

## ABSTRACT

Title of Document: FLUORESCENCE AND NMR STUDIES OF  
THE ROLE OF METAL IONS IN HIV-1  
GENOMIC RNA DIMERIZATION AND  
MATURATION

Hui-Wen Lee, Ph.D., 2009

Directed By: Professor D. Fushman, Department of Chemistry  
and Biochemistry

The dimerization initiation site (DIS) is an essential RNA element responsible for dimerization of HIV-1 genomic RNA through a kissing loop interaction. The DIS loop contains six auto-complementary nucleotides stabilized by 5'- and 3'-flanking purines. NCp7 chaperone protein catalyzes conversion of an intermediate DIS kissing dimer to a more thermodynamically stable extended duplex dimer in the presence of  $Mg^{2+}$ . Sequence constructs intended to model the extended duplex, (DIS 21), and the kissing dimer, DIS23(GA)•DIS23(HxUC), were designed to examine the structural information and biochemical behaviors during maturation. We introduced the fluorescent labeling, 2-aminopurine (2-AP) into these RNA constructs, to finely probe structural transition and local dynamics accompanied by the formation of the DIS dimer. The 2-AP nucleotides were inserted either in the DIS loop or junction to study loop-loop interaction or purine base stacking conformation at the junction responding to the metal ion effect. High resolution NMR methods were then used to probe structural changes associated with mono versus divalent cation binding to the DIS dimers and also determine the  $Mg^{2+}$

binding sites. Significant chemical shift perturbations (CSP) were found upon  $\text{Mg}^{2+}$  binding and used to map structural changes. Further  $\text{Mn}^{2+}$  paramagnetic relaxation enhancement (PRE) experiments provided evidence for specific  $\text{Mg}^{2+}$  ion binding are localized around the 5' purine bases in both the extended duplex and kissing dimers with profound line broadening effects. Mapping the CSP and PRE data onto the available X-ray crystal and NMR solution structures allowed localization of specific  $\text{Mg}^{2+}$  ions at binding sites on the DIS dimers created by the unpaired flanking DIS loop purine nucleotides. Our data indicates that the conformations that are metal cation dependent. These findings are consistent with previous results that suggested a role for divalent metal cations in stabilizing the DIS kissing dimer structure and influencing its maturation to an extended duplex form through interactions with the DIS loop.

FLUORESCENCE AND NMR STUDIES OF THE ROLE OF METAL IONS IN  
HIV-1 GENOMIC RNA DIMERIZATION AND MATURATION

By

Hui-Wen Lee

Dissertation submitted to the Faculty of the Graduate School of the  
University of Maryland, College Park, in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy  
2009

Advisory Committee:  
Professor David Fushman, Chair  
Professor Stephen M. Mount  
Professor Jeffery Davis  
Professor Jason Kahn  
Professor John P. Marino

© Copyright by  
Hui-Wen Lee  
2009



## **Dedication**

I lovingly dedicate this doctoral dissertation to my parents for giving me so many wonderful opportunities through out my life.

## Acknowledgements

I would like to acknowledge the inspirational instruction and guidance of Dr. John Marino. I appreciate the opportunity he gave me through NMR studies to realize the beauty of the structural biology of RNA. He devotes his knowledge, time, and energy to the world of RNA which motivates many others to dedicate their lives to scientific research.

Dr. David Frushman, my co-advisor, is a well respected researcher in protein chemistry NMR studies. Besides the strong foundation he gave me for my work with NMR, I thoroughly enjoyed the course's mathematical and physical focus.

Dr. Jason Kahn provided a very detailed critique of my dissertation. He was extremely helpful in the preparation of my presentation formats. His opinions were tremendously useful and helped me to complete my research.

Dr. Stephen Mount and Dr. Jeffery Davis served on my committee, and I am sincerely grateful for their valuable assistance.

I would like to thank Dr. Nese Sari, Dr. Robert Brinson, and Dr. Raviprasad Aduri who instructed me in the very complicated NMR lab techniques which were a very important part of my research. I thank Dr. Justyna Mysliwy, Dr. Nadik Abdoulaev, Dr. Gary Giulian, and Katie Briggs who contributed their ideas, feedback, and advice not only in the academic arena but also provided cultural enlightenment.

To everyone else who contributed their efforts to this project, I offer my respectful appreciation and gratitude.

Certainly not least, I thank my parents for their endless support, encouragement, patience, and their open mindedness that allow me to reach this goal. They have my undying love and respect.

# Table of Contents

|  |     |
|--|-----|
| Dedication.....  | ii  |
| Acknowledgements.....  | iii |
| Table of Contents.....   | v   |
| List of Tables.....  | ix  |
| List of Figures.....   | x   |
| List of Abbreviations.....   | xv  |
| Chapter 1 Introduction .....   | 1   |
| 1.1 Overview of HIV-1 Biology.....   | 1   |
| 1.1.1 Replication Cycle.....   | 1   |
| 1.1.2 Genomic RNA Dimerization and Maturation.....   | 3   |
| 1.1.2.1 Dimerization Initiation Site (DIS) Structure.....  | 3   |
| 1.1.2.2 Packaging and Assembly .....   | 8   |
| 1.1.2.3 Nucleocapsid Protein (NCp7).....   | 11  |
| 1.1.2.4 DIS Kissing Dimer – Extended Duplex Conversion .....   | 14  |
| 1.2 Role of Metal Ions in RNA Structure and Function .....   | 15  |
| 1.2.1 Di- versus Mono-Valent Cation Interactions.....  | 15  |
| 1.2.2 Metal Binding Pockets versus Diffuse Cloud Interactions.....   | 17  |
| Chapter 2 2-AP Fluorescence Detection of RNA Structural Changes in Response to<br>Metal Binding the Dimerization Initiation Stem-Loop Kissing Complex..... | 19  |
| 2.1 Abstract.....  | 19  |
| 2.2 Introduction.....  | 19  |
| 2.2.1 Fluorescence Emission / Quenching Experiments.....   | 19  |

|           |  |    |
|-----------|--|----|
| 2.2.2     | Fluorescence versus UV Detected Thermal RNA Melting Curves .....   | 21 |
| 2.2.3     | Sample Preparation & 2-AP Labeling .....   | 22 |
| 2.3       | Material and Methods .....   | 23 |
| 2.3.1     | Design of the DIS RNA Hairpins .....   | 25 |
| 2.3.2     | Fluorescence Experiments .....   | 28 |
| 2.4       | Results.....   | 29 |
| 2.4.1     | Fluorescence Melting Curve Analysis of the DIS Kissing Complexes —<br>Distinguishing Different Helical and Junction Melting Events.....  | 29 |
| 2.4.2     | Measurement of the Site Specific 2-AP Responses to Metal Ion<br>Interactions: Response of the Two Unpaired Purine Bases on the 5' Side of DIS<br>Loop in Kissing Complex ..... | 41 |
| 2.5       | Discussion.....  | 49 |
| 2.6       | Conclusion .....   | 52 |
| Chapter 3 | Resonance Assignment of the DIS Extended Duplex and Kissing Dimer<br>Structures Using High-Resolution NMR.....   | 53 |
| 3.1       | Abstract.....  | 53 |
| 3.2       | Introduction.....  | 54 |
| 3.2.1     | Nuclear Magnetic Resonance Spectroscopy Applied to RNA .....   | 54 |
| 3.2.1.1   | Sample Preparation & Isotope Labeling.....   | 54 |
| 3.2.1.2   | NMR Experiments for Resonance Assignment in RNA .....  | 55 |
| 3.2.1.3   | NMR Structure Determination of RNA.....  | 60 |
| 3.2.2     | Analysis of the DIS Extended Duplex and Kissing Dimer Structures<br>Using High- Resolution NMR.....  | 62 |

|   |     |
|---|-----|
| 3.3 Material & Methods.....   | 65  |
| 3.4 Results and Discussion .....  | 68  |
| 3.4.1 Structural Assignments .....  | 68  |
| 3.5 Conclusion .....  | 105 |
| Chapter 4 Characterization of Mg <sup>2+</sup> Binding and Associated Conformational<br>Changes in the DIS Extended Duplex and Kissing Dimer Structures.....              | 106 |
| 4.1 Abstract.....   | 106 |
| 4.2 Introduction.....   | 107 |
| 4.2.1 Detecting Metal Binding to RNA .....  | 107 |
| 4.2.1.1 Chemical Shift Perturbation.....  | 107 |
| 4.2.1.2 Paramagnetic Relaxation Enhancement (Manganese Based PRE-<br>NMR) .....   | 108 |
| 4.2.1.3 Cobalt Hexamine .....   | 109 |
| 4.2.2 Identification and Comparison of Sites Found for High-Affinity Mg <sup>2+</sup><br>Binding on the DIS Extended Duplex and Kissing Dimer Structures .....            | 110 |
| 4.3 Material and Methods .....  | 113 |
| 4.4 Results and Discussion .....  | 114 |
| 4.4.1 NMR Methods to Identify and Compare Sites Found for High-Affinity<br>Mg <sup>2+</sup> Binding Sites on the DIS Extended Duplex and Kissing Dimer<br>Structures..... | 114 |
| 4.4.1.1 Chemical Shift Perturbation (CSP).....  | 114 |
| 4.4.1.2 Paramagnetic Relaxation Enhancement –NMR .....  | 121 |
| 4.5 Conclusion .....  | 144 |

|                                   |     |
|-----------------------------------|-----|
| Chapter 5 Concluding Remarks..... | 145 |
| Bibliography .....                | 150 |

## List of Tables

|           |  |     |
|-----------|--|-----|
| Table 2.1 | Melting temperatures of DIS24(GA)-12ap • DIS23(HxUC) kissing complex: metal ions, metal ion concentration, and pH dependence.....                      | 50  |
| Table 2.2 | Summary of fluorescence of the kissing complexes DIS24(GA)-9ap • DIS23(HxUC) and DIS24(GA)-10ap • DIS23(HxUC): metal ion dependence and pH effect..... | 51  |
| Table 3.1 | Chemical shift ranges for imino and amino resonances.....  | 57  |
| Table 3.2 | Chemical shift range for non-exchangeable proton.....  | 58  |
| Table 3.3 | Assignments of homonuclear and heteronuclear correlations.....   | 69  |
| Table 3.4 | Assignments of DIS21 duplex with Na <sup>+</sup> .....   | 84  |
| Table 3.5 | Assignments of DIS21 duplex with Mg <sup>2+</sup> and Na <sup>+</sup> .....  | 85  |
| Table 3.6 | Assignments of DIS23(GA) in the DIS23(GA)•DIS23(HxUC) with Mg <sup>2+</sup> and Na <sup>+</sup> .....  | 104 |
| Table 4.1 | Summary of RMS values for differences in chemical shifts of DIS21 duplex induced by Mg <sup>2+</sup> .....   | 119 |
| Table 4.2 | Summary of peak height ratios at the local DIS junction of DIS21 duplex.....   | 130 |
| Table 4.3 | Summary of peak height ratios at the local DIS junction of DIS23(GA)•DIS23(HxUC).....  | 140 |



## List of Figures

|         |   |    |
|---------|---|----|
| Fig 1.1 | The HIV-1 life cycle.....   | 3  |
| Fig 1.2 | The secondary structure of DIS (Mal) and DIS (Lai).....   | 4  |
| Fig 1.3 | X-ray crystal structures of the DIS dimers.....   | 6  |
| Fig 1.4 | The sequence and secondary structure of the 5' end of HIV-1.....  | 9  |
| Fig 1.5 | The $\Psi$ -site packaging sequence in the HIV-1 genomic RNA.....   | 10 |
| Fig 1.6 | The fragments of gag protein.....   | 10 |
| Fig 1.7 | The sequence and secondary structure of NCp7 protein.....   | 13 |
| Fig 2.1 | The adenosine-uracil and the 2-aminopurine-uracil base pairs.....   | 23 |
| Fig 2.2 | RNA sequences and secondary structures of the DIS homo- and hetero-<br>dimers.....                                    | 27 |
| Fig 2.3 | Fluorescence melting curves of DIS24(GA)-12ap in the presence of<br>$Mg^{2+}$ versus $Na^+$ .....                     | 31 |
| Fig 2.4 | Fluorescence melting curve of DIS24(GA)-12ap•DIS23(HxUC) kissing<br>complex with 0.5 mM $Mg^{2+}$ .....               | 33 |
| Fig 2.5 | The sigmoidal melting curve DIS24(GA)-12ap • DIS23(HxUC) kissing<br>complex with 0.5 mM $Mg^{2+}$ .....               | 34 |
| Fig 2.6 | Normalized derivative of melting curve of DIS24(GA)- 12ap•<br>DIS23(HxUC) kissing complex with 0.5 mM $Mg^{2+}$ ..... | 35 |
| Fig 2.7 | Fluorescence melting curve of DIS24(GA)-12ap•DIS23(HxUC)<br>kissing complex with 200 mM $Na^+$ .....                  | 36 |
| Fig 2.8 | Normalized derivative of melting curve of DIS24(GA)-12ap•   |    |

|          |  |    |
|----------|--|----|
|          | DIS23(HxUC) kissing complex with 200 mM Na <sup>+</sup> .....  | 37 |
| Fig 2.9  | Fluorescence melting curves of DIS24(GA)-12ap•DIS23(HxUC) kissing complex with 0.5 mM Mg <sup>2+</sup> and 5 mM Mg <sup>2+</sup> .....             | 38 |
| Fig 2.10 | Normalized derivatives of melting curves of DIS24(GA)-12ap•DIS23(HxUC) kissing complex with 0.5 mM Mg <sup>2+</sup> and 5mM Mg <sup>2+</sup> ..... | 39 |
| Fig 2.11 | Fluorescence melting curves of DIS24(GA)-12ap•DIS23(HxUC) kissing complex with 0.5 mM Mg <sup>2+</sup> at different pH values.....                 | 40 |
| Fig 2.12 | Fluorescence melting curves of DIS24(GA)-12ap•DIS23(HxUC) kissing complex with 200 mM Na <sup>+</sup> at different pH.....                         | 41 |
| Fig 2.13 | The secondary structure of the DIS kissing complexes with 2-AP labeling at position 9 and 10 on the DIS24(GA) side.....                            | 42 |
| Fig 2.14 | The fluorescence of the kissing complex DIS24(GA)-10ap • DIS23(HxUC) with 0.5 mM Mg <sup>2+</sup> at pH 6.5.....                                   | 43 |
| Fig 2.15 | The fluorescence of the kissing complex DIS24(GA)-10ap • DIS23(HxUC) with 0.5 mM Mg <sup>2+</sup> at pH 8.0.....                                   | 44 |
| Fig 2.16 | The fluorescence of the kissing complex DIS24(GA)-10ap • DIS23(HxUC) with 200 mM Na <sup>+</sup> at pH 6.5.....                                    | 45 |
| Fig 2.17 | The fluorescence of the kissing complex DIS24(GA)-10ap • DIS23(HxUC) with 200 mM Na <sup>+</sup> at pH 8.0.....                                    | 46 |
| Fig 2.18 | The fluorescence of DIS24(GA)-9ap • DIS23(HxUC) at different pH values with 200 mM Na <sup>+</sup> .....   | 48 |
| Fig 3.1  | Chemical shift ranges for exchangeable and non-exchangeable proton...  | 57 |
| Fig 3.2  | The representation of <i>J</i> -coupling geometry.....   | 61 |

|          |  |     |
|----------|--|-----|
| Fig 3.3  | Secondary structure of DIS21.....  | 64  |
| Fig 3.4  | Schematic diagram of the formation of the heterodimer DIS23(GA)•<br>DIS23(HxUC).....   | 64  |
| Fig 3.5  | 1D imino proton spectra of the DIS21 duplex with NaCl vs. MgCl <sub>2</sub> .....  | 71  |
| Fig 3.6  | 2D NOESY spectra of the DIS 21 with Mg <sup>2+</sup> vs. Na <sup>+</sup> .....   | 73  |
| Fig 3.7  | 2D COSY/TOCSY spectra of DIS21 duplex with Mg <sup>2+</sup> vs. Na <sup>+</sup> .....  | 75  |
| Fig 3.8  | <sup>1</sup> H <sup>13</sup> C-HSQC spectrum of DIS21 duplex with Mg <sup>2+</sup> and Na <sup>+</sup> .....                       | 77  |
| Fig 3.9  | 2D <u>H</u> <u>C</u> <u>C</u> <u>H</u> -COSY spectrum of the DIS21 duplex with Mg <sup>2+</sup> and Na <sup>+</sup> ...            | 79  |
| Fig 3.10 | 2D <u>H</u> <u>C</u> <u>C</u> <u>H</u> -COSY experiment of DIS21 duplex with Mg <sup>2+</sup> and Na <sup>+</sup> ..               | 80  |
| Fig 3.11 | The NOE walk of DIS21 duplex.....  | 83  |
| Fig 3.12 | 1D imino proton spectra of DIS21 homodimer and mixture of kissing<br>and duplex.....   | 87  |
| Fig 3.13 | <sup>1</sup> H, <sup>15</sup> N-HSQC of the uniformly labeled <sup>13</sup> C <sup>15</sup> N-DIS23(GA).....                       | 88  |
| Fig 3.14 | COSY of DIS23(GA) hairpin loop.....  | 89  |
| Fig 3.15 | Amino-imino proton region of the DIS23(HxUC).....  | 91  |
| Fig 3.16 | 2D TOCSY spectrum of DIS23(HxUC).....  | 92  |
| Fig 3.17 | 1D imino proton spectra of DIS23(GA), DIS23(HxUC), and kissing<br>complex.....   | 94  |
| Fig 3.18 | The comparison of <sup>13</sup> C <sup>15</sup> N-DIS23(GA)•DIS23(HxUC) and hairpins:<br><sup>1</sup> H, <sup>15</sup> N-HSQC..... | 97  |
| Fig 3.19 | The imino-imino proton region of DIS23(GA)•DIS23(HxUC).....  | 98  |
| Fig 3.20 | 2D TOCSY of DIS23(GA)•DIS23(HxUC) with Na <sup>+</sup> and Mg <sup>2+</sup> .....  | 100 |
| Fig 3.21 | 2D CT-HSQC of <sup>13</sup> C <sup>15</sup> N-DIS23(GA)•DIS23(HxUC) with Na <sup>+</sup> and                                       |     |

|          |   |     |
|----------|---|-----|
|          | Mg <sup>2+</sup> .....  | 101 |
| Fig 3.22 | <sup>13</sup> C edited 2D NOESY of labeled DIS23(GA) and <sup>12</sup> C filtered<br>2D NOESY.....  | 102 |
| Fig 3.23 | NOE walk of DIS23(GA) side in the DIS23(GA)•DIS23(HxUC).....  | 103 |
| Fig 4.1  | A model of Na <sup>+</sup> -Mg <sup>2+</sup> -Mn <sup>2+</sup> -Mg <sup>2+</sup> competition at the specific metal ion<br>binding sites on the DIS stem-loop kissing complex..... | 112 |
| Fig 4.2  | 2D NOESY spectra of DIS21 duplex with and without Mg <sup>2+</sup> .....  | 116 |
| Fig 4.3  | Summary of Mg <sup>2+</sup> induced chemical shift perturbations of DIS21<br>duplex for each nucleobase.....  | 118 |
| Fig 4.4  | Schematic representation of bases which are perturbed by Mg <sup>2+</sup><br>in DIS21 duplex.....   | 120 |
| Fig 4.5  | 1D imino titration of Mn <sup>2+</sup> of the DIS21 duplex.....   | 123 |
| Fig 4.6  | GU wobble base pair.....  | 124 |
| Fig 4.7  | Overlay of 2D NOESY spectra with Mn <sup>2+</sup> and without Mn <sup>2+</sup> of the<br>DIS21 duplex.....  | 125 |
| Fig 4.8  | Summary of peak height ratios of DIS21 duplex due to<br>paramagnetic line broadening effects for each nucleobase.....   | 127 |
| Fig 4.9  | Schematic representation of Mg <sup>2+</sup> binding sites of DIS21 duplex.....   | 131 |
| Fig 4.10 | Divalent metal cation-binding sites in solution mapped on the<br>X-ray structure of DIS homo-duplex.....  | 134 |
| Fig 4.11 | 1D imino proton Mg <sup>2+</sup> induced chemical shift perturbation of<br>DIS23(GA)•DIS23(HxUC) kissing complex.....   | 136 |
| Fig 4.12 | Summary of peak height ratios of DIS23(GA)•   |     |

|          |   |     |
|----------|---|-----|
|          | DIS23(HxUC) due to paramagnetic line broadening effect.....   | 139 |
| Fig 4.13 | Schematic representation of $Mg^{2+}$ binding sites of<br>DIS23(GA) side of DIS23(GA)•DIS23(HxUC) kissing complex.....                              | 141 |
| Fig 4.14 | Divalent metal cation-binding sites in solution mapped on the<br>X-ray structure of DIS kissing dimer and solution RNA<br>structure comparison..... | 143 |

## List of Abbreviations

|          |  |
|----------|--|
| 2-AP     | 2-Aminopurine  |
| CA       | Capsid   |
| CD4      | Cluster of differentiation 4                         |
| CHARMM   | Chemistry at Harvard molecular mechanics simulations |
| COSY     | Correlation spectroscopy                             |
| CSP      | Chemical shift perturbation                          |
| CT-HSQC  | Constant time Heteronuclear single quantum coherence |
| DIS      | Dimerization initiation site                         |
| DLS      | Dimer linkage structure                              |
| DNA      | Deoxyribonucleic acid                                |
| DQF-COSY | Double quantum filtered correlation spectroscopy     |
| Env      | Envelope glycoprotein                                |
| FRET     | Fluorescence energy transfer                         |
| gp120    | Glycoprotein (120 kilodaltons)                       |
| gp41     | Glycoprotein (41 kilodaltons)                        |
| HIV      | Human immunodeficiency virus                         |
| HMQC     | Heteronuclear multiple quantum correlation           |
| MA       | Matrix   |
| mRNA     | Messenger ribonucleic acid                           |
| NC       | Nucleocapsid   |
| Nef      | Negative regulatory factor                           |
| NMR      | Nuclear magnetic resonance                           |

|                |  |
|----------------|--|
| NOESY          | Nuclear overhauser enhancement spectroscopy    |
| PAGE           | Preparative polyacrylamide gel electrophoresis |
| PBS            | Primer binding site                            |
| PR             | Protease                                       |
| PRE            | Paramagnetic relaxation enhancement            |
| Rev            | Regulator of virion protein expression         |
| RNA            | Ribonucleic acid                               |
| RRE            | Rev responsive element                         |
| SD             | Splice donor                                   |
| SL             | SL stem-loop                                   |
| T <sub>1</sub> | Spin lattice relaxation time                   |
| T <sub>2</sub> | Spin-spin relaxation time                      |
| Tat            | Trans-activator of transcription               |
| T <sub>m</sub> | Melting temperature                            |
| TOCSY          | Total correlation spectroscopy                 |
| UV             | Ultraviolet-visible spectroscopy               |
| Vif            | Virion infectivity factor                      |
| Vpr            | Viral protein r                                |
| Vpu            | Viral protein u                                |

# **Chapter 1 Introduction**

## **1.1 Overview of HIV-1 Biology**

### **1.1.1 Replication Cycle**

Human immunodeficiency virus type I (HIV-1) is a retrovirus, meaning that RNA carries its genetic information. Retroviruses use reverse transcriptase to convert RNA into DNA, which is then integrated into the host cell's genome. The genomic RNA in HIV-1 is composed of three main genes encoding Gag (core protein), pol (enzymes), and Env (envelope glycoproteins). Other accessory gene products, Tat, Rev, Vif, Vpu, Vpr, and Nef, are also encoded in the HIV-1 genome and carry out important functions in HIV replication. Gag functions as a core protein which is comprised of capsid protein (p24), matrix, and nucleocapsid. Pol encodes the viral enzymes protease, reverse transcriptase, and integrase. Env consists of the structural proteins such as an external glycoprotein (gp120) and a transmembrane glycoprotein (gp41). Rev is essential for nuclear export of viral RNA and binds to the Rev responsive element (RRE) RNA. Tat is a transcriptional regulator interacting with the transactivating responsive (TAR) RNA element.

During the initial stages of viral infection, the envelope protein, gp120, of HIV binds to the CD4 receptor on the surfaces of the helper T cell. Two copies of the HIV genome, as well as various enzymes such as a reverse transcriptase, are then fused from virions into the host's cell in a process mediated by HIV gp41 protein which is bound to the cell membrane. Viral genomic RNA is used as a template by the viral reverse transcriptase (RT) to produce a single stranded DNA. Single stranded DNA subsequently serves as a template to produce a cyclic double stranded DNA which then enters the



host's nucleus and is integrated into the host's DNA. The viral DNA that is integrated into the host's DNA is termed a provirus and integration is catalyzed by the viral enzyme integrase. The provirus is transcribed into mRNA which is then exported out of the nucleus and translated into viral proteins in the cytoplasm. The new generation of HIV core proteins, enzymes, and genomic RNA associate together and assemble to form a new virus particle near the cell surface. The newly formed virus is released by budding from the host's cell as an immature viral particle which then undergoes further structural changes, in a process termed maturation, to become a fully infectious virus. Maturation involves HIV-1 protease processing of the gag polyprotein to release the structural domains (matrix, capsid, etc.) and the nucleic acid chaperone protein NCp7. NCp7 release is postulated to be involved in many viral functions among which is RNA genome refolding associated with genome dimer stabilization (1).

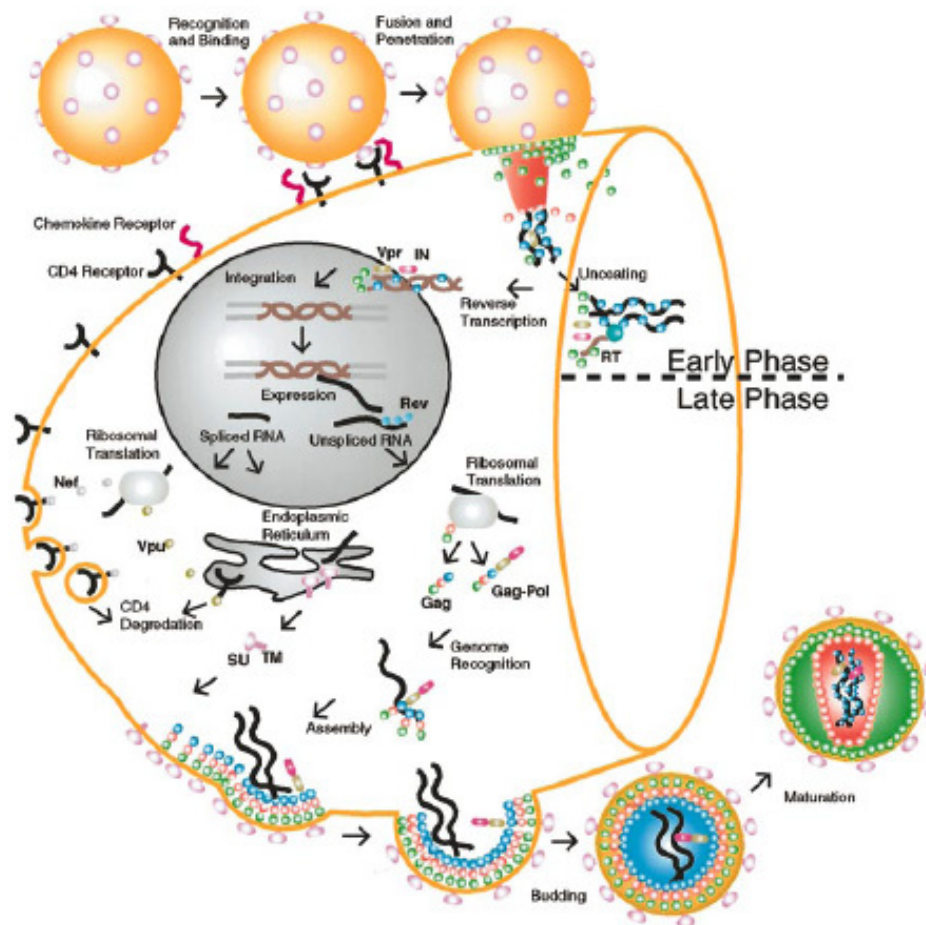


Fig 1.1. The HIV-1 life cycle (reprinted with permission (2)).

## 1.1.2 Genomic RNA Dimerization and Maturation

### 1.1.2.1 Dimerization Initiation Site (DIS) Structure

The DIS RNA sequence, which is primarily responsible for initiation of the dimerization of genomic RNA, forms a stem-loop secondary structure in the dimer linkage structure (DLS) (3-5). The DIS stem-loop can form a non-covalent kissing loop dimer through its autocomplementary hexanucleotide loop. Two specific hexanucleotide sequences, GUGCAC (Mal variant or subtype A) and GCGCGC (Lai variant or subtype

B), are observed for DIS with one 3' and two 5' unpaired flanking purines (Fig 1.2). Two steps have been proposed for the dimerization mechanism. First, recognition binding occurs through the two self-complementary hexanucleotide palindromic sequences to form a metastable kissing loop dimer (6). Second, the DIS dimer matures to form the extended duplex complex by a strand exchange (7). Structural isomerization and thermal stability of the RNA dimers at two different temperatures have been determined from analyzing complex stability using 1.5% agarose gel electrophoresis as a function of temperature (8). The results imply dimer isoforms at 37 °C and 55 °C are kinetic and thermostable isomers, respectively. At temperature 37 °C, a less stable complex, the kissing complex, is formed via base pairing of the complementary loop sequences. At temperature 55 °C, the Watson-Crick stem base pairs in the kissing dimer are broken and form an extended double stranded dimer which is denatured only at temperature greater than 75 °C.

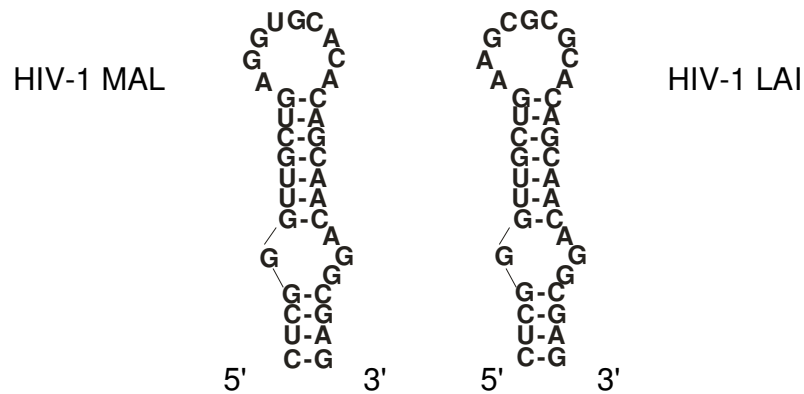


Fig 1.2. The secondary structure of DIS Mal variant and DIS Lai variant of HIV-1 RNA.

In order to understand the relation between the structure of DIS and its function, several methods have been used, including obtaining the crystal structure of the DIS stem loop dimer in both the kissing and mature forms. The X-ray structures of two truncated DIS dimers, a kissing-loop complex and an extended duplex, were determined from crystals grown at 293 and 303 K, respectively (9). The crystal structures of both the kissing complex and the extended duplex in DIS subtype B show that the unpaired purine adenosines (A272, A273) on the two strands of the duplex bulge out of the helix (10). Similarly, the 5' flanking purines A272, A273 in DIS subtype B (Lai variant) or A272, G273 in DIS subtype A (Mal variant) form bulge out and stack in crystal structures of the kissing complexes ((11, 12), Fig 1.3).

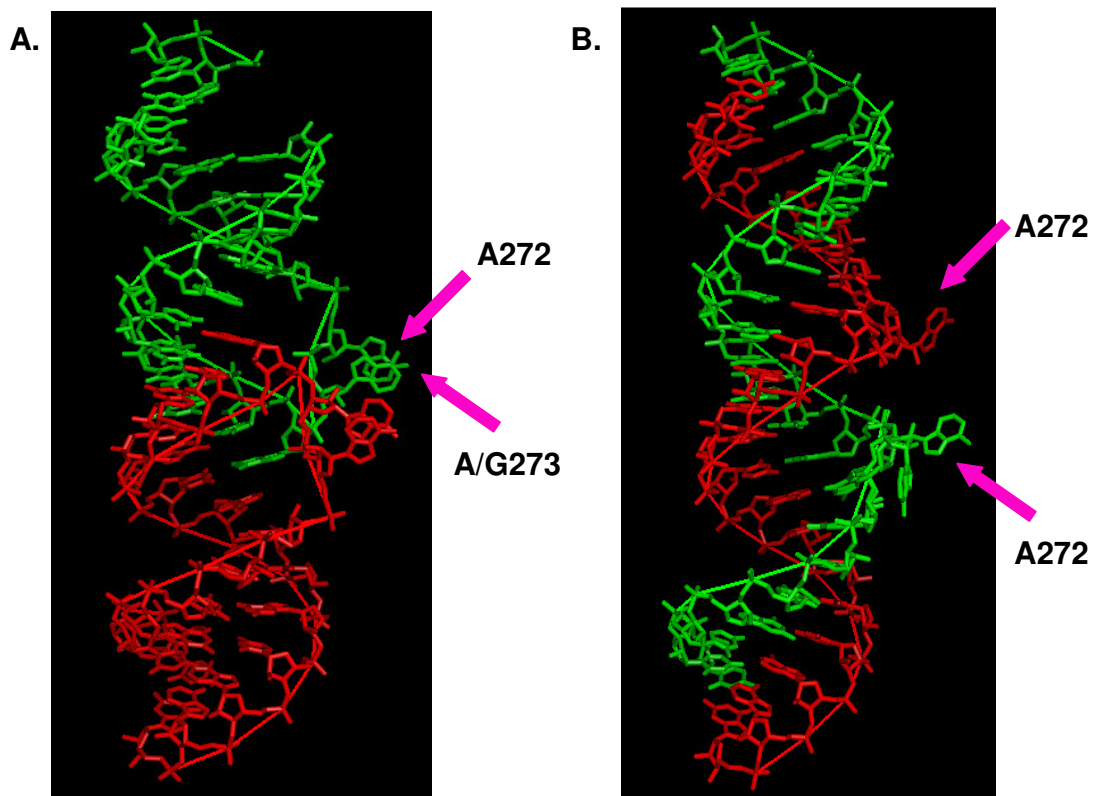


Fig 1.3. X-ray crystal structures of the (A) DIS kissing-loop complex and (B) extended duplex. The complexes are composed with two RNA strands represented in red and green. A272 and A273 (for subtype DIS-B) or G273 (for subtype DIS-A) are found in a bulge-out conformation in crystal structures of the kissing complexes. The conserved purines, A272 and A/G273, are stacked in a bulged-out conformation in the crystal structure of the extended duplex.

While the crystal structure has provided important insights into the conformation of the DIS dimers, this technique is limited in explaining function of dynamic biological molecules, especially flexible loops which are static within a lattice. In the RNA crystal, one typically observes only one possible structure while the structure may fluctuate from

one metastable state to another state under physiological conditions in solution. The X-ray crystal structures also do not directly indicate how RNA refolding occurs during dimer maturation. In order to address these questions, complementary NMR and fluorescence techniques can be used in solution to probe the RNA structure, dynamics and folding. In this respect, unlike the 5' flanking purines which are bulged-out and stack in pairs in both duplex and kissing X-ray structures, all available NMR structures for DIS isomers show that the flanking 5' purines are bulged within the helix stem (13-16). Further NMR studies by Takahashi et al. revealed that the stem bulge on DIS plays an important role in the packaging of the viral replication (17). Deletion of the bulge residues from DIS reduces the ability to form extended duplex DIS dimer. These results suggest that the bulge destabilizes the stem as part of the initiation of strand exchange during the two-step dimerization process with the stem-bulge-stem loop motif including NCp7 (17).

Modeling studies can further complement experimental results by providing information on structure, dynamics, and conformational transitions of RNA (18, 19). The CHARMM (Chemistry at HARvard Molecular Mechanics) simulations program provides a wide range of molecular simulations. CHARMM27 is especially useful for studying molecular dynamics of DNA, RNA, and lipids in solution (20). Solvent molecular dynamics (MD) simulations with the CHARMM force field showed that 5' flanking purines tend to be in bulged-in conformation, which is consistent with NMR solution structure (21). The simulations suggest that in solution the flanking purines are flexible due to molecular dynamics with different conformational states which are bulged-in and bulged-out conformations.

### **1.1.2.2 Packaging and Assembly**

The packaging of retroviral genomes, such as HIV-1, as dimers is known to be important for various critical events in the viral life cycle. RNA genome dimerization has been shown to increase the rate of genomic recombination (22) and has been suggested to increase the genetic diversity and adaptability of retroviruses (23). Dimerization of retroviral RNA genomes occurs prior to packaging into virions (7, 24) and is a central feature of all retroviruses involved in encapsidation, reverse transcription, and translation. Retroviral genome dimers are linked primarily through noncovalent interaction between 5' leader RNA sequences. The 5'-leader untranslated sequences of HIV-1 contains TAR, poly A signal, primer binding site (PBS), dimerization initiation site (DIS), splice donor (25), core packaging signal ( $\psi$ ) and AUG initiation codon of the gene Gag ((26), Fig 1.4).

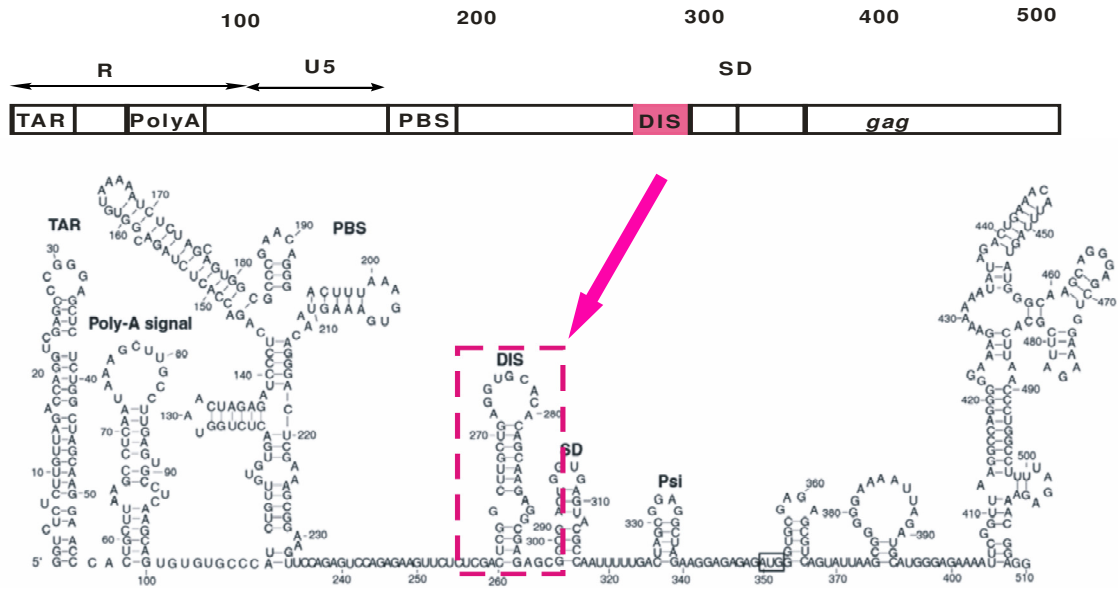


Fig 1.4. The sequence (A) and the secondary structure schematic (B) of the 5' end of HIV-1 genomic RNA. Dimerization initiation site is highlighted in a red box.

The sequences responsible for the RNA dimerization initiation overlap with the viral packaging element sequences. Specifically, the dimer initiation site stem-loop (so called DIS or SL1) of HIV-1 is located within the  $\Psi$ - packaging signal. The dimerization of two copies of the RNA genome for packaging is initiated by the DIS stem-loop which contains a G + C rich loop sequence (13). The deletion of both SL1 and SL3 also reduces genomic packaging more significantly than the deletion of either SL1 or SL3 alone. In addition, any mutations of each stem-loop SL1, SL2, SL3, or SL4 has been found to have significant effects on packaging and dimerization ((27), Fig 1.5).





The C terminal domain of CA, p1 (spacer peptide), and NC are the most important structural sequence regions in assembly. The assembly process starts by formation of a Gag p55 complex (Gag multimerization) which requires the NC region, the C-terminal of CA and the spacer peptide p2. Then, the Gag p55 complex binds to genomic RNA and is followed by formation of Gag/Gag-Pol complexes. The complexes of Gag, Gag-Pol and RNA form preassembled complexes containing small viral proteins (Vif and Vpr) and host cell proteins. Finally, preassembled complexes are transported to the plasma membrane of host cells for budding.

#### **1.1.2.3 Nucleocapsid Protein (NCp7)**

The multifunctional nucleocapsid protein (NCp7) is involved in several different steps in the HIV-1 replication cycle and has been shown to enhance the efficiency of reverse transcription, to accelerate the rate of genomic RNA dimerization, and to incorporate genomic RNA into virions for correct assembly and maturation. During viral maturation, NCp7 is cleaved from the gag protein by the retroviral protease before acting as a nucleic acid chaperone of the DIS dimer structural isomerization (31). NCp7 is a prime target for antiviral therapy due to its crucial role involved in many stages of viral replication cycle.

NCp7 domain is also responsible for annealing tRNA(Lys) with one specific amino acid residue, lysine, to the primer binding site (PBS) on the genomic RNA in reverse transcription. The interaction between tRNA(Lys) and the primer binding site during annealing has been measured directly by fluorescence energy transfer (FRET) by introducing fluorescence probes in the tRNA acceptor stem. Based on Förster theory, the

requirement of dipole-dipole interaction between donor and acceptor leads to a strong dependence of energy transfer on the distance between the two fluorophores. The efficiency of the energy transfer is then obtained from the peak fluorescence intensity of donor in the absence of the acceptor and donor in the presence of acceptor. The ends of tRNA spread out to 40 Å when tRNA anneals to PBS (32). Comparing the distance of donor and acceptor in tRNA before and after binding to PBS implies conformational change during annealing. Heteronuclear NMR spectroscopy has also been used to study the mechanism of tRNA binding to PBS by monitoring the step by step process of forming complexes in real time. For example, Heteronuclear Multiple Quantum Correlation (HMQC) spectroscopy has been used to compare shifting and/or disappearance of imino resonances before and after tRNA binds to PBS. It was then used to detect NCp7 interactions with the tRNA and the initiation of melting two base pairs, G6-U67 and T54-A58, which are the starting points in the unwinding of tRNA that is required for binding to the PBS (33).

The structure of NC contains two zinc finger motifs, Cys-X2-Cys-X4-His-X4-Cys, where X is variable amino acids. This zinc finger motifs of NCp7 are essential for packaging, annealing a primer tRNA to the viral RNA primer binding site, and maturation of viral RNA (Fig 1.7).

