in similar binding sites revealed by pervious studies and stabilize the three-way junction by a similar mechanism proposed according to the NMR study[17] (Figure 4.6).

However, the effect of Prp24 on structural dynamics of U2/U6 may be due to alternative mechanisms. Furthermore, the function of each RRM is unclear. More single molecule and biochemical experiments with wild type and mutant Prp24, which is going on in the lab now, will clarify these questions (See future directions in Chapter 5).

4.5 References

1. Burge, C. B.; Tuschl, T. A.; Sharp, P. A., Splicing of precursors to mRNAs by the spliceosomes. *in The RNA World 2nd edition, R. F. Gesteland, T. R. Cech, J. F. Atkins (eds.), Cold Spring Harbor Laborotory Press, Cold Spring Harbor, NY, 1999, 525-560.* **1999**.

2. Moore, M. J., Query, C.C., and Sharp, P.A., Splicing of precursors to mRNA by the spliceosome. *In TheRNA World, R. Gesteland and J. Atkins, eds. (New York: Cold Spring Harbor Laboratory Press)* **1993**, pp. 303–357.

3. Sharp, P. A., On the origin of RNA splicing and introns. *Cell* **1985,** *42* (2), 397-400.

4. Staley, J. P.; Guthrie, C., Mechanical devices of the spliceosome: motors, clocks, springs, and things. *Cell* **1998,** *92* (3), 315-26.

5. Valadkhan, S., The spliceosome: a ribozyme at heart? *Biol Chem* **2007,** *388* (7), 693-7.

6. Collins, C. A.; Guthrie, C., The question remains: is the spliceosome a ribozyme? *Nat Struct Biol* **2000,** *7* (10), 850-4.

7. Madhani, H. D.; Guthrie, C., Dynamic RNA-RNA interactions in the spliceosome. *Annu Rev Genet* **1994,** *28*, 1-26.

8. Butcher, S. E.; Brow, D. A., Towards understanding the catalytic core structure of the spliceosome. *Biochem Soc Trans* **2005,** *33* (Pt 3), 447-9.

9. Nilsen, T. W., RNA-RNA Interactions in Nuclear Pre-mRNA Splicing. *In RNA Structure and Function, R. Simons and M. Grunberg-Manago, eds. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press),* **1998**, pp. 279–307.

10. Yang, K.; Zhang, L.; Xu, T.; Heroux, A.; Zhao, R., Crystal structure of the beta-finger domain of Prp8 reveals analogy to ribosomal proteins. *Proc Natl Acad Sci U S A* **2008,** *105* (37), 13817-22.

11. Ritchie, D. B.; Schellenberg, M. J.; Gesner, E. M.; Raithatha, S. A.; Stuart, D. T.; Macmillan, A. M., Structural elucidation of a PRP8 core domain from the heart of the spliceosome. *Nat Struct Mol Biol* **2008,** *15* (11), 1199-205.

12. Abelson, J., Is the spliceosome a ribonucleoprotein enzyme? *Nat Struct Mol Biol* **2008,** *15* (12), 1235-7.

13. Rhode, B. M.; Hartmuth, K.; Westhof, E.; Luhrmann, R., Proximity of conserved U6 and U2 snRNA elements to the 5' splice site region in activated spliceosomes. *EMBO J* **2006,** *25* (11), 2475-86.

14. Guo, Z.; Karunatilaka, K. S.; Rueda, D., Single-molecule analysis of protein-free U2-U6 snRNAs. *Nat Struct Mol Biol* **2009,** *16* (11), 1154-9.

15. Madhani, H. D.; Guthrie, C., A novel base-pairing interaction between U2 and U6 snRNAs suggests a mechanism for the catalytic activation of the spliceosome. *Cell* **1992,** *71* (5), 803-17.

16. Bae, E.; Reiter, N. J.; Bingman, C. A.; Kwan, S. S.; Lee, D.; Phillips, G. N., Jr.; Butcher, S. E.; Brow, D. A., Structure and interactions of the first three RNA recognition motifs of splicing factor prp24. *J Mol Biol* **2007,** *367* (5), 1447-58.

17. Martin-Tumasz, S.; Reiter, N. J.; Brow, D. A.; Butcher, S. E., Structure and functional implications of a complex containing a segment of U6 RNA bound by a domain of Prp24. *RNA* **2010,** *16* (4), 792-804.

18. Wahl, M. C.; Will, C. L.; Luhrmann, R., The spliceosome: design principles of a dynamic RNP machine. *Cell* **2009,** *136* (4), 701-18.

19. Raghunathan, P. L.; Guthrie, C., A spliceosomal recycling factor that reanneals U4 and U6 small nuclear ribonucleoprotein particles. *Science* **1998,** *279* (5352), 857-60.

20. Kwan, S. S.; Brow, D. A., The N- and C-terminal RNA recognition motifs of splicing factor Prp24 have distinct functions in U6 RNA binding. *RNA* **2005,** *11* (5), 808-20.

