

3-3-2008

# RNA-metal Ion Interactions and Metal Ion-Induced Conformational Change in the Spliceosomal U2-U6 snRNA Complex Studied by Lanthanide Ion Luminescence and Resonance Energy Transfer Techniques

Faqing Yuan  
*Florida State University*

Follow this and additional works at: <http://diginole.lib.fsu.edu/etd>

---

## Recommended Citation

Yuan, Faqing, "RNA-metal Ion Interactions and Metal Ion-Induced Conformational Change in the Spliceosomal U2-U6 snRNA Complex Studied by Lanthanide Ion Luminescence and Resonance Energy Transfer Techniques" (2008). *Electronic Theses, Treatises and Dissertations*. Paper 933.

This Dissertation - Open Access is brought to you for free and open access by the The Graduate School at DigiNole Commons. It has been accepted for inclusion in Electronic Theses, Treatises and Dissertations by an authorized administrator of DigiNole Commons. For more information, please contact [lib-ir@fsu.edu](mailto:lib-ir@fsu.edu).

FLORIDA STATE UNIVERSITY  
COLLEGE OF ARTS AND SCIENCES

RNA-METAL ION INTERACTIONS AND METAL ION- INDUCED CONFORMATIONAL  
CHANGE IN THE SPLICEOSOMAL U2-U6 SNRNA COMPLEX STUDIED BY LANTHANIDE  
ION LUMINESCENCE AND RESONANCE ENERGY TRANSFER TECHNIQUES

BY  
FAQING YUAN

A Dissertation submitted to the  
Department of Chemistry & Biochemistry  
in partial fulfillment of the  
requirements for the degree of  
Doctor of Philosophy

Degree Awarded:  
Spring Semester, 2008

The members of the Committee approve the dissertation of Faqing Yuan defended on March 3<sup>rd</sup>, 2008.

---

Nancy L. Greenbaum  
Professor Co-Directing Dissertation

---

Geoffrey F. Strouse  
Professor Co-Directing Dissertation

---

P. Bryant Chase  
Outside Committee Member

---

Hong Li  
Committee Member

---

Qingxiang Amy Sang  
Committee Member

Approved:

---

Joseph B. Schlenoff, Department Chair, Department of Chemistry & Biochemistry

---

Joseph Travis, Dean, College of Arts & Sciences

The Office of Graduate Studies has verified and approved the above named committee members.

Dedicated to my parents for their unconditional love

## ACKNOWLEDGEMENTS

Looking back the six and half years I have spent in graduate school, a lot of emotions come to me. In addition to a PhD degree, the best thing I have done is that I have grown from a young person with superficial understanding of life to a mature man. I have learned tolerance, patience, and persistence. These qualities along with independent thinking and understanding of science will go a long way for my future life and career.

PhD study is not an easy task. Without the support of many significant people in my life, I would have given up. I would like to give my most sincere appreciation to my parents. Without their support, nothing is meaningful. Because I was born in a small village in Central China, the chance of going to college, least of all, coming to America, was slim. They have supported me to the fullest extent by overcoming numerous economical and social obstacles. I am grateful that I have had the best education possible because of them. I am sorry that I have been alone in America all these years and that I could not spend even a single Spring Festival with them. I know I can only work hard to make them proud.

I would also like to thank my PhD advisor Dr. Nancy Greenbaum for her continuous support over all these years. I know the advisor-student relationship is a kind of love-hate relationship. We have had some rough times and good times together. Graduating with a PhD degree concludes our relationship with all the good memories by erasing all the bad ones. I wish you all the best, Nancy!

I would also like to thank Dr. Geoffrey Strouse for being my co-advisor when Dr. Greenbaum was at Hunter College; Dr. Ken Goldby for providing me a place to stay and work; and Dr. Bryant Chase, Dr. Hong Li, and Dr. Amy Sang for being my helpful committee. I especially would like to thank Dr. Sang for her help and support. She recruited me to FSU with the recommendation of Dr. Yun-Bo Shi from NIH, and she gave me spiritual support when I was in a bad mood. I would also like to thank Dr. Tim Logan, Dr. John Dorsey, Dr. Lei Zhu, Dr. Brain Miller for their encouragement and support.

The department of Chemistry and Biochemistry has been really helpful throughout these years. It not only supports me financially, but also is a big help in academic and life matters. The faculty and staff are the real assets of the Florida State University. The chairman of the department, Dr. Joseph Schlenoff, deserves a lot of credit for the completion of my PhD study for his support.

In Greenbaum Lab and FSU, I have established a lot of friendships with so many great people. The first person is Dr. Kersten Schroeder, who is really a good friend and always helpful and fun to hang out with. Milena Popović, Joy Nelson, Masud Monwar, LauraJane Phelps, Laura Griffin...., these names are in my memory and I will miss all of you. I also would like to thank Dr. Volker Buschmann for his friendship and expert advice, and Mani Singh for being so nice with the lifetime measurements in chapter four.

I wish all my friends the best. I will work hard to make you proud! Thank you.

# TABLE OF CONTENTS

List of tables	-----viii
List of figures	----- ix
List of equations	-----xi
List of abbreviations and symbols	-----xiii
Abstract	-----xv
1. Introduction	----- 1
1.1 Introduction	-----1
1.2 Nucleic acids	-----2
1.3 Versatile roles of RNA in biological processes	-----5
1.4 Pre-mRNA splicing and the spliceosome	-----6
1.5 Splicing-related catalysis by protein-free spliceosomal snRNAs	-----7
1.6 Metal ions and U2-U6 snRNA complex	-----14
1.7 Specific aims	-----17
2. Use of a novel FRET method to localize site-bound terbium ions in the U2–U6 snRNA complex	-----19
2.1 Introduction	-----19
2.1.1 FRET method	-----19
2.1.2 An extension of FRET method	-----21
2.1.3 Specific metal ion binding on the U2-U6 snRNA complex	-----21
2.1.4 Goal of this chapter	-----23
2.2 Materials and methods	-----23
2.2.1 RNA design and synthesis	-----23
2.2.2 Non-denaturing gel electrophoresis	-----24
2.2.3 Measurement of Tb(III) luminescence	-----24
2.2.4 Measurement of metal ion binding stoichiometry to RNA constructs	-----25
2.2.5 Measurements of Tb(III) luminescence decay	-----25
2.2.6 FRET experiments using the Cy3-Cy5 pair	-----26
2.3 Results	-----26
2.3.1 Design of RNA samples and confirmation of base pairing	-----26
2.3.2 Stoichiometry of site-bound ions	-----26
2.3.3 Lifetime measurements of Tb(III) site-bound to unlabeled RNA	-----31
2.3.4 Determination of the angle between Helix I and U6 ISL	-----31
2.3.5 Lifetime measurements with Cy3-labeled RNA constructs at pH 7.2	-----33
2.3.6 Calculation of Cy3-Tb(III) distances and localization of ion binding sites	-----36
2.3.7 Confirmation of ion-binding locations by use of mutations	-----36
2.3.8 pH-dependence of ion binding	-----38
2.4 Discussion	-----39
3. Characterization of terbium ions bound on RNA and investigation of their suitability as FRET	

donors	-----43
3.1 Introduction	-----43
3.1.1 RNA structure and metal ions	-----43
3.1.2 Probing RNA-metal ion binding by Tb(III)	-----43
3.1.3 Tb(III) bound on RNA as FRET donors	-----44
3.1.4 Properties of Tb(III) bound on RNA and their suitability as FRET donors	-----46
3.1.5 Specific aims	-----48
3.1.6 Model system: GUAA loop	-----49
3.2 Materials and Methods	-----49
3.2.1 Construct design	-----49
3.2.2 Tb(III) luminescence intensity measurements	-----51
3.2.3 Tb(III) luminescence lifetime measurements	-----51
3.2.4 QY of Tb(III) bound on RNA measurements	-----51
3.3 Results	-----52
3.3.1 Steady state signal accuracy: compensation factor	-----52
3.3.2 Degree of quenching: H <sub>2</sub> O vs D <sub>2</sub> O	-----55
3.3.3 Binding constant of the GUAA RNA and Tb(III)	-----59
3.3.4 Tb(III) luminescence lifetime and free Tb(III) fraction	-----66
3.3.5 Fix the lifetime of Tb(III) bound on RNA	-----67
3.3.6 QY of Tb(III) bound on RNA	-----65
3.3.7 Diffusional energy transfer process	-----68
3.3.8 Structure distortion by a group II intron branch site	-----71
3.4 Discussion	-----73
3.4.1 Compensation factor	-----73
3.4.2 “Diffusional equilibrium” hypothesis	-----73
3.4.3 Suitable conditions for Tb(III) bound on RNA to be FRET donors	-----75
3.4.4 Using Tb(III) bound on RNA as FRET donors	-----75
3.4.5 Possibility of designing a new class of FRET donors	-----76
4. Measurement of metal ion-induced conformational change in the spliceosomal U2-U6 snRNA complex	-----78
4.1 Introduction	-----78
4.1.1 Motivation of research: folding hypothesis	-----78
4.1.2 Possible factors which may induce the conformational change	-----79
4.1.3 Two possible folding schemes	-----79
4.1.4 Nano-surface energy transfer (NSET)	-----80
4.1.5 Goal of this chapter	-----82
4.2 Materials and methods	-----83
4.2.1 Design of U2-U6 snRNA constructs	-----83
4.2.2 RNA sample preparation	-----86
4.2.3 Intron design and synthesis	-----88
4.2.4 Proteins expression and purification	-----90
4.2.5 FRET and NSET instrumentation	-----90
4.3 Results	-----90

4.3.1 Confirmation of complex formation	-----90
4.3.2 Distance measurements results	-----94
4.4 Discussion	-----101
5. Conclusions and future direction	-----103
5.1 Conclusions	-----103
5.1.1 Tb(III) binding sites in the U2-U6 snRNA complex	-----103
5.1.2 Tb(III)-RNA binding	-----104
5.1.3 Tb(III) bound on RNA as FRET donors	-----104
5.1.4 Conformational change of the U2-U6 snRNA complex	-----105
5.2 Future direction	-----105
5.2.1 Design a new class of FRET donors	-----105
5.2.2 Measure conformational changes in the U2-U6 snRNA complex	-----106
Appendix: Common lab protocols	-----108
References	-----112
Biographical sketch	-----120

## LIST OF TABLES

Table 1.1 Structural properties of nucleic acid helices	----- 4
Table 2.1 Tb(III) binding stoichiometry of RNA constructs at pH 7.2	-----30
Table 2.2 Distances between Cy3 and Cy5 in different constructs	-----32
Table 2.3 Summary of Tb(III) lifetimes (in milliseconds) in the presence of acceptor ( $\tau_{DA}$ ) at different pH values	-----34
Table 3.1 Magnesium, Terbium and Europium ion size and coordination number comparison	- 4 5
Table 3.2 Intensity and lifetime of Tb(III) luminescence in H <sub>2</sub> O and D <sub>2</sub> O	-----57
Table 3.3 Lifetime data from mixing different concentrations of RNA and Tb(III)	-----63
Table 3.4 Energy transfer efficiencies measured in constructs dsGUAA and bpGUAA	-----72
Table 4.1 Lifetime and energy transfer efficiency results on design C	-----100

## LIST OF FIGURES

Figure 1.1 Chemical compositions of DNA and RNA	-----3
Figure 1.2 Regular Watson-Crick base pairs in RNA	-----4
Figure 1.3 Splicing is a two-step trans-esterification reaction	-----8
Figure 1.4 Proposed human U2 and U6 snRNA complex secondary structure	-----9
Figure 1.5 Interactions between U2-U6 snRNA complex and the intron	-----11
Figure 1.6 Secondary structure of a group II intron	-----12
Figure 1.7 the protein-free U2-U6 snRNA complex catalyzes splicing like reactions	-----13
Figure 1.8 Cryo-EM structure of pre-mRNA, SF3b155 and p14	-----15
Figure 1.9 Cartoon of “diffuse” ions and “chelated” ions	-----16
Figure 2.1 A FRET system with Tb(III) as donor and Cy3 as acceptor	-----22
Figure 2.2 The simplified construct and its mutations used in the FRET study	-----27
Figure 2.3 Demonstration of base pairing between top and bottom strands by a non-denaturing gel electrophoresis	-----28
Figure 2.4 Measurement of RNA metal-ion-binding stoichiometry by Job plot	-----30
Figure 2.5 Calculation of the angle between helix I and U6 ISL using the acceptor normalization method	-----32
Figure 2.6 (Figure next page) The exponential decay fitting process used in the study.	-----34

Figure 3.1 Terbium(III) energy level and spectra of excitation and emission	-----45
Figure 3.2 Three RNA sequences used in this chapter and its comparison with the sequence of ai57 group II intron	-----50
Figure 3.3 Lanthanide luminescence signal collection with a pulsed excitation	-----53&54
Figure 3.4 plot intensity in D <sub>2</sub> O/ $\tau_{D2O}$ vs intensity in H <sub>2</sub> O/ $\tau_{H2O}$	-----58
Figure 3.5 Binding constant calculations	-----61
Figure 3.6 Lifetimes against percentage of free Tb(III)	-----64
Figure 3.7 Quantum yield of Tb(III) bound on RNA luminescence	-----67
Figure 3.8 Free Tb(III) has an impact on diffusional energy transfer efficiency	-----69
Figure 4.1 Two possible folding patterns in U2-U6 snRNA complex	-----81
Figure 4.2 Design A	-----84
Figure 4.3 Design B	-----85
Figure 4.4 Design C: three-piece RNA system	-----87
Figure 4.5 Schematic sequence of an intron	-----89
Figure 4.6 Demonstration of invariant bottom U2-48 formed base pairs with variant top strands in design B	-----92
Figure 4.7 Confirmation of the triplex formation in design C	-----93
Figure 4.8 Lifetime of Cy3 in the Cy3-Cy5 labeled WT construct of design A	-----95
Figure 4.9 Acceptor normalization process for FAM-Cy3 FRET measurement	-----97
Figure 4.10 Lifetime of AF488 in the absence of AF555 in the WT of design B	-----98

## LIST OF EQUATIONS

- Equation 2.1** 
$$E = \frac{R_0^6}{R_0^6 + R^6}$$
- Equation 2.2** 
$$R_0 = 0.21 * \sqrt[6]{J * q_D * n^{-4} * k^2} \quad (\text{in } \text{\AA} \text{ angstroms})$$
- Equation 2.3** 
$$k^2 = (\cos \theta_{DA} - 3 \cos \theta_D \cos \theta_A)^2$$
- Equation 2.4** 
$$E = 1 - \frac{I_{DA}}{I_D}$$
- Equation 2.5** 
$$E = \frac{I_{AD}}{I_{AD,100\%}}$$
- Equation 2.6** 
$$E = 1 - \frac{\tau_{DA}}{\tau_D}$$
- Equation 2.7** 
$$I = I_0 + \sum_i A_i e^{\frac{-t}{\tau_i}}$$
- Equation 3.1** 
$$I = B_1 [Ln^{3+}] + B_2 [Ln^{3+}] \approx B [Ln^{3+}]$$
- Equation 3.2** 
$$\int_{0 \rightarrow \infty} I_0 * e^{-t/\tau} = -I_0 \tau e^{-t/\tau} = I_0 \tau$$
- Equation 3.3** 
$$\int_{T \rightarrow T_2} I_0 * e^{-t/\tau} = -I_0 \tau e^{-t/\tau} = I_0 \tau e^{-T/\tau} - I_0 \tau e^{-T_2/\tau}$$
- Equation 3.4** 
$$\int_{T \rightarrow T_2} I_0 * e^{-t/\tau} \approx \int_{T \rightarrow \infty} I_0 * e^{-t/\tau} \equiv I_0 \tau e^{-T/\tau}$$
- Equation 3.5** 
$$CF = \frac{\int_{0 \rightarrow \infty} I_0 * e^{-t/\tau}}{\int_{T \rightarrow \infty} I_0 * e^{-t/\tau}} = e^{T/\tau}$$
- Equation 3.6** 
$$CF_0 = \frac{e^{T/\tau}}{\tau}$$

**Equation 3.7** 
$$y = \frac{B}{2} * (C + x + k - \sqrt{(C + x + k)^2 - 4Cx})$$

**Equation 3.8** 
$$\tau = 660 \times (1 - A) + 427 \times A \text{ (\mu s)}$$

**Equation 3.9** 
$$\tau = \tau_1 \times (1 - A) + \tau_2 \times A$$

**Equation 3.10** 
$$cf = \frac{e^{T/\tau}}{\tau * QY}$$

**Equation 3.11** 
$$E = \frac{\frac{I_{AD}}{\tau_{AD} Q_a}}{\frac{I_{AD}}{\tau_{AD} Q_a} + \frac{I_{DA}}{\tau_{DA} Q_D}}$$

**Equation 3.12** 
$$Q_D = Q_a * \frac{I_{DA}}{I_{AD}} * \frac{\tau_D - \tau_{AD}}{\tau_{AD}}$$

## LIST OF ABBREVIATIONS AND SYMBOLS

Ψ	pseudouridine, modified RNA nucleoside
Å	angstrom, $10^{-10}$ meters
ε	extinction coefficient ( $M \cdot cm$ ) <sup>-1</sup> ; used in Beer-Lambert Law ( $A = \epsilon bc$ )
°C	degrees Celsius
μg	$10^{-6}$ grams
μL	$10^{-6}$ liters
μM	$10^{-6}$ molar (moles per liter)
μs	microsecond
2'-OH	ribose functional group at the 2' position, nucleophile in ribozyme
3'-ss	3'-splice site
<sup>32</sup> P	radioactive phosphorous-32
5'-ss	5'-splice site
AF	Alexa Fluor, if used as a dye, such as AF488, AF555, AF647 etc.
ATP	adenosine triphosphate
A.U.	arbitrary units
CF or cf	compensation factor
Cryo-EM	electron cryo-microscopy
d <sub>0</sub>	NSET radius
D <sub>2</sub> O	deuterium oxide
DEPC	diethyl pyrocarbonate
DNA	deoxyribonucleic acid
dsRNA	double-stranded RNA
DTT	dithiothreitol
EDTA	ethylene diamine tetraacetic acid
FAM	fluorescein
FRET	Förster Resonance Energy Transfer
HEPES	4-(2-Hydroxyethyl) piperazine-1-ethanesulfonic acid, buffer
ISL	intramolecular stem loop
ITC	isothermal titration calorimetry
k <sub>d</sub>	dissociation constant (a measure of binding affinity)
MES	2-(N-Morpholino) ethanesulfonic Acid
MeSH	Medical Subject Headings
mg	$10^{-3}$ grams
Ln(III)	lanthanide tripositive ions
miRNA	MicroRNA
mL	$10^{-3}$ liters

mm	$10^{-3}$ meters
mM	$10^{-3}$ molar (moles per liter)
MOPS	3-(N-morpholino)-propanesulfonic acid buffer
mRNA	messenger RNA
ms	millisecond
nm	$10^{-9}$ meters
NMR	nuclear magnetic resonance
ns	nanosecond
NSET	Nano-Surface Energy Transfer
O-H	Hydroxyl
PAGE	polyacrylamide gel electrophoresis
pmol	$10^{-12}$ moles
PMT	photomultiplier tube
pre-mRNA	precursor messenger RNA
ps	picosecond
$R_0$	Förster radius
RNA	ribonucleic acid
RNAase	ribonuclease
RNAi	RNA interference
RNP	ribonucleoprotein particle
rRNA	ribosomal RNA
SDS	sodium dodecyl sulfate
snRNA	small nuclear RNA
snoRNA	small nucleolar RNA
TCEP	Tris(2-carboxyethyl)phosphine
tRNA	transfer RNA
Tris	Tris (hydroxymethyl) aminomethane, buffer
UMP	uridine monophosphate
UV	ultraviolet
WC	Watson-Crick base pair
<b>Nucleosides</b>	
A	Adenosine
C	Cytidine
G	Guanine
U	Uridine
T	Thymine

## ABSTRACT

RNA splicing is an integral step of gene expression. In eukaryotic cells, it is carried out by a dynamic RNA-protein assembly called the spliceosome. Among its RNA components, U2 and U6 snRNAs are most conserved and form a complex with extensive base pairs. This complex is able to catalyze splicing-related reactions in the absence of proteins. Metal ions are required for the activity both *in vivo* and *in vitro*. The goal of this dissertation is to study the metal ion binding properties of the U2-U6 snRNA complex and elucidate the impact of metal ions on the complex structure. Spectroscopic informative Tb(III) is used to study RNA and metal ion binding properties. Specific metal ion binding sites and their pH-dependence on the complex are mapped out by using Tb(III) bound on RNA as FRET donors. The results not only confirm the metal ion binding sites suggested by other methods, but also reveal a new pH-dependent metal ion binding site. Proper experimental conditions of using Tb(III) as FRET donors are suggested. By using FRET and NSET methods, the structural impact of metal ions on the U2-U6 snRNA complex is monitored. Under the conditions tested in the study, no conformational changes have been observed upon the addition of metal ions. Our results introduce the use of Tb(III) bound on RNA as FRET donor for the first time; suggest proper conditions of how to use Tb(III) as FRET donor; and show that there is no conformational changes observed by the discussed experimental designs and methods.

# CHAPTER 1

## INTRODUCTION

### 1.1 Introduction

Ribonucleic acid (RNA) is a versatile molecule fulfilling numerous roles in many biological processes. The biological function of a molecule is tightly related to its structure; in the case of RNA, it is under intense investigation. Compared to its chemical-related counterpart deoxyribonucleic acid (DNA), RNA has a 2'-OH group on its backbone ribose which is missing in DNA. This hydroxyl group makes RNA highly labile, but also provides RNA the capacity of additional hydrogen bonding and metal ion coordination, and makes it more chemically active.

The electrostatic interactions between metal ions and the RNA backbone and other polar chemical groups have significant impact on RNA structure and function. Finding out what these impacts are will help us understand the various roles RNA plays and design research and therapeutic agents for various purposes.

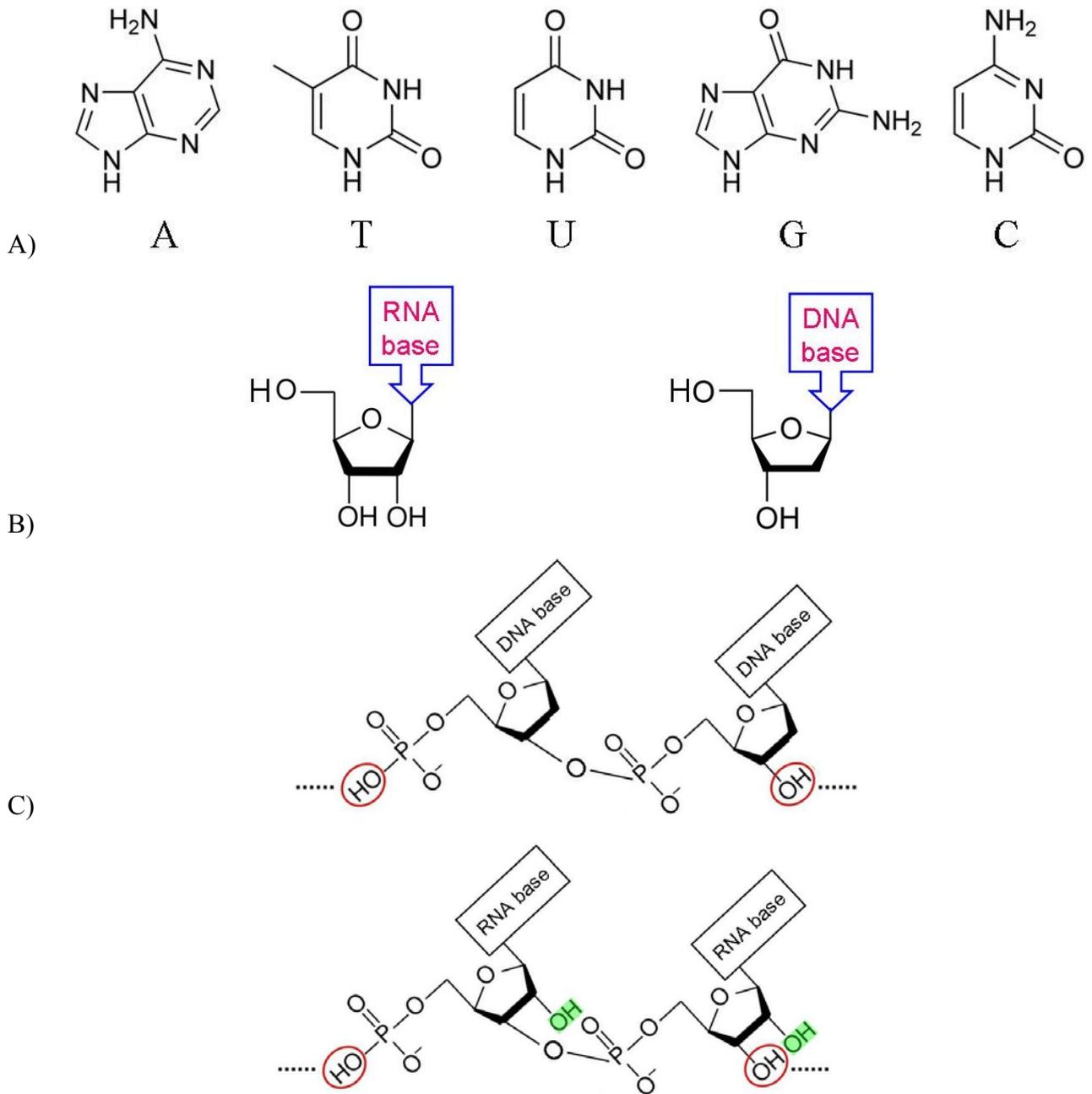
In the research described herein, the spliceosomal U2-U6 small nuclear (sn)RNA complex is used as a model system to study the role of metal ions on RNA structure and function. RNA splicing is an integral step in the maturation of precursor messenger RNA (pre-mRNA) and is carried out by a RNA-protein complex called the spliceosome in eukaryotic cells. The highly conserved U2 and U6 snRNAs base pair to form a complex which is thought to be at the catalytic core of the spliceosome. Metal ions are crucial for splicing activities. Two aspects of RNA-metal ion interactions are being investigated. The first objective is to determine how and where the metal ions bind on RNA. Lanthanide ions [ $\text{Ln}^{3+}$ ] are used as probes because  $\text{Ln}^{3+}$  react with RNA in a similar fashion as  $\text{Mg}^{2+}$  and unlike their biological counterpart  $\text{Mg}^{2+}$ , some  $\text{Ln}^{3+}$  are spectroscopically informative. The second objective is to elucidate how the U2-U6 snRNA complex changes its

conformations upon binding metal ions in order to facilitate the splicing reaction. Förster Resonance Energy Transfer (FRET) and Nano-Surface Energy Transfer (NSET) are utilized to test for the possible conformational changes within the complex.

## 1.2 Nucleic acids

The basic units of nucleic acids (DNA and RNA) are nucleotides. There are four nucleobases in both DNA and RNA. In DNA, the four bases are adenine (A), guanine (G), thymine (T) and cytosine (C). In RNA, uracil (U) is present instead of thymine (T) (Fig 1.1A). A nucleobase and a ribose (in RNA) or deoxyribose (in DNA) sugar ring form a nucleoside. Nucleosides A, G, U, and C in RNA are called adenosine, guanosine, uridine and cytidine; nucleosides A, G, T, and C in DNA are called deoxyadenosine, deoxyguanosine, deoxythymidine and deoxycytidine (Fig 1.1B). A nucleoside and one or more phosphate groups form a nucleotide. Based on the number of phosphate groups, they are called *name of nucleoside + mono/di/tri* phosphate. For example, deoxyadenosine with one, two or three phosphate groups in DNA are called deoxyadenosine monophosphate (dAMP), deoxyadenosine diphosphate (dADP) and deoxyadenosine triphosphate (dATP) respectively; adenosine with one, two or three phosphate groups in RNA is called adenosine monophosphate (AMP), adenosine diphosphate (ADP) and adenosine triphosphate (ATP). Nucleotides are linked with phosphodiester bonds between 5'-OH of the phosphates on the previous nucleotide and 3'-OH of the sugars on the next nucleotide to form a linear polymer (Fig 1.1C).

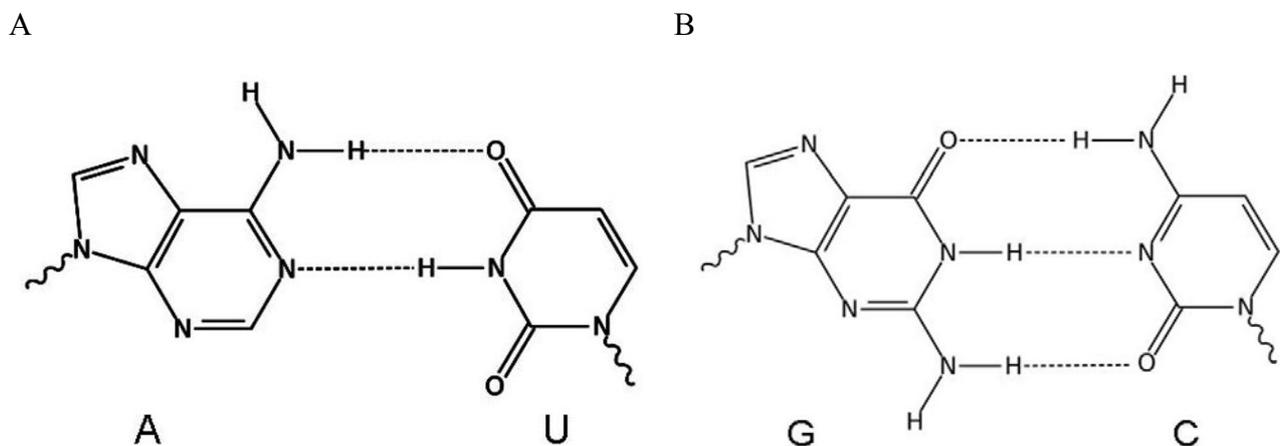
Following the discovery of DNA B-type double helix <sup>1</sup>, additional helical structures have been discovered (table 1.1). Complementary double helix originates from extensive hydrogen bonding formed by nucleobases (called Watson-Crick bases pairs). Regular Watson-Crick base pairs in RNA are shown in Fig 1.2. There are two hydrogen bonds between A and U and three between G and C. The 2'-OH of backbone sugar, which is absent in DNA, has a profound impact on RNA structure and function. The 2'-OH is prone to hydrolysis (water mediated degradation), thus RNA is less stable than DNA. Structurally, the presence of the 2'-OH favors RNA double helix to be A-type in its natural form. DNA, on the other hand, is B-type helix. RNA is predominately present in short and single-stranded forms and DNA in long and double-stranded forms. The 2'-OH also provides RNA additional capacities for hydrogen bonding, metal ion interaction, and catalysis. Various non-Watson-Crick base pairs are also common in RNAs.



**Figure 1.1 Chemical compositions of DNA and RNA.** **A)** The five standard nucleobases in nucleic acids. A, T, G, and C are in DNA and A, U, G, and C in RNA. **B)** Nucleoside contains a ribose in RNA while a deoxyribose in DNA. **C)** Nucleotides linked with phosphodiester bonds between 5'-OH of the phosphates and 3'-OH of the sugars form a nucleic acid polymer. Circled OH groups are available for further polymerization; 2'-OH groups on ribose in RNA which are missing in DNA are highlighted.

**Table 1.1 Structural properties of nucleic acid helices.** Three types of helices (A-, B and Z) are formed depending on their chemical compositions and buffer conditions <sup>2</sup>.

Helix Type	Complementary Strands	Diameter	Helix pitch (rise per turn)	Base pairs per helical turn	Helix rise per base pair
A	RNA-RNA DNA-RNA	~26 Å	28 Å	11	2.6 Å
B	DNA-DNA	~20 Å	34 Å	10	3.4 Å
Z	DNA-DNA	~18 Å	45 Å	12 (6 dimers)	3.7 Å



**Figure 1.2 Regular Watson-Crick base pairs in RNA. A)** A:U base pair: there are two hydrogen bonds between adenine and uracil. **B)** G:C base pair: there are three hydrogen bonds between guanine and cytosine. Dotted lines indicate the hydrogen bonds. Tails indicate the linkage to backbone ribose. Pictures were drawn in ChemDraw.

Because of the unique properties of RNA, it plays far more roles in various biological processes compared to DNA whose predominate role is to store genetic information.

### **1.3 Versatile roles of RNA in biological processes**

The chemical stability and the ability of readily forming double stranded helix make DNA a good storage material for genetic information<sup>3</sup>. In contrast, RNA, which is known to carry genetic information in only some viruses, such as Human Immunodeficiency Virus (HIV) and Hepatitis B Virus (HBV), plays a more versatile role in many biological processes. RNA is also present in nature in more complicated forms.

**Messenger RNA (mRNA):** the functions of most RNA molecules were first not realized, as the only function known by then was being a messenger between the DNA genetic information and their protein products. These RNAs are called mRNA<sup>4,5</sup>. Transcription of DNA to mRNA by RNA polymerase is the first step in gene expression. Prokaryotic mRNA is mature upon transcription except in rare cases. On the other hand, eukaryotic pre-mRNA requires extensive processing, including 5' cap addition, 3' polyadenylation, splicing and editing. In eukaryotic cells, mRNAs are processed in the nucleus and later transported into the cytoplasm; mature mRNAs are translated into amino acid chains in ribosome.

**Transfer RNA (tRNA):** translation is the process that one specific amino acid is added at a time into protein according to the sequence of mRNA in ribosome. Each amino acid is carried by a specific tRNA. tRNA is a single stranded RNA with less than 100 bases and it has a defined tertiary structure. Translation accuracy is ensured by the specific tRNA which can only carry one type of amino acid and transfer it to the right codon sequence. The tRNA specificity is determined by its tertiary structure, which contains a three nucleotide anticodon region complementing the coding region of mRNA and an acceptor stem attaching only one specific type of amino acid.

**Ribosomal RNA (rRNA):** the ribosome is the protein manufacturing machinery in all cells<sup>6</sup>. Ribosome is a RNA-protein complex. Its RNA component is called rRNA. In a typical cell, rRNA makes about 65% of ribosome mass and 80% of all RNA mass. It is conserved phylogenetically. It is rRNA which translates mRNA to amino acid chains in ribosome; proteins only stabilize RNA structure. Thus, ribosome is now classified as a ribozyme<sup>7</sup>.

**Small nuclear RNA (snRNA):** a class of small RNA molecules found in the nucleus of

eukaryotic cells is called snRNA. They are commonly associated with specific proteins to form a complex called Ribonucleoprotein (RNP). snRNAs participate in RNA splicing, which is the subject of this dissertation and will be discussed further. In addition to RNA splicing, snRNAs are involved in many other important processes such as regulating transcription factors or RNA polymerase II and maintaining the telomeres <sup>8</sup>.

Small nucleolar RNAs (snoRNAs) found in the nucleus are belong to a subgroup of snRNA under the definition of Medical Subject Headings (MeSH). They play an essential role in RNA biogenesis and guiding chemical modifications of other RNAs, such as rRNA, tRNA and snRNA <sup>9</sup>.

**Catalytic RNA:** a biological enzyme was first defined as a “catalytic protein”. Carl Woese, Francis Crick and Leslie Orgel speculated about the possible catalytic activity of RNA because they had complex secondary structures <sup>10</sup>. In the early 1980s, the first two RNA enzymes (ribozymes) were discovered independently: group I intron by Thomas R. Cech <sup>11</sup> and RNase P by Sidney Altman <sup>12</sup>. After these findings, many ribozymes were discovered such as rRNA as described earlier and a series of small ribozymes (such as hammerhead, hairpin, HDV and VS ribozymes, they all cut a RNA phosphodiester backbone to yield a 5'-OH and a 2', 3' cyclic phosphodiester as product) <sup>13</sup>. The fact that RNA molecules can perform all the catalytic functions of proteins leads to the hypothesis of “RNA world” theory. In this theory, RNA acts as both the genetic carrier and a catalyst before the emergence of DNA and proteins which possess more specialized functions <sup>14</sup>.

**Other non-coding RNAs:** there are many RNA molecules which do not code for proteins in cells. Many of them were first classified as “junk genes” but we now know they have various functions as we described earlier. There are still many RNAs without known functions. This area of research draws tremendous interest and we expect more RNA functions to be discovered. One example of a success story in this area is the discovery of MicroRNAs (miRNAs). miRNAs were discovered in 1993 <sup>15</sup> but their RNAi function was not revealed until 1998 <sup>16</sup>. RNAi is a mechanism for RNA-guided regulation during gene expression. RNAi inhibits gene expression by causing the degradation of specific RNA molecules or by hindering the transcription of specific genes. Exogenous dsRNA and miRNA can both initiate RNAi.

#### **1.4 Pre-mRNA splicing and the spliceosome**

The mRNA molecule directly copied from the DNA template by transcription is called

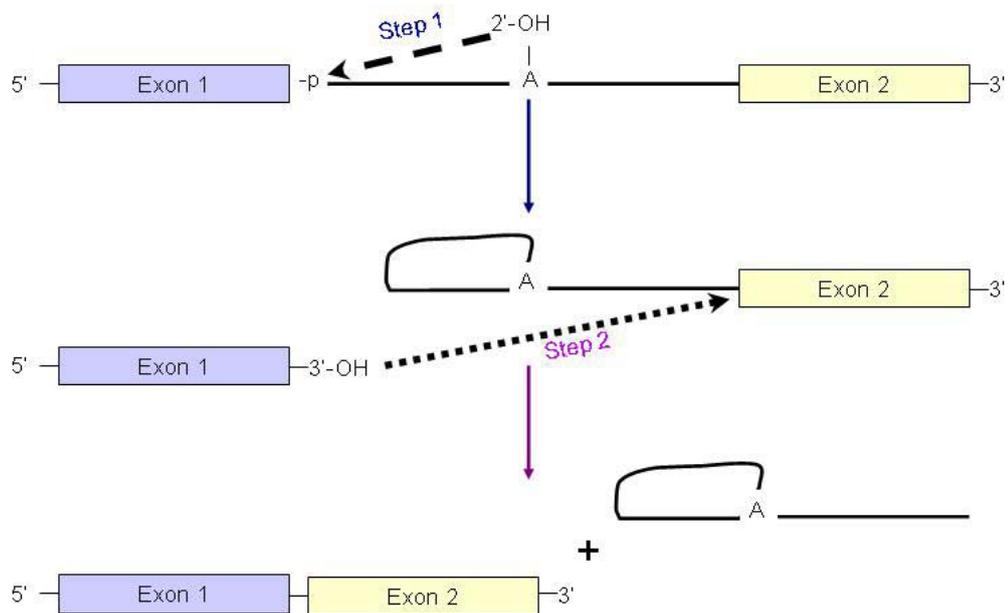
precursor (pre-)mRNA. As discussed, prokaryotic mRNA is usually mature upon transcription. In eukaryotic cells, however, pre-mRNA will usually go through multiple processing steps; one of which is splicing. Thus, splicing is an integral step of gene expression in eukaryotic cells. During splicing, the non-coding regions on pre-mRNA (called introns) are removed and coding regions (called exons) are ligated together. Splicing at alternative sites provides a mechanism to generate alternative gene products from the same gene sequences for gender-, cell-, and development stage-specific purposes. Splicing defects are linked to certain diseases such as heart diseases and cancers .

There are different mechanisms for splicing reactions. Eukaryotic RNA splicing is catalyzed by the spliceosome in the nucleus. By comparison, self-splicing group II intron found in certain protozoa and eukaryotic organelles catalyzes the same splicing reaction on its own. The chemical reaction for both systems is identical, a two-step trans-esterification reaction (Fig 1.3). At the first step, the 2'-OH of a helical bulged adenosine (called branch site) attacks the 5'-splice site (5'-ss). The phosphate group will form a 5', 2'- cyclic phosphodiester bond with branch site 2'-OH; 3'-OH on 5'-exon is released. The second step is that the newly freed 3'-OH on 5'-exon is attacking 3'-splice site (3'-ss). Two exons are joined together and the intron is released as a lariat.