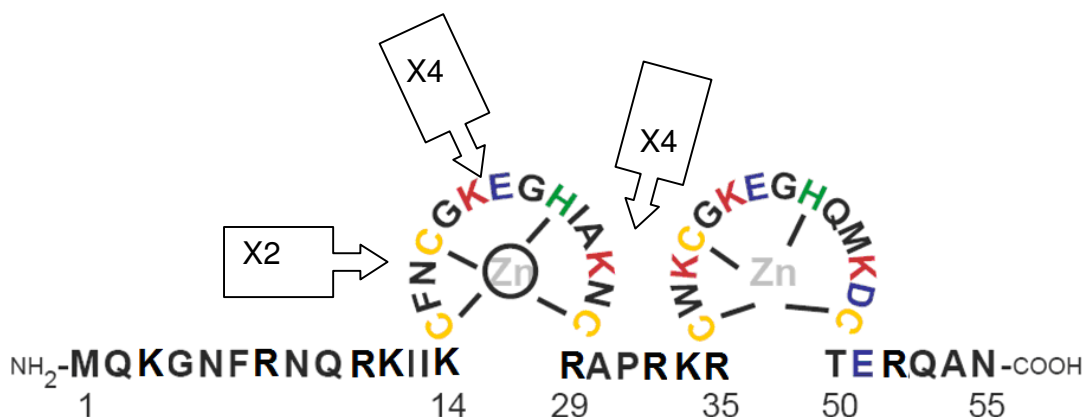


Fig 1.7. The sequence and secondary structure of NCp7 protein with the two zinc finger motifs Cys-X2-Cys-X4-His-X4-Cys indicated. The X2 and X4 represent the consensus spacing between the conserved Cys and His residues in the zinc finger motif (31).

While NC interacts with stem-loops (SL) in the  $\Psi$  site including DIS, SD,  $\psi$ , and gag initiation codon (AUG) during packaging of viral RNA, it is found to interact particularly tightly with the G-rich SL3 ( $\psi$ ). The dissociation constant of a 1:1 complex composed of a 20-nucleotide  $\psi$  containing sequence of SL3 and NCp7 is 100 nM. Furthermore, the binding site of the N-terminal domain of NC from Lys<sup>3</sup> to Arg<sup>10</sup> forms a 3<sub>10</sub> helix which binds in the major groove of the RNA stem and the C-terminal zinc finger binds to the loop region of the stem-loop (34). Mutations of zinc binding amino acids or alterations in the linker sequences in the CCHC motifs in both sites result in reduced viral infection, showing that the specific sequences of the zinc finger motifs are required. For example, mutation of His to Cys in the first zinc finger and replacement of Trp37 with Leu decreases infectious activity. The decreased infectivity of the Trp37 mutant suggests

that the indole group on tryptophan suitably orients NC and makes important nucleic acids contacts to NC that are spatially optimized (35).

In maturation, gag is cleaved by protease to produce NC, which then acts as a nucleic acid chaperone of the DIS dimer structural isomerization. Under physiological salt and pH conditions at 37 °C *in vitro*, the kissing-loop dimer is a kinetically stable complex. The rate of structural isomerization is increased at higher temperature and/or in the presence of NCp7 at 37 °C, allowing to overcome the activation energy barrier between these two isomers. This result can be shown by examining the effect of dimerization process as a function of the temperature with or without NCp7 (8).

#### **1.1.2.4 DIS Kissing Dimer – Extended Duplex Conversion**

The mechanism of kissing loop to duplex conversion has been studied by a combination of gel electrophoresis, fluorescence, UV melting, and NMR. Dimerization is initiated via loop-loop interaction of the 6-base pair palindromes within the kissing loop dimer (36). The ratio of DIS dimer to monomer has been found to be a function of ionic strength. The higher the ionic concentration, the more the equilibrium shifts to the dimer state, since higher salt concentrations shield the negative charges on phosphate backbones to relieve the electrostatic repulsion between two complementary strands in the DIS dimer. Typically, mono and multivalent cations are crucial to stabilize the RNA structure (37). The formation of the kissing dimer is metal ion dependent (17, 38) and the binding of NCp7 chaperone protein facilitates its conversion into the extended duplex dimer (8). Gel electrophoresis showed that the formation of the kissing – loop dimer depends on the RNA concentration and the  $Mg^{2+}$  concentration (17). The kissing-loop

dimer is relatively unstable in the absence of  $\text{Mg}^{2+}$  at submillimolar RNA concentration. This result reinforces that the kissing-loop dimer with  $\text{Mg}^{2+}$  is a kinetically trapped dimer. In contrast, the extended duplex dimer can be formed without  $\text{Mg}^{2+}$  because of higher stability from the large numbers of interstrand base pairs.

## **1.2 Role of Metal Ions in RNA Structure and Function**

### **1.2.1 Di- versus Mono-Valent Cation Interactions**

Metal ions are required for RNA folding and catalysis (39, 40). RNA molecules are negatively charged polyelectrolytes due to the negatively charged backbone phosphate. Metal cations act as counter cations to neutralize the multiple charges during the process of RNA folding. Metal ions are also required for stability and activity of RNA enzymes. For example, divalent metal cations in group I self-splicing introns play a structural role as a catalytic metal (41). The most commonly used ions in RNA folding are mono-valent and di-valent metal ions, such as  $\text{K}^+$  (42) and  $\text{Mg}^{2+}$  (43).  $\text{Na}^+$  is known to be important for stabilizing the tertiary structure of tRNA (44). The radius of group II (alkaline earth) ions is smaller than the group I (alkali metal) ions. Due to the smaller radius and higher positive charge, group II metal ions have the greater charge density and a stronger stabilizing effect on RNA tertiary structure compared to monovalent ions. In physiological conditions, multivalent ions (~2 mM) stabilize folding the HDV and *Tetrahymena* ribozymes more efficiently than monovalent ions (~200-500 mM) (45, 46). Transition metal ions, such as  $\text{Mn}^{2+}$  and  $\text{Cd}^{2+}$  can substitute for  $\text{Mg}^{2+}$ . Some of these

transition metals, like  $\text{Mn}^{2+}$ , are paramagnetic and can induce significant line-broadening of resonances in NMR experiments.

One example of a tertiary structure that is stabilized by chelated  $\text{K}^+$  at specific binding sites is the structure of a variant of a 58-nucleotide *E. coli* rRNA fragment (42). The stabilization ability of a monovalent cation depends on its unhydrated ionic radius.

$\text{Mg}^{2+}$  has much greater affinity for oxygen binding than for sulfur.  $\text{Mg}^{2+}$  is a hard metal ion which forms more stable complexes with ligands containing the hard donor oxygen. The coordination geometry is an octahedron with the common arrangement, hexa-aqua magnesium ( $\text{Mg}[\text{H}_2\text{O}]_6^{2+}$ ). Manganese ( $\text{Mn}^{2+}$ ) has very similar chemical characteristics and hexa-aqua coordination. It has been used as a  $\text{Mg}^{2+}$  analogue in spectroscopic studies due to its paramagnetic properties.  $\text{Mn}^{2+}$  is considered a soft metal as it has a greater affinity for nitrogen and tends to form inner-sphere coordination (47). Zinc ( $\text{Zn}^{2+}$ ) ions have approximately the same size as  $\text{Mg}^{2+}$  but different coordination behavior. Unlike  $\text{Mg}^{2+}$ ,  $\text{Zn}^{2+}$  tends to form 4-, 5-, and 6- coordinate complexes with equal stabilities. When the coordination number is 4 or 5, zinc binds nitrogen and sulfur, whereas with a coordination number of 6, oxygen is a preferred ligand (48).

From the thermodynamic point of view, three energetic factors are considered to discriminate among mono and divalent ions. The dehydration energy of divalent ions is about four times larger than monovalent ions with the same radius. The repulsion between a divalent chelated ion and other RNA-associated ions is more unfavorable. At the same position in an electrostatic field, the electrostatic free energy of a divalent ion gains two times more than the energy of a monovalent ion. The compensation of the more intense electrostatic field is required due to the large dehydration and ion repulsion

penalties for divalent ions (49). It explains that the most electronegative RNA ligands, oxygen atoms, from non-bridging phosphate oxygens are bound to  $\text{Mg}^{2+}$  more than  $\text{K}^+$  or  $\text{Na}^+$ .

### **1.2.2 Metal Binding Pockets versus Diffuse Cloud Interactions**

Several factors affecting the relative affinity of different metal ions for a specific ion binding pocket are ionic radius, hardness of ions, types of coordinating ligands, coordination geometry, and hydration number of ions (50). There are two distinct modes of binding: diffuse binding and site specific binding. In diffuse binding, hydrated monovalent and divalent metal ions bind to the RNA through long range electrostatic interactions (51). The positively charged layer provides charges to screen the repulsion between RNA backbone and phosphodiesteres. Outer-sphere and inner-sphere interactions are two different types of site binding modes that depend on the hydration state of the metals. Outer-sphere interaction is when metal ions bind to specific sites without direct contacts to RNA (52). Outer-sphere bound metal ions are strongly attracted to nearby electronegative RNA ligands where they share solvation shells. In inner-sphere binding, metal ions interact directly with RNA such that at least one metal-bound water is replaced by a ligand from RNA. Outer-sphere binding is considered to be nonspecific binding, whereas inner-sphere binding of partially dehydrated metal ions localize in specific binding pockets. The common binding sites on RNA are phosphoryl oxygens, the ribose 2'-OH, N7 of purines, and base keto groups, especially O6 of guanosine and O4 of uracil (39).



In outer-sphere binding, nucleobases coordinate to metal ion via a water-bridge. Softer Lewis acids, such as  $\text{Mn}^{2+}$ , tend to exhibit more inner-sphere interactions than  $\text{Mg}^{2+}$  (53). A  $\text{Mn}^{2+}$  ion directly coordinates with two phosphate oxygens and with nitrogen atoms of one nucleobases (54). Typically, the change of chemical shift upon metal cation binding to labeled nitrogen ( $^{15}\text{N}$ ) in nucleobase and the  $J$ -coupling ( $^1J_{\text{M}}^{(\text{n}+)-\text{N}}$ ) between the metal and coordination site on labeled nitrogen atoms can be used to detect formation of an inner-sphere coordination of the metal (55).

Three approaches have been mainly used to study ion-RNA interactions: direct structural methods (X-ray crystallography and NMR), replacing the metal ion with an analog which is more sensitive to a biophysical technique, and changing the local metal-binding characteristics by specific modifications of the RNA. While it can provide the highest resolution for describing metal coordination, one limitation of crystallography is that may provide only a limited picture of RNA associated ions. For example, in the haloarcula 50S ribosome subunit, only a limited number of mono- and divalent ions can be identified in the crystal structure, enough to neutralize only about 11% of the phosphate charges. In these cases, the crystallographic picture can't describe the importance of the 'invisible' metal ions and how these ions reduce the RNA electrostatic free energy compared to the resolved ions (56). Solution NMR technique can provide additional insight in these cases and allow identification of metal sites.

## **Chapter 2    2-AP Fluorescence Detection of RNA Structural Changes in Response to Metal Binding the Dimerization Initiation Stem-Loop Kissing Complex**

### **2.1    Abstract**

2-Aminopurine (2-AP) fluorescence detection methods have been used in this study to characterize metal binding in the dimerization initiation stem-loop kissing complexes. The fluorescence quantum yield of 2-AP is sensitive to its microenvironment and can be a generally non-perturbing nucleotide analogue when inserted at a specific site in nucleic acids. Therefore, 2-AP fluorescence detected methods can be ideal for studying the DIS isomerization chaperoned by NCp7, secondary structure transition of DIS stem-loop denaturation, and loop-loop dissociation. In this chapter, fluorescence methods will be described to detect the structural transitions and local dynamics associated with DIS dimer formation and structural conversion chaperoned by nucleocapsid protein.

### **2.2    Introduction**

#### **2.2.1    Fluorescence Emission / Quenching Experiments**

Fluorescence spectroscopy is a useful method to study conformational changes, binding interactions of ligands, and determination of stoichiometry for a complex. Fluorescence occurs when a molecule relaxes to the ground state after being excited. Fluorescence quenching is a non-radiative relaxation in which the excitation energy is

released as heat instead of light to a ground state. Fluorescence quenching is a process which decreases the fluorescence intensity. Interaction between a fluorophore and its corresponding quencher usually perturbs the fluorescence intensity. Fluorescence quenching results from the changes in its microenvironment. Several fluorophore and quencher combinations have been used to study long range interactions in biomacromolecules. Formation of a complex between two molecules can result in static quenching. However, direct physical interaction between a fluorophore and a ligand is not necessary to cause fluorescence quenching. Fluorescence resonance energy transfer (FRET) relies on the distance between a donor fluorophore and an acceptor fluorophore. A FRET transfer occurs when a donor fluorophore is excited by incident light and an acceptor is in close proximity to receive this nonradiative energy (57). The energy can be transferred over a distance range from 10 to 100 Å (58).

Fluorescent probe, 2-AP, which has been selectively incorporated at an adenosine position can be statically quenched through titration of a known molecular quencher, like acrylamide. Fluorescence quenching can take place when a non-fluorescent complex is formed at the ground state without emission of a photon from the  $\pi\pi^*$  excited state (59). The association constant for complex formation is given by,

$$K_s = \frac{[F - Q]}{[F][Q]} \quad (1)$$

Where [F] is the concentration of the fluorophore not a complex, [Q] is the concentration of the quencher, and [F-Q] is the concentration of the complex with no fluorescence.

The total concentration of the fluorophore is [F<sub>0</sub>].

$$[F_0] = [F] + [F - Q] \quad (2)$$

Then, the association constant can be rearranged as follows,

$$K_s = \frac{[F_0] - [F]}{[F][Q]} = \frac{[F_0]}{[F][Q]} - \frac{1}{[Q]} \quad (3)$$

The fluorophore concentration can be substituted for fluorescence intensities. Rearrange equation 3 yielding

$$\frac{I_0}{I} = 1 + K_s[Q] \quad (4)$$

### 2.2.2 Fluorescence versus UV Detected Thermal RNA Melting Curves

Melting transitions can be detected by UV absorbance, circular dichroism (CD), NMR, and differential scanning calorimetry. UV melting is by far the most commonly used technique to monitor the transition in solution. RNA undergoes the unfolding process as a function of temperature under quasi-stationary conditions from native to denaturing states. The temperature at the midpoint of the transition is called  $T_m$ . Upon unfolding of the RNA, UV melting shows hyperchromicity of unstacked bases with results in an increasing UV absorbance. Many UV melting curves can consist of more than one transition which results in a complex melting profile. In this case, the low temperature transition is usually attributed to the unfolding of the tertiary structure. In contrast, to unfold secondary structure helices, higher thermal energy is required to unzip stacked bases resulting in higher temperature transition. The ideal buffer used for melting curves should not absorb UV light and the pKa should not change with temperature. For example, the typical standard buffer for RNA samples is sodium cacodylate at pH 7. The advantage of using UV detection is its high sensitivity which means only small amount of sample is required. The enthalpy ( $\Delta H^\circ$ ) and entropy ( $\Delta S^\circ$ ) of the transition can be derived from UV-detected melting curves using the van't Hoff

equations. However, UV melting examines the global melting process as an ensemble average of the whole nucleic acid structure.

### **2.2.3 Sample Preparation & 2-AP Labeling**

2-Aminopurine (2-AP) is a fluorescent nucleotide base analog of adenine (60). The nucleotide 2-AP can form base pairs with cytosine in a wobble configuration (RNA and DNA) (61, 62), Watson-Crick type base pairs with either uracil (RNA) or thymidine (DNA) (63, 64). The comparison of the Watson-Crick adenosine-uracil base pair and the 2-aminopurine-uracil base pair are shown in Fig 2.1. Site-specific labeling with 2-aminopurine (2-AP) has been widely used to study the folding, structural dynamics, and molecular interactions of nucleic acids (38). The application of fluorescent probe 2-AP to monitor local melting process can overcome the ambiguity in the UV melting profile observed for complex nucleic acid structures which contain numbers of secondary structure motifs (65). 2-AP has a red shifted absorption spectrum which can be distinguished from the excitation of other nucleic acids and proteins. The changes of fluorescence intensities monitor the secondary structure transitions as a function of temperature including hairpin melting, loop opening, and base stacking, etc. The quantum yield of 2-AP is extremely sensitive to its microenvironment. 2-AP is highly quenched due to stacking interactions with neighboring nucleotide bases (66), but its fluorescence intensity can increase dramatically if the 2-AP base is flipped out of the RNA helix and exposed to solvent. In our laboratory, 2-AP has been incorporated into specific sites of the dimerization initiation site (DIS) dimers to serve as a fluorescence probe to examine the structural conversion of the DIS dimer from a kissing to an extended duplex mediated by NCp7 (67, 68). The fluorescence response of 2-AP to

local conformational changes in RNA has also been utilized to distinguish thermal melting transitions of the DIS dimer and to characterize local structural conformations of unpaired dynamic bases including stacked positions in the DIS loop.

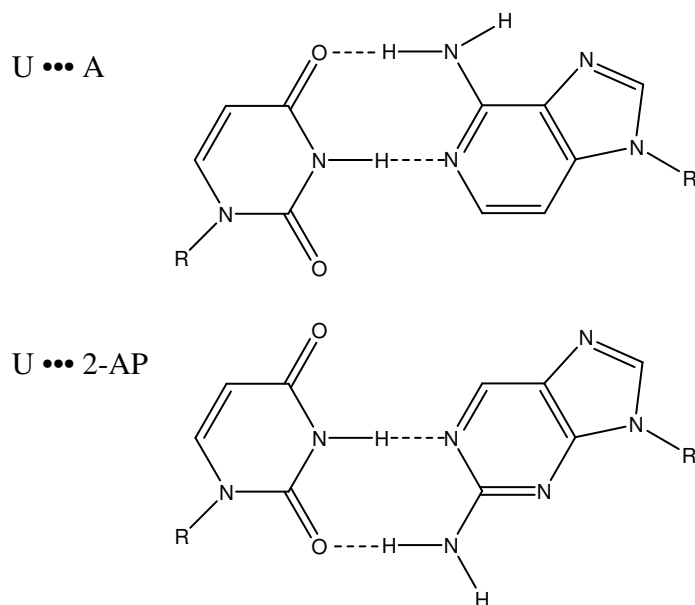


Fig 2.1. Comparison of the Watson-Crick adenosine-uracil base pair with the 2-aminopurine-uracil base pair.

## 2.3 Material and Methods

*Synthetic RNA Oligonucleotides.* The chemical synthesis of RNA is more difficult than DNA due to its reactive 2'OH group which introduces chemical interactions that can lead to RNA decomposition. In this regard, RNA nucleotides must be protected during chemical synthesis. The method to synthesize the RNA oligonucleotides is essentially described by Rist and Marino (38). 2-Aminopurine 2'-O-methylriboside (2-AP) incorporated at adenosine positions in RNA oligonucleotides are synthesized on solid phase via standard [(triisopropylsilyl)oxy]methyl (69)-protected phosphoramidite

chemistry to prevent partial formation of 2'-5'-linked phosphodiester bonds on a MerMade6 DNA/RNA Oligonucleotide Synthesizer (Bioautomation, Plano, TX (70)). The standard MerMade6 synthesis was followed by the procedures from Glen Research (Sterling, VA). De-blocking is carried out to remove DMT resulting in a free 5' hydroxyl group on the first base. Ethylthiotetrazole (ETT) was used as activator in a 3 minute coupling time (3 consecutive 1 minute cycles). Nucleoside phosphoramidites including 2-aminopurine riboside were purchased from Glen Research (Sterling, VA). Deprotection of the TOM group at the end of synthesis is achieved without affecting RNA integrity for RNA fluorescent labeling. The synthesis of unlabeled samples prepared by *in vitro* T7 polymerase run-off transcription and purification of both 2-AP fluorescent labeled and unlabeled samples are mentioned in chapters 1 and 3.

*Sample Preparation.* Oligonucleotide concentrations were derived from extinction coefficients at 260 nm of 231400 L / (mole·cm) for DIS24(GA)-9ap, 10ap, and 12ap and 208900 L / (mole·cm) for DIS23(HxUC) (OligoAnalyzer: IDT, Coralville, IA). RNA oligonucleotides with hairpin secondary structures were prepared by heating at 90 °C for 3 min followed by snap-cooling in ethanol with dry ice to bring them to - 70 °C for 5 min and then left at room temperature for 20 min before use.

### 2.3.1 Design of the DIS RNA Hairpins

The DIS kissing loop complex can rearrange to form a more thermodynamically stable duplex via base pairing between complementary nucleotides in each stem-loop. In addition, RNA concentrations and specific experimental conditions can result in different secondary structures and dimeric conformations. The idea of using heterodimer-forming DIS stem-loops was therefore introduced by our laboratory (38) to allow a finer control of dimer state to allow examination of the structural characteristics and biochemical behaviors of kissing loop and duplex individually. To inhibit the tendency of interconversion between kissing loop complex and extended duplex, two unique DIS stem-loops have been designed through point mutations in the palindromic sequence [U275 to A275 and A278 to G278 to form the DIS(GA) stem-loop and U275 to C275 and A278 to U278 to form the complementary DIS(HxUC) stem-loop with base sequence swapping on the DIS stem helix (Fig 2.2)]. This design favors the formation of a hetero-kissing complex by DIS23(HxUC) and DIS24(GA) as the predominant conformation in the solution. The higher stability of these heterocomplexes through complementary hexamer loop sequences ensures that stable formation of homodimers such as DIS (HxUC-HxUC) is less favored due to the mismatches in these loop-loop interactions (67, 71). The distinction between the DIS monomer and dimer states can be performed by native polyacrylamide gel electrophoresis (PAGE). However, standard gel electrophoresis is unable to easily distinguish between two DIS dimer isomers, kissing and extended duplex dimers, due to their similar conformations and same molecular weight (10, 11). Distinctions using PAGE can be made based on the fact that both kissing and extended duplex dimers are stably formed in the presence of  $Mg^{2+}$ , while the



kissing dimer formed in the absence of  $\text{Mg}^{2+}$  is not observed on gels. The design of 2-AP labeled DIS stem-loop provides an alternative and more robust method to distinguish these two isomers.

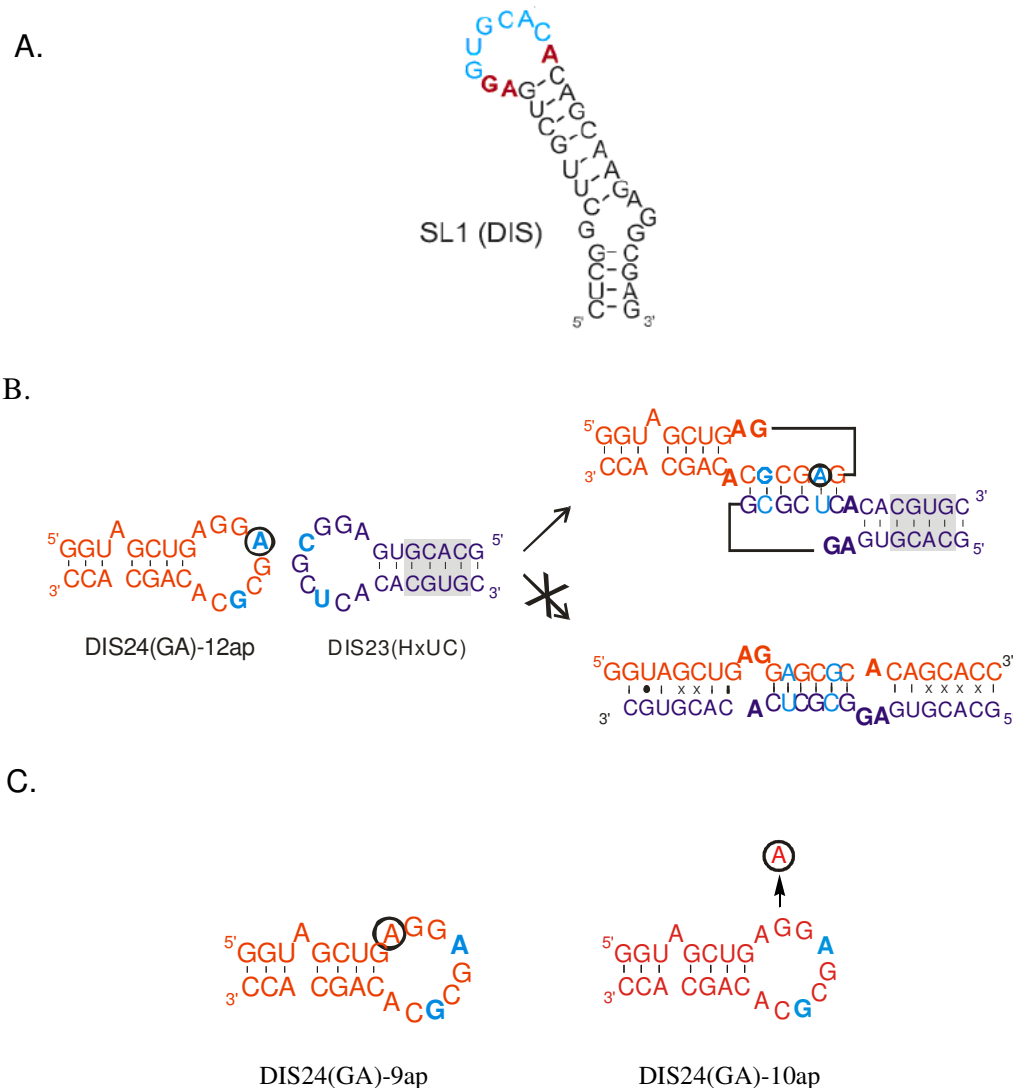


Fig 2.2. (A) RNA sequence and secondary structure of the wild type DIS (SL1) stem-loop from subtype-A of HIV-1 (38). (B) RNA sequence and secondary structure of the DIS stem-loop with point mutations in the hexanucleotide sequence [U12 to A12 and A15 to G15 to form the DIS(GA) stem-loop and U11 to C11 and A14 to U14 to form the complementary DIS(HxUC) stem-loop]. The sequence in the shaded grey box corresponds to the modified stem sequence that disfavors the formation of an extended duplex. Nucleotide positions in the DIS(GA) stem-loop replaced by the fluorescent probe 2-AP are circled in black. (C) RNA sequence and secondary structure of the DIS stem-loop at two specific fluorescent labeled positions, DIS24(GA)-9ap and DIS24(GA)-10ap. Nucleotide G10 was substituted with 2-AP circled in black.

### 2.3.2 Fluorescence Experiments

2-AP labels were inserted at three specific positions in the DIS24(GA) stem-loop. DIS24(GA)-12ap was substituted with 2-AP at the A12 loop position. DIS24(GA)-9ap was labeled with 2-AP at the A9 DIS junction and the nucleotide 10G was replaced by 2-AP on the DIS junction to make DIS24(GA)-10ap stem-loop. Three hetero-kissing complexes with 2-AP labels on different nucleotide positions coupled with the exchanged stem helix of DIS23(HxUC) were formed under different metal ion conditions. Based on the preliminary results, the DIS hetero- and homo- kissing dimers were both found to form kinetically trapped complexes in the presence of  $\text{Mg}^{2+}$  (~ 5 mM) without further conversion to the DIS duplex form. Based on the preliminary data, the empirically equivalent concentration of  $\text{Na}^+$  to  $\text{Mg}^{2+}$  is around 300 mM under standard buffer conditions where kissing complexes were created. However, mixtures of kissing and duplex homodimers were found in the presence of  $\text{Na}^+$  where the conversion reaction occurred spontaneously.

*Fluorescence-Monitored Melting* — The fluorescence melting curves of the AP-containing RNA oligonucleotides samples (100 nM) were performed using a Fluoromax - 2 spectrofluorometer (Jobin Yvon-Spex, a division of Instruments SA) with an excitation wavelength at 310 nm and emission wavelength from 330 nm to 460 nm in 1 nm increments. Melting curves generated by heating the RNA from 20 to 80 °C at the rate of 1 °C /min were measured for the change in 2-AP fluorescence intensity as a function of temperature. All thermal melts of the 2-AP labeled RNA samples were recorded in a 1 cm quartz cuvette. The RNA solutions in the cuvettes were allowed to reach thermal equilibrium in a programmable thermal bath at 20 °C for 15 minutes before initiation of a

temperature gradient occurred. The local melting temperature ( $T_m$ ) of each transition was derived from the derivative curve of the melting curve using KalediaGraph fitting.

*Fluorescence Spectra Measurements* — Fluorescence spectra of the RNA samples were measured on a SPEX Fluoromax-3 spectrofluorometer recorded between 330 and 460 nm. The excitation and emission wavelengths were 310 nm and 370 nm, respectively with an interval of 1 nm. Microcuvette (light-path: 3 mm × 3 mm) with a volume of 150  $\mu$ L was used in fluorescence measurements. Fluorescence experiments were performed at a constant temperature with a thermostated controlled water bath between 10 °C and 37 °C.

## **2.4 Results**

### **2.4.1 Fluorescence Melting Curve Analysis of the DIS Kissing Complexes — Distinguishing Different Helical and Junction Melting Events**

To specify the individual transition corresponding to a particular structural change within a complex RNA molecule, a fluorescence melting curve analysis with 2-AP labels is an ideal method. A 2-AP probe was inserted into the complementary loop on the hetero-kissing complex DIS24(GA)-12ap • DIS23(HxUC) as shown in Fig 2.2. Using this construct, the melting of the loop- loop helix in the DIS hetero-dimer was examined under varied metal ion and pH conditions.

*Effect of different metal ions*— The  $\text{Na}^+$  and  $\text{Mg}^{2+}$  stabilized DIS kissing dimers were used to examine different metal ion effects on the DIS dimer stability. The melts of DIS hairpin stem-loop and DIS kissing complex dissociation were monitored with mono- and di-valent metal ions in standard buffer using a construct with 2-AP at position A12 of DIS24(GA).

(1) DIS stem loop melting — Initially, the 2-AP labeled DIS(GA) stem-loop was measured as the change in 2-AP fluorescence intensity with increasing temperature from 20 to 80 °C in the presence of  $\text{Na}^+$  or  $\text{Mg}^{2+}$ . Next, melting profiles for a mixture of the DIS24(GA)-12ap and DIS23(HxUC) stem-loops without metal ions at pH 6.5 as well as the DIS24(GA)-12ap hairpin in the presence of  $\text{Mg}^{2+}$  or  $\text{Na}^+$  were acquired. These melting curves were broad and did not indicate a melting of a distinct secondary structure element as would be expected for the transition of a base in a loop sequence as the hairpin unfolds. The melting curve for DIS24(GA) hairpin stabilized by either  $\text{Na}^+$  or  $\text{Mg}^{2+}$  represents a single strand melt. Without metal ions, stable formation of kissing dimer is not observed. Instead, a linear like melting curve was observed (Fig 2.3) comprised of a mixture of two DIS hairpin loops corresponding to the melting transition of the loop as the hairpin unfolds.

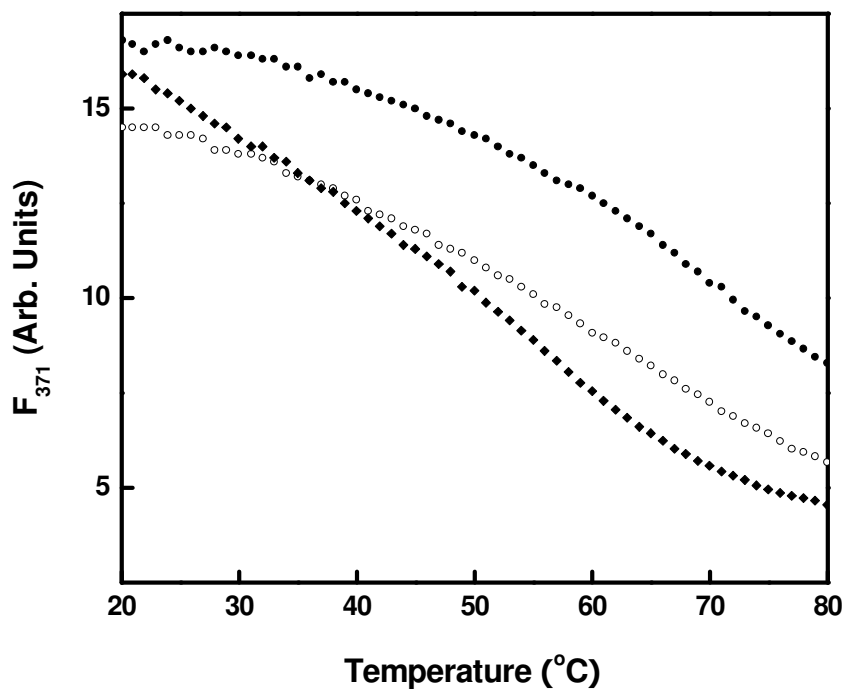


Fig 2.3. Plot of the fluorescence melting curves of 100 nM DIS24(GA)-12ap stem-loop (filled black circle) monitored by change in emission at 371 nm as a function of temperature in standard buffer (0.5 mM  $\text{Mg}^{2+}$ , pH 6.5); 100 nM DIS24(GA)-12ap stem-loop in the presence of 200 mM  $\text{Na}^+$  (open circle); and 100 nM DIS24(GA)12ap•DIS23(HxUC) kissing complex in the absence of metal ions (filled diamond).

(2) Kissing complex melting curves containing loop-loop denaturing and stem-loop melting — In the presence of  $\text{Mg}^{2+}$ , the DIS kissing dimer is found to be kinetically trapped. An approximately 3 fold decrease in 2-AP fluorescence emission occurs upon the formation of a kissing complex where the 2-AP substituted base is well stacked in the kissing loop-loop region before initiation of a temperature gradient from 20 °C (data not shown). As the temperature rises up to 71 °C, the

fluorescence intensity of the 2-AP probe of DIS24(GA)-12ap is found to increase around 2.5 fold followed by quenching of fluorescence (Fig 2.4). To distinguish the fluorescence changes resulting from the loop-loop thermal denaturation from the melting of the DIS stem-loop, the kissing dimer melting curve which is comprised of these two transitions needs to be normalized by taking the ratio of the intensity at each point in the melting curves from the melting curve of the individual DIS stem-loop (Fig 2.5). This sigmoidal melting curve represents the main melting transition from the kissing complex to the loop-loop dissociation. The approach to acquire the melting  $T_m$  is to take a first derivative of the sigmoidal melting curve which is the normalized fluorescence intensities. Two melting transitions ( $T_m$ ) are observed from the derivative plot at ~29 °C and 48 °C, respectively (Fig 2.6).

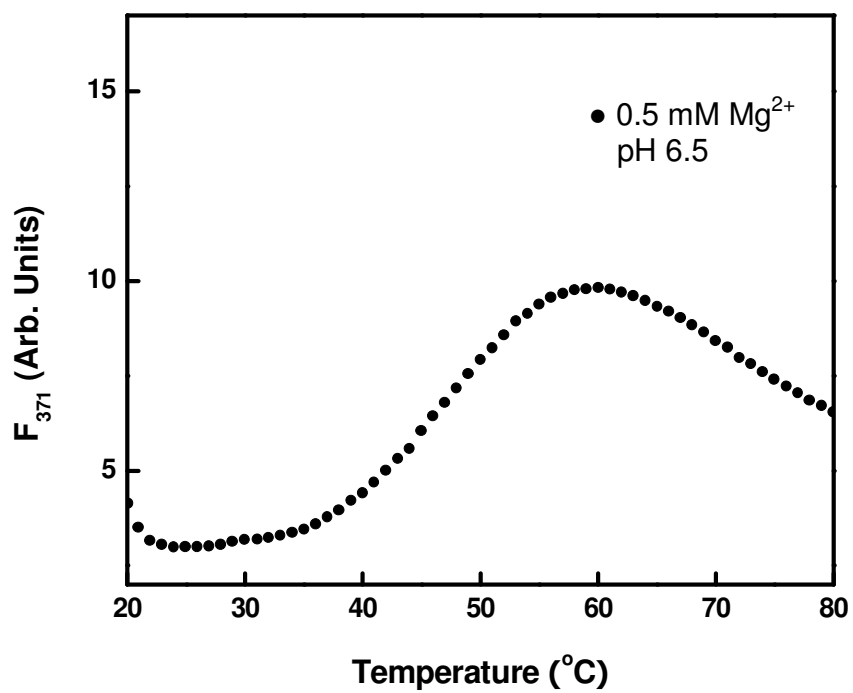


Fig 2.4. Fluorescence melting curve of 100 nM DIS24(GA)-12ap•DIS23(HxUC) kissing complex in the presence of 0.5 mM  $\text{Mg}^{2+}$  in standard buffer at pH 6.5.



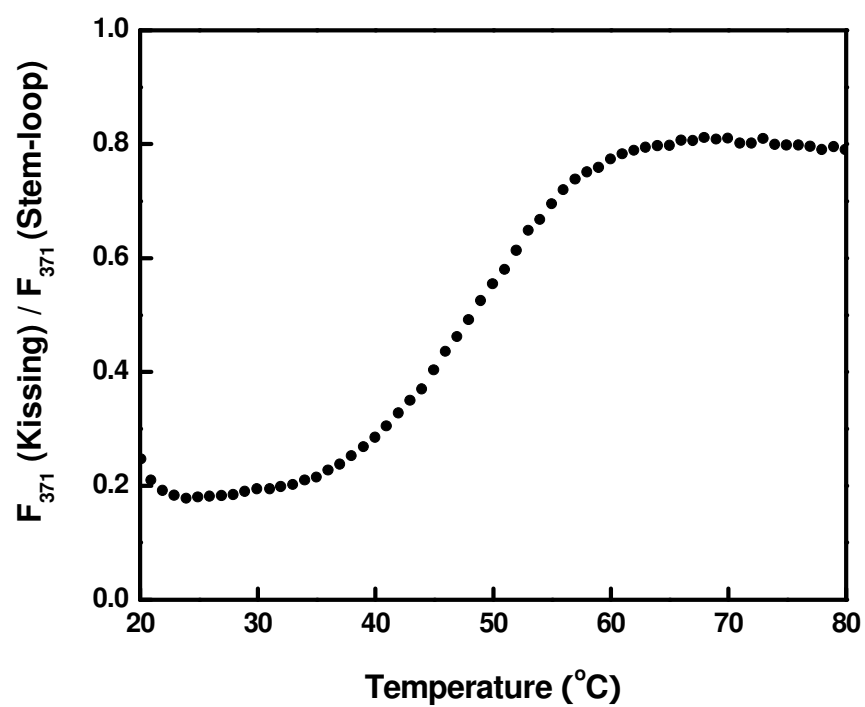


Fig 2.5. The sigmoidal curve pertaining to the loop-loop thermal denaturation process of 100 nM DIS24(GA)-12ap • DIS23(HxUC) kissing complex with 0.5 mM  $\text{Mg}^{2+}$  in standard buffer at pH 6.5.

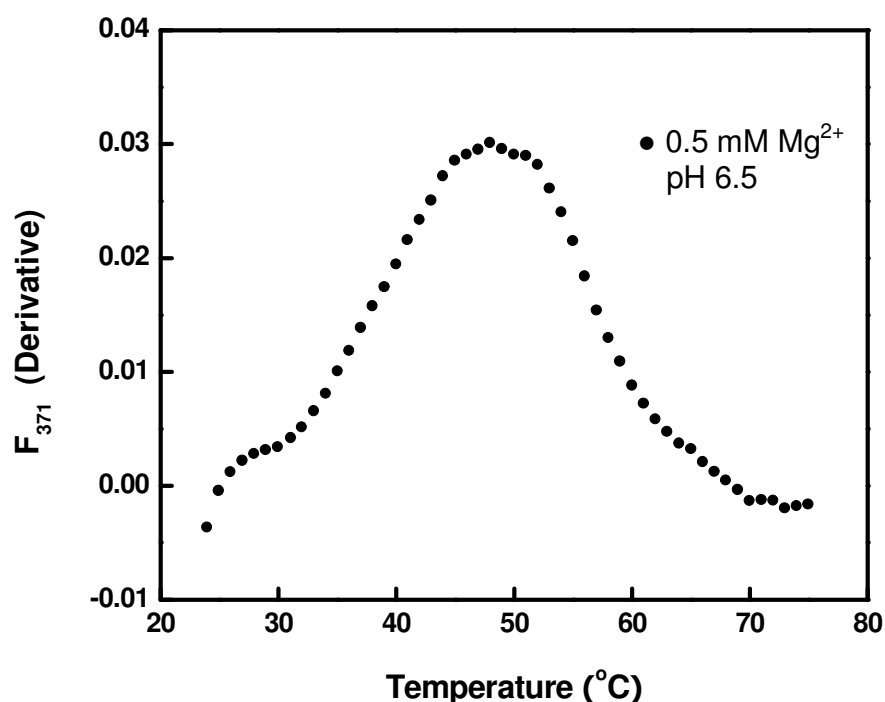


Fig 2.6. Normalized derivative of melting curve of 100 nM DIS24(GA)-12ap• DIS23(HxUC) kissing complex in the presence of 0.5 mM Mg<sup>2+</sup> in standard buffer at pH 6.5.

The DIS Stem-loops are also observed to associate as kissing dimers in the presence of monovalent cations with the mismatched stem-loop DIS23(HxUC) to form a hetero kissing complex (67). As shown in Fig 2.7, the melting curve in the presence of 200 mM Na<sup>+</sup> exhibits a different behavior in stabilizing DIS kissing complex. One lower temperature transition corresponding to the loop-loop dissociation at 27 °C is around the same range compared to the complex with 0.5 mM Mg<sup>2+</sup> (29 °C). In contrast, the main transition temperature referring to the melting of the stem-loop is shifted to ~ 33.5 °C indicating that the behaviors of DIS junctions of the hairpins are different with

mono- and di-valent cations (Fig 2.8). In addition, the initial fluorescence quenching from DIS24(GA)-12ap stem-loop to heterodimeric DIS24(GA)-12ap • DIS23(HxUC) with 0.5 mM  $\text{Mg}^{2+}$  has a more significant change than  $\text{Na}^+$  (~ 1.4 fold).

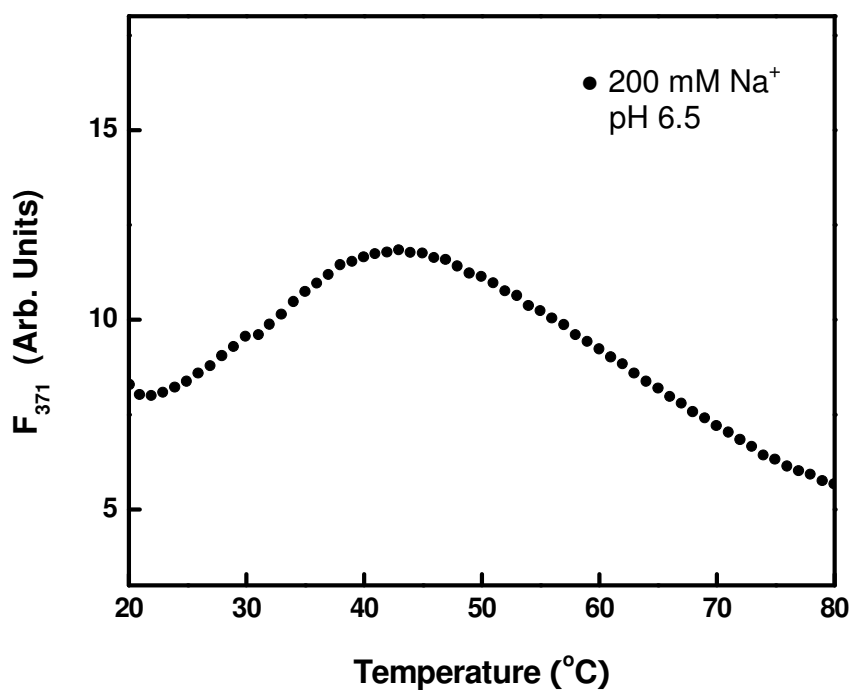


Fig 2.7. Fluorescence melting curve of 100 nM DIS24(GA)-12ap•DIS23(HxUC) kissing complex with 200 mM  $\text{Na}^+$  in standard buffer at pH 6.5.