21. Ghetti, A.; Company, M.; Abelson, J., Specificity of Prp24 binding to RNA: a role for Prp24 in the dynamic interaction of U4 and U6 snRNAs. *RNA* **1995,** *1* (2), 132-45.

22. Achsel, T.; Brahms, H.; Kastner, B.; Bachi, A.; Wilm, M.; Luhrmann, R., A doughnut-shaped heteromer of human Sm-like proteins binds to the 3'-end of U6 snRNA, thereby facilitating U4/U6 duplex formation in vitro. *EMBO J* **1999,** *18* (20), 5789-802.

23. Verdone, L.; Galardi, S.; Page, D.; Beggs, J. D., Lsm proteins promote regeneration of pre-mRNA splicing activity. *Curr Biol* **2004,** *14* (16), 1487-91.

24. Ryan, D. E.; Stevens, S. W.; Abelson, J., The 5' and 3' domains of yeast U6 snRNA: Lsm proteins facilitate binding of Prp24 protein to the U6 telestem region. *RNA* **2002,** *8* (8), 1011-33.

25. Jandrositz, A.; Guthrie, C., Evidence for a Prp24 binding site in U6 snRNA and in a putative intermediate in the annealing of U6 and U4 snRNAs. *EMBO J* **1995,** *14* (4), 820-32.

26. Stevens, S. W.; Ryan, D. E.; Ge, H. Y.; Moore, R. E.; Young, M. K.; Lee, T. D.; Abelson, J., Composition and functional characterization of the yeast spliceosomal penta-snRNP. *Mol Cell* **2002,** *9* (1), 31-44.

27. Vidaver, R. M.; Fortner, D. M.; Loos-Austin, L. S.; Brow, D. A., Multiple functions of Saccharomyces cerevisiae splicing protein Prp24 in U6 RNA structural rearrangements. *Genetics* **1999,** *153* (3), 1205-18.

28. Karaduman, R.; Fabrizio, P.; Hartmuth, K.; Urlaub, H.; Luhrmann, R., RNA structure and RNA-protein interactions in purified yeast U6 snRNPs. *J Mol Biol* **2006,** *356* (5), 1248-62.

29. Lamichhane, R.; Solem, A.; Black, W.; Rueda, D., Single-molecule FRET of protein-nucleic acid and protein-protein complexes: Surface passivation and immobilization. *Methods* **2010**.

30. Yean, S. L.; Wuenschell, G.; Termini, J.; Lin, R. J., Metal-ion coordination by U6 small nuclear RNA contributes to catalysis in the spliceosome. *Nature* **2000,** *408* (6814), 881-4.

31. Sontheimer, E. J.; Sun, S.; Piccirilli, J. A., Metal ion catalysis during splicing of premessenger RNA. *Nature* **1997,** *388* (6644), 801-5.

32. Fabrizio, P.; Abelson, J., Thiophosphates in yeast U6 snRNA specifically affect pre-mRNA splicing in vitro. *Nucleic Acids Res* **1992,** *20* (14), 3659-64.

33. Roy, R.; Hohng, S.; Ha, T., A practical guide to single-molecule FRET. *Nat Methods* **2008,** *5* (6), 507-16.

34. Cornish, P. V.; Ha, T., A survey of single-molecule techniques in chemical biology. *ACS Chem Biol* **2007,** *2* (1), 53-61.

35. Gordon, P. M.; Sontheimer, E. J.; Piccirilli, J. A., Metal ion catalysis during the exon-ligation step of nuclear pre-mRNA splicing: extending the parallels between the spliceosome and group II introns. *RNA* **2000,** *6* (2), 199-205.

36. Karaduman, R.; Dube, P.; Stark, H.; Fabrizio, P.; Kastner, B.; Luhrmann, R., Structure of yeast U6 snRNPs: arrangement of Prp24p and the LSm complex as revealed by electron microscopy. *RNA* **2008,** *14* (12), 2528-37.

CHAPTER 5

Conclusions and future directions

Splicing is one of the essential steps in the maturation of the pre-mRNA, during which the non-coding sequences (introns) are removed and the sequences decoding proteins (exons) are ligated together[1,2]. Spliceosomes catalyze splicing in organisms ranging from yeast to humans. The spliceosome is a highly dynamic ribonucleoprotein complex, composed of 5 small nuclear RNAs (U1, U2, U4, U5 and U6 snRNAs) and about 80 proteins in yeast and 170 in human[3,4]. Its catalytic core comprises three small nuclear RNAs (U2, U5 and U6) involved in substrate positioning and catalysis. U5 is showed to be functionally dispensable for the catalysis *in vitro*[5,6], so it is highly possible that the U2 and U6 snRNAs form the catalytic core of the spliceosome and directly play a pivotal role in catalysis[7,8]. It has been proposed, but never shown experimentally, that the U2/U6 complex adopts multiple conformations that reflect different activation states of the spliceosome[9,7]. The active structure of U2/U6 has been a matter of debate over recent years. Genetics study in yeast revealed that the AGC triad forms base pairs with the GCU in U2 yielding the catalytically important helix IB and the U2/U6 complex adopts a three-way junction structure[7]. Helix IB was shown to be very important for the second step of splicing, so this three-way junction structure may represent the active site of the second step of splicing[7,10]. However, according to a more recent NMR study, the U2/U6 adopts a four-way junction structure with an extended U6 ISL that may be the active structure of the first step of splicing[9] (Figure 1.5B). My single molecule study of U2/U6 provides some insights into this issue.