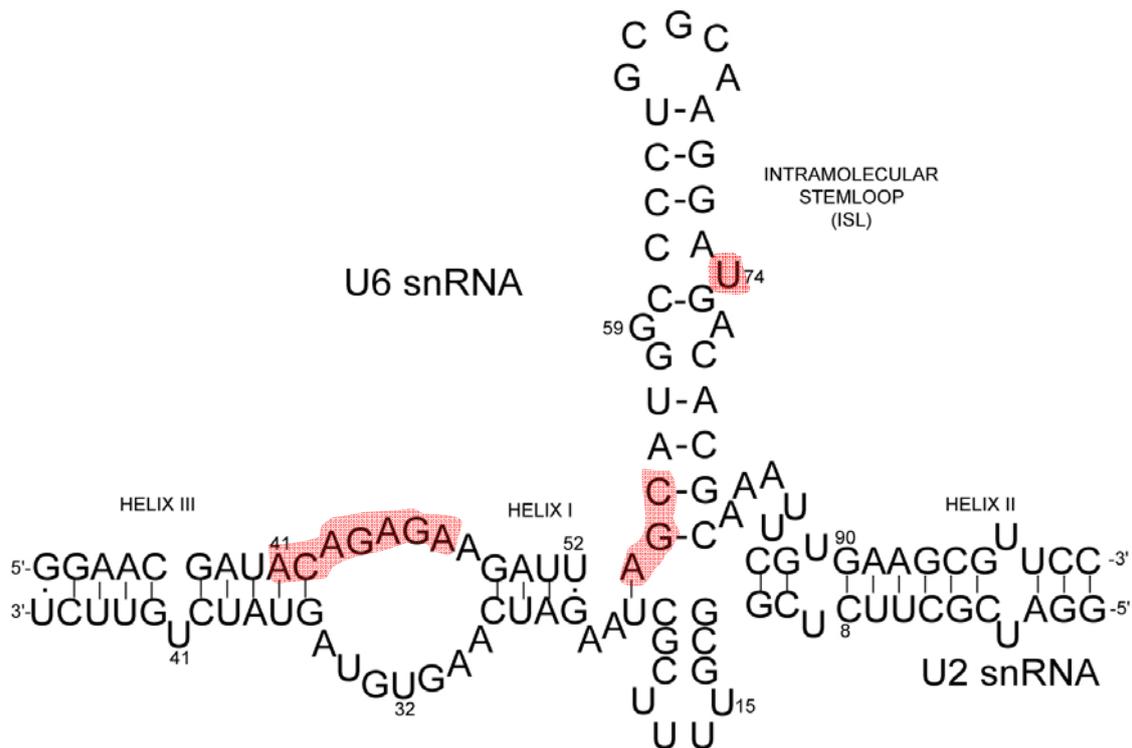
The spliceosome is a dynamic assembly of five small nuclear snRNAs (U1, U2, U4, U5 and U6) and over 100 proteins; it is the largest and one of the most complicated cellular machines known<sup>17</sup>. RNA-RNA, RNA-protein and protein-protein interactions form an extensive network within the spliceosome<sup>18</sup>. Specific inter- and intra-molecular interactions involving snRNAs and pre-mRNA are required for the assembly of the catalytic core of the spliceosome and the definition of the splice sites.

### **1.5 Splicing-related catalysis by protein-free spliceosomal snRNAs**

In spliceosome, U2 and U6 snRNAs are the only snRNAs out of the five snRNAs required for both steps of splicing. Their sequences are highly conserved phylogenetically<sup>19</sup>. The two snRNAs have extensive intra- and inter-molecular base pairs and form a complex (Fig 1.4). The U6 snRNA contains an intramolecular stem loop (U6 ISL); the U2 and U6 snRNAs form three intermolecular helices, called helix I, helix II and helix III<sup>20, 21</sup>. These regions are conserved and are important for splicing activity, but sequence conservation is not strict. Conversely, several sequences are found to be strictly conserved. They are ACAGAGA loop located between helices I



**Figure 1.3 Splicing is a two-step trans-esterification reaction.** The first step of splicing is the 2'-OH group of branch point adenosine attacks the 5'-splice site. 3'-OH on 5'-ss and exon 1 are freed. The second step of splicing is that the 3'-OH on exon 1 attacks the 3'-splice site. A phosphodiester bond is formed between exons 1 and 2 and intron is released as a lariat.



**Figure 1.4 Proposed human U2 and U6 snRNA complex secondary structure**<sup>22</sup>. The two snRNAs form extensive base pairs with several invariant secondary structures. These invariant structures include the intramolecular stem loop (ISL) and helices I, II, and III. Conserved sequences are highlighted; they are the ACAGAGA loop between helices I and III, the AGC triad near the four way junction, and U74 on the U6 ISL. They are ion-binding sites implicated by phosphorothioate substitution experiments.

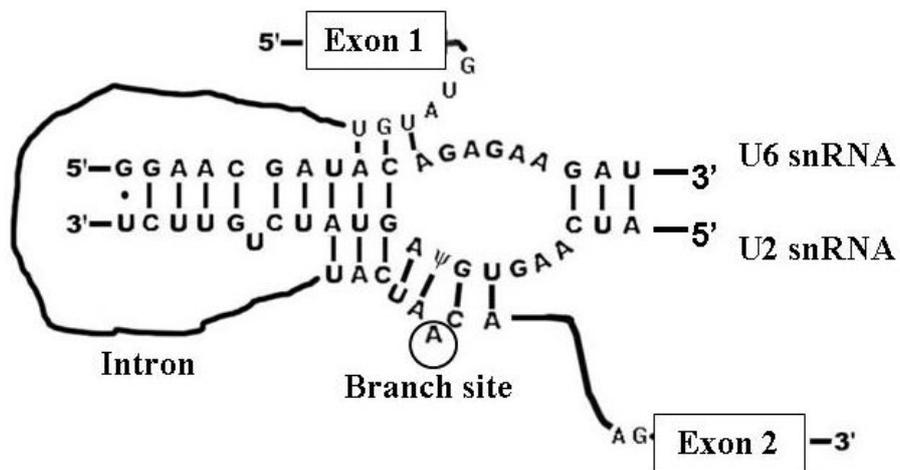
and II; AGC triad within the hinge of the four-way junction, and the U74 (U80 in yeast, U74 in human) within the U6 ISL labeled in Fig 1.4<sup>23-26</sup>.

The first step of splicing involves a bulged branch site adenosine nucleophilic- attack the 5'-ss of the intron (Fig 1.3). These two reactants both have interactions with the U2-U6 snRNA complex. The branch site region of the intron forms a defined duplex with 4-7 base pairs. A conserved modified base pseudouridine on U2 snRNA help position the branch site adenosine so that the 2'-OH group of the adenosine is more accessible for attacking the 5'-ss<sup>27,28</sup>. The 5'- splice site of the intron interacts with the U6 ACAGAGA loop; but they only form small number of base pairs. This indicates that in order to bind and position the 5'-ss, the U6 and U2 snRNA complex may need help from other RNA/protein components. These interactions are summarized in Fig 1.5.

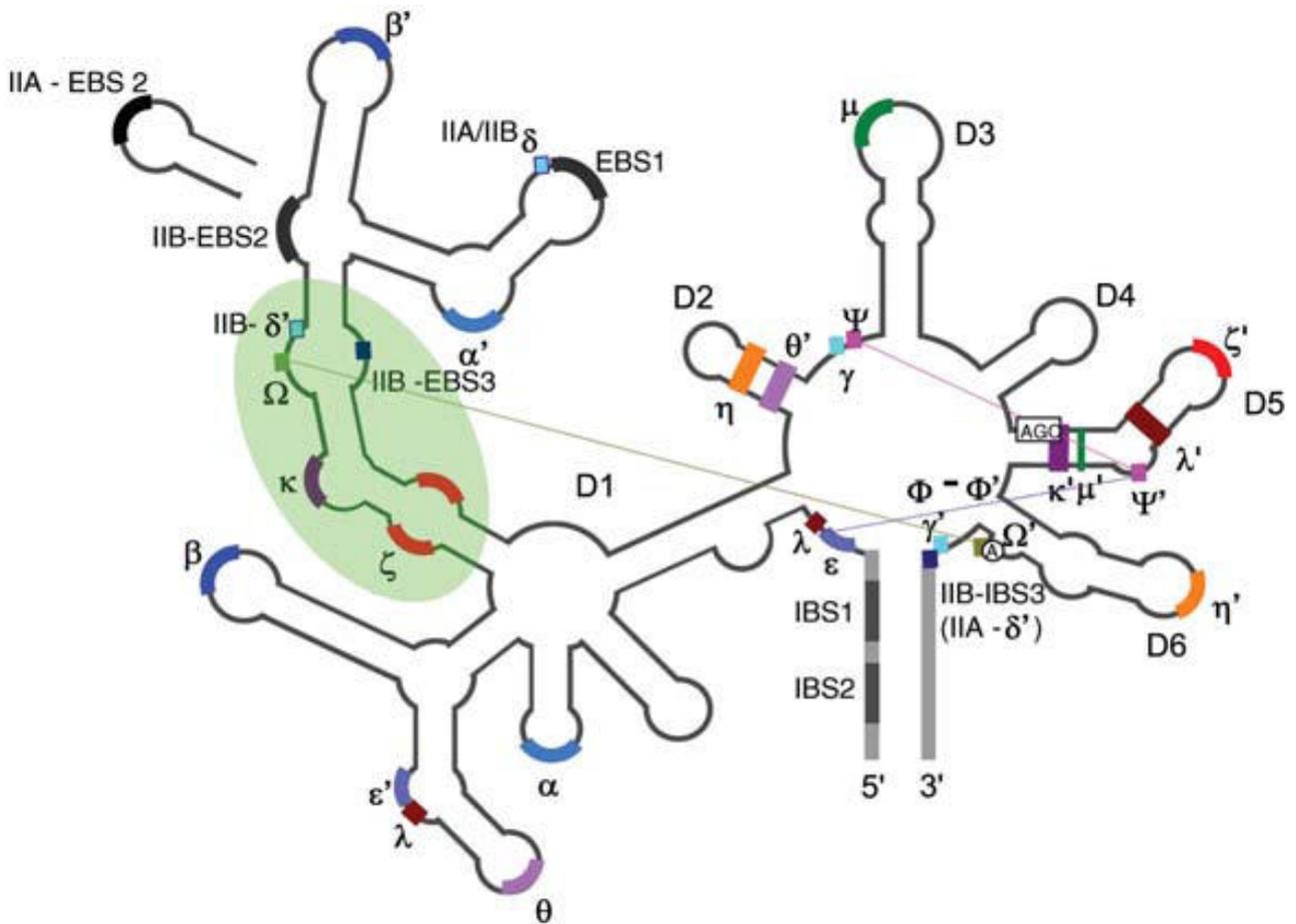
Very interestingly, the self-splicing group II introns and the spliceosomal U2-U6 snRNA complex have extensive mechanistic and structural similarities. Group II intron was one of the first ribozymes discovered<sup>29</sup>. The splicing reaction catalyzed by group II intron does not involve any proteins *in vitro* and has identical stereochemistry as reactions catalyzed by spliceosome. Both group II intron and U2-U6 snRNA complex bind to a catalytically essential metal ion by an AGC trinucleotide<sup>23,30</sup>. Both have an unpaired flanking branch site residue; usually adenosine. The U6 ISL has a similar structure as the D5 of the group II intron. The U2-branch site may also have a similar structure as D6 of the group II intron (Fig 1.6).

A report by Valadkhan and Manley in 2001 showed that a protein-free U2-U6 snRNA complex could perform a reaction similar to the first step of splicing further exemplifies that they were involved in the direct catalytic reactions<sup>22</sup>. In this report, a small RNA intron strand containing the branch site was covalently linked with the AGC triad after incubating the intron and U2-U6 snRNA complex together (Fig 1.7A). The reaction was only similar to the first step of splicing; and the yield of 0.1 % after 20 hours indicated an inefficient and slow catalysis.

A more recent report showed that designing a system which contained an intron with the branch site and 5'-ss resulted a reaction more similar to the first step of splicing<sup>31</sup>. The intron 5'-ss did not bind stably with the ACAGAGA loop. So a 5'-ss sequence was covalently linked to U6 snRNA sequence. When the RNAs formed a secondary structure, 5'-ss was stably attached to the ACAGAGA loop. After the reaction, a covalent bond between 5'-ss and branch site adenosine was formed and the exon was released (Fig 1.7B). Even though the yield of 0.2 % after 15 hours was



**Figure 1.5 Interactions between U2-U6 snRNA complex and the intron.** The 5'-ss of the intron forms small number of base pairs with the complex. The branch site forms extensive base pairs with the complex. The branch site adenosine is bulged out by its opposite pseudouridine base ( $\psi$ ) on U2 snRNA as shown in a NMR structure by Newby and Greenbaum <sup>27</sup>.



**Figure 1.6 Secondary structure of a group II intron.** The figure is from Fedorova and Zingler 2007 <sup>29</sup>. D1-D6 are abbreviated for domain 1, domain 2 etc. The circled adenosine on D6 is the branch site; the boxed AGC on D5 is the conserved AGC triad. D5 has similar structure and function as the U6 ISL in spliceosome; D6 has similar structure and function as the U2-intron branch site helix. For other labels please refer to the original reference.



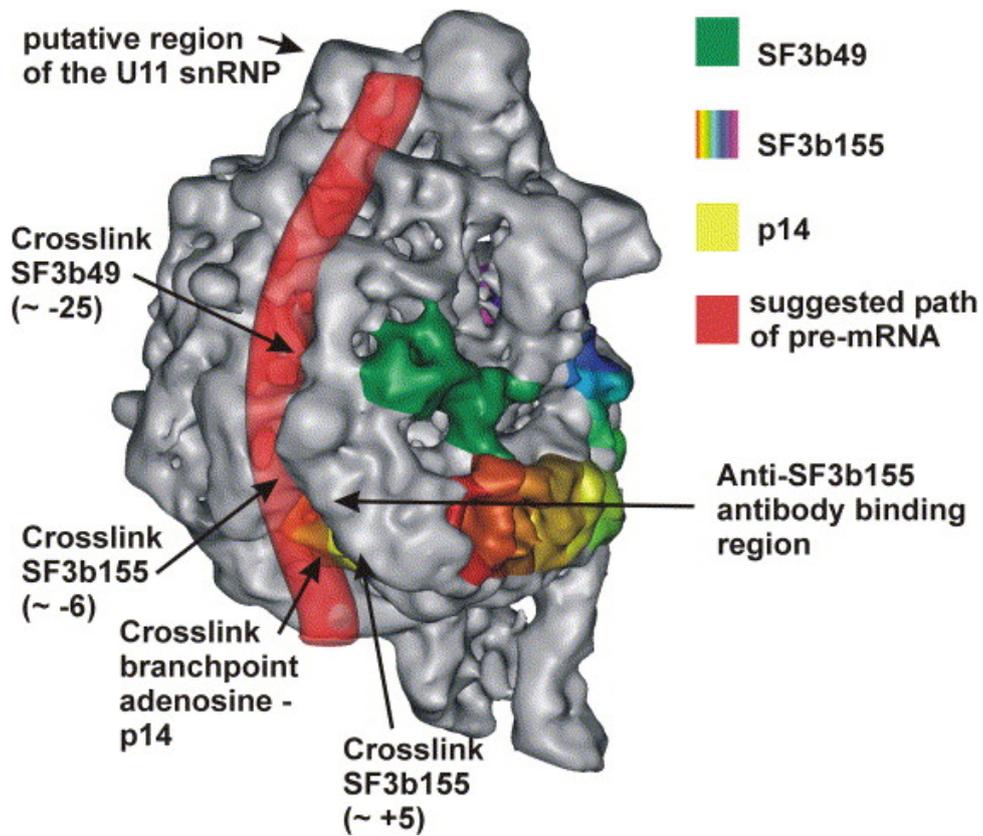
improved, it still indicated a very inefficient and slow reaction.

Based on these results discussed above, the U2-U6 snRNA complex is thought to be at the center of the catalytic core of the spliceosome. Even though the spliceosome has about 97% molecular mass as proteins, spliceosome may be a ribozyme<sup>17</sup>. However, the extremely low yield and efficiency compared to protein-assisted splicing, and the abundance of protein components, lead us to understand that even if spliceosome is a ribozyme; protein components may still play important roles in the spliceosome activity. One example is that two protein components in U2 snRNP, SF3b155 and p14, are in close proximity of the pre-mRNA branch site as shown in Fig 1.8<sup>32</sup>. The positioning may play an important role in the assembly of the U2 snRNP. Thus, the U2-U6 snRNA complex not only juxtaposes 5'-ss and the branch site, but also may provide a scaffold for the assembly of the spliceosome.

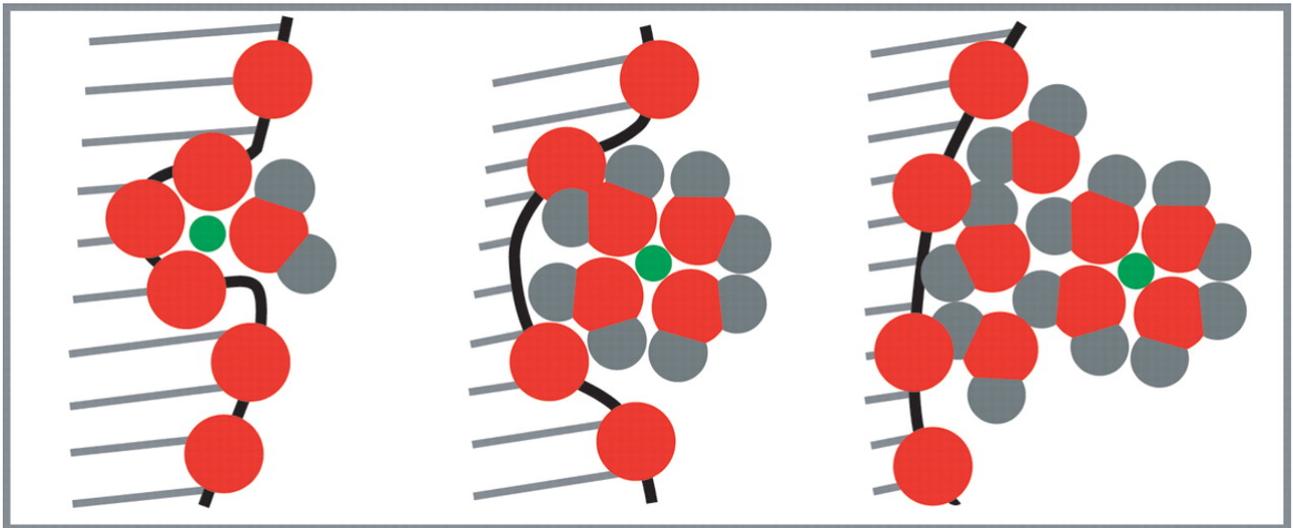
### **1.6 Metal ions and U2-U6 snRNA complex**

Metal ions have a great impact on the folded RNA tertiary structure due to the electrostatic interactions with the RNA backbone and other polar chemical groups. There are two types of RNA-ion interactions as summarized<sup>33</sup>: “diffuse” ions which interact with RNA electrostatic field at a hydrated state, and “chelated” ions which interact directly with specific RNA surface locations and are held in position by electrostatic forces at a dehydrated state (Fig 1.9). Both of these interactions are crucial to the stability of RNA structures. Metal ions may also participate in direct catalysis by activating nucleophilic oxygens during oxygen–phosphorus bond formation or stabilize oxyanion leaving groups during oxygen–phosphorus bond breakage<sup>34</sup>.

Spliceosome assembly and activity are dependent on “hard” metal ions interactions. “Hard” metal ions, which are characterized with small ionic radii or high charge, prefer “hard” ligands, which do not give up electron density. They are held together by electrostatics. On the other side, “soft” metal ions prefer “soft” ligands. They possess polarizable electron clouds and are prone to share electron density. Specifically, the splicing catalysis by protein-free U2-U6 snRNA complex is dependent on metal ions, generally  $Mg^{2+}$ , a “hard” ion. The phosphorothioate substitution for phosphate experiments suggested several metal ion binding sites essential for splicing catalysis. The splicing activity was obliterated upon the substitutions (oxygen atom, a “hard” ligand, to sulfur atom, a “soft” ligand) but rescued by using “soft” metal ion  $Mn^{2+}$ . These sites include U74 residue



**Figure 1.8 Cryo-EM structure of pre-mRNA, SF3b155 and p14.** The proteins p14 and SF3b155 are in proximity of the branch site adenosine as shown in the Cryo-EM structure as well as by cross-linking experiments. The figure is from Golas *et al.*, 2003<sup>32</sup>.



**Figure 1.9 Cartoon of “diffuse” ions and “chelated” ions.** The figure is from Draper *et al.*<sup>35</sup>. Small green dots are  $\text{Mg}^{2+}$ ; red dots are oxygens; grey dots are hydrogens; two hydrogens and one oxygen form a water molecule; red dots on RNA backbones are nonbridging phosphate oxygens. **Left)** a “chelated” ion loses some inner-sphere water molecules and stays in a specific location on RNA backbone. **Right)** a “diffuse” ion retains its outer shell and interacts with RNA backbone by electrostatic forces mediated by outer-shell water molecule. **Middle)** inner-sphere water molecules of a fully hydrated ion interact directly with RNA backbone.

on ISL, the A residue in the AGC triad and the two G residues in the ACAGAGA loop (Fig 1.4)<sup>23, 25, 26, 36</sup>. UV cross-linking experiments also showed that these regions had metal ion dependent interactions<sup>37</sup>.

The binding to the U74 of ISL is of particular interest to us because its pH-dependency. The metal ion binding is supported by phosphorothioate substitution results<sup>36</sup>, but the overall structure of U6 ISL is not altered upon metal ion binding<sup>38</sup>. The metal ion binding on this site is also pH-dependent. Protonation of the neighboring C-A base pair and binding of metal ions were found to be mutually antagonistic. At higher pH, the C-A pair was deprotonated; U80 was stacked in the helix and formed an ion binding site<sup>38</sup>.

## 1.7 Specific aims

Understanding how the spliceosome catalyzes the splicing reaction will help us gain insights in this important step of gene expression with possible medical and engineering applications. Interactions between the U2-U6 snRNA complex and metal ions are crucial for the splicing catalysis. Thus, the study of how they interact and what structural impact metal ions have on the U2-U6 snRNA complex is the focus of this dissertation.

The goal of chapter two is to find out the locations of the “chelated” metal ions in the U2-U6 snRNA complex. The phosphorothioate substitution experiments suggested possible specific metal ion binding sites on U2-U6 snRNA complex; but direct observations of bound metal ions were lacking. The biological metal ion  $Mg^{2+}$  is spectroscopically uninformative, thus, a lanthanide ion  $Tb^{3+}$  was used to mimic  $Mg^{2+}$  by taking advantage of its luminescence and  $Mg^{2+}$ -like chemical and physical properties. A new approach which utilized RNA bound  $Tb^{3+}$  as a FRET donor was taken; specific metal ion binding sites were successfully mapped out spectroscopically.

The success in chapter two encouraged me to evaluate whether this new type of FRET donors could be used in different conditions. This type of FRET donors, unlike previous FRET donors which have to be covalently linked to target molecules, do not require covalent coupling. For example, RNA bound  $Tb^{3+}$  could act as FRET donors because its luminescence was enhanced by neighboring RNA. If  $Tb^{3+}$  whose luminescence was enhanced by RNA or protein which they “chelated” bind can be a class of general FRET donors, labeling procedures would be much easier and there could be numerous applications involving these FRET donors.

Thus, in chapter three, the properties of these RNA bound  $Tb^{3+}$  were studied and the optimal conditions in which these donors could work best were investigated. This chapter laid a guidance for the practice of using RNA bound  $Tb^{3+}$  as FRET donors broadly accepted and useful.

In chapter four, the question of what were the structural impacts of metal ions on the U2-U6 snRNA complex was investigated. The hypothesis was that in order for the complex to carry out splicing reactions, key components were needed to come into proximity during the catalytic reaction. By placing FRET and NSET probes on different locations on the U2-U6 snRNA complex, conformational changes within the complex could be monitored. Both steady-state and time-resolved FRET and NSET were used.

The results will help us better understand specific metal ion binding on the U2-U6 snRNA complex. Additionally, the introduction of using RNA bound  $Tb^{3+}$  as FRET donors may broaden the use of FRET methodology in new areas.

## CHAPTER 2

# USE OF A NOVEL FRET METHOD TO LOCALIZE SITE-BOUND TERBIUM IONS IN THE U2–U6 SNRNA COMPLEX

### 2.1 Introduction

#### 2.1.1 FRET method

FRET is commonly used as a ‘molecular ruler’<sup>39</sup> in biomolecular structural studies. Initially, the acronym FRET referred specifically to Fluorescence Resonance Energy Transfer because the donor and acceptor pairs were both fluorophores. In a variant experiment, the luminescence energy of chelated lanthanide ions was transferred directly to an acceptor molecule<sup>40</sup>, in which case the approach was called LRET (Luminescence Resonance Energy Transfer). Lanthanide ion luminescence is formally neither fluorescence (singlet-to-singlet transition) nor phosphorescence (triplet-to-singlet transition), but as a result of atomic configuration where emission arises predominantly from electric dipole transitions<sup>41</sup>. For this reason, both forms of energy transfer obey the same fundamental principles; and it has been suggested that the acronym FRET applies both to the specific (fluorescence) and general (Förster) resonance energy transfers<sup>42</sup>.

FRET originates from non-radiative interaction between a donor and an acceptor exhibiting a spectral overlap between donor emission and acceptor excitation spectra. The relationship between the energy transfer efficiency (E) and the FRET pair distance (R) is:

$$E = \frac{R_0^6}{R_0^6 + R^6} \quad (\text{Equation 2.1})$$

Where  $R_0$  is the Förster distance between the FRET pair when  $E=50\%$ .  $R_0$  depends on the overlap integral of the donor emission spectrum with the acceptor absorption spectrum, quantum yield of the donor and their relative molecular orientation as expressed by the following equation:

$$R_0 = 0.21 \sqrt[6]{J \cdot q_D \cdot n^{-4} \cdot k^2} \quad (\text{in } \text{\AA} \text{ angstroms}) \quad (\text{Equation 2.2})^{41}$$

Where  $J$  is the normalized spectral overlap of the donor emission and acceptor absorption;  $q_D$  is the quantum yield of the donor;  $n$  is the refraction index (1.33 for water, 1.29 for many organic solvents);  $k^2$  is the orientation factor describing the relative orientation of the transition dipoles of the donor and acceptor.  $k^2$  is often a source of uncertainty, being defined as:

$$k^2 = (\cos \theta_{DA} - 3 \cos \theta_D \cos \theta_A)^2 \quad (\text{Equation 2.3})$$

Where  $\theta_{DA}$  is the angle between the donor and acceptor transition dipole moments;  $\theta_D$  and  $\theta_A$  are the angles between the donor (acceptor) transition dipole moment and the vector joining the two dyes<sup>41</sup>.

In recent years, advancements in photochemistry instrumentation enable FRET to be at the center stage of biomolecular studies with accurate and sensitive detection. Single-molecule FRET<sup>43</sup> and FRET microscopy<sup>44</sup> can reveal information at the level of single molecules and living cells.

Many FRET donors and acceptors have been discovered and developed<sup>45</sup>. They can be classified into four categories: 1) Organic fluorophores, which are organic small molecules (generally < 1kDa); 2) Fluorescent proteins, such as GFP and its derivatives; 3) Luminescent lanthanide ions, such as  $\text{Tb}^{3+}$  and  $\text{Eu}^{3+}$ ; and 4) Quantum dots. Among them, organic fluorophores are the most common because they are readily available with high absorption coefficients and quantum yields. Other categories are routinely used as well with a lot of new developments.

Lanthanide ions have unique luminescence spectra and  $\text{Tb}^{3+}$  and  $\text{Eu}^{3+}$  have been used as FRET donors in several applications<sup>41</sup>. In these applications,  $\text{Tb}^{3+}$  or  $\text{Eu}^{3+}$  are chelated with EDTA-like molecules which are covalently linked to target molecules.

The energy transfer efficiencies ( $E$ ) can be calculated by several methods:<sup>46</sup> 1) using donor

fluorescence intensity decrease,  $E = 1 - \frac{I_{DA}}{I_D}$  (Equation 2.4), where  $I_D$  and  $I_{DA}$  are donor

intensities in the absence and presence of acceptor respectively. 2) using acceptor fluorescence

intensity increase,  $E = \frac{I_{AD}}{I_{AD,100\%}}$  (Equation 2.5), where  $I_{AD}$  is the acceptor intensity in the presence

of donor;  $I_{AD,100\%}$  is the acceptor intensity in the presence of donor when energy transfer efficiency

is 100%. 3) using lifetime decrease of donor fluorescence,  $E = 1 - \frac{\tau_{DA}}{\tau_D}$  (**Equation 2.6**),  $\tau_{DA}$  and  $\tau_D$  are the lifetimes of donor in the presence and absence of acceptor respectively.

Using lifetimes to measure energy transfer efficiency is most accurate because it is derived directly from the energy transfer mechanism. Time-resolved FRET can also differentiate more than one species because of the exponential decay nature of lifetime data. Thus, multiple-distance resolution can be achieved at one time.

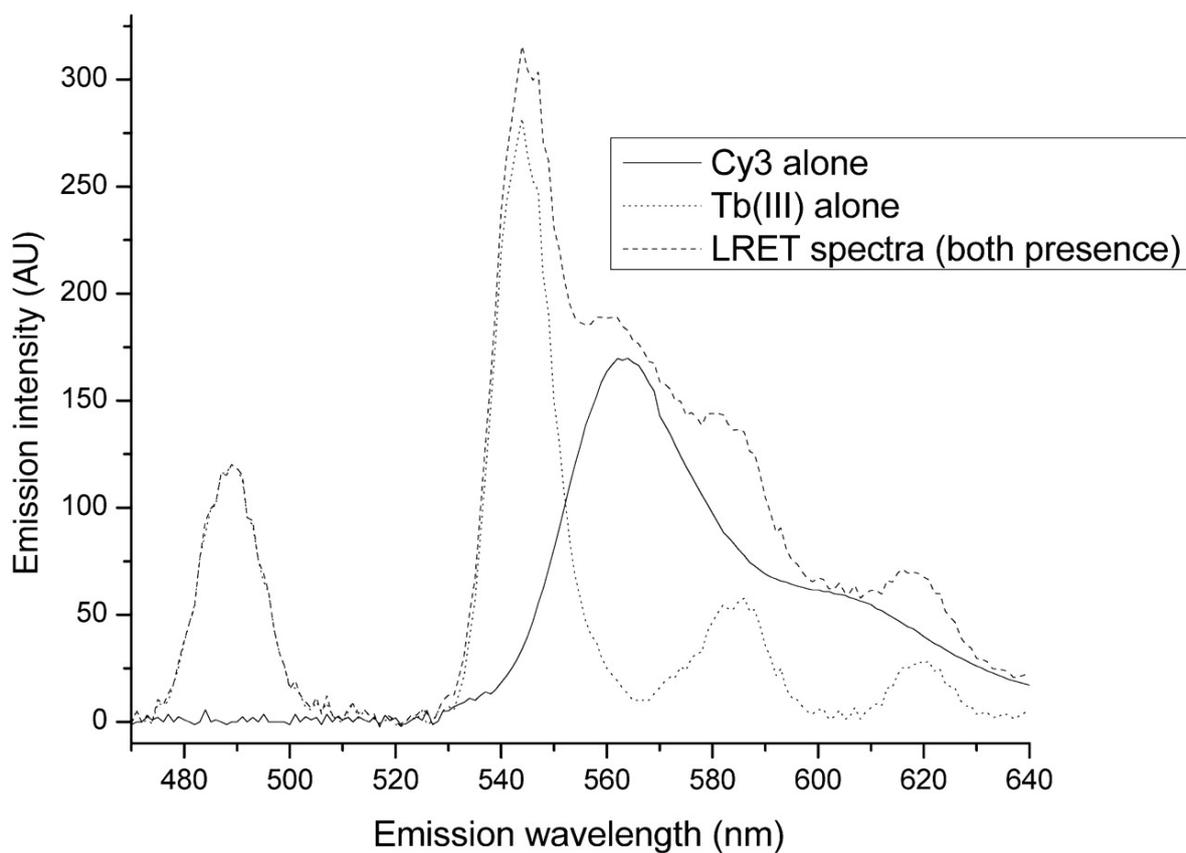
### 2.1.2 An extension of FRET method

Lanthanide ions have been used as FRET donors. They are chelated to an EDTA-like molecule which contains an antenna group and are covalently linked to target molecules<sup>41</sup>. The spectra of a FRET system with chelated  $Tb^{3+}$  as donor and Cy3 as acceptor are shown in Fig 2.1. Free  $Ln^{3+}$  ions are not practical FRET donors because of their low absorption coefficients ( $\epsilon \sim 1 M^{-1}cm^{-1}$  compared to  $\epsilon \sim 10^4 M^{-1}cm^{-1}$  for nucleotide)<sup>41</sup>.

$Tb^{3+}$  luminescence is highly enhanced by energy transfer from RNA aromatic groups. However, the luminescence of up to 1mM  $Tb^{3+}$  in aqueous solution was undetectable on a Cary Eclipse spectrophotometer while the luminescence of less than 1  $\mu M$  RNA bound  $Tb^{3+}$  was detectable under the same condition. Luminescence was still not detectable after adding EDTA to  $Tb^{3+}$  solution or dissolve  $Tb^{3+}$  in  $D_2O$ , suggesting the luminescence enhancement was mainly from RNA sensitizing, not from the removal of inner sphere water molecules. This enhancement is similar as the effects of the antenna groups on organic chelate LRET groups<sup>41</sup>. The RNA “chelated”  $Tb^{3+}$  may be used as FRET donors. The new aspect of the method is no covalent chemical chelate modifications are done on target molecules.

### 2.1.3 Specific metal ion binding on the U2-U6 snRNA complex

Metal ions play a significant role in RNA folding and ribozyme catalysis<sup>47</sup>. Bound metal ions assist in folding and stabilization of RNA stem-loop structures by increasing the number of favorable intra- and inter-strand interactions. Divalent metal ions, particularly  $Mg^{2+}$ , are essential for spliceosomal assembly and activity. Several lines of evidence support the obligatory role of divalent metal ions in the activity of the U2-U6 snRNA complex. Replacement of phosphate groups with phosphorothioate groups, which do not strongly bind hard metal ions such as  $Mg^{2+}$ , at



**Figure 2.1 A FRET system with  $Tb^{3+}$  as donor and Cy3 as acceptor.** The LRET spectrum is de-convoluted to  $Tb^{3+}$  spectrum and Cy3 spectrum.  $Tb^{3+}$  spectrum is obtained by normalized the peak at 490 nm with a standard  $Tb^{3+}$  spectrum (Fig 2.1). Cy3 spectrum is the difference between LRET and normalized  $Tb^{3+}$  spectra. This figure is from Yuan *et al.*, 2007<sup>48</sup>.

specific sites on the U6 snRNA backbone resulted in the obliteration of splicing activity. These sites include the backbone at U74 of U6 snRNA, which forms part of an internal loop within the intramolecular stem loop (ISL)<sup>26</sup>, the two G residues in the invariant ACAGAGA segment between helices I and III<sup>25</sup>, and the A residue of the invariant AGC triad near the four-way-junction<sup>23</sup>. Some of the splicing activity was rescued by Mn<sup>2+</sup>, a softer metal ion, which coordinates more readily with sulfur atoms. Moreover, UV-induced cross-linking experiments demonstrated metal ion dependent interaction between these regions<sup>37</sup>. Protein-free RNA splicing activity was also shown to be dependent upon Mg<sup>2+</sup><sup>22,31</sup>.

Evidence for pH-dependent metal ion binding of a bulged uridine of the U6 ISL (U80 in the yeast sequence, which corresponds to U74 in the human sequence) comes from NMR studies<sup>38</sup>. Protonation of the neighboring C-A base pair and binding of metal ions were found to be mutually antagonistic. At higher pH, the C-A pair was deprotonated; U80 was stacked in the helix and formed an ion binding site. The possibility of a pH-dependent mechanism for the regulation of spliceosomal activity was suggested<sup>38</sup>. Although it is possible that other sites exhibit pH-dependent ion coordination behavior, their pH-dependence has not yet been examined.

#### **2.1.4 Goal of this chapter**

The evidence that there are several specific metal ion binding sites on the U2-U6 snRNA complex and pH values could turn them on and off encouraged us to apply the new FRET practice discussed in section 2.1.2 for identifying the locations of these site-bound multivalent ions and the pH dependence of the binding. Thus, the goal of this chapter is to identify the locations of site-bound multivalent ions within the U2-U6 snRNA complex and ascertain the pH dependence of ion binding behavior by using “chelated” Tb<sup>3+</sup> as a FRET donor. Time-resolved FRET was used between the site-bound luminescent lanthanide ions and a covalently attached organic dye. The novelty of this study was that the FRET donors were labeled without the use of chelates because they were lanthanide ions which were coordinated to the specific sites in the RNA complex.

## **2.2 Materials and methods**

### **2.2.1 RNA design and synthesis**

All experiments were carried out on RNA sequences of human U2-U6 snRNA (Fig 1.4)<sup>22</sup>. For measurements with single 5'-Cy3 and double 5'-Cy3, 3'-Cy5 labels, the top strand was

purchased from Dharmacon (Lafayette, CO). Otherwise, unlabeled RNA strands were made by *in vitro* transcription under standard conditions using T7 RNA polymerase with synthesized DNA templates from IDTDNA (Coralville, IA). DNA templates were designed with a standard T7 promoter sequence. T7 RNA polymerase for transcription reactions was expressed and isolated according to a published protocol <sup>49</sup>. Transcripts were purified by electrophoresis on a denaturing polyacrylamide gel and eluted with an electro-eluter (Schleicher and Schuell, Whatman Inc.). RNA concentrations were determined from the absorbance at 260 nm, and the integrity of the final RNA samples was confirmed by the appearance of a single band on a 15% denaturing PAGE gel.

### 2.2.2 Nondenaturing gel electrophoresis

RNA samples were loaded on a 10% non-denaturing gel in Tris-HEPES buffer. 0.01 nmol (~1 µg) each of top and bottom strand were combined, heated at 80 °C for 45 seconds and cooled at room temperature for 10 minutes prior to loading to the gel. The gel was run in a 4 °C cold room under 100V. The gel was stained with Ethidium Bromide for UV visualization.

### 2.2.3 Measurement of Tb<sup>3+</sup> luminescence

Tb<sup>3+</sup> stock solution was made by dissolving Tb<sub>2</sub>O<sub>3</sub> in HClO<sub>4</sub> (both purchased from Sigma Aldridge). Concentration was measured by titration with EDTA as described previously <sup>50</sup>. The excitation wavelength was 280 nm.

Lanthanide ions can hydrolyze RNA molecules nonspecifically at high concentrations, pH and temperatures <sup>51</sup>. In the study, 7 µM Tb<sup>3+</sup> did not hydrolyze the RNA samples under the described conditions for days (data not shown). Also, high buffer pH could change the binding properties of lanthanide ions by converting them to lanthanide hydroxides or changing hydration status <sup>52</sup>. However, we observed no difference in lifetimes of Tb<sup>3+</sup> bound to RNA at pH 5.6-7.2, suggesting its overall hydration/binding status was not altered significantly.

RNA samples (1µM) were titrated with Tb<sup>3+</sup> until the luminescence did not increase further. The dissociation constant of Tb<sup>3+</sup> binding to U2-U6 snRNA complex was calculated by fitting the plot of the saturation plot with a standard binding isotherm for three equivalent sites yielded a mean K<sub>d</sub> of approximately 3 µM. The moderate binding affinity allows us to use the stoichiometry measurements in the next section with minimal assumptions. Bound Tb<sup>3+</sup> could be competed off the RNA by Mg<sup>2+</sup>, resulting in decreasing luminescence to zero. This showed that Mg<sup>2+</sup> occupied the

same or similar sites as  $Tb^{3+}$ . However,  $Mg^{2+}$  had about 200 times weaker affinity to U2-U6 complex than  $Tb^{3+}$ .

## 2.2.4 Measurement of metal ion binding stoichiometry to RNA constructs

The continuous titration approach of measuring  $Tb^{3+}$ :RNA binding stoichiometry, or Job plot method, was used<sup>53</sup>, employing a repetitive dilution approach for the titration. In a 200  $\mu$ L cuvette, the total [RNA +  $Tb^{3+}$ ] concentration was 2.0  $\mu$ M. For an RNA fraction of 0.9, 180  $\mu$ L of 2  $\mu$ M RNA and 20  $\mu$ L of 2  $\mu$ M  $Tb^{3+}$ , both in buffer, were mixed; for an RNA fraction of  $(0.9)^2$ , 20  $\mu$ L of the previous mixture was removed and 20  $\mu$ L of 2.0  $\mu$ M  $Tb^{3+}$  in buffer was added to the cuvette. The RNA fraction of  $(0.9)^n$  can be achieved using this approach. Measurement of  $Tb^{3+}$  emission was performed on a Cary Eclipse spectrophotometer (Varian Inc.), with excitation at 280nm. The emission intensity at 545 nm was plotted against the corresponding RNA mole fraction and fit to a Gaussian function in SigmaPlot 8.0 (Systat Software, Inc.). The number of ion binding sites is equivalent to  $(1-x)/x$ , where x is the RNA fraction at the peak of the Gaussian curve<sup>50</sup>.

## 2.2.5 Measurements of $Tb^{3+}$ luminescence decay

Lifetime measurements were carried out on a Cary Eclipse fluorophotometer. The RNA concentration was 0.7  $\mu$ M; titrated  $Tb^{3+}$  concentrations were 3.5  $\mu$ M to 7  $\mu$ M. Photomultiplier tube (PMT) power was 800 V if not specified otherwise. Excitation pulses had a width of 2-3  $\mu$ s at 280 nm and excitation/emission slits were 10 nm.  $Tb^{3+}$  luminescence was measured at 565 nm when Cy3-labeled RNA was used. With unlabeled RNA,  $Tb^{3+}$  luminescence at 545nm was recorded. Delay time was 0.1 ms. Buffers contained 150 mM NaCl and 10 mM MES for pH 5.6, 6.1, 6.6 or 10 mM MOPS for pH 7.2. The pH was adjusted by NaOH and HCl. A high concentration of NaCl (150 mM) was used to minimize nonspecific binding of  $Tb^{3+}$ . Decay curves were fit with OriginLab using least-square-goodness-of-fit algorithm to equation 2.7. A maximum of three independent decay components could be accurately obtained.

$$I = I_0 + \sum_i A_i e^{-\frac{t}{\tau_i}} \quad \text{(Equation 2.7)}$$

Where  $I_0$  is system background,  $A_i$  are the preexponential amplitudes for each exponential decay which can be used to identify the population ratio of different decays, and  $\tau_i$  are the individual decay lifetimes.