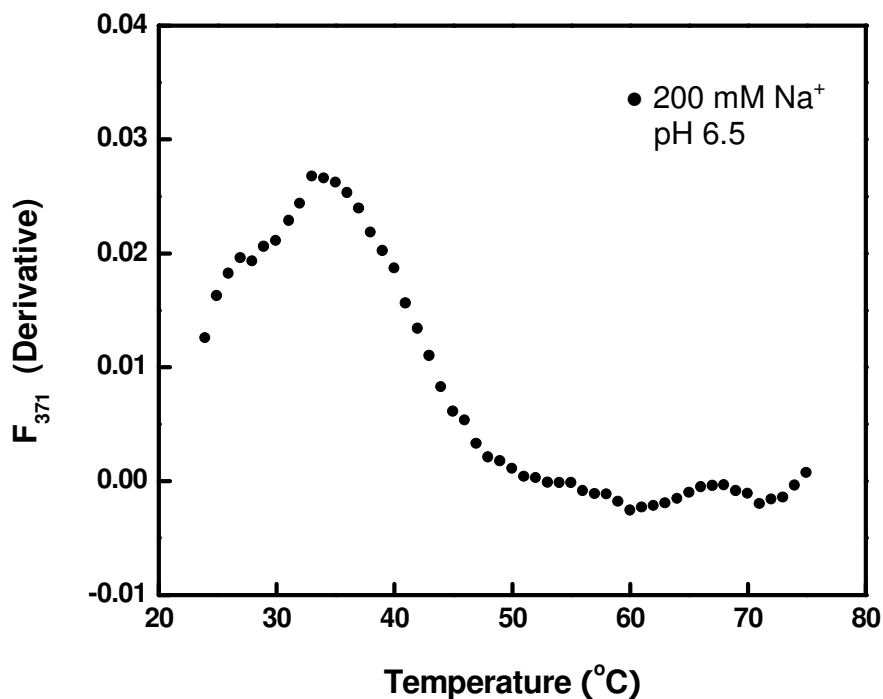


Fig 2.8. Normalized derivative of melting curve of 100 nM DIS24(GA)-12ap•DIS23(HxUC) kissing complex with 200 mM Na<sup>+</sup> in standard buffer at pH 6.5.

*Effect of metal ion concentration* — The fluorescence melting transition of the heterodimeric kissing complex is dependent on the metal ion concentration. As the concentration of Mg<sup>2+</sup> increases to 5.0 mM, a shift in the melting curve of DIS24(GA)-12ap•DIS23(HxUC) kissing dimer is found in Fig 2.9. The denaturation of the DIS stem-loop of kissing complex with 5.0 mM Mg<sup>2+</sup> is shifted to 65  $^{\circ}\text{C}$  (Fig 2.10). This would be expected based on the general stabilization of cations by Mg<sup>2+</sup> on RNA structures.

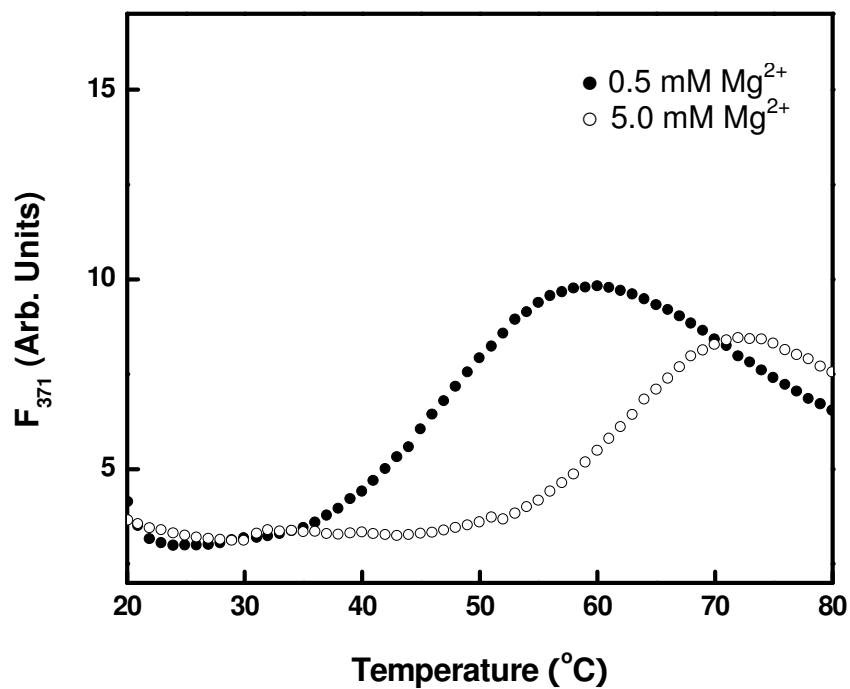


Fig 2.9. Plot of the fluorescence melting curves of 100 nM DIS24(GA)-12ap•DIS23(HxUC) kissing complex in the presence of 0.5 mM  $\text{Mg}^{2+}$  (filled circle) was monitored by change in emission at 371 nm as a function of temperature and the kissing complex with 5.0 mM  $\text{Mg}^{2+}$  (open circle) in standard buffer (pH 6.5).

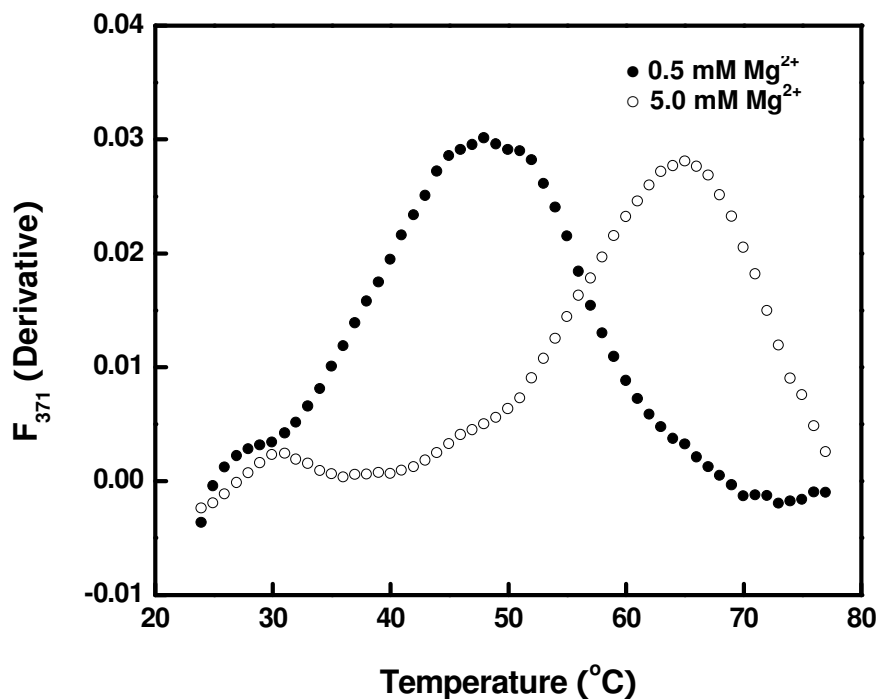


Fig 2.10. Normalized derivative of melting curve of 100 nM DIS24(GA)-12ap•DIS23(HxUC) kissing complex with 0.5 mM Mg<sup>2+</sup> (filled circle) and 5.0 mM Mg<sup>2+</sup> (open circle) in standard buffer at pH 6.5.

*Effect of pH* — A proton-coupled dynamic conformational switch has been identified in the DIS kissing complex at near-physiological pH which modulates the rate of conversion of the dimerization initiation site and the purine junction stability (71). Despite this observation, protonation/deprotonation on the DIS purine junction of kissing complex doesn't have a profound effect on the melting transitions of the stem-loop and loop-loop dissociation in the presence of either Mg<sup>2+</sup> or Na<sup>+</sup> (Fig 2.11, Fig 2.12).

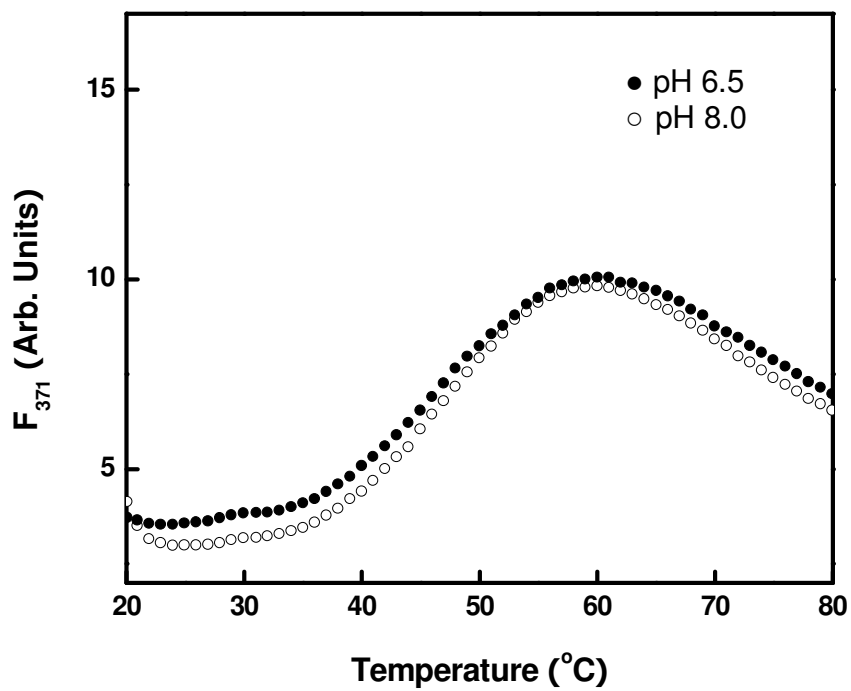


Fig 2.11. Fluorescence melting curve of 100 nM DIS24(GA)-12ap•DIS23(HxUC) kissing complex with 0.5 mM Mg<sup>2+</sup> in standard buffer at pH 6.5 (filled circle) and pH 8.0 (open circle).

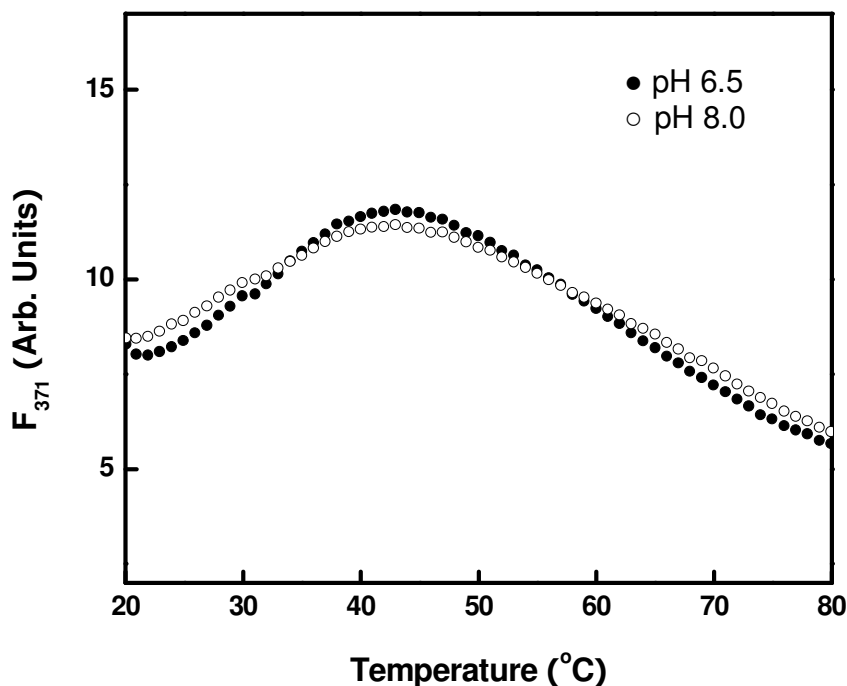


Fig 2.12. Fluorescence melting curve of 100 nM DIS24(GA)-12ap•DIS23(HxUC) kissing complex with 200 mM Na<sup>+</sup> in standard buffer at pH 6.5 (filled circle) and pH 8.0 (open circle).

#### 2.4.2 Measurement of the Site Specific 2-AP Responses to Metal Ion Interactions: Response of the Two Unpaired Purine Bases on the 5' Side of DIS Loop in Kissing Complex

2-AP is commonly used to study local base stacking and dynamics inserted within the loop, unpaired base, or inside base paired nucleotides. To probe changes in stacking behavior in response to metal ion perturbation and pH dependence, the two purine junction nucleotides at position 9 and 10 were substituted with 2-AP, respectively (Fig 2.13). It has been found that the rate of NCp7 catalyzed maturation of the DIS kissing



dimer is directly correlated with a proton-coupled conformation dynamics localized around these bases (71). Consequently, to probe the dynamics and stacking of these purine nucleotides becomes crucial to characterize their structural roles in stabilizing the kissing dimer and facilitating NCp7 conversion.

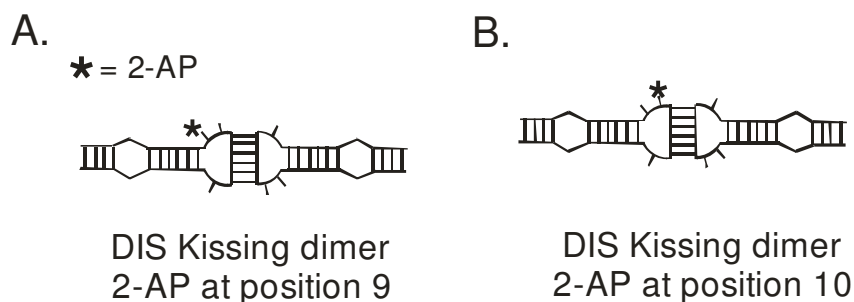


Fig 2.13. The secondary structure of the DIS kissing complexes with 2-AP labeling at 9 and 10 on DIS24(GA) side (72).

In these experiments, 2-AP fluorescent bases were found to become more unstacked in the kissing complexes which lead to the increase in fluorescence emission intensity compared to the conformation found in the stem-loop. The fluorescence intensity of the 2-AP inserted at the position 10 of the DIS24(GA) loop increases more than the position 9 right on the top of the DIS stem. A series of experiments have been obtained to examine the pH dependency and metal ion dependence on these two purine nucleotides on the DIS junction. In Fig 2.14 and Fig 2.15, a comparison of fluorescence measurements between at pH 6.5 and pH 8.0 in the presence of 0.5 mM  $\text{Mg}^{2+}$  has been performed. To investigate whether the specific protonation of the N1 base nitrogen of the DIS loop results in a fluorescence change that may be related to 2AP unstacking occurring in the formation of the kissing complex, a comparison of pH measurement in

the presence of 200 mM Na<sup>+</sup> at pH 6.5 and pH 8.0 is shown in Fig 2.16 and Fig 2.17. Overall, the magnitude of fluorescence increase obtained with Na<sup>+</sup> is smaller than that obtained with Mg<sup>2+</sup> in the kissing complex.

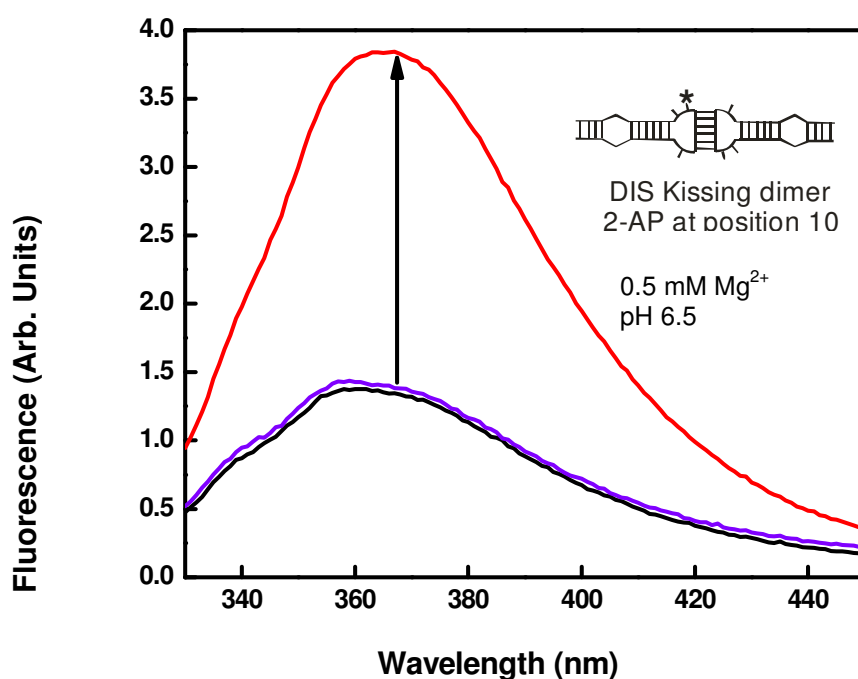


Fig 2.14. Plot of the fluorescence increase in the formation of kissing complex, 100 nM DIS24(GA)-10ap • DIS23(HxUC) with 0.5 mM Mg<sup>2+</sup> at pH 6.5 (25°C) in standard buffer condition. The order of addition each species in the solution is as follows: DIS24(GA)-10ap (black solid line), DIS23(HxUC) (blue solid line), and 0.5 mM Mg<sup>2+</sup> (red solid line). The black arrow indicates that the increase in fluorescence emission intensity upon the formation of the kissing complex.

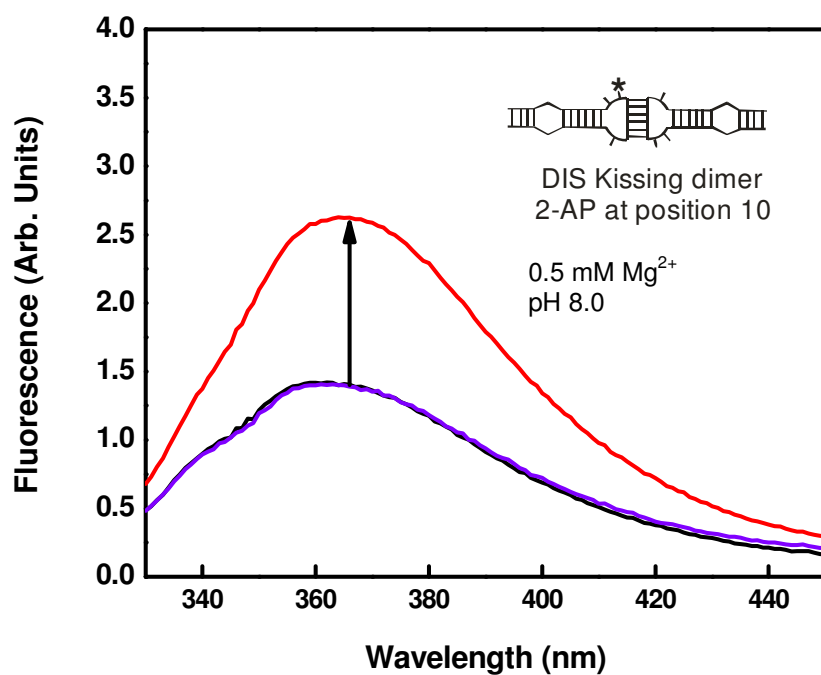


Fig 2.15. Fluorescence increase in the formation of kissing complex, 100 nM DIS24(GA)-10ap • DIS23(HxUC) with 0.5 mM  $\text{Mg}^{2+}$  at pH 8.0. The order of addition each species in the solution is as follows: DIS24(GA)-10ap (black solid line), DIS23(HxUC) (blue solid line), and 0.5 mM  $\text{Mg}^{2+}$  (red solid line). The increase in fluorescence emission intensity is labeled with the black arrow after formation of the DIS kissing complex.

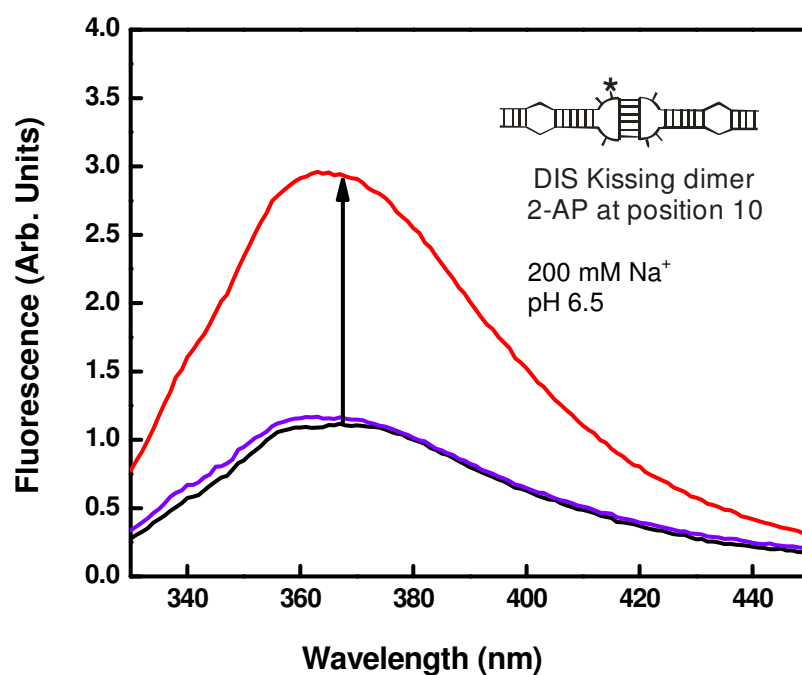


Fig 2.16. Fluorescence increase in the formation of kissing complex, 100 nM DIS24(GA)-10ap • DIS23(HxUC) with 200 mM Na<sup>+</sup> at pH 6.5. DIS24(GA)-10ap (black solid line), DIS23(HxUC) (blue solid line), and 200 mM Na<sup>+</sup> (red solid line). The increase in fluorescence emission intensity is labeled with the black arrow after formation of the DIS kissing complex.

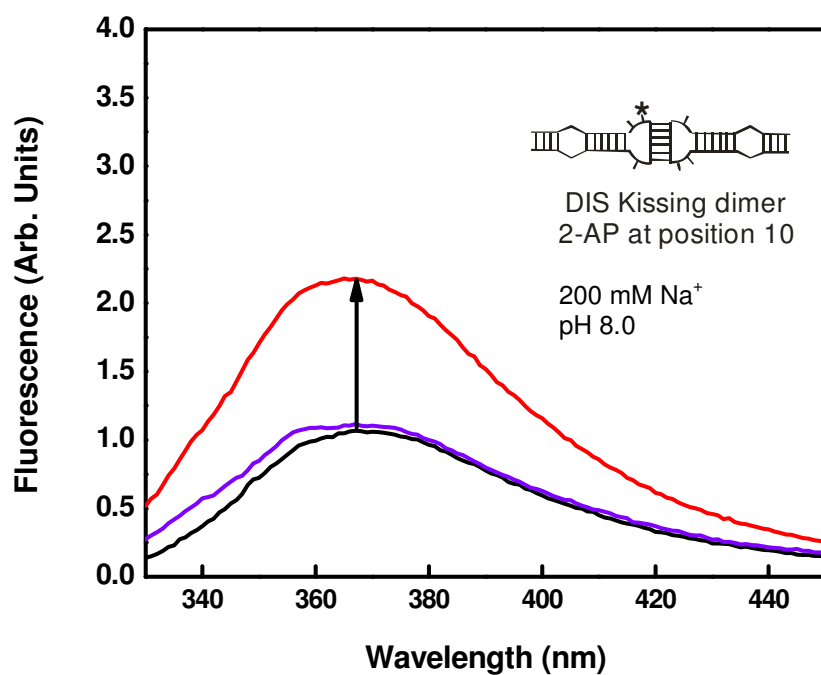


Fig 2.17. Fluorescence increase in the formation of kissing complex, 100 nM DIS24(GA)-10ap • DIS23(HxUC) with 200 mM Na<sup>+</sup> at pH 8.0. DIS24(GA)-10ap (black solid line), DIS23(HxUC) (blue solid line), and 200 mM Na<sup>+</sup> (red solid line). The increase in fluorescence emission intensity is labeled with the black arrow after formation of the DIS kissing complex.

The heterokissing complex with 2AP labeling at the A9 loop position DIS24(GA)-9ap basically follows a similar trend although with a smaller increase of fluorescence intensity when compared to the A10 loop position substituted with 2AP. The increase of fluorescence intensity in the presence of 200 mM Na<sup>+</sup> with pH 6.5 (Fig 2.18 (A)) and pH 8.0 (Fig 2.18 (B)) are presented.

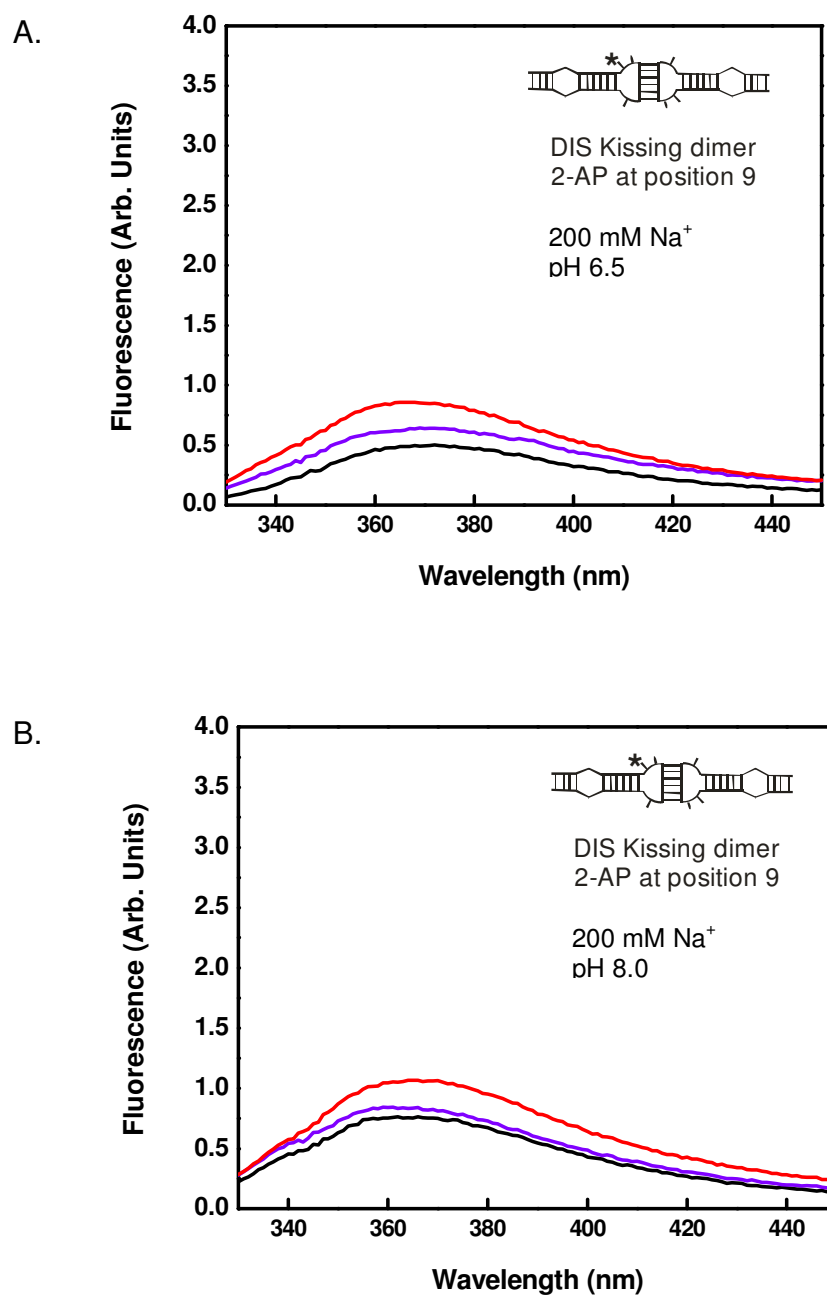


Fig 2.18. The increase of fluorescence intensity after the addition of 200 mM Na<sup>+</sup> to form the kissing complex, DIS24(GA)-9ap • DIS23(HxUC) at different pH. (A) pH 6.5 and (B) pH 8.0 DIS24(GA)-9ap (black solid line), DIS23(HxUC) (blue solid line), and 200 mM Na<sup>+</sup> (red solid line).

## 2.5 Discussion

Metal ion dependence, concentration dependence of metal ions, and pH effect have been studied in this fluorescence melting curve analysis with the fluorescent probe 2-AP inserted at different positions in the DIS loop. pH effect doesn't effect the stability of the kissing dimer as measured by loop-loop dissociation and stem-loop denaturation in the presence of mono- (Table 2.1 (d and e)) and di-valent metal ions (Table 2.1 (a and b)) since there is only a subtle change in each melting temperature. At higher magnesium concentration, the major melting transition corresponding to the DIS loop-loop dissociation is increased by ~17 °C (Table 2.1 (a and c)). As might be expected, the higher magnesium concentration stabilizes the kissing dimers in solutions. Based on two distinct derivatives of fluorescent melting curves (Fig 2.6, Fig 2.8),  $Mg^{2+}$  and  $Na^+$  appear to play different roles in stabilization the structures of kissing complexes. The higher melting temperature for loop-loop thermal denaturation implies that the extra stability of the RNA secondary structure in the presence of 0.5 mM  $Mg^{2+}$  compared to 200 mM  $Na^+$  (Table 2.1 (a and d)).



| Kissing complex                      | $T_m / ^\circ\text{C}$ (minor) | $T_m / ^\circ\text{C}$ (major) |
|--------------------------------------|--------------------------------|--------------------------------|
|                                      | Loop-Loop dissociation         |                                |
| (a) 0.5 mM $\text{Mg}^{2+}$ (pH 6.5) | 29                             | 48                             |
| (b) 0.5 mM $\text{Mg}^{2+}$ (pH 8.0) | 29                             | 47                             |
| (c) 5 mM $\text{Mg}^{2+}$ (pH 6.5)   | 30                             | 65                             |
| (d) 200 mM $\text{Na}^+$ (pH 6.5)    | 29                             | 33.5                           |
| (e) 200 mM $\text{Na}^+$ (pH 8.0)    | 27                             | 35                             |

Table 2.1. The melting temperatures of 100 nM DIS24(GA)-12ap • DIS23(HxUC) kissing complex corresponding to the local transitions monitored using fluorescence melting to summarize the effects of metal ions, metal ion concentration, and pH dependence.

Table 2.2 shows the different fluorescence responses to kissing dimer formation of 2-AP at the position 9 and 10 on the 5' end of the DIS loop in the presence of 0.5 mM  $\text{Mg}^{2+}$  and 200  $\text{Na}^+$  at pH 6.5 and 8.0. For both positions with 2-AP insertion, the increase observed in the fluorescence emission spectrum suggests that both bases become more unstacked in the complex when compared to the conformation found in the stem-loop, with AP10 showing a more profound difference when compared to AP9 which is the first base on the 5' side of DIS loop. In addition, it is consistently found that the fluorescence intensities increase more when the kissing complexes are stabilized by  $\text{Mg}^{2+}$  for both positions and different pH (Table 2.2). Noticeably, the purine junction which is the

position AP9 has been identified to be stacked into the helix in the NMR solution structures. However, the X-ray structures show that this purine is bulged-out. Our fluorescence observation supports a model wherein these bases on the DIS junction are likely in a conformationally dynamic exchange with a stacked-in conformation being dominant in the solution state.

| DIS24(GA)-Xap • DIS23(HxUC)<br>FL (fold, 371 nm, 298K) | pH 6.5 | pH 8.0 |
|--|--------|--------|
| ap10 + 0.5 mM Mg <sup>2+</sup>                         | 2.9    | 2.6    |
| ap10 + 200 mM Na <sup>+</sup>                          | 1.9    | 2.0    |
| ap9 + 0.5 mM Mg <sup>2+</sup>                          | 1.7    | 1.9    |
| ap9 + 200 mM Na <sup>+</sup>                           | 1.4    | 1.7    |

Table 2.2. Summary of fluorescence increase regarding to metal ion dependence and pH effect of the kissing complexes, DIS24(GA)-9ap • DIS23(HxUC) and DIS24(GA)-10ap • DIS23(HxUC).

## 2.6 Conclusion

2-AP fluorescence methods have been used to investigate local structural stability for the DIS kissing dimer. Fluorescence melting curves with 2-AP labeled nucleotide on the DIS loop show each local transition of DIS kissing dimer denaturation. Overall, the smaller change in intrinsic fluorescence upon the formation of kissing dimers with 2-AP labeled on unpaired junction base right on the top of the stem (AP9) suggests that it is more stacked in the helix compared to the second position in AP10. These results somewhat contradict the x-ray structures where both bases are observed to be extrahelical and unstacked. The fluorescence is however consistent with a poorly stacked, but conformational bulged in observed by solution NMR.  $\text{Mg}^{2+}$  has more profound effects on stabilizing both two positions in the kissing dimer compared to  $\text{Na}^+$ .

## Chapter 3 Resonance Assignment of the DIS Extended Duplex and Kissing Dimer Structures Using High-Resolution NMR

### 3.1 Abstract

The HIV-1 dimerization initiation site is primarily responsible for initiation of dimerization of the HIV-1 genomic RNA *via* a loop-loop kissing complex formation. To study this RNA-RNA interaction, we have used model RNA hairpin constructs that form homo and hetero dimeric DIS complexes and compared the structures of DIS extended duplex and kissing dimers as a function of mono- versus divalent metal ions. DIS sequence constructs capable of forming homodimeric kissing and extended duplex dimers as well as heterodimeric kissing complexes, have been constructed as models for the DIS stem loop complexes. All DIS complexes have the same sequence at and around the DIS junction where the largest response to metal ion binding was expected. As a first step in our studies, the resonance assignments of the extended duplex homodimer DIS21 and the kissing dimer DIS23(GA)•DIS23(HxUC) in the presence of  $Mg^{2+}$  are discussed in this chapter and compared with assignments for the  $Na^+$  stabilized RNA structures. The sequence specific resonance assignments were carried out to allow atomic level resolution probing of divalent metal binding sites in the DIS dimer structures by utilizing NMR methods, including chemical shift perturbation and paramagnetic relaxation enhancement.

## **3.2 Introduction**

### **3.2.1 Nuclear Magnetic Resonance Spectroscopy Applied to RNA**

#### **3.2.1.1 Sample Preparation & Isotope Labeling**

Enzymatic in vitro transcription using with T7 RNA polymerase and chemically synthesized DNA as a template is used to synthesize unlabeled and uniformly isotope labeled  $^{13}\text{C}$  and  $^{15}\text{N}$  oligonucleotides (73, 74). T7 RNA polymerase transcription is a standard method to produce a high yield of RNA oligonucleotides. Normally, a 5 ml scale transcription can produce milligram amounts of RNA. The polymerase requires G-rich sequences on the 5' end of the templates such as GGG or GGA to increase the transcription yield. The technique of isotope labeling has been utilized to resolve overlapped resonances in multidimensional spectra due to a narrow chemical shift range for the sugar proton resonances. For instance, isotope edited multidimensional experiments allow selective observation of signal in the intermolecular complexes of the labeled or unlabeled subdomain (75). In this reaction, additional non-templated nucleotides may be added (N+1 and N+2) beyond the desired sequence (N) and small abortive initiation products are also observed. This can be typically separated from the desired product by using preparative polyacrylamide gel electrophoresis (PAGE). Overall, the desired uniformly labeled RNA products can be synthesized by enzymatic transcription with high yield.

### 3.2.1.2 NMR Experiments for Resonance Assignment in RNA

Nuclear magnetic resonance (NMR) spectroscopy has been used as a premier analytical tool for analysis of RNA structures, dynamics, and RNA-ligand interactions. NMR is a physical phenomenon to measure nuclear magnetic moments. Those nuclei with an odd number of protons or odd number of neutrons have nuclear spin, such as  $^1\text{H}$  (99.98 %),  $^2\text{H}$  (0.01 %),  $^{13}\text{C}$  (1.11%),  $^{15}\text{N}$ (0.37%),  $^{19}\text{F}$  (~100%), and  $^{31}\text{P}$  (~100%). Ribonucleic acids (RNA) are composed of nucleotides which consist of a nucleobases, ribose sugars, and phosphates. Structure determination has been explicitly studied by a series of NMR experiments. The molecular mass limit of RNA NMR analysis is up to 100 kDa complex of packaging (76).

Initially, the base pairing pattern is identified by 1D proton spectroscopy to examine the guanines and uracil imino proton resonances, which are from exchangeable protons and are located in the spectral region from 10 to 15 ppm. The number of imino proton resonances corresponds to the number of base pairs in RNA complexes (Table 3.1 and Fig 3.1). Homonuclear 2D Nuclear Overhauser Enhancement Spectroscopy (NOESY) provides further information about base pairing from sequential assignment of imino proton resonances in water. The NOE is a through space correlation (i.e. dipole-dipole interaction). The NOE depends on the inverse of the distance to the sixth power between two spins and is therefore typically observed for protons that are within 6 Å of each other. The chemical shifts of the imino proton resonances of the RNA are predominantly influenced by base stacking, aromatic ring current effects, and chemical environment of the imino protons having hydrogen bonding to solvent molecules (77). A:U base pairs are readily identified with strong NOE cross peak intensities between the

H2 proton of adenine and the uracil H3 imino proton. In G:C base pairs, the H1 imino proton of guanine has a strong NOE interaction with amino protons on the base paired cytosine. Additional information from NOESY cross peaks contains guanine H1 imino protons and cytosine H5 protons due to spin diffusion. The cytosine H5 protons can be further used to identify the H5-H6 correlation on non-exchangeable spectra in helical RNA. The NOE magnetization is transferred between protons during mixing time in two dimensional experiments. In general, mixing times are around 200 to 400 msec depending on molecular weight of molecules for NMR structural assignments. If the NOEs are used for quantitative distance restraints for use in structural calculations, then mixing time can't be too long because of spin diffusion (78).

NOE based-assignment which is based on a structural model is not sufficient to fully assign the RNA resonances. Triple resonance, heteronuclear multi-dimensional experiments are essential to provide data for correlating different protons within individual nucleobases, correlating all sugar resonances, connecting ribose sugar and nucleobase resonances, and connecting sequential resonance in the phosphodiester backbone (79, 80). Assignments and correlations of resonance in the spectrum to particular nuclei can be used to identify residue specific base and sugar spin systems. The nucleobase spin system includes NH2, H2, H5, H6, and H8. The aromatic resonances including purine H8, pyrimidine H6, and adenosine H2 are located between 6.5 and 8.5 ppm. Pyrimidine H5 resonances generally appear between 5 and 6.3 ppm (Table 3.2). However, unstacked pyrimidine H5 resonances which are on the terminal or in bulge and loop of RNA tend to be shifted downfield of this range (81).

| Base pair | Base-pair atoms   | Chemical-shift regions [ppm] | Distance [Å] |
|-----------|-------------------|------------------------------|--------------|
| G:C       | G N1-H1 ••• N3 C  | 12-13.5                      | 1.89         |
|           | G N2-H2 ••• O2 C  | 8-9                          | 2.08         |
|           | G O2 ••• H4-N4 C  | 8-9                          | 1.71         |
|           | G N2-H2           | 6.5-7                        | -            |
|           | C N4-H4           | 6.5-7                        | -            |
| U:A       | U N3-H3 ••• N1 A  | 13-15                        | 1.93         |
|           | U O4•••H6-N6 A    | 7.5-8.5                      | 1.82         |
|           | A N6-H6           | 6.5-7                        | -            |
| G:U       | G N1-H1•••O2(O4)U | 10-12                        | 1.76         |
|           | G O2•••H3-N3 U    | 11-12                        | 1.96         |

Table 3.1. Chemical shift ( $\delta$ ) range for imino and amino resonances in Watson-Crick and wobble base pairs (82).

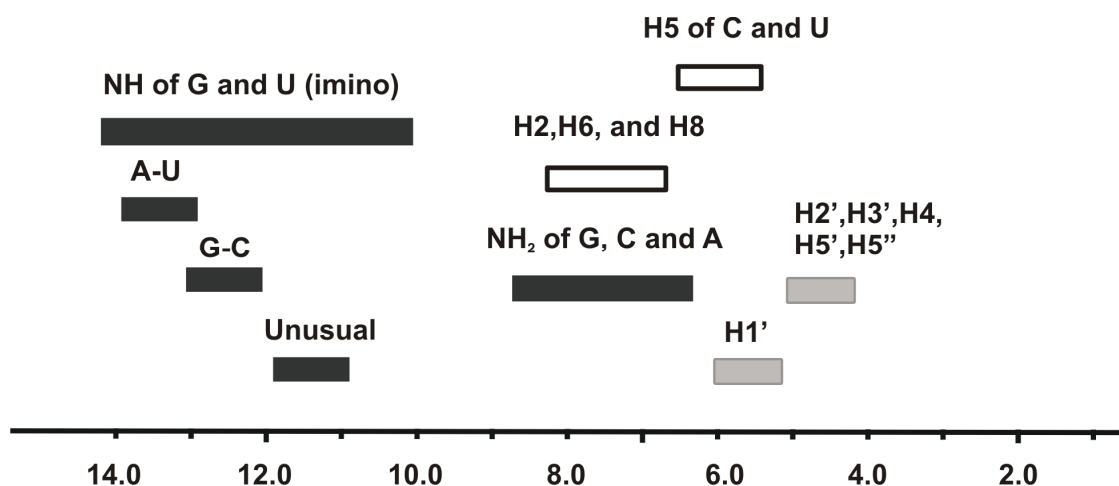


Fig 3.1. Typical  $^1\text{H}$  chemical shift ranges for exchangeable imino and amino protons and non-exchangeable base and ribose protons are indicated (83).



| Atoms  | $\delta^1\text{H}$ [ppm] | $\delta^{13}\text{C}$ [ppm] |
|--------|--------------------------|-----------------------------|
| C1'H1' | 4.4-6.5                  | 89-95                       |
| C2'H2' | 4-5                      | 70-80                       |
| C3'H3' | 3.8-5                    | 70-80                       |
| C4'H4' | 3.8-4.8                  | 81-86                       |
| C5'H5' | 2.5-5                    | 62-70                       |
| C2H2   | 6.5-8.5                  | 150-154                     |
| C5H5   | 5.0-6.3                  | 167-170                     |
| C6H6   | 7.0-7.7                  | 137-140                     |
| C8H8   | 7.0-8.0                  | 133-140                     |

Table 3.2. Chemical shift ( $\delta$ ) range for non-exchangeable proton and  $^{13}\text{C}$  resonances (82).