In this thesis work, I have developed a single-molecule FRET (smFRET) assay to

probe the structural dynamics of a protein-free U2–U6 RNA complex from yeast in solution[11]. I tested the folding dynamics of the wild type U2/U6 and mutants of highly conserved regions of U2/U6. My data show the presence of at least three distinct conformations in equilibrium[11]. Both the three-way junction and the four-way junction structures may exist in solution. The minimal folding pathway consists of a two-step process with an obligatory intermediate[11]. The first step is strongly magnesium dependent, and we provide evidence suggesting that the second step corresponds to the formation of the genetically conserved helix IB. Site-specific mutations in the highly conserved AGC triad and the U80 base in U6 suggest that the observed conformational dynamics correlate with residues that have an important role in splicing[11]. Our data supports the hypothesis that U2–U6 adopts multiple conformations at various splicing stages[12], which may or may not reflect the presence of unique active sites for each step[13]. In the high FRET conformation (N), the AGC triad extends the ISL as predicted for the four-helix structure[9], and the ACAGAGA loop and U80 in the ISL are brought into close proximity by a stabilizing tertiary contact. Because of the expected relationship between these three regions during the first splicing step, it is possible that this conformation resembles the active conformation of the first step of splicing. In the mid-FRET conformation, the AGC triad still forms the extended ISL (four-helix structure)[9], but the tertiary contact between the ACAGAGA loop and U80 is no longer formed, as its stability does not depend on the U80 mutations (Chapter 3). Our current data, however, do not allow us to conclude whether this conformation has a more direct role in splicing. In the low FRET conformation, a junction migration takes place, resulting in the formation of helix IB (three-helix junction)[7]. Helix I is important for both steps of

splicing, and a conformational rearrangement must precede each step[14]. This result is in agreement with our mutational data that link the formation of the low FRET conformation to an activating step between the first and second splicing steps. It is tempting to propose that the observed structural dynamics also have an activating role in the first splicing step, but the mutations tested here do not provide enough support for this conclusion.

Furthermore, according to the folding of the some ribozymes like hairpin ribozymes and the NMR structure of the protein free U2/U6 complex, Butcher and coworkers proposed that the U2/U6 complex folds to the four-way junction structure autonomously and the formation of helix IB may need the assistance of proteins[9,15]. In the presence of an essentially constant supply of Mg^{2+} ions *in vivo*, it is possible that the role of at least some of the spliceosomal proteins is to adjust the relative stability of these conformations to time the U2–U6 structural dynamics for accurate and efficient splicing.

I also report the first direct structural evidence that supports the existence of the base triples in the spliceosomal snRNA U2/U6 in this thesis work. These interactions were proposed according to a corresponding set of base triple interactions discovered in the recent published crystal structure of the self-splicing group II intron[16,17]. We proposed that these base triples existing in the spliceosomal RNA U2/U6 complex are in the same family of the ones found in crystal structure of the self-splicing group II intron given the extensive similarities between the spliceosome and the group II intron. Our data agree very well with the hypothesis. The AGC triad in U6 is highly conserved and most mutations within it lead to a severe deficiency of splicing[7,18,10]. In contrast, mutations of their base pairing partners in U2 or in U6 ISL do not[7,19,10]. These data indicate that the

AGC triad may have multiple roles besides forming base pairs in extended U6 ISL or helix IB. The recent crystal of group II intron suggest that the AGC triad may participate in base triple interactions that are important in maintaining the compact and delicate structure of the active core[17]. The putative base triples in U2/U6 may have a very significant role in the function of the spliceosome. They are able to draw essential parts for the first step of splicing in close proximity, which are otherwise far away from each other in primary sequence. The base triple G86-C61•U80 brings the catalytic important Mg^{2+} close to the proposed active site for the first step of splicing, containing the 5'splice site bound with ACAGAGA loop, branch site and the catalytic triad AGC. Both U88-A59•A53 and U87•G60•G52 bring the 5'splice site and the branch site to the active core (Figure 3.1A). But in contrast to the other two that play a central role in maintaining the overall structure and catalysis, the U88-A59•A53 base triple may only have a supporting role in stabilizing the active structure, similar to its counterpart in the group II intron. This may be because it is on the bottom of this base triple sandwich and not important for the stacking. The G60 is invariant in U6 and so is the analogous G359 in the group II intron. This may be able to explain the failure of rescue of this base triple. Any change made in this base triple disrupts the overall structure, which might be able to be attributed to the breaking of the stacking interactions among three base triples.