## 2.2.6 FRET experiments using the Cy3-Cy5 pair

Measurements were performed on the Cary Eclipse. Excitation of Cy3 was at 520 nm with excitation/emission slits of 5 nm and the temperature set at 22 °C. The emission scan was collected from 540-700 nm. Equal molar amount (0.7 μM each) of top (labeled) and bottom strands were annealed prior to measurements. RNA samples labeled with Cy3 only, Cy5 only, and both dyes were measured.

## 2.3 Results

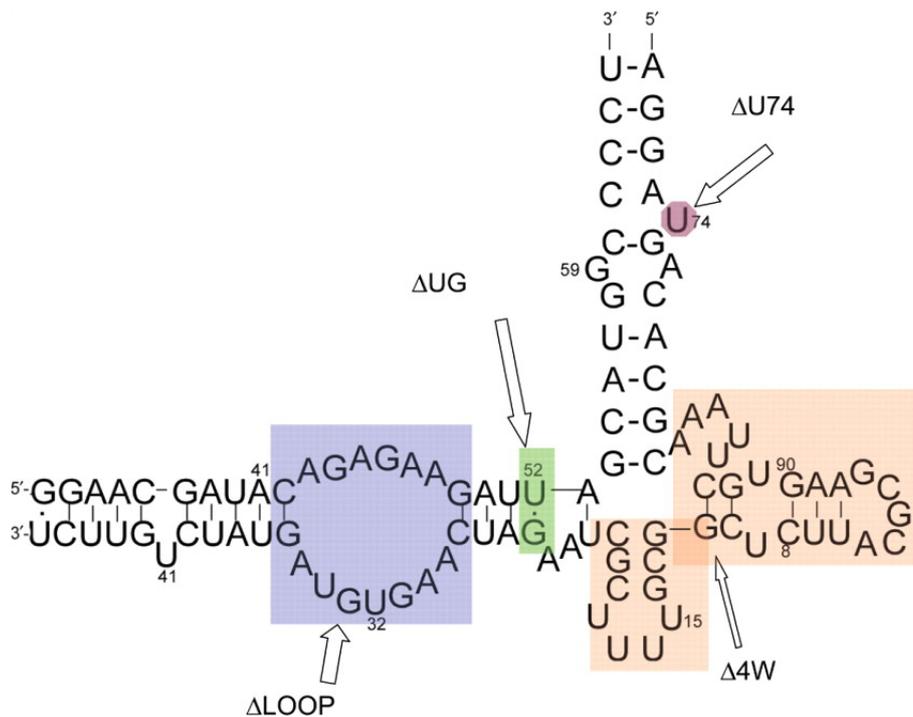
### 2.3.1 Design of RNA samples and confirmation of base pairing

The proposed secondary structure of human U2-U6 snRNA complex is shown in Fig 1.4. It contains three intermolecular helices and a U6 intramolecular stem loop (ISL)<sup>22</sup>. In order to simplify construction of samples suitable for our studies, a model pairing was designed in which the invariant top strand (the shorter RNA in Fig 2.2) and a variable bottom strand represented the native U2-U6 RNA pairing and its variants. In order to make this change, the 5-base loop within the ISL was omitted and a terminal loop was added to helix II. This pentaloop was demonstrated in separate experiments not to bind metal ions (data not shown). Variations of the bottom strand were designed so that when paired with the top strand, suspected ion binding sites were deleted by the formation of complementary base paired segments (Fig 2.2). The top strand included nucleotides 33-64 of the U6 sequence. The bottom strand was a chimera of fragments of U2 and U6 snRNAs.

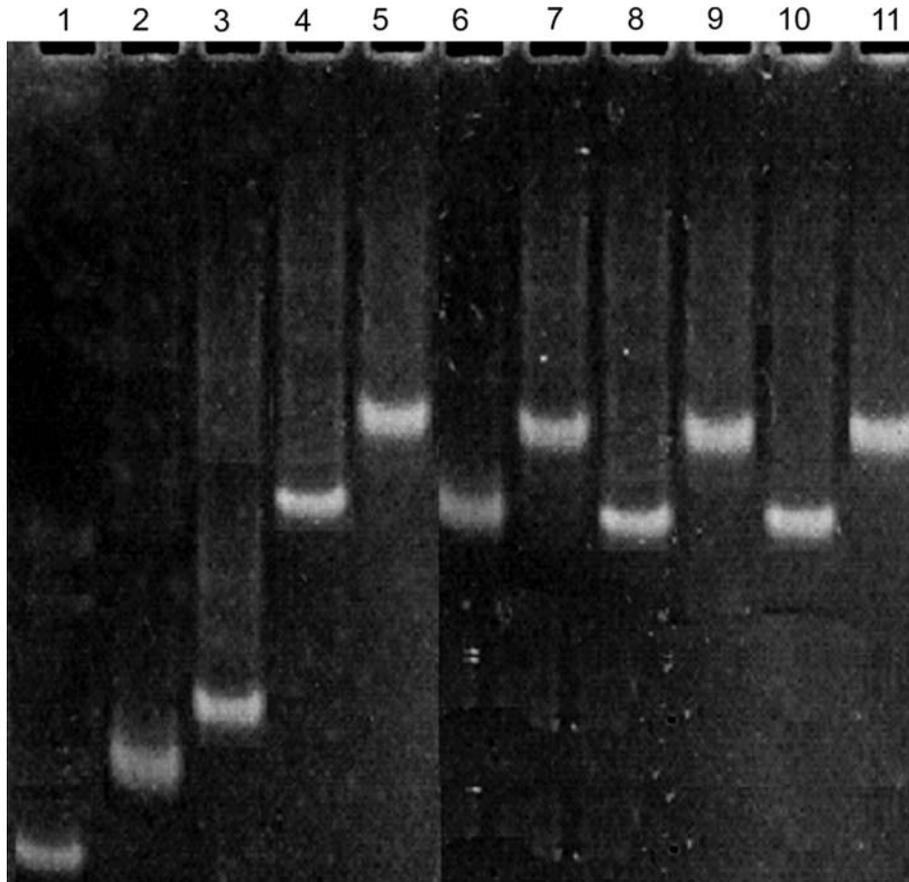
In order to confirm that the bottom strands formed base pairs with the invariant top strand, we monitored migration of each paired RNA sample on a non-denaturing gel. In each case, the disappearance of both single-stranded top and bottom bands and the appearance of a new band which migrated slower than either of the single bands (Fig 2.3) suggested that the constructs formed complexes by the pairing of top and bottom strands.

### 2.3.2 Stoichiometry of site-bound ions

In order to estimate the number of site-bound terbium ions per RNA complex, Tb<sup>3+</sup> luminescence for different ratios of [Tb<sup>3+</sup>]:[RNA] was recorded and plotted using a Job plot<sup>53</sup>. In this context, “site-bound” denotes residence of the metal ion at a specific site accompanied by the loss of one or more water molecules. In the case of Tb<sup>3+</sup>, it is associated with increased



**Figure 2.2 The simplified construct and its mutations used in the FRET study.** The whole sequence represents the simplified original construct (called WT), which includes an invariant 32-nt top strand and a 69-nt bottom strand. Fluorescent dyes are attached at the 3' or 5' end of the 32-nt strand. The four mutation areas are shaded. All mutations were created by variations of segments of the bottom strand.  $\Delta$ LOOP eliminates the ACAGAGA loop by mutating the bottom sequence to result in a complementary stem.  $\Delta$ UG mutates the U-G base pair to U-A.  $\Delta$ 4W eliminates the four-way junction by replacing the shaded area with three adenosines.  $\Delta$ U74 deletes the shaded uridine from the U6 ISL sequence. This figure is from Yuan *et al.* 2007<sup>48</sup>.



**Figure 2.3 Demonstration of base pairing between top and bottom strands by non-denaturing gel electrophoresis.** Lane 1: the invariant top 32 mer. Lane 2:  $\Delta 4W$  bottom strand (40 nt) alone; Lane 3:  $\Delta 4W$  bottom with top strand. Both the brightest bands in lanes 1 and 2 disappeared in lane 3, indicating pairing of the strands. Lanes 4 and 5: WT bottom (69 nt) alone and with top strand. Lanes 6 and 7:  $\Delta LOOP$  bottom (68 nt) alone and with top strand. Lanes 8 and 9:  $\Delta U74$  bottom (68 nt) alone and with top strand. Lanes 10 and 11:  $\Delta UG$  bottom (68 nt) alone and with top 32 strand. The retardation of paired strands indicated the formation of base-paired complexes. This figure is from Yuan *et al.* 2007<sup>48</sup>.

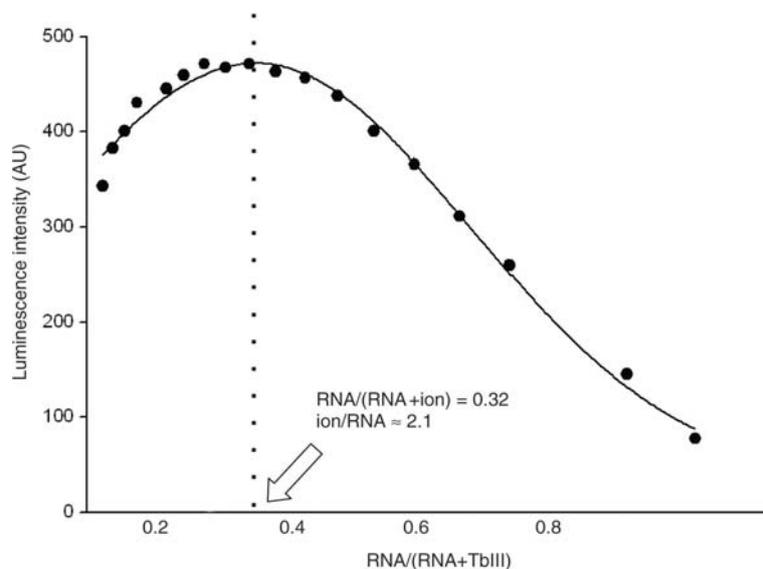
luminescence. The luminescence intensity represents the amount of RNA-Tb<sup>3+</sup> complex by assuming a linear relationship between binding and the luminescence intensity. The principle of this method is that when the combined Tb<sup>3+</sup> and RNA concentration is kept constant, at the maximum intensity, the metal binding site number corresponds to the mole ratio [Tb<sup>3+</sup>]:[RNA].

Figure 2.4 shows the fit process of Tb<sup>3+</sup> binding to the  $\Delta$ U74 mutant complex at pH 7.2 as an example. The luminescence intensity at 545 nm vs. RNA fraction for each measurement was plotted. The data were fit by a Gaussian distribution. The peak of the curve was at the RNA fraction of 0.32. The stoichiometry of ions bound to RNA was calculated as  $n = (1-x)/x$ , where  $x = 0.32$  for  $\Delta$ U74 at pH 7.2. This fraction corresponds to 2.1 binding sites, or approximately two site-bound ions per RNA complex.

Using the same process, we found 3.0 binding sites in the construct WT at pH 7.2. The stoichiometry was the same as that of the native U2-U6 construct (3.1) (sequence in Fig 2.2A), which validated that moving the loop did not affect the number of its site-bound ions. Compared to the native construct (U2-U6) and construct WT,  $\Delta$ LOOP (1.8) and  $\Delta$ U74 (2.1) each had one fewer binding site,  $\Delta$ 4W (2.9) and  $\Delta$ UG (3.3) both had three binding sites (Table 2.1).

The loss of one binding site in the  $\Delta$ LOOP construct as compared with the original three sites occurred upon replacement of the loop between helices I and III by Watson-Crick base pairs. This suggested that this loop had one Tb<sup>3+</sup> binding site. Similarly, the deletion of the U74 nucleotide from the ISL ( $\Delta$ U74 mutant) resulted in the loss of one binding site, suggesting that U74 of U6 ISL participated in formation of one binding site. We could not determine the location of the other binding site from these experiments alone because the  $\Delta$ 4W and  $\Delta$ UG mutations did not eliminate any binding sites.

We note that the use of this approach includes certain assumptions, such as fairly tight binding and a linear relationship between luminescence intensity and the amount of RNA-metal complex, and that measurements derived from them are therefore only approximate. However, our results are likely to be reasonable estimates of Tb<sup>3+</sup> binding behavior in the system because they were internally consistent. Our spectroscopic data on the native and mutant segments thus support the tentative conclusion that there are three metal ion binding sites on U2-U6 complex at pH 7.2: one in the ACAGAGA loop, one in the vicinity of U74 in the U6 ISL, and one not determined.



**Figure 2.4 Measurement of RNA metal-ion-binding stoichiometry by Job plot.** By way of example, the Job plot for determination of ion-binding stoichiometry of the  $\Delta$ U74 mutant complex (from which U74 had been deleted from the ISL) at pH 7.2 is shown. Luminescence was recorded for samples with a range of RNA/Tb<sup>3+</sup> ratios, with total [RNA + Tb<sup>3+</sup>] = 2  $\mu$ M (see Materials and methods section for details), and was plotted versus RNA/(RNA + ion). Gaussian curve was fit to the data points and the peak of the curve was at RNA/[RNA + Tb<sup>3+</sup>] = 0.32. The binding site number (ion/RNA) was  $n = (1-0.32)/(0.32) \sim 2.1$ . The stoichiometry for other constructs was calculated with the same method<sup>50</sup>. This figure is from Yuan *et al.* 2007<sup>48</sup>.

**Table 2.1 Tb<sup>3+</sup> binding stoichiometry of RNA constructs at pH 7.2.** Construct 'U2-U6' includes the original U2 and U6 sequences shown in Fig 2.2. Abbreviations for other constructs and experimental details are provided in the Materials and methods section. Table is from Yuan *et al* 2007<sup>48</sup>.

Construct	Tb <sup>3+</sup> /RNA
U2-U6	3.1 $\pm$ 0.1
WT	3.0 $\pm$ 0.2
$\Delta$ 4W	2.9 $\pm$ 0.1
$\Delta$ LOOP	1.8 $\pm$ 0.2
$\Delta$ U74	2.1 $\pm$ 0.1
$\Delta$ UG	3.3 $\pm$ 0.3

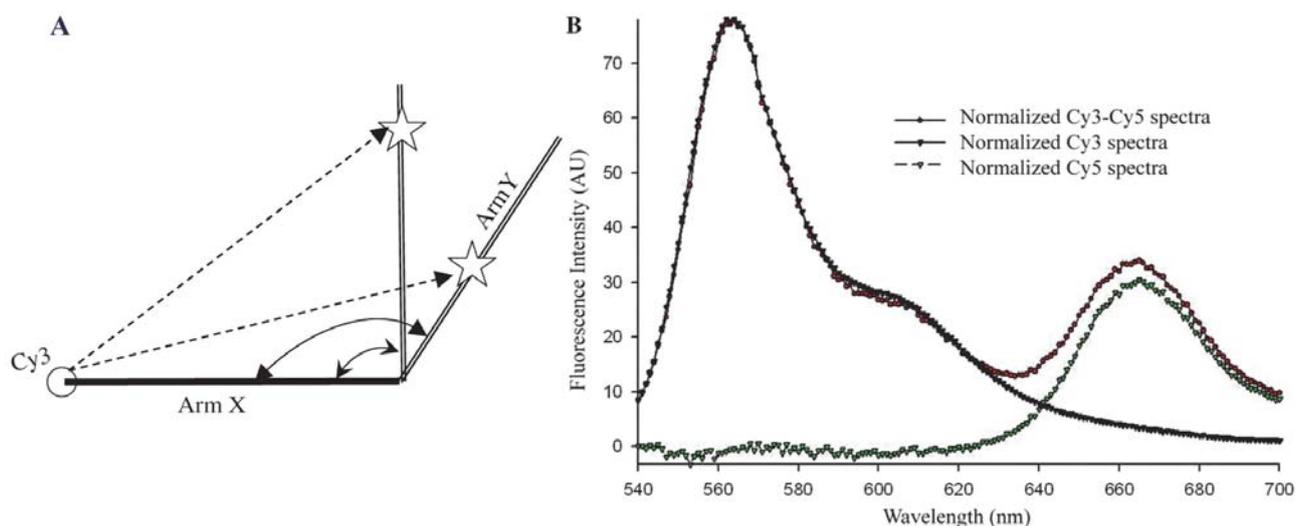
### 2.3.3 Lifetime measurements of Tb<sup>3+</sup> site-bound to unlabeled RNA

For measurement of Tb<sup>3+</sup> lifetime without acceptor,  $\tau_D$ , luminescence decay of Tb<sup>3+</sup> was monitored at 545 nm following a 280nm flash. RNA concentration was 0.7  $\mu$ M. As discussed earlier in section 2.1.2, essentially no luminescence was detected for Tb<sup>3+</sup> solution without RNA concentrations up to 1 mM. Therefore, all signals were from the RNA-bound ions. Tb<sup>3+</sup> luminescence decay with each of the unlabeled RNA constructs were measured. Analysis of each of the decays revealed a single exponential lifetime. The lifetime was relatively constant at  $\tau_D = 0.77 \pm 0.05$  ms, regardless of construct or pH (in the range of 5.6-7.2). The observation of a single population of lifetimes suggests that the metal ions were in similar environments (hydration, *etc.*) in each of the constructs.

### 2.3.4 Determination of the angle between Helix I and U6 ISL

In order to formulate distances obtained from LRET measurements into ion residency locations, the lengths of “arms” X and Y, and the angle between them, are needed (Fig 2.5A). As an approximation, we make the assumption that Helices I and III form a continuous arm between Cy3 and the four-way-junction structure (arm X in Fig 2.5A); the U6 ISL from the four-way-junction to Cy5 is arm Y. Cy3 and Cy5 were attached to the ends of the X and Y arms by means of covalent bonding (Fig 2.2B, 5' and 3' ends of invariant top strand). In order to find the angle between the two arms, we measured FRET between the two dyes. The samples were excited at 520 nm and emission scans were collected from 520 nm to 700 nm. The acceptor normalization method was used to determine the distance between the two dye molecules<sup>54</sup>. A scheme of the normalization process is shown in Fig 2.5B.

The distances between Cy3 and Cy5 in different constructs were summarized in Table 2.2.  $\Delta$ LOOP complex was examined first because it has a complementary paired X arm. There was a distance of  $79 \pm 2$  Å between the two dyes in  $\Delta$ LOOP. When covalently bound to RNA helix, the fluorophore Cy3 was stacked onto the end of helix like an extra base<sup>55</sup>. We assume Cy5 has the same behavior as Cy3 due to their structural similarities. The X arm of  $\Delta$ LOOP has 21 continuous base pairs and the Y arm has 11 base pairs. Assuming each forms a regular A-type helix, with an average rise of 2.6 Å/bp, the length of the X arm is  $2.6 \text{ Å} \cdot (21+1) \approx 57 \text{ Å}$ , and the Y arm is  $2.6 \text{ Å} \cdot (11+1) \approx 31 \text{ Å}$ . The distance between the ends of X and Y arms was 79 Å. The angle between the



**Figure 2.5 Calculation of the angle between helix I and U6 ISL using the acceptor normalization method.** (A) The metal-ion-binding site on arm Y cannot be determined without knowing the angle between arms X and Y. As shown, the distance between Cy3 and a site-bound  $Tb^{3+}$  ion (dashed line from Cy3 to star) will correspond to different sites on arm Y if the angle between arms X and Y is unknown. (B) ‘Acceptor normalization method’ normalization process. Normalized Cy5 fluorescence, which is the acceptor intensity, is calculated from normalized Cy3–Cy5 emission spectra minus normalized Cy3 scan. This figure is from Yuan *et al.*, 2007<sup>48</sup>.

**Table 2.2 Distances between Cy3 and Cy5 in different constructs.** The invariant top strand was labeled with Cy3 and Cy5 on 5’ end and 3’ end respectively. In each construct, by pairing the 32-nt labeled oligomer with its bottom strand, the distance between the two dyes was measured by the acceptor normalization method described in the text and shown in Fig 2.5B. This figure is from Yuan *et al.*, 2007<sup>48</sup>.

Construct	% Efficiency	Cy3 - Cy5 Distance
WT	$14 \pm 1$	$73 \pm 1 \text{ \AA}$
$\Delta 4W$	$13 \pm 1$	$74 \pm 1 \text{ \AA}$
$\Delta LOOP$	$9.4 \pm 1$	$79 \pm 2 \text{ \AA}$
$\Delta U74$	$13 \pm 1$	$74 \pm 1 \text{ \AA}$
$\Delta UG$	$12 \pm 1$	$75 \pm 1 \text{ \AA}$

X and Y arms was triangulated to be  $124 \pm 5^\circ$ , *i.e.*, essentially tetrahedral. There are several assumptions made in these calculations that may contribute to some inaccuracy in the overall measurement, *e.g.*: linear A-type helix throughout the X and Y arms; and random orientation of the dyes on their linkers. Assumption of randomness in the orientation of dye linkers was generally made, and was considered in the  $\kappa^2$  measurement as well as the distance measurement. The X arm of the  $\Delta$ LOOP construct comprises all Watson-Crick or GU base pairs, which are anticipated to obey standard A-type helical geometry; similarly, the NMR-derived structural model of the ISL indicates an essentially collinear A-type helix for this stem loop<sup>38</sup> (the Y arm).

We then used the same FRET technique to measure the Cy3-Cy5 distance in each of the complexes that contained the native ACAGAGA loop sequence in the X arm. FRET measurements yielded a mean Cy3-Cy5 distance of  $74 \pm 2 \text{ \AA}$ . Assuming that the angle between the two arms remained constant in all constructs in which the four-way junction was not altered, and using the same values for the triangulation ( $31 \text{ \AA}$  for the Y arm,  $124^\circ$  between the X and Y arms and Cy3-Cy5 distance of  $74 \text{ \AA}$ ), the mean length of the X arm with the native internal loop was calculated as  $52 \pm 1 \text{ \AA}$ , as compared with  $57 \text{ \AA}$  in the construct with the complementary base paired ACAGAGA region. These data suggest that that presence of the internal loop shortens the overall length of the stem. Although it is possible that presence of the highly flexible internal loop in the Helix I-III regions may induce some deviation from pure A-type helix or perfect collinearity in the X arm, we note that the addition of a fragment representing the intron (which pairs with the bottom strand opposing the ACAGAGA sequence) did not alter the measurement (data not shown).

### **2.3.5 Lifetime measurements with Cy3-labeled RNA constructs at pH 7.2**

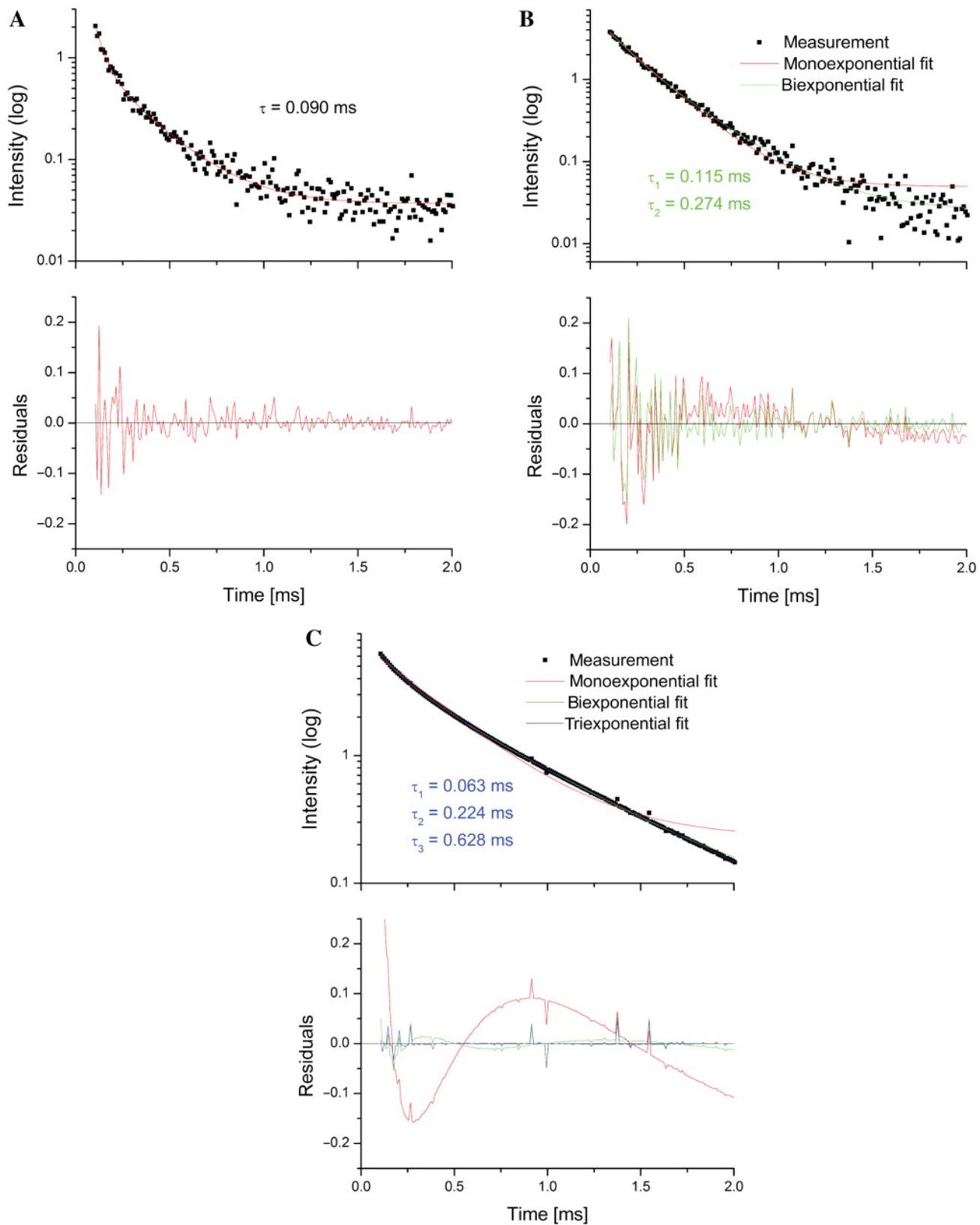
We next measured the distances between Cy3 and  $\text{Tb}^{3+}$  at pH 7.2. In order to accomplish this, we analyzed decay curves following FRET from site-bound  $\text{Tb}^{3+}$  ions to Cy3. FRET contributions from individual  $\text{Tb}^{3+}$  ions to Cy3 yield distance information between the organic dye and each of the luminescent metal ions. The resulting emission decay curve was analyzed for the best fit by one, two, or three (the maximal number of curves that could be reliably evaluated) components. For data acquired at pH 7.2, the luminescence decay curve was best fit by the sum of three independent decay curves with lifetimes  $\tau_{\text{DA1-3}}$  of  $\tau_{\text{DA1}} = 0.06 \pm 0.005 \text{ ms}$ ;  $\tau_{\text{DA2}} = 0.22 \pm 0.019 \text{ ms}$  and  $\tau_{\text{DA3}} = 0.63 \pm 0.031 \text{ ms}$ , with an overall  $R^2 = 0.998$  (Fig 2.6C). To estimate the quality of the

**Figure 2.6 (Figure next page) The exponential decay fitting process used in the study.** The exponential decay data were fitted with mono, bi- or tri-exponential decay curves. An example of the best fit in each category (1, 2 or 3 lifetimes) is presented. **(A)** WT construct at pH 5.6 best fitted to one-lifetime decay.  $\tau \approx 0.09$  ms. **(B)** Construct  $\Delta U74$  at pH 7.2 best fit to two-lifetime exponential decay.  $\tau_1 \approx 0.12$  ms;  $\tau_2 \approx 0.27$  ms. **(C)** Construct WT at pH 7.2 best fitted to three-lifetime exponential decay.  $\tau_1 \approx 0.06$  ms;  $\tau_2 \approx 0.22$  ms;  $\tau_3 \approx 0.63$  ms. The residuals represent the differences between experimental data and fitting curves, which showed the goodness of fit. The choice of the number of lifetimes is based on the  $R^2$  and  $\chi^2$  values. The improvement of the fit is shown by comparing other fits and the residue figures. Residuals refer to the difference between theoretical and experimental data. The figures are from Yuan *et al.* 2007<sup>48</sup>.

**Table 2.3 Summary of  $Tb^{3+}$  lifetimes (in milliseconds) in the presence of acceptor ( $\tau_{DA}$ ) at different pH values.** The fitting process is described in the text and examples are shown in Fig 2.6.  $\tau_1$ ,  $\tau_2$  and  $\tau_3$  are three lifetimes ( $\pm$ SD included in text). The Cy3– $Tb^{3+}$  distances corresponding to the three lifetimes are  $\sim 32$ ,  $\sim 50$  and  $\sim 67$  Å, respectively.  $\chi^2$  values (reported as reduced  $\chi^2$  values) from deconvolution of compound curves are listed with the lifetimes. From Yuan *et al.* 2007<sup>48</sup>.

	pH 5.6				pH 6.1				pH 6.6				pH 7.2			
	$\tau_1$	$\tau_2$	$\tau_3$	$\chi^2$												
WT	0.09	-	-	0.00024	0.07	0.27	-	0.00073	0.13	0.32	-	0.00092	0.06	0.22	0.63	0.00021
$\Delta 4W$	0.14	0.34	-	0.00078	0.14	0.34	-	0.00065	0.05	0.24	-	0.00108	0.05	0.32	0.60	0.00064
$\Delta LOOP$	-	-	-	-	-	0.30	-	0.00074		0.30	-	0.00083	-	0.35	0.60	0.0008
$\Delta U74$	0.06	-	-	0.00033	0.10	0.26	-	0.00075	0.10	0.26	-	0.00085	0.12	0.27	-	0.00056
$\Delta UG$	0.09	-	-	0.0005	0.08	0.29	-	0.00047	0.08	0.29	-	0.00089	0.05	0.32	0.60	0.00017

Figure 2.6, caption in page 34



fit, apart from the  $R^2$ -value, the residuals (difference between expected and experimental data) were considered. The three-exponential fit is consistent with three independent decay curves, corresponding to three independent ion binding sites, the location of which can be mapped from their distance from Cy3.

As further support for the finding of three discrete site-bound  $Tb^{3+}$  ions at pH 7.2, the stoichiometry measurements acquired under steady state conditions, also yielded approximately three  $Tb^{3+}$  ions for the WT construct under the same conditions (Table 2.1).

### **2.3.6 Calculation of Cy3- $Tb^{3+}$ distances and localization of ion binding sites.**

From the measured values for  $\tau_{DA1}$ ,  $\tau_{DA2}$ , and  $\tau_{DA3}$  and the relationships in equation 2.6, transfer efficiencies  $E_1$ ,  $E_2$ , and  $E_3$  were calculated as  $92 \pm 1\%$ ,  $71 \pm 1\%$ , and  $18 \pm 1\%$ , respectively, from which (using equation 2.1) we calculated distances ( $R_1$ ,  $R_2$ , and  $R_3$ ) corresponding to 36 Å, 46 Å, and 69 Å from the Cy3 dye. Using the three distances, the approximate length of the X and Y arms, and the calculated X-Y angle of  $\sim 124^\circ$ , we estimated the locations of the three binding sites within the WT construct. Again, using the value of  $\sim 2.6$  Å rise per base pair in an A-type helix and for Cy3, the first site, measured as 36 Å from Cy3, corresponds to a site located  $(36-2.6)/2.6 \approx 12.8$  base pairs away from Cy3. This approximate measurement places the ion within the center part of the ACAGAGA loop, *i.e.*, in the vicinity of the first G if helical geometry is maintained, or in the GAG sequence if the unpaired helix is compacted. The second site,  $\sim 46$  Å from the dye, is near the end of the X arm (length of X arm  $\approx 52$  Å with Cy3); therefore the ion is located in the vicinity of the four-way-junction structure, which includes the AGC triad. In order to place the third ion, which was a distance of 69 Å from the Cy3 dye, we exploited the angle between the X and Y arms (estimated by FRET methods to be  $\sim 124^\circ$ ) as well as the length of the X arm (52 Å with Cy3) to solve for the length of the third side of the triangle, which was 25 Å from the X-Y junction (Fig 2.5A). Arm Y, corresponding to the U6 ISL, has already been shown to adopt essentially A-type helical geometry throughout<sup>38</sup>, which (assuming the human ISL adopts similar overall structure to that of yeast) allows us to place the third ion approximately  $25/2.6 \approx 9.6$  bases pairs away from the four-way-junction. This placed the site in the vicinity of U74 within the ISL.

### **2.3.7 Confirmation of ion-binding locations by use of mutations.**

In order to confirm the validity of the three-lifetime fit of the luminescence data and to assist

with determination of precise ion-binding locations, we repeated the  $\text{Tb}^{3+}$ -Cy3 FRET studies on variant U2-U6 complexes. As a means to test for the presence of each independent site (which would have a corresponding luminescence transfer process), we designed a specific U2-U6 pair in which the suspected site was mutated or deleted by formation of Watson-Crick base pairs in place of the bulge, loop, or noncanonical pairing.

We first probed binding with the U6 ISL. The proposed site in the vicinity of U74 has been shown to bind metal ions<sup>38</sup>.  $\Delta\text{U74}$  has the deletion of U74. Samples were excited at 280 nm, started recording with a gate of 0.1 ms at 565 nm. The  $\Delta\text{U74}$  decay data were analyzed and two lifetimes were found (Fig 2.6B), 0.12 ms and 0.27 ms, with  $R^2 = 0.996$ . These two lifetimes and calculated distances were very similar to the ones resulting from the ACAGAGA loop site and 4way junction site of WT. The disappearance of the longest luminescence lifetime with the modification suggested that the structure caused by U6 ISL U74 was responsible for the metal ion binding on the Y arm.

To test the hypothesis that the loop between helices I and III is responsible for binding, we performed lifetime FRET on  $\Delta\text{LOOP}$  in which the bottom strand was mutated to form Watson-Crick base pairs in place of ACAGAGA loop. The decay data were recorded and analyzed as described. In comparison with the WT construct, two lifetimes were found, 0.35 ms and 0.60 ms ( $R^2 = 0.989$ ). These lifetimes and the calculated distances corresponded to those in the four-way junction site and the ISL site of WT. The disappearance of the first site suggested that the loop was responsible for binding at the first site.

Finally, I attempted to determine which region was responsible for ion-binding in the proximity of the four-way junction. We first replaced the four-way junction (shaded area in Fig 2.2B) with a three nucleotide “hinge” in  $\Delta\text{4W}$ . The FRET decay curve of  $\Delta\text{4W}$  was similar to that of WT (best fit with lifetimes of 0.05, 0.32, 0.60 ms,  $R^2 = 0.993$ ), implying similar ion binding sites as that of WT. These results suggested that either the four-way junction itself was not responsible for the second binding site, or that binding was replaced by three adenosine residues at the hinge.

It has been shown that the U-G non-Watson-Crick base pair has many important biological functions including the potential to bind metal ions<sup>56,57</sup>. To test whether U-G was responsible for binding metal ions, a U-G to U-A base pair mutation was made in  $\Delta\text{UG}$ . The exponential decay analysis suggested it still had three binding sites, and its binding properties were similar to WT (best fit with lifetimes of 0.05, 0.32, 0.60 ms,  $R^2 = 0.993$ ). From this, we concluded that mutating the

U-G base pair did not change the second binding site.

In summary, we have identified three binding sites on the U2-U6 complex at pH 7.2. Using mutations, we have defined the regions responsible for the first and third binding sites. However, the second binding site, in the region of four way junction, remains poorly characterized.

### 2.3.8 pH-dependence of ion binding.

We evaluated the pH-dependence of Tb<sup>3+</sup>-binding behavior at each of these sites by performing the lifetime FRET at different pH environments. We excited the Tb<sup>3+</sup>-RNA sample at 280nm and recorded the resulting emission at 565nm as described previously. All results are listed in Table 2.3.

#### *pH 5.6*

The resulting decay of construct WT was best fit with one lifetime ( $\tau = 0.09$  ms,  $R^2 = 0.997$ , Fig 2.6A), which corresponds to a distance of about 39 Å between Tb<sup>3+</sup> and Cy3. This distance from Cy3 also places the Tb<sup>3+</sup> in the ACAGAGA loop. The absence of emission components corresponding to the other two sites suggests that terbium ion is binding only at the ACAGAGA loop, *i.e.*, binding at other locations is pH-dependent. Similarly, the decay curves following the excitation of Tb<sup>3+</sup> bound to the  $\Delta U74$  or  $\Delta UG$  constructs were best fit with a single lifetime ( $\tau = 0.06$  ms,  $R^2 = 0.933$ ;  $\tau = 0.09$  ms,  $R^2 = 0.987$ , respectively). Interestingly, decay data for  $\Delta 4W$  was best fit by two lifetimes ( $\tau_1 = 0.14$  ms,  $\tau_2 = 0.34$  ms,  $R^2 = 0.962$ ), which corresponded to the ACAGAGA and junction regions. These data suggested that replacing the junction for the hinge may create a non-pH-dependent ion binding site, whether by creating a new U-G base pair, or by changing the binding property of the AGC triad. It has been shown that U-G non-Watson-Crick base pairs have the potential to bind metal ion<sup>56,57</sup>. Excitation of  $\Delta LOOP$  did not yield signal and suggested no ions were bound to this construct at this pH, thus reinforcing the conclusion that the one site at pH 5.6 was the ACAGAGA loop.

#### *pH 6.1 and pH 6.6*

At pH 6.1, the decay curves of construct WT were best fit with two lifetime ( $\tau_1 = 0.07$  ms,  $\tau_2 = 0.27$  ms,  $R^2=0.995$ ). The calculated distances between Tb<sup>3+</sup> and Cy3 suggested the locations of

the ion sites on the ACAGAGA loop and the junction. Similarly, the decay curves of the  $\Delta 4W$ ,  $\Delta U74$  or  $\Delta UG$  constructs were best fit with similar double lifetimes ( $\tau_1 = 0.14$  ms,  $\tau_2 = 0.34$  ms,  $R^2 = 0.979$ ;  $\tau_1 = 0.10$  ms,  $\tau_2 = 0.26$  ms,  $R^2 = 0.987$ ;  $\tau_1 = 0.08$  ms,  $\tau_2 = 0.29$  ms,  $R^2 = 0.992$ ; respectively), suggesting the same two ion-bound locations. The decay curve of  $\Delta LOOP$  has one lifetime ( $\tau = 0.30$  ms,  $R^2 = 0.946$ ), indicating the ACAGAGA loop site was disrupted. At pH 6.6, the two pH environments yielded similar lifetime results. The lifetime results of different constructs at different pH values were summarized in Table 2.3.

In summary, we have identified three regions of site-binding by  $Tb^{3+}$  in the U2-U6 snRNA complex, all of which can be competed by  $Mg^{2+}$ . One is located at the ACAGAGA loop; it is not pH-dependent in the pH range of 5.6-7.2. The second is located in the vicinity of the four-way-junction structure, possibly associated with the AGC triad in the context of the intact junction structure; it binds metal ions at pH > 6 only. The third ion-binding site is located in the vicinity of U74 of U6 ISL; it binds metal ions at pH > 7 and does not bind at pH lower than 7.

## 2.4 Discussion

FRET techniques have long been used as a valuable tool in structural biology. As an extension of traditional FRET experiments using two organic dyes, experiments using lanthanide ions as donors<sup>40</sup> have several advantages over conventional FRET: 1) large difference in lifetimes between donor and acceptor minimizes inaccuracies stemming from incomplete incorporation of an organic dye; 2) the orientation factor of the site-bound metal ion becomes unimportant, as compared to organic dyes, which are more dependent on the rotational mobility of the fluorophores<sup>41</sup>; 3) the millisecond lifetimes of lanthanide luminescence, as compared with nanoseconds for organic dyes, allow for ease of time-resolved data acquisition by a standard fluorophotometer; and 4) the  $R_0$  values of the  $Ln^{3+}$ -organic dye pairs are generally larger than for dye-dye pairs, enabling measurements over a longer distance<sup>41</sup>.

In contrast with previous lanthanide FRET experiments, which used chelated ions as exogenous probes<sup>40</sup>, we have now taken advantage of the relative similarity in binding properties between  $Mg^{2+}$ , the spectroscopically uninformative native ion, and  $Tb^{3+}$ , a luminescent lanthanide ion, to exploit energy transfer between excited site-bound  $Tb^{3+}$  and a covalently attached fluorophore to target locations where  $Tb^{3+}$  is situated in RNA complexes.

Mg<sup>2+</sup> and Ln<sup>3+</sup> ions have been shown to occupy overlapping sites in a number of RNAs<sup>58-60</sup>. In the U2-U6 snRNA construct used in these studies, competition of Tb<sup>3+</sup> luminescence by Mg<sup>2+</sup> implied that each of the ions was capable of occupying a site in the same vicinity, although detailed properties of binding may be different<sup>61,62</sup>.

This is the first time a site-bound metal ion has been used as a FRET probe. Misra and Draper have described three general modes of multivalent ion binding to RNA: “diffuse binding” that screens charges between RNA backbone segments “outer sphere site binding” involving specific coordination of anionic ligands to hydrated Mg<sup>2+</sup>, and “inner sphere site binding”, involving one or more direct contacts between the ion and a ligand without intervening water molecules<sup>63</sup>. All modes are important for the stabilization of RNA structure. However, a key feature that makes this measurement possible is that emission of Ln<sup>3+</sup> is largely quenched by the -OH oscillators of water<sup>64</sup> and RNA sensitizing, so that luminescence is only observed for those ions bound *via* some degree of inner sphere coordination.