2D DQF-COSY (Double Quantum Filtered Correlation SpectroscopY), TOCSY (Total Correlation SpectroscopY), and NOESY experiments can detect non-exchangeable proton H5 and H6 correlation. Magnetization is transferred by scalar coupling (through bond coupling) in COSY experiments. The essential phenomenon detected using the DQF-COSY technique is double quantum filtered single quantum coherence. Being a phase sensitive experiment, DQF-COSY can adjust absorption line shape for both diagonal and cross peaks in order to obtain more refined spectra (84). In the TOCSY experiment, all protons of a coupled spin system are correlated. One important feature of TOCSY is the use of a spin lock, a continuous low-power pulse of constant phase applied in a period of time, to transfer magnetization between coupled proton pairs through bond correlation. As the size of a biological molecule increases, the line widths become larger as a result of the slower correlation rate. Partial cancellation of the cross peaks result in weak signals from anti-phase COSY spectrum. Unlike COSY,

TOCSY techniques provide a net transfer of in-phase magnetization with stronger cross peak intensities (85). The 2D CT-HSQC (Constant Time – Heteronuclear Single Quantum Coherence) experiment is used to identify adenine H2 resonances from  $^{13}\text{C}$ - $^1\text{H}$  correlation spectra. C2 signals resonate between 145 and 155 ppm and are separated from other carbon resonances. H2 resonances also show on NOESY spectra in  $\text{D}_2\text{O}$ . H2-H8 correlation can be achieved by HCCH-TOCSY spectroscopy (86). C6 and C8 resonances in HSQC spectra are found between 135 and 145 ppm and they both appear as singlets which are easy to distinguished two types of resonances (Table 3.2).

The protons of the ribose sugar are H1', H2', H3', H4', H5', and H5". The anomeric H1' proton appears at 6.25-5.25 ppm, where H5 resonances also appear. TOCSY or DQF-COSY can eliminate the ambiguity between H5 resonances and ribose protons. In a fully  $^{13}\text{C}$  labeled sample, the H1'-C1'-C2'-H2' spin systems of ribose sugars can be assigned using an HCCH-COSY experiment. 3D HCCH-COSY-TOCSY experiment is used to assign crowded ribose spin systems with uniformly labeled samples. Base to base protons can be assigned using a combination of HNCCH and HCCNH-TOCSY experiments. These experiments correlate non-exchangeable proton resonances with exchangeable imino and amino proton resonances. Base to ribose experiments such as H(C)N-type experiments correlate H1' to N1/N9-C8/C6-H8/H6 (87). Ribose to phosphate backbone H(C)P type experiments give rise to sequential assignment of adjacent nucleotides (88). Magnetization is transferred from 3' end correlating  $\text{H3}'_i\text{C3}'_i/\text{H4}'_i\text{C4}'_i$  with  $\text{P}_{i+1}$  to 5' end  $\text{H5}'_i\text{H5}''\text{C5}'_i/\text{H4}'_i\text{C4}'_i$  to  $\text{P}_i$  correlation. 3D HCP experiments can couple with HCCH-TOCSY to have further correlation of phosphorous and C1'H1' resonances (89). Combinations of isotope-filtered/ edited NOESY

experiments and nucleotide specific isotopically labeled RNA can be utilized to further identify and assign resonances (90).

### 3.2.1.3 NMR Structure Determination of RNA

RNA structure determinations benefit from the availability of methods for isotope labeling that allow heteronuclear experiments to be applied to mitigate cross peak overlap found in the proton spectra. After the sequence specific assignments of RNAs are obtained, NOESY experiments can be utilized to obtain proton-proton distance restraints for determination of the structure of RNA. Structure calculation based on NOE-derived distance restraints are generally defined based on the intensity of NOESY cross peaks, strong, medium, and weak to very weak. The distances are typically set from 3.0 Å for the most intense NOEs to 7.0 Å for the weakest NOEs in H<sub>2</sub>O experiments.

Five backbone torsion angles define the ribose sugar ( $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\epsilon$ , and  $\zeta$ ). Torsion angle restraints can be obtained from scalar coupling constants. Usually, ribose sugars in A-form RNA helices adopt the C3' endo pucker, with small coupling constants (~1 Hz) (91). The  $\alpha$  (O3'-P-O5'-C5') and  $\zeta$  (C3'-O3'-P-O5') dihedral angles can't be determined using J-coupling measurements, since <sup>16</sup>O nuclei which have no spin (1/2) isotope to allow the measurements. <sup>31</sup>P chemical shifts are the main source for structure calculation, but the correlations between <sup>31</sup>P chemical shifts and the phosphodiester backbones are not well determined due to the loose constraints on these torsions (92). The  $\beta$  (P-O5'-C5'-C4') and  $\epsilon$  (C4'-C3'-O3'-P) torsions can be constrained by measuring <sup>13</sup>C-<sup>31</sup>P and <sup>1</sup>H-<sup>31</sup>P coupling constants. <sup>3</sup>J<sub>H4'H5'</sub> and <sup>3</sup>J<sub>H4'H5''</sub> couplings are used to define  $\gamma$  (O5'-C5'-C4'-C3') torsion (93).

Hydrogen bond restraints are usually applied as a pair of distance restraints. With  $^{15}\text{N}$  labeled RNA samples, heteronuclear 2D HSQC and HNN-COSY experiments can distinguish base pairing and clarify scalar spin-spin couplings, respectively. Since uracil and guanine imino resonances are spaced apart by 10 ppm, it becomes easier to distinguish two imino resonances in 2D  $^{15}\text{N}$  HSQC experiments. Trans H-bond, scalar spin-spin couplings,  $^2J_{(\text{N},\text{N})}$  and  $^1J_{(\text{H},\text{N})}$  can be measured by HNN-COSY experiments (94). Base pairing can be confirmed and distance information obtained from these  $^hJ_{(\text{D},\text{A})}$  (D: hydrogen bond donor; A: hydrogen bond acceptor) couplings which are observed between the nitrogen nucleus and proton of the hydrogen bond donor and the nitrogen nucleus of the hydrogen acceptor (95). Scalar couplings across a second type of hydrogen bonds  $\text{N-H} \cdots \text{O} = \text{C}$  can also be measured (96), however, the magnitude of this type of scalar couplings is much smaller than  $\text{N-H} \cdots \text{N}$  type hydrogen bonds, making observation and precise measurement more challenging (Fig 1.9).

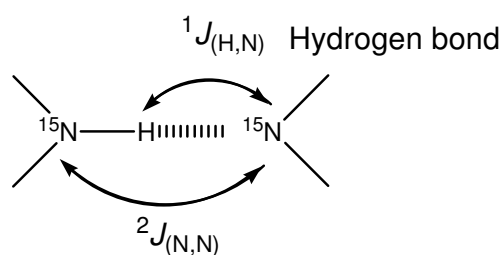


Figure 3.2. The typical representation of  $J$ -couplings through bonds is shown in this geometry.

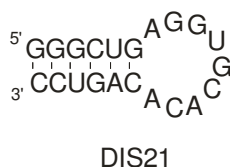
A limitation in nuclear overhauser effects, scalar coupling distance, and angular measurements is that these restraints are restricted to defining local geometry  $< 6 \text{ \AA}$  (97). In this respect, Residual Dipolar Coupling (RDC) measurements provide an ideal tool for determining global conformations such as the orientation of two helical axes of RNA stems. The specific mutations causing conformational change can be probed by RDCs (98). RDCs for RNA samples are typically measured using Pf1 phage to generate partial alignment, since principally the  $^1\text{H}$ - $^1\text{H}$ ,  $^1\text{H}$ - $^{13}\text{C}$ , and  $^1\text{H}$ - $^{15}\text{N}$  couplings are averaged to zero in isotropic solution (99). Projection restraints, relative to a global alignment tensor, can be determined from dipolar coupling data when the distance for a directly bonded pair of nuclei is known.

### **3.2.2 Analysis of the DIS Extended Duplex and Kissing Dimer Structures Using High- Resolution NMR**

The dimeric retroviral RNA genome is crucial in the HIV-1 life cycle. Dimerization is initiated *via* a loop-loop kissing interaction between the two monomers facilitated by the dimerization initiation site located near the 5' end of the viral RNA. The formation of a kissing loop-loop complex is a necessary intermediate prior to duplex formation (10). Maturation is primarily dependent on the structure and dynamics of the DIS kissing dimer and results in the conversion of DIS through stem strand exchange to form extended duplex dimers. To fully understand the structural differences between both the DIS duplex and kissing dimers in response to mono- and divalent metal ions, the homodimeric duplex dimer formed by a 21-mer DIS stem-loop, as well as a

heterodimeric kissing complex formed by DIS(GA) and DIS(HxUC) (Fig 3.4), have been designed for NMR studies. The isomerization of DIS21 homodimers involved in conversion of the DIS21 kissing loop dimer to extended duplex isoforms is shown in Fig 3.3 (B). Kinetic analysis of DIS dimerization and maturation from our lab showed that in the presence of monovalent ions ( $\text{Na}^+$  and  $\text{K}^+$ ) or  $\text{Mg}^{2+}$ , the structures of DIS kissing complexes are different and interact differently with NCp7. DIS kissing dimers are generated upon snap-cooling and kinetically trapped in the presence of  $\text{Mg}^{2+}$ . The heterokissing dimer, DIS23(GA)•DIS23(HxUC), is formed by RNA hairpins with point mutations in the autocomplementary loop sequence and non-complementary stems. The creation of the heterokissing dimer with altered sequences inhibits structural interconversion of the DIS isomers between duplex and kissing dimers. Both homo- and hetero- kissing dimers containing the same sequence on the DIS junctions are designed to study localized conformational dynamics upon metal ion binding to obtain a detailed understanding of structural rearrangement from kissing to duplex dimers. Moreover, protonation of the DIS purine junctions results in the dynamic conformation of DIS kissing complexes that promotes maturation. The different pKa values of the DIS purine junction determined from the C8 carbon chemical shift as a function of metal ion indicate that different structures occurred in the conformation of the junction purines are metal ion dependent (71). In this chapter, high-resolution NMR methods will be used to obtain sequence specific resonance assignment of the homo and hetero DIS dimers in the presence of  $\text{Na}^+$  and  $\text{Mg}^{2+}$  cations to characterize the functional role for the structural rearrangement of DIS dimers pertaining to metal binding.

A.



B.

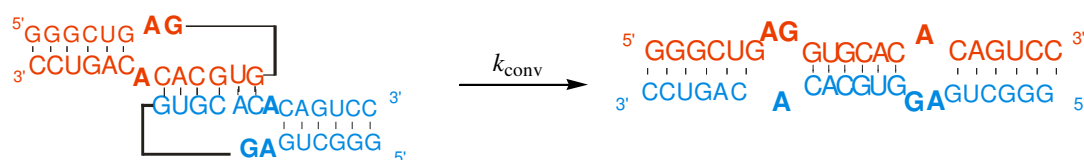


Fig 3.3. (A) Secondary structure of DIS21 stem-loop. (B) The structural conversion of the DIS (45) homodimer from a kissing-loop complex to extended duplex conformation. Two identical DIS 21 RNA strands are shown in red and blue for both conformations.

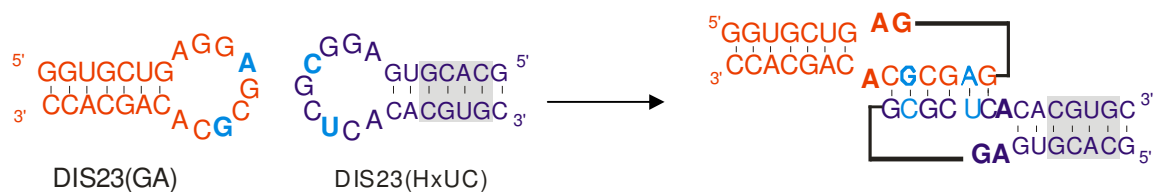


Fig 3.4. Schematic diagram of the formation of the heterodimer, DIS23(GA)•DIS23(HxUC), with the palindromic hexanucleotide DIS loop sequence of the Mal variant and exchanged stem sequence in DIS23(HxUC).

### 3.3 Material & Methods

*RNA Preparation* — RNA samples were prepared by *in vitro* transcription with T7 RNA polymerase using chemically synthesized DNA templates containing top strand and bottom strand. Unlabeled and isotope  $^{13}\text{C}$  and  $^{15}\text{N}$  labeled RNA samples can be made by enzymatic *in vitro* transcription with T7-RNA polymerase using chemically synthetic DNA as templates containing top strand and bottom strand (73). The template top strand contains the -17 to -1 of the promoter region and the bottom strand contains both promoter and templates. DNA templates were prepared by heating two top and bottom DNA strands together at 90 °C for 3 minutes and then cooling slowly for 20 minutes. The optimal concentration of  $\text{Mg}^{2+}$  needs to be determined before large scale synthesis of RNA samples. The concentrations of  $\text{Mg}^{2+}$  used in several small scale reactions were 20 mM, 25 mM, 30 mM, and 35 mM (40  $\mu\text{L}$ ). The standard transcription was performed in 40 mM Tris buffer [pH 8.3], 1 mM spermidine, 5 mM DTT (dithiothreitol), and 0.01 % (v/v) Triton X-100. Each NTP is added to a final concentration of 4 mM (3 mM  $^{13}\text{C}$  and  $^{15}\text{N}$  labeled NTP), DNA template to a final concentration of 0.8  $\mu\text{M}$  and T7 polymerase to a final concentration of 0.1 mg/mL, with water added to the final volume of the large scale reaction (5 mL) incubated overnight at 37 °C. All the stock solutions must be brought to room temperature prior to addition of T7 polymerase in order to avoid precipitation of the DNA template with spermidine and  $\text{MgCl}_2$ . Addition of EDTA at twice the concentration of  $\text{Mg}^{2+}$  was used to quench the reaction. The RNA sequence started with one or more guanine residues at the 5' end to increase the efficiency of T7 polymerase transcription. Preparative denaturing polyacrylamide gel electrophoresis (PAGE) containing 8 M urea is used to purify the desired RNA from abortive



transcription products and N+1, N+2 extension products. The RNA samples are recovered by excision of gel bands followed by electrophoretic elution. Microdialysis (Pierce Instruments, Rockford, IL) was the final step to provide further purification and exchange into standard buffer (1 mM cacodylate, pH 6.5, 25 mM NaCl) through extensive dialysis.

Commercial  $^{13}\text{C}^{15}\text{N}$ -rATP, -UTP, -CTP, and -GTP nucleotides (Sigma-Aldrich, St. Louis, MO) were used to synthesize a uniformly  $^{13}\text{C}^{15}\text{N}$ -labeled DIS23(GA) stem-loop. The DIS homodimer duplexes were formed either without  $\text{MgCl}_2$  or stabilized by 2 mM  $\text{MgCl}_2$ . For the DIS homo-dimeric kissing-loop dimer, the RNA solution was diluted to 10  $\mu\text{M}$  followed by snap-cooling and  $\text{MgCl}_2$  was added at 5 mM to kinetically trap the kissing dimer. Then the RNA sample was reconcentrated to 0.4 mM followed by microdialysis against NMR buffer with 100  $\mu\text{M}$   $\text{MgCl}_2$  at pH 6.5 for a final 250  $\mu\text{L}$  volume for the RNA sample Shigemi susceptibility matched NMR tubes (Shigemi, Allison Park, PA). The  $^{13}\text{C}^{15}\text{N}$ -DIS23(GA) and DIS23(HxUC) stem-loops were snap-cooled separately before mixing the two hairpins in a 1:1 ratio. The DIS kissing complex was stabilized by using 200-250 mM NaCl and 200  $\mu\text{M}$   $\text{MgCl}_2$ . The final concentration of the DIS kissing complex was around 0.3 mM.

*NMR Spectroscopy* — NMR spectra were collected on Bruker DRX 500 and DRX-600 NMR instruments at 298K. Spectrometers are equipped with a z-gradient triple-resonance cryoprobe or with a triple resonance 3-axis gradient probe. Samples were in  $\text{H}_2\text{O}$  (90%  $\text{H}_2\text{O}$ /10%  $\text{D}_2\text{O}$ ) or  $\text{D}_2\text{O}$  (99.96%  $\text{D}_2\text{O}$ , Cambridge Isotope Laboratories). 1D proton spectra were collected by using a jump-and-return pulse sequence or  $^{15}\text{N}$ -selected jump-and-return pulse sequence. Assignments of the imino proton resonances were

performed by  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC (100), 2D  $2J_{\text{NN}}$ -HNN-COSY, and 2D  $^{15}\text{N}/^{14}\text{N}$  edited / filtered NOESY spectra (71). Most spectra were recorded in States-TPPI mode using a gradient water flip-back pulse for solvent suppression in  $\text{D}_2\text{O}$  samples (101). Two dimensional NOESY spectra of  $\text{D}_2\text{O}$  samples were collected with sweep widths of 5389.9 Hz in the  $t_2$  and 5400.3 Hz in the  $t_1$  dimensions,  $4096t_2$  by  $400t_1$  complex points, and 250 ms mixing time. The 2D  $^1\text{H}^{13}\text{C}$  Constant-Time HSQC (CT-HSQC) experiments on the DIS21 homoduplex and  $^{13}\text{C}^{15}\text{N}$ -DIS23(GA)•DIS23(HxUC) heterodimeric kissing complex were recorded with 128 scans of 2048 points in the acquisition dimension and 64 increments in the indirect dimension. The constant time evolution period was set to 25 ms to have the same sign of the cross peaks (102). 2D  $^1\text{H}^{13}\text{C}$ -HCCH-COSY spectra were collected with  $^1\text{H}$  and  $^{13}\text{C}$  spectral widths of 6068 Hz and 5999 Hz, respectively, and  $1024 t_2$  by  $108 t_1$  complex points (103). Three dimensional HCCH-TOCSY experiment was collected with spectral widths of 10.1, 39.8, 8.0 ppm ( $f_3, f_2, f_1$ ) and  $1024t_3 \times 48t_2 \times 48t_1$ .  $^{13}\text{C}$ -edited /  $^{12}\text{C}$ -filtered 2D NOESY was recorded with 192 scans with a mixing time of 250 ms. Spectra were processed with nmrPipe (104) and analyzed with Sparky (105).

## **3.4 Results and Discussion**

### **3.4.1 Structural Assignments**

The strategy to assign specific resonances of particular interest involves the performance of homonuclear and heteronuclear NMR experiments. Exchangeable resonance assignments can be obtained from 2D NOE experiments in H<sub>2</sub>O and provide information about imino-imino correlations, imino-amino correlations, and imino-H5 correlations due to spin diffusion. Non-exchangeable resonance assignments relying on <sup>13</sup>C and <sup>15</sup>N based correlations which are through-bond can be recorded with labeled RNAs. Through-space correlations from the NOESY experiments can also be used to obtain aromatic and ribose resonance assignments in regions of canonical A-form structure where NOE patterns have been well established. The experiments corresponding to certain correlations and assignments are listed in Table 3.3.

| Experiment (unlabeled RNA samples) | Correlations                   | Assignments  |
|------------------------------------|--------------------------------|--|
| 2D NOESY (H <sub>2</sub> O)        | <sup>1</sup> H- <sup>1</sup> H | Imino N1H, N3H<br>Amino H41, H42<br>Pyrimidine H5<br>Purine H2 |
| 2D NOESY (D <sub>2</sub> O)        | <sup>1</sup> H- <sup>1</sup> H | Aromatic H2, H8<br>Pyrimidine H5, H6<br>Ribose H1'             |
| 2D DQ-COSY                         | <sup>1</sup> H- <sup>1</sup> H | Pyrimidine H5, H6  |
| 2D TOCSY                           | <sup>1</sup> H- <sup>1</sup> H | Pyrimidine H5, H6  |

| Experiment ( <sup>13</sup> C <sup>15</sup> N-labeled RNA samples) | Correlations   | Assignments  |
|---|--|--|
| <sup>13</sup> C-CT-HSQC (ribose)                                  | <sup>1</sup> H- <sup>13</sup> C                                  | Ribose <sup>1</sup> H, <sup>13</sup> C                         |
| <sup>13</sup> C-CT-HSQC (aromatic)                                | <sup>1</sup> H- <sup>13</sup> C                                  | Base <sup>1</sup> H, <sup>13</sup> C                           |
| 2D <u>H</u> <u>C</u> <u>C</u> <u>H</u> -COSY                      | <sup>1</sup> H- <sup>13</sup> C- <sup>13</sup> C- <sup>1</sup> H | Ribose <sup>1</sup> H, <sup>13</sup> C                         |
| 2D <u>H</u> <u>C</u> <u>C</u> <u>H</u> -COSY                      | <sup>1</sup> H- <sup>13</sup> C- <sup>13</sup> C- <sup>1</sup> H | Ribose <sup>1</sup> H, <sup>1</sup> H and<br>Pyrimidine H5, H6 |
| 3D <u>H</u> <u>C</u> <u>C</u> <u>H</u> -TOCSY                     | <sup>1</sup> H- <sup>13</sup> C- <sup>13</sup> C- <sup>1</sup> H | Ribose <sup>1</sup> H, <sup>13</sup> C, <sup>1</sup> H         |
| <sup>1</sup> H <sup>15</sup> N-HSQC                               | <sup>1</sup> H- <sup>15</sup> N                                  | Imino <sup>1</sup> H, <sup>15</sup> N                          |

Table 3.3. The assignments of homonuclear and heteronuclear correlations refer to corresponding experiments.

For homodimers and heterokissing complexes, all RNA samples are initially saturated by Na<sup>+</sup> as reference data and Mg<sup>2+</sup> is then added empirically to occupy the higher affinity divalent binding sites without causing overall line broadening.

### **DIS21 Extended Duplex Homodimer $\text{Mg}^{2+}$ Free State Versus $\text{Mg}^{2+}$ Bound State**

Three states of the DIS21 homo-complexes were analyzed: (1) The DIS21 extended duplex without  $\text{Mg}^{2+}$  (50 mM  $\text{Na}^+$ ), (2) the DIS21 extended duplex with  $\text{Mg}^{2+}$  at concentrations of 0.1 mM ( $^{13}\text{C}$  and  $^{15}\text{N}$  labeled sample) and 2.0 mM as well as 50 mM  $\text{Na}^+$ , and (3) DIS21 kissing dimer in the presence of 0.1 mM  $\text{Mg}^{2+}$  and 50 mM  $\text{Na}^+$ . Two identical DIS 21mer stem-loop strands (Fig 3.3 (A)) can form a mature homodimeric duplex (Fig 3.3 (B)) either with  $\text{Mg}^{2+}$  or without  $\text{Mg}^{2+}$  in the presence of  $\text{Na}^+$  at lower concentration ( $\sim 50$  mM). In 50 mM NaCl, the energetically favorable formation of the homoduplex occurred spontaneously and depends on the RNA concentrations. Because the DIS21 duplex is a symmetric homodimers, the number of observable resonances is cut in half due to the symmetry relationship which simplifies assignment problem.

Significant chemical shift perturbations are found in the DIS complexes as a result of  $\text{Mg}^{2+}$  binding (Fig 3.5). The imino proton signal for G6 disappeared upon  $\text{Mg}^{2+}$  addition, which suggests that the binding of  $\text{Mg}^{2+}$  to the DIS 21 duplex causes dynamics near the DIS junction (Fig 3.5 (B, C)). That is, the apical stem G-C base pair appears to be disrupted, or at least destabilized, in the presence of  $\text{Mg}^{2+}$ . The difference in chemical environment of the imino protons leads to more well separated resonances of U5 and U10 with 5 mM  $\text{Mg}^{2+}$ . In addition, in the presence of  $\text{Mg}^{2+}$  the imino resonance of G9 appears to be shifted and becomes to be overlapped with G11 (Fig 3.5 (C)). Overall, the changes of imino resonances indicate that the structure of the DIS21 duplex is significantly perturbed upon  $\text{Mg}^{2+}$  binding, and the perturbation is localized to the unpaired bases at the junctions.

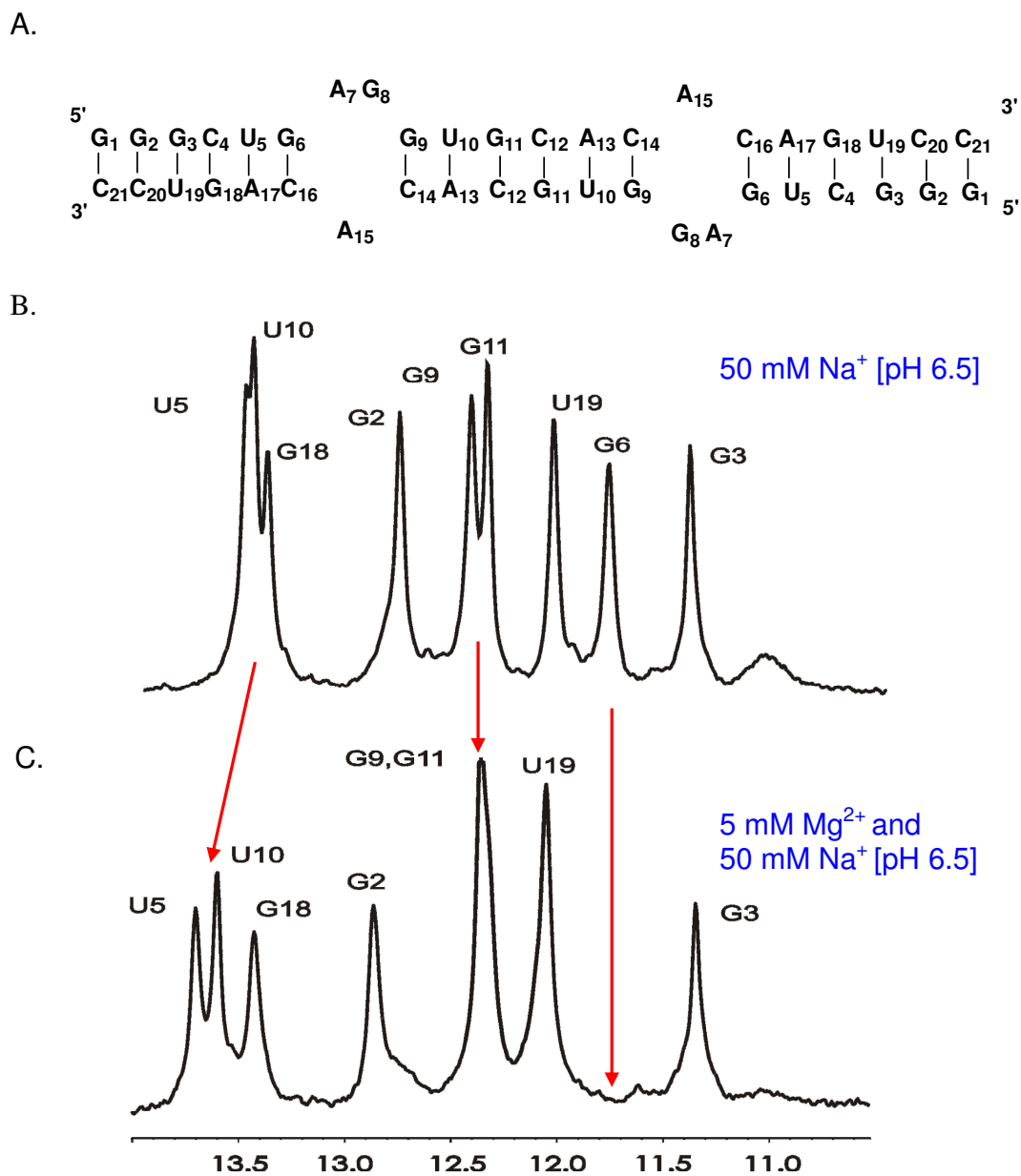


Fig 3.5. (A) The sequence of DIS21 homo duplex (B) 1D imino proton spectra of the DIS21 mature homodimer in the presence of 50 mM NaCl and (C) 5 mM MgCl<sub>2</sub> and 50 mM Na<sup>+</sup> (Mihailescu & Marino, unpublished data). The shifted resonances are indicated by arrows.

**Comparison of Resonance Assignments of DIS21 Extended Duplex Homodimer in the Presence of Na<sup>+</sup> versus Mg<sup>2+</sup>**

*Assignments of Exchangeable Protons* — The assignments of exchangeable protons were measured by 2D NOESY with a jump-and-return pulse sequence to obtain good water suppression. The imino proton resonances of guanine and uracils appear between 10-15 ppm. The 1D imino proton resonances match the diagonal proton peaks of the 2D NOESY spectrum. The noncanonical base pair, GU wobble pair, has the most upfield chemical shift at 11.2 ppm (Fig 3.6). The strong intensity of G3 and U19 cross peaks connecting to two diagonal peaks suggests that two imino protons are within 5 Å. G6 cross peak was not observed in the presence of Mg<sup>2+</sup> upon binding. Imino resonance of the terminal nucleotide, G1, was not detectable in 2D NOESY due to the fast exchange of the imino proton with the bulk water solvent (97).

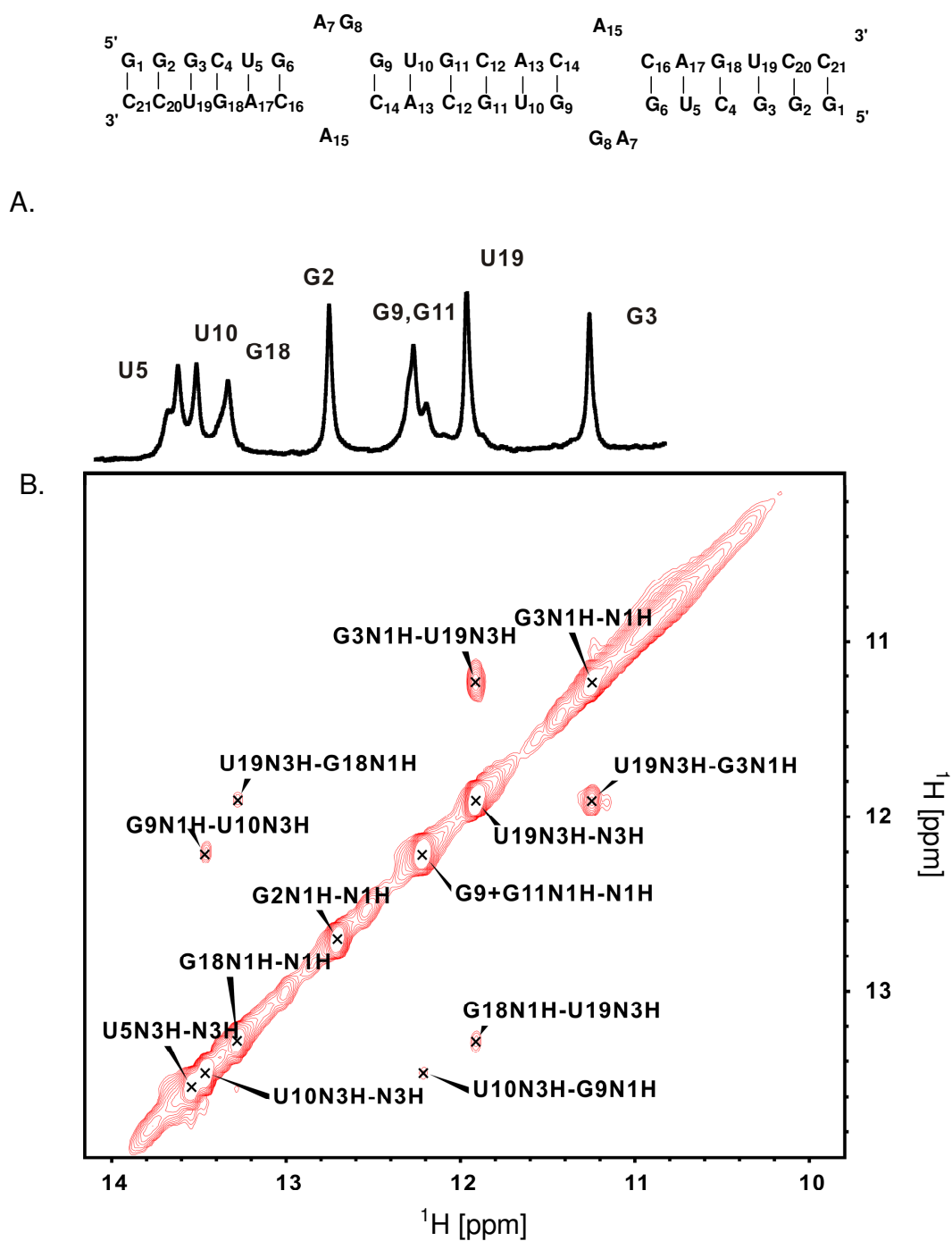


Fig 3.6. (A) The 1D imino proton and (B) 2D NOESY spectrum of the DIS21 homodimer duplex in NMR buffer with 2.0 mM  $\text{Mg}^{2+}$  and 50 mM  $\text{Na}^+$  (298K, pH 6.5).



*Assignments of Non-Exchangeable Protons* — The assignments of non-exchangeable protons and carbon resonances of the DIS21 duplex were derived from 2D DQ-COSY, 2D TOCSY,  $^{13}\text{C}$ -CT (constant time) HSQC —aromatic and ribose regions, HCCH-COSY — ribose proton resonances, 2D NOE connectivity walks, 3D NOESY, and 3D HCCH-TOCSY with  $^{13}\text{C}^{15}\text{N}$ -DIS21 duplex at 298K and pH 6.5 in NMR buffer.

First, individual base hydrogen spin systems (pyrimidine H5/H6) were identified by 2D DQ-COSY and 2D-TOCSY experiments. Presence of 50 mM  $\text{Na}^+$ , nine H5-H6 correlation cross peaks were observed in the 2D DQ-COSY spectrum in the fingerprint region, as would be expected for a C-2 symmetric homodimer (Fig 3.7 (A)). Upon  $\text{Mg}^{2+}$  addition, the anti-phase peaks were found to have weaker cross peak intensities due to the cancellation effects from anti-phase peaks in COSY experiment (data not shown). In this regard, 2D TOCSY experiment was performed to obtain correlations through an in-phase magnetization and avoid anti-phase signal cancellation associated with COSY(85). The binding of  $\text{Mg}^{2+}$  to the DIS21 duplex causes dynamics at the G6-C16 junction (Fig 3.7 (B)) as evidenced by the fact that the intensity of the C16 cross peak was weaker compared to the  $\text{Na}^+$  (50 mM) bound state of DIS21.



The aromatic, ribose hydrogen, and carbon spin systems were identified by the  $^{13}\text{C}$ -constant-time HSQC (CT-HSQC) experiments, which give better resolution by removing heteronuclear C-C coupling along the  $F_1$  indirect domain (102, 106). Moreover, in the experiment with 0.1 mM  $\text{Mg}^{2+}$  addition, it was found to produce the best spectra allowing occupancy of the higher affinity sites without overall broadening. Adenine H2 (AH2) protons of A•U pairs were obtained using these experiments to further confirm NOE assignments (Fig 3.8 (B)). The aromatic H6-C6 and H8-C8 resonances fall between 132 – 142 ppm. In the presence of 0.1 mM  $\text{Mg}^{2+}$  and 50 mM  $\text{Na}^+$ , the H6-C6 and H8-C8 correlations of G6, A7, G8, G9, and C16 bases are weak or absent in  $^1\text{H}$ ,  $^{13}\text{C}$ -HSQC spectrum. Three cross peaks were unassignable due to broadening effect (Fig 3.8 (A)).

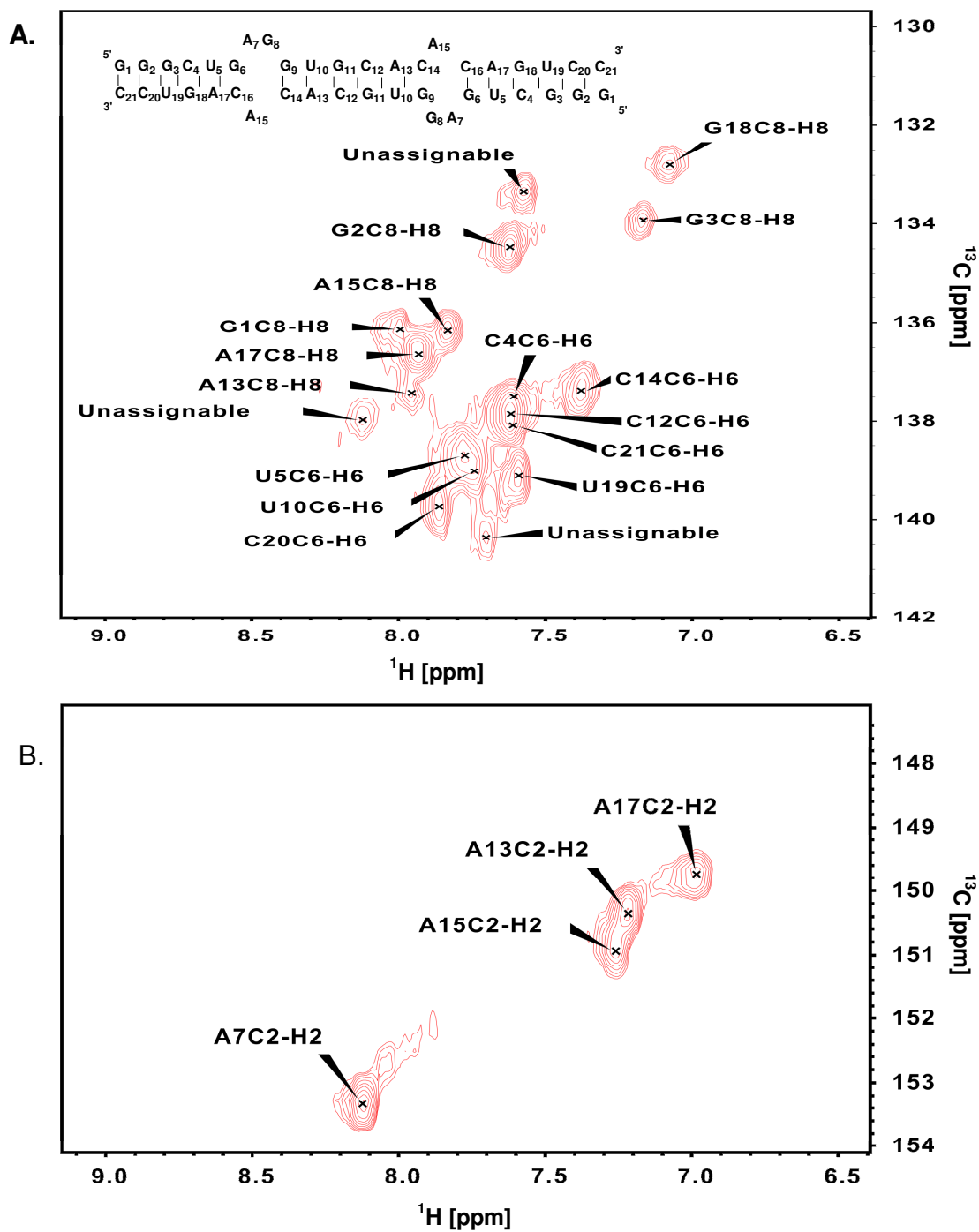


Fig 3.8.  $^1\text{H}^{13}\text{C}$ -HSQC spectrum showing the aromatic (A) H8-C8 and H6-C6 (B) H2-C2 correlation of DIS21 duplex in the presence of 0.1 mM  $\text{Mg}^{2+}$  and 50 mM  $\text{Na}^+$  (pH 6.5, 298K).

The CH groups of individual ribose spin systems, C1'-H1', C1'-H2', C2'-H1', and C2'-H2', were linked using 2D  $\text{HCCH-COSY}$  experiments on the  $^{13}\text{C}$ -labeled DIS21 duplex with 0.1 mM  $\text{Mg}^{2+}$  and 50 mM  $\text{Na}^+$  (Fig 3.9). A7, G8, and G9 purine base resonances were overlapped with other resonances, which presented assignment. The same pattern of the ribose H1' to C1' correlated region of the DIS21 duplex was acquired by both  $^1\text{H}^{13}\text{C}$   $\text{HCCH-COSY}$  and  $^1\text{H}^{13}\text{C}$  CT-HSQC (data not shown). Ribose protons were correlated *via* the 2D  $\text{HCCH-COSY}$  experiment to obtain H1'-H2' correlations (Fig 3.10).

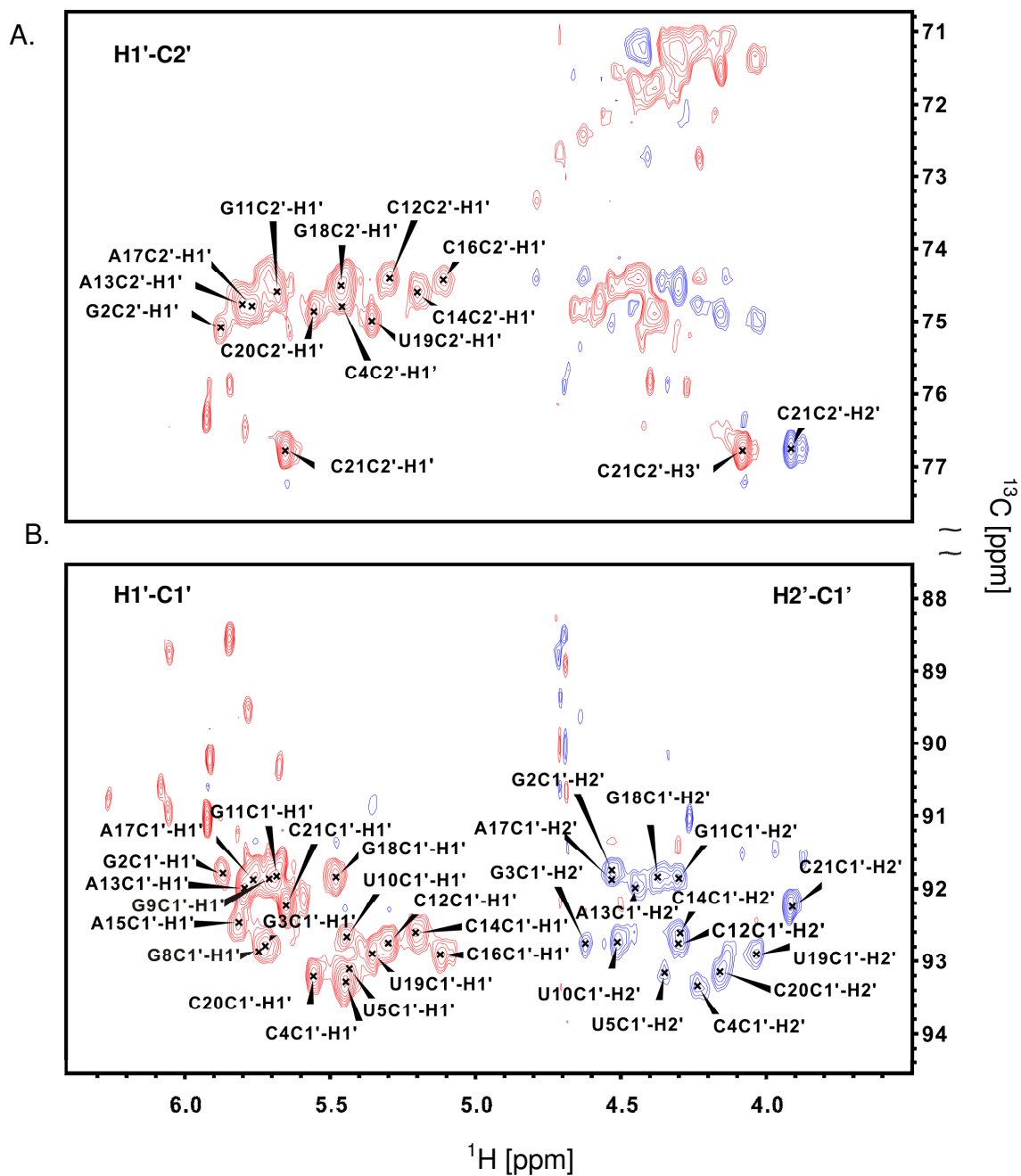


Fig 3.9. The H1'-C1', H2'-C1', and H1'-C2' spin systems of the sugar specific 2D  $\text{HCCH-COSY}$  spectrum of the  $^{13}\text{C}$ -DIS21 duplex with 0.1 mM  $\text{Mg}^{2+}$  and 50 mM  $\text{Na}^+$  at 298K and pH 6.5.