The base triples may also play a significant role in position catalytically important Mg^{2+} ions for the two-metal ion catalysis proposed for the spliceosome. Most ribozymes require divalent ions such as Mg^{2+} as cofactors for catalysis[20]. Numerous biochemical and structural studies of the group II intron suggest a two-metal ion mechanism for the splicing reaction[21,22]. And the spliceosome may also utilize the same mechanism[23]. It was

revealed that catalytically important divalent ions bind in the DV bulge and the catalytic triad[16, 24]. The crystal structure of the group II intron demonstrated that the triple helix provides a platform for metal ion binding with a 3.9 Å distance between two metal ions, which agrees very well with the two-metal ion model (Figure 3.2)[16, 24b,23]. Analogous regions in U6 has been shown to bind Mg^{2+} ions involved in structure and catalysis[18,25,26]. U80 binds a Mg^{2+} ion essential for first step of splicing reaction[25]. Several positions in the ACAGAGA loop and AGC triad may have a role in metal ion binding[18]. The hypothetic triple helix in U2/U6 that brings these regions together may also be able to position two metal ions in the active core of the spliceosome for the catalysis in a similar way as in the group II intron.

Our data support a model that extended U6 ISL with the AGC base pair with GUU of U6 that participates in the triple helix formation[9]. If the extended U6 ISL is disrupted, the high FRET state is also decreased. In Chapter 2, we demonstrated that under low $MgCl_2$ condition (5mM), most of U2/U6 complexes prefer the high FRET four-way junction structure with an extended U6 ISL, and with high concentration of $MgCl_2$ (40mM-100mM), the intermediate state with a medium FRET is stabilized drastically and enable the formation of low FRET three-way helix structure with helix IB[11] (Figure 3.10). This is surprising and people may expect the opposite. The base triple model proposed in this report offers an explanation for these data (Figure 3.10). According to a NMR study of U6 ISL, with low concentration of $MgCl_2$, U80 prefer to flip out of the helix of U6 ISL[27], which enables the base triple formation between U80 and C61-G86. This is also observed for DV bulge in the crystal structure of the group II intron[16]. Higher concentration of $MgCl_2$ stabilizes the conformation with U80 stacked in

U6 ISL[27] and disrupts the base triples (Figure 3.10). This conformation is highly possible to be the intermediate state in our folding pathway of U2/U6[11]. In addition, according to the folding nature of the four-way junction structure, it may be able to adopt a conformation in which the ACAGAGA loop and the U80 are in close proximity (Figure 3.1D)[28,9,29-30]. However, because the triple interactions mainly occur between the AGC triad and a third strand (U80, G52 and A53) and the base-paring partner of AGC triad (G86-U88 in U6 or G21-U23 in U2) does not directly participate in the interactions. So we cannot totally rule out the possibility that the AGC forms a triplex in the form of helix IB. It is possible the base triples form during both steps of splicing and play pivotal structural and functional role in catalysis, similar to proposed role of base triples of group II intron[17,31]. Interestingly, both in the group II intron and spliceosomal snRNA U2/U6, the catalytic triad AGC is highly sensitive to mutations, while their base paring partners only lead to minor defects in splicing[32,33].

It is highly possible that spliceosomal proteins are able to stabilize the U6 ISL and base triples *in vivo*. For example, prp8 has been shown to interact extensively with snRNAs in the active core and is required for catalysis. This leads to the hypothesis that both RNA and protein components of spliceosome are involved directly in catalysis[34]. Proteins that bind U6 and are essential for the first step of splicing, for example Cwc 25, may stabilize the triple helix *in vivo*[35].

In this thesis work, I only tested the structural dynamics of protein free U2/U6 complex. However, it is more interesting to observe the dynamics of U2/U6 in the whole spliceosome. The more than 70 proteins in the spliceosome may interact with the U2/U6 and change the structural dynamics. Dr. Amanda Solem is working on this in the lab now.

She obtained the splicing extract from yeast and replaced the native U6 of the spliceosome with Cy3-Cy5 labeled U6. She observed very interesting dynamics of U6 in spliceosome. This approach has great potential to understanding the structural dynamics of snRNAs and the assembly of the spliceosome under a biological relevant condition. It is interesting to test with this approach that whether the structural dynamics I observed with protein-free U2/U6 is also exist in the spliceosome. It is also very important to test the mutations that affect the folding dynamics of protein U2/U6 in the whole spliceosome to further link the dynamics of U2/U6 to the splicing reaction.

There are a large number of proteins in the spliceosome that play vital structural and catalytic roles[3,36]. RNA Chaperones, like Prp24, help the structural rearrangements of spliceosomal snRNAs[37,38]. The spliceosomal protein Prp24 is an essential component of the U6 snRNP in yeast and helps U6 to remodel during assembly and catalysis of the spliceosome[39,40]. We report single molecule FRET data showing that spliceosomal protein Prp24 can induce a conformational change in U2/U6 complex. This RNA:protein interaction is Mg^{2+} and protein concentration dependent. I also proposed a mechanism by which Prp24 modulates the structural dynamics of U2/U6.