Use of Tb<sup>3+</sup> in a FRET study allows us to map binding sites directly without altering its native structure or mutating chemical groups. As a result of the relatively large R<sub>0</sub> of the Tb<sup>3+</sup>-dye pair, long lifetime, and high signal-to-noise ratio, ion binding site can be probed over a wide range of distances and multiple populations could be resolved from the measurements with greater confidence. The number of sites is limited to the fitting process of the luminescence decay curve; in this case, we found that deconvolution into three independent curves was the limit of resolution. However, the fact that there were only three site-bound Tb<sup>3+</sup> ions contributing to energy transfer to Cy3 was supported by separate stoichiometry measurements of each of the U2-U6 snRNA complexes. Measurements for each construct in which a suspected binding region was replaced by a complementary paired stem resulted in one less site-bound Tb<sup>3+</sup> ion.

From Tb<sup>3+</sup>-Cy3 FRET data, we found that lifetimes of the three curves (whether calculated by deconvolution or measured individually in mutant constructs) fell into three groups at pH 7.2:  $\tau_{DA1} \approx 0.07\text{ms}$ ;  $\tau_{DA2} \approx 0.31\text{ms}$  and  $\tau_{DA3} \approx 0.61\text{ms}$ , corresponding to distances from Cy3 (at the 5' terminus of the U6 snRNA strand) of 36Å, 46Å and 69Å, respectively. Using the angle of 124° determined between the two arms, calculated from FRET between Cy3 and Cy5 on termini, we translated these distances into locations corresponding to Tb<sup>3+</sup> binding sites in the central part of the

ACAGAGA sequence, the four-way-junction, and the internal loop (including U74) of the U6 intra-molecular stem loop.

The fact that we observed the same patterns in the native and mutant sequences supports the robustness of the fitting process. Lifetime measurements on each mutation resulted in a decay curve missing the specific component corresponding to the mutation without perturbing the remaining components. The one exception was the slightly longer distance between Cy3 and the four-way junction site in the  $\Delta$ LOOP construct (35 Å, as compared with ~28 Å), which is consistent with the longer continuous stem created by the complementary pairing between helices I and III in place of the flexible ACAGAGA loop.

Divalent cations play both catalytic and structural roles in ribozymes<sup>47</sup>. U2 and U6 snRNAs are the only RNAs required for both splicing reactions and their sequences are conserved phylogenetically. A protein-free U2-U6 complex has been shown to catalyze an intramolecular reaction similar to the first step of splicing<sup>22</sup>. This evidence strongly suggests that the spliceosome is a ribozyme. The obliteration of splicing activities following substitution of phosphorothioate for phosphate in the backbone at U74, the two G residues in the invariant ACAGAGA segments, or the A of the invariant AGC triad, and the rescue of activity with  $Mn^{2+}$ , support the model that binding of  $Mg^{2+}$  or other hard metal ion in these regions is essential for structural or chemical activities associated with RNA splicing<sup>23, 25, 26, 65</sup>. NMR experiments displayed changes in proton chemical shifts for U80, the yeast equivalent of U74, upon addition of  $Mg^{2+}$  to an RNA stem loop representing the ISL from *S. cerevisiae*<sup>38</sup>, implying that this was a region of ion interaction. Our results obtained by FRET from site-bound  $Tb^{3+}$  are entirely consistent with the conclusions of these other investigations.

Using the same  $Tb^{3+}$ -Cy3 FRET technique, we also evaluated the pH-dependence of each of the luminescence contributions to the total decay signal. First, we found that the signal with the shortest lifetime that associated with binding of  $Tb^{3+}$  in the ACAGAGA loop, exhibited no pH dependence, *i.e.* we observed a luminescence component from this site at all pH values examined. It is important to note that, however, under *in situ* conditions, behavior at this site may be modulated by increased structure associated with formation of the branch site motif upon pairing of the intron strand with U2 snRNA<sup>27</sup>.

The luminescence component with the longest lifetime, corresponding to the site within the ISL disappeared at pH below 7, suggesting that  $Tb^{3+}$  does not bind to this site below pH 7. The remaining compound luminescence decay curve was identical to that of the  $\Delta U74$  mutant. This conclusion is completely consistent with findings of Huppler *et al.* on the ISL sequence of yeast<sup>38</sup>, who had observed changes in NMR chemical shifts to show the pH-dependent metal ion binding of the U6 ISL. The similarity (but not identity) of yeast and human sequences is no guarantee of equivalent ion-binding behavior. Therefore, the current demonstration of pH-dependent site-binding in the human sequence *via* a new method, at a location identified by phosphorothioate substitution experiments to require interaction with hard metal ions to achieve splicing activity<sup>26</sup>, is of biological, as well as technical, importance. The similarity in behavior to the yeast construct<sup>66</sup> suggests that the ion-binding behavior at this site is conserved phylogenetically.

The ion-bound site located in the four-way-junction has previously been suggested<sup>23</sup>, but the pH-dependence of the site has not been analyzed. Our stoichiometry data suggest only one binding site in the vicinity of the four-way junction (Table 1). At pH 5.6, the  $\Delta 4W$  mutant (in which the four-way junction was replaced by three adenosines) contributed luminescence to the compound signal, implying that there was a  $Tb^{3+}$  ion bound at this site. In contrast, any of the constructs containing the native four-way junction exhibited evidence of binding the ion only at higher pH values. NMR studies of the four-way junction have shown that it is flexible<sup>66</sup>; although it is possible that pH-dependent site in the junction has been replaced by a pH-independent site in the adenosine bulge, we consider it more likely that the lesser flexibility or perturbation of local structure induced by the bulge modifies metal ion binding properties by a native site in the area. The most likely candidate is the AGC triad, whose backbone was implicated in the binding of hard metal ions essential for splicing by phosphorothioate substitution experiments<sup>23</sup>. We speculate under native conditions, a pK shift in the adenosine of the triad allows it to be protonated at pH <6, preventing ion binding. If so, removal of the flexible junction perturbs specific structural or electrostatic features surrounding the triad by artificially stabilizing other noncanonical interactions. Although the environment in cells is maintained at relatively constant pH, shifted nucleotide pK<sub>a</sub> in small ribozymes have been found to be surprisingly frequent, and have been associated with various ion-binding and catalytic mechanisms<sup>67</sup>. The two pH-dependent metal ion binding sites we identified here may contribute to such mechanisms in the regulation of spliceosomal activity.

# CHAPTER 3

## CHARACTERIZATION OF TERBIUM IONS BOUND ON RNA AND INVESTIGATION OF THEIR SUITABILITY AS FRET DONORS

### 3.1 Introduction

#### 3.1.1 RNA structure and metal ions

Metal ions are important for the stability of folded RNA tertiary structure by electrostatic interactions. Both “diffuse” ions and “chelated” ions are important for RNA structure and function (Fig 1.9) <sup>33</sup>. In this study, “chelated” ions are the focus. Specific metal ion binding (interaction between “chelated” ion and RNA) is very important for RNA structural conformation and stability and for ribozyme catalytic activity <sup>63, 68</sup>.

Throughout this study, RNA samples were dissolved in buffer solutions with high salt concentration (100~200 mM, as specified in each case) monovalent ion ( $\text{Na}^+$ ) to avoid multivalent ions ( $\text{Mg}^{2+}$  or  $\text{Tb}^{3+}$ ) “diffuse” binding. Addition of  $\text{Tb}^{3+}$  in the RNA samples will result in “chelated” binding monitored by luminescence.

In chapter two,  $\text{Tb}^{3+}$  bound on RNA were used as FRET donors to localize specific metal ion binding sites. In this chapter, the properties of these  $\text{Tb}^{3+}$  ions are under careful investigation. It has to be noted that some of following introduction is the theoretic foundation of chapter two.

#### 3.1.2 Probing RNA-metal ion binding by $\text{Tb}^{3+}$

Lanthanide ions [ $\text{Ln}^{3+}$ ] are used for probing metal ion binding properties to metalloproteins, metalloenzymes, and metallo-RNAs since 1970s <sup>69</sup>. The physiological ions such as  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$  are spectroscopically uninformative; thus luminescent metal ions in solution are attractive tools to study metal ion interactions with biomolecules. Aqueous luminescent metal ion solutions are quite

rare. It is advantageous to use lanthanide ions because they are with interesting spectral or magnetic properties<sup>70</sup>. Among them, only Tb<sup>3+</sup> and Eu<sup>3+</sup> are capable of emitting luminescence in aqueous solution at room temperature. Even though there is no biological role attributed to them, because of the chemical and metric similarities between Ln<sup>3+</sup> and Ca<sup>2+</sup>/Mg<sup>2+</sup> (table 3.1), Ln<sup>3+</sup> are used to replace Ca<sup>2+</sup>/Mg<sup>2+</sup> to study metal binding in certain proteins and RNAs<sup>50, 61, 71, 72</sup>.

Like Mg<sup>2+</sup>, Ln<sup>3+</sup> are “hard” acids preferring “hard” base ligands such as oxygen atoms. In comparison, Mn<sup>2+</sup> is “soft” preferring “soft” ligands such as sulfur atoms<sup>73</sup>. Ln<sup>3+</sup> have been shown to occupy the identical sites as Mg<sup>2+</sup> in some x-ray structures of RNA<sup>74</sup>. In the case of GNRA tetraloop which will be used as model system, the NMR structural study shows that Mg<sup>3+</sup> and Ln<sup>3+</sup> do occupy the same site (Monwar, M. and Greenbaum, NL. unpublished data). However, it has to be noted that Mg<sup>2+</sup> may bind RNA differently from Ln<sup>3+</sup>. The ionic radii and coordination numbers are different. As an example, a report by Mundoma and Greenbaum shows that not only Ln<sup>3+</sup> has much stronger binding affinity to RNA than Mg<sup>2+</sup> does; but also Ln<sup>3+</sup> binds RNA with inner-sphere coordination while Mg<sup>2+</sup> with outer sphere coordination<sup>61, 75</sup>.

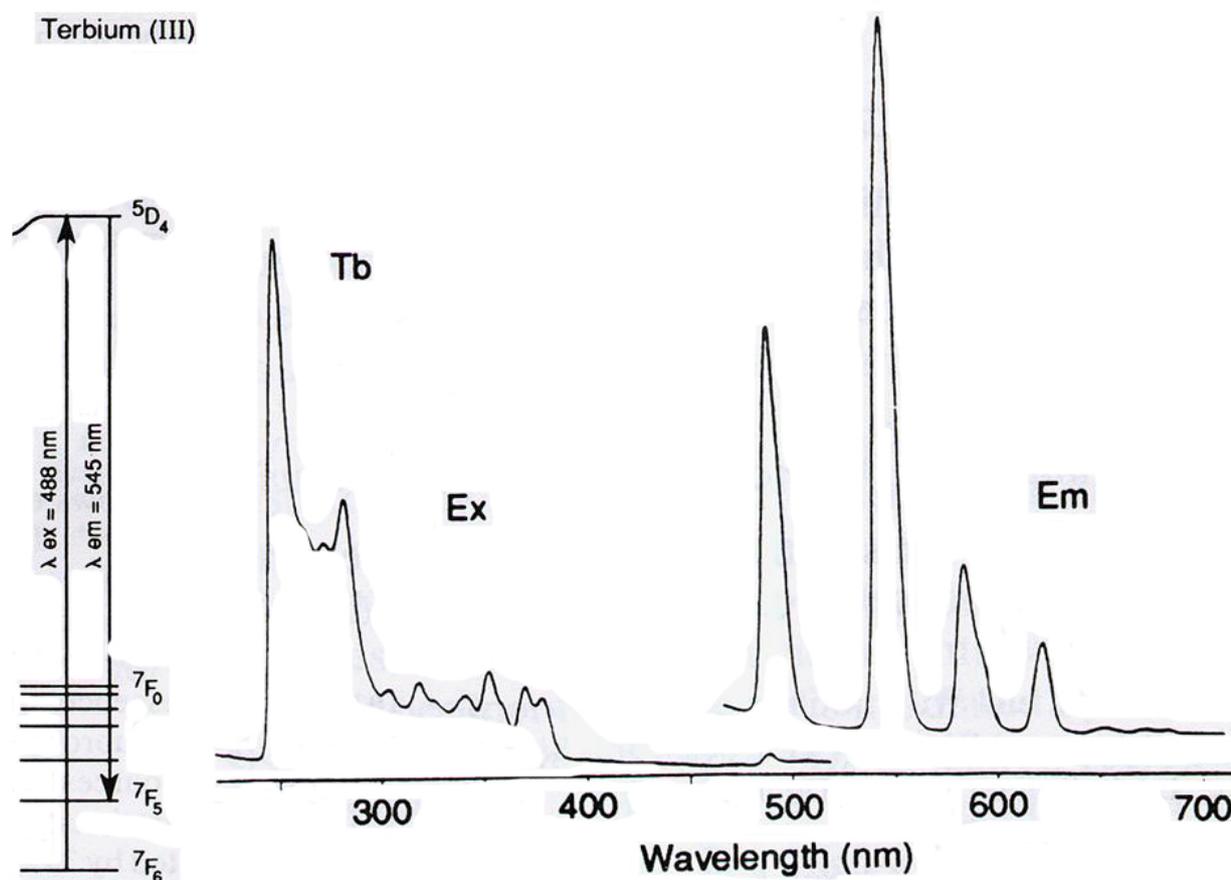
Tb<sup>3+</sup> (atomic number of 65, 159 Da) is a “hard” acid as Mg<sup>2+</sup>. Tripositive terbium ion Tb<sup>3+</sup> in aqueous solution has luminescence as shown in Fig 3.1. Tb<sup>3+</sup> maximal absorption is around 260-280 nm, where Eu<sup>3+</sup> does not absorb. Thus, Tb<sup>3+</sup> is the choice of Ln<sup>3+</sup> in this research. The luminescence spectrum has four sharply spiked peaks at 488nm, 542nm, 588nm, and 620nm arising from electric dipole transitions with long excited-state lifetimes lasting in the range of milliseconds. Long luminescence lifetime results from the forbidden <sup>5</sup>D<sub>4</sub> to <sup>7</sup>F<sub>J</sub> (where J = 0–6) electronic transitions. Tb<sup>3+</sup> emission is formally neither fluorescence (singlet-to-singlet transition) nor phosphorescence (triplet-to-singlet transition)<sup>41</sup>. The emission is primarily from electric dipole transitions.

### 3.1.3 Tb<sup>3+</sup> bound on RNA as FRET donors

In chapter two, Tb<sup>3+</sup> bound on RNA were used as FRET donors in the U2-U6 snRNA complex to locate the Tb<sup>3+</sup> binding sites as well as their pH-dependence<sup>48</sup>. This is the first time an unchelated Ln<sup>3+</sup> is used as a FRET donor. Lanthanide ions were previously used as FRET donors in chelated forms<sup>40</sup>. In addition to the advantages of luminescence resonance energy transfer (LRET)<sup>41</sup>, such as long lifetimes, relatively bigger R<sub>0</sub>, high signal-to-noise ratio, using RNA bound Ln<sup>3+</sup> as FRET donors allows us to measure relatively strong lanthanide ion binding sites directly without

**Table 3.1: Magnesium, Terbium and Europium ion size and coordination number comparison.** The lanthanide ions  $Tb^{3+}$  and  $Eu^{3+}$  have larger ion radii than  $Mg^{2+}$ , accompanied by higher charges.<sup>76</sup> typically,  $Mg^{2+}$  has 6 inner-sphere coordination while  $Ln^{3+}$  has up to 9<sup>61</sup>.

Metal Ion	Ionic Radii (Å)	Geometry	Inner-sphere coordination
$Mg^{2+}$	0.65	Hexahedral	6
$Tb^{3+}$	0.923	Octahedral/Enneahedral	Up to 9
$Eu^{3+}$	0.950	Octahedral/Enneahedral	Up to 9



**Figure 3.1  $Tb^{3+}$  energy level and spectra of excitation and emission.** Left: Terbium(III) excited state is  $5D_4$ ; when it comes back to  $7F$  state, it emits photons. Right: excitation spectrum of  $Tb^{3+}$  on the left; emission spectrum on the right. Figures are from Horrocks 1993<sup>77</sup>.

altering its native structure or mutating chemical groups.

The enhancement of  $Tb^{3+}$  luminescence when “chelated” binding to RNA makes  $Tb^{3+}$  attractive FRET donors. The reason for the enhancement, as we discussed in section 2.1.2, is that RNA aromatic bases are able to absorb and transfer energy from incoming excitation photons to their neighboring  $Tb^{3+}$ . The removal of part or all of  $Tb^{3+}$  inner-sphere water molecules, whose O-H bonds quench  $Tb^{3+}$  luminescence<sup>75</sup>, contribute several magnitudes lower for the enhancement. Inner-sphere water molecules decrease luminescence intensity by decreasing  $Tb^{3+}$  luminescence lifetimes. However, they do not change the absorption coefficient. Thus, the removal of them is not a major contributor for the enhancement. The enhancement of luminescence, however, is always accompanied by the removal of inner-sphere water molecules because  $Tb^{3+}$  has to be in the proximity of RNA aromatic groups (see Fig 1.9left for a cartoon illustration).

### 3.1.4 Properties of $Tb^{3+}$ bound on RNA and their suitability as FRET donors

Characterizing the properties of the  $Tb^{3+}$  bound on RNA will provide a theoretical foundation for the broader application of using them as FRET donors. There are some areas remaining for clarification regarding the properties of  $Tb^{3+}$  bound on RNA.

**1: Accuracy of using luminescence steady state intensity to represent  $Tb^{3+}$  bound on RNA concentration.** In chapter two, a simplified relationship between the concentrations of RNA- $Tb^{3+}$  complex and luminescence intensity was used to calculate RNA binding constant and stoichiometry:

$$I = B_1[Ln^{3+*}] + B_2[Ln^{3+}] \approx B[Ln^{3+*}] \quad (\text{Equation 3.1})$$

$I$  is luminescence intensity;  $Ln^{3+*}$  is  $Tb^{3+}$  bound on RNA;  $Ln^{3+}$  is free  $Tb^{3+}$ ;  $B_1$ ,  $B_2$  and  $B$  are normalization factors to convert concentrations to luminescence intensity. Here,  $B_2 \ll B_1$ ,  $B_1 \approx B$ .

The first concern regarding this relationship is that the intensity measured may not accurately represent the real intensity of the  $Tb^{3+}$  bound on RNA luminescence. An accurate intensity is the prerequisite to use this relationship. A careful examination reveals that some of the lanthanide luminescence intensity is lost by the instrumentation because of  $Tb^{3+}$  long lifetimes.

The second concern is that a linear relationship may not be able to accurately describe the relationship between the concentration of  $Tb^{3+}$  bound on RNA and its luminescence intensity. There is little disagreement that  $B_2 \ll B_1$  because free  $Tb^{3+}$  luminescence is quenched by water molecules

and is not enhanced by RNA aromatic groups (section 2.1.2). However, the assumption of the normalization factor B to be constant is taken into question because observed luminescence lifetime of lanthanide ions is changing. Luminescence lifetimes of lanthanide ions are related to their hydration status<sup>78</sup>.  $Tb^{3+}$  in aqueous solution usually has up to 9 water molecules in its inner-sphere (table 3.1). When  $Tb^{3+}$  “chelated” binds to RNA molecules, some of the coordinated water molecules are replaced by RNA (Fig 1.9). Hydroxyl (O-H) of water molecules vibrations provide a channel for internal conversion of the lanthanide excitation energy, thus the luminescence of lanthanide ions is quenched by water coordination<sup>75</sup>. The hydration results in the shortening of luminescence lifetimes. With long (millisecond range)  $Tb^{3+}$  luminescence lifetime and moderate binding constant ( $\mu M$  range), bound  $Tb^{3+}$  is readily exchanging with solvent. While still excited, bound  $Tb^{3+}$  may 1) bound on RNA without getting quenched; 2) leave RNA, gain inner-sphere water coordination and get quenched; or 3) anything between 1 and 2 and get partially quenched. Thus,  $Tb^{3+}$  luminescence intensity is quenched by a very complicated mechanism, and a linear relationship in equation 3.1 may not be adequate.

**2: Degree of quenching.** The changing  $Tb^{3+}$  luminescence lifetimes indicate the degree of quenching is different at various environments. Which determines the degree of quenching, the  $Tb^{3+}$  concentration, the RNA concentration, or both concentrations? By answering this question, a clear picture of how RNA and  $Tb^{3+}$  interact may be found.

As for the suitability of  $Tb^{3+}$  bound on RNA as FRET donors, the relatively weak binding constant for  $Tb^{3+}$ -RNA will make free  $Tb^{3+}$  present. Minimal free  $Tb^{3+}$  is present in EDTA chelates. Binding constant for EDTA- $Tb^{3+}$  is in  $10^{-15}$  M range; for RNA- $Tb^{3+}$ , in  $10^{-6}$  M range.

### **1: Change of $Tb^{3+}$ luminescence lifetime**

In order to calculate energy transfer efficiency by tr-FRET (equation 2.6),  $\tau_D$  [lifetime of  $Tb^{3+}$  bound on RNA luminescence in the absence of FRET acceptor] should be constant. However,  $Tb^{3+}$  luminescence lifetime with RNA has been found to be variant. A solution has to be found in order to overcome this problem.

### **2: Quantum yield of $Tb^{3+}$ may change so that Förster distance changes**

The Förster distance is calculated in equation 2.2. The shape of the  $Tb^{3+}$  luminescence does not change upon binding on RNA, thus, the spectral overlap is fixed. The refraction index of water is assumed to be 1.33. The relative orientation for  $Tb^{3+}$  based FRET donor is simple because the

rotation of  $Tb^{3+}$  is free, thus the direction of luminescence is random. Orientation factor  $\kappa^2$  is assumed to be 2/3 based on equation 2.3. Thus, in the case of  $Tb^{3+}$  bound on RNA, the Förster distance  $R_0$  is only related to their quantum yield (QY). The QY, however, has not been experimentally tested. Experiments will be done to prove whether it is constant.

### **3. Diffusional energy transfer may interfere with energy transfer measurements**

Protein-bound  $Tb^{3+}$  was first used as diffusional energy transfer donors <sup>79</sup>. Diffusional energy transfer makes the apparent distance between FRET donor and acceptor to be the smallest distance they can get during the diffusional process. The  $Tb^{3+}$  binding groups in this class of FRET donors have moderate binding affinity and long lifetimes. There will be always a considerable fraction of free  $Tb^{3+}$ .

The effect of diffusion on energy transfer depends on  $D\tau_0/s^2$ , where  $D$  is the sum of the diffusion coefficients of the donor and acceptor;  $\tau_0$  is the lifetime of the donor in the absence of acceptors;  $s$  is the mean distance between donors and acceptors. When  $D\tau_0/s^2 \ll 1$ , the energy transfer is static and diffusional energy transfer is very limited <sup>79</sup>. However, because of the presence of free  $Tb^{3+}$  and long lifetimes,  $D\tau_0/s^2$  may be bigger than 1. In this case, there would be considerable diffusional energy transfer. The diffusional energy transfer could be too strong to make measuring the static energy transfer between the bound  $Tb^{3+}$  and acceptor impossible.

#### **3.1.5 Specific aims**

A systematic study of these issues will broaden our understanding about RNA metal ion interactions and encourage the practice of using lanthanide ion luminescence to probe RNA metal ion interactions. Proper conditions for  $Tb^{3+}$  bound on RNA as practical FRET donors will provide a guide for the new approach of this FRET method.

1) I will study the origin of data inaccuracy and propose a compensation factor to adjust luminescence intensity so that it can accurately represent  $Tb^{3+}$  bound on RNA concentration.

2) I will attempt to find out the degree of quenching for lanthanide ion luminescence in aqueous solution by comparing luminescence profiles in  $H_2O$  and  $D_2O$ . In  $D_2O$ , luminescence can be considered as unquenched. From the result, I will propose a linear relationship between the preexponential amplitude and the “chelated”  $Tb^{3+}$  concentration. I will also propose a model to explain the process of  $Tb^{3+}$  and RNA interaction and luminescence quenching.

3) I will attempt to find proper conditions to limit the effect of variation of  $Tb^{3+}$  luminescence lifetime and find a way to overcome it.

4) QY of  $Tb^{3+}$  bound on a GUAA loop RNA will be calculated under different conditions to see whether it is fixed.

5) I will attempt to find the concentration ranges for both  $Tb^{3+}$  and RNA so that the diffusional energy transfer is at a minimal level.

### **3.1.6 Model system: a GUAA RNA tetraloop**

Loop sequences in RNA provide a means of chain reversal within RNA molecules, act as nucleation sites, and participate in essential specific tertiary interactions<sup>80</sup>. Loop structure is stabilized by base stacking and the formation of hydrogen bonds involving base functionalities, ribose 2'OH, or backbone non-bridging phosphate oxygen atoms<sup>81</sup>. Metal ions may also interact with exposed chemical groups to further stabilize the loop structure.

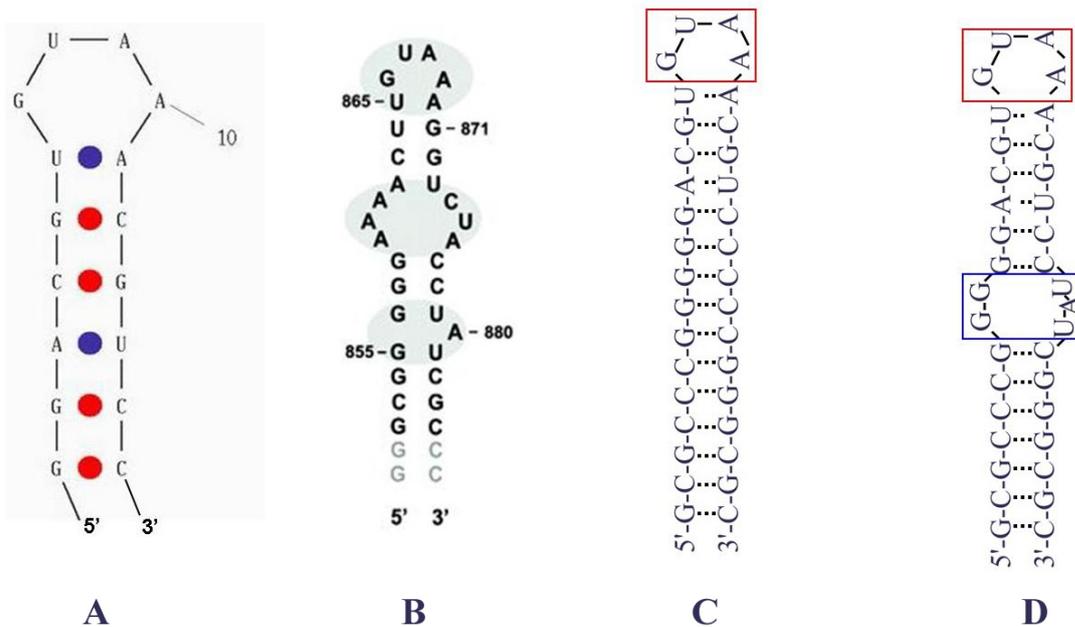
The GNRA hairpin loops (where N is any base and R is a purine) are a highly conserved family in both prokaryotes and eukaryotes<sup>82</sup>. They may participate in loop-loop interactions or metal ion coordination. They have unusual stability from base stacking and hydrogen bonding network formed by the four nucleotides. The metal ion binding properties of the GNRA have been well studied by lanthanide ion luminescence<sup>50, 61, 72</sup>. Even though  $Tb^{3+}$  binds to the loop differently from  $Mg^{2+}$ , it is still useful as a model system.

GUAA loop is also present in splicing RNA systems. Group II intron is an example. As we discussed in chapter one, domain 5 is very important for groups II intron splicing catalysis just as U6 ISL to spliceosome. Domain 5 contains a GUAA loop (red area labeled with  $\zeta'$  in Fig 1.6) and it has interactions with a region in domain 1 (red area labeled with  $\zeta$  in Fig 1.6).

## **3.2 Materials and Methods**

### **3.2.1 Construct design**

Three RNA constructs were designed. Their sequence choices were partially from the upper part of group II intron D6 (Fig 3.2B). The GUAA construct (Fig 3.2A) contained a GUAA loop and six base pairs; the dsGUAA construct (Fig 3.2C) contained the GUAA loop and 15 Watson-Crick base pairs; the bpGUAA construct (Fig 3.2D) contained the GUAA loop and a bulged Watson-Crick



**Figure 3.2 Three RNA sequences used in this chapter and its comparison with the sequence of ai5Y group II intron. A:** the GUAA RNA construct contains a loop and six base pairs. The number indicates sequence from 5'- of this construct. **B:** the sequence of ai5Y group II intron D6<sup>83</sup>. The numbers indicate sequence in original group II intron. **C:** the dsGUAA construct which contains Watson-Crick base pairs and the GUAA loop (red circled). **D:** the bpGUAA construct with contains the GUAA loop and base paired stem from dsGUAA except a branch site resembling the one from group II intron in B.

base pairs. The bulged region was inserted in the sequence to gain information about how a bulge which resembles group II intron branch site alters the orientation of the neighboring helices and whether the new FRET donor can measure small structural changes within an RNA molecule. Cy3 is labeled at the 5'-end of both the dsGUAA and bpGUAA constructs when mentioned. The RNA strands were purchased from Dharmacon (Lafayette, CO).

### **3.2.2 Tb<sup>3+</sup> luminescence intensity measurements**

A Tb<sup>3+</sup> stock solution was made by dissolving Tb<sub>2</sub>O<sub>3</sub> in HClO<sub>4</sub> (both purchased from Sigma Aldridge). Concentration was measured by titration with EDTA as described previously<sup>50</sup>. Tb<sup>3+</sup> has minimal luminescence in water solution as we discussed earlier. Upon binding to RNA, its luminescence is greatly enhanced by means of water displacement and RNA antenna effect (sensitized excitation). Buffer solution contains 10 mM MES, 150 mM NaCl, pH 6.1. The setting on the Cary Eclipse fluorescence spectrometer for steady state Tb<sup>3+</sup> luminescence is as follows: data mode phosphorescence; delay time 200 μs; gate time 9 ms; slit 5/5nm; PMT 800V; room temperature ~20°C; excitation 270nm; emission 300-800nm.

Lanthanide ions can hydrolyze RNA molecules nonspecifically at high concentrations, pH and temperatures<sup>84</sup>. In this study, Tb<sup>3+</sup> up to 150 μM did not hydrolyze the RNA samples under the described conditions in the timeframe of the experiments (data not shown).

### **3.2.3 Tb<sup>3+</sup> luminescence lifetime measurements**

Lifetime measurements were carried out on a Cary Eclipse fluorophotometer. RNA and titrated Tb<sup>3+</sup> concentrations were varied as described in the context. Photomultiplier tube (PMT) power was 800 V if not specified otherwise. The excitation xenon flash lamp pulse had a width of 2-3 μs at 280 nm and excitation/emission slits were 10 nm. Tb<sup>3+</sup> luminescence decay was measured at 545nm was recorded. Delay time was 150 μs. Decay curves were fit with OriginLab using least-square-goodness-of-fit algorithm. The decay equation is equation 2.7.

In this chapter, all exponential decays were monoexponential. Thus, the lifetimes were presented without discussing the data fitting parameters. Refer to Fig 2.6 for fitting procedures.

### **3.2.4 QY of Tb<sup>3+</sup> bound on RNA measurements**

RNA samples used were unlabeled GUAA, dsGUAA and bpGUAA constructs. Similar results were obtained for all three RNAs. Alexa Fluor 555 (AF555, from Invitrogen, same spectral

shape as Cy3) was used as the acceptor. RNA between 1~10  $\mu\text{M}$  was mixed with free AF555 from 2~20  $\mu\text{M}$ ;  $\text{Tb}^{3+}$  concentration between 1 to 20  $\mu\text{M}$  was titrated. The equation for QY calculation is equation 3.12. intensity and lifetime were measured by the protocol in Xiao, Selvin 2001<sup>85</sup>.

### 3.3 Results

#### 3.3.1 Steady state signal accuracy: compensation factor

There are two major differences between fluorescence and phosphorescence steady state instrumentation. First, fluorescence has shorter lifetimes (ns) than phosphorescence (ms or longer). Second, fluorescence is stronger than phosphorescence. Even though  $\text{Ln}^{3+}$  luminescence is neither fluorescence nor phosphorescence as discussed in section 3.1.2, it is measured the same way as phosphorescence. For steady state fluorescence measurements, there are two types of excitation light source: continuous and pulsed. If a constant illumination is used, steady state of the fluorescence is reached immediately upon excitation and recorded. If a pulsed source is used, with an open emission window, the recorded intensity is the same as using continuous excitation:

$$\int_{0 \rightarrow \infty} I_0 * e^{-t/\tau} = -I_0 \tau e^{-t/\tau} = I_0 \tau \quad (\text{Equation 3.2})$$

Where I is the intensity;  $I_0$  is the preexponential amplitude;  $\tau$  is the lifetime; t is the time.

For phosphorescence, however, a pulsed source has to be used because phosphorescence has longer lifetimes and it is usually much weaker. In order to allow short lifetime components and background to fade away (Fig 3.3A; the width of the excitation peak is comparably very small), a delay time is used. The delay time is usually 100~150  $\mu\text{s}$ , but as short as 15  $\mu\text{s}$  was built<sup>41</sup>.

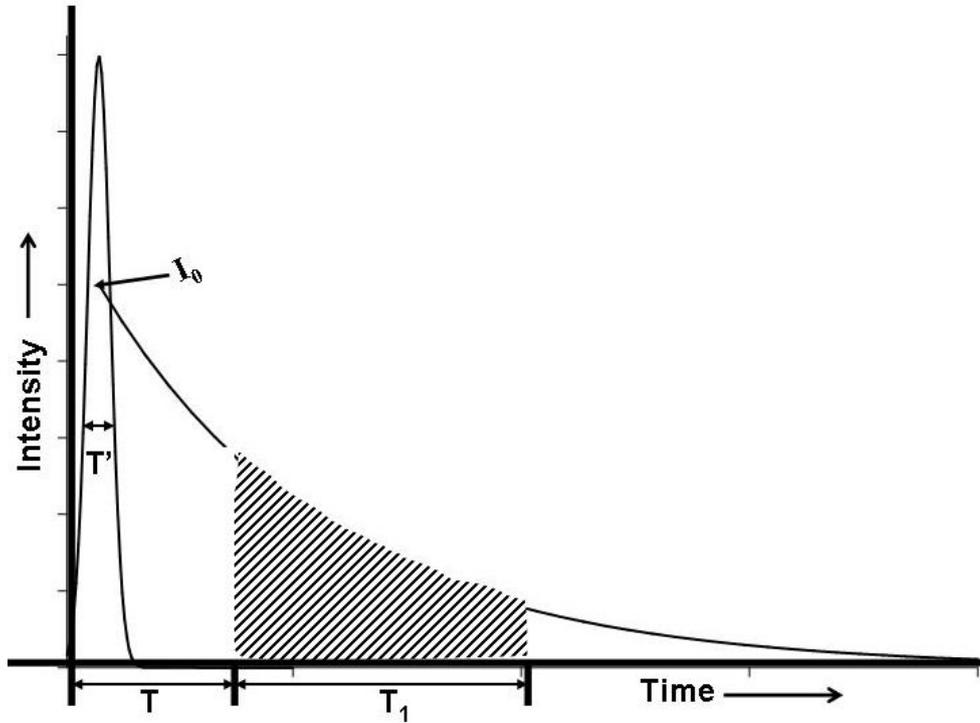
The steady state recorded for lanthanide ion luminescence is the integral of the shadowed area in Fig 3.3A. Let us compare two hypothetical scenarios. In Fig 3.3B and 3.3C, the gray area is the signal recorded; the black area is the signal not recorded. Using the same delay time 0.1 ms, apparently from the figures, a luminescence with 0.1 ms lifetime lost most of its signal during the signal recording process (Fig 3.3B) while a luminescence with 1 ms lifetime lost only a small fraction (Fig 3.3C).

The measured intensity which is only the grey area in fig 3.3A can be calculated:

$$\int_{T \rightarrow T_2} I_0 * e^{-t/\tau} = -I_0 \tau e^{-t/\tau} = I_0 \tau e^{-T/\tau} - I_0 \tau e^{-T_2/\tau} \quad (\text{Equation 3.3})$$

Figure 3.3

A.



B.

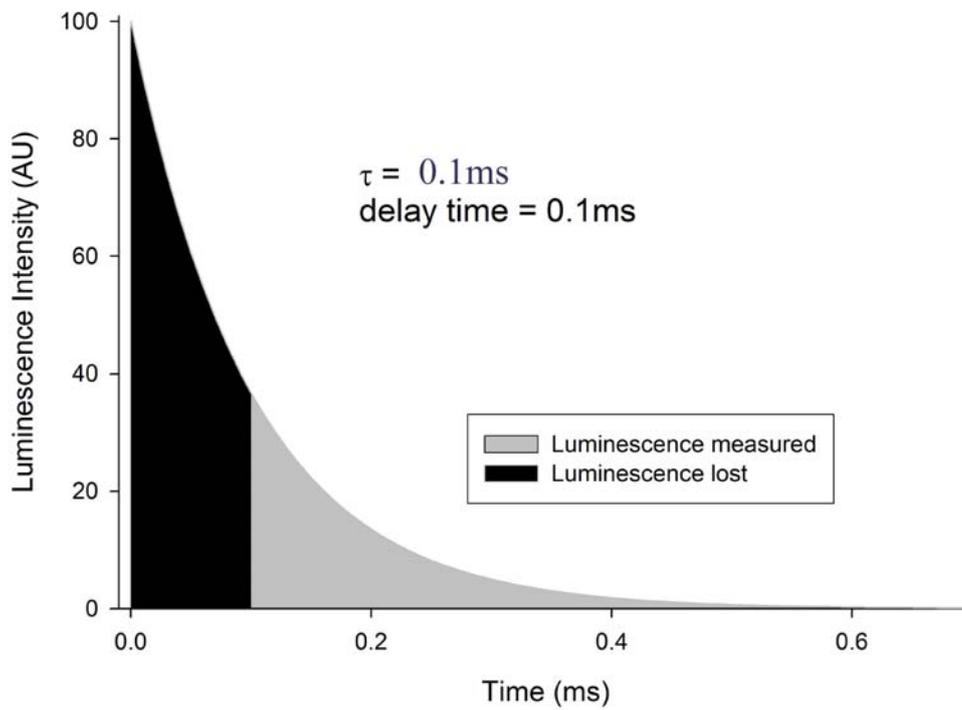
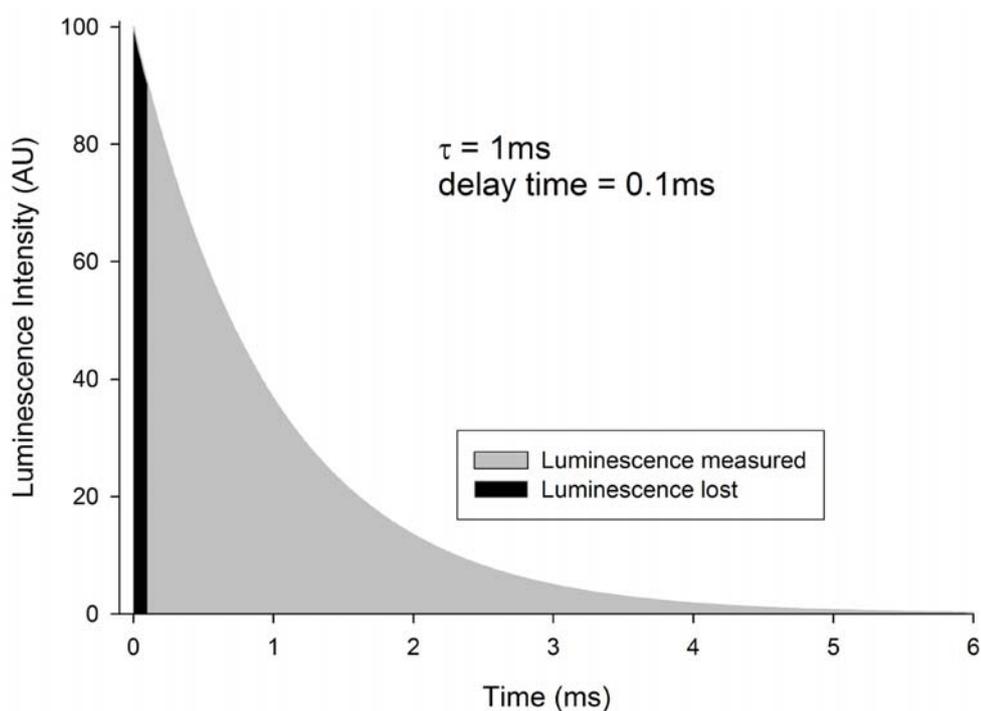


Figure 3.3-continued

C.



**Figure 3.3 Lanthanide luminescence signal collection with a pulsed excitation.** A) Schematic diagram of luminescence excitation and recording.  $T'$  is the width of excitation source;  $T$  is the delay time which allows shorter components to die out;  $T_1$  is the detection gate time;  $I_0$  is the maximal luminescence intensity. B) An example of lifetime 0.1 ms; delay time 0.1 ms. Black area is the signal lost due to the delay time. Grey area is the signal measured by the instrument. The total luminescence which can represent the actual intensity is both combined. C) An example of lifetime 1 ms; delay time 0.1 ms.