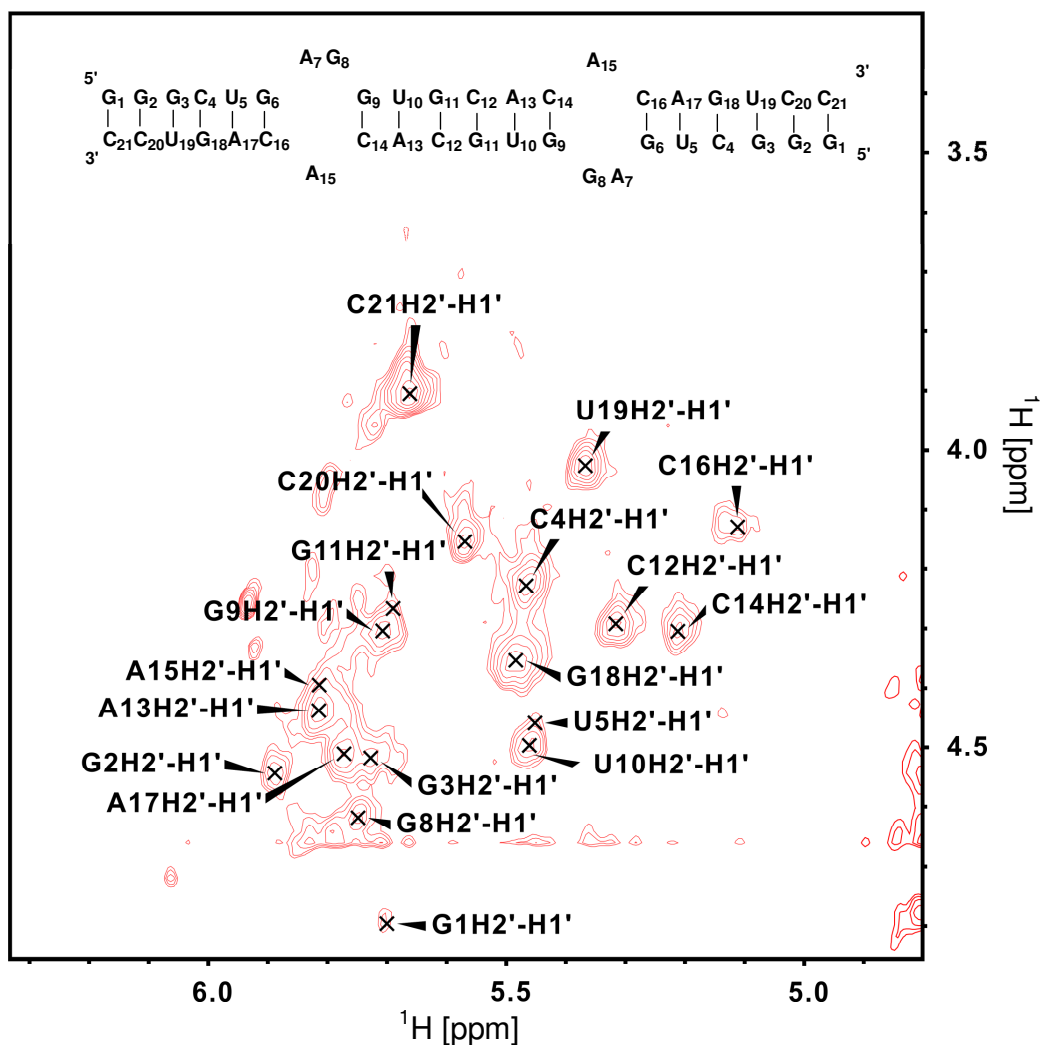
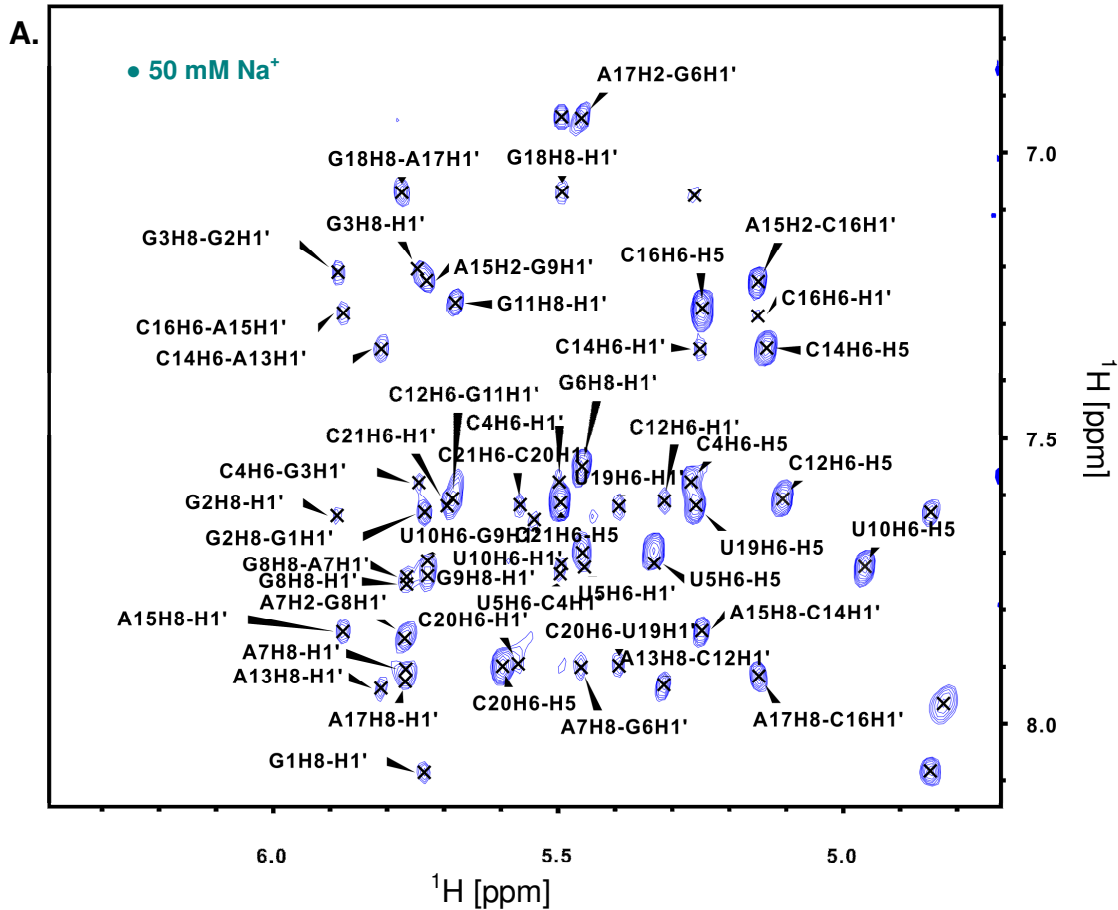
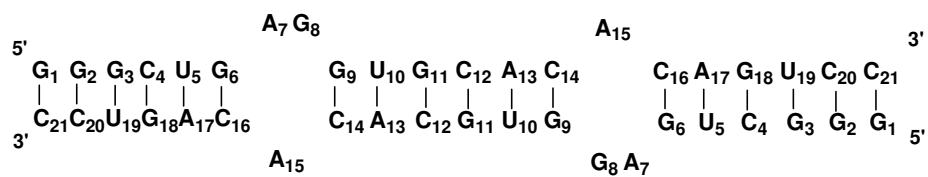


Fig 3.10. The 2D  $\text{HCCCH}$ -COSY experiment for the assignment of ribose H1'-H2' correlation of the DIS21 duplex with 0.1 mM  $\text{Mg}^{2+}$  and 50 mM  $\text{Na}^+$  (pH 6.5, 298K).

The nucleoside spin systems were mapped onto the sequence to obtain sequential assignments *via* through-space correlated NOESY spectroscopy (78, 107). The pyrimidine H6-H5 cross peaks observed from the DQ-COSY and TOCSY experiments also appeared in NOESY spectrum in  $\text{D}_2\text{O}$  at 250 ms mixing-time and could be immediately identified (Fig 3.11). These cross peaks were more resolved in the

presence of 2 mM  $\text{Mg}^{2+}$  than in the presence of 0.1 mM  $\text{Mg}^{2+}$  in the NOESY spectrum (Fig 3.11 (B)). The results are summarized in Fig 3.11 (A) in the presence of 50 mM  $\text{Na}^+$  and Fig 3.11 (B) for the presence of 2 mM  $\text{Mg}^{2+}$  and 50 mM  $\text{Na}^+$ , with NOE walks obtained for the A-form RNA helical regions of the structure. An example of NOE walk connecting two residues A15 and C16 involved in intra- and inter-residue NOE interactions was shown in Fig. 3.11 (B) with annotated blue lines. Sequential purine H8 or pyrimidine H6 to the H1' assignments were determined by the NOE walks along the backbone of the DIS21 helix from 5' to 3'. These protons are closest in an A-form helix within 5 Å. Table 3.4 and Table 3.5 summarize the assignments of the DIS21 duplex in the presence of 50 mM  $\text{Na}^+$  and with the addition of 2 mM  $\text{Mg}^{2+}$ , respectively.





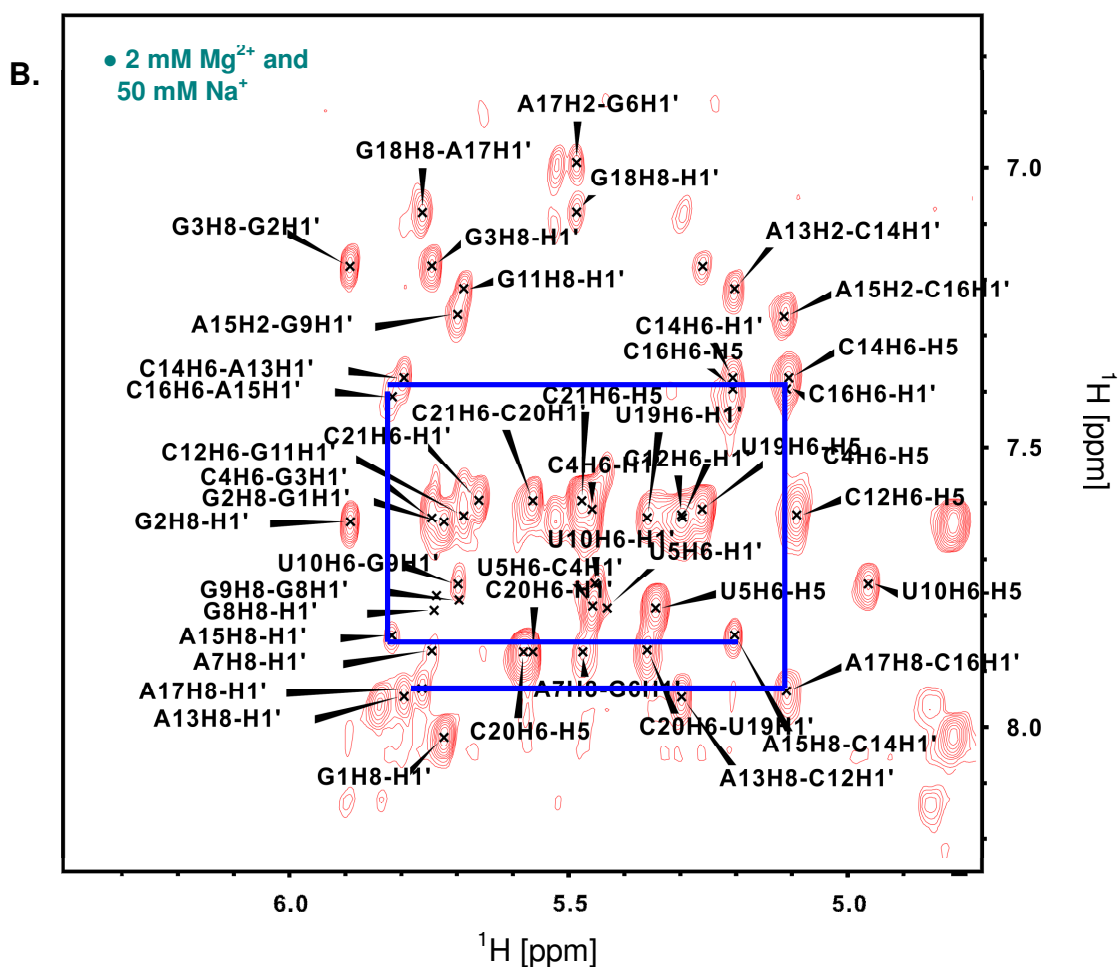


Fig 3.11. Sequential assignments for the DIS21 duplex in the presence of 50 mM  $\text{Na}^+$  (A) as well as in the presence of 2 mM  $\text{Mg}^{2+}$  and 50 mM  $\text{Na}^+$  (B) were obtained using the NOESY spectrum in  $\text{D}_2\text{O}$  with a mixing time of 250 ms. Assignments based on NOE walks are annotated for two residues, A15 and C16, with inter-and intra connectivities shown in blue lines.

| Atom<br>Base \ | H1' | H2' | H2  | H5  | H6  | H8  | C1'  | C2'  | C2    | C6    | C8    |
|----------------|-----|-----|-----|-----|-----|-----|------|------|-------|-------|-------|
| G1             | 5.7 | 4.9 |     |     |     | 8.1 | -    | -    |       |       | 138.0 |
| G2             | 5.9 | 4.6 |     |     |     | 7.6 | 90.0 | 73.4 |       |       | 134.5 |
| G3             | 5.7 | 4.6 |     |     |     | 7.2 | 91.1 | u.a. |       |       | 133.9 |
| C4             | 5.5 | 4.3 |     | 5.3 | 7.6 |     | 91.8 | 73.3 |       | 137.5 |       |
| U5             | 5.5 | 4.5 |     | 5.3 | 7.7 |     | 91.7 | u.a. |       | 138.7 |       |
| G6             | 5.5 | 4.4 |     | -   | 7.6 |     | 90.5 | 72.9 |       |       | -     |
| A7             | 5.8 | -   | 7.9 |     |     | 7.9 | -    | -    | 153.3 |       |       |
| G8             | 5.7 | 4.7 |     |     |     | 7.8 | 91.2 | 72.9 |       |       | u.a.  |
| G9             | 5.7 | 4.4 |     |     |     | 7.8 | 89.8 | 73.6 |       |       | u.a.  |
| U10            | 5.5 | 4.6 |     | 5.0 | 7.7 |     | 91.2 | u.a. |       | 139.1 |       |
| G11            | 5.7 | 4.4 |     |     |     | 7.3 | 90.4 | 73.0 |       |       | u.a.  |
| C12            | 5.3 | 4.4 |     | 5.1 | 7.6 |     | 91.3 | 72.9 |       | 137.8 |       |
| A13            | 5.8 | 4.5 | 7.2 |     |     | 7.9 | 90.5 | 73.2 | 150.3 |       |       |
| C14            | 5.3 | 4.4 |     | 5.1 | 7.3 |     | 91.0 | 73.1 |       | 137.4 |       |
| A15            | 5.9 | 4.4 | 7.2 |     |     | 7.8 | 90.5 | 73.5 | 150.9 |       |       |
| C16            | 5.2 | 4.3 |     | 5.2 | 7.3 |     | 91.4 | 72.8 |       | -     |       |
| A17            | 5.8 | 4.6 | 7.0 |     |     | 7.9 | 90.3 | 73.2 | 149.7 |       |       |
| G18            | 5.5 | 4.4 |     |     |     | 7.1 | 90.5 | 72.9 |       |       | 132.8 |
| U19            | 5.4 | 4.0 |     | 5.3 | 7.6 |     | 91.5 | 73.4 |       | 139.1 |       |
| C20            | 5.6 | 4.2 |     | 5.6 | 7.9 |     | 91.6 | 73.2 |       | 139.7 |       |
| C21            | 5.7 | 3.9 |     | 5.5 | 7.6 |     | 90.5 | 76.7 |       | 138.1 |       |

Table 3.4. Assignments ( $^1\text{H}$  and  $^{13}\text{C}$ , ppm) for the DIS21 duplex in the presence of 50 mM  $\text{Na}^+$ . “-“ represents cross peaks that are not observed and “u.a.” is noted as unassigned peaks.

| Atom<br>Base \ | H1' | H2' | H2  | H5  | H6  | H8   | C1'  | C2'  | C2    | C6    | C8    |
|----------------|-----|-----|-----|-----|-----|------|------|------|-------|-------|-------|
| G1             | 5.7 | 4.8 |     |     |     | 8.1  | -    | -    |       |       | 138.0 |
| G2             | 5.9 | 4.6 |     |     |     | 7.6  | 91.8 | 75.1 |       |       | 134.5 |
| G3             | 5.7 | 4.5 |     |     |     | 7.2  | 92.8 |      |       |       | 133.9 |
| C4             | 5.5 | 4.2 |     | 5.3 | 7.6 |      | 93.3 |      |       | 137.5 |       |
| U5             | 5.4 | 4.3 |     | 5.3 | 7.8 |      | 93.1 |      |       | 138.7 |       |
| G6             | -   | -   |     |     |     | -    | -    | -    |       |       | -     |
| A7             | 5.7 | -   | 8.1 |     |     | 7.9  | -    | -    | 153.3 |       |       |
| G8             | 5.7 | -   | -   |     |     | 7.8  | 92.9 | -    |       |       | -     |
| G9             | 5.7 | -   |     |     |     | 7.8  | 91.9 | u.a. |       |       | -     |
| U10            | 5.5 | 4.5 |     | 4.9 | 7.7 |      | 92.6 | u.a. |       | 139.1 |       |
| G11            | 5.7 | 4.3 |     |     |     | u.a. | 91.7 | 74.6 |       |       | -     |
| C12            | 5.3 | 4.3 |     | 5.1 | 7.6 |      | 92.7 | 74.4 |       | 137.8 |       |
| A13            | 5.8 | 4.4 | 7.2 |     |     | 7.9  | 92.0 | 74.8 | 150.3 |       |       |
| C14            | 5.2 | 4.3 |     | 5.1 | 7.4 |      | 92.6 | 74.6 |       | 137.4 |       |
| A15            | 5.8 | -   | 7.8 |     |     |      | 92.5 | u.a. | 150.9 |       |       |
| C16            | 5.1 | 4.1 |     | 5.2 | 7.4 |      | 92.9 | 74.4 |       | -     |       |
| A17            | 5.8 | -   | 7.0 |     |     | 7.9  | 92.0 | 74.8 | 149.7 |       |       |
| G18            | 5.5 | 4.4 |     |     |     | 7.1  | 91.8 | 74.5 |       |       | 132.8 |
| U19            | 5.4 | 4.0 |     | 5.3 | 7.6 |      | 92.9 | 74.9 |       | 139.1 |       |
| C20            | 5.6 | 4.2 |     | 5.6 | 7.9 |      | 93.2 | 74.8 |       | 139.7 |       |
| C21            | 5.7 | 3.9 |     | 5.5 | 7.6 |      | 92.2 | 76.7 |       | 138.1 |       |

Table 3.5. Assignments ( $^1\text{H}$  and  $^{13}\text{C}$ , ppm) for the DIS21 duplex in the presence of 2 mM  $\text{Mg}^{2+}$  and 50 mM  $\text{Na}^+$ . “-” represents cross peaks that are not observed and “u.a.” is noted as unassigned peaks.

### **DIS21 Kissing Homodimer**

The DIS homo-dimeric kissing-loop dimer was diluted to 10  $\mu\text{M}$  then snap-cooled, and  $\text{MgCl}_2$  to 5 mM was added to kinetically trap the kissing dimer. The RNA sample was then reconcentrated to 0.4 mM followed by microdialysis against NMR buffer with 100  $\mu\text{M}$   $\text{MgCl}_2$  and 50 mM  $\text{Na}^+$  at pH 6.5. As judge from the imino and other  $^1\text{H}$  2D spectra, particularly U5, U10, and G18 resonances, a mixture of homodimeric kissing and duplex dimers was obtained using this procedure (Fig 3.12 (B, D)). The heterokissing dimers with altered stem sequences were therefore used to model the kissing heterodimer.

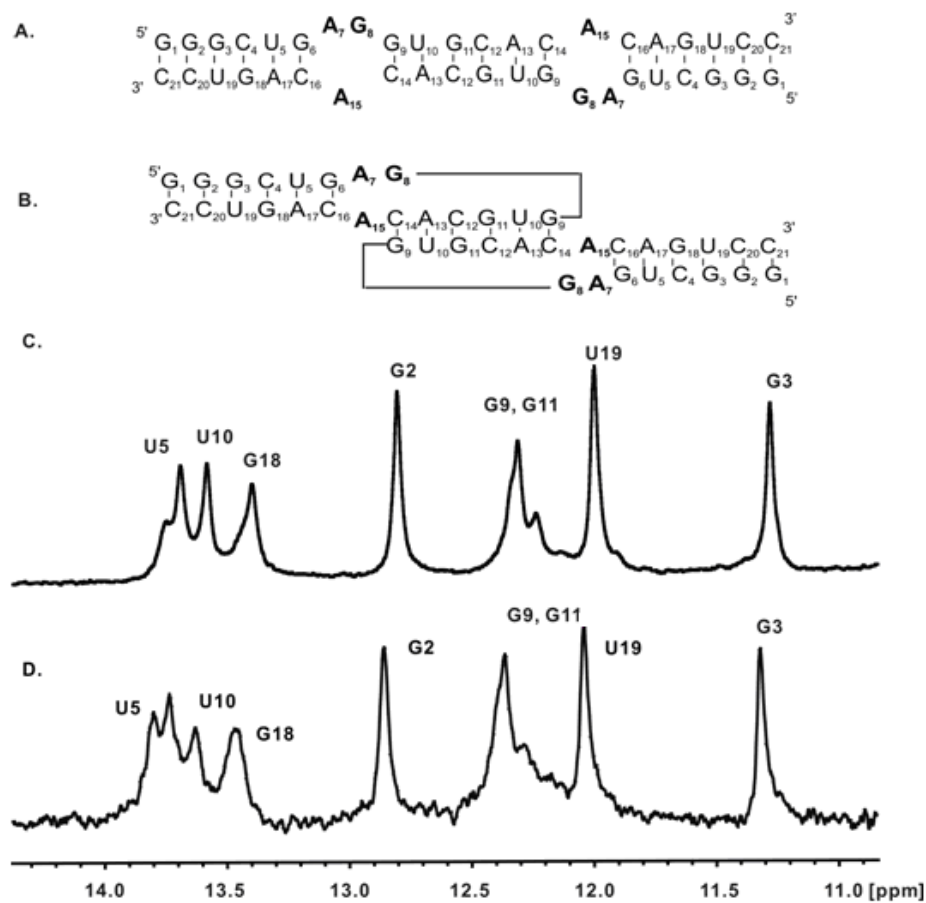


Fig 3.12. The secondary structure of DIS21 duplex (A) and kissing dimer (B). 1D imino proton resonances of DIS21 duplex (C) and mixture of duplex and kissing dimer (D) in the presence of 0.1 mM  $Mg^{2+}$  and 50 mM  $Na^+$  at pH 6.5 and 298K.

### $^{13}\text{C}$ $^{15}\text{N}$ -DIS23(GA) Hairpin

*Assignments of Exchangeable Protons* — Uracil and guanine imino resonances involved in hydrogen bonds in base pairs were obtained from  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC spectrum of the uniformly labeled DIS23(GA) (Fig 3.13). The U N3-H3 resonances appear at upper field ( $\sim 13.4$ - $13.8$  ppm).

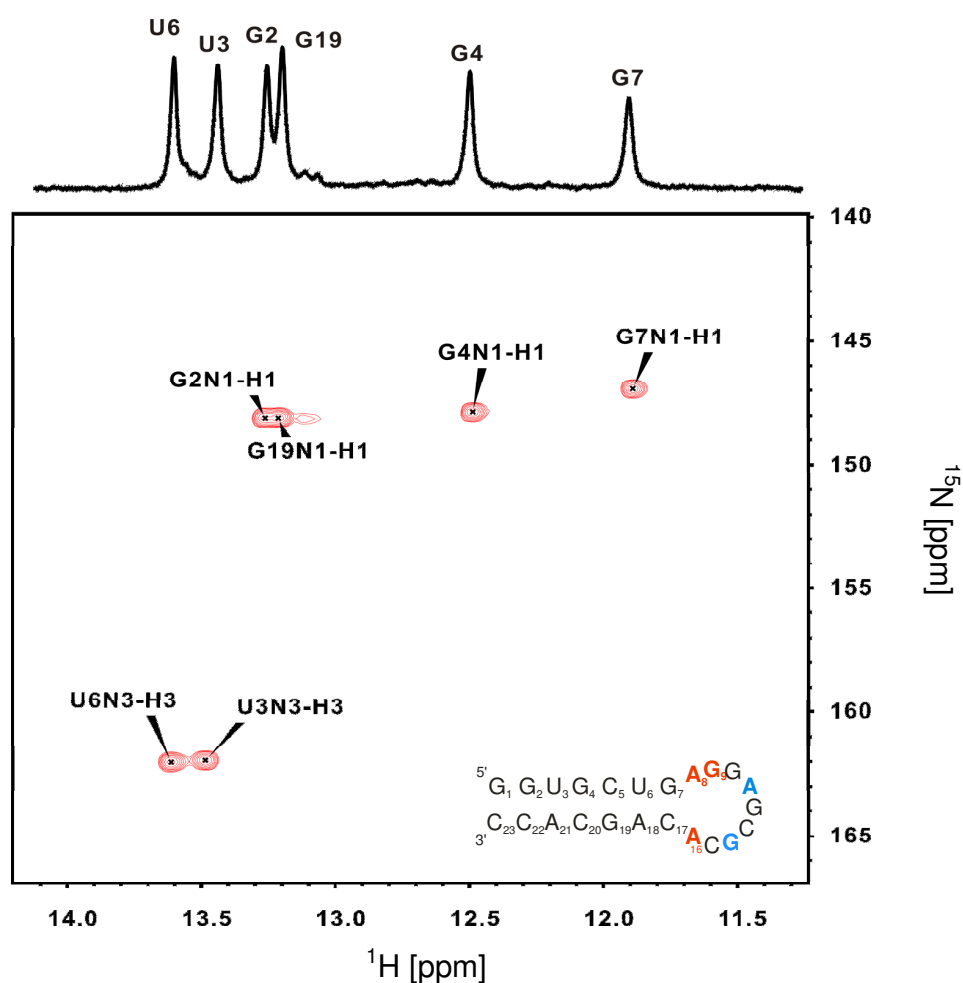


Fig 3.13.  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC of the uniformly labeled  $^{13}\text{C}$   $^{15}\text{N}$ -DIS23(GA) hairpin loop recorded in NMR buffer in the presence of 200 mM  $\text{Na}^+$  at 298K. A and G nucleotides which are point mutations in the loop are highlighted in blue. Purine bases on the DIS junction are highlighted in red.

The H5-H6 cross peaks of DIS23(GA) hairpin loop were identified by a COSY experiment shown in Fig 3.14. As expected, seven H6-H5 cross peaks in the stem of the hairpin are found. Two loop bases, C13 and C15, are downfield shifted.

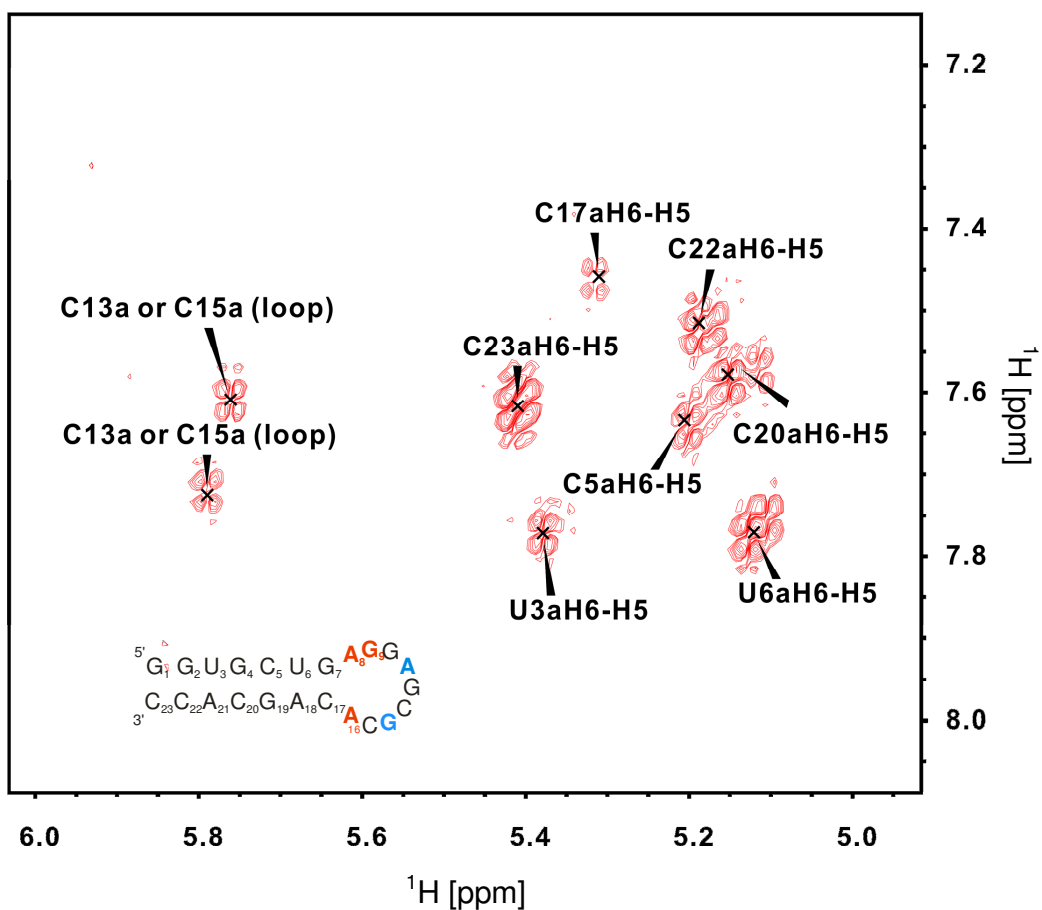


Fig 3.14. Chemical shifts of pyrimidine H5-H6 proton cross peaks of the DIS23(GA) hairpin loop. The lower case “a” refers to DIS23(GA) side.



### **DIS23(HxUC) Hairpin**

*Assignments of Exchangeable Protons* — In order to avoid a mixture of monomer and homodimer in the DIS23(HxUC) sample, a diluted RNA sample was made followed by snap-cooling and then concentration, to give predominantly the hairpin-loop conformation in low salt NMR buffer. (A minor fraction of homodimers is still observed). 2D-NOE spectroscopy was employed using a jump-and-return pulse sequence to obtain NOE interactions between amino-imino region and pyrimidine H5-imino region of DIS23(HxUC) hairpin as shown in Fig 3.15. Spin diffusion through the amino proton resonances can provide the connection to individual cytosine H5-H6 resonances on the DIS stem.

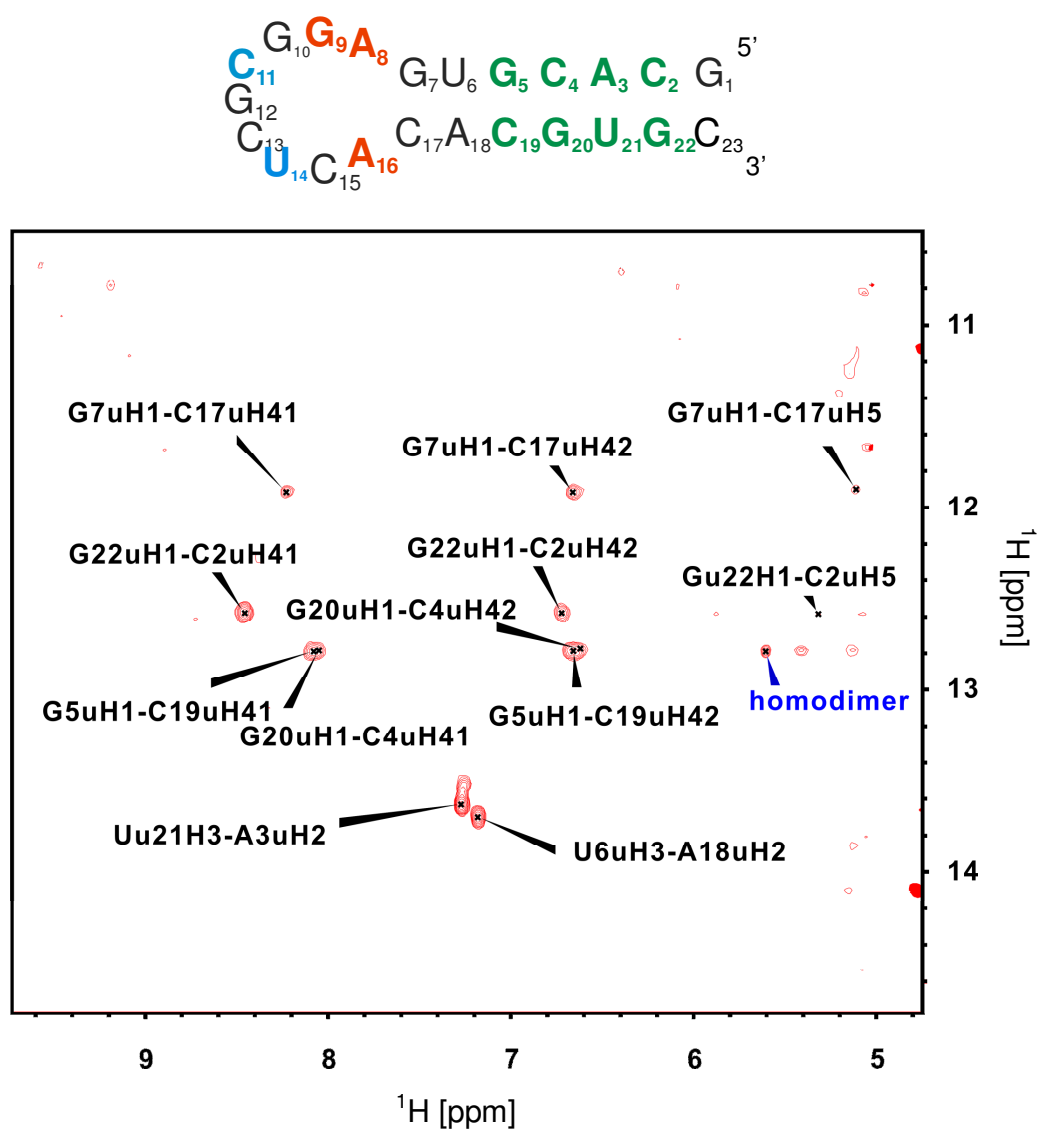


Fig 3.15. Amino-imino proton region of the DIS23(HxUC) hairpin loop in a 2D NOESY spectrum in NMR buffer at 298K and pH 6.5. The lower case “u refers to DIS23(HxUC) side. U and C nucleotides which are point mutations in the loop are highlighted in blue. Purine bases on the DIS junction are highlighted in red.

*Assignments of Non-Exchangeable Protons* — The positions of pyrimidine H5-H6 cross peaks were confirmed by the TOCSY experiment to provide a starting point for NOE sequential assignment. Three cytidine cross peaks on the loop are downfield shifted in the DIS23(HxUC) hairpin (Fig 3.16).

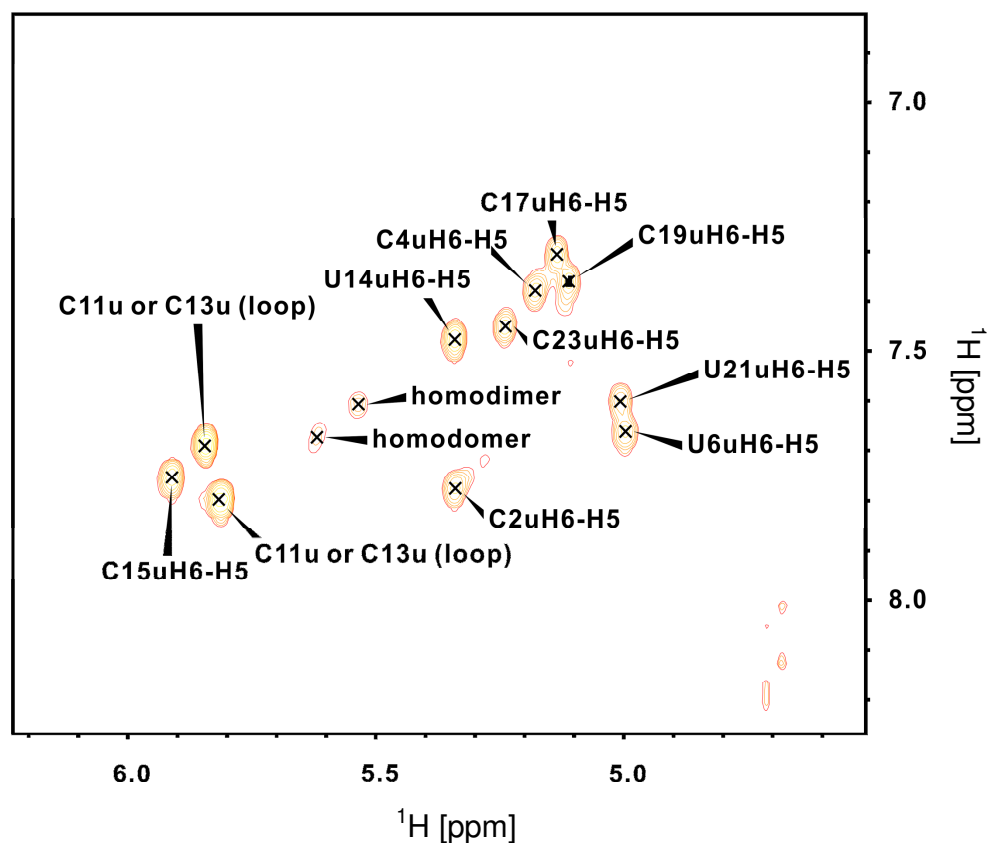


Fig. 3.16. 2D TOCSY experiment was used to reveal H5-H6 connectivities for DIS23 (HxUC) in NMR buffer (25 mM  $\text{Na}^+$ , pH 6.5, and 298K).

### **$^{13}\text{C}^{15}\text{N}$ -DIS23(GA)•DIS23(HxUC) Heterokissing Complex**

NMR studies for identification of the  $\text{Mg}^{2+}$  binding sites on the DIS stem-loop kissing complex were carried out by forming a kissing dimer with uniformly  $^{13}\text{C}^{15}\text{N}$ -labeled DIS23(GA) and unlabeled DIS23(HxUC). The assignments of the 1D imino proton resonances of  $^{13}\text{C}^{15}\text{N}$ -DIS23 (GA) and DIS23(HxUC) are shown in Fig 3.17 (B) and (C), respectively. The complete assignment of the imino proton spectra obtained at pH 7.5 shows that the imino resonances of G10, G12, and G14 (in red,  $^{13}\text{C}^{15}\text{N}$ -DIS23(GA) side) and G10, G12, and U14 (in blue, DIS23(HxUC) side) are associated with new base pairs in the loop-loop of the DIS kissing dimer. The heterokissing complex is stabilized with high concentration of  $\text{Na}^+$  (200 mM) to saturate the non-specific metal binding sites. The chemical shifts of the rest of resonances from the stems also varied upon the formation of the kissing complex (Fig 3.17 (D)). The resonances of  $^{13}\text{C}^{15}\text{N}$ -labeled DIS23(GA) of the DIS kissing complex can be easily observed by 2D heteronuclear experiments (Fig 3.17 (E)).