Mg^{2+} ions help the binding of Prp24 to U2/U6, but in contrast, Prp24 inhibits the binding Mg^{2+}. The Mg^{2+} binding sites have been proposed in the highly conserved regions, the ACAGAGA loop, AGC triad and U80[18,41,42,25], overlapping with binding sites of Prp24[43,44]. Prp24 may bind these regions and inhibit binding of incoming Mg^{2+} ions.

Prp24 has been shown to be able to anneal U4 and U6 to form a complex[38,37]. A recent NMR study with a segment of U6 and RRM2 of Prp24 revealed possible binding

sites of Prp24 in U6[43]. In the resulting model, RRM2 of Prp24 binds the GAGA box in the ACAGAG loop in a sequence-specific fashion and RRM1 interacts with U6 via charge-charge interactions. Binding of Prp24 may be able to unwind part of the internal helix of U6 snRNA and expose bases for annealing with U4[43]. According to that model and our single molecule data, we proposed a mechanism by which Prp24 modulates the structural dynamics of U2/U6. As described in Chapter 3, with low concentration of Mg^{2+}, there are base triples that stabilize the high FRET state. Prp24 RRM1 and RRM2 bind the nucleotides that participate in the base triples (G52, A53 and AGC triad) and the binding of Prp24 may protect these nucleotides and disrupt the base triples and the high FRET state. At the same time, helix Ia and part of the U6 ISL may be unwound and this partial open conformation of U2/U6 may correspond to the intermediate FRET state observed. The opening of the U6 ISL facilitates the formation of helix IB and Prp24 may stay bound after the formation of three-helix structure with helix IB, stabilizing it form going back to the intermediate state and high FRET state. This may explain why with low concentrations of Prp24, the intermediate state becomes less defined. If the protein dissociates form U2/U6 before annealing, the partial open complex may adopt random conformations. In addition, low FRET state is less populated and also become less defined with low concentration of Prp24, suggesting that the binding of Prp24 stabilizes this conformation.

It has been shown that the first RRMs of Prp24, RRM1 and RRM2, are important for high affinity binding of Prp24 to U6 snRNA and the function of RRM3 and RRM4 is unclear. It is possible that RRM3 and RRM4 are able to bind U2 and stabilize the complex formed between U2 and U6. The reason why Prp24 prefers to stabilize the

three-way helix structure of U2/U6 may be that binding of Prp24 disrupts the base-triple interactions that stabilize the four–way helix structure. Furthermore, we showed evidence that destabilizing of the base triples may favor the formation of intermediate state and enable the transitions to the three-way helix conformation in Chapter 3. Many functions of Prp24 have been suggested and proposed, including annealing and unwinding of U4/U6[43,37,38], modulating the structural dynamics of the U2/U6. To conduct these multiple functions that are sometimes conflicting with each other, there must be some regulation mechanism for Prp24 function. Lsm proteins interact with Prp24 and enhance the annealing of U4/U6[45,46,47]. In addition, other spliceosomal protein factors, especially regulation factors, may modulate the function of Prp24 at different stages of splicing.

According to genetic studies, Prp24 promotes the annealing of U4/U6[37,38]. Structural studies with full length and truncated U6 and Prp24 revealed the possible binding sites of Prp24 to U6[39,44]. More recently, NMR studies revealed the sequence-specific binding of RRM2 of Prp24 and proposed that Prp24 can open partially the U6 ISL to facilitate the annealing of U4/U6. Because U2/U6 is believed to be the catalytic core of the spliceosome and Prp24 is an essential component of U6 snRNP[40,7], we hypothesize that Prp24 also play an important role in modulating the structural dynamics of U2/U6. Our bulk and single molecule data is consistent with this idea. Prp24 may bind U6 in similar binding sites revealed by pervious studies and stabilize the three-way junction by a similar mechanism proposed according to the NMR study[43] (Figure 4.6).

To further test this model, more experiments are performing in the lab now by Chandani Wanarsooriya. She is testing the binding of Prp24 with U6 only. This experiment is able to reveal that whether Prp24 modulates the conformational change of

U2/U6 or unwinds the U2 from U6. In addition, Prp24 titrations to U2/U6 with various Prp24 concentrations are conducting in the lab. The single molecule K_D and more details about the binding of Prp24 to U2/U6 will be obtained. More bulk titrations with various concentrations of Prp24 and Mg^{2+} should be conducted to further understand the effect of Prp24 on Mg^{2+} binding. Furthermore, the effect of mutant Prp24 constructs (deletion of one or more RRMs), like Prp24 234C, N123, 23 and 12, on the folding dynamics of U2/U6 will be tested to assess the function of each RRM of Prp24. This study will lead to an important mechanism by which the Prp24 modulates the structure of U2/U6.

Part of Chapter 5 is adopted from paper: Guo, Z.; Karunatilaka, K. S.; Rueda, D., Single-molecule analysis of protein-free U2-U6 snRNAs. Nat Struct Mol Biol 2009, 16 (11), 1154-9.