Where T is the gate time, T<sub>2</sub> is the stop time for data collection (9 ms was used in my measurements). Tb<sup>3+</sup> lifetimes are generally much smaller than 9 ms, T<sub>2</sub> >> τ, so:

$$\int_{T \rightarrow T_2} I_0 * e^{-t/\tau} \approx \int_{T \rightarrow \infty} I_0 * e^{-t/\tau} \equiv I_0 \tau e^{-T/\tau} \quad \text{(Equation 3.4)}$$

The measured intensity thus should multiply a compensation factor (CF) in order to represent the real intensity.

$$CF = \frac{\int_{0 \rightarrow \infty} I_0 * e^{-t/\tau}}{\int_{T \rightarrow \infty} I_0 * e^{-t/\tau}} = e^{T/\tau} \quad \text{(Equation 3.5)}$$

If the lifetime (τ) is long or the delay time (T) is short, the compensation factor would be close to 1. In this case, the recorded intensity and the actual intensity are similar. For example, in the case in Fig 3.3C,  $CF = e^{\frac{0.1ms}{1ms}} = e^{0.1} = 1.105$ , lost signal only accounts 9.5%. If the lifetime (τ) and the delay time (T) are comparable, the compensation factor would be big, lost signal will account high percentage. For example, in the case of Fig 3.3B,  $CF = e^{\frac{0.1ms}{0.1ms}} = e^1 = 2.718$ , lost signal is 63.2%.

This inconsistency of lanthanide luminescence intensity data been observed<sup>86</sup>. However, the reasoning behind it was ambiguous. It is pointed out here that for lanthanide luminescence (or phosphorescence) intensity measurements, lifetime data should be available. If lifetime is comparable with the delay time, a compensation factor in equation 3.5 should be used.

### 3.3.2 Degree of quenching: H<sub>2</sub>O vs D<sub>2</sub>O

By combining different concentrations of RNA and Tb<sup>3+</sup> together, seemly random lifetimes were observed (table 3.2, column 3). The phenomenon indicated that excited Tb<sup>3+</sup> luminescence was quenched at different levels. To find out the degree of quenching, parallel RNA and Tb<sup>3+</sup> binding experiments were done in H<sub>2</sub>O and D<sub>2</sub>O.

D<sub>2</sub>O has similar chemical properties as H<sub>2</sub>O but the quenching effects on Tb<sup>3+</sup> luminescence is minimal<sup>77</sup>. By comparing luminescence intensity in H<sub>2</sub>O and D<sub>2</sub>O, the degree of quenching in H<sub>2</sub>O solution can be found out.

20 μM GUAA loop RNA was combined with different concentrations of Tb<sup>3+</sup> (1-100 μM) in

10 mM MES, 150 mM NaCl, pH 6.1 buffer. For D<sub>2</sub>O solution, RNA stock solution was dried down in speed-vacuum, and re-dissolved in D<sub>2</sub>O-MOPS buffer. Tb<sup>3+</sup> stock was obtained by dissolving in D<sub>2</sub>O three times<sup>50</sup>. The lifetime of each sample was fit in OriginLab with monoexponential decay. The results were listed in table 3.2.

Lifetimes in H<sub>2</sub>O were changing: the higher the Tb(III) concentration, the shorter the lifetime as shown in table 3.2. Lifetimes in D<sub>2</sub>O, however, are relatively constant to be  $3.1 \pm 0.1$  ms. This constant lifetime for Tb(ClO<sub>4</sub>)<sub>3</sub> is in agreement of the number reported for Tb(CH<sub>3</sub>COO)<sub>3</sub> in literature<sup>87</sup>.

At first glance, the lifetimes and the intensities seem to have a random relationship. However, when the ratios of intensity and lifetime in H<sub>2</sub>O were compared with their counterparts in D<sub>2</sub>O, a linear relationship was found (Fig 3.4). The  $y = x$  relationship indicated the ratios were the same at the same RNA and Tb<sup>3+</sup> concentration regardless of in H<sub>2</sub>O or D<sub>2</sub>O.

The understanding of this relationship is straightforward. Recall Equation 3.2, the intensity of fluorescence/phosphorescence decay is  $I = I_0 \times \tau$ . The intensities listed in table 3.2 are already compensated using the compensation factor. Thus, the ratio of intensity/lifetime is actually the preexponential amplitude in Equation 3.2. For the same RNA and Tb<sup>3+</sup> concentrations, they are the same regardless of whether in D<sub>2</sub>O or H<sub>2</sub>O.

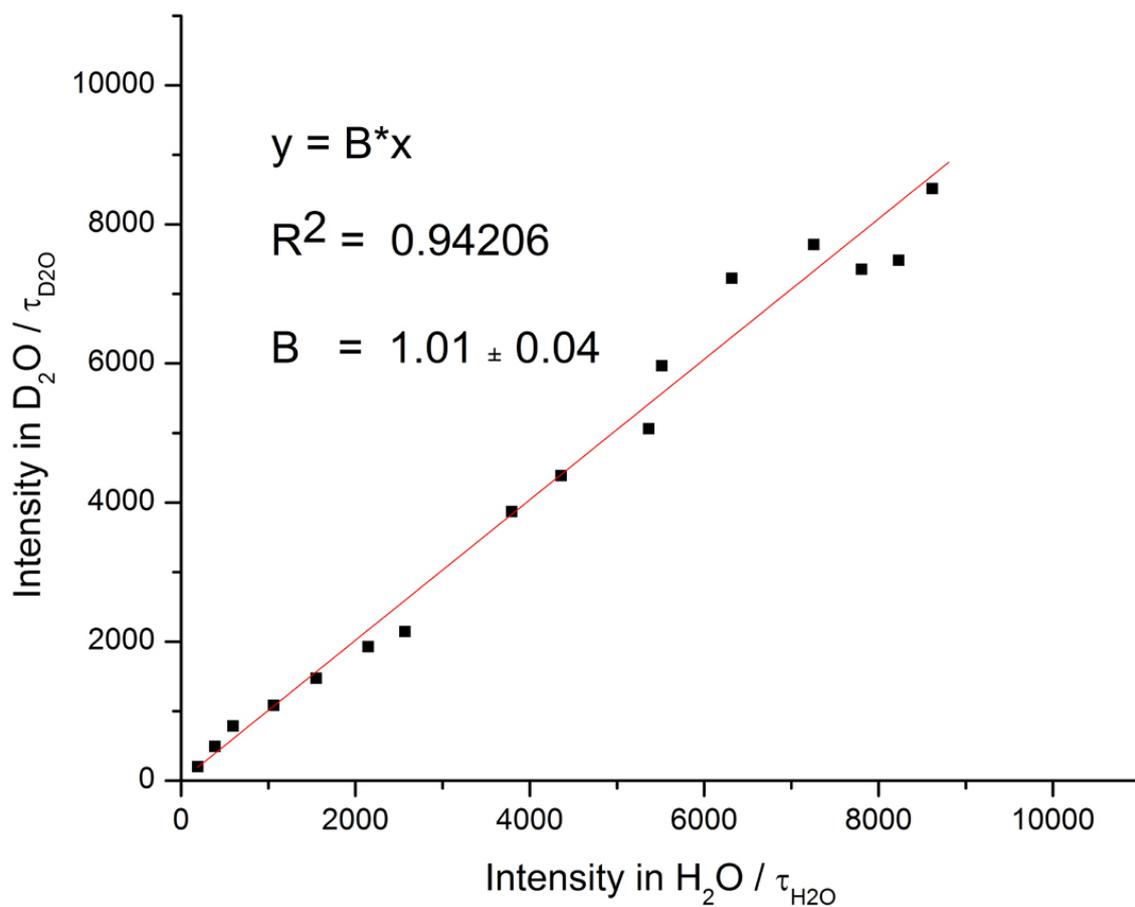
D<sub>2</sub>O does not change the binding properties between RNA and Tb<sup>3+</sup>. The same RNA and Tb<sup>3+</sup> concentrations will result in the same concentration of the RNA-Tb<sup>3+</sup> complex in D<sub>2</sub>O and H<sub>2</sub>O. Thus, the preexponential amplitude is a more accurate way of deriving the concentration of the RNA-Tb<sup>3+</sup> complex.

I propose that this conclusion applies to other fluorescence/phosphorescence measurements as well. This hypothesis is supported by a report which found good agreement between the population data deduced from NMR experiments and those derived from the preexponential amplitudes of the fluorescence decay<sup>88</sup>. As a matter of fact, preexponential amplitudes are routinely used to calculate population fraction in time-resolved FRET experiments<sup>41</sup>. However, it is more important to use the preexponential amplitudes for lanthanide ion luminescence because their lifetimes are changing.

Now let us go back to Equation 3.1. The assumption is not accurate because there is no linear

**Table 3.2 Intensity and lifetime of Tb<sup>3+</sup> luminescence in H<sub>2</sub>O and D<sub>2</sub>O.** Lifetimes were obtained by fitting with monoexponential decay in OriginLab. Intensities have already been compensated using the compensation factor.

[GUAA] ( $\mu$ M)	[Tb <sup>3+</sup> ] ( $\mu$ M)	$\tau_{H_2O}$ (ms)	$\tau_{D_2O}$ (ms)	H <sub>2</sub> O (AU)	D <sub>2</sub> O (AU)
20	1.28	0.570	3.1 $\pm$ 0.1	108	630
	2.56	0.561		218	1530
	3.85	0.553		330	2430
	5.13	0.536		569	3350
	6.41	0.537		832	4570
	8.97	0.531		1140	5970
	11.5	0.522		1340	6650
	14.1	0.520		2790	10100
	16.7	0.517		1960	12000
	19.2	0.530		2310	13600
	24.4	0.515		2840	18500
	29.5	0.505		3190	22400
	34.6	0.492		3570	23900
	39.7	0.483		3770	22800
44.9	0.475	3910	23200		
70.5	0.469	4040	26400		



**Figure 3.4** plot intensity in  $D_2O/\tau_{D_2O}$  vs intensity in  $H_2O/\tau_{H_2O}$ . A data point was from using intensity in  $D_2O/\tau_{D_2O}$  as y axis and intensity in  $H_2O/\tau_{H_2O}$  as x axis at each  $Tb^{3+}$  concentration. By fitting the data in OriginLab using a function of  $y = B \cdot x$ , B was found to be close to one ( $1.01 \pm 0.04$ ).

relationship between the luminescence intensity and the concentration of the luminescence probe, the  $Tb^{3+}$  bound on RNA complex. To further compensate the inaccuracy, in the basis of Equation 3.5, a new compensation factor ( $CF_0$ ) should be used to get the preexponential amplitude:

$$CF_0 = \frac{e^{T/\tau}}{\tau} \quad (\text{Equation 3.6})$$

Where T is the delay time,  $\tau$  is the lifetime of the luminescence decay.

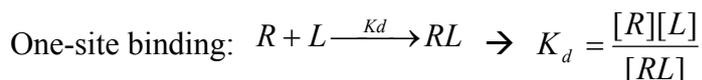
Using this compensation factor, and defining  $[Ln^*]$  as the preexponential amplitude for excited lanthanide ion luminescence, not its concentration, equation 3.1 is now more accurate.

The compensation factors do not abandon the usefulness of intensity measurements. For CF in equation 3.5, if there is no delay time (that is the case for all fluorescence measurements) or delay time is much shorter than the lifetime (instrumentation development), we do not have to consider it because it is close to 1. For  $CF_0$ , if lifetimes do not change in the intensity measurements,  $CF_0$  is constant and does not influence the data accuracy. The results after using compensation ( $CF_0$ ) matched the stoichiometry data reported in chapter two because of the above reasons.

### 3.3.3 Binding constant of the GUAA RNA and $Tb^{3+}$

The GNRA loop is well studied in literature<sup>61, 72</sup> and its binding constant has been studied by using unbiased methods such ITC (Isothermal Titration Calorimetry) to be  $k_d \approx 7 \mu M$ <sup>50</sup>. Stoichiometric measurement of  $Tb^{3+}$  and GUAA loop was carried out on ITC as well with Job plot (details in chapter 2) in the MES buffer. The stoichiometry was found to be 1, indicating one  $Tb^{3+}$  ion to one RNA molecule.

The following calculation is to deduce a formula to calculate binding constant for a single site. For simplicity, RNA,  $Tb^{3+}$  and RNA- $Tb^{3+}$  complex are short as R, L and RL; C is the concentration of total RNA; x is the concentration of total  $Tb^{3+}$ ; B is a conversion factor (see equation 3.1); y is the compensated luminescence intensity; binding constant  $k_d$  is short as k:



$$[R] = c - \frac{y}{B} \quad \text{and} \quad [L] = x - \frac{y}{B} \quad \therefore k = \frac{(C - \frac{y}{B}) * (x - \frac{y}{B})}{\frac{y}{B}}$$

Solve y as a function of x  $\rightarrow$

$$y = \frac{B}{2} * (C + x + k - \sqrt{(C + x + k)^2 - 4Cx}) \quad (\text{Equation 3.7})$$

Different concentrations of RNA (from 1-20  $\mu\text{M}$ ) and  $\text{Tb}^{3+}$  (1-100 $\mu\text{M}$ ) were mixed together in MES pH 6.1 buffer. Both luminescence intensity and corresponding lifetimes were measured. By plotting y values against x values with known C values in OriginLab software, k value was solved.

Y values were compensated with the compensation factor  $CF_0 = \frac{e^{T/\tau}}{\tau}$ . In Fig 3.5, uncompensated and partial compensated results were compared.

The  $k_d$  was found to be  $k_d = 9.01 \pm 1.08 \mu\text{M}$ . This value was close to the value reported in literature by ITC ( $k_d \approx 7 \mu\text{M}$ )<sup>61</sup> even though the sequences were different. For uncompensated data points,  $K_d = 6.83 \pm 3.16 \mu\text{M}$ ; for  $CF = e^{T/\tau}$  compensated data points,  $k_d \approx 7.71 \pm 2.71 \mu\text{M}$ . Using the compensation factor improved the fit. The three binding constants did not differ too much because first, if lifetime did not change too much, compensation factor for each point was similar; and binding constant results would be similar too. Second, lifetimes was long compared to the delay time (0.4-0.6 ms vs 0.15 ms), the impact of compensation was limited.

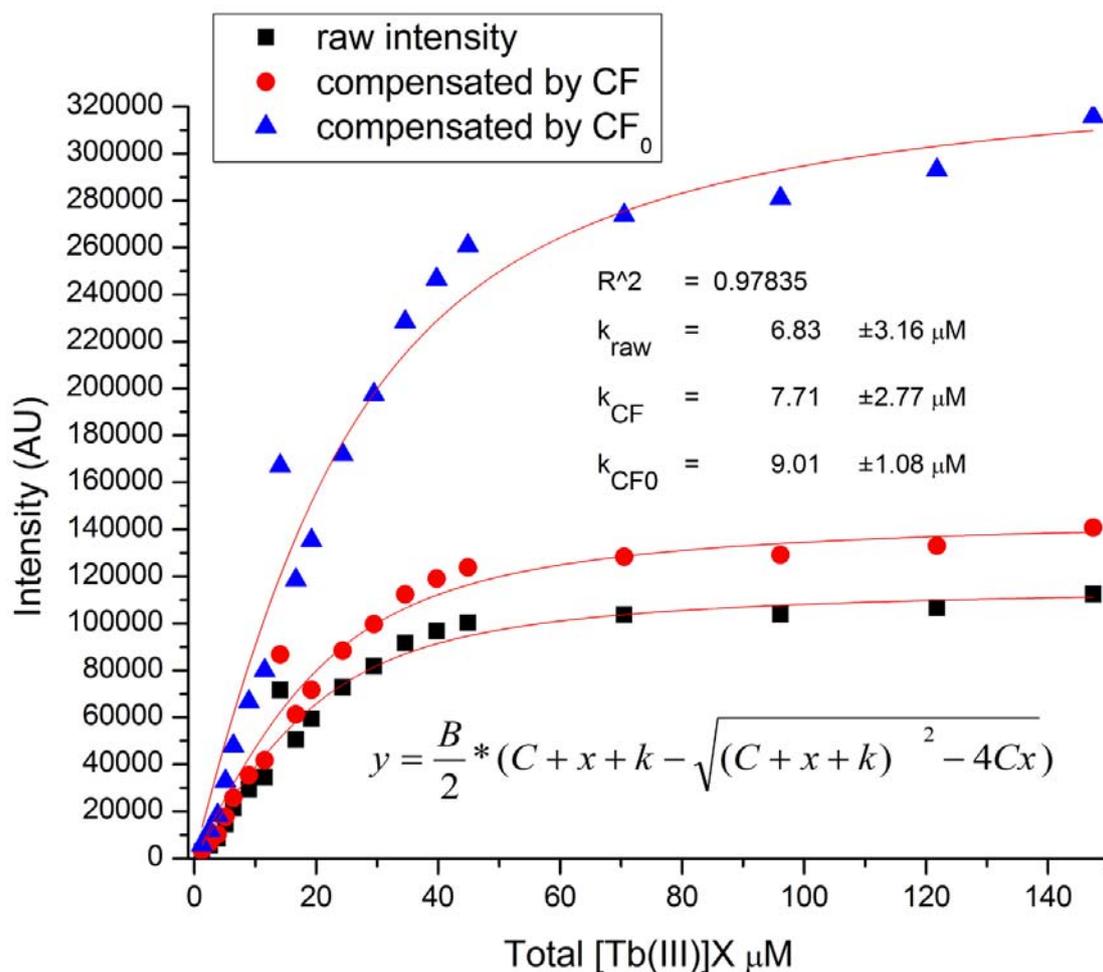
### 3.3.4 $\text{Tb}^{3+}$ luminescence lifetime and free $\text{Tb}^{3+}$ fraction

In the last three sections, we drew the conclusion that there was a linear relationship (positive slope) between the preexponential amplitude and RNA- $\text{Tb}^{3+}$  complex concentration; and a linear relationship (negative slope) between the degree of luminescence quenching and lifetimes. The next question we asked was what determined the lanthanide ion luminescence lifetimes? In other words, what controlled the degree of the lanthanide ion luminescence quenching by  $\text{H}_2\text{O}$ ?

A set of lifetime data were obtained by mixing different concentrations of GUAA construct and  $\text{Tb}^{3+}$  in MES buffer. The concentrations of the RNA ranged from 1 to 20  $\mu\text{M}$ ; and  $\text{Tb}^{3+}$  ranged from 1 to 100 $\mu\text{M}$ . Decay data were analyzed by fitting with OriginLab. All exponential decays were monoexponential. These lifetimes ranged from 420 to 580  $\mu\text{s}$  (table 3.3).

The measurable values chosen for each measurement were: concentration of total  $\text{Tb}^{3+}$  and concentration of total  $\text{Tb}^{3+}$ . The following parameters were calculated by using  $k_d = 9 \mu\text{M}$  and equation 3.7: concentration of bound  $\text{Tb}^{3+}$ ; concentration of free  $\text{Tb}^{3+}$ ; concentration of free RNA; fraction of free RNA out of total RNA; fraction of free  $\text{Tb}^{3+}$  out of total  $\text{Tb}^{3+}$  (table 3.3).

By plotting lifetimes at the Y axis and other measurable values at the X axis, the only



**Figure 3.5 Binding constant calculations.** Three different colored shapes represent raw intensity data, CF normalized data and CF<sub>0</sub> normalized data. A non-linear least square fitting on the listed equation in Origin Lab was used. The compensation processes increased binding constant accuracy. Bigger changes are predicted if lifetimes are significantly different. Binding constant was 9.01±1.08 μM.

relationship we could establish was a linear relationship between lifetimes and fractions of free  $Tb^{3+}$  out of total  $Tb^{3+}$  (Fig 3.6).

$$\tau = 660 \times (1 - A) + 427 \times A \text{ } (\mu s) \text{ (Equation 3.8)}$$

Here,  $\tau$  is  $Tb^{3+}$  luminescence lifetime;  $A$  is the fraction of free  $Tb^{3+}$  out of total  $Tb^{3+}$ .  $(1-A)$  thus is the fraction of bound  $Tb^{3+}$  out of total  $Tb^{3+}$ .

I hypothesized that the 660  $\mu s$  lifetime component was the  $Tb^{3+}$  bound on RNA; 427  $\mu s$  was the lifetime of  $Tb^{3+}$  in free aqueous solution. This was supported by a report that  $Tb^{3+}$  luminescence lifetime in buffer was measured at about 430  $\mu s$ <sup>86</sup>.

Similar experiments were carried out with the dsGUAA and bpGUAA constructs. They gave a relationship similar to equation 3.8. To broaden the equation 3.8, a general relationship was:

$$\tau = \tau_1 \times (1 - A) + \tau_2 \times A \text{ (Equation 3.9)}$$

Where  $\tau$  is the apparent lifetime of lanthanide ion luminescence;  $A$  is the fraction of free  $Tb^{3+}$  the same as equation 3.8;  $\tau_1$  and  $\tau_2$  represent the luminescence lifetime of  $Tb^{3+}$  bound on RNA and  $Tb^{3+}$  in free aqueous solution, respectively.

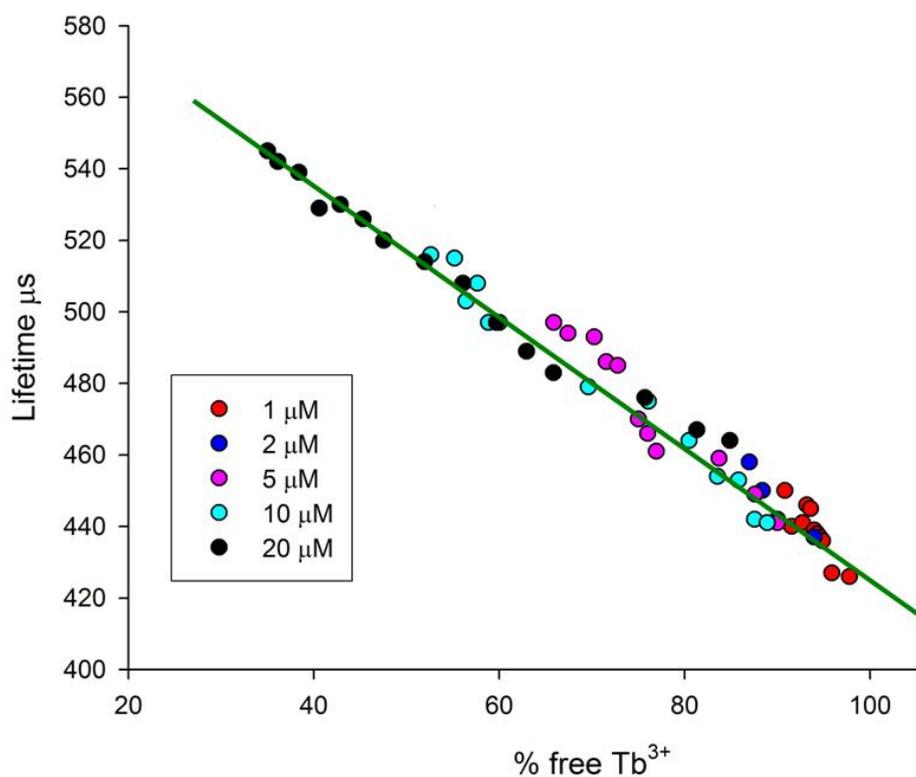
This result suggested that the lanthanide ion luminescence lifetime was linearly determined by the fraction of free  $Tb^{3+}$ . In section 3.3.2, it was clear that the degree of quenching is linearly related to its lifetime. Thus, the degree of quenching was linearly related to fraction of free  $Tb^{3+}$ .

From above results, a “diffusional equilibrium” theory is proposed for the process of RNA- $Tb^{3+}$  binding: only  $Tb^{3+}$  “chelated” bound on RNA can be sensitized by RNA (section 2.1.2). After the pulse excitation, excited  $Tb^{3+}$  exchanges with free  $Tb^{3+}$  during its long luminescence lifetime. The equilibrium is quickly reached so that at any give time after excitation, based on the fractions of bound  $Tb^{3+}$   $(1-A)$  and free  $Tb^{3+}$   $(A)$ , the fraction  $(1-A)$  of the total  $Tb^{3+}$  stays on RNA without further getting quenched by  $H_2O$ . The fraction  $(A)$  of the total  $Tb^{3+}$  leaves the “chelated” site and gets quenched further by more  $H_2O$ . The fraction is determined by Boltzmann’s distribution law.

The measured exponential decay of  $Tb^{3+}$ -RNA luminescence lifetime data is single exponential decay, not double lifetimes containing two components of bound  $Tb^{3+}$  and free  $Tb^{3+}$ . This suggests that during the lifetime of luminescence, bound and free  $Tb^{3+}$  exchange at a fast rate and the excited  $Tb^{3+}$  are indistinguishable by displaying a uniform lifetime.

**Table 3.3 Lifetime data from mixing different concentrations of RNA and Tb<sup>3+</sup>.** RNA and Tb<sup>3+</sup> concentrations were total. Lifetimes were obtained by fitting in OriginLab single exponential decay; concentrations of the RNA-Tb<sup>3+</sup> complex, free Tb<sup>3+</sup> and free RNA, percentage of free RNA and Tb<sup>3+</sup> were all calculated based on  $K_d = 9 \mu\text{M}$  in equation 2.7.

[RNA] $\mu\text{M}$	[Tb(III)] $\mu\text{M}$	Lifetime $\mu\text{s}$	[complex] $\mu\text{M}$	Free RNA $\mu\text{M}$	free RNA %	Free Tb $\mu\text{M}$	free Tb %
1	1	450	0.0916	0.908	90.8	0.908	90.8
	5	446	0.341	0.658	65.8	4.65	93.1
	8	438	0.455	0.544	54.4	7.54	94.3
	9	437	0.486	0.513	51.3	8.51	94.5
	10	436	0.513	0.486	48.6	9.48	94.8
	15	427	0.615	0.384	38.4	14.3	95.8
	36	426	0.796	0.203	20.3	35.2	97.7
2	3	479	0.442	1.557	77.8	2.55	85.2
	5	458	0.651	1.348	67.4	4.34	86.9
	7	450	0.814	1.185	59.2	6.18	88.3
	10	442	1.00	1.00	50.0	9.00	90.0
	23.6	437	1.42	0.577	28.8	22.1	93.9
5	1	497	0.341	4.65	93.1	0.658	65.8
	5	486	1.42	3.57	71.5	3.57	71.5
	10	461	2.30	2.69	53.9	7.69	76.9
	20	459	3.25	1.74	34.9	16.7	83.7
	30	449	3.72	1.27	25.5	26.2	87.5
	40	441	4.00	1.00	20.0	36.0	90.0
	50	431	4.17	0.820	16.4	45.8	91.6
10	1	550	0.513	9.48	94.8	0.486	48.6
	5	530	2.30	7.69	76.9	2.69	53.9
	10	497	4.00	6.00	60.0	6	60.0
	20	479	6.07	3.92	39.2	13.9	69.6
	50	454	8.22	1.77	17.7	41.7	83.5
	60	453	8.51	1.48	14.8	51.4	85.8
	70	442	8.71	1.28	12.8	61.2	87.5
	80	441	8.87	1.12	11.2	71.1	88.9
20	1.3	572	0.883	19.1	95.5	0.416	32.0
	9	539	5.54	14.4	72.2	3.45	38.3
	11.5	529	6.83	13.1	65.8	4.66	40.5
	19.2	520	10.0	9.92	49.6	9.12	47.5
	34.6	497	13.9	6.06	30.3	20.6	59.7
	44.9	483	15.3	4.66	23.3	29.5	65.8
	70.5	476	17.1	2.88	14.4	53.3	75.7
	121.8	464	18.3	1.60	8.00	103	84.8
147.4	463	18.6	1.30	6.53	128	87.3	



**Figure 3.6 Lifetimes against percentage of free Tb<sup>3+</sup>.** Lifetimes were obtained by fitting on OriginLab single exponential decay. Percentage of free Tb<sup>3+</sup> out of total Tb<sup>3+</sup> was calculated. Y axis were the lifetimes in μs; X axis were against percentage of free Tb<sup>3+</sup>. Different RNA concentrations (1, 2, 5, 10, 20 μM) were labeled with different colors.

### 3.3.5 Fix the lifetime of Tb<sup>3+</sup> bound on RNA

The luminescence lifetime of Tb<sup>3+</sup> bound on RNA changes with free Tb<sup>3+</sup> fraction as stated in equation 3.9. In order to calculate accurate energy transfer efficiency by time-resolved FRET (tr-FRET) when using Tb<sup>3+</sup> bound on RNA as FRET donors, accurate lifetime should be obtained. In other tr-FRET experiments, the lifetime of donors in the absence of energy transfer is generally fixed. Data inaccuracy will occur if same practice is used as for other tr-FRET measurements.

However, to measure luminescence lifetime at each matching concentration of Tb<sup>3+</sup> with unlabeled RNA and labeled RNA is tedious. During Tb<sup>3+</sup> titrations, it is tedious to match every data point for the measurements in the absence of acceptors and in the presence of acceptors.

The solution to this problem is to calculate lifetime. By knowing the binding constant, RNA total concentration, Tb<sup>3+</sup> total concentration for a tr-FRET condition in the presence of acceptors, free Tb<sup>3+</sup> fraction can be calculated by using equation 3.7. If we know  $\tau_1$  and  $\tau_2$ , lifetimes for that condition can be calculated by using equation 3.9.

Thus, in order to do tr-FRET experiments using Tb<sup>3+</sup> bound on RNA as FRET donors, some preparation experiments should be done. The first parameter to obtain is the binding constant. The procedure is discussed in details in section 3.3.3. The other two parameters needed are  $\tau_1$  and  $\tau_2$  in equation 2.9, which can be measured as discussed in section 3.3.4.

Generally tr-FRET experiments do not depend on concentrations, but exact RNA and Tb<sup>3+</sup> concentrations have to be recorded when using Tb<sup>3+</sup> bound on RNA as tr-FRET donors.

The limited number of RNA based FRET donors will allow us to measure their binding constants,  $\tau_1$ , and  $\tau_2$ . Accurate lifetimes can be calculated based on a calculation.

### 3.3.6 QY of Tb<sup>3+</sup> bound on RNA

In section 3.3.2, we draw the conclusion that the concentration of excited Tb<sup>3+</sup> can be represented by its preexponential amplitude. However, this conclusion is based on an invariant QY. When QY is considered, equation 3.6 should be revised as:

$$cf = \frac{e^{-T/\tau}}{\tau * QY} \quad (\text{Equation 3.10})$$

Let us discuss the relationship between the QY of donor and acceptor.  $I_{DA}$  is the intensity of donor in the presence of acceptor;  $I_{AD}$  is the intensity of acceptor in the presence of donor;  $\tau_{DA}$  is the

lifetime of donor in the presence of acceptor;  $\tau_{AD}$  is the lifetime of acceptor in the presence of donor;  $\tau_D$  is the lifetime of donor in the absence of acceptor;  $Q_a$  is the QY of acceptor;  $Q_D$  is the QY of donor.

Number of excited acceptors via energy transfer is:  $\frac{I_{AD}}{\tau_{AD}Q_a}$ ; number of remaining donors

is  $\frac{I_{DA}}{\tau_{DA}Q_D}$ ; Energy transfer efficiency

$$E = \frac{\frac{I_{AD}}{\tau_{AD}Q_a}}{\frac{I_{AD}}{\tau_{AD}Q_a} + \frac{I_{DA}}{\tau_{DA}Q_D}} \quad (\text{Equation 3.11})$$

This provides a new way of measuring QY of a fluorophore if its FRET pair counterpart QY is known. In lanthanide ion FRET experiments,  $\tau_{AD} = \tau_{DA}$ , where  $\tau_{AD}$  is generally measured without the interference of donor (Fig 2.1)<sup>40</sup>. Thus, equation 3.11 can also be rewritten as:

$$E = \frac{\frac{I_{AD}}{Q_a}}{\frac{I_{AD}}{Q_a} + \frac{I_{DA}}{Q_D}}$$

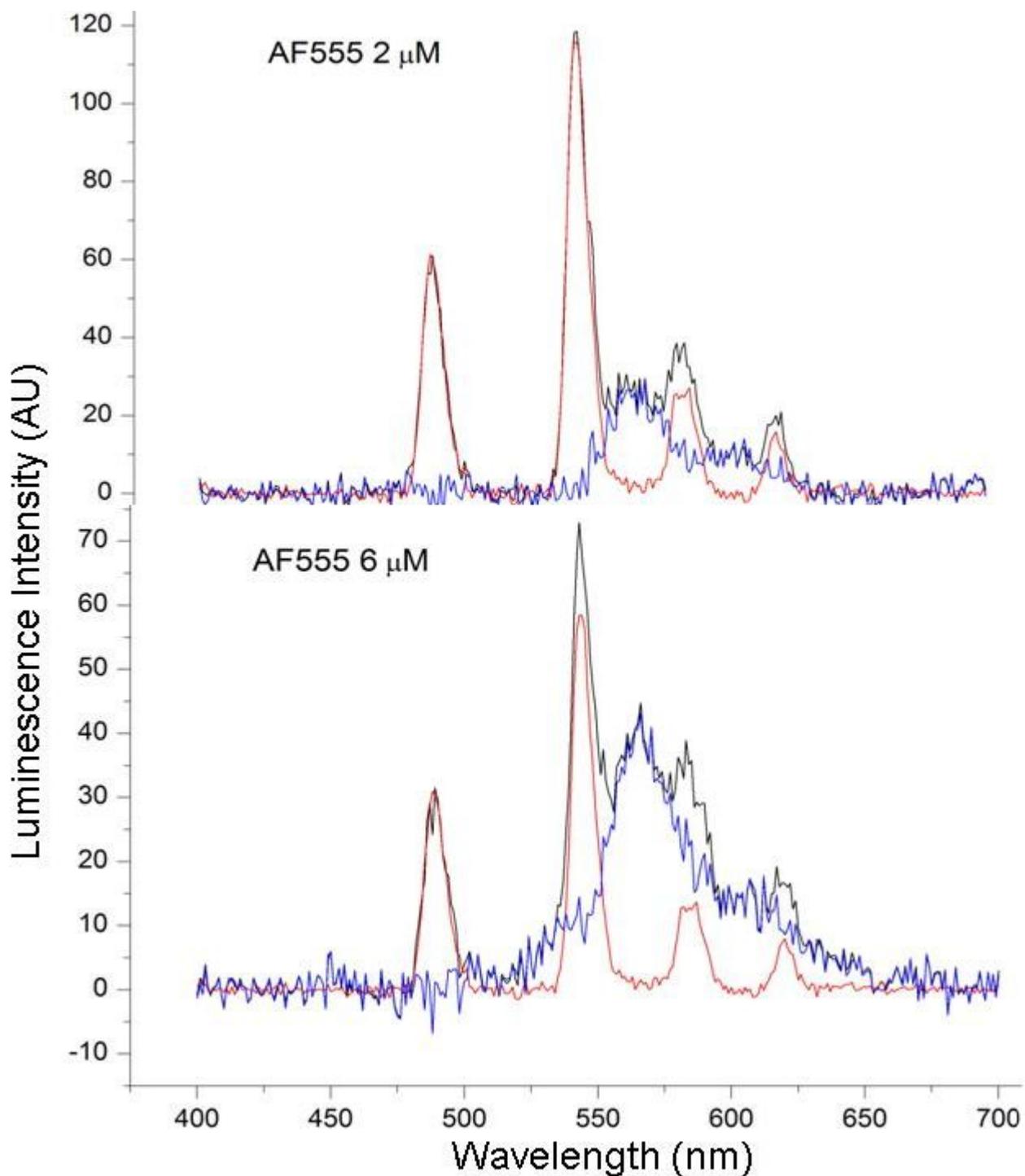
Using  $E = 1 - \frac{\tau_{AD}}{\tau_D}$ ,  $Q_D$  can be calculated as:

$$Q_D = Q_a * \frac{I_{DA}}{I_{AD}} * \frac{\tau_D - \tau_{AD}}{\tau_{AD}} \quad (\text{Equation 3.12}).$$

Where  $Q_D$  is the QY of  $Tb^{3+}$  bound on RNA;  $Q_a$  is the QY of the acceptor;  $I_{DA}$  is the donor emission in the presence of acceptor;  $I_{AD}$  is the acceptor emission in the presence of donor;  $\tau_D$  is the donor lifetime in the absence of acceptor;  $\tau_{AD}$  is the lifetime of acceptor in the presence of donor.

This is exactly the same equation as in Xiao and Selvin 2001, in which they used different method of deducing the equation<sup>85</sup>.

In each experiment, the fraction of free  $Tb^{3+}$  was calculated based on equation 3.7. The lifetime of donor ( $\tau_D$ ) for each experimental condition was calculated with equation 3.8.  $Tb^{3+}$



**Figure 3.7 Quantum yield of Tb<sup>3+</sup> bound on RNA luminescence.** The black trace is the spectrum of both Tb<sup>3+</sup> and AF555. The red trace is the spectrum of Tb<sup>3+</sup> by itself. It is normalized by using peak @ 490nm and a standard Tb<sup>3+</sup> luminescence spectrum. The blue trace is the spectrum of AF555 by itself. It is normalized by deducting the red trace from the black trace.

luminescence spectra with unlabeled RNA were standard Tb<sup>3+</sup> luminescence curve. The peak at 490 nm at the curve is used to normalize the luminescence only from Tb<sup>3+</sup> luminescence. By excluding the normalized Tb<sup>3+</sup> luminescence from total spectra, spectra of AF555 spectrum are obtained. The lifetime measured at 565nm is the lifetime of acceptor in the presence of donor ( $\tau_{AD}$ ). Figure 3.7 showed of the process of extracting parameters with 2 and 6  $\mu$ M AF555.

QY for free AF555 is  $Q_a = 0.1$  (from Molecular Probes, Invitrogen);  $I_{DA}$  is the integral intensity under normalized Tb<sup>3+</sup> spectrum;  $I_{AD}$ ; is the integral intensity under AF555 spectrum.

Using equation 3.12, in the concentration range, QY of Tb<sup>3+</sup> was found to be constant:

$$QY_{RNA\ bound\ Tb^{3+}} = 0.14$$

$R_0$  was calculated by equation 2.2 to be 50 Å based on this quantum yield. The constant quantum yield indicated that quantum yield of Tb<sup>3+</sup> luminescence was only dependent on the RNA-Tb<sup>3+</sup> binding environment. The exchange process did not change the quantum yield.

### 3.3.7 Diffusional energy transfer process

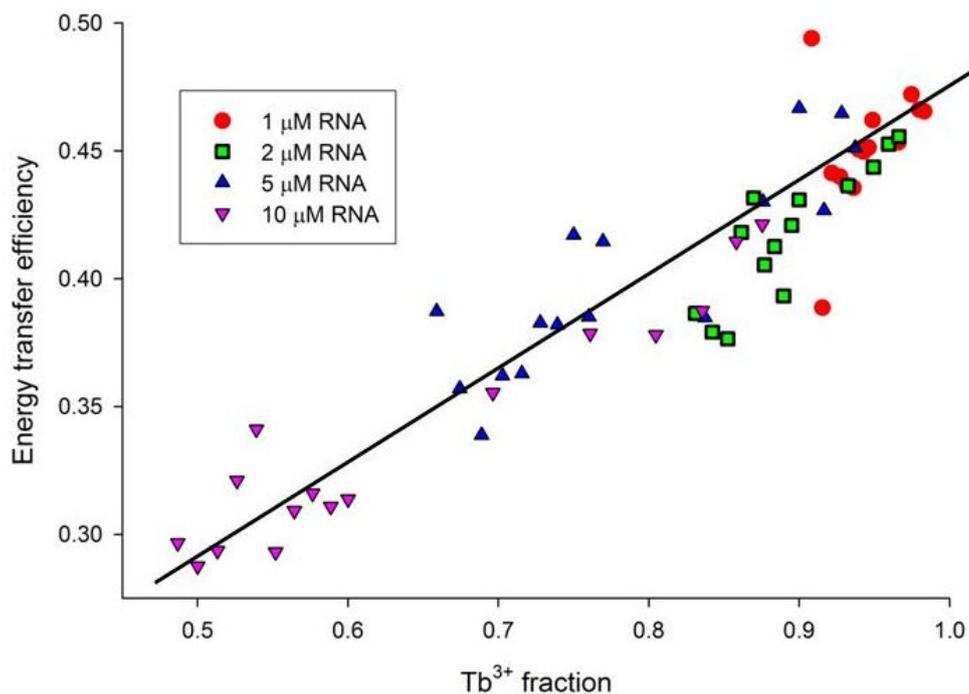
When both donor and acceptor are free, and lifetime of donor is long,  $D\tau_0/s^2 \gg 1$ , the energy transfer is dominated by diffusional energy transfer. In this section, I want to test whether the free Tb<sup>3+</sup> interferences the diffusional energy transfer processes.

The first question is to answer whether free Tb<sup>3+</sup> contributes to diffusional energy transfer at a fixed donor concentration. In 1  $\mu$ M RNA and 3  $\mu$ M AF55 sample, different concentrations of Tb<sup>3+</sup> were titrated. Free Tb<sup>3+</sup> concentration and fraction were calculated based on the binding constant. Lifetime of Tb<sup>3+</sup> by itself ( $\tau_D$ ) was calculated from equation 3.8. Lifetime of the acceptor in the presence of donor ( $\tau_{DA}$ ) is measured at 565 nm (Cy3 and AF55 spectra are identical). The energy transfer efficiency is calculated by equation 2.6.