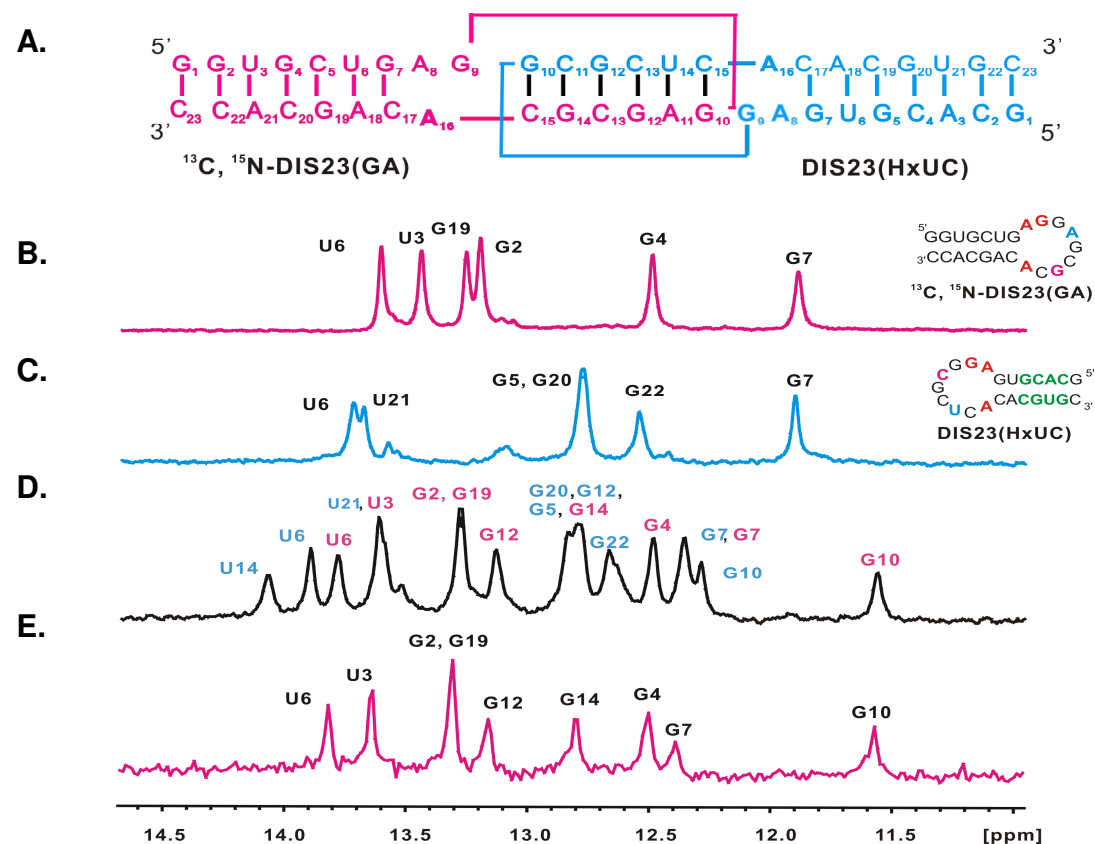
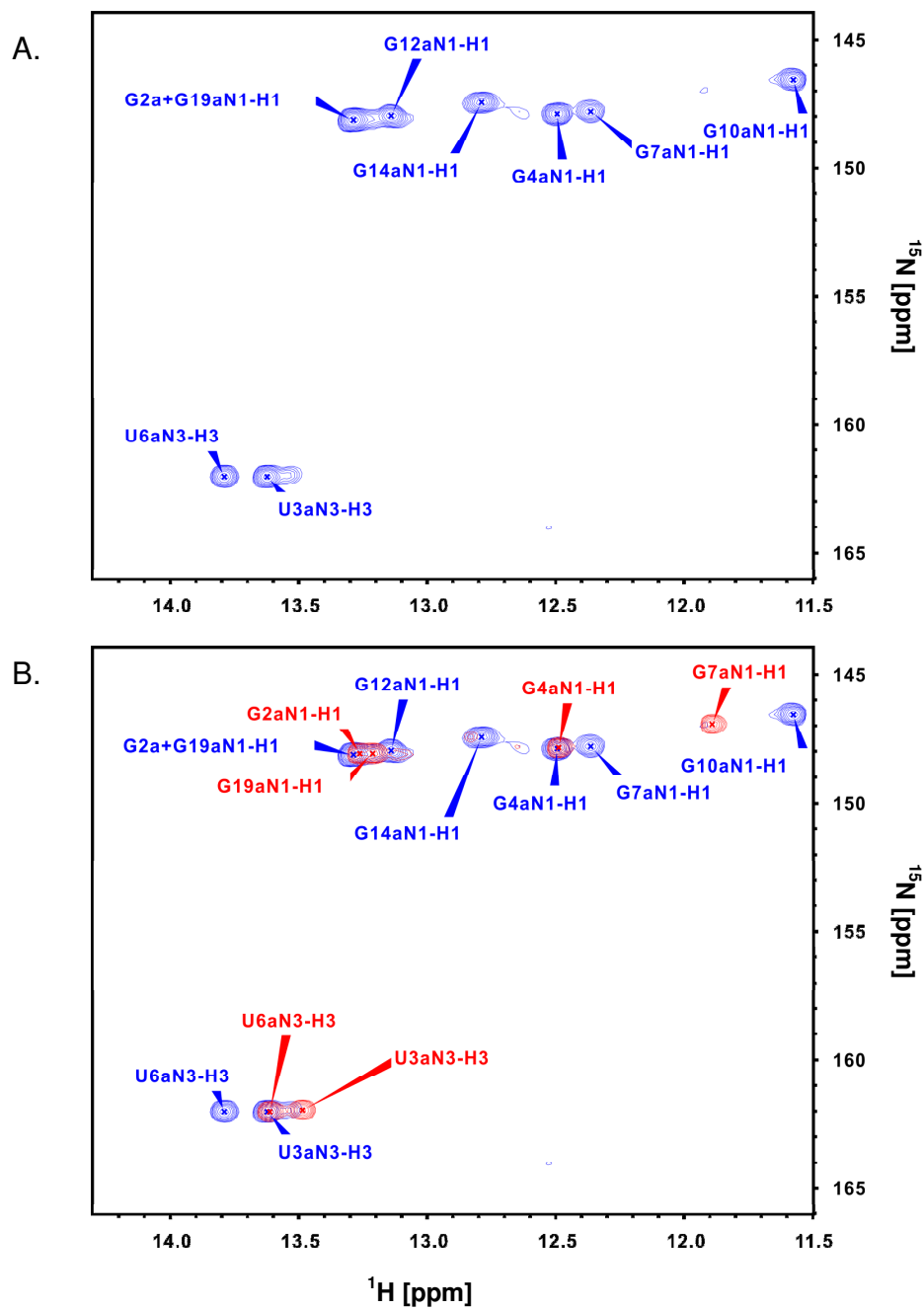
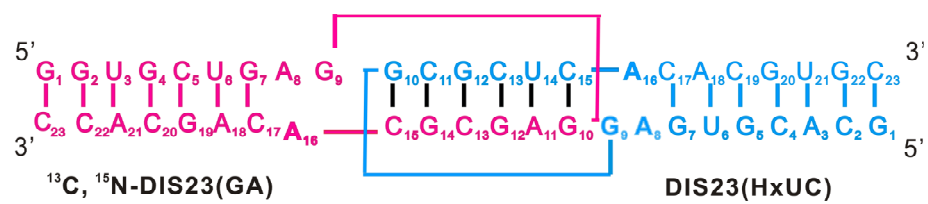


Fig. 3.17. Sequence and secondary structure of the DIS23(GA)•DIS23(HxUC) heterodimeric kissing complex formed by uniformly  $^{13}\text{C}^{15}\text{N}$ -labeled DIS23(GA) in red and unlabeled DIS23(HxUC) with mismatched stem sequence in blue. Assigned imino proton resonances (B) DIS23(GA); (C) DIS23(HxUC); (D) [DIS23(GA) in red and DIS23(HxUC) in blue] of the DIS kissing dimer ; and (E) the 1<sup>st</sup> FID of a 2D  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC experiment on  $^{13}\text{C}^{15}\text{N}$ -labeled DIS23(GA) site of the DIS kissing complex in NMR buffer at pH 7.5 and 200 mM  $\text{Na}^+$  (298K).

### **$^{13}\text{C}$ $^{15}\text{N}$ -DIS23(GA)•DIS23(HxUC) Kissing Complex**

*Assignments of Exchangeable Protons* — The imino nitrogen atoms corresponding to the complete  $^{13}\text{C}^{15}\text{N}$ -labeled DIS23(GA) resonances in the  $^{13}\text{C}^{15}\text{N}$ -DIS23(GA)•DIS23(HxUC) kissing complex was obtained by the  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC experiment shown in Fig 3.18 (A). The comparison of two  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC spectra between  $^{13}\text{C}^{15}\text{N}$ -DIS23(GA)•DIS23(HxUC) kissing complex and  $^{13}\text{C}^{15}\text{N}$ -DIS23(GA) hairpin is shown in Fig 3.18 (B) in the presence of 200 mM  $\text{Na}^+$ . The amino-imino resonances of G10, G12, and G14 were associated with formation of base pairs in the loop-loop of the DIS kissing dimer on the DIS23(GA) hairpin. In addition, the amino-imino resonances of U3, U6, and G7 were found to be downfield shifted in the formation of the heterokissing complex. The addition of 200  $\mu\text{M}$   $\text{Mg}^{2+}$  allows  $\text{Mg}^{2+}$  to substitute for  $\text{Na}^+$  at the higher affinity divalent binding sites without overall line broadening (Fig 3.18 (C)). The unambiguous assignments of all the imino proton resonances and associated adenine resonances (H2, data not shown) were obtained from the unlabeled kissing complex in the presence of 200 mM  $\text{Na}^+$  (Fig 3.19).



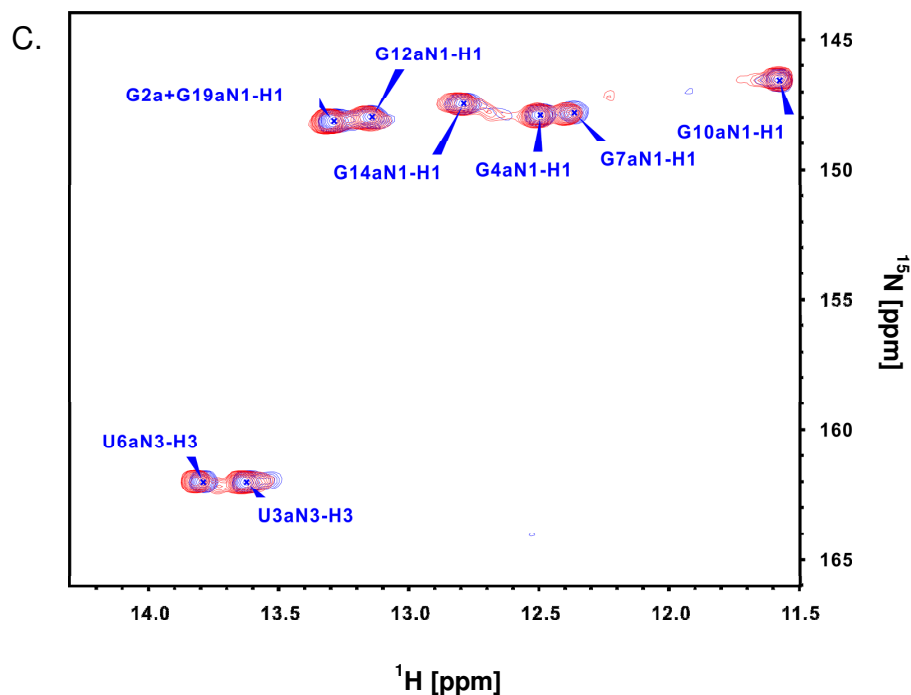


Fig 3.18. (A)  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC of the uniformly labeled  $^{13}\text{C}^{15}\text{N}$ -DIS23(GA)•DIS23(HxUC) kissing complex was recorded in NMR buffer in the presence of 200 mM  $\text{Na}^+$  (298K, pH 7.5). The lower case “a” denotes DIS23(GA) in the kissing complex. (B) Overlay of two  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC spectra of the  $^{13}\text{C}^{15}\text{N}$ -DIS23(GA)•DIS23(HxUC) kissing complex (labeled in blue) and  $^{13}\text{C}^{15}\text{N}$ -DIS23(GA) DIS hairpin (labeled in red). (C) Overlay of two  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC spectra of  $^{13}\text{C}^{15}\text{N}$ -DIS23(GA)•DIS23(HxUC) kissing complex with 200 mM  $\text{Na}^+$  (labeled in blue) and  $^{13}\text{C}^{15}\text{N}$ -DIS23(GA)•DIS23(HxUC) kissing complex with 200  $\mu\text{M}$   $\text{Mg}^{2+}$  and 200 mM  $\text{Na}^+$  (labeled in red).



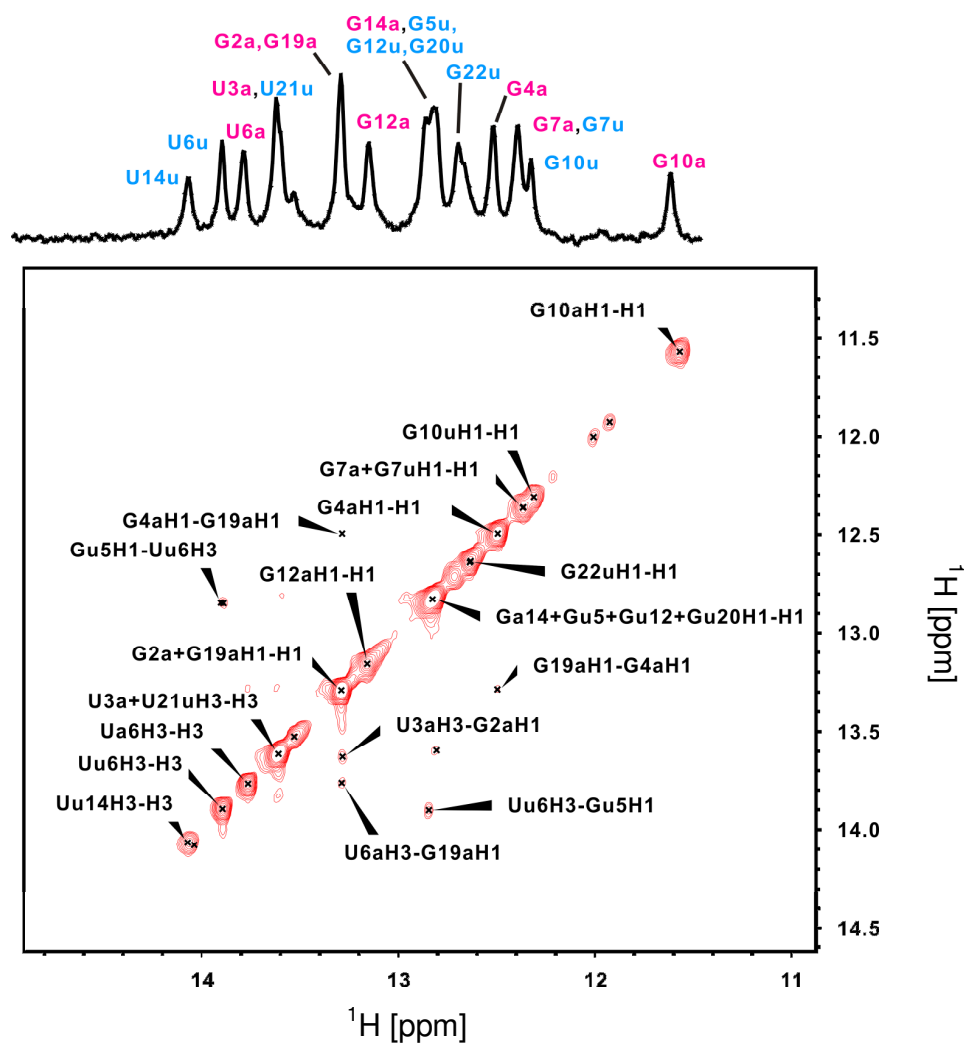


Fig 3.19. The imino-imino proton cross peak region and of the DIS23(GA)•DIS23(HxUC) kissing complex in a 2D NOESY spectrum in NMR buffer in the presence of 200 mM Na<sup>+</sup> (298K and pH 7.5). The lower case “a” denotes DIS23(GA) and “u” denotes DIS23(HxUC) in the kissing complex.

*Assignments of Non-Exchangeable Protons* — The assignments of pyrimidine H5-H6 cross peaks are shown in the TOCSY experiment (Fig 3.20). H2, H6, and H8 resonances obtained from  $^1\text{H}$ ,  $^{13}\text{C}$  CT-HSQC aromatic region were used to assign the cross peaks from 2D NOESY experiment (Fig 3.21). The cross peaks corresponding to NMR active isotope ( $^{13}\text{C}$ ) labeled DIS23(GA) were obtained by  $^{13}\text{C}$  edited 2D NOESY experiment (Fig 3.22 (A)), whereas, the unlabeled DIS23(HxUC) hairpin in the kissing complex was measured by  $^{12}\text{C}$  filtered 2D NOESY experiment (Fig 3.22 (B)). A combination of isotope editing and filtering spectra presents the cross peaks conforming to the DIS23(GA)•DIS23(HxUC) kissing complex (Fig 3.23). In the case of severe signal overlap, it can allow the resolution of signals from the labeled DIS23(GA) from the unlabeled DIS23(HxUC). The sequential assignments for labeled DIS23(GA) were simplified by extracting correlations from the selective observations in this filter experiment applied to the partially labeled kissing complex (Fig 3.23). A summary of resonance assignments of the labeled DIS23(GA) kissing loop in the heterokissing complex was presented in Table 3.6.

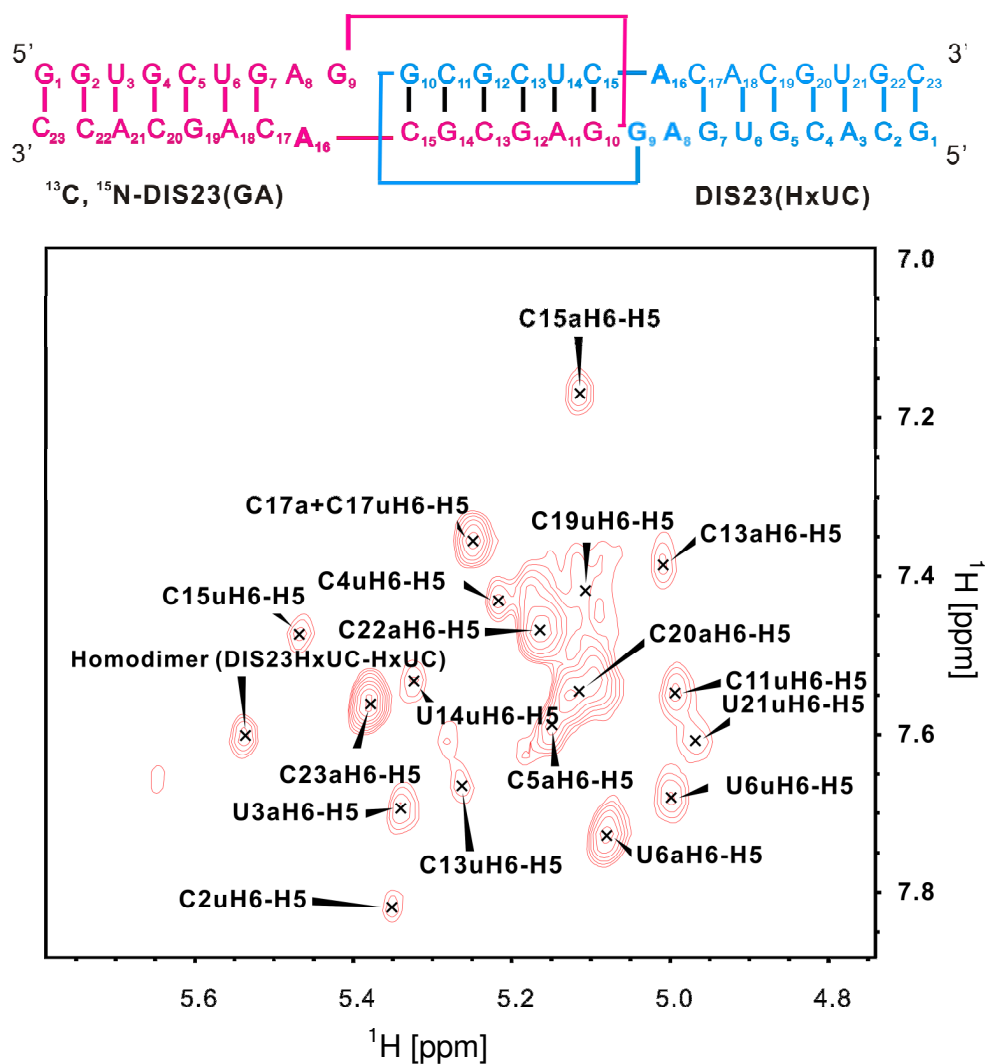


Fig 3.20. H5-H6 cross peaks of unlabeled DIS23(GA)•DIS23 (HxUC) kissing complex in the TOCSY spectrum in the presence of 200 mM  $Na^+$  and 200  $\mu M$   $Mg^{2+}$  at 298K and pH 7.5.

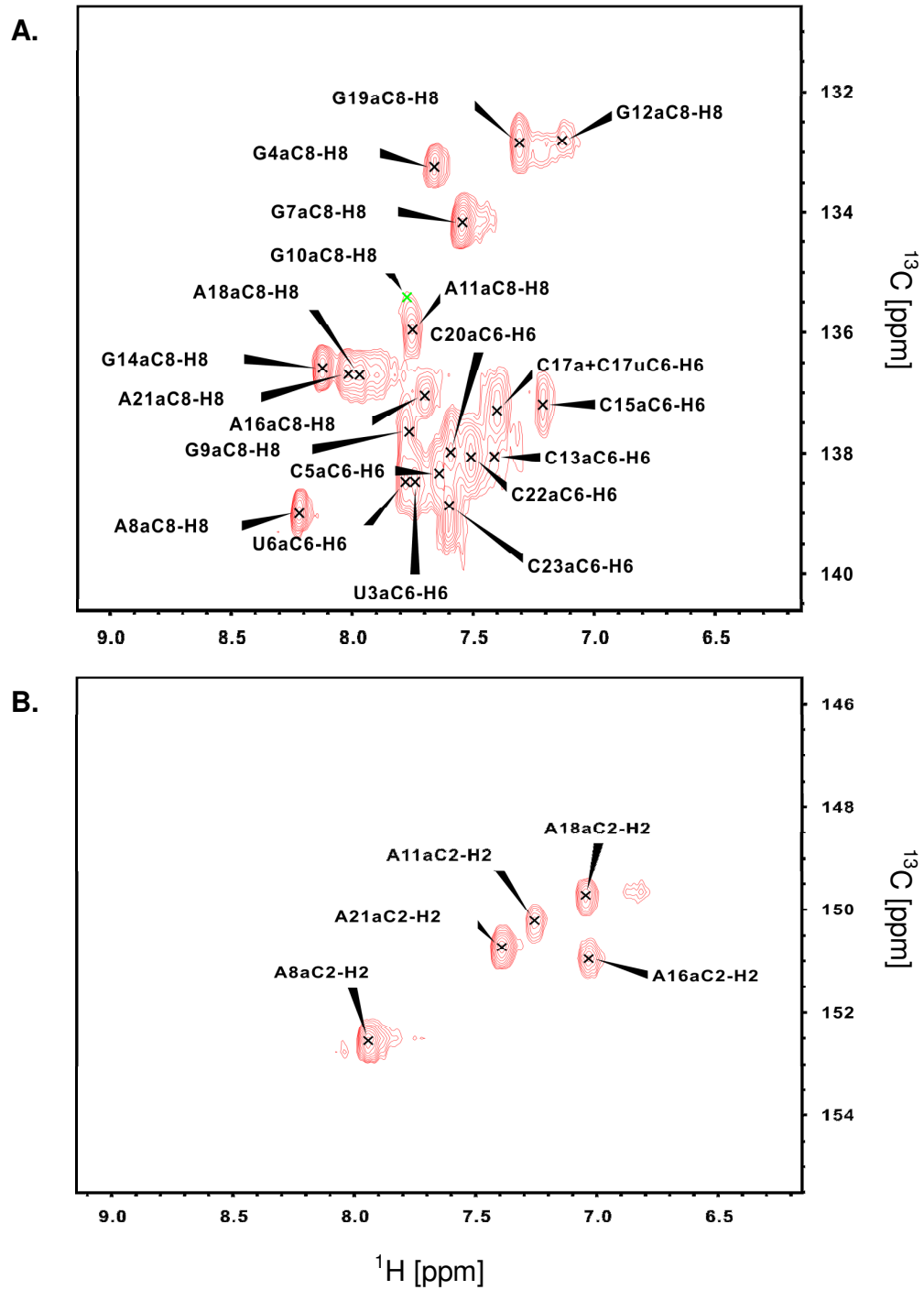


Fig 3.21. A 2D CT-HSQC experiment shows (A) H8-C8 and H6-C6 (B) H2-C2 correlations of  $^{13}\text{C}^{15}\text{N}$ -DIS23(GA)•DIS23(HxUC) kissing complex in the presence of 200 mM  $\text{Na}^+$  and 200  $\mu\text{M}$   $\text{Mg}^{2+}$  (pH 7.5, 298K).

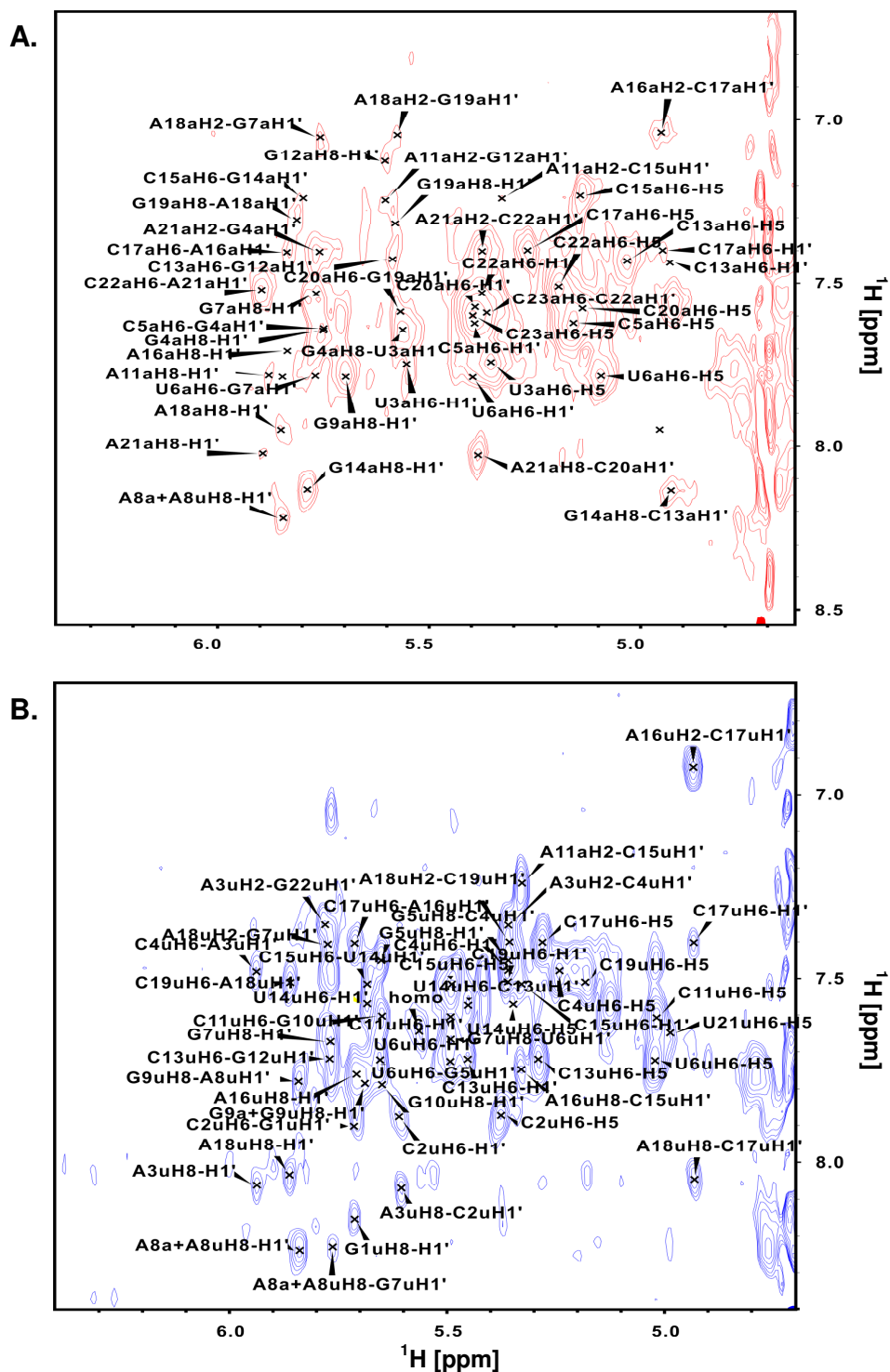


Fig 3.22. (A)  $^{13}\text{C}$  edited 2D NOESY spectrum of labeled DIS23(GA); (B)  $^{12}\text{C}$  filtered 2D NOESY spectrum of unlabeled DIS23(HxUC) in the  $^{13}\text{C}^{15}\text{N}$ -DIS23(GA)•DIS23(HxUC) kissing complex.

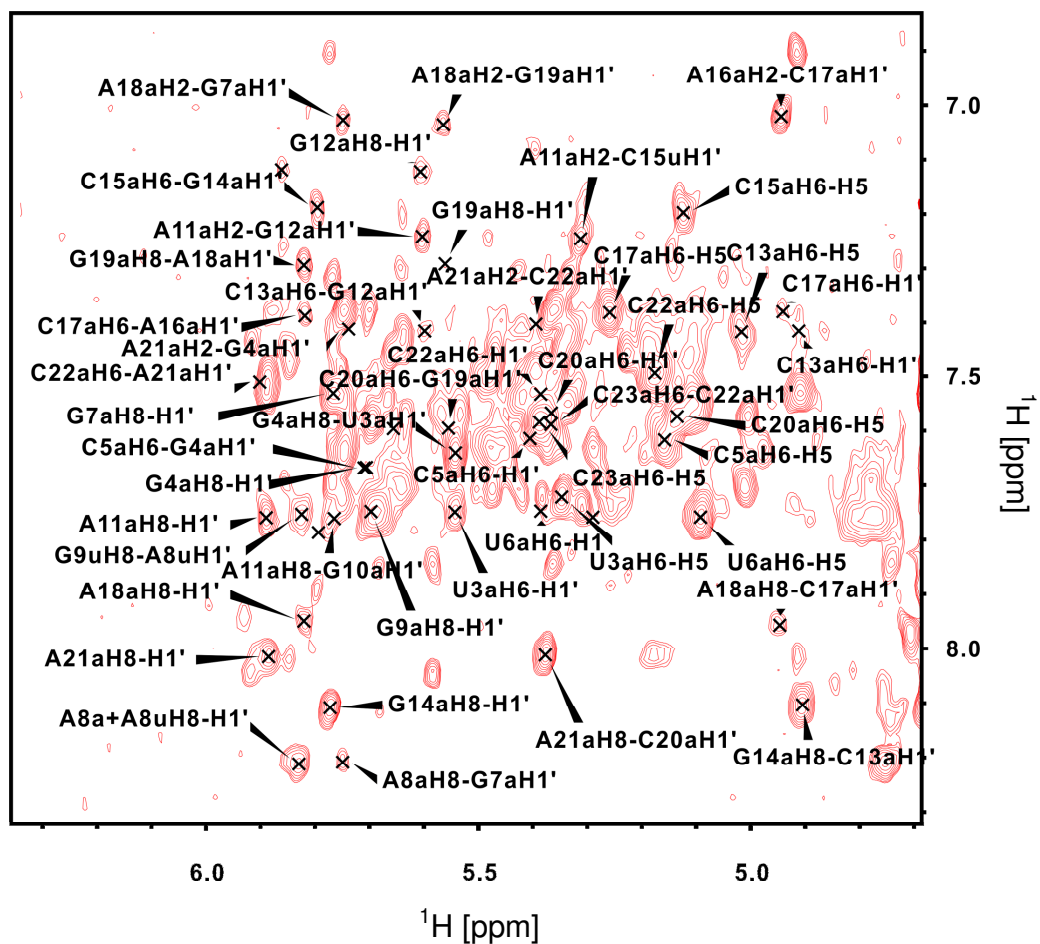


Fig 3.23. The sequential assignments of DIS23(GA) side in the DIS23(GA)•DIS23(HxUC) kissing complex. The assignments were retrieved from a  $^{13}\text{C}$  edited 2D NOESY experiment in the presence of 200 mM  $\text{Na}^+$  and 200  $\mu\text{M}$   $\text{Mg}^{2+}$  (pH 7.5, 298K). by sequential NOE contacts.

| Base \ Atom | H1' | H2  | H5  | H6  | H8  | C2    | C6    | C8    |
|-------------|-----|-----|-----|-----|-----|-------|-------|-------|
| U3a         | 5.5 |     | 5.1 | 7.8 |     |       | 138.5 |       |
| G4a         | 5.7 |     |     |     | 7.6 |       |       | 133.2 |
| C5a         | 5.3 |     | 5.2 | 7.6 |     |       | 138.4 |       |
| U6a         | 5.4 |     | 5.0 | 7.7 |     |       | 138.5 |       |
| G7a         | 5.7 |     |     |     | 7.5 |       |       | 134.2 |
| A8a         | 5.8 | -   |     |     |     | 152.6 |       |       |
| G9a         | 5.7 |     |     |     | 7.8 |       |       | 137.7 |
| G10a        | 5.8 |     |     |     | 7.7 |       |       | 135.4 |
| A11a        | 5.9 | 7.2 |     |     | 7.7 | 150.2 |       |       |
| G12a        | 5.6 |     |     |     | 7.1 |       |       | 132.8 |
| C13a        | 5.5 |     | 5.0 | 7.4 |     |       | 138.1 |       |
| G14a        | 5.8 |     |     |     | 8.1 |       |       | 136.6 |
| C15a        | -   |     | 5.1 | 7.2 |     |       | 137.2 |       |
| A16a        | -   | 7.0 |     |     |     | 150.9 |       |       |
| C17a        | 4.9 |     | 5.3 | 7.4 |     |       | 137.3 |       |
| A18a        | 5.8 | 7.0 |     |     | 8.0 | 149.7 |       |       |
| G19a        | 5.6 |     |     |     | 7.3 |       |       | 132.9 |
| C20a        | 5.4 |     | 5.1 | 7.6 |     |       | 138.0 |       |
| A21a        | 5.9 | 7.4 |     |     | 8.0 | 150.7 |       |       |
| C22a        | 5.4 |     | 5.1 | 7.6 |     |       | 138.1 |       |

Table 3.6. The assignments of the DIS23(GA) side in the  $^{13}\text{C}^{15}\text{N}$ -DIS23(GA)•DIS23(HxUC) kissing complex for each base in the presence of 200 mM  $\text{Na}^+$  and 200  $\mu\text{M}$   $\text{Mg}^{2+}$  (pH 7.5, 298K). “-” represents cross peaks that are not observed.

### 3.5 Conclusion

The resonances of two model RNA hairpin system for the extended duplex as well as kissing DIS complex [DIS23(GA)•DIS23(HxUC)] were assigned by NMR methods in the presence of Na<sup>+</sup> and Mg<sup>2+</sup>. Homonuclear and heteronuclear experiments were utilized to measure aromatic and ribose proton resonances. Specifically, isotope labeled DIS23(GA) was used the kissing complex in order to solve the spectral overlap problem. In the case of the mature DIS21 duplex dimer and DIS23(GA)•DIS23(HxUC) heterokissing complex, the apical stem G-C base pairs appear to be disrupted in the presence of Mg<sup>2+</sup>. Upon Mg<sup>2+</sup> binding, distinct structural changes were found in homo and hetero dimers. Based on these assignments, we can use complementary NMR methods such as chemical shift perturbations and paramagnetic relaxation enhancement to probe specific Mg<sup>2+</sup> binding sites on the DIS complexes.



## **Chapter 4: Characterization of $\text{Mg}^{2+}$ Binding and Associated Conformational Changes in the DIS Extended Duplex and Kissing Dimer Structures**

### **4.1 Abstract**

The assignments of the extended duplex (DIS21) and the kissing dimer (DIS23(GA)•DIS23(HxUC)) have been described in Ch 3. In this chapter, NMR methods are described which probe for divalent metal binding sites in the DIS dimer structures. We describe  $\text{Mg}^{2+}$  induced chemical shift perturbations (CSP) that were used to detect structural changes in the RNA in response to the metal binding. We also present paramagnetic relaxation enhancement (PRE) technique observed using  $\text{Mn}^{2+}$  substitution to localize  $\text{Mg}^{2+}$  binding sites in the DIS complexes. Our results suggest that the unpaired purine bases in the DIS loop junction in the DIS extended duplex form a primary  $\text{Mg}^{2+}$  metal binding site through a bulged in conformation. The heterokissing complex, purine bases in the DIS junction are found to have a similar geometry, stacking between the stem loop and kissing loop-loop helix and form an analogue pocket for  $\text{Mg}^{2+}$  ion binding.

## **4.2 Introduction**

### **4.2.1 Detecting Metal Binding to RNA**

Several techniques have been used to identify metal-binding sites including X-ray crystallography and NMR methods. Metal ion-induced cleavage experiments can also be used to probe higher affinity ion binding sites (108). NMR spectroscopy provides dynamic information about metal ion-RNA interactions in solution. Metal-ion binding can be characterized by the information of chemical shift changes, paramagnetic line broadening, and intermolecular NOE contacts from NMR spectroscopic approaches.  $\text{Mn}^{2+}$  is known to compete with  $\text{Mg}^{2+}$  due to its greater electron density.  $\text{Mn}^{2+}$  is more readily identified in the electron density maps which provide additional information for assigning  $\text{Mg}^{2+}$  chelated peaks (109).

#### **4.2.1.1 Chemical Shift Perturbation**

The chemical shifts of RNA resonances can change upon metal binding. Upon introduction of only one species,  $\text{Mg}^{2+}$ , to the system, chemical shift perturbations (CSP) can be used to directly monitor the response of the RNA to this ion. Chemical shifts can provide information not only on specific metal ion binding but may also detect conformational changes induced upon addition of  $\text{Mg}^{2+}$  to RNA. The resonances of the proton, nitrogen, carbon, and phosphorus nuclei can all report chemical shift changes induced by metal ion association with RNA. For example, one of the  $\text{Mg}^{2+}$  binding sites with higher affinity in hammerhead ribozyme has been identified by  $^{31}\text{P}$  NMR (110). In order to detect direct interactions among metal ions and nucleobases, manganese induced

paramagnetic relaxation enhancement (PRE) and NOE based detection using cobalt hexamine are often complementary methods to chemical shift mapping to probe specific  $\text{Mg}^{2+}$  binding sites in bound states.

#### 4.2.1.2 Paramagnetic Relaxation Enhancement (Manganese Based PRE-NMR)

A common NMR approach is to use the paramagnetic metal  $\text{Mn}^{2+}$  as an analogue to compete with  $\text{Mg}^{2+}$  at specific binding sites. Paramagnetic  $\text{Mn}^{2+}$  ion has five unpaired *d* electrons ( $S = 5/2$ ) with high spin ligands ( $\text{H}_2\text{O}$ ) in a weak-field octahedral complex.  $\text{Mn}^{2+}$  can displace  $\text{Mg}^{2+}$ , resulting in paramagnetic line broadening for NMR signals in close proximity to divalent metal binding sites with a distance dependence of  $r^{-6}$  (111). Three types of the relaxation of the nuclear spin by the paramagnetic ion include Fermi contact, dipolar coupling, and Curie relaxation (112). The Curie spin relaxation also called magnetic susceptibility relaxation is usually more important for  $T_2$  than for  $T_1$  and thus leads to line broadening. Due to the high number of unpaired electrons at the paramagnetic metal center, such as  $\text{Mn}^{2+}$  ( $S=5/2$ ), Curie relaxation dominates. The observed relaxation rates which depend on the dipolar coupling between the unpaired electrons are different in the presence and absence of paramagnetic ions. The paramagnetic ion will cause an enhancement in  $^{13}\text{C}$  or  $^{15}\text{N}$   $T_1$  and  $T_2$  relaxation rates, which is the reciprocal of the 6<sup>th</sup> power of the metal ion and heteronuclei distance dependent in the presence of paramagnetic metal ions (69). Therefore, the distance between a specific metal binding site and surrounding heteronuclei can be extracted from the information of this relaxation time based on detectable paramagnetic effects. Notably, paramagnetic relaxation enhancements are quite similar to NOE restraints with

respect to this distance information. Transverse nuclear relaxation rate ( $T_2$ ) also has been used for structural calculation in spin-labeled RNA to determine the structure of protein-RNA complexes (113).

In addition to the paramagnetic line broadening, the observed chemical shift affected by the presence of unpaired electrons also can provide an important source of information regarding to spin density distribution from metal center and the dynamic behavior of the binding site (114).  $Mn^{2+}$  titrations can be performed to discover the potential binding sites by chemical shift analysis and the chemical shift changes for imino resonances as a function of  $Mn^{2+}$  concentration can be fit to determine a binding isotherm (115).

#### **4.2.1.3 Cobalt Hexamine**

Cobalt hexamine is a second useful magnesium analog that can be used to locate specific metal binding sites in RNA. It provides the opportunity for direct measurement of intermolecular NOE contacts between the hexamine ligand protons and RNA because the amino groups of cobalt hexamine do not exchange with the bulk solvent. The coordinate geometry is octahedral via outer-sphere coordination. The modeling of the exact metal ion binding position from NOE contacts between ammine protons and RNA are more precise than X-ray structures in the major groove of the RNA (50). Due to the strongly positive charge on Co (III), the same octahedral geometry as  $[Mg(H_2O)_6]^{2+}$ , and the inert ammonium ligands which resemble water ligands in magnesium hexahydrate,  $[Co(NH_3)_6]^{3+}$  is considered to be an ideal mimic for  $Mg^{2+}$  to detect the metal binding sites directly in solution (116).

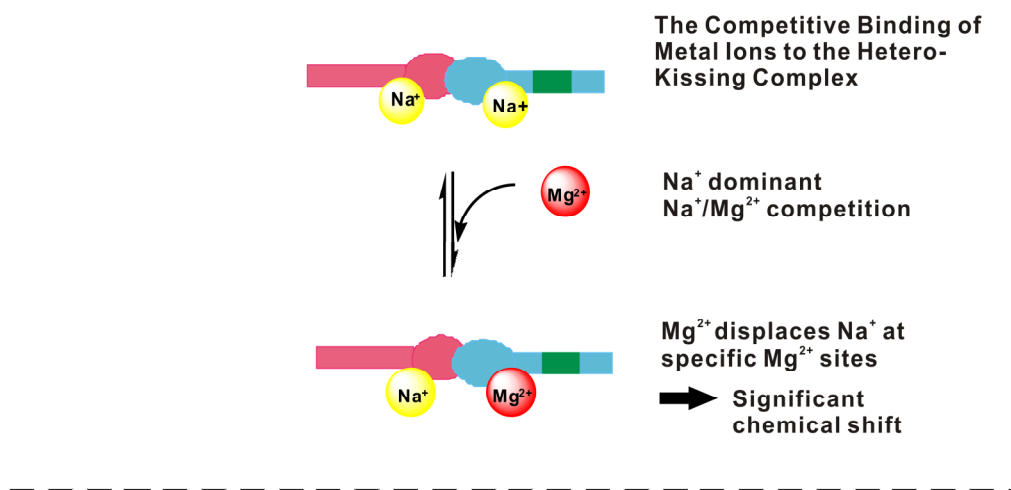
In solution studies of the P5b helix and stem loop, Kieft et al. (1997) determined through observation of NOE cross peaks between the ammine protons and RNA protons in the major groove that cobalt (III) hexammine is located in a specific binding pocket (117). The NMR solution technique is an ideal tool to detect the dynamic behavior of metal-RNA complexes compared to crystal structures. The single cobalt (III) hexammine proton resonance is at 3.65 ppm, and all RNA proton resonances are sharp in the presence of  $\text{Co}^{3+}$ . These observations suggest that there is a fast equilibrium between bound and unbound states. In addition to intermolecular NOE contacts,  $\text{Co}^{3+}$  induced chemical shift changes provide an additional information to characterize metal ion binding sites. For instance, the apparent chemical shifts are located around the G-U base pair in the P5b stem loop implies that  $\text{Co}^{3+}$  ion may bind to this region (117).