Reference

1. Burge, C. B.; Tuschl, T. A.; Sharp, P. A., Splicing of precursors to mRNAs by the spliceosomes. *in The RNA World 2nd edition, R. F. Gesteland, T. R. Cech, J. F. Atkins (eds.), Cold Spring Harbor Laborotory Press, Cold Spring Harbor, NY, 1999, 525-560.* **1999**.

2. Sharp, P. A., On the origin of RNA splicing and introns. *Cell* **1985,** *42* (2), 397-400.

3. Wahl, M. C.; Will, C. L.; Luhrmann, R., The spliceosome: design principles of a dynamic RNP machine. *Cell* **2009,** *136* (4), 701-18.

4. Madhani, H. D.; Guthrie, C., Dynamic RNA-RNA interactions in the spliceosome. *Annu Rev Genet* **1994,** *28*, 1-26.

5. O'Keefe, R. T.; Norman, C.; Newman, A. J., The invariant U5 snRNA loop 1 sequence is dispensable for the first catalytic step of pre-mRNA splicing in yeast. *Cell* **1996,** *86* (4), 679-89.

6. Segault, V.; Will, C. L.; Polycarpou-Schwarz, M.; Mattaj, I. W.; Branlant, C.; Luhrmann, R., Conserved loop I of U5 small nuclear RNA is dispensable for both catalytic steps of pre-mRNA splicing in HeLa nuclear extracts. *Mol Cell Biol* **1999,** *19* (4), 2782-90.

7. Madhani, H. D.; Guthrie, C., A novel base-pairing interaction between U2 and U6 snRNAs suggests a mechanism for the catalytic activation of the spliceosome. *Cell* **1992,** *71* (5), 803-17.

8. Parker, R.; Siliciano, P. G.; Guthrie, C., Recognition of the TACTAAC box during mRNA splicing in yeast involves base pairing to the U2-like snRNA. *Cell* **1987,** *49* (2), 229-39.

9. Sashital, D. G.; Cornilescu, G.; McManus, C. J.; Brow, D. A.; Butcher, S. E., U2-U6 RNA folding reveals a group II intron-like domain and a four-helix junction. *Nat Struct Mol Biol* **2004,** *11* (12), 1237-42.

10. Hilliker, A. K.; Staley, J. P., Multiple functions for the invariant AGC triad of U6 snRNA. *RNA* **2004,** *10* (6), 921-8.

11. Guo, Z.; Karunatilaka, K. S.; Rueda, D., Single-molecule analysis of protein-free U2-U6 snRNAs. *Nat Struct Mol Biol* **2009,** *16* (11), 1154-9.

12. Rhode, B. M.; Hartmuth, K.; Westhof, E.; Luhrmann, R., Proximity of conserved U6 and U2 snRNA elements to the 5' splice site region in activated spliceosomes. *EMBO J* **2006,** *25* (11), 2475-86.

13. Moore, M. J.; Sharp, P. A., Evidence for two active sites in the spliceosome provided by stereochemistry of pre-mRNA splicing. *Nature* **1993,** *365* (6444), 364-8.

14. Mefford, M. A.; Staley, J. P., Evidence that U2/U6 helix I promotes both catalytic steps of pre-mRNA splicing and rearranges in between these steps. *RNA* **2009,** *15* (7), 1386-97.

15. Rupert, P. B.; Ferre-D'Amare, A. R., Crystal structure of a hairpin ribozyme-inhibitor complex with implications for catalysis. *Nature* **2001,** *410* (6830), 780-6.

16. Toor, N.; Keating, K. S.; Taylor, S. D.; Pyle, A. M., Crystal structure of a self-spliced group II intron. *Science* **2008,** *320* (5872), 77-82.

17. Keating, K. S.; Toor, N.; Perlman, P. S.; Pyle, A. M., A structural analysis of the group II intron active site and implications for the spliceosome. *RNA* **2010,** *16* (1), 1-9.

18. Fabrizio, P.; Abelson, J., Thiophosphates in yeast U6 snRNA specifically affect pre-mRNA splicing in vitro. *Nucleic Acids Res* **1992,** *20* (14), 3659-64.

19. Ryan, D. E.; Abelson, J., The conserved central domain of yeast U6 snRNA: importance of U2-U6 helix Ia in spliceosome assembly. *RNA* **2002,** *8* (8), 997-1010.

20. Cech, T. R.; Bass, B. L., Biological catalysis by RNA. *Annu Rev Biochem* **1986,** *55*, 599-629.

21. Pyle, A. M.; Lambowitz, A. M., Group II Introns: Ribozymes That Splice RNA and Invade DNA. *in The RNA World, 3rd edition (ed. R.F. Gesteland et al.), pp. 469-505. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.* **2006**.