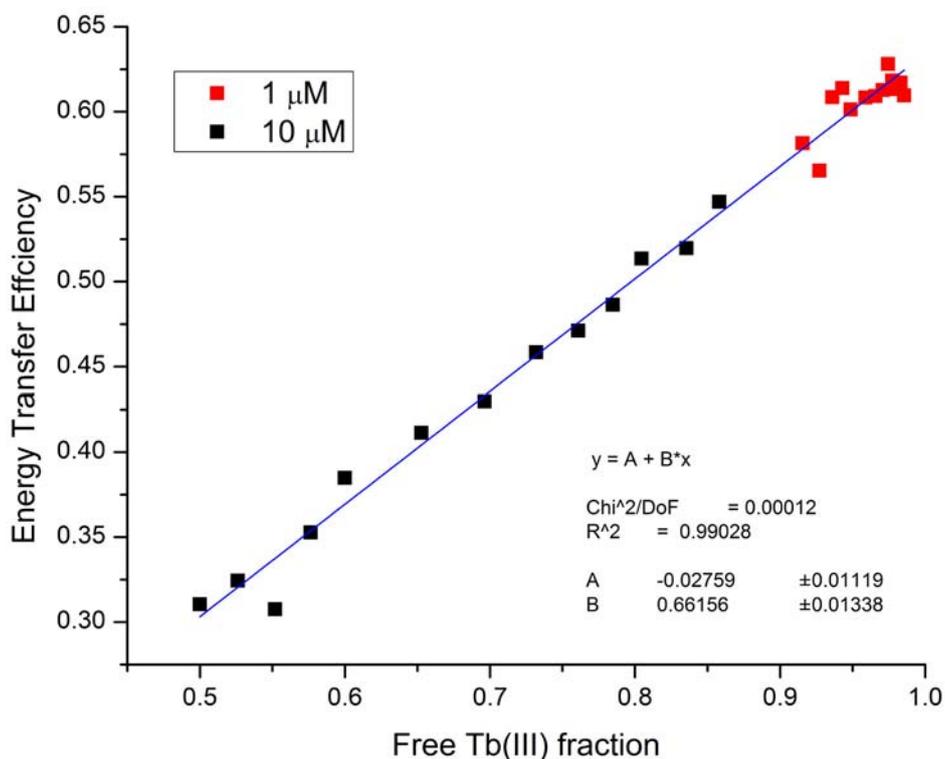
Same experiments except different RNA concentrations (2, 5, 10  $\mu$ M) were done.

If free Tb<sup>3+</sup> has no contribution to diffusional energy transfer, calculated concentration of Tb<sup>3+</sup> bound on RNA will have a correlation with the energy transfer efficiency. However, the only correlation I found was that the energy transfer efficiency and free Tb<sup>3+</sup> percentage has a linear relationship. The fitted equation was:  $E = 0.45 \times A$  (Figure 3.8A), where A is the fraction of free Tb<sup>3+</sup>. Using the same experiments except with 5  $\mu$ M AF 555 (Fig 3.8B), the relationship was  $E = 0.66 \times A$ .

A



B.



**Figure 3.8 Free  $Tb^{3+}$  has an impact on diffusional energy transfer efficiency. A)** Diffusional energy transfer efficiency with 3  $\mu M$  AF 555 in the solution. Best fit equation is  $E = 0.45 \times A$ . **B)** With 5  $\mu M$  AF555.  $E = 0.66 \times A$ . Where E is the energy transfer efficiency; A is the free  $Tb^{3+}$  fraction.

These two equations provide us two conclusions. For the FRET donor  $Tb^{3+}$ , diffusional energy transfer efficiency is dependent on the fraction of free  $Tb^{3+}$ ; the more fraction of free  $Tb^{3+}$ , the higher the energy transfer efficiency (dependent on the fraction A). For the FRET acceptor, increased acceptor concentration increases the energy transfer efficiency (0.45 and 0.66).

In rapid diffusional limit, energy transfer occurs when donor/acceptor are at the closest distance<sup>79</sup>. Thus, the acceptor AF555 contribution can be explained solely by its concentration. The higher concentration, the closer they get.

If free  $Tb^{3+}$  does not contribute to diffusional energy transfer, we can predict that the efficiency will only dependent on the concentration of  $Tb^{3+}$  bound on RNA. From the two equations, free  $Tb^{3+}$  contributes to the diffusional energy transfer in a linear fashion. Free  $Tb^{3+}$  will always be in the system because the binding constant is relatively low. Thus, to minimize the contribution of free  $Tb^{3+}$  to diffusional energy transfer, some measures have to been used.

The diffusional energy transfer efficiency will be at a certain level, which is determined by the free  $Tb^{3+}$  fraction and acceptor concentration. If the diffusional energy transfer efficiency is lower than the static energy transfer resulted from fixed distance between donor and acceptor, diffusional energy transfer will have no impact on the outcome the static energy transfer will dominate the signal. On the other hand, if the diffusional energy transfer efficiency is higher than the energy transfer efficiency resulted from the static energy transfer, the diffusional energy transfer will dominate the energy transfer signal and static energy transfer will be hard to measure.

To reduce the contribution of diffusional energy transfer, two methods can be utilized to overcome the problem. One way is to lower acceptor concentration. This can be achieved by using low concentration of acceptor labeled RNA. The other way is to lower the fraction of free  $Tb^{3+}$ . This can be achieved by decreasing titrated  $Tb^{3+}$  concentration. It has to be noted that the better way to study diffusional energy transfer effect on fixed-distance energy transfer is to use a model system with both diffusional energy transfer and fixed-distance energy transfer. The current model system only includes diffusional energy transfer. However, the experimental results reveal ways of minimizing diffusional energy transfer. In our recent publication<sup>48</sup>, FRET measurements were carried out with low concentrations of labeled RNA (Cy3 labeled RNA concentration = 0.7  $\mu$ M) and low  $Tb^{3+}$  concentrations. These two measures minimized diffusional energy transfer. Thus, diffusional energy transfer did not interference with static energy transfer.

### 3.3.8 Structure distortion by a group II intron branch site

The branch site contains an adenosine which is one reactant of the first step of splicing (Fig 1.3). The structure of the branch site in both spliceosome<sup>27</sup> and group II intron<sup>83</sup> suggested that in order for the adenosine to attach the 5'-ss, the 2'OH of the adenosine has to be pushed out of the solution. Here we designed two constructs; one contains (dsGUAA) a GUAA loop and all Watson-Crick base pairs helix and one containing a GUAA loop and a similar helix inserted with the branch site. The GUAA loop acted as a FRET donor when Tb<sup>3+</sup> bound. By comparing these two RNAs, we can not only test the validity of the concept of using Tb<sup>3+</sup> bound on RNA as FRET donors, but also gain some insights of how the branch site changes the orientation of the helix.

Cy3 was attached to the 5'-end of each construct. RNA concentrations were kept between 1-5  $\mu$ M; Tb<sup>3+</sup> concentration were kept between 1-20 $\mu$ M.

**dsGUAA:** the energy transfer efficiency was measured by tr-FRET (equation 2.6).  $\tau_D$  was calculated for each concentration;  $\tau_{DA}$  was measured at 565 nm. E was constant for both RNA and Tb<sup>3+</sup> ranges to be 77 $\pm$ 1% (table 3.4). Using equation 2.1 and  $R_0=50$  Å, R was calculated to be 41 $\pm$ 1Å. The dsGUAA contained 15 base pairs between Cy3 and GUAA tetraloop, and Cy3 was about an extra base pair long, predicated A-type helix was 2.6 Å \*(15+1) = 41.6 Å. The experimental result was very consistent with the A-type helix predication.

**bpGUAA:** the energy transfer efficiency was measured to be 83 $\pm$ 1%; R was calculated to be 38 $\pm$ 1Å. Compared to dsGUAA, even though bpGUAA was one nucleotide longer, the length of the helix was about 3 $\pm$ 2 Å shorter. This indicates that introducing a branch point to the helix disrupt the local regular Watson-Crick double helix. There is a bend at the branch site. We can triangulate the angle of the bend. The side containing Cy3 was 8 nucleotides, its length was (1+8)\*2.6=23.4 Å; the side containing the GUAA loop was 8 nucleotides, its length was 8\*2.6=20.8 Å; the side opposite to the angle was 38 Å. If we assume the helical structure was not disturbed, the bend angle could be triangulated to be 120°. However, a NMR structure showed that the branch site alone did not create an extrahelical branch site but the local helix got broader and shorter<sup>89</sup>. Thus, the insertion of this branch point introduced a bend >120°.

These results also show that using Tb<sup>3+</sup> bound on RNA as FRET donor can measure small energy transfer efficiency changes because of the distance's sixth power dependence.

**Table 3.4 Energy transfer efficiencies measured in constructs dsGUAA and bpGUAA.** RNA and Tb<sup>3+</sup> concentrations were initial concentrations before mixing. The lifetimes of donor in the absence of acceptor ( $\tau_D$ ) are calculated by using  $K_d = 9 \mu\text{M}$  and equation 3.8. The lifetimes of donor in the presence of acceptor ( $\tau_{DA}$ ) are measured at 565 nm. Energy transfer efficiency is calculated with equation 2.6.

Construct	[RNA] ( $\mu\text{M}$ )	[Tb <sup>3+</sup> ] ( $\mu\text{M}$ )	(A) Free Tb <sup>3+</sup> fraction	Calculated $\tau_D$ ( $\mu\text{s}$ )	Measured $\tau_{DA}$ $\mu\text{s}$ at 565 nm	E (%)
dsGUAA	1	1	0.908	448	105	76.5
		5	0.932	443	102	77.0
		10	0.949	439	107	75.6
		15	0.959	437	108	75.3
		20	0.966	435	100	77.1
	5	1	0.659	506	116	77.0
		5	0.716	493	110	77.8
		10	0.770	481	115	76.0
		15	0.809	472	111	76.5
		20	0.837	465	107	77.0
bpGUAA	1	1	0.908	448	75.8	83.1
		5	0.932	443	77.5	82.5
		10	0.949	439	75.5	82.8
		15	0.959	437	73.8	83.1
		20	0.966	435	74.4	82.9
	5	1	0.659	506	86.6	82.9
		5	0.716	493	83.8	83.0
		10	0.770	481	85.6	82.2
		15	0.809	472	79.2	83.2
		20	0.837	465	74.4	84.0

## 3.4 Discussion

In this chapter, two aspects regarding  $Tb^{3+}$  bound on RNA were studied. The one aspect is to study the  $Tb^{3+}$  bound on RNA properties. The results will provide a foundation if  $Tb^{3+}$  bound on RNA to be used as FRET donors. The results here also may improve data accuracy of using  $Tb^{3+}$  as probes for studying RNA metal ion bindings. The other aspect is to study the suitability of  $Tb^{3+}$  bound on RNA as FRET donors. The results could provide a guidance of how to carry out this type of experiments.

### 3.4.1 Compensation factor

The assumption in equation 3.1, which is practiced widely, is found not to be accurate. The preexponential amplitude shown as  $I_0$  in both equations 2.7 and 3.2 is a more accurate value to represent the concentration of the excited lanthanide ion. In order to make steady state intensity data accurate, they must multiply a compensation factor shown in equation 3.6.

In situations such as no delay times, comparable short delay times and long lifetimes, and constant stable lifetimes, the compensation factor is mathematically constant. The steady state intensity, still not theoretically accurate, can represent the concentration of excited lanthanide ion (RNA “chelated”). In this case, only the normalization factor (B) in equation 3.1 is changed.

Thus, in order to get concentration information from RNA lanthanide ion samples by luminescence method, steady state is not enough to be accurate. One of the reasons that lanthanide ion luminescence intensity assay is used to gain concentration information is because of its simplicity. Fortunately, lifetime measurements for lanthanide ion luminescence are relatively simple because of its long lifetime. For example, all the lifetime measurements in this chapter were done on a commonly used Cary Eclipse fluorophotometer. Preexponential amplitude can also be found by lifetime measurements alone by lifetime decay measurements.

The linear relationship between the preexponential amplitude and the excited lanthanide concentration is only true for a specific binding species under specific environments. Different RNA  $Tb^{3+}$  pair, buffer conditions and instrument setup may change the preexponential amplitude.

### 3.4.2 “Diffusional equilibrium” hypothesis

When  $Tb^{3+}$  “chelated” binds to RNA and loses one or more inner-sphere water molecules, some of its emission energy is still quenched by remaining inner-sphere water molecule. The degree

of quenching can be calculated by comparing the lifetime in D<sub>2</sub>O. Emission energy can be expressed as  $E = I_0 \times \tau$ . Without being quenched in D<sub>2</sub>O the energy is  $E' = I_0 \times \tau_{D_2O}$ , with quenching in H<sub>2</sub>O the energy is  $E'' = I_0 \times \tau_{H_2O}$ . The degree of quenching thus is  $\frac{\tau_{D_2O} - \tau_{H_2O}}{\tau_{D_2O}}$ . Since lifetime in D<sub>2</sub>O is constant, the degree of quenching is linear (negative slope) with its lifetime.

Some of the emission energy is further lost because Tb<sup>3+</sup> luminescence is further decreased by adding more Tb<sup>3+</sup>. A linear relationship was found by plotting the lifetimes against the fraction of free Tb<sup>3+</sup> (equation 3.9).  $\tau = \tau_1 \times (1 - A) + \tau_2 \times A$ .

From these results, a “diffusional equilibrium” theory about the Tb<sup>3+</sup> ion excitation and quenching process when it “chelated” binds a RNA molecule is proposed:

Only “chelated” Tb<sup>3+</sup> is excited at the time of the excitation pulse (280nm) because of RNA aromatic groups’ sensitized energy transfer is in a short range. Free Tb<sup>3+</sup> at the time is only excited at a minimal level because of its low molar absorptivity. After the pulse, equilibrium is reached almost immediately between excited and unexcited Tb<sup>3+</sup> because of its long luminescence lifetime. For excited Tb<sup>3+</sup>, some will stay “chelated”, displaying  $\tau_1$  lifetime. At any given time, their fraction is (1-A) because of random diffusion. Others will go “free” and further get quenched by H<sub>2</sub>O, displaying  $\tau_2$  lifetime. Their fraction at any given time is (A). The bound Tb<sup>3+</sup> species has a maximal lifetime in H<sub>2</sub>O for a certain Tb<sup>3+</sup>-RNA pair. For example, GUAA loop and Tb<sup>3+</sup> has  $\tau_1=660\mu s$ . Free Tb<sup>3+</sup> species has a minimal lifetime in H<sub>2</sub>O for a certain lanthanide ion in buffer. For example, Tb<sup>3+</sup> in MES buffer has  $\tau_2=427\mu s$ .

Energy (represented by intensity  $E = I_0 \times \tau$  in equation 3.2) and concentration (represented by preexponential amplitude  $I_0$ ) are two easily confusing concepts. It may be helpful to clarify their differences by using the “diffusional equilibrium” theory on equation 3.9.

Assume  $I_0$  is the preexponential amplitude (proportional to bound Tb<sup>3+</sup> concentration); A,  $\tau$ ,  $\tau_1$ , and  $\tau_2$  are described earlier. E,  $E_1$ ,  $E_2$  are energies described below within,  $I_1$  and  $I_2$  are preexponential amplitudes described below within.

After excitation and equilibrium, the overall emission energy is:  $E = I_0 \times \tau$ , where  $\tau$  is the apparent lifetime. There are two components of Tb<sup>3+</sup> after the equilibrium. One is the “chelated”

Tb<sup>3+</sup>, its energy is:  $E_1 = I_1 \times \tau_1$ , (1-A) is the fraction of “chelated” Tb<sup>3+</sup>, so  $I_1 = (1 - A) \times I_0$  because of the diffusional equilibrium. The other is free Tb<sup>3+</sup>, its energy is  $E_2 = I_2 \times \tau_2$ , where  $I_2 = I_0 \times A$ , A is the fraction of free Tb<sup>3+</sup>. The overall energy should be the total of its two components:  $E = E_1 + E_2$ . Equation 3.9 thus is deduced.

### 3.4.3 Suitable conditions for Tb<sup>3+</sup> bound on RNA to be FRET donors

Free Tb<sup>3+</sup> is always in RNA-Tb<sup>3+</sup> binding system because of their relative weak binding constant (in 10<sup>-6</sup> M range). Three potential problems emerge because of this. The first is that the lifetime of Tb<sup>3+</sup> bound on RNA is not fixed. This lifetime is directly related to the calculation of energy transfer efficiency (equation 2.6). The solution to this problem is to record concentrations of RNA and Tb<sup>3+</sup> for each measurement. Lifetimes can be accurately calculated by using binding constant and equation 3.9.

The second concern is that the quantum yield may not be fixed. This is not a concern anymore because the result shows that quantum yield is indeed fixed.

The third is the contribution of diffusional energy transfer. Diffusional energy transfer could be overwhelming because of the presence of free Tb<sup>3+</sup>. Low acceptor (labeled RNA) and low Tb<sup>3+</sup> concentrations are recommended to minimize the diffusional energy transfer.

### 3.4.4 Using Tb<sup>3+</sup> bound on RNA as FRET donors

Tb<sup>3+</sup> bound on RNA was utilized as FRET donors in the U2-U6 snRNA complex model system. The experiments were successful and the Tb<sup>3+</sup> binding sites as well as their pH-dependence in the U2-U6 snRNA complex were found<sup>48</sup>. This is the first time an unchelated lanthanide ion is used as a FRET donor. Lanthanide ions were previously used as FRET donors in chelated forms<sup>40</sup>. In addition to the advantages of luminescence resonance energy transfer (LRET)<sup>41</sup>, such as long lifetimes, relatively bigger R<sub>0</sub>, high signal-to-noise ratio, using Tb<sup>3+</sup> bound on RNA as FRET donors allows us to measure relatively strong lanthanide ion binding sites directly without altering its native structure or mutating chemical groups.

The enhancement of Tb<sup>3+</sup> luminescence when “chelated” binding to RNA makes them attractive FRET donors. The reason for the enhancement, as we discussed in chapter two, is that RNA aromatic bases are able to absorb and transfer energy from incoming excitation photons to

their neighboring  $Tb^{3+}$ . The remove of part or all of  $Tb^{3+}$  inner-sphere water molecules, whose O-H bonds quench  $Tb^{3+}$  luminescence <sup>75</sup>, is not necessary a requirement for the enhancement. Inner-sphere water molecules decrease luminescence intensity by decreasing  $Tb^{3+}$  luminescence lifetimes. However, they do not change the preexponential amplitude. Thus, the quenching is not too severe. The enhancement of luminescence, however, is always accompanied by the removal of inner-sphere water molecules because  $Tb^{3+}$  has to be in the proximity of RNA aromatic groups (see Fig 1.9 for cartoon illustration).

There are two possible far-reaching implications associated with the success of our work. One application is that one could map out relatively strong metal ion binding sites in target molecules not only limited to RNA molecules, but also proteins. Many RNAs can “chelated” bind with  $Tb^{3+}$  and enhance their luminescence <sup>35</sup>. Some metallo-proteins also show the enhancement of specific bound  $Tb^{3+}$  luminescence. One example is the EF hand motifs <sup>90</sup>. After all, metal ion binding and aromatic amino acids are common in many proteins. By using similar approaches as used in chapter three, “native”  $Tb^{3+}$  binding sites can be mapped out. The word “native” here does not mean  $Tb^{3+}$  binding occurs in biological conditions; it means  $Tb^{3+}$  can bind to these sites without modifications.

### **3.4.5 Possibility of designing a new class of FRET donors**

The other possible application is to develop a new class of FRET donors. These FRET donors contain two components, a  $Tb^{3+}$  ion and a  $Tb^{3+}$  binding group. The  $Tb^{3+}$  binding group can be an organic molecule, a RNA motif, a DNA loop or a protein domain. The only requirement is that  $Tb^{3+}$  should have relatively tight binding with  $Tb^{3+}$  and the binding results in enhanced  $Tb^{3+}$  luminescence.

In this sense, Selvin’s application of using EDTA-chelated  $Tb^{3+}$  as FRET donors is one approach of the second possibility <sup>41</sup>. The  $Tb^{3+}$  binding groups contain two parts, an EDTA-like chelate group and an antenna aromatic group. They bind to  $Tb^{3+}$  with excellent affinity and luminescence enhancement. The drawback is that they have to be covalently linked to the target molecules.

The candidates for the  $Tb^{3+}$  binding group part of the FRET donors are DNA, RNA or protein. When the  $Tb^{3+}$  binding group is incorporated into its respective target molecule at a carefully

chosen location, a FRET donor is created without covalent coupling. For example, a RNA based  $Tb^{3+}$  binding group can be incorporated into a carefully selected site on a RNA molecule by convenient ways of making RNA, such *in vitro* transcription or chemical synthesis; a DNA based  $Tb^{3+}$  binding group can be incorporated into a carefully selected DNA sequence; a protein based  $Tb^{3+}$  binding group can be incorporated into a protein sequence. The technology advancements of making these biological molecules help us incorporate these  $Tb^{3+}$  binding groups more easily than chemical modification.

This class of donors contains two components,  $Tb^{3+}$  and a  $Tb^{3+}$  binding group.  $Tb^{3+}$  has poor absorption coefficient, thus, it is not a practical FRET donor by itself. Via sensitized excitation, lanthanide ion luminescence can be enhanced 4-5 orders of magnitude. The enhancement make  $Tb^{3+}$  attractive FRET donors.

The advantages of using “native” specific metal ion binding sites as FRET donors are two fold. On one hand, metal ion binding sites can be mapped out on a molecule as our previous practice<sup>48</sup>. On the other hand, protein or RNA based  $Tb^{3+}$  binding groups can be designed. They can be conveniently inserted into target molecules.

The search for the lanthanide luminescence enhancement by DNA on Pubmed does not yield any results, suggesting using lanthanide ions to study DNA metal ion interaction is rare. The reason is that the predominately double helix forms and little structural variability of DNA. Using  $Tb^{3+}$  to probe RNA and protein metal ion interactions, however, are common. For example, a series of  $Tb^{3+}$  binding peptides were developed<sup>91</sup>. In this chapter, only RNA based  $Tb^{3+}$  binding group is discussed.  $Tb^{3+}$  is used as a representative of  $Tb^{3+}$ .

With the advancement in aptamers (oligonucleotides or peptide molecules that bind a specific target molecule) research, new strong  $Tb^{3+}$  binding groups can be found.

The two major disadvantages are: 1) as we discussed, tr-FRET experiments for  $Tb^{3+}$  and  $Tb^{3+}$  binding group with moderate binding affinity are relatively tedious. Normal tr-FRET experiments do not depend on concentrations but the information about concentrations and binding constant is required here. 2) The insertion should be carefully chosen so that it does not introduce unwanted structural changes. However, these can be easily overcome.

# CHAPTER 4

## MEASUREMENT OF METAL ION-INDUCED CONFORMATIONAL CHANGE IN THE SPLICEOSOMAL U2-U6 snRNA COMPLEX

### 4.1 Introduction

#### 4.1.1 Motivation of research: folding hypothesis

The sequences of U2 and U6 snRNAs are most conserved phylogenetically among all spliceosomal snRNAs<sup>19,92</sup>. U2 snRNA contains an intramolecular stem loop called U2 ISL<sup>93</sup>; U6 snRNA contains a conserved U6 ISL<sup>94</sup>. In addition to these intramolecular helices, the two snRNAs form several conserved intermolecular helices, including helix I<sup>20</sup>, helix II<sup>20</sup> and helix III<sup>93</sup>. Several sequences, such as the ACAGAGA loop, the AGC triad and the U74 on U6 ISL, are highly conserved<sup>23-26</sup>. These conserved secondary structures are shown in Fig 1.4. A protein-free U2-U6 snRNA complex is able to carry out a reaction similar to the first step of splicing<sup>22,31</sup> (Fig 1.7).

Most conserved regions are essential for splicing catalytic activity. For example, Helix I may be involved in 5' splice site selection<sup>95</sup>, and it may also have redundant function as Helix II as deletion of helix I or helix II does not disrupt splicing but deletion of both does<sup>96</sup>. Helix III may help juxtapose the 5' splice site and the branch site<sup>93</sup>. The U6 ISL is required for spliceosomal catalytic activity; its structural feature is required for splicing but not its sequence<sup>97</sup>; it forms a structure similar to group II intron D5 (Fig 1.6)<sup>66</sup>. It not only contains a possible metal ion binding site as suggested earlier, but also is likely to participate in spliceosomal assembly and activation<sup>98</sup>. The invariant ACAGAGA loop, the AGC triad and the U74 on U6 ISL are important regions for specific metal ion binding as discussed in Chapter 2.

We hypothesize that in order for the U2-U6 snRNA complex to perform splicing catalysis; these key areas have to be brought into proximity. The secondary structure as shown in Fig 1.4

could not indicate any tertiary interactions. Thus, its conformation should be induced by some factors to arrange the key components into their active positions.

#### **4.1.2 Possible factors that may induce the conformational change**

Metal ions are required for spliceosome activity *in vivo* and *in vitro*<sup>17</sup>; that is also the case for many ribozymes<sup>99</sup>. Metal ions were also shown to induce UV-cross linking between distant nucleotides in cross-linking experiments. Considering the absolute requirement for splicing activity, as well as proven the specific binding sites on the U2-U6 snRNA complex<sup>48</sup>, metal ions are the primary candidates which could induce a conformational change to arrange key components into proximity in the U2-U6 snRNA complex. The metal ion we studied is Mg<sup>2+</sup>, the ion most likely to be associated with physiological function of RNA.

Another factor considered is the intron strand. One of the models about protein enzyme and substrate interaction is “induced fit”. Substrates induce protein enzymes to change their conformations to better fit with each other<sup>100</sup>. The intron substrate may induce the U2-U6 snRNA complex to conformational changes because the branch-site adenosine; the 5'-splice site and the 3'-splice site of the intron directly interact with U2-U6 snRNA complex (Fig 1.5). Structural study regarding the pairing between intron branch site and U2 showed that the intron branch site adenosine, by pairing with U2 snRNA, was exposed into the solution. This expose favored the first step of splicing. A conserved pseudouridine in U2 snRNA was required for the extrahelical position of the branch site adenosine<sup>27, 28, 101</sup>. The strong interaction between the intron and the U2-U6 snRNA complex may suggest that the presence of intron changes the conformation of the complex.

The third factor considered is protein. The protein-free splicing systems carry out the splicing-like reaction at extremely low efficiency and yield. Taking aside the fact that there is still a debate about whether the spliceosome is a ribozyme<sup>17</sup>, it is safe to conclude that protein components of spliceosome play important roles for the catalysis. the roles are Most likely for structural stability. Some protein factors may help position U2-U6 snRNA complex.

#### **4.1.3 Two possible folding schemes**

In order for the splicing reaction to be carried out, the reacting parts of the intron are required to be brought together. Four-helix junction is known to juxtapose catalytically essential elements by form coaxial helical stacks<sup>20</sup>. The NMR data from Butcher and coworkers<sup>66</sup> showed that the

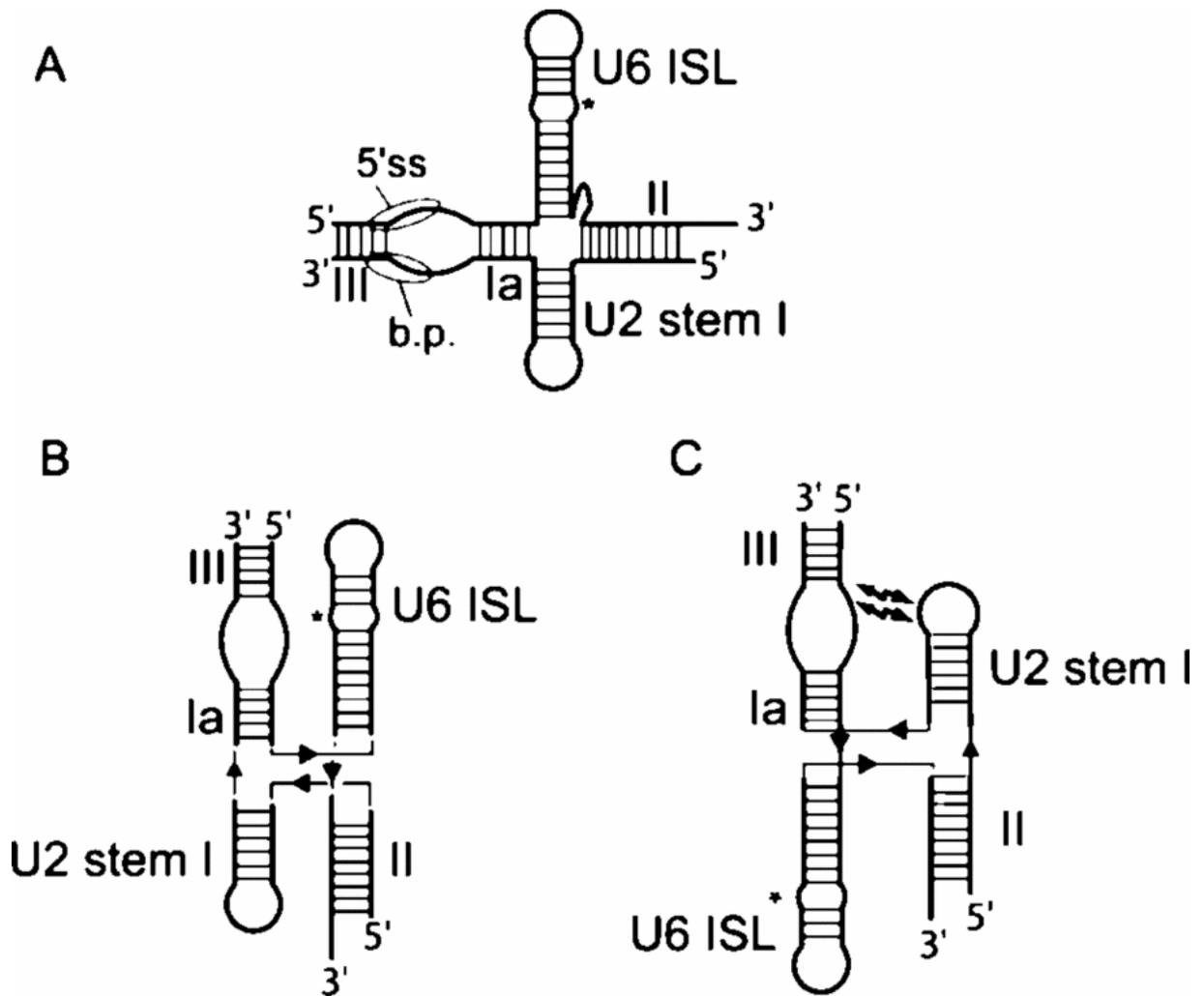
U2-U6 snRNA complex contained a four-way junction at the AGC triad region (Fig 1.4). Because several key components for the splicing reaction have to have direct contact in order for the reaction to take place and they have direct interactions with the U2-U6 snRNA complex, it is thus safe to hypothesize that the U2-U6 snRNA complex may have conformational changes during catalysis.

There are two possible folding schemes proposed in the literature<sup>102</sup>. The first scheme is that the U6 ISL is brought into proximity to the branch site and 5'-splice site (Fig 4.1B). This is supported by modeling the four-way junction based on a known crystal structure of a hairpin ribozyme four-way junction<sup>103</sup>. The second possible folding scheme is that U2 stem I (U2 ISL) is brought into proximity to the branch site and 5'-splice site (Fig 4.1C). This is supported by results of UV cross-linking studies<sup>37, 104</sup>. These results were inconsistent with the first model but could be explained by the second. They also supported the interaction between helix I and the ACAGAGA loop which was a previously identified genetic interaction<sup>105</sup>.

#### 4.1.4 Nano-surface energy transfer (NSET)

FRET is an efficient way of measuring distances with a limit of within 10 nm (discussed in section 3.1.1). Electron microscopy is able to measure extended distances, but it lacks the ability to measure dynamic conformational changes without altering biological functions. NSET is an energy transfer between a fluorophore and a metal surface. A successful demonstration of dipole-surface energy transfer shows that NSET is able to measure distance up to 20 nm and distance dependence is of  $\frac{1}{R^4}$ <sup>106</sup>. In this report, optically excited organic fluorophores are energy transfer donors and small metal nano-particles are energy transfer acceptors. This method significantly extends the FRET distance measurements.

The reason why FRET follows  $R^6$  relationship while NSET follows  $R^4$  is because FRET probes are two single dipoles while NSET probes are one single dipole and one 2D dipole. Both FRET and NSET physically originate from the weak electromagnetic coupling of two dipoles. According to the Fermi Golden Rule in the dipole approximation of energy transfer, the energy transfer rate ( $k_{\text{EnT}}$ ) is the product of the interaction elements of the donor ( $F_{\text{D}}$ ) and acceptor ( $F_{\text{A}}$ ), *i.e.*,  $k_{\text{EnT}} \approx F_{\text{D}}F_{\text{A}}$ . Their separation distance ( $d$ ) dependencies are sole functions of their geometric arrangement. For single dipoles,  $F \approx 1/d^3$ , for a 2D dipole array,  $F \approx 1/d$ , for a 3D dipole array,  $F$  is a constant<sup>106</sup>. FRET donor and acceptor are both single dipoles, thus  $k_{\text{FRET}} \approx F_{\text{D}}F_{\text{A}} \approx (1/d^3)(1/d^3) \approx$



**Figure 4.1 Two possible folding patterns in U2-U6 snRNA complex from <sup>102</sup>.** A) U2-U6 snRNA complex forms a four-helix junction. B) Folding with U6 ISL in the proximity of the ACAGAGA loop. C) Folding with U2 ISL in the proximity of the ACAGAGA loop.

$1/d^6$ . NSET, on the other hand, contains a single dipole (a common fluorophore) and a 2D dipole (a nanoparticles surface). Surface energy transfer rate ( $k_{\text{SET}}$ ) is:  $k_{\text{SET}} \approx F_{\text{D}}F_{\text{A}} \approx (1/d^3)(1/d) \approx 1/d^4$ .

Both FRET and NSET energy transfer efficiency are calculated by equations 3.4, 3.5, 3.6. Both FRET and NSET also follow the relationship expressed in equation 4.1.

$$E = \frac{1}{1 + \left(\frac{R}{d_0}\right)^n} \quad \text{(Equation 4.1)}$$

where E is the energy transfer efficiency; R is the distance between energy transfer donor and acceptor;  $d_0$  is the characteristic distance when  $E = 50\%$ . For FRET transfer,  $n = 6$ ; for NSET transfer,  $n = 4$ . The Förster radius,  $R_0$ , is expressed in equation 3.2 and it is generally around 50-60 Å. The nano-surface energy transfer radius ( $d_0$ ) is expressed in equation 4.2 (details in reference <sup>106</sup>) is generally larger than Förster radius:

$$d_0 = \sqrt[4]{\frac{0.525c^3\Phi_D}{\omega^2\omega_f k_f}} \quad \text{(Equation 4.2)}^{106}$$

where  $\Phi_D$  is the donor quantum efficiency;  $\omega$  is the frequency of the donor electronic transition;  $\omega_f$  is the Fermi frequency of metal;  $k_f$  is Fermi wave vector of the metal <sup>107</sup>. FAM-1.5 nm NanoGold (NG) has  $d_0 = 76.3$  Å as reported in <sup>108</sup>. In that report,  $\Phi_D = 0.8$ ,  $\omega = 3.8 \times 10^{15} \text{ s}^{-1}$ ,  $\omega_f = 8.4 \times 10^{15} \text{ s}^{-1}$ , and  $k_f = 1.2 \times 10^8 \text{ cm}^{-1}$ . For Cy5,  $\Phi_D = 0.4$ ,  $\omega = 2.91 \times 10^{15} \text{ s}^{-1}$ ,  $d_0 = 73$  Å <sup>109</sup>. For AF555,  $\Phi_D = 0.40$  for AF555 bound on RNA,  $\omega = 3.4 \times 10^{15} \text{ s}^{-1}$ ,  $d_0 = 67.6$  Å. NSET not only has bigger SET radius, the energy transfer efficiency is also less sensitive to distance changes because it depends on 4<sup>th</sup> power compared to 6<sup>th</sup> power for FRET. Thus, NSET has wider range of measurable distances. When a nanometer metal is used as energy transfer acceptor, energy transfer follows NSET rules <sup>106</sup>.

#### 4.1.5 Goals of this chapter

The goals of this chapter are to illustrate how the U2-U6 snRNA complex in solution changes its conformations upon being activated. The activation factors tested are metal ions, proteins and intron strands. FRET and NSET pairs were labeled in different sites of the U2-U6 snRNA complex. By measuring distance changes between these probes within the complex by using FRET and NSET methods, the impact of magnesium ions, intron strands and protein on the conformational change of the complex can be found. The results may help us choose one of the two models described in section 4.1.3; or help us propose a new folding scheme.

## 4.2 Materials and methods

### 4.2.1 Design of U2-U6 snRNA constructs

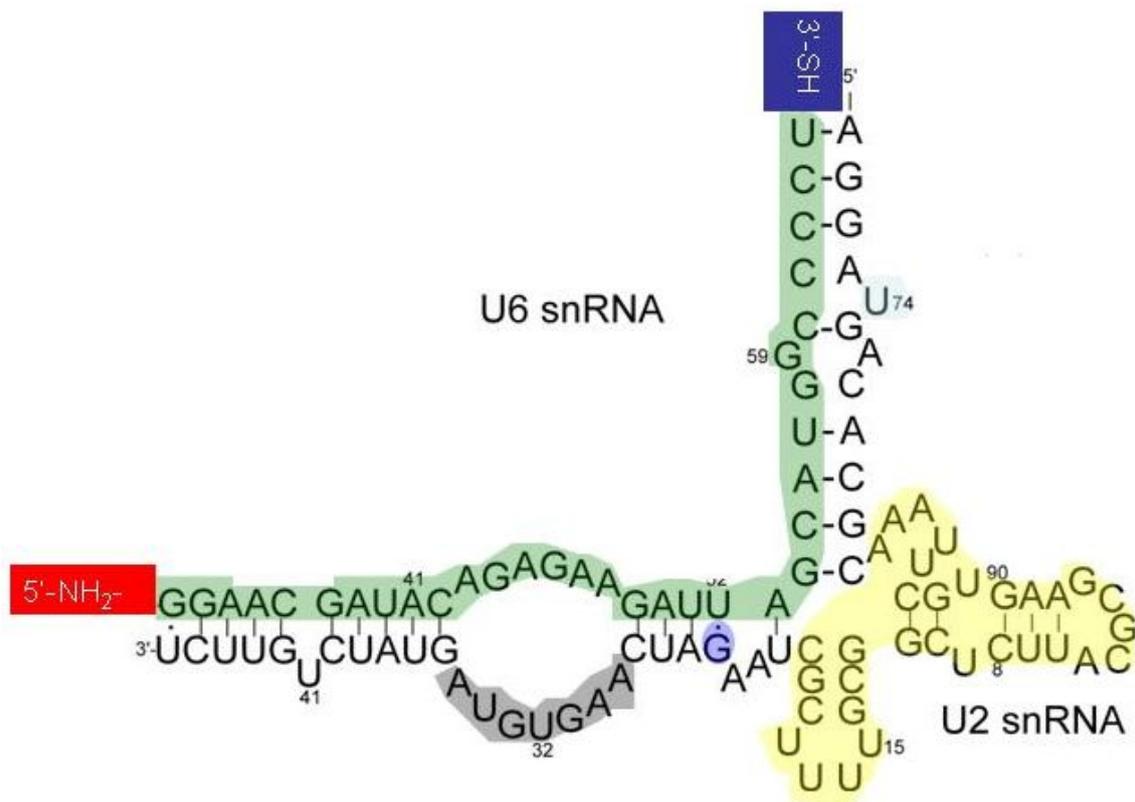
In order to test these two folding schemes, distances between energy transfer probes within the complex were measured. There were three sets of constructs designed. For simplicity, they were called designs A, B and C. In this chapter, each construct will be discussed in parallel.

#### **Design A for FRET**

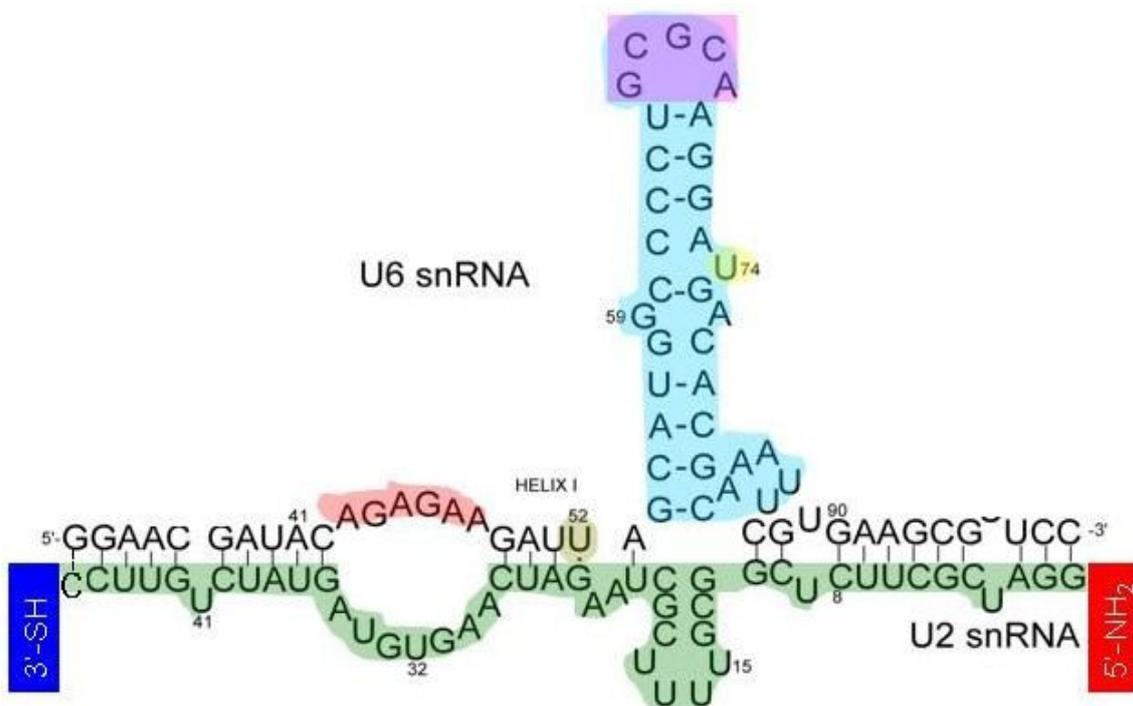
Design A is to test the first folding scheme (Fig 4.1B). In this folding scheme, the U6 ISL has significant movements relative to the two reacting components of the first step of splicing (branch point adenosine and 5'-splice site, both have contact with the conserved ACAGAGA region, refer to Fig 1.5) in order to carry out the splicing catalysis. Thus, a two-piece RNA construct series was designed (Fig 4.2). For sequence simplicity, a pentaloop on U6 ISL was removed; this pentaloop was used to linked U2 and U6 on the helix II side. Separate experiments showed the pentaloop did not change the behavior of the complex. In this series, an amine reactive group was labeled near the ACAGAGA region of the invariant top strand (named as U6-32); a thiol reactive group was labeled on U6 ISL of the invariant top strand. The two reactive groups were available for probe labeling. The colored regions represented mutations as discussed in chapter three (Fig 2.2). Other than the different FRET pairs, this series was identical to the construct design used in chapter 2.