#### **4.2.2 Identification and Comparison of Sites Found for High-Affinity $\text{Mg}^{2+}$ Binding on the DIS Extended Duplex and Kissing Dimer Structures**

$\text{Mg}^{2+}$  ions bind to specific sites on RNA and neutralize the phosphate charges of RNA backbone to help form more compact RNA structures. Previous kinetic analysis of DIS dimerization and maturation showed that DIS kissing complexes are structurally different in the presence of monovalent and divalent cations. The first objective in this study is to analyze the metal dependence of the DIS kissing loop and mature duplex structures. The other objective is to identify  $\text{Mg}^{2+}$  binding sites in the two DIS dimer conformations using the DIS21 extended duplex and DIS23(GA)•DIS23(HxUC) heterokissing complex. Our focus is on the local conformational dynamics upon metal binding and the sequence dependence at the DIS unpaired junction bases (A, G, and A).

Fig 4.1 illustrates a scheme for the identification of  $\text{Mg}^{2+}$  binding sites of the DIS heterokissing complex using chemical shift perturbations and paramagnetic relaxation enhancement methods. The hetero-kissing complex is initially stabilized with  $\text{Na}^+$  to saturate nonspecific divalent metal binding sites.  $\text{Mg}^{2+}$  is then added to compete with  $\text{Na}^+$  at the sites with higher affinity for divalent ions. The unique  $\text{Mg}^{2+}$  induced chemical shift perturbations identify sites involved in conformational change and metal binding.  $\text{Mn}^{2+}$  is finally added in the complex to substitute for  $\text{Mg}^{2+}$  at the specific binding sites with the same affinity. The ratio of metal ion concentration is  $\text{Na}^+:\text{Mg}^{2+}:\text{Mn}^{2+} = 20000:20:1$ . The paramagnetic relaxation effect results in selective line broadening allowing specific divalent metal binding sites to be identified. A rapid exchange between  $\text{Mn}^{2+}$  and  $\text{Mg}^{2+}$  exists exchange rate  $\sim 10,000 \text{ s}^{-1}$  and results in a mixture of  $\text{Mn}^{2+}$  and  $\text{Mg}^{2+}$  bound states of the DIS heterokissing complexes.

## Chemical Shift Perturbation (CSP)



## Paramagnetic Relaxation Enhancement (PRE)

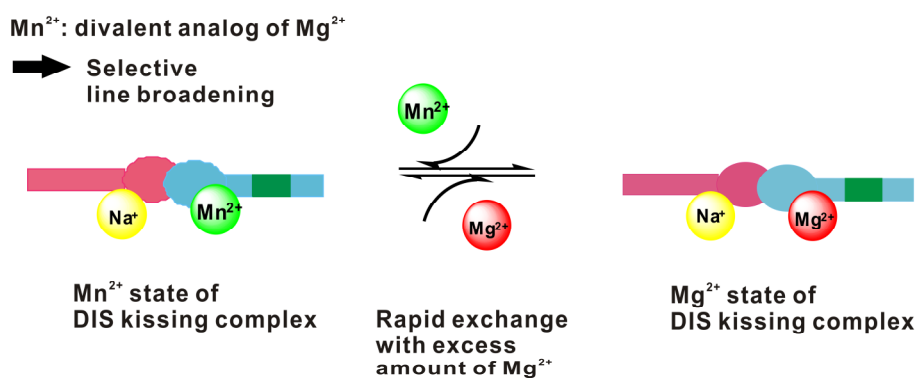


Fig 4.1. A model of Na<sup>+</sup>-Mg<sup>2+</sup>-Mn<sup>2+</sup>-Mg<sup>2+</sup> competition at the specific metal ion binding sites on the DIS stem-loop kissing complex, DIS23(GA)•DIS23(HxUC).

In the previously obtained X-ray crystal structures of homodimeric DIS extended duplex and kissing dimers, unpaired purine bases in the DIS loop junction of the DIS duplex and kissing dimers adopt a flipped out conformation soaked in  $\text{Mg}^{2+}$ . However, the NMR solution structures in the absence of  $\text{Mg}^{2+}$  have shown as a bulged-in conformation for the same nucleotides. By using chemical shift perturbation and paramagnetic relaxation enhancement methods in our NMR studies, the  $\text{Mg}^{2+}$  ions are found to stabilize the DIS loop junction *via* specific binding. It has already been found that the rate of structural NCp7 (nucleocapsid protein) maturation is 10-20 fold faster for the divalent metal ion stabilized DIS kissing dimer than the monovalent metal ion stabilized kissing structure.  $\text{Mg}^{2+}$  plays a unique functional role in structuring the RNA for interaction with NCp7 to promote structural isomerization.

### 4.3 Material and Methods

*RNA Preparation* — DIS 21 homodimer and both labeled and unlabeled DIS23(GA)•DIS23(HxUC) RNAs were transcribed from chemically synthesized DNA templates *in vitro* by using T7 RNA polymerase as described in Chapter 3.3. The RNA hairpin loop conformations are made by the heat and snap-cool method. For the hetero kissing complex, the same annealing procedures were followed before mixing unlabeled and labeled RNAs in a 1:1 ratio.

*NMR Spectroscopy* — NMR spectra were recorded on Bruker DRX 500 and DRX-600 NMR instruments. Spectra were processed using NMRPipe (104) and were assigned by Sparky (105). The experiments were described in Chapter 3.3.



## **4.4 Results and Discussion**

### **4.4.1 NMR Methods to Identify and Compare Sites Found for High-Affinity $Mg^{2+}$ Binding Sites on the DIS Extended Duplex and Kissing Dimer Structures**

#### **4.4.1.1 Chemical Shift Perturbation (CSP)**

Two complementary techniques to characterize divalent cation-binding sites in the DIS complexes have been used in this study: (1) chemical shift perturbation (CSP) and (2) paramagnetic relaxation enhancement (PRE) using  $Mn^{2+}$  ion as a probe (Fig 4.1). Secondary structures and a schematic diagram of the formation of both homo and heterokissing complex are shown in Fig 3.3 and Fig 3.4. Previous data collected on both the homodimeric mature duplex formed by the 21-mer DIS stem-loop and the heterodimeric kissing complex formed by DIS23(GA) and DIS23(HxUC) in the presence of  $Mg^{2+}$  display significant changes in chemical shift upon  $Mg^{2+}$  binding.  $Mg^{2+}$  not only coordinates the unpaired purine binding pockets but also stabilizes the loop-loop helix to make DIS complexes more stable. In the DIS21 homoduplex dimer, the 1D imino protons in the presence of  $Mg^{2+}$  showed that the apical stem G-C base pair appeared to be disrupted (Fig 3.5). In addition, the imino protons G9, G11, U5, and U10 experienced different coordination environments with  $Mg^{2+}$  addition, resulting in significantly different imino resonances. To limit the nonspecific aggregation, the optimized concentration of magnesium ion was added within the range of 0.1 – 5 mM in order to acquire good quality on NMR data. The non-hydrogen bonded imino protons of the junction guanine bases are not observed in 1D imino spectrum. The data of chemical shift perturbations of non-exchangeable protons from DIS21 duplex were collected

through 2D-TOCSY (H6/H5 correlations), 2D NOESY (H8, H6, H5, H2, and H1' resonances), 2D  $\underline{\text{HCCH}}$ -COSY/ $\underline{\text{HCCH}}$ -COSY (C1', C2', H1', and H2' resonances), and  $^{13}\text{C}$  CT-HSQC (ribose region) experiments to compare bound and unbound  $\text{Mg}^{2+}$  complexes.  $\text{Mg}^{2+}$  induced chemical shift perturbations suggest not only metal ion binding but also structural changes. In 2D NOESY spectrum, the effect of  $\text{Mg}^{2+}$  induced chemical shifts on overall conformation of the RNA appears to be selective. The titrated sample of the DIS21 duplex with  $\text{Mg}^{2+}$  results in shifting certain cross peaks while other cross peaks remains unchanged. One example of significant chemical shift of A15H8-H1' cross peak induced by  $\text{Mg}^{2+}$  was shown in Fig 4.2, while C12H6-H5 resonance didn't show significant shift upon addition of 2 mM  $\text{Mg}^{2+}$ .

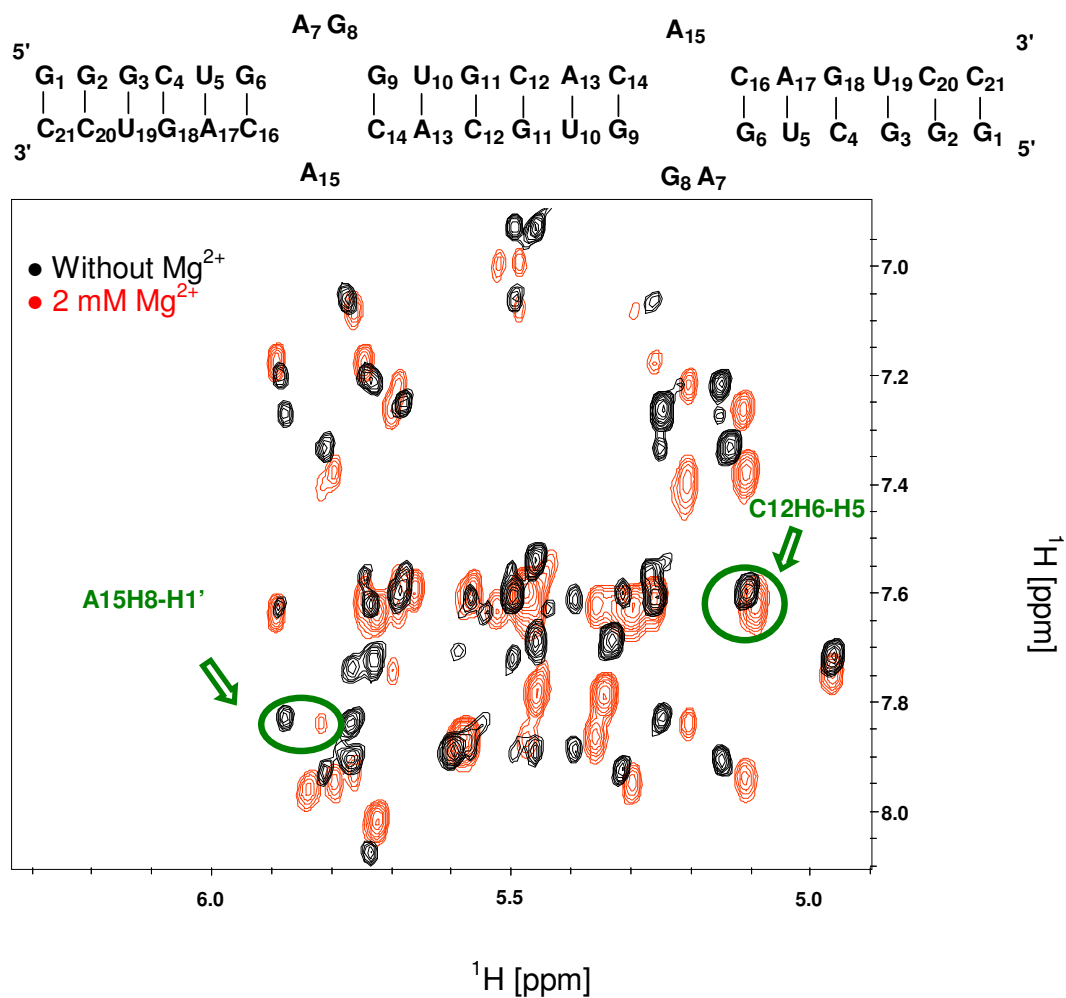
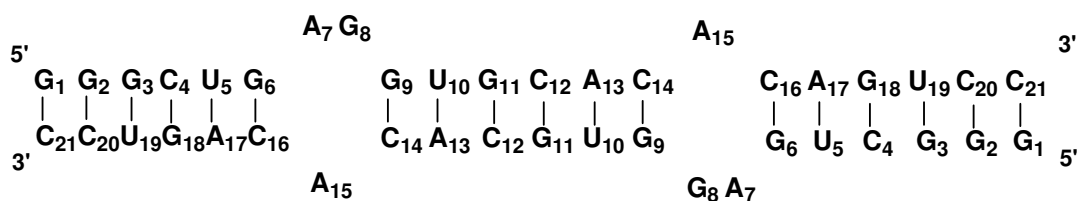


Fig 4.2.  $\text{Mg}^{2+}$  induced chemical shift perturbations associated with DIS21 duplex. Overlay of the non-exchangeable proton region of 2D NOESY spectra. Free  $\text{Mg}^{2+}$  bound duplex and 2 mM  $\text{Mg}^{2+}$  bound DIS21 duplex are labeled in black and red, respectively.

The significant chemical shift perturbations of non-exchangeable protons induced by  $\text{Mg}^{2+}$  with respect to each nucleobases are summarized in Fig 4.3 and Table 4.1. The average difference in chemical shifts of proton resonances resulted from  $\text{Mg}^{2+}$  perturbation is diminished due to the negative and positive values of chemical shift cancellation. A root mean square calculation is therefore used to obtain the magnitude of chemical-shift difference. The differences in chemical shifts of ribose and base proton resonances were analyzed separately for each nucleobase. H1' and H2' are taken into account for chemical-shift difference of ribose protons. H2, H5, H6, and H8 protons are utilized to compare the difference in chemical shifts of nucleobase protons. A significant chemical shift perturbation upon addition of 2 mM  $\text{Mg}^{2+}$  to the DIS 21 duplex showed that G6 cross peaks were missing beyond detection. This result was consistent with the absence of imino resonance of G6-C16 base pair in the presence of  $\text{Mg}^{2+}$ . G-U wobble base pair is well known as an excellent binding site for a  $\text{Mg}^{2+}$  in the major groove (118). Both root mean square of chemical shifts of ribose and base proton resonances were observed to shift significantly for G3-U19 wobble base pair. It suggests that GU mismatch is potential for creating metal ion binding site in the DIS21 homodimer helix (119). Significant chemical shift perturbation for base protons were also observed for the junction base A7 and stem-loop junction base C16. Significant CSP for base proton resonance was also identified for U5 located in the helix loop next to the stem-loop junction (Fig 4.4).



Root Mean Square (RMS) of  
Mg<sup>2+</sup> induced chemical shift  
perturbations (ppm)

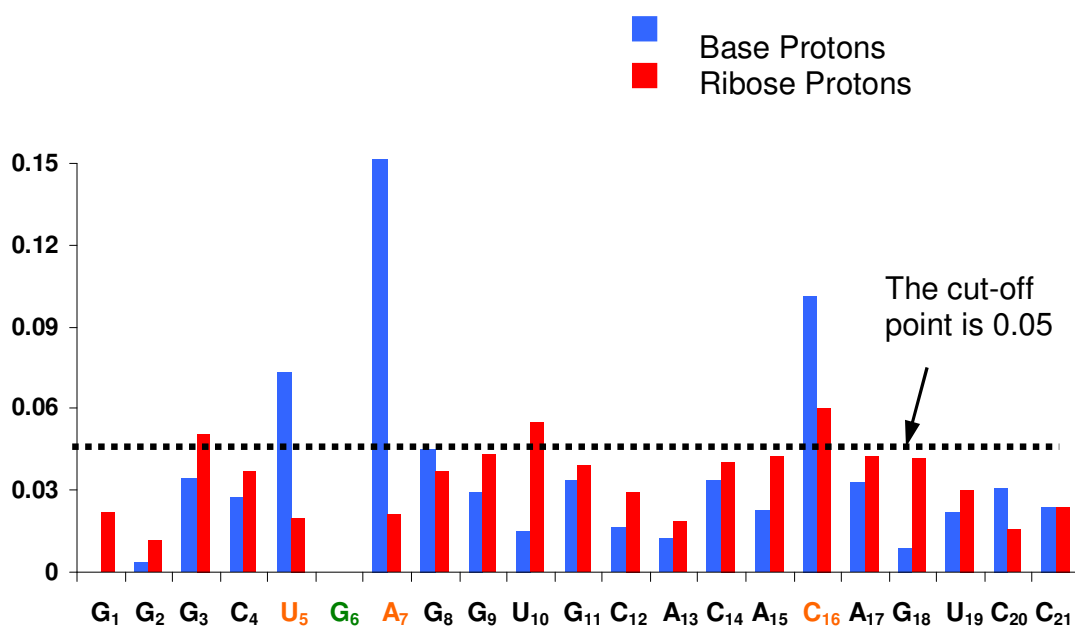


Fig 4.3. The summary of Mg<sup>2+</sup> induced chemical shift perturbations in DIS21 duplex with respect to each nucleobase. Comparisons of the root mean square difference in chemical shifts for base protons and ribose protons are highlighted in blue and red, respectively. The cross peaks of G6 were broadened beyond detection (green). The chemical shift changes of base proton resonances of U5, A7, and C16 were the most significant (highlighted in orange).

| DIS21 Duplex<br>(Base) | RMS of aromatic<br>proton resonances:<br>H2, H5, H6, and H8 | RMS of ribose proton<br>resonances: H1' and H2' |
|------------------------|---|---|
| <b>G1</b>              | —   | 0.022   |
| <b>G2</b>              | 0.003   | 0.011   |
| <b>G3</b>              | 0.035   | 0.051   |
| <b>C4</b>              | 0.027   | 0.037   |
| <b>U5</b>              | 0.074   | 0.020   |
| <b>G6</b>              | —   | —   |
| <b>A7</b>              | 0.157   | 0.021   |
| <b>G8</b>              | 0.045   | 0.037   |
| <b>G9</b>              | 0.029   | 0.043   |
| <b>U10</b>             | 0.015   | 0.055   |
| <b>G11</b>             | 0.033   | 0.039   |
| <b>C12</b>             | 0.016   | 0.029   |
| <b>A13</b>             | 0.012   | 0.018   |
| <b>C14</b>             | 0.033   | 0.040   |
| <b>A15</b>             | 0.023   | 0.042   |
| <b>C16</b>             | 0.101   | 0.060   |
| <b>A17</b>             | 0.033   | 0.042   |
| <b>G18</b>             | 0.009   | 0.042   |
| <b>U19</b>             | 0.022   | 0.030   |
| <b>C20</b>             | 0.031   | 0.016   |
| <b>C21</b>             | 0.024   | 0.024   |

Table 4.1. The summary of the RMS (root mean square) of difference in chemical shifts for base protons and ribose protons induced by  $\text{Mg}^{2+}$  in DIS21 duplex with respect to each nucleobase. “—” represents cross peaks that are not observed.

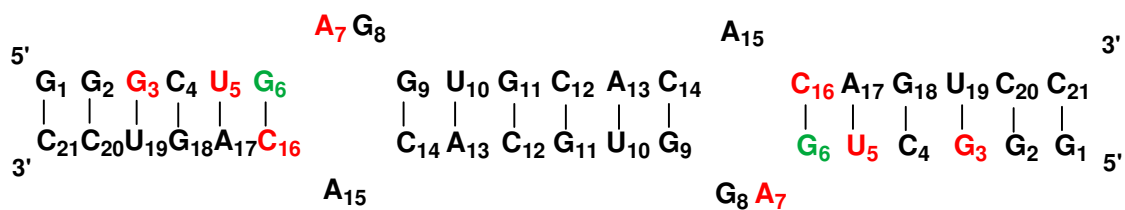


Fig 4.4. Schematic representation of bases which are perturbed upon  $\text{Mg}^{2+}$  binding in DIS21 duplex. The resonances of G6 highlighted in green were broadened beyond detection.  $\text{Mg}^{2+}$  binding causes dynamics with disruption G6 and C16 base pair. The base proton resonances of U5, A7, and C16 highlighted in red were perturbed more profoundly by  $\text{Mg}^{2+}$ . The GU wobble base pair is a well known  $\text{Mg}^{2+}$  binding site. Base and ribose proton resonances of G3 were perturbed less than bases on the DIS junction.

#### 4.4.1.2 Paramagnetic Relaxation Enhancement –NMR

*DIS21 homo-duplex dimer* — PRE NMR methods were used to probe for divalent metal binding sites in the DIS dimer structures. The paramagnetic ion,  $\text{Mn}^{2+}$ , was used to substitute for  $\text{Mg}^{2+}$  at the higher affinity divalent metal binding sites, leading to line broadening of resonances in close proximity. The enhanced relaxation rates which have a  $1/r^6$  dependence on the distance between the heteronucleus and the ion are induced by  $\text{Mn}^{2+}$  (120). Initially, DIS21 duplex is saturated with  $\text{Mg}^{2+}$  to occupy the non-specific and specific divalent metal ion binding sites. By performing  $\text{Mn}^{2+}$  titration experiments, the empirically determined optimal concentration of  $\text{Mn}^{2+}$  ion for PRE is obtained. The ideal rates which depend on metal concentration binding limit showed result in selective line broadening effects rather than non-specific binding of the paramagnetic ion. After the DIS21 duplex free form saturated with 2 mM  $\text{Mg}^{2+}$ , the end point of  $\text{Mn}^{2+}$  titration was determined by the selective broadening of the imino protons. The ratio of  $\text{Mn}^{2+}$  addition to  $\text{Mg}^{2+}$  is around 1: 37 for the DIS21 duplex. Ideally, the effect of paramagnetic line broadening is to compare the absolute values of either linewidth increases (119) or peak broadening in terms of volume changes induced by  $\text{Mn}^{2+}$ . Line width is characterized by the full width at half maximum of a resonance. Peak volumes are determined by using Sparky, a spectral display and analysis program, to integrate peak volumes in 2D NMR spectra. However, due to extensive overlap of signals caused by paramagnetic line broadening, it is difficult to fit the peak volume in Sparky by Gaussian or Lorentzian lineshape fitting (25). To calculate the absolute values concerning to the peak broadening, the peak height which is proportional to the intensity can be fit by Sparky quantitatively to each proton resonance from a 2D NOE experiment.



The peak height ratio is determined by comparing the peak height before and after addition of 54  $\mu\text{M}$   $\text{Mn}^{2+}$  to the DIS21 duplex in the presence of 2 mM  $\text{Mg}^{2+}$  in Fig 4.5. All the imino signals were broadened considerably. The U19 imino proton resonance in the G3-U19 wobble was more significantly broadened ( $\sim 31\%$ ) than the G3 ( $\sim 70\%$ ). Selective peak broadening with  $\text{Mn}^{2+}$  addition is more profound for the U imino proton than for the G imino proton in GU wobble base pairs (121). The major groove side of a GU wobble pair is the most common divalent cation binding site (Fig 4.6).

The paramagnetic line broadening of the non-exchangeable resonances provides further investigation for the bases in the junction. Fig 4.7 presents examples of line broadening due to  $\text{Mn}^{2+}$ . In the presence of  $\text{Mn}^{2+}$ , the U10 H6-H5 resonance was broadened beyond detection, whereas a significant line broadening of the G11 H8-H1' cross peak was not found. In this regard, the selective line broadening induced by  $\text{Mn}^{2+}$  suggests that  $\text{Mn}^{2+}$  can displace  $\text{Mg}^{2+}$  from the higher affinity binding sites of the DIS21 duplex. 2D NOE in  $\text{H}_2\text{O}$ , 2D TOCSY, and 2D NOESY in  $\text{D}_2\text{O}$  were performed to identify the specific divalent ion binding sites in the presence of  $\text{Mn}^{2+}$  at the atomic level.

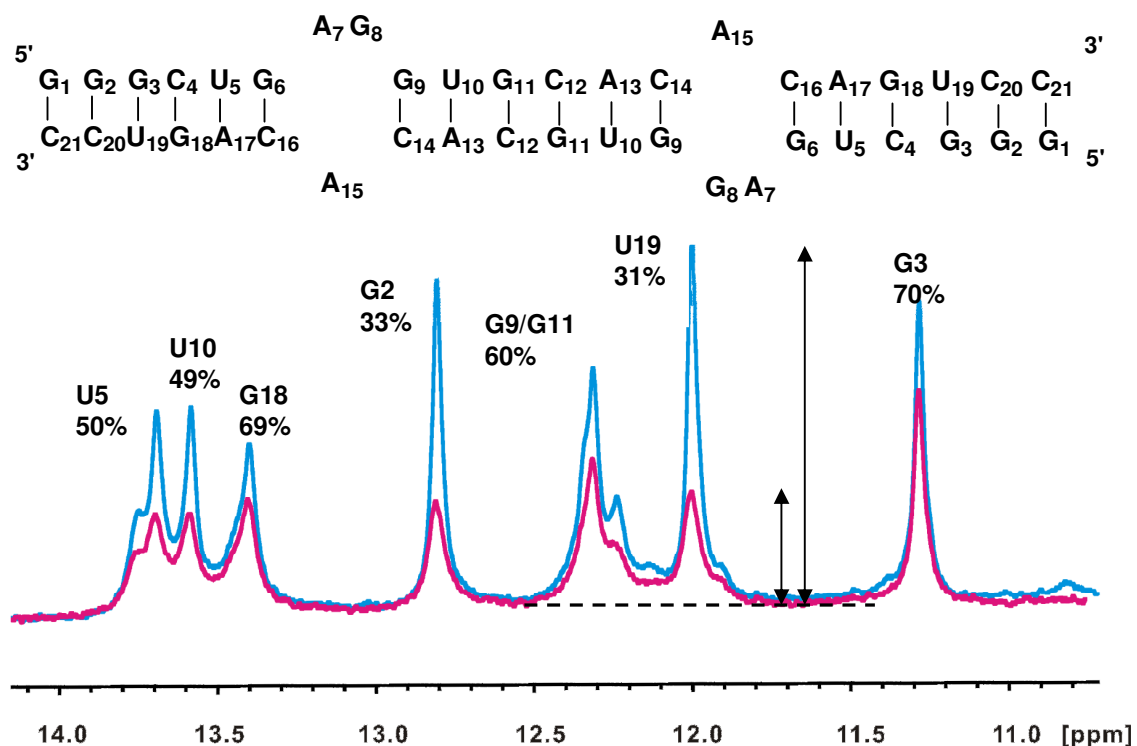
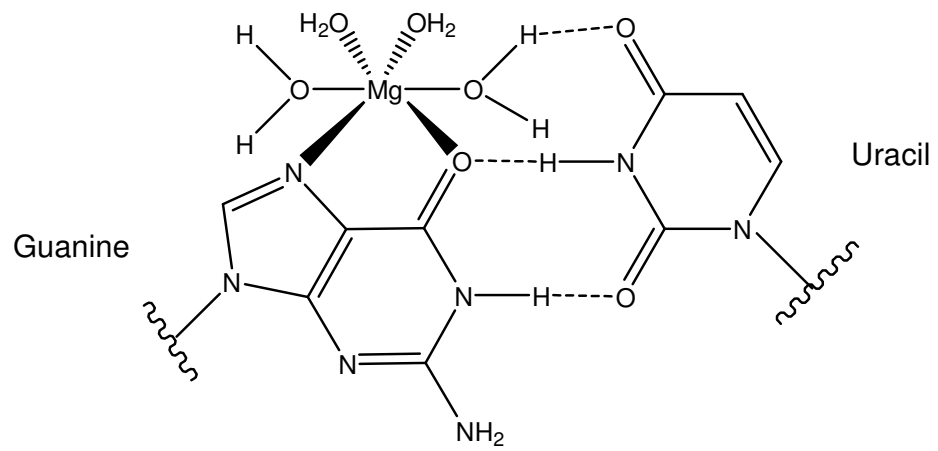


Fig 4.5. Imino proton region of the  $^1\text{H}$ -NMR spectra of the DIS21 duplex without  $\text{Mn}^{2+}$  (blue) and with  $54\ \mu\text{M}\ \text{Mn}^{2+}$  (120) DIS21 duplex ( $2\ \text{mM}\ \text{Mg}^{2+}$ , duplex concentration  $275\ \mu\text{M}$ , pH 6.5, 298K). The peak height ratio is given by comparing two peak heights before and after addition of  $\text{Mn}^{2+}$  as shown with double arrows and the fit baseline.

major groove



minor groove

Fig 4.6. GU wobble base pair is commonly observed in RNA helices.  $Mg^{2+}$  cations bind to the major groove at the GU pairs (122).

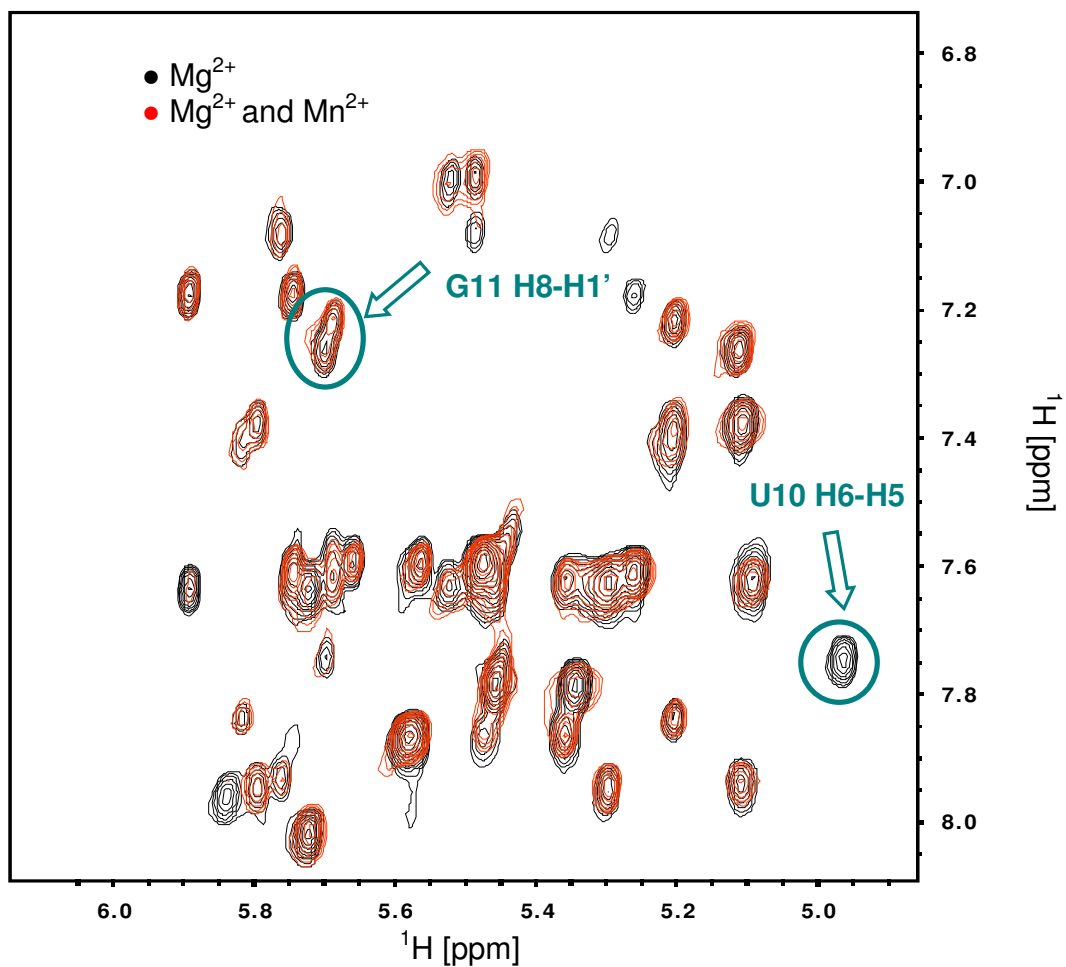


Fig 4.7. Overlay of the non-exchangeable proton resonances of a 2D NOESY spectrum of the DIS21 duplex with 2 mM  $\text{Mg}^{2+}$  (black) and in the presence of 2 mM  $\text{Mg}^{2+}$  and 54  $\mu\text{M}$   $\text{Mn}^{2+}$  (120). The examples of the paramagnetic line broadening effect by  $\text{Mn}^{2+}$  are labeled with open green circles. U10 resonance was broadened in the presence of  $\text{Mn}^{2+}$ .

In order to obtain the more accurate peak height extracting from 2D NOESY spectra, the fit height is determined after integrating by fitting the peak positioned at the local maximum from Sparky. Each cross peak referring to each proton resonance was fit to obtain the peak height in the presence of 2 mM  $\text{Mg}^{2+}$  comparing to the spectrum in the presence of both 2 mM  $\text{Mg}^{2+}$  and 54  $\mu\text{M}$   $\text{Mn}^{2+}$ . The peak height ratio is as follows,

$$\text{Peak height ratio} = \frac{\text{Peak height (Mn}^{2+} + \text{Mg}^{2+})}{\text{Peak height (Mg}^{2+})} \quad (4.1)$$

The G6, A7, G8, G9, and U10 cross peaks broadened beyond detection upon addition of  $\text{Mn}^{2+}$  (Fig 4.8). These cross peaks were substantially weak compared to other nucleobases.  $\text{Mn}^{2+}$  ions compete selectively to the above binding sites.

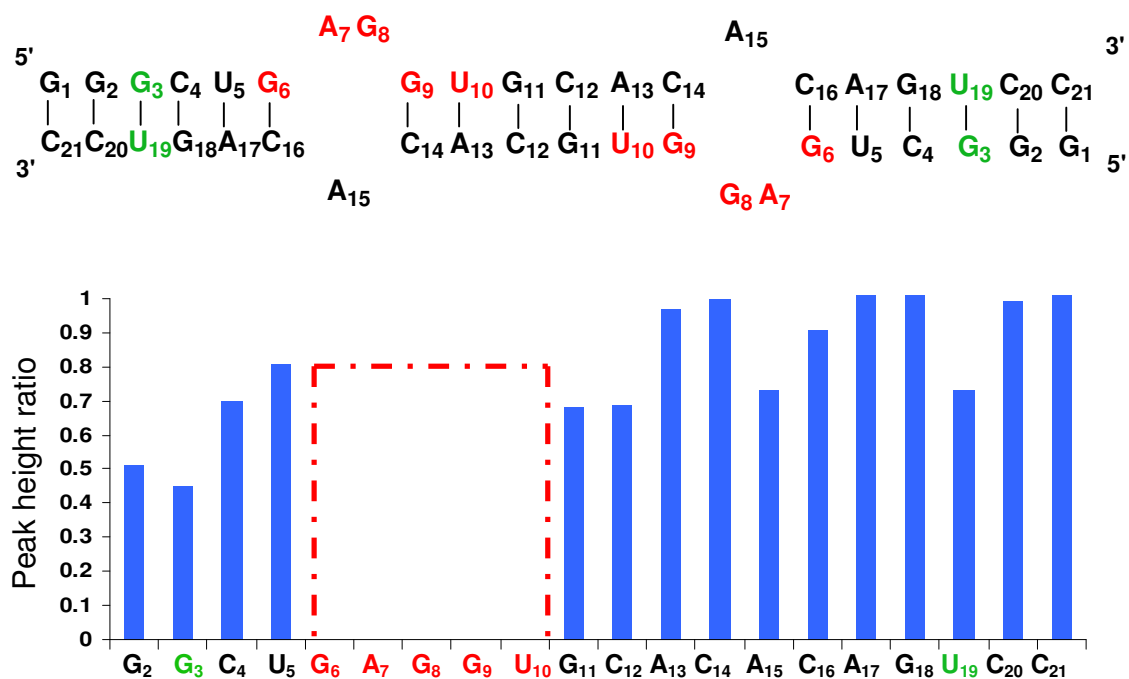


Fig 4.8. The summary of peak height ratio for each nucleobase considering each proton resonance from 2D NOESY spectra. G6, A7, G8, G9, and U10 (highlighted in red) were broadened beyond detection by  $\text{Mn}^{2+}$ . G3-U19 (highlighted in green) was found to be the  $\text{Mg}^{2+}$  binding site from chemical shift perturbation experiment and 1D imino  $\text{Mn}^{2+}$  titration.

The results of paramagnetic relaxation measurements have great potential for the precise localization of metal binding sites at the atomic level. The displayed contour level from Sparky, a graphical NMR assignment and integration program, was adjusted to the lower order to examine individual atom localized on the base from G6 to U10 as well as two junction bases A15 and C16 which were observed with more significant chemical shift perturbations induced by  $Mg^{2+}$ . Peak height ratios are summarized in Table 4.2 and Fig 4.9. Although, the displayed contour level was adjusted to the lower level, G6 and A7 cross peaks were still undetectable due to selective paramagnetic line broadening. G8 (H1'), G9 (H8 and H1'), U10 (H6 and H5), and A15 (H2) have more significant broadening effects in the close proximity to the  $Mn^{2+}$  ion with smaller peak height ratio. Two NMR methods, CSP and PRE, suggest that a  $Mg^{2+}$  binding site is localized around G6 and A7 nucleotides since these display the most significant perturbations that might be expected due to direct binding of a divalent cation.  $Mg^{2+}$  binding disrupts G6-C6 base pairing and causes dynamics in the stem helices. Here, we use the solution and X-ray structures of the DIS duplex as templates for mapping  $Mg^{2+}$  ion binding sites based on the CSP and PRE results at the atomic level. The solution structure containing the loop-stem region was determined using NMR methods. It also indicates that the junction purine A7 adopts a stacked in geometry towards the stem loop helix and metal binding pocket in the solution structure (Fig 4.10 (A); solution structure is in the absence of  $Mg^{2+}$ ). Eight  $Mg^{2+}$  sites have been identified in the DIS duplex crystal structures which were grown in 5 mM magnesium (Fig 4.10 (B)). Four magnesium cations symmetrically localize around G6, G8, G9, and A15. The PRE results showed that  $Mg^{2+}$  ions are bound around these residues with more significant line broadening effects. One magnesium

binding site localized around U10 is consistent with the PRE data. Interestingly, one  $\text{Mg}^{2+}$  binding site in the crystal structure is around the purine junction where A7 was identified in a bulged out position. The flanking purine, A7, was found to be close to  $\text{Mg}^{2+}$  ions based on the significant CSP and selective PRE results which suggest exhibiting a bulge-in conformation.



| <b>H (Mn<sup>2+</sup>+Mg<sup>2+</sup>)/H (Mg<sup>2+</sup>)</b> | H1'  | H2   | H5   | H6   | H8   |
|--|------|------|------|------|------|
| G6H8-H1'   | —    |      |      |      | —    |
| A7H8-G6H1'   | —    |      |      |      | —    |
| A7H8-H1'   | —    |      |      |      | —    |
| A7H2-G8H1'   | —    | —    |      |      |      |
| G8H8-H1'   | 1.10 |      |      |      | 0.71 |
| G9H8-H1'   | 0.65 |      |      |      | 0.70 |
| U10H6-G9H1'  | 0.88 |      |      | 0.80 |      |
| U10H6-H5   |      |      | 0.40 | 0.15 |      |
| U10H6-H1'  | 0.98 |      |      | 0.85 |      |
| A15H8-C14H1'   | 0.83 |      |      |      | 1.15 |
| A15H8-H1'  | 0.81 |      |      |      | 1.14 |
| A15H2-C16H1'   | 0.67 | 0.50 |      |      |      |
| A15H2-G9H1'  | 0.58 | 0.91 |      |      |      |
| C16H6-A15H1'   | 0.89 |      |      | 2.00 |      |
| C16H6-H5   |      |      | 1.40 | 0.72 |      |
| C16H6-H1'  | 0.81 |      |      | 0.72 |      |

Table 4.2. The summary of peak height ratios. “H” represents the peak height of the resonances. The peaks broadened beyond detection by Mn<sup>2+</sup> are denoted with “—”. The estimate error based on S/N is (~ 0.09). Three random positions in the NMR spectra in Sparky with where there are no resonance signals were chosen. Take the average of these three noise positions for the error determination associated with the average peak height of resonance signals.