22. Lambowitz, A. M.; Zimmerly, S., Mobile group II introns. *Annu Rev Genet* **2004,** *38*, 1-35.

23. Steitz, T. A.; Steitz, J. A., A general two-metal-ion mechanism for catalytic RNA. *Proc Natl Acad Sci U S A* **1993,** *90* (14), 6498-502.

24. (a) Gordon, P. M.; Piccirilli, J. A., Metal ion coordination by the AGC triad in domain 5 contributes to group II intron catalysis. *Nat Struct Biol* **2001,** *8* (10), 893-8; (b) Toor, N.; Keating, K. S.; Fedorova, O.; Rajashankar, K.; Wang, J.; Pyle, A. M., Tertiary architecture of the Oceanobacillus iheyensis group II intron. *RNA* **2010,** *16* (1), 57-69.

25. Yean, S. L.; Wuenschell, G.; Termini, J.; Lin, R. J., Metal-ion coordination by U6 small nuclear RNA contributes to catalysis in the spliceosome. *Nature* **2000,** *408* (6814), 881-4.

26. Yu, Y. T.; Maroney, P. A.; Darzynkiwicz, E.; Nilsen, T. W., U6 snRNA function in nuclear pre-mRNA splicing: a phosphorothioate interference analysis of the U6 phosphate backbone. *RNA* **1995,** *1* (1), 46-54.

27. Huppler, A.; Nikstad, L. J.; Allmann, A. M.; Brow, D. A.; Butcher, S. E., Metal binding and base ionization in the U6 RNA intramolecular stem-loop structure. *Nat Struct Biol* **2002,** *9* (6), 431-5.

28. Butcher, S. E.; Brow, D. A., Towards understanding the catalytic core structure of the spliceosome. *Biochem Soc Trans* **2005,** *33* (Pt 3), 447-9.

29. Valadkhan, S.; Manley, J. L., Intrinsic metal binding by a spliceosomal RNA. *Nat Struct Biol* **2002,** *9* (7), 498-9.

30. Murchie, A. I.; Thomson, J. B.; Walter, F.; Lilley, D. M., Folding of the hairpin ribozyme in its natural conformation achieves close physical proximity of the loops. *Mol Cell* **1998,** *1* (6), 873-81.

31. de Lencastre, A.; Hamill, S.; Pyle, A. M., A single active-site region for a group II intron. *Nat Struct Mol Biol* **2005,** *12* (7), 626-7.

32. Boulanger, S. C.; Belcher, S. M.; Schmidt, U.; Dib-Hajj, S. D.; Schmidt, T.; Perlman, P. S., Studies of point mutants define three essential paired nucleotides in the domain 5 substructure of a group II intron. *Mol Cell Biol* **1995,** *15* (8), 4479-88.

33. Peebles, C. L.; Zhang, M.; Perlman, P. S.; Franzen, J. S., Catalytically critical nucleotide in domain 5 of a group II intron. *Proc Natl Acad Sci U S A* **1995,** *92* (10), 4422-6.

34. Abelson, J., Is the spliceosome a ribonucleoprotein enzyme? *Nat Struct Mol Biol* **2008,** *15* (12), 1235-7.

35. Warkocki, Z.; Odenwalder, P.; Schmitzova, J.; Platzmann, F.; Stark, H.; Urlaub, H.; Ficner, R.; Fabrizio, P.; Luhrmann, R., Reconstitution of both steps of Saccharomyces cerevisiae splicing with purified spliceosomal components. *Nat Struct Mol Biol* **2009,** *16* (12), 1237-43.

36. Staley, J. P.; Guthrie, C., Mechanical devices of the spliceosome: motors, clocks, springs, and things. *Cell* **1998,** *92* (3), 315-26.

37. Raghunathan, P. L.; Guthrie, C., A spliceosomal recycling factor that reanneals U4 and U6 small nuclear ribonucleoprotein particles. *Science* **1998,** *279* (5352), 857-60.

38. Vidaver, R. M.; Fortner, D. M.; Loos-Austin, L. S.; Brow, D. A., Multiple functions of Saccharomyces cerevisiae splicing protein Prp24 in U6 RNA structural rearrangements. *Genetics* **1999,** *153* (3), 1205-18.

39. Kwan, S. S.; Brow, D. A., The N- and C-terminal RNA recognition motifs of splicing factor Prp24 have distinct functions in U6 RNA binding. *RNA* **2005,** *11* (5), 808-20.

40. Bae, E.; Reiter, N. J.; Bingman, C. A.; Kwan, S. S.; Lee, D.; Phillips, G. N., Jr.; Butcher, S. E.; Brow, D. A., Structure and interactions of the first three RNA recognition motifs of splicing factor prp24. *J Mol Biol* **2007,** *367* (5), 1447-58.

41. Gordon, P. M.; Sontheimer, E. J.; Piccirilli, J. A., Metal ion catalysis during the exon-ligation step of nuclear pre-mRNA splicing: extending the parallels between the spliceosome and group II introns. *RNA* **2000,** *6* (2), 199-205.

42. Sontheimer, E. J.; Sun, S.; Piccirilli, J. A., Metal ion catalysis during splicing of premessenger RNA. *Nature* **1997,** *388* (6644), 801-5.