#### **Design B for FRET**

Design B is to test the second folding scheme (Fig 4.1C). In this hypothesis, the helix II has significant movements to the conserved ACAGAGA region in order to carry out the splicing reaction. Thus, another two-piece RNA construct series was designed (Fig 4.3). In this series, an amine and thiol terminal labeled strand was chosen to be invariant (highlighted strand in Fig 4.3, named U2-48). The top strands were variations of U6 snRNA. There were five mutations on the U6 snRNA sequence. The non-mutant sequence was called WT; the  $\Delta$ LOOP construct contained replacement of the red highlighted ACAGAGA loop with complementary Watson-Crick sequence with the bottom sequence; the  $\Delta$ UG construct contains the highlighted U $\rightarrow$ C mutation; the  $\Delta$ ISL construct deletion of a nucleotide from the ISL (highlighted); the  $\Delta$ ISLCAP construct included deletion of the highlighted ISL pentaloop; the  $\Delta$ U74 construct contained the deletion of highlighted U74. The rational of mutation design was similar as with design A.



**Figure 4.2 Design A.** The invariant top strand is highlighted in green (U6-32). For sequence simplicity, a pentaloop on U6 ISL was removed; this pentaloop was used to linked U2 and U6 on the helix II side. Separate experiments showed the pentaloop did not alter the behavior of this construct. The amine and thiol groups are labeled in red and blue; other highlighted areas are the mutational sites. Construct WT contains unmodified top and bottom strands; in construct  $\Delta 4W$  the yellow region was replaced by three adenosines; in construct  $\Delta LOOP$  the black labeled region was replaced by complimentary sequence to the top strand; in construct  $\Delta U74$  the highlighted U74 was deleted; in construct  $\Delta UG$  the highlighted G was replaced by adenosine. The sequences were identical to the research in chapter 3.



**Figure 4.3 Design B.** In this design, the highlighted and amine in red and thiol in blue terminal labeled invariant strand was chosen to be the highlighted U2 snRNA (called U2-48). The top strand was the U6 snRNA. There were five mutations on the U6 snRNA sequence. The non-mutant sequence was called WT; the  $\Delta$ LOOP construct contained replacement of the red highlighted ACAGAGA loop with complimentary Watson-Crick sequence with the bottom sequence; the  $\Delta$ UG construct contains highlighted U $\rightarrow$ C mutation; the  $\Delta$ ISL construct contained the total deletion of highlighted ISL; the  $\Delta$ ISLCAP construct contained the deletion of the highlighted ISL pentaloop; the  $\Delta$ U74 construct contained the deletion of highlighted U74.

### **Design C for FRET and NSET**

Design C is to test the two folding schemes together in one design (Fig 4.1 B, C). The designs A and B lacked the ability to simultaneously monitoring these two possible folding patterns in one experiment. Thus, a three-piece RNA design was made to accomplish this (Fig 4.4).

Each of the three highlight RNA pieces (left portion of U6 snRNA called DC32; the right portion of U6 snRNA called DC28; the U2 part called DC46) were 5'-labeled. DC32 was labeled with 5'-thiol, which was available for NanoGold (NG) labeling; DC28 was labeled with 5'-amine, which was available for Cy3 labeling; DC46 was labeled with 5'-FAM (Fig 4.4).

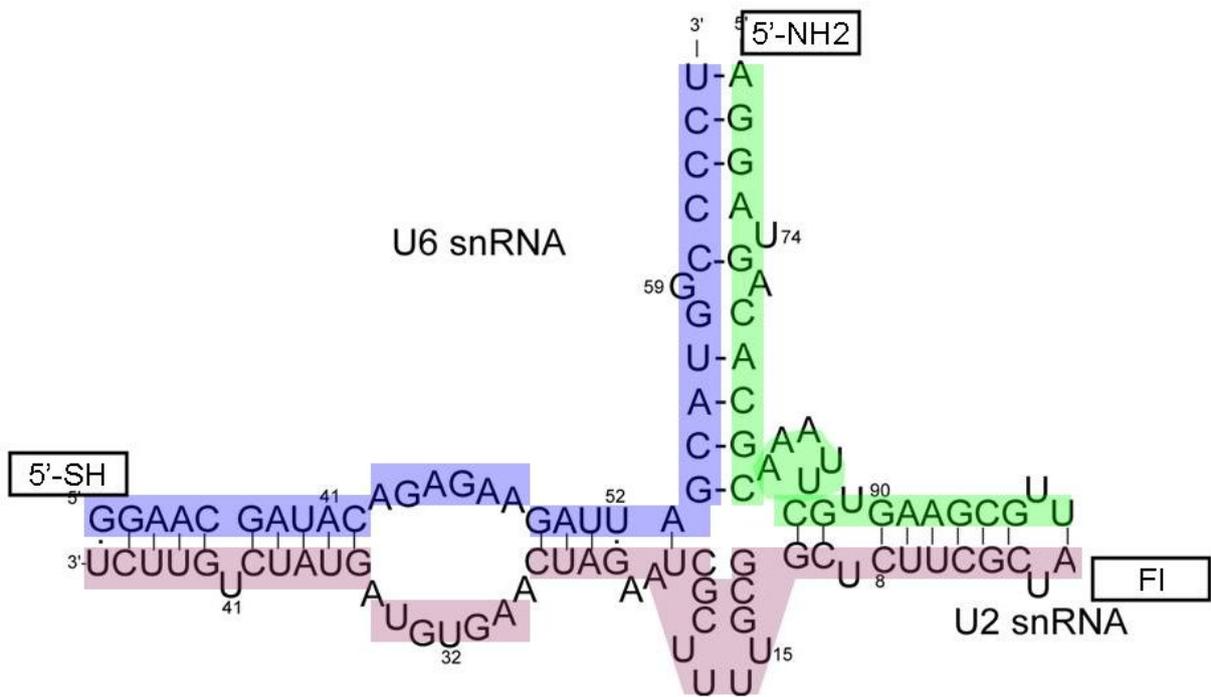
### **4.2.2 RNA sample preparation**

**Design A:** All unlabeled RNAs were made by *in vitro* transcription or purchased as described in section 2.2.1. For Cy3/Cy5 pair labeled samples, dye-labeled DC32 RNA was directly purchased from either IDT technology or Dharmacon.

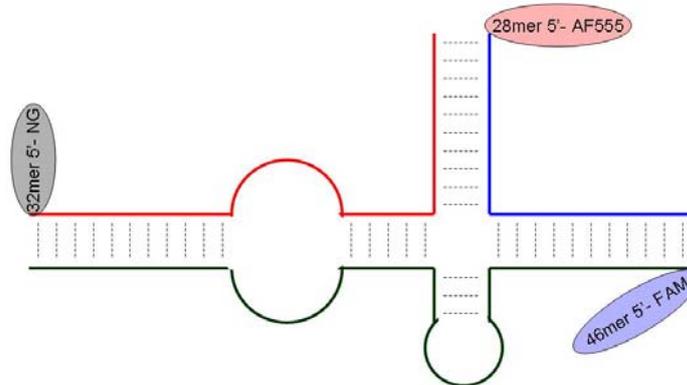
In other dye labeled RNA strands, Alexa Fluor dyes (AF for short thereafter, from Invitrogen-Molecular Probes) were conjugated to purchased amine and thiol labeled DC32 in the lab by following the companion instructions. The Alexa Fluor dyes used were amine-reactive AF488 carboxylic acid,2,3,5,6-tetrafluorophenyl ester (TFP) or succinimidyl esters is; thiol reactive AF555 C<sub>2</sub>-maleimide and AF647 C<sub>2</sub>-maleimide. AF488/AF555 and AF488/AF647 are the two FRET pairs.

RNA purification: all unlabeled *in vitro* transcripts, unlabeled purchased RNA, labeled purchased RNA, self-labeled purchased RNA were purified by 10% denaturing PAGE gel. Gel slices containing the desired RNA were eluted by an electro-eluter (Schleicher & Schuell, Whatman Inc.). Unlabeled RNA concentrations were determined by absorption at 260 nm. The RNA concentration and probe concentration of labeled RNA strands were determined by absorption at 260 nm (for RNA, 1 OD<sub>260</sub> Unit = 40 µg/ml for single-stranded RNA, or more accurately using calculated sequence-dependent extinction coefficient  $\epsilon$  on IDTDNA website <http://www.idtdna.com/analyzer>). Dye concentrations were determined by absorption at 550 nm (Cy3,  $\epsilon = 150000 \text{ M}^{-1}\text{cm}^{-1}$ ), 650 nm (Cy5,  $\epsilon = 250000 \text{ M}^{-1}\text{cm}^{-1}$ ), 500 nm (FAM,  $\epsilon = 92300 \text{ M}^{-1}\text{cm}^{-1}$ ), 494 nm (AF488,  $\epsilon = 78000 \text{ M}^{-1}\text{cm}^{-1}$ ), 555 nm (AF555,  $\epsilon = 158000 \text{ M}^{-1}\text{cm}^{-1}$ ) and 649 nm (AF647,  $\epsilon = 250000 \text{ M}^{-1}\text{cm}^{-1}$ ). The 1:1 ratio of RNA and probe on it indicated pure labeled RNA.

A



B



**Figure 4.4 Design C: three-piece RNA system.** A) The three RNA pieces (DC32 in blue; DC28 in green; DC46 in plum) were highlighted in colors. The 5'- thiol is labeled on DC32; the 5'- amine is labeled on DC28; the 5'- FAM (FI) is labeled on DC46. B) the schematic representation of labeling in design C. The DC32 RNA was labeled with 1.5 nm NanoGold by reacting with 5'-SH; DC28 with AF555 by reacting with 5'-NH<sub>2</sub>; DC46 with 5'-FAM purchased.

**Design B:** Sample preparation for this set of RNA strands was very similar as for design A. Unlabeled RNAs were made by *in vitro* transcription. Amine and thiol labeled U2-48 was purchased. For FAM/Cy3 pair, dye-labeled U2-48 RNA was purchased. For AF488/AF555 and AF488/AF647 pairs, they were labeled in the same way as for design A. RNA purification and quantification were done by the same procedure as for design A.

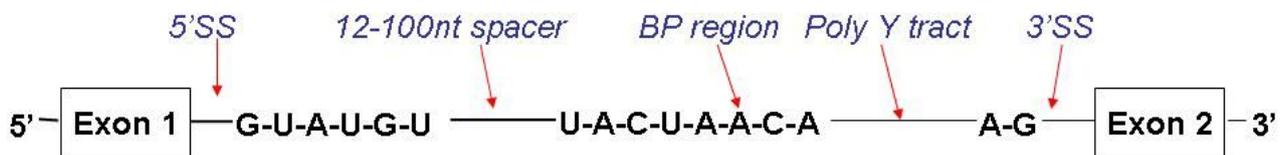
**Design C:** The reason for this design was that in both designs A and B, two probes were on one RNA strand. This caused the labeling yield to be low. In design C, however, only one probe was on one RNA strand. This simplified the labeling and purification process.

The unlabeled DC28, DC32 and DC46, the 5'-thiol labeled DC32, the 5'-amine labeled DC28 and the 5'-FAM labeled DC46 were purchased. Two different approaches were done on the amine or thiol reactive RNAs. In one approach, 5'-NH<sub>2</sub> labeled DC28 was labeled with amine-reactive 1.4 nm NanoGold (from Nanoprobes, Cat # 2025A, Mono-Sulf-NHS-NanoGold); 5'-SH labeled DC32 was labeled with AF555 by reacting with thiol-reactive Alexa Fluor 555 C<sub>2</sub>-maleimide. In the other approach, 5'-SH DC32 was labeled with thiol-reactive 1.5 nm NanoGold (from Strouse lab <sup>109</sup>), 5'-NH<sub>2</sub> DC28 was labeled with amine-reactive Alexa Fluor 555. Due to its better coupling efficiency and availability, 1.5 nm NanoGold was favored over 1.4 nm NanoGold. Thus, the results of only one approach (1.5 nm on DC32, Fig 4.4B) were discussed.

RNA purification: the unlabeled and dye labeled RNAs were purified by the same procedure as for design A and B. The 1.4 nm and 1.5 nm NanoGold labeled RNAs, however, were treated differently. RNA and NanoGold were reacted for 24 hours before going through ethanol precipitation. The precipitation was washed with 70% ethanol two times. Concentrations of the sample were estimated from original RNA sample by assuming 100% labeling efficiency. All the subsequence experiments had to be finished within a week to preserve NanoGold stability.

### 4.2.3 Intron design and synthesis

An intron which can be the substrate of the spliceosome has a sequence in Fig 4.5. Three intron constructs were designed. The 26-nucleotide intron strand included the branchpoint region (the same sequence as used protein-free splicing substrate Valadkhan and Manley 2001 <sup>22</sup>); its sequence was (5'- GGG CCU GGC UUU UUU UAC UGA CAC UU-3'), as the branch site underlined. It was called In26. The In45 intron contained the spacer region, the branchpoint and



**Figure 4.5 Schematic sequence of an intron.** The pointed regions are conserved and recognized by spliceosome. These region include the 5'-splice site (5'SS), the 12-100 nt spacer region, the branchpoint (BP) region, the polypyrimidine (poly Y) tract and the 3'-splice site (3'SS).

the poly Y (pyrimidine) tract; its sequence was (5'-GGG AGU CUU GUC GUC UAC UGA CAC UUC UUC UUC UCU CUU UUU CCC -3'). It was called In45. The In27 intron contained 5'-ss and the branchpoint; its sequence was (5'- CAG GUA AGU UUU UCU UUA CUG ACA CUU-3'), as the 5'-splice site and branchpoint boxed. It was called In27.

#### **4.2.4 Proteins expression and purification**

As for proteins, because there are over 200 proteins<sup>110</sup> in spliceosome, the only proteins used in this chapter were p14 and SF3b155 (1-462). They have been shown to be in the proximity of the branch site region (Fig 1.8). They were provided by Dr. Kersten Schroeder in Greenbaum lab; the expression and purification procedure were discussed in detail in his dissertation<sup>111</sup>.

#### **4.2.5 FRET and NSET instrumentation**

Steady-state FRET and NSET experiments were carried out on a Cary Eclipse fluorophotometer as discussed in section 3.2.6. The information about excitation and emission wavelength, PMT power etc. will be included in the "Results" section of this chapter.

Time-resolved FRET and NSET measurements were carried out by Mani Singh in Dr. Strouse's laboratory. The detailed instructions were in sections 3.2.4, 4.2.4 and 6.2 on Dr. Travis Jennings' dissertation<sup>112</sup>. In brief, lifetime data were acquired on diluted samples using the output of a Nd:VO<sub>4</sub>-pumped (Spectra-Physics Vanguard, 2 W, 532 nm, 76 MHz, 10 ps) R6G dye laser (Coherent 702-1). Cavity dumping of the dye laser to 1.9 MHz dropped the pulse train to 1 pulse every 2  $\mu$ s. In the case of FAM and AF488, the laser was frequency doubled for excitation at 300 nm. In the case of Cy3 and AF555, the pulsed output was adjusted for excitation at 560 nm. Samples were excited with <1  $\mu$ W at a right angle geometry relative to a Chromex 500is 0.5 m imaging monochromator with 50 g/mm grating, 0.5 nm resolution. Output of the monochromator is focused into a Hamamatsu C5680 streak camera operating at 20 ns window (FAM, AF488) or 10 ns window (Cy3, AF555). Lifetimes were measured by binning intensity vs time for a 20 nm spectral range about the emission maximum for the dye.

### **4.3 Results**

#### **4.3.1 Confirmation of complex formation and further sample purification**

In order to confirm that the RNA strands in the complex form desired base pairs,

non-denaturing gels were used to monitor the RNA mobility (experimental part in section 3.2.2).

**Design A:** Prior to all experiments, all labeled and unlabeled RNA strands were purified on denaturing 15% denaturing polyacrylamide gels. Samples of U6-32 top strand with different bottom strands were heated up at 80 °C for 45 seconds and cooled at room temperature for 10 minutes prior to loading to the gel. They were run on a 10% non-denaturing gel. Since the sequences were the same as in chapter three, Fig 3.3 showed U6-32 and each of the bottom strands formed complexes with retarded mobility. This showed that they formed RNA dimers.

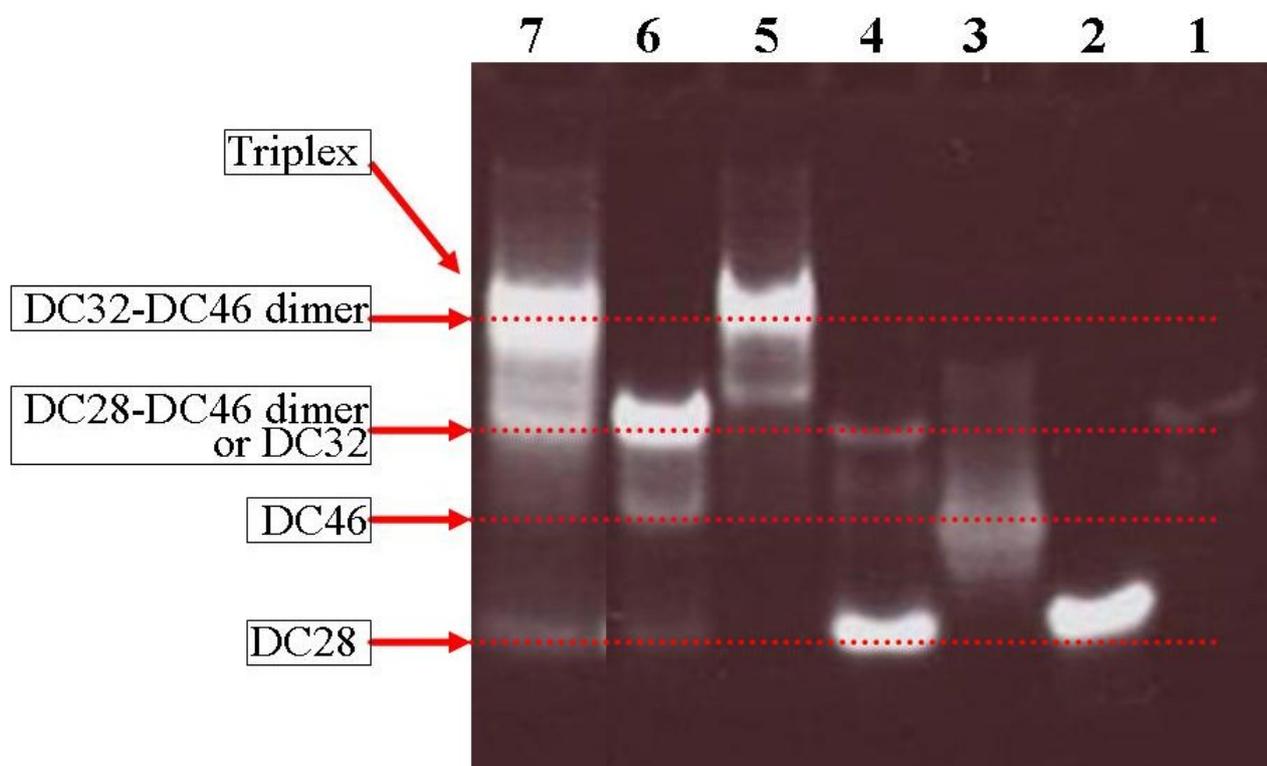
**Design B:** Verification of the complex formation was done by running non-denaturing gels. In this gel, U2-48 was radioactively labeled with  $\gamma$ -ATP with standard procedure. In short, U2-48 was linked with a  $\gamma$ -ATP by T4 Polynucleotide Kinase (PNK, from USB Corporation). The appearance of a new band indicated that U2-48 formed a complex with each of the variant top strands (Fig 4.6).

**Design C:** Samples of DC32, DC28 and DC46 were mixed and heated up at 80 °C for 45 seconds and cooled at room temperature for 10 minutes prior to loading to the gel or measurements. To confirm the three pieces to form a complex, a 15% non-denaturing gel was run. The three RNA pieces were called DC32, DC28 and DC46 for simplicity (Fig 4.4). Samples of DC32, DC28, DC46, DC32+DC28, DC32+DC46, DC28+DC46 and DC32+DC28+DC46 were run on the gel. The gel was visualized on a PDI gel scanner. From the gel, DC32+DC28 did not form a dimer; both DC32+DC46 and DC28+DC46 formed dimers. In the mixture of three RNA strands, DC32+DC46 and DC28+DC46 dimers existed while DC32+DC28+DC46 triplex was questionable. The most likely band was the top band (Fig 4.7). In the following FRET and NSET experiments, the top band was cut and purified as the triplex. Potential problems will be discussed in the discussion part of this chapter.

The non-denaturing gel experiments were done with different forms of DC32, DC28 and DC46. For example, DC32 could be NanoGold, FAM or AF555 labeled, or unlabeled; DC28 could be FAM labeled or unlabeled etc. Results were the same regardless of whether the RNA was labeled or unlabeled. If the RNA was fluorescent dye labeled, it would be bright; otherwise, Ethidium Bromide staining was used and it would not be as bright. In Fig 4.7, DC32 was unlabeled, DC28 was labeled with FAM and DC46 was FAM labeled. That is the reason why DC32 was not as bright as DC28 and DC46.



**Figure 4.6 Demonstration of invariant bottom U2-48 formed base pairs with variant top strands in design B.** Lanes 1: U2-48 by itself; lane 2: U2-48+ $\Delta$ U74 top strand; lane 3: U2-48+ $\Delta$ ISLCAP top strand; lane 4: U2-48+ $\Delta$ ISL top strand; lane 5: U2-48+ $\Delta$ LOOP top strand. The appearance of a retarded band in each of lanes 2-5 indicated that the two RNA pieces formed complex which had slower mobility on a non-denaturing PAGE gel than  $^{32}$ p labeled U2-48.



**Figure 4.7 Confirmation of the triplex formation in design C.** Sequences of DC32, Dc28 and DC46 are indicated in Fig 4.4. Lane 1: DC32; lane 2: DC28; lane 3: DC46; lane 4: DC32 + DC28; lane 5: DC32 + DC46; lane 6 :DC28 + DC46; lane 7 : all three. By comparing the mobility of these bands, the dimers and triplex were identified on the gel.

### 4.3.2 Distance measurements results

**Design A:** The Cy3-Cy5 steady-state FRET experimental results on design A were discussed in section 3.3.4. In short, in the wild type construct (original sequence in Fig 4.2, 5'-NH<sub>2</sub> was labeled with Cy3; 3'-SH was labeled with Cy5), the energy transfer efficiency was  $14 \pm 1$  % between Cy3 and Cy5. By using  $R_0$  of 54 Å and equation 3.1, the distance was calculated to be  $73 \pm 1$  Å. The mutations did not seem to alter the distance, with only  $\Delta$ LOOP construct extended the distance by about 4 Å. This was due to the elongation of a helix instead of a loop (section 3.2.4).

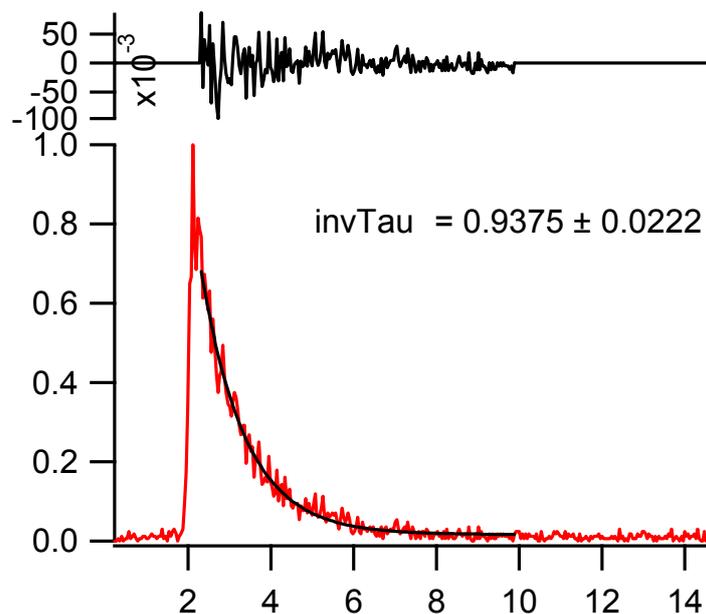
$Mg^{2+}$  up to 20 mM and 1:1 introns strands (including In26, In27 and In45 described in section 4.2.3) were added to the complex constructs. There were no measurable FRET efficiency changes after adding the metal ion, the introns, or both. This suggested that there was no conformational change in the U2-U6 snRNA complex induced by  $Mg^{2+}$  and introns.

Time-resolved FRET experiments were carried out on the wild-type Cy3-Cy5 construct. Lifetime of donor Cy3 was measured to be  $\tau_D = 1.24$  ns; lifetime of Cy3 in the presence of Cy5 was  $\tau_{DA1} = 1.07$  ns (Fig 4.8A); lifetime of Cy3 in the presence of Cy5 with 20mM  $Mg^{2+}$   $\tau_{DA2} = 1.11$  ns (Fig 4.8B). Adding intron strands of different lengths (In26, In27 or In45) before or after  $Mg^{2+}$  had no effect on the lifetimes. The energy transfer  $\tau_{DA1}$  has energy transfer efficiency of  $E = 1 - \frac{1.07}{1.24} = 13.4\%$ , corresponding to a distance of 73.7 Å, a value consistent with steady-state measurements; upon adding  $Mg^{2+}$ ,  $\tau_{DA2}$  the energy transfer efficiency was measured to be  $E = 1 - \frac{1.11}{1.24} = 10.5\%$ , corresponding to a distance of 77.2 Å. Thus, the changes in energy transfer efficiency and distance were very small. Because U2-U6 snRNA complex was a very dynamic system, there were two possibilities. One was that the distance change was very small. The other possibilities were that a small population at any time may be changing in response to metal ion. But the measurements could not distinguish these two.

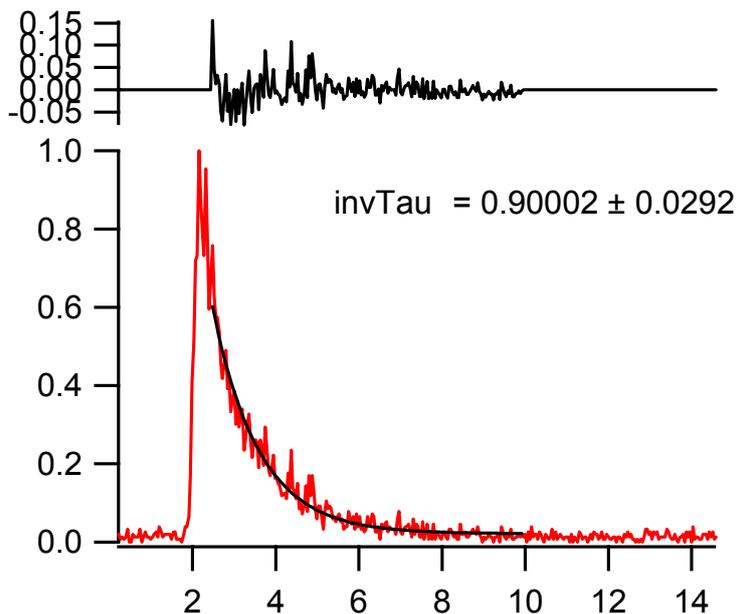
**Addition of proteins p14 and SF3b155:** Proteins p14 and SF3b are in the proximity of the branch site of the spliceosome (Fig 1.8). Adding either of the proteins to the wild type construct of design A with intron strand did not change energy transfer efficiency measured by steady-state FRET, suggesting that conformational changes were not induced by these protein components.

Steady state intensity FRET measurements with FRET pairs AF488/AF555 ( $R_0 = 70$  Å) and AF488/AF647 ( $R_0 = 56$  Å) gave us similar results.

A.



B.



**Figure 4.8 Lifetime of Cy3 in the sample of Cy3-Cy5 labeled WT construct of design A.** A) Cy3

lifetime was measured to be  $\tau_{DA1} = \frac{1}{0.9375 \pm 0.0222} ns = 1.07 \pm 0.03 ns$  in the absence of  $Mg^{2+}$ . Buffer was MOPS 10 mM, NaCl 150 mM, pH 7.2. B) Cy3 lifetime was measured to be

$\tau_{DA2} = \frac{1}{0.90002 \pm 0.0292} ns = 1.111 \pm 0.035 ns$  in the presence of 20 mM  $Mg^{2+}$  in a buffer of MOPS 10 mM, NaCl 150 mM, pH 7.2.

**Design B:** FAM and Cy3 labeled U2-48 and wild-type top strand were paired (Fig 4.3). The excitation wavelength was 440 nm; an acceptor normalization process was used similarly as in Fig (Fig 4.9). The big peak for Cy3 was due to direct excitation by 440 nm photons, not *via* FRET transfer. This was confirmed by a measurement that the same intensity was obtained by exciting a Cy3 acceptor-alone sample of the same concentration at 440 nm in the same conditions. For the FAM-Cy3 pair,  $R_0 = 53 \text{ \AA}$ , the energy transfer efficiency was very low. This small number indicated the distance was too far for the FAM-Cy3 pair to measure accurately.

The AF488-AF555 FRET pair has a Förster radius of  $70 \text{ \AA}$ . It should therefore be useful for solving the distance between these two sites. Using the same acceptor normalization process, the energy transfer efficiency was measured at  $\sim 18\%$ . This corresponded to be  $\sim 90 \text{ \AA}$  between AF488 and AF555.

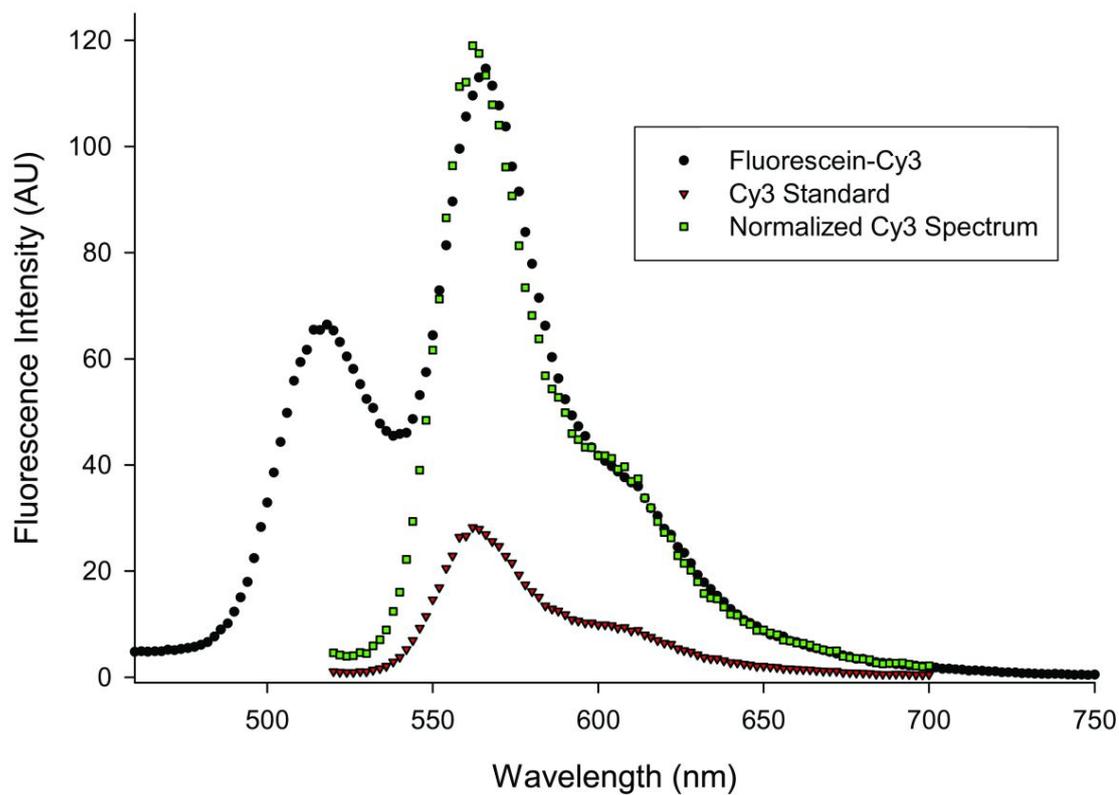
$\text{Mg}^{2+}$  ions were added to the complex samples up to a final concentration of 20 mM. There was no measurable FRET efficiency change after adding the metal ions. No measurable FRET efficiency change was observed after introns (In26, In27, and In45, concentration of them were 1:1 to complex concentration) were added before or after  $\text{Mg}^{2+}$ . These results suggested that there was no measurable conformational change induced by  $\text{Mg}^{2+}$  or intron strands in design B. Helix II was not induced to come closer to the ACAGAGA loop.

A lifetime experiment on the AF488/AF555 labeled sample of wild-type design B was carried out. AF488 was the FRET donor; AF555 was the acceptor. The lifetime of AF488 in the absence of AF555 was measured to be  $\tau_D = 4.0 \text{ ns}$  (Fig 4.10).

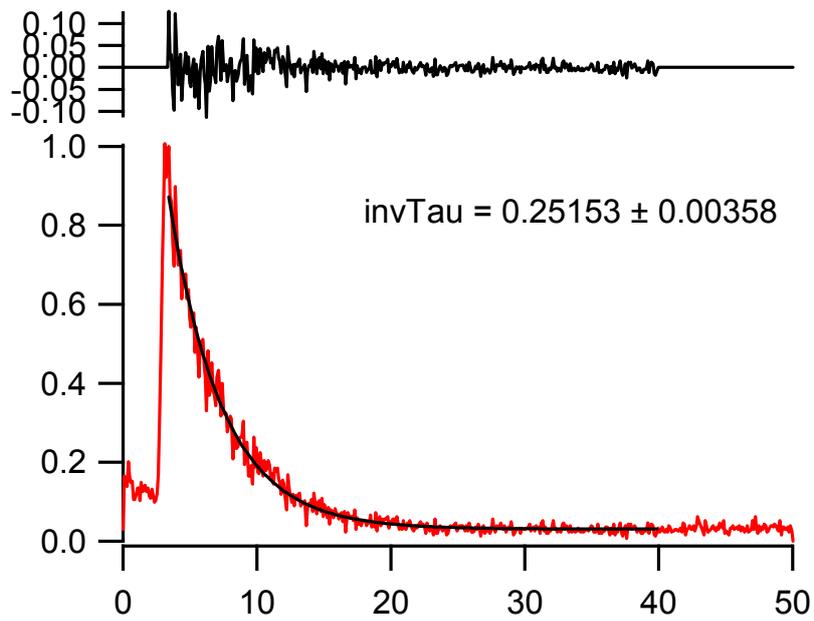
The lifetime of AF488 in the presence of AF555 was measured to be  $\tau_{DA} = 3.3 \text{ ns}$ . This gave similar energy transfer efficiency (18%) as the results from the steady state FRET, suggesting the distance separating the two probes was about  $90 \text{ \AA}$ . There were no observable conformational changes induced by  $\text{Mg}^{2+}$  or introns under the described conditions.

**Design C:** The experiments for the designs A and B showed that the distances between the two arms (U6 ISL, helix II) and the ACAGAGA loop did not change upon addition of  $\text{Mg}^{2+}$ , with or without the intron strand. However, these two distances were too large for most FRET measurements. To overcome this problem, NSET method was used by taking advantage of its larger  $d_0$  as well as its  $R^4$  dependency.

The RNA construct was illustrated in Fig 4.4B. DC32 RNA was labeled with 1.5 nm



**Figure 4.9 Acceptor normalization process for FAM-Cy3 FRET measurement.** The black dots were results from FAM-Cy3 sample. Normalized Cy3 spectrum was obtained by normalizing the region from 650-700 against a standard Cy3 spectrum.



**Figure 4.10** Lifetime of AF488 in the absence of AF555 in the wild-type construct of design B. the lifetime was measured to be about 4.0 ns. Adding  $Mg^{2+}$  or intron substrates did not change it.

NanoGold labeled; DC28 was labeled with AF555; DC46 was labeled with FAM.

The purity of DC28-AF555 and DC46-FAM were verified by a single band for each on a denaturing PAGE gel. In order to assure the three pieces to form a triplex with 1:1:1 stoichiometry, samples were run on a 10% non-denaturing gel. The top band (Fig 4.7) was most likely the triplex because it moved most slowly. It was cut out and purified. Five different combinations of probes were used: in the triplex, AF555 on DC28; FAM on DC46; NSET pair of AF555-NG; NSET pair of FAM-NG and three-probe combination. Their cartoon representation is shown in table 4.1. On the gel, the retardation effect of the bulky NG was neutralized by its negative charges.

Triplex T5 (table 4.1) in design C had two organic fluorophores and one metal. Energy transfer between three-interacting probes were mathematically elucidated<sup>113</sup>. Thus, in design C, it was also possible to triangulate the three distances at a time. However, it is still under investigation.

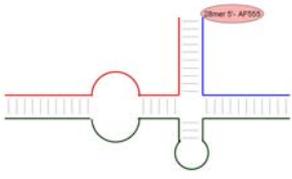
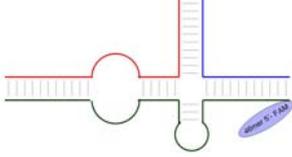
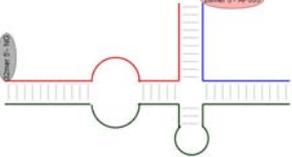
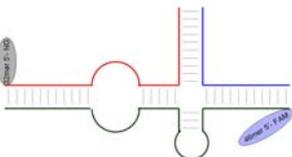
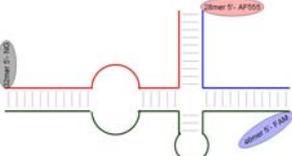
Lifetime measurements were acquired on triplex T1 with AF555 on DC28; triplex T2 with FAM on DC46; triplex T3 with AF555 on DC28 and NG on DC32; triplex T4 with NG on DC32 and FAM on DC46 and triplex T5 with NG on DC32, AF555 on DC28 and FAM on DC46. The lifetime results were summarized in table 4.1. The results were very confusing. AF555 on triplex T1 with AF555 on DC28 had a lifetime of  $0.73 \pm 0.03$  ns; in the presence of NanoGold triplex T3, the lifetime was increased to  $0.79 \pm 0.02$  ns. This was certainly against the rule of FRET/NSET because in the presence of acceptor (NanoGold), the lifetime of donor should be decreased. The only explanation was that the environment the donor AF555 was in was changed. One possibility was that AF555 in triplex T1 had a different triplex forming efficiency as in triplex T3. The NanoGold particle may interfere with the triplex formation. When the triplex was dissolved in buffer, they dissociated. Triplex T1 may have more or less dissociation than triplex T3.

FAM lifetimes were equally confusing. The triplex T2 represented the FAM lifetime as  $3.60 \pm 0.09$ ; in the presence of NanoGold, the lifetime was increased to  $3.65 \pm 0.15$  ns.

To test whether  $Mg^{2+}$  or introns could cause conformational changes of the triplex, in the sample of triplex T3 in table 4.1, 20 mM  $Mg^{2+}$  or intron strand (in a 1:1 ratio to triplex) were added. There was no observed change of the AF55 lifetime. These data suggested that  $Mg^{2+}$  and intron both did not cause a measurable distance change between the NanoGold and AF555.

In the sample of triplex T4, 20 mM  $Mg^{2+}$  or intron strand (in a 1:1 ratio to triplex) were added as well. There was no observed change of the FAM lifetime. These data suggested that  $Mg^{2+}$

**Table 4.1 Lifetime and energy transfer efficiency results on design C.** The symbol x indicated no results were reported. Lifetime of FAM was measured at the excitation of 300nm and emission at 520 nm; lifetime of AF555 was measured at excitation of 560 nm and emission of 580 nm.

Pairing Combination	$\tau$ FAM (ns)	$\tau$ AF555 (ns)	E
T1 	×	$0.73 \pm 0.03$	×
T2 	$3.60 \pm 0.09$	×	×
T3 	×	$0.79 \pm 0.02$	<0?
T4 	$3.65 \pm 0.15$	×	<0?
T5 	$3.58 \pm 0.19$	$0.64 \pm 0.07$	?

and intron both did not cause a measurable distance change between the NanoGold and FAM.

## 4.4 Discussion

The objective of this chapter was to study the conformational changes of the U2-U6 snRNA complex upon binding with  $Mg^{2+}$ , intron substrate and proteins which were known to interact with the branch site. The evidence for the conformational changes was sufficient for the hypothesis that key components in the complex were brought together in order for the splicing reaction to occur. There was a two-step plan regarding this research. The first step was to observe conformational change under a condition. The second step was to compare results from variant conditions to the condition under which conformational change happened. However, in the conditions tested for the three designs, distances did not change between the FRET and NSET pairs.

From a negative result, it is intrinsically difficult to find out the reasons behind the failure to obtain a positive result. The possible reasons of why conformational changes were not observed are discussed as follows.

**Flaws in the experimental design:** To gain insights in the possible complex conformational changes, three designs have been used. However, each design has its own flaws. For design A illustrated in Fig 4.2, there were two differences between design A and the natural U2-U6 snRNA sequence. First, the pentaloop on the U6 ISL was cut for the sake of probe labeling. This, on one hand may destabilize the complex; on the other hand, the pentaloop may possess some functions not known yet. RNA loops are known to play essential roles in interactions, such as RNA loop-loop interaction or metal ion binding. However, this was unlikely because in design B, no conformational changes even though the pentaloop was there. Second, the helix II region of the complex was enclosed with a loop. Again, this may also have two-fold influence. The loop may stabilize the complex. Four-way junctions are known to have flexible structures<sup>114</sup>, especially in the U2-U6 snRNA complex because a NMR structure has already shown the flexibility<sup>66</sup>. On the other hand, the new loop may introduce an unwanted interaction.

For design B illustrated in Fig 4.3, it followed the exact sequence of the natural U2-U6 snRNA sequence. However, the distance between the two probes was too long ( $\sim 90\text{\AA}$ ) to measure for ordinary FRET pairs. It may also not change because even there were two possible folding schemes; the first possibility was favored due to function relevance. The importance of the U6 ISL

was well known<sup>38</sup> while the function of helix II was still debatable<sup>20</sup>.

For design C illustrated in Fig 4.4, in addition to the removal of the U6 ISL pentaloop, the major problem was from the triplex formation of these three pieces. From the gel (Fig 4.7), the triplex formation efficiency was low. Even if the band cut from the gel was purified carefully, upon dissolving in buffer solutions, the three pieces may dissociate from each other again. From gel (Fig 4.7), lane 7, there were several bands. The major band was DC32-DC46 dimer. The pure triplex would be dissociated again after it was dissolved in buffer. Thus, even the sample contained 1:1:1 of all three pieces, only a small portion was the triplex. This was also likely based on the lifetime results (Table 4.1) because donor lifetimes should go down if their environment was the same.

**The folding is hard to measure:** The spliceosome is very complicated and dynamic and contains numerous proteins. As discussed in section 1.5, the protein-free splicing has a very low yield and a very slow rate of catalysis<sup>22, 31</sup>. The low yield and slow rate may make bulk phase energy transfer methods not suitable to monitor the conformational changes. Only a small fraction of RNA underwent conformational changes. This made conformational changes hard to observe. Even though some proteins (p14 and SF3b155) were used in some of the measurements, it was far from constituting a natural spliceosome. We could conclude that p14 and this small segment of SF3b155 were insufficient to facilitate folding. However, including other factors to the study is a reasonable plan for the future, but is out of the scope of this current research.

## CHAPTER 5

### CONCLUSIONS AND FUTURE DIRECTION

#### 5.1 Conclusions

##### 5.1.1 Tb<sup>3+</sup> binding sites in the U2-U6 snRNA complex

Metal ions play both catalytic and structural roles in ribozymes<sup>47</sup>. The protein-free U2 and U6 snRNAs complex could catalyze reactions resembling the first step of splicing in the presence of metal ions<sup>22,31</sup>. The backbone at U74, the two G residues in the invariant ACAGGAGA segments, and the A of the invariant AGC triad were suggested to bind specific magnesium ions by phosphorothioate substitution experiments. These results support that binding of Mg<sup>2+</sup> in these regions is essential for structural or chemical activities associated with RNA splicing<sup>23, 25, 26, 65</sup>. NMR experiments displayed changes in proton chemical shifts for U80, the yeast equivalent of U74, upon addition of Mg<sup>2+</sup> to an RNA stem loop representing the ISL from *S. cerevisiae*<sup>38</sup>, implying that this was a region of ion interaction. In the same report, pH-dependent metal ion binding of a bulged uridine of the U6 ISL was also suggested<sup>38</sup>.

Mg<sup>2+</sup> and Ln<sup>3+</sup> ions have been shown to occupy overlapping sites in a number of RNAs<sup>58-60</sup>. In chapter two, we took advantage of the luminescence of Tb<sup>3+</sup> by using Tb<sup>3+</sup> bound on RNA as FRET donors. The metal ion binding sites are spectroscopically mapped out.

The following is the binding scheme. When Tb<sup>3+</sup> binds to a specific sites, its luminescence is enhanced several magnitudes because of sensitized excitation by RNA. Tb(III) which is not bound on RNA is hardly luminescent because of its low extinction coefficient. This excited Tb<sup>3+</sup> then can transfer energy to an organic dye. By measuring the energy transfer efficiency, the distance between Tb<sup>3+</sup> bound on RNA and the organic dye can be calculated. Because tr-FRET can resolve multiple lifetimes in one measurement, up to three metal ion binding sites were measured. The

pH-dependence of metal ion binding can also be measured because pH could turn on or off the metal ion binding. If the binding site is pH-dependent, at some pH values, the site can be identified; at other pH values, the site can not be identified.

Using this approach, three regions of site-binding by  $Tb^{3+}$  in the U2-U6 snRNA complex were identified. One is located at the ACAGAGA loop; it is not pH-dependent in the pH range of 5.6-7.2. The second is located in the vicinity of the four-way-junction structure, possibly associated with the AGC triad in the context of the intact junction structure; it binds metal ions at  $pH > 6$  only. The third ion-binding site is located in the vicinity of U74 of U6 ISL; it binds metal ions at  $pH > 7$  and does not bind at pH lower than 7.

### 5.1.2 $Tb^{3+}$ -RNA binding

It is the first time that  $Tb^{3+}$  bound on RNA were used as FRET donors (chapter 2). In chapter three, characterization of the type of  $Tb^{3+}$  is investigated by using a GUAA loop as a model RNA system.  $Tb^{3+}$  and RNA binding has moderate binding affinity and  $Tb^{3+}$  luminescence is long. Several properties of the  $Tb^{3+}$  bound on RNA are found in the research in chapter three.

The first property is that in order to use  $Tb^{3+}$  luminescence to represent the concentration of  $Tb^{3+}$  bound on RNA, which is in equation 3.1, a compensation factor ( $CF_0 = \frac{e^{\tau/\tau}}{\tau}$  equation 3.6) should be used. After the compensation,  $Tb^{3+}$  luminescence is the preexponential amplitude shown as  $I_0$  in equation 2.7.

However, most luminescence intensity can be approximately represent the concentration of  $Tb^{3+}$  bound on RNA because equation 3.6 do not change in situations such as no delay times, comparable short delay times and long lifetimes, and constant stable lifetimes.

The other property is that there is a “diffusional equilibrium” process for the  $Tb^{3+}$  ion excitation and quenching process when it “chelated” binds a RNA molecule. At the time of the excitation, only “chelated”  $Tb^{3+}$  is excited (280nm) because of RNA aromatic groups act as sensitizers. Free  $Tb^{3+}$  at the time is only excited at a minimal level because of its low molar absorptivity. After the pulse, an equilibrium is reached between excited and unexcited  $Tb^{3+}$ .  $Tb^{3+}$  displays a lifetime expressed by equation 3.9.

### 5.1.3 $Tb^{3+}$ bound on RNA as FRET donors

In chapter three, the use of of  $Tb^{3+}$  bound on RNA as FRET donors is evaluated. Because of

moderate binding affinity between RNA and  $Tb^{3+}$ , there is always some free  $Tb^{3+}$  in the system. The free  $Tb^{3+}$  will contribute to the diffusional energy transfer process. To minimize the effects, the practice of using low RNA concentration and low  $Tb^{3+}$  concentration is recommended. Doing this type of tr-FRET experiments also need to record RNA and  $Tb^{3+}$  concentrations to ensure the data accuracy.

Using  $Tb^{3+}$  bound on RNA as FRET donors can help map out relatively strong metal ion binding sites in target molecules. Many RNAs can “chelated” bind with  $Tb^{3+}$  and enhance their luminescence<sup>35</sup>. Some metallo-proteins also show the enhancement of specific bound  $Tb^{3+}$  luminescence. By using similar approaches as used in chapter three, “native”  $Tb^{3+}$  binding sites can be mapped out. The word “native” here does not mean  $Tb^{3+}$  binding occurs in biological conditions; it means  $Tb^{3+}$  can bind to these sites without modifications.

#### **5.1.4 Conformational change of the U2-U6 snRNA complex**

In chapter 4, the conformational changes of the U2-U6 snRNA complex upon binding with  $Mg^{2+}$ , intron substrate and proteins which were known to interact with the branch site are studied. By labeling FRET and NSET probes in different sites of the complex, distance changes are monitored. In the conditions tested, distances did not change between the FRET and NSET pairs.

### **5.2 Future directions**

#### **5.2.1 Design a new class of FRET donors**

The success of using  $Tb^{3+}$  bound on RNA as FRET donors encourages one to imagine developing a new class of FRET donors. These FRET donors contain two components, a  $Tb^{3+}$  ion and a  $Tb^{3+}$  binding group. The  $Tb^{3+}$  binding group can be an organic molecule, a RNA motif, a DNA loop or a protein domain. The only requirement is that  $Tb^{3+}$  should have relatively tight binding with  $Tb^{3+}$  and the binding results in enhanced  $Tb^{3+}$  luminescence.

In this sense, Selvin’s application of using EDTA-chelated  $Tb^{3+}$  as FRET donors is one approach of the second possibility<sup>41</sup>. The  $Tb^{3+}$  binding groups contain two parts, an EDTA-like chelate group and an antenna aromatic group. They bind to  $Tb^{3+}$  with excellent affinity and luminescence enhancement. The drawback is that they have to be covalently linked to the target molecules.

The candidates for the  $Tb^{3+}$  binding group part of the FRET donors are DNA, RNA or protein.

When the  $Tb^{3+}$  binding group is incorporated into its respective target molecule at a carefully chosen location, a FRET donor is created without covalent coupling. For example, a RNA based  $Tb^{3+}$  binding group can be incorporated into a carefully selected site on a RNA molecule by convenient ways of making RNA, such *in vitro* transcription or chemical synthesis; a DNA based  $Tb^{3+}$  binding group can be incorporated into a carefully selected DNA sequence; a protein based  $Tb^{3+}$  binding group can be incorporated into a protein sequence. The technology advancements of making these biological molecules help us incorporate these  $Tb^{3+}$  binding groups more easily than chemical modification.

This class of donors contains two components,  $Tb^{3+}$  and a  $Tb^{3+}$  binding group.  $Tb^{3+}$  has poor absorption coefficient, thus, it is not a practical FRET donor by itself. Via sensitized excitation, lanthanide ion luminescence can be enhanced 4-5 orders of magnitude. The enhancement make  $Tb^{3+}$  attractive FRET donors.

The advantages of using “native” specific metal ion binding sites as FRET donors are two fold. On one hand, metal ion binding sites can be mapped out on a molecule as our previous practice<sup>48</sup>. On the other hand, protein or RNA based  $Tb^{3+}$  binding groups can be designed. They can be conveniently inserted into target molecules.

The search for the lanthanide luminescence enhancement by DNA on Pubmed does not yield any results, suggesting using lanthanide ions to study DNA metal ion interaction is rare. The reason is that the predominately double helix forms and little structural variability of DNA. Using  $Tb^{3+}$  to probe RNA and protein metal ion interactions, however, are common. For example, a series of  $Tb^{3+}$  binding peptides were developed<sup>91</sup>. In this chapter, only RNA based  $Tb^{3+}$  binding group is discussed.  $Tb^{3+}$  is used as a representative of  $Tb^{3+}$ .

With the advancement in aptamers (oligonucleotides or peptide molecules that bind a specific target molecule) research, new strong  $Tb^{3+}$  binding groups can be found.

### **5.2.2 Measure conformational changes in the U2-U6 snRNA complex**

There are no conformational changes observed in chapter four. There are two possible reasons. One is that there are flaws in the experimental designs. For design A, there were two differences between design A and the natural U2-U6 snRNA sequence. First, the pentaloop on the U6 ISL was cut for the sake of probe labeling. This, on one hand may destabilize the complex; on

the other hand, the pentaloop may possess some functions not known yet. Second, the helix II region of the complex was enclosed with a loop. Again, this may also have two-fold influence. The loop may stabilize the complex. Four-way junctions are known to have flexible structures<sup>114</sup>, especially in the U2-U6 snRNA complex because a NMR structure has already shown the flexibility<sup>66</sup>. On the other hand, the new loop may introduce an unwanted interaction.

For design B illustrated in Fig 4.3, it followed the exact sequence of the natural U2-U6 snRNA sequence. However, the distance between the two probes was too long ( $\sim 90\text{\AA}$ ) to measure for ordinary FRET pairs. It may also not change because even there were two possible folding schemes; the first possibility was favored due to function relevance. The importance of the U6 ISL was well known<sup>38</sup> while the function of helix II was still debatable<sup>20</sup>. For design C illustrated in Fig 4.4, in addition to the removal of the U6 ISL pentaloop, the major problem was from the triplex formation of these three pieces. The triplex formation efficiency was low. Even if the band cut from the gel was purified carefully, upon dissolving in buffer solutions, the three pieces may dissociate from each other again. From gel (Fig 4.7), lane 7, there were several bands. The major band was DC32-DC46 dimer. The pure triplex would be dissociated again after it was dissolved in buffer. Thus, even the sample contained 1:1:1 of all three pieces, only a small portion was the triplex. This was also likely based on the lifetime results because donor lifetimes should go down if their environment was the same. Thus, to overcome this problem, new constructs should be designed.

The other problem is that the folding is hard to measure: The spliceosome is very complicated and dynamic and contains numerous proteins. As discussed in section 1.5, the protein-free splicing has a very low yield and a very slow rate of catalysis<sup>22, 31</sup>. The low yield and slow rate may make bulk phase energy transfer methods not suitable to monitor the conformational changes. Only a small fraction of RNA underwent conformational changes. This made conformational changes hard to observe. Even though some proteins (p14 and SF3b155) were used in some of the measurements, it was far from constituting a natural spliceosome. We could conclude that p14 and this small segment of SF3b155 were insufficient to facilitate folding. Thus, the future direction about this problem is to include other factors to the study this project.

## APPENDIX: COMMON LABORATORY PROTOCOLS

### I PAGE (polyacrylamide gel electrophoresis)

Denaturing PAGE:

1: make 15% denaturing gel stock solution:

Acrylamide: 145 g

Bis-Acrylamide: 5 g

Urea: 420.4 g (to 8M)

Tris base: 12.11 g (to 90 mM)

Boric Acid: 5.56 g (to 90 mM)

EDTA: 0.37 g (to 1 mM)

Add water while adjusting pH at ~7 to 1 L.

Acrylamide:Bis-Acrylamide should be 29:1.

Percentage is calculated by  $(145+5)/1000 \times 100\% = 15\%$ .

Other % can be made by varying the amounts of Acrylamide and Bis-Acrylamide.

2: made 10% APS

Dissolve ammonia persulfate 10 g in 100 mL water; store in -20 °C.

3: make 15% PAGE gel

Gel mix from step 1: 10 mL

80 µL APS from step 2

7.6 µL TEMED

Run the gel in TBE buffer

Non-denaturing PAGE:

1: make 30% Acrylamide/Bis-Acrylamide solution (generally commercial available)

Acrylamide: 145 g

Bis-Acrylamide: 5 g

Add water to 1 L.

2: make 15% gel

30% gel stock from step 1: 5mL

Water: 4 mL

10×Tris-HEPES buffer: 1 mL (final 1×)

10% APS: 80 µL

TEMED: 7.6 µL

Other % just change gel stock volume, adjusting with water.

Run the gel in Tris-HEPES buffer in cold room.

## II RNA *in vitro* transcription

1: make 10×OMg buffer, store in -20 °C

0.4 M Tris-HCl pH7.5

10mM Spermidine

0.1% Triton-100

50mM DTT

DEPC-H<sub>2</sub>O

2: transcription condition

OMg 1×

NTP 3-5 mM

GMP 5 mM

MgCl<sub>2</sub> 15-25 mM

DNA template 0.2 μM

T7 RNA polymerase 1×

T7 promotor sequence

5'-GTAATACGACTCACTATA[GGGC]-3'

3'-CATTATGCTGAGTGATATCCCG-5'

Make NTP solution:

Dissolve in 10 mM Tris-HCl, pH8.0,

Then use NaOH to adjust to pH7.0

## III buffers and solutions

PBS, pH7.4

NaCl 8g

KCl 0.2g

NaH<sub>2</sub>PO<sub>4</sub> 1.44g

KH<sub>2</sub>PO<sub>4</sub> 0.24g

H<sub>2</sub>O 800 mL

Adjust to pH7.4, to 1L.

TE, pH8.0

Tris-HCl 10mM

EDTA 1mM

EDTA solution

Dissolve in pH8.0,

Any sulfonic acid buffers are sterilized by filters, not by autoclave

RNA loading buffer	
Deionized formamide	95% (V/V)
Bromophenol Blue	0.025% (V/V)
Xylen Cyanol FF	0.025% (V/V)
EDTA	5mM
SDS	0.025% (V/V)

DEPC-H<sub>2</sub>O: RNase free water

1: add 0.1 mL to 100 mL water, shake vigorously

2: incubate for 12 hours at 37°C.

3: autoclave to break down DEPC.

Tris buffer: dissolve Tris buffer in DEPC-H<sub>2</sub>O, can not mix DEPC and Tris directly

Make Tb<sup>3+</sup> solution

1: Tb<sub>2</sub>O<sub>3</sub> dissolve in HClO<sub>4</sub>, heat and stir, boil down

2: dissolve in water, vacuum-dry down three times, dissolve in water, keep at room temperature, stock concentration should be lower than 0.1 M, pH 2~3

3: put 1 mL Tb<sup>3+</sup> solution in a beaker, add solid powder of hexamethylene-tetraamine as buffer, stir, adjust to pH5~6

4: add a drop of 0.1 % Xylenol Orange as indicator

5: titrate pH8 EDTA solution, end point: red to yellow suddenly

## **VI ethanol precipitation of nucleic acids**

1: add 10% sample volume of 3M sodium acetate

2: add 100% ethanol 2-2.5 times of sample volume to DNA or 3 times to RNA

3: vortex, 30-60 min at -20 °C.

4: centrifuge 10k rpm for 60 min

5: remove supernant, keep precipitate, wash with cold 70% ethano

6: centrifuge 10k rpm for 5 min

7: keep precipitate, air dry for 10 min

8: redissolve in DEPC-water and keep in -20 °C.

## **V couple probes to amine or thiol labeled RNA**

i: pre-treat amine or thiol labeled RNA

1: dissolve in DEPC water, chloroform extraction three times using phage lock gel

2: ethanol precipitation

3: dissolve to concentration of 25 µg/µL water, store in -20 °C.

ii: label dyes from Molecular Probe to amine modified RNA

1: make fresh 0.1 M sodium tetraborate buffer by dissolving 0.038g of sodium tetraborate decahydrate for every mL of water, adjust pH to 8.5. store at -20 °C if not in use.

2: freshly dissolve 250 µg dye in 14 µL DMSO

3: add the following to the tube in step 2, mix these gently at room temperature for at least 6 hours

\*7 $\mu$ L water

\*75 $\mu$ L buffer

\*4  $\mu$ L 25  $\mu$ g/ $\mu$ L RNA

4: purification by PAGE

iii: label dyes from Molecular Probe to thiol modified RNA

1: dissolve thiol modified RNA at 50-100  $\mu$ M at room temperature in suitable buffer (10-100 mM phosphate, tris or HEPES buffer pH 7-7.5).

2: add DTT (or TCEP if using maleimide) to keep thiol in reduced form, 24hours

3: Nap-5 desalting/buffer exchange column to clean up DTT (TCEP)

4: make fresh 1-10 mM dye solution in DMSO or water.

5: mix and react at 4°C or room temperature for two hours

5: purification by PAGE

## REFERENCES

1. Watson, J. D.; Crick, F. H., Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **1953**, 171, (4356), 737-8.
2. Voet, D.; Voet, J. G.; Pratt, C. W., *Fundamentals of Biochemistry*. First ed.; John Wiley @ Sons, Inc.: 1999; p 728.
3. Watson, J. D.; Crick, F. H., Genetical implications of the structure of deoxyribonucleic acid. *Nature* **1953**, 171, (4361), 964-7.
4. Volkin, E.; Astrachan, L., Phosphorus incorporation in Escherichia coli ribo-nucleic acid after infection with bacteriophage T2. *Virology* **1956**, 2, (2), 149-61.
5. Jacob, F.; Monod, J., Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol* **1961**, 3, 318-56.
6. Palade, G. E., A small particulate component of the cytoplasm. *J Biophys Biochem Cytol* **1955**, 1, (1), 59-68.
7. Rodnina, M. V.; Beringer, M.; Wintermeyer, W., How ribosomes make peptide bonds. *Trends Biochem Sci* **2007**, 32, (1), 20-6.
8. Kiss, T.; Fayet, E.; Jady, B. E.; Richard, P.; Weber, M., Biogenesis and intranuclear trafficking of human box C/D and H/ACA RNPs. *Cold Spring Harb Symp Quant Biol* **2006**, 71, 407-17.
9. Bachellerie, J. P.; Cavaille, J.; Huttenhofer, A., The expanding snoRNA world. *Biochimie* **2002**, 84, (8), 775-90.
10. Woese, C., *The Genetic Code*. Harper and Row: New York, 1967.
11. Kruger, K.; Grabowski, P. J.; Zaug, A. J.; Sands, J.; Gottschling, D. E.; Cech, T. R., Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. *Cell* **1982**, 31, (1), 147-57.
12. Guerrier-Takada, C.; Gardiner, K.; Marsh, T.; Pace, N.; Altman, S., The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **1983**, 35, (3 Pt 2), 849-57.
13. McKay, D.; Wedekind, J., Small ribozymes. In *The RNA world, second edition*, Gesteland, R.; Cech, T. R.; Atkins, W. M., Eds. Cold Spring Harbor Laboratory Press: 1999.

14. Gilbert, W., Origin of life: The RNA world. *Nature* **1986**, 319, (618), doi:10.1038/319618a0.
15. Lee, R. C.; Feinbaum, R. L.; Ambros, V., The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* **1993**, 75, (5), 843-54.
16. Fire, A.; Xu, S.; Montgomery, M. K.; Kostas, S. A.; Driver, S. E.; Mello, C. C., Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **1998**, 391, (6669), 806-11.
17. Valadkhan, S., The spliceosome: a ribozyme at heart? *Biol Chem* **2007**, 388, (7), 693-7.
18. Brow, D. A., Allosteric cascade of spliceosome activation. *Annu Rev Genet* **2002**, 36, 333-60.
19. Brow, D. A.; Guthrie, C., Spliceosomal RNA U6 is remarkably conserved from yeast to mammals. *Nature* **1988**, 334, (6179), 213-8.
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. 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.
22. Valadkhan, S.; Manley, J. L., Splicing-related catalysis by protein-free snRNAs. *Nature* **2001**, 413, (6857), 701-7.
23. 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.
24. 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.
25. 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.
26. Fabrizio, P.; Abelson, J., Thiophosphates in yeast U6 snRNA specifically affect pre-mRNA splicing in vitro. *Nucleic Acids Res* **1992**, 20, (14), 3659-64.
27. Newby, M. I.; Greenbaum, N. L., Sculpting of the spliceosomal branch site recognition motif by a conserved pseudouridine. *Nat Struct Biol* **2002**, 9, (12), 958-65.
28. Newby, M. I.; Greenbaum, N. L., A conserved pseudouridine modification in eukaryotic U2 snRNA induces a change in branch-site architecture. *RNA* **2001**, 7, (6), 833-45.
29. Fedorova, O.; Zingler, N., Group II introns: structure, folding and splicing mechanism. *Biol Chem* **2007**, 388, (7), 665-78.

30. 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.
31. Valadkhan, S.; Mohammadi, A.; Wachtel, C.; Manley, J. L., Protein-free spliceosomal snRNAs catalyze a reaction that resembles the first step of splicing. *RNA* **2007**, 13, (12), 2300-11.
32. Golas, M. M.; Sander, B.; Will, C. L.; Luhrmann, R.; Stark, H., Molecular architecture of the multiprotein splicing factor SF3b. *Science* **2003**, 300, (5621), 980-4.
33. Draper, D. E., A guide to ions and RNA structure. *RNA* **2004**, 10, (3), 335-43.
34. Fedor, M. J.; Williamson, J. R., The catalytic diversity of RNAs. *Nat Rev Mol Cell Biol* **2005**, 6, (5), 399-412.
35. Draper, D. E.; Grilley, D.; Soto, A. M., Ions and RNA folding. *Annu Rev Biophys Biomol Struct* **2005**, 34, 221-43.
36. 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.
37. Valadkhan, S.; Manley, J. L., A tertiary interaction detected in a human U2-U6 snRNA complex assembled in vitro resembles a genetically proven interaction in yeast. *RNA* **2000**, 6, (2), 206-19.
38. 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.
39. Stryer, L.; Haugland, R. P., Energy transfer: a spectroscopic ruler. *Proc Natl Acad Sci U S A* **1967**, 58, (2), 719-26.
40. Selvin, P. R.; Hearst, J. E., Luminescence energy transfer using a terbium chelate: improvements on fluorescence energy transfer. *Proc Natl Acad Sci U S A* **1994**, 91, (21), 10024-8.
41. Selvin, P. R., Principles and biophysical applications of lanthanide-based probes. *Annu Rev Biophys Biomol Struct* **2002**, 31, 275-302.
42. Periasamy, A.; Day, N., *Molecular Imaging: FRET Microscopy and Spectroscopy*. Oxford University Press: New York, Oxford, 2005.
43. Zhuang, X., Single-molecule RNA science. *Annu Rev Biophys Biomol Struct* **2005**, 34, 399-414.
44. Jares-Erijman, E. A.; Jovin, T. M., Imaging molecular interactions in living cells by FRET microscopy. *Curr Opin Chem Biol* **2006**, 10, (5), 409-16.
45. Giepmans, B. N.; Adams, S. R.; Ellisman, M. H.; Tsien, R. Y., The fluorescent toolbox for assessing protein location and function. *Science* **2006**, 312, (5771), 217-24.

46. Clegg, R. M., Fluorescence resonance energy transfer and nucleic acids. *Methods Enzymol* **1992**, 211, 353-88.
47. Fedor, M. J., The role of metal ions in RNA catalysis. *Curr Opin Struct Biol* **2002**, 12, (3), 289-95.
48. 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.
49. He, B.; Rong, M.; Lyakhov, D.; Gartenstein, H.; Diaz, G.; Castagna, R.; McAllister, W. T.; Durbin, R. K., Rapid mutagenesis and purification of phage RNA polymerases. *Protein Expr Purif* **1997**, 9, (1), 142-51.
50. Greenbaum, N. L.; Mundoma, C.; Peterman, D. R., Probing of metal-binding domains of RNA hairpin loops by laser-induced lanthanide(III) luminescence. *Biochemistry* **2001**, 40, (4), 1124-34.
51. Matsumura, K.; Komiyama, M., Enormously fast RNA hydrolysis by lanthanide(III) ions under physiological conditions: eminent candidates for novel tools of biotechnology. *J Biochem (Tokyo)* **1997**, 122, (2), 387-94.
52. Costa, D.; Burrows, H. D.; da Graca Miguel, M., Changes in hydration of lanthanide ions on binding to DNA in aqueous solution. *Langmuir* **2005**, 21, (23), 10492-6.
53. Huang, C. Y., Determination of binding stoichiometry by the continuous variation method: the Job plot. *Methods Enzymol* **1982**, 87, 509-25.
54. Clegg, R., Fluorescence Resonance Energy Transfer and Nucleic Acids. *Methods enzymol.* **1992**, 211, 352-388.
55. Norman, D. G.; Grainger, R. J.; Uhrin, D.; Lilley, D. M., Location of cyanine-3 on double-stranded DNA: importance for fluorescence resonance energy transfer studies. *Biochemistry* **2000**, 39, (21), 6317-24.
56. Xu, D.; Greenbaum, N. L.; Fenley, M. O., Recognition of the spliceosomal branch site RNA helix on the basis of surface and electrostatic features. *Nucleic Acids Res* **2005**, 33, (4), 1154-61.
57. Chen, X.; McDowell, J. A.; Kierzek, R.; Krugh, T. R.; Turner, D. H., Nuclear magnetic resonance spectroscopy and molecular modeling reveal that different hydrogen bonding patterns are possible for G.U pairs: one hydrogen bond for each G.U pair in r(GGCGUGCC)(2) and two for each G.U pair in r(GAGUGCUC)(2). *Biochemistry* **2000**, 39, (30), 8970-82.
58. Wolfson, J. M.; Kearns, D. R., Europium as a fluorescent probe of transfer RNA structure. *Biochemistry* **1975**, 14, (7), 1436-44.
59. Marciniec, T.; Ciesiolka, J.; Wrzesinski, J.; Krzyzosiak, W. J., Identification of the magnesium, europium and lead binding sites in E. coli and lupine tRNAPhe by specific metal ion-induced

- cleavages. *FEBS Lett* **1989**, 243, (2), 293-8.
60. Hargittai, M. R.; Musier-Forsyth, K., Use of terbium as a probe of tRNA tertiary structure and folding. *RNA* **2000**, 6, (11), 1672-80.
61. Mundoma, C.; Greenbaum, N. L., Sequestering of Eu<sup>3+</sup> by a GAAA RNA tetraloop. *J Am Chem Soc* **2002**, 124, (14), 3525-32.
62. Feig, A. L.; Panek, M.; Horrocks, W. D., Jr.; Uhlenbeck, O. C., Probing the binding of Tb<sup>3+</sup> and Eu<sup>3+</sup> to the hammerhead ribozyme using luminescence spectroscopy. *Chem Biol* **1999**, 6, (11), 801-10.
63. Misra, V. K.; Draper, D. E., On the role of magnesium ions in RNA stability. *Biopolymers* **1998**, 48, (2-3), 113-35.
64. Horrocks, W. D., Jr.; Sudnick, D. R., Lanthanide ion probes of structure in biology. Laser induced luminescence decay constants provide a direct measure of the number of metal coordinated water molecules. *J Am Chem Soc* **1979**, (101), 334-340.
65. Blad, H.; Reiter, N. J.; Abildgaard, F.; Markley, J. L.; Butcher, S. E., Dynamics and metal ion binding in the U6 RNA intramolecular stem-loop as analyzed by NMR. *J Mol Biol* **2005**, 353, (3), 540-55.
66. 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.
67. Butcher, S. E., Structure and function of the small ribozymes. *Curr Opin Struct Biol* **2001**, 11, (3), 315-20.
68. Ferre-D'Amare, A. R., The hairpin ribozyme. *Biopolymers* **2004**, 73, (1), 71-8.
69. Evans, C. H., *Biochemistry of the Lanthanides*. Plenum Press: New York, 1990.
70. Bunzli, J.; Piguet, C., Taking advantage of luminescent lanthanide ions. *Chem Soc Rev* **2005**, 34, (12), 1048-1077.
71. Lee, L.; Sykes, B. D., Use of Lanthanide-Induced Nuclear Magnetic-Resonance Shifts for Determination of Protein-Structure in Solution - Ef Calcium-Binding Site of Carp Parvalbumin. *Biochemistry* **1983**, 22, (19), 4366-4373.
72. Mundoma, C.; Greenbaum, N. L., Binding of europium(III) ions to RNA stem loops: role of the primary hydration sphere in complex formation. *Biopolymers* **2003**, 69, (1), 100-9.
73. Pan T., L. D. M., and Uhlenbeck O.C., Divalent metal ions in RNA folding and catalysis. In *The RNA World*, Gesteland R. F., A. J. F., Ed. Cold Spring Harbor Press: New York, 1993; pp 271-302.

74. Jack, A.; Ladner, J. E.; Rhodes, D.; Brown, R. S.; Klug, A., Crystallographic Study of Metal-Binding to Yeast Phenylalanine Transfer-Rna. *Journal of Molecular Biology* **1977**, 111, (3), 315-328.
75. Horrocks, W. D., Jr.; Schmidt, G. F.; Sudnick, D. R.; Kittrell, C.; Bernheim, R. A., Laser-induced lanthanide ion luminescence lifetime measurements by direct excitation of metal ion levels. A new class of structural probe for calcium-binding proteins and nucleic acids. *J Am Chem Soc* **1977**, 99, (7), 2378-80.
76. Cotton, F.; Wilkinson, G.; Grimes, R., *Advanced Inorganic Chemistry*. Fifth ed.; John Wiley & Sons Inc: 1988; p 1488.
77. Horrocks, W. D., Jr., Luminescence spectroscopy. *Methods Enzymol* **1993**, 226, 495-538.
78. Horrocks, W. D., Jr.; Sudnick, D. R., Time-resolved europium(III) excitation spectroscopy: a luminescence probe of metal ion binding sites. *Science* **1979**, 206, (4423), 1194-6.
79. Thomas, D. D.; Carlsen, W. F.; Stryer, L., Fluorescence energy transfer in the rapid-diffusion limit. *Proc Natl Acad Sci U S A* **1978**, 75, (12), 5746-5750.
80. Hermann, T.; Patel, D. J., RNA bulges as architectural and recognition motifs. *Structure* **2000**, 8, (3), R47-54.
81. Tinoco, I., Jr.; Bustamante, C., How RNA folds. *J Mol Biol* **1999**, 293, (2), 271-81.
82. Woese, C. R.; Gutell, R.; Gupta, R.; Noller, H. F., Detailed analysis of the higher-order structure of 16S-like ribosomal ribonucleic acids. *Microbiol Rev* **1983**, 47, (4), 621-69.
83. Schlatterer, J. C.; Crayton, S. H.; Greenbaum, N. L., Conformation of the Group II intron branch site in solution. *J Am Chem Soc* **2006**, 128, (12), 3866-7.
84. Matsumura, K.; Komiyama, M., Enormously fast RNA hydrolysis by lanthanide(III) ions under physiological conditions: eminent candidates for novel tools of biotechnology. *J Biochem* **1997**, 122, (2), 387-94.
85. Xiao, M.; Selvin, P. R., Quantum yields of luminescent lanthanide chelates and far-red dyes measured by resonance energy transfer. *J Am Chem Soc* **2001**, 123, (29), 7067-73.
86. Dadlez, M.; Goral, J.; Bierzynski, A., Luminescence of peptide-bound terbium ions. Determination of binding constants. *FEBS Lett* **1991**, 282, (1), 143-6.
87. Barja, B.; Baggio, R.; Garland, M.; Aramendia, P.; Pena, O.; Pereg, M., Crystal structures and luminescent properties of terbium(III) carboxylates. *Inorganica Chimica Acta* **2003**, 346, (25), 187-196.
88. Ross, J. B.; Wyssbrod, H. R.; Porter, R. A.; Schwartz, G. P.; Michaels, C. A.; Laws, W. R., Correlation of tryptophan fluorescence intensity decay parameters with <sup>1</sup>H NMR-determined rotamer conformations: [tryptophan<sup>2</sup>]oxytocin. *Biochemistry* **1992**, 31, (6), 1585-94.

89. Erat, M. C.; Zerbe, O.; Fox, T.; Sigel, R. K., Solution Structure of Domain 6 from a Self-Splicing Group II Intron Ribozyme: A Mg(2+) Binding Site is Located Close to the Stacked Branch Adenosine. *ChemBiochem* **2007**, 8, (3), 306-314.
90. Jain, S.; Welch, J. T.; Horrocks, W. D., Jr.; Franklin, S. J., Europium luminescence of EF-hand helix-turn-helix chimeras: impact of pH and DNA-binding on europium coordination. *Inorg Chem* **2003**, 42, (24), 8098-104.
91. Sculimbrenne, B. R.; Imperiali, B., Lanthanide-binding tags as luminescent probes for studying protein interactions. *J Am Chem Soc* **2006**, 128, (22), 7346-52.
92. Yan, D.; Ares, M., Jr., Invariant U2 RNA sequences bordering the branchpoint recognition region are essential for interaction with yeast SF3a and SF3b subunits. *Mol Cell Biol* **1996**, 16, (3), 818-28.
93. 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.
94. Fortner, D. M.; Troy, R. G.; Brow, D. A., A stem/loop in U6 RNA defines a conformational switch required for pre-mRNA splicing. *Genes Dev* **1994**, 8, (2), 221-33.
95. Luukkonen, B. G.; Seraphin, B., A role for U2/U6 helix Ib in 5' splice site selection. *RNA* **1998**, 4, (8), 915-27.
96. Field, D. J.; Friesen, J. D., Functionally redundant interactions between U2 and U6 spliceosomal snRNAs. *Genes Dev* **1996**, 10, (4), 489-501.
97. 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.
98. McManus, C. J.; Schwartz, M. L.; Butcher, S. E.; Brow, D. A., A dynamic bulge in the U6 RNA internal stem-loop functions in spliceosome assembly and activation. *RNA* **2007**, 13, (12), 2252-65.
99. Lilley, D. M., The origins of RNA catalysis in ribozymes. *Trends Biochem Sci* **2003**, 28, (9), 495-501.
100. Koshland, D. E., Application of a Theory of Enzyme Specificity to Protein Synthesis. *Proc Natl Acad Sci U S A* **1958**, 44, (2), 98-104.
101. Newby, M. I.; Greenbaum, N. L., Investigation of Overhauser effects between pseudouridine and water protons in RNA helices. *Proc Natl Acad Sci U S A* **2002**, 99, (20), 12697-702.
102. Butcher, S. E.; Brow, D. A., Towards understanding the catalytic core structure of the spliceosome. *Biochem Soc Trans* **2005**, 33, (Pt 3), 447-9.
103. 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.

104. Ryan, D. E.; Kim, C. H.; Murray, J. B.; Adams, C. J.; Stockley, P. G.; Abelson, J., New tertiary constraints between the RNA components of active yeast spliceosomes: a photo-crosslinking study. *RNA* **2004**, 10, (8), 1251-65.
105. Madhani, H. D.; Guthrie, C., Randomization-selection analysis of snRNAs in vivo: evidence for a tertiary interaction in the spliceosome. *Genes Dev* **1994**, 8, (9), 1071-86.
106. Yun, C. S.; Javier, A.; Jennings, T.; Fisher, M.; Hira, S.; Peterson, S.; Hopkins, B.; Reich, N. O.; Strouse, G. F., Nanometal surface energy transfer in optical rulers, breaking the FRET barrier. *Journal of the American Chemical Society* **2005**, 127, (9), 3115-3119.
107. Chance, R.; Prock, A.; R, S., Molecular fluorescence and energy transfer near interfaces. *Advances in Chemical Physics* **1978**, 37, 1-65.
108. Jennings, T. L.; Schlatterer, J. C.; Singh, M. P.; Greenbaum, N. L.; Strouse, G. F., NSET molecular beacon analysis of hammerhead RNA substrate binding and catalysis. *Nano Letters* **2006**, 6, (7), 1318-1324.
109. Jennings, T. L.; Singh, M. P.; Strouse, G. F., Fluorescent lifetime quenching near  $d = 1.5$  nm gold nanoparticles: probing NSET validity. *J Am Chem Soc* **2006**, 128, (16), 5462-7.
110. Will, C.; Luhrmann, R., Spliceosome structure and function. In *The RNA world*, Gesteland, R. F., Cech, T. R., and Atkins, J. F., Ed. 2006; pp pp.369-400.
111. Schroeder, K. T. Structural investigation of RNA-RNA and RNA-protein interactions involving the pre-mRNA branch site region of the functional core of the spliceosome. Dissertation, Florida State University, Tallahassee, Florida, 2006.
112. Jennings, T. L.; Schlatterer, J.; Strouse, G. F.; Greenbaum, N. L., Using nanomaterials to hammerhead a new frontier. *Abstracts of Papers of the American Chemical Society* **2005**, 230, U522-U522.
113. Watrob, H. M.; Pan, C. P.; Barkley, M. D., Two-step FRET as a structural tool. *J Am Chem Soc* **2003**, 125, (24), 7336-43.
114. Lilley, D. M.; Clegg, R. M., The structure of the four-way junction in DNA. *Annu Rev Biophys Biomol Struct* **1993**, 22, 299-328.

## BIOGRAPHICAL SKETCH

Faqing Yuan was born in a small village called “Yuan Jia Yan” in Yunmeng County, Hubei Province, China to Zhengxin Yuan and Dongmei Leng on a hot summer day of 1978. He has a brother who is 9 years older. He received education in his hometown until he was 19 years old before attending Wuhan University in 1997. After receiving a Bachelor’s degree in biochemistry (biotechnology) in 2001, he came to USA to pursue a PhD degree in the Department of Chemistry and Biochemistry at the Florida State University later that year. His mentor is Dr. Nancy L. Greenbaum who has expertise in RNA structure research. In the summer of 2007, Dr. Greenbaum accepted a position in Hunter College of the City University of New York and Dr. Geoffrey F. Strouse kindly agreed to be his co-advisor. Faqing has accepted a postdoctoral position with Dr. Peixuan Guo in the University of Cincinnati and will work there for a few years before developing his own independent career.