43. Martin-Tumasz, S.; Reiter, N. J.; Brow, D. A.; Butcher, S. E., Structure and functional implications of a complex containing a segment of U6 RNA bound by a domain of Prp24. *RNA* **2010,** *16* (4), 792-804.

44. Karaduman, R.; Fabrizio, P.; Hartmuth, K.; Urlaub, H.; Luhrmann, R., RNA structure and RNA-protein interactions in purified yeast U6 snRNPs. *J Mol Biol* **2006,** *356* (5), 1248-62.

45. Verdone, L.; Galardi, S.; Page, D.; Beggs, J. D., Lsm proteins promote regeneration of pre-mRNA splicing activity. *Curr Biol* **2004,** *14* (16), 1487-91.

46. Ryan, D. E.; Stevens, S. W.; Abelson, J., The 5' and 3' domains of yeast U6 snRNA: Lsm proteins facilitate binding of Prp24 protein to the U6 telestem region. *RNA* **2002,** *8* (8), 1011-33.

47. Karaduman, R.; Dube, P.; Stark, H.; Fabrizio, P.; Kastner, B.; Luhrmann, R., Structure of yeast U6 snRNPs: arrangement of Prp24p and the LSm complex as revealed by electron microscopy. *RNA* **2008,** *14* (12), 2528-37.

48. Toor, N.; Rajashankar, K.; Keating, K. S.; Pyle, A. M., Structural basis for exon recognition by a group II intron. *Nat Struct Mol Biol* **2008,** *15* (11), 1221-2.

ABSTRACT

SINGLE MOLECULE STUDIES OF SPLICEOSOMAL snRNAs U2-U6

by

ZHUOJUN GUO

DECEMBER 2010

Advisor: Dr. David Rueda

Major: Chemistry

Degree: Doctor of Philosophy

Spliceosomes catalyze the maturation of precursor mRNAs in organisms ranging from yeast to humans. Their catalytic core comprises three small nuclear RNAs (U2, U5 and U6) involved in substrate positioning and catalysis. It has been postulated, but never shown experimentally, that the U2–U6 complex adopts at least two conformations that reflect different activation states. We have used single-molecule fluorescence to probe the structural dynamics of a protein-free RNA complex modeling U2–U6 from yeast and mutants of highly conserved regions of U2–U6. Our data show the presence of at least three distinct conformations in equilibrium. The minimal folding pathway consists of a two-step process with an obligatory intermediate. The first step is strongly magnesium dependent, and we provide evidence suggesting that the second step corresponds to the formation of the genetically conserved helix IB. Site-specific mutations in the highly conserved AGC triad and the U80 base in U6 suggest that the observed conformational dynamics correlate with residues that have an important role in splicing. We also report the first direct structural evidence that supports the existence of the base triples in the

spliceosomal snRNA U2/U6. These interactions were proposed according to a corresponding set of base triple interactions discovered in the recent published crystal structure of the self-splicing group II intron.[19-16, 24b-17, 48] We proposed that these base triples existing in the spliceosomal RNA U2/U6 complex are in the same family of the ones found in crystal structure of the self-splicing group II intron given the extensive similarities between the spliceosome and the group II intron. Our data agree very well with the hypothesis. There are a large number of proteins in the spliceosome that play vital structural and catalytic roles[3,36]. RNA Chaperones, like Prp24, help the structural rearrangements of spliceosomal snRNAs. We report single molecule FRET data showing that spliceosomal protein Prp24 can induce a conformational change in U2/U6 complex. This RNA:protein interaction is Mg^{2+} and protein concentration dependent and inhibits the binding of Mg^{2+} to U2/U6.

AUTOBIOGRAPHICAL STATEMENT

EDUCATION

2005-Present Chemistry *Wayne State University* Ph.D.
Thesis: Single Molecule Studies of Spliceosomal snRNAs U2-U6 (Advisor: David Rueda)

2001-2005 Chemistry *Nankai University, China* BS

PRESENTATIONS

Oral Presentation

Spring 2009 Splicing Mechanisms: Lessons from Single Molecule Spectroscopy
Michigan RNA Conference Albion, MI

Posters

Summer 2009	RNA Society	Madison, WI
Spring 2009	ASBMB	New Orleans, LA
Summer 2008	Gordon Research Conference	Newport, RI
Summer 2007	RNA Society	Madison, WI

PROFESSIONAL AFFILIATIONS

RNA Society, ACS, Phi Lambda Upsilon National Chemical Honor Society

PUBILICATIONS

1. **Guo, Z.**, Karunathilaka, K.S. and Rueda, D. Single-Molecule Folding of Protein Free U2/U6 snRNAs. **Nat. Struc. Mol. Biol**, 2009 Nov; 16(11): 1154-9.
2. **Guo, Z.**, and Wanarsooriya, C., Rueda, D. Single Molecule Study of protein free U2/U6 Reveals Existence of Base Triples. In preparation.