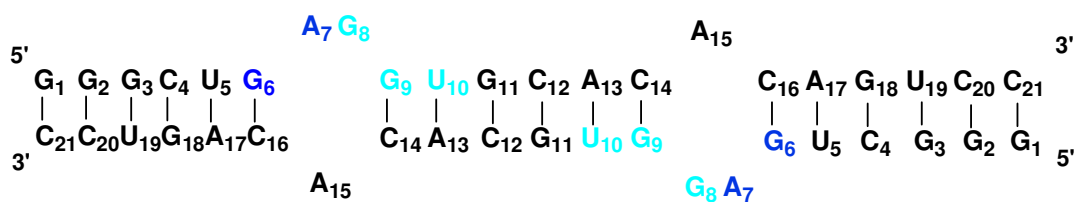
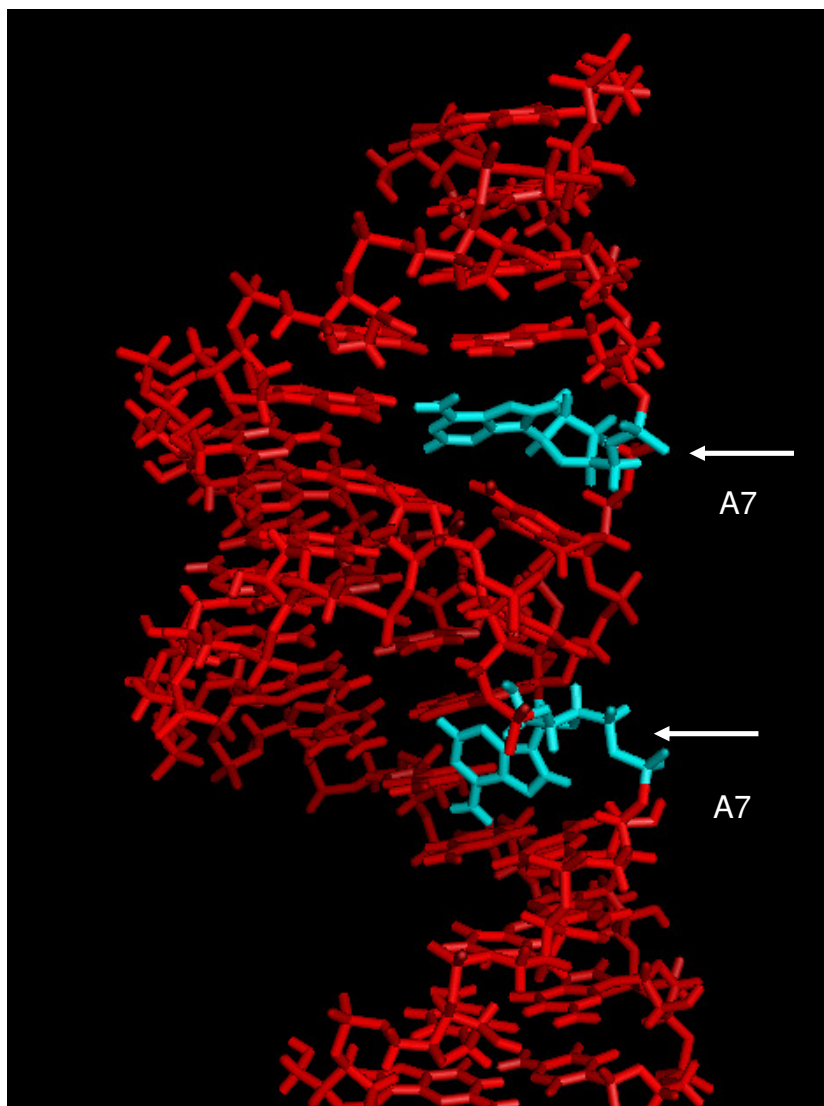
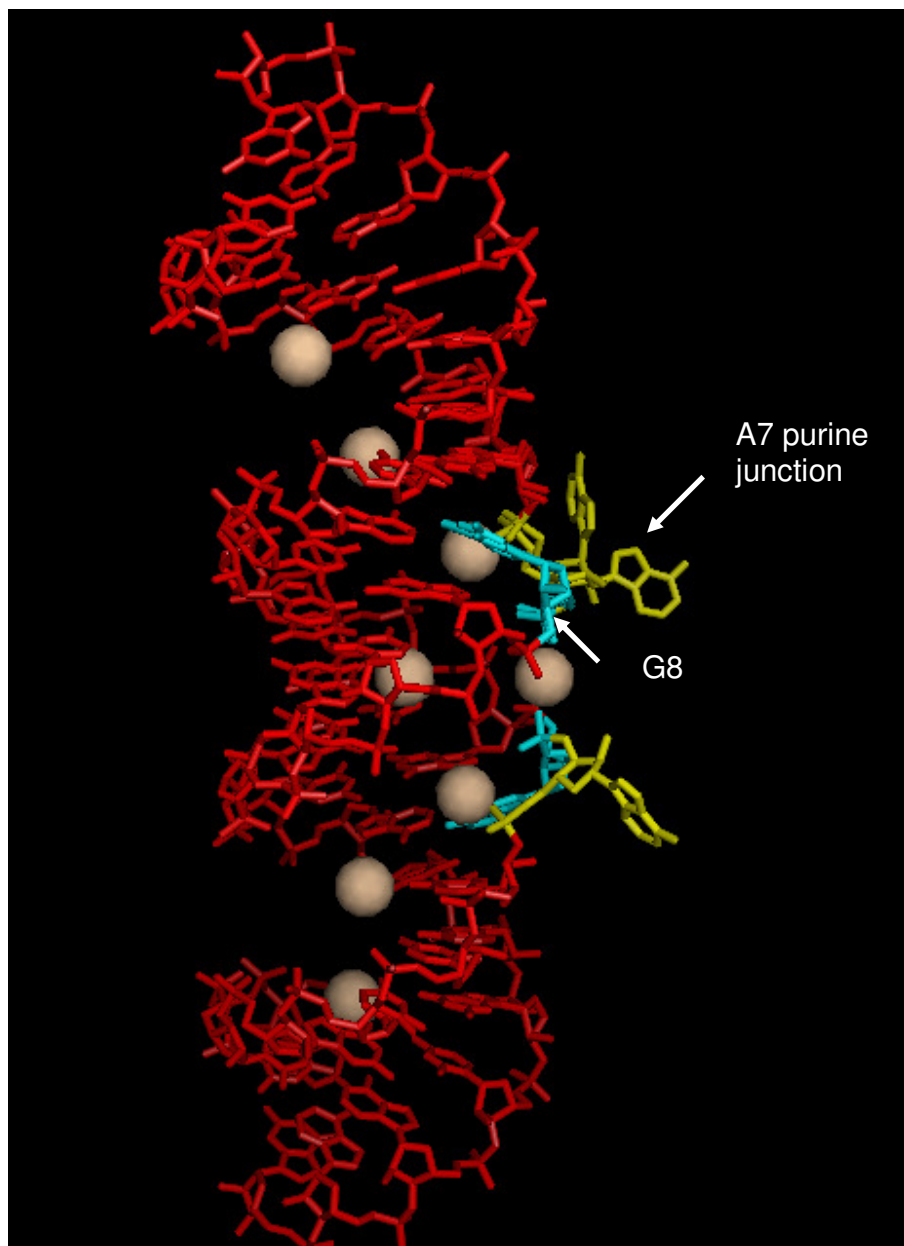


Fig 4.9. Two levels of gradation of line broadening effects were shown here. The cross peaks of G6 and A7 were completed broadening away in the presence of  $Mn^{2+}$  displayed at the lower order of contour level from Sparky (highlighted in dark blue). G8, G9, and U10 experienced lesser paramagnetic line broadening compared to G6 and A7 (light blue).

A.



B.



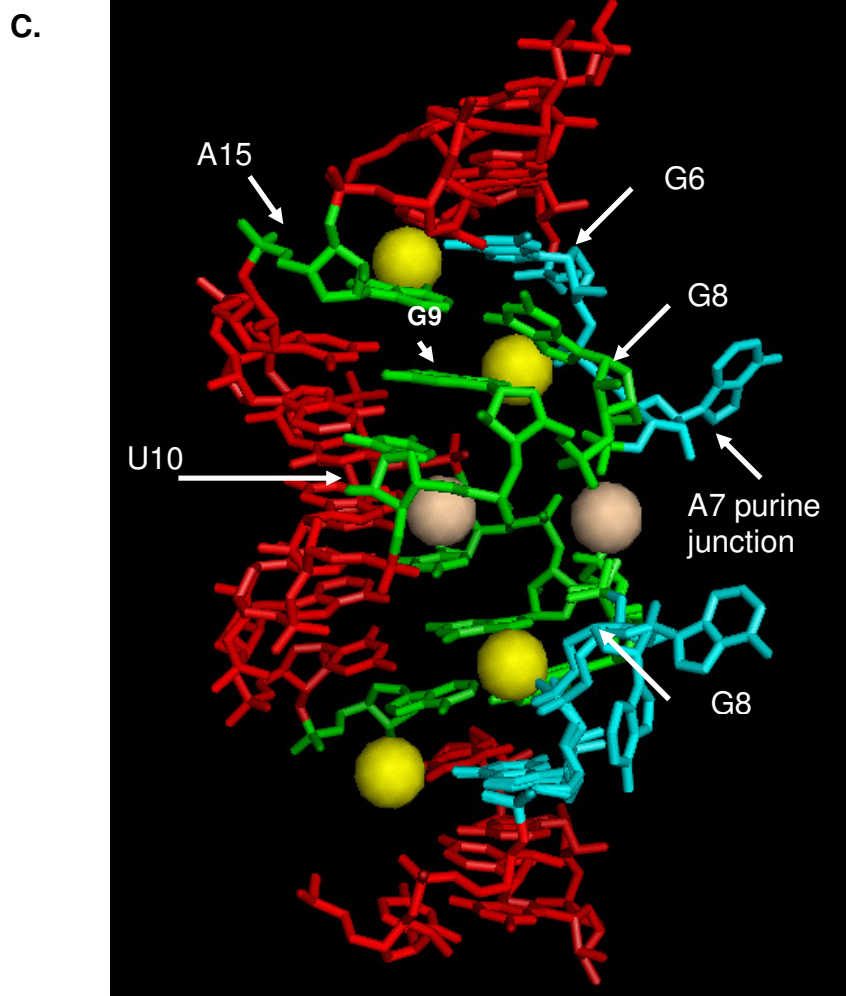


Fig 4.10. (A) NMR solution structure of DIS extended duplex in the absence of  $\text{Mg}^{2+}$ . The conserved purine A7 at the 5' side was suggested to be a bulged-in orientation (highlighted in cyan; PDB code: 2D19). (B) X-ray structure of DIS extended duplex with the purine bulged out residues (A7: a bulged out purine highlighted in yellow; G8: purine base at the junction highlighted in cyan). Eight  $\text{Mg}^{2+}$  ions were found in the structure (PDB code: 1Y99). (C) Divalent metal cation-binding sites in solution mapped on the X-ray structure of DIS junction. G6 and A7 (highlighted in cyan) were found to be close to the  $\text{Mg}^{2+}$  binding sites due to paramagnetic line broadening effect in the solution structure. The  $\text{Mg}^{2+}$  binding sites represented as yellow spheres are most consistent with the PRE data. The  $\text{Mg}^{2+}$  ions represented as yellow spheres were bound around G8 and G9 (highlighted in green). However, A7 adopts a bulged geometry which is not consistent with CSP and PRE data.

*<sup>13</sup>C <sup>15</sup>N-DIS23(GA)•DIS23(HxUC) hetero-kissing complex* —Both Na<sup>+</sup> and Mg<sup>2+</sup> can stabilize DIS complexes. Na<sup>+</sup> at high concentrations can saturate nonspecific divalent metal binding sites. The hetero-kissing complex was initially saturated with 200 mM Na<sup>+</sup>. 200 μM Mg<sup>2+</sup> (Na<sup>+</sup>:Mg<sup>2+</sup>=1000:1) was then added to occupy the high affinity divalent binding sites. Then, 10 μM Mn<sup>2+</sup> was added in the complex to substitute for Mg<sup>2+</sup> at the specific binding sites. Mg<sup>2+</sup> and Mn<sup>2+</sup> bind with about the same affinity. The paramagnetic relaxation effect can be identified. Initially, the experiment was performed in the addition of 200 mM Na<sup>+</sup> followed by small amount of Mn<sup>2+</sup>. However, there was almost no line broadening effect observed. It suggests that Na<sup>+</sup> could not compete the divalent sites even at 200 mM concentration (data not shown). In addition, there are no significant Mg<sup>2+</sup> induced chemical shift perturbations in the hetero-kissing complex compared to homodimers duplex (Fig 4.11).

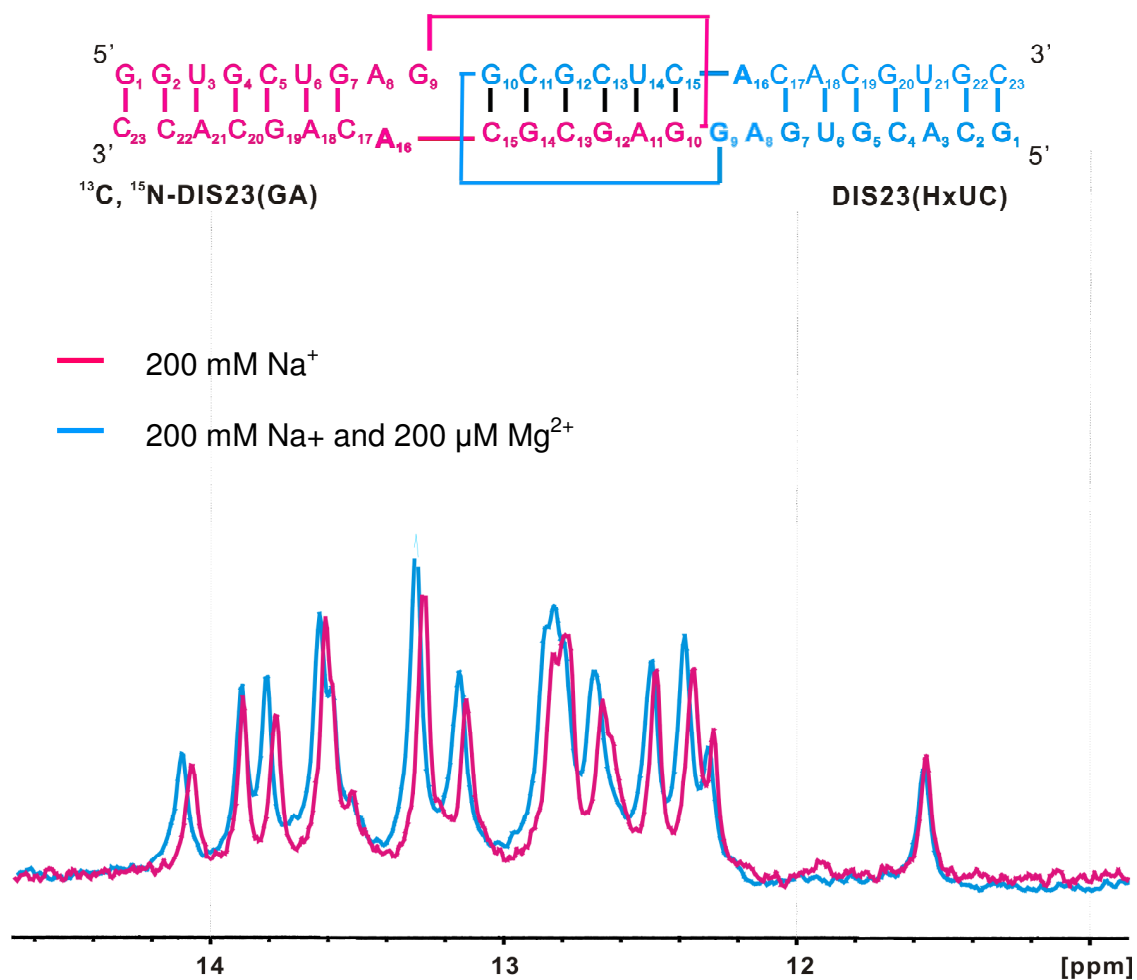


Fig 4.11. 1D imino proton  $Mg^{2+}$  induced chemical shift perturbation of the imino proton resonances of  $^{13}C^{15}N$ -DIS23(GA)•DIS23(HxUC) kissing complex (298K and pH 7.5). Overlay of the imino region of 1D spectra in the presence of 200 mM  $Na^+$  (pink) and both 200 mM  $Na^+$  and 200  $\mu M$   $Mg^{2+}$ .

To identify the specific  $\text{Mg}^{2+}$  binding sites on the DIS stem-loop kissing complex, the paramagnetic  $\text{Mn}^{2+}$  as an analogue was used to displace  $\text{Mg}^{2+}$ . The ratio of species in the mixture is DIS kissing complex: $\text{Na}^+:\text{Mg}^{2+}:\text{Mn}^{2+}= 25:20000:20:1$ . The DIS stem-loop kissing complex is fully saturated with  $\text{Na}^+$ . The specific divalent metal binding sites were occupied by  $\text{Mg}^{2+}$  with higher affinity. Small amount of  $\text{Mn}^{2+}$  was added to swapped  $\text{Mg}^{2+}$  ions on the specific binding sites on the DIS stem-loop kissing complex. The summary of peak height ratio respecting to each nucleobase on the DIS23(GA) side in the kissing complex was compared before and after addition of  $\text{Mn}^{2+}$  (Fig 4.12). The cut off point was 0.5 suggesting that at least 50 % peaks were broadened by the presence of  $\text{Mn}^{2+}$ . Seven resonances showed selective paramagnetic line broadening induced by the presence of  $\text{Mn}^{2+}$  including G7a, A8a, G9a, G10a, C13a, A16a, and C17a (Fig 4.13, Table 4.3). Specific divalent metal binding identification with PRE measurements suggests bases A8a, G9a, and A16a have the greater effect of divalent metal interaction. Interestingly, in the protonated state of the kissing complex, the observed exchange broadening of resonances were involved in A8a, G9a, and A16a. G10a was also found broadening in the  $^1\text{H}^{13}\text{C}$ -HSQC experiment due to protonation of residue A8a resulting in dynamics in the DIS kissing complex purine junction (71). Fig 4.14 presents X-ray and NMR solution structures displaying two different conformations (14). The crystal structure of the kissing loop complex was obtained by fast-cooling method and then crystallization at 37 °C without nucleocapsid protein. In the X-ray crystal structure, the unpaired purines, A8 and G9, at the 5' side in DIS-A form are in a bulged-out conformation (12). Two bases are stacking through the  $\pi$ - $\pi$  interaction between two aromatic rings (Fig 4.14 (A)). Three  $\text{Mg}^{2+}$  ions binding sites are around residues G7,



G10, and A16. Our NMR data also show the same divalent metal binding sites. NMR structure adopting a bulge-in conformation in the absence of  $\text{Mg}^{2+}$  ions (14) was found to be consistent with our results.  $\text{Mg}^{2+}$  ions were localized in the purine junction around A8, G9, and A16 based on PRE data which are consistent with the solution structure where these purine bases stack in the helix (Fig 4.14 (B)). The solution structure obtained by Baba et al. (2006) was in the absence of  $\text{Mg}^{2+}$ .

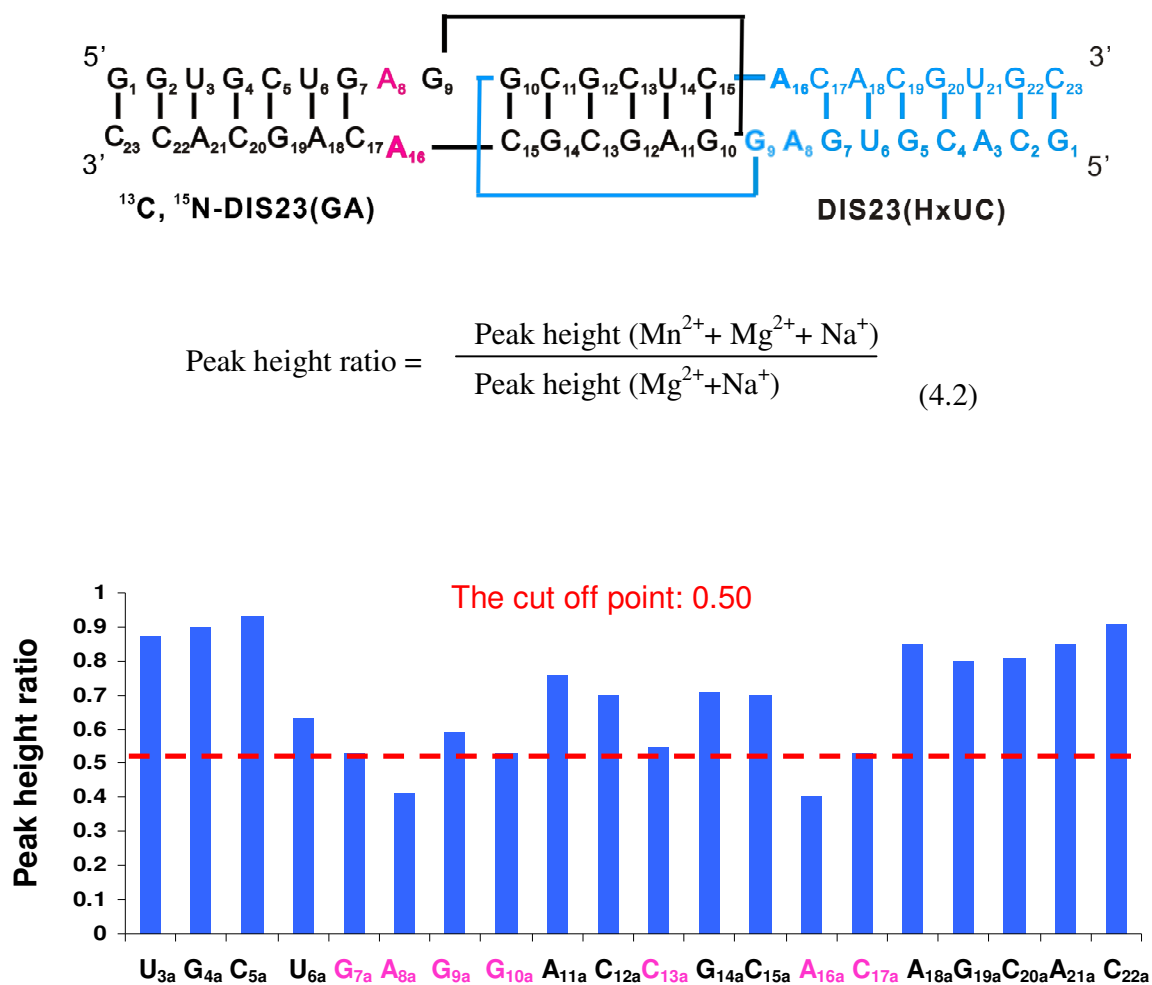


Fig 4.12. The summary of peak height ratio respects to each nucleobase considering the average of aromatic and ribose proton resonances (H2, H5, H6, H8, H1') from 2D NOESY spectra on the DIS23(GA) side.

| <b>H(Na<sup>+</sup>+Mg<sup>2+</sup>+Mn<sup>2+</sup>)/H(Na<sup>+</sup>+Mg<sup>2+</sup>)</b> | <b>H1'</b> | <b>H2</b> | <b>H5</b> | <b>H6</b> | <b>H8</b> |
|--|------------|-----------|-----------|-----------|-----------|
| G7aH8-U6aH1'   | 1.11       |           |           |           | 0.81      |
| G7aH8-H1'  | 0.59       |           |           |           | 0.7       |
| A8aH8-G7aH1'   | —          |           |           |           | —         |
| A8aH8-H1'  | 0.78       |           |           |           | 0.88      |
| G9aH8-A8aH1'   | —          |           |           |           | —         |
| G9aH8-H1'  | 0.88       |           |           |           | 0.79      |
| G10aH8-G9aH1'  | 0.69       |           |           |           | 0.76      |
| G10aH8-H1'   | —          |           |           |           | —         |
| C13aH6-G12aH1'   | —          |           |           | —         |           |
| C13aH6-H5  |            |           | 0.69      | 0.62      |           |
| C13aH6-H1'   | 0.9        |           |           | 0.49      |           |
| A16aH8-H1'   | —          |           |           |           | —         |
| A16aH2-C17aH1'   | 0.79       | 0.79      |           |           |           |
| C17aH6-A16aH1'   | 0.75       |           |           | 0.63      |           |
| C17aH6-H5  |            |           | 0.93      | 1.48      |           |
| C17aH6-H1'   |            |           |           | —         | —         |

Table 4.3. The summary of peak height ratio considering each broadening effect of the individual atom localized on the DIS junction of DIS23(GA) side in the kissing complex. “H” represents the peak height of the resonances. The missing peaks beyond detection induced by the presence of Mn<sup>2+</sup> are denoted with “—”.

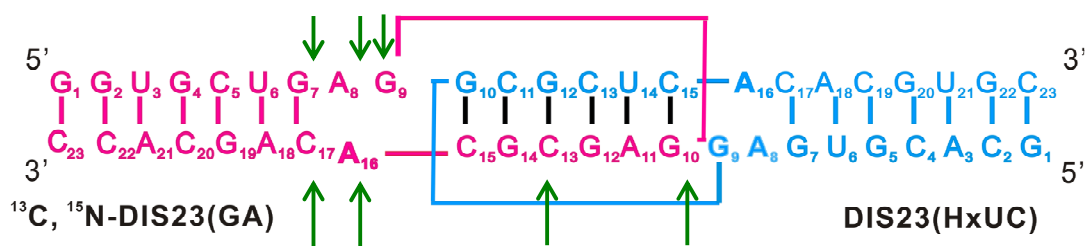
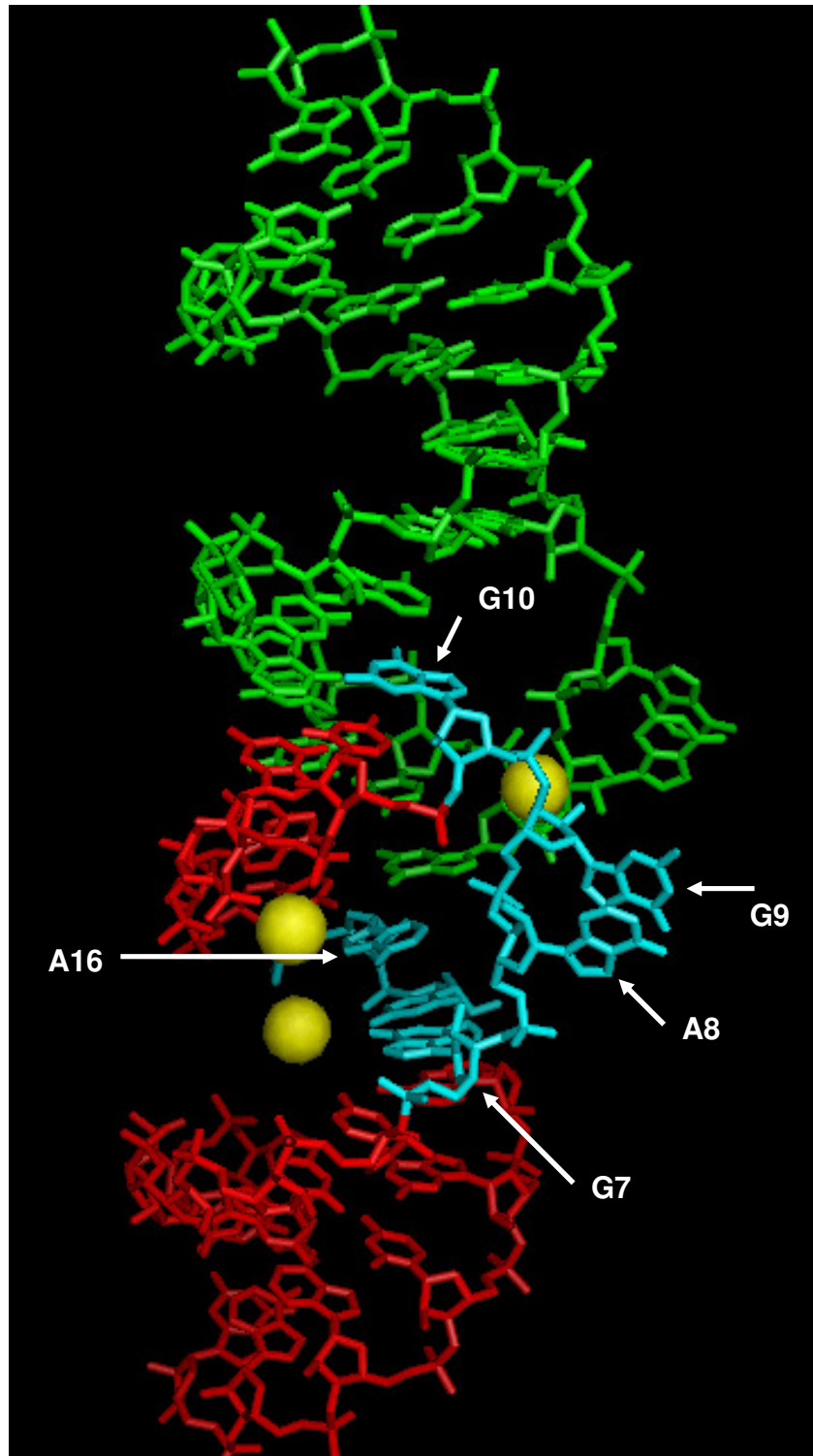


Fig 4.13. The divalent metal binding sites including G7a, A7a, G9a, G10a, C13a, A16a, and C17a on the DIS23(GA) side in the kissing complex, indicating greater line broadening in the presence of  $\text{Mn}^{2+}$ .

A.



B.

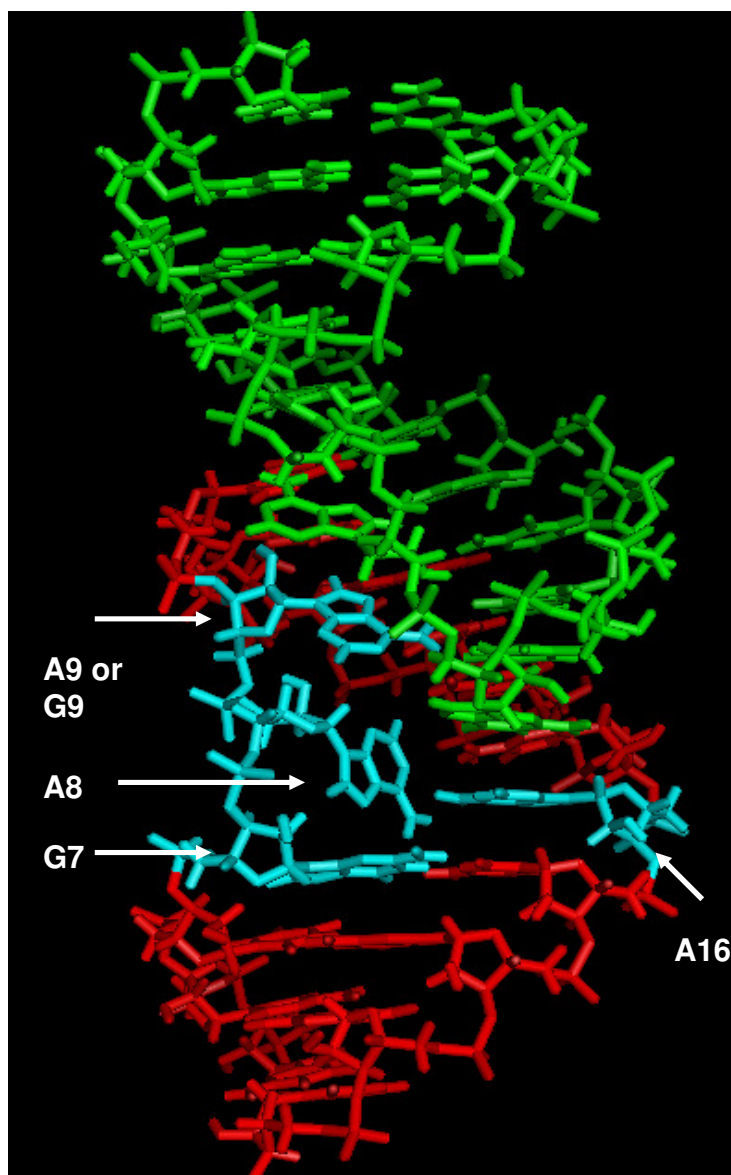


Fig 4.14. (A) X-ray structure of DIS kissing dimer structure with a bulged out purine (A8 and G9) conformation (12). Three  $\text{Mg}^{2+}$  ions are represented as yellow spheres. The NMR data characterized from PRE to identify the  $\text{Mg}^{2+}$  binding sites around or formed by bases G7, G10, and A16 (highlighted in cyan) are consistent with crystal structures. (B) Solution RNA structure model of the DIS kissing loop dimer with a bulge in conformation (without metal ions). Purine bases, A8 and G9/A9, are stacked towards the stem helix (PDB code: 1XPF). Our PRE NMR data showed that A8 and G9 (highlighted in cyan) were found to be close to the  $\text{Mg}^{2+}$  binding sites adopting the same geometry with the NMR solution structure.

## 4.5 Conclusion

The divalent metal ion binding sites in solution were identified initially by  $\text{Mg}^{2+}$  induced chemical shift perturbations which suggest conformational changes in the DIS structures as well as direct binding effect. Paramagnetic line broadening induced by  $\text{Mn}^{2+}$  leads to the identification of the specific binding sites. In DIS21 homodimer duplex, G3, U5, G6, A7, and C16 are perturbed upon addition of  $\text{Mg}^{2+}$ , suggesting in long range conformational change and metal binding. The specific divalent binding sites were identified by  $\text{Mn}^{2+}$  resulting in line broadening on G6, A7, G8, G9, and U10. These resonances were missing beyond detection upon  $\text{Mg}^{2+}$  binding. The specific divalent binding sites on the DIS23(GA) side in the hetero DIS stem-loop kissing complex,  $^{13}\text{C}^{15}\text{N}$ -DIS23(GA)•DIS23(HxUC), G7a, A8a, G9a, and G10a from 5' end of the RNA; C13a in the loop; and A16a and C17a from 3' end of the complex were found selective line broadening induced by the presence of  $\text{Mn}^{2+}$ .  $\text{Mg}^{2+}$  ions specifically localize in this region to stabilize these junction bases. DIS 21 duplex and hetero kissing complex show that purine bases are stacked in the stem helix resulting from the direct  $\text{Mg}^{2+}$  binding.

## Chapter 5 Concluding Remarks

Understanding the mechanism of dimerization of HIV-1 RNA is essential since it is involved in various critical events in the HIV life cycle. Each HIV virus carries two identical copies of the single-stranded (+) RNA. The advantage of the two copies is for a higher rate of recombination during infection and consequently an increase in genetic diversity of HIV. Dimerization is initiated through loop-loop kissing interaction between two genomic RNA strands associated by six self-complementary base pairs on the dimerization initiation site (DIS) within the 5' end of retroviral genomes. The wild type of DIS is a highly conserved hairpin containing a bulge and an apical loop consisting of nine bases with six palindromic bases. Mutation in the DIS results in a reduction of the replication rate, the infectivity of the HIV-1 (26, 123), and the level of genomic RNA packaging (7). Additionally, the deletion of the stem-loop structure reduces the infectivity of the HIV-1 up to 1000 fold (124). Therefore, the kissing-loop complex has been considered to be a potential target for anti-HIV drugs. For example, aminoglycosides, widely used antibiotics, block mRNA and tRNA translocation and inhibit ribosome recycling (125). The unpaired purine residues of the loop are essential to the dimerization process especially for divalent metal ion binding.

The dimeric nature of the RNA genome also plays a structural role in regulating specific stages of replication. Dimerization is required for proper RNA packaging with the assistance of viral nucleocapsid protein (NCp7), which binds to SL1 (stem-loop, DIS), SL2 (stem-loop, SD: splice donor), and SL3 of the  $\psi$  signal (126). The highly basic NCp7 protein containing two CX<sub>2</sub>CX<sub>4</sub>HX<sub>4</sub>C-type zinc fingers activates refolding of DIS from metastable kissing complex to thermodynamically stable extended duplex.



NCp7 protein plays a critical role in virus assembly and infection by chaperoning RNA dimerization and packaging. NCp7 accelerates the rate of strand exchange between the two hairpin stem helices with a 2:1 stoichiometry for NCp7 binding to the DIS kissing dimer (38).

2-Aminopurine, a fluorescent base analog, has been used to study the mechanism of nucleocapsid protein catalyzed structural isomerization of the DIS in our group. The 2-aminopurines were inserted either on the DIS stem or the loop to unambiguously distinguish their different fluorescent properties corresponding to mature duplex and kissing dimer, respectively. The identification of the activity of nucleocapsid protein was characterized by using 2-AP labeled sequence on the stem of DIS hairpin associated with fluorescence quenching when the DIS mature duplex dimer is formed. In the addition of NCp7, the rate of isomerization from the  $Mg^{2+}$  stabilized DIS kissing dimer to the extended duplex is 10-20 fold faster than the  $Na^+$  stabilized metastable DIS kissing dimer. The observed mechanism using the stem-inserted 2-AP fluorescent probe proves that the isomerization is involved in melting of intramolecular stem base pairs and reannealing of intermolecular base pairing to form a duplex. Moreover, the NCp7 catalyzed maturation rates are found to be faster when A272 is protonated at pH 6.0 compared to pH 7.2. The flanking purines at the DIS junction are dynamic in solution having fluctuational motility. It leads to base pair opening at the junction stimulated by interaction with the NCp7 protein to unwind two complementary strands within the stem in the kissing complex at low  $Mg^{2+}$  concentration.

To better understand the role of metal ions in HIV-1 genomic RNA dimerization and maturation facilitated by the NCp7 chaperone protein, 2-AP labeled fluorescence and

high resolution NMR experiments were carried out to examine the local transition of DIS kissing dimer at the purine junction. Metal ions are essential for the folding of RNA into a stable and compact conformation when the dimerization and maturation take place. The first stage is the formation of RNA secondary structure referring to the DIS hairpins stimulated by the presence of monovalent and divalent cations. The second stage is the formation of RNA tertiary structure including DIS kissing and extended duplex dimers. The RNA molecules in the folded conformation are readily stabilized by interacting with  $Mg^{2+}$  based on our fluorescence and NMR results. Fluorescence melting curve analysis corresponding to a 2-AP probe inserted into the DIS loop of the kissing complex confirmed that the  $Mg^{2+}$  stabilized conformation is more stable than the  $Na^+$  stabilized RNA structure. Additionally, pH had little effect of the kissing dimer with respect to dissociation in the presence of  $Na^+$  or  $Mg^{2+}$ . Two unpaired junction bases with 2-AP labeling on DIS loop, AP9 and AP10, showed that AP9 is more stacked in the helix in the formation of kissing dimer with a smaller increase in fluorescence in comparison with AP10. The  $Mg^{2+}$  ion is likely to bind at the highly conserved stacked AP9 purine base in the DIS loop. Overall,  $Mg^{2+}$  ions stabilize DIS kissing complexes more than  $Na^+$  ions accompanied with the higher increase in fluorescence for both unpaired 2-AP labeled bases. The different conformations are expected to contribute to the different NCp7 binding interaction and chaperone activity with these two fluorescent labels at the purine junction.

After characterizing the fluorescence properties in terms of metal ion binding on DIS purine junction, NMR studies were used to specifically identify the local  $Mg^{2+}$  binding on DIS dimers. Homo and hetero DIS complexes with stem-loop secondary

structures were designed to examine the structural information depending on metal ions and  $Mg^{2+}$  binding sites. The conformations of DIS complexes in the presence of monovalent and divalent metal ions are different.  $Mg^{2+}$  binding not only stabilizes DIS complexes more than  $Na^+$  but also facilitates RNA strand annealing stimulated by NCp7 chaperone protein (127). Our data showed that  $Mg^{2+}$  ion binding results in specific proton chemical shift changes at DIS21 homo duplex. Three junction bases, G6, A7, and C16, are close to  $Mg^{2+}$  binding sites with significant chemical shift perturbations. Paramagnetic line broadening by  $Mn^{2+}$  suggests that  $Mg^{2+}$  ions are localized in G6, A7, G8, and G9 on the 5' end of the DIS duplex. Hetero DIS stem-loop kissing complex was designed by constructing with an exchanged stem sequence, DIS23(HxUC). Likewise, the selective paramagnetic line broadening has a stronger effect on G7a, A8a, G9a, A16a, and C17a of the DIS23(GA) side in the kissing complex. In conclusion,  $Mg^{2+}$  ions are localized around the 5' purine bases in the extended duplex, whereas the  $Mg^{2+}$  binding pocket is near both 5' and 3' purine bases in the kissing-loop dimer.

The X-ray crystal structures and NMR solution structures of DIS have been solved from group Ennifar et al. (10) and Kawai et al. (14), respectively. The  $Mg^{2+}$  dependent crystal structures of DIS kissing-loop dimers showed that the unpaired purines (A8 and G9) on the junction are flipped out, whereas the purine bases are stacked within the helical stem in solution structures in the absence of  $Mg^{2+}$ . The adenine base, A16, was found to be perfectly stacked within the helix in crystal and solution structures. The discrepancy between crystal and solution structures of DIS centers is primarily on difference in the observed A8 and G9 stacking conformation at the loop junctions. Our DIS extended duplex and kissing dimers were analyzed in the presence of  $Mg^{2+}$  in

solution. The  $\text{Mg}^{2+}$  dependent crystal structures and  $\text{Mg}^{2+}$  free solution structures were used as templates for mapping ion binding sites with the CSP and PRE results. From our results, major  $\text{Mg}^{2+}$  binding sites are localized in the DIS junction around A8, G9, and A16 in the kissing dimer. This is in agreement with the NMR solution structures. Although A8 and G9 might be flexible and dynamic in solution and flip out of the helix in the absence of  $\text{Mg}^{2+}$ , our data suggest that these two bases are stacked in the helix upon  $\text{Mg}^{2+}$  binding. Similarly, a divalent metal binding site localized around A8 base in the  $\text{Mg}^{2+}$  stabilized extended duplex is found to be consistent with the solution structure.

$\text{Mg}^{2+}$  binding interaction affects not only the conformation but also dynamics at the purine junctions in the kissing dimer. These flanking purine bases are believed to stabilize the kissing dimer and assist the initiation of maturation stimulated by NCp7 chaperone protein. In the case of the extended duplex,  $\text{Mg}^{2+}$  binding effect results in the significant chemical shift changes.  $\text{Mg}^{2+}$  ions also stabilize the unpaired purines as part of maturation of the RNA or as part of NCp7 binding during maturation. Although the sequence of purine bases are the same around the DIS junction in both the extended duplex and kissing-loop dimers, the backbone trajectory around the junctions is particularly different in the structure. The different  $\text{Mg}^{2+}$  binding effects in these two dimers are likely to explain the differences associated with NCp7 chaperone activity in the dimerization.

## Bibliography

1. Paillart, J. C., Shehu-Xhilaga, M., Marquet, R., and Mak, J. (2004) Dimerization of retroviral RNA genomes: an inseparable pair, *Nature reviews* 2, 461-472.
2. Turner, B. G., and Summers, M. F. (1999) Structural biology of HIV, *Journal of molecular biology* 285, 1-32.
3. Bender, W., and Davidson, N. (1976) Mapping of poly(A) sequences in the electron microscope reveals unusual structure of type C oncornavirus RNA molecules, *Cell* 7, 595-607.
4. Kung, H. J., Hu, S., Bender, W., Bailey, J. M., Davidson, N., Nicolson, M. O., and McAllister, R. M. (1976) RD-114, baboon, and woolly monkey viral RNA's compared in size and structure, *Cell* 7, 609-620.
5. Bender, W., Chien, Y. H., Chattopadhyay, S., Vogt, P. K., Gardner, M. B., and Davidson, N. (1978) High-molecular-weight RNAs of AKR, NZB, and wild mouse viruses and avian reticuloendotheliosis virus all have similar dimer structures, *Journal of virology* 25, 888-896.
6. Skripkin, E., Paillart, J. C., Marquet, R., Ehresmann, B., and Ehresmann, C. (1994) Identification of the primary site of the human immunodeficiency virus type 1 RNA dimerization in vitro, *Proceedings of the National Academy of Sciences of the United States of America* 91, 4945-4949.
7. Paillart, J. C., Berthoux, L., Ottmann, M., Darlix, J. L., Marquet, R., Ehresmann, B., and Ehresmann, C. (1996) A dual role of the putative RNA dimerization

- initiation site of human immunodeficiency virus type 1 in genomic RNA packaging and proviral DNA synthesis, *Journal of virology* 70, 8348-8354.
8. Muriaux, D., De Rocquigny, H., Roques, B. P., and Paoletti, J. (1996) NCp7 activates HIV-1 RNA dimerization by converting a transient loop-loop complex into a stable dimer, *The Journal of biological chemistry* 271, 33686-33692.
  9. Yusupov, M., Walter, P., Marquet, R., Ehresmann, C., Ehresmann, B., and Dumas, P. (1999) Crystallization of the dimerization-initiation site of genomic HIV-1 RNA: preliminary crystallographic results, *Acta crystallographica* 55, 281-284.
  10. Ennifar, E., Yusupov, M., Walter, P., Marquet, R., Ehresmann, B., Ehresmann, C., and Dumas, P. (1999) The crystal structure of the dimerization initiation site of genomic HIV-1 RNA reveals an extended duplex with two adenine bulges, *Structure* 7, 1439-1449.
  11. Ennifar, E., Walter, P., Ehresmann, B., Ehresmann, C., and Dumas, P. (2001) Crystal structures of coaxially stacked kissing complexes of the HIV-1 RNA dimerization initiation site, *Nature structural biology* 8, 1064-1068.
  12. Ennifar, E., and Dumas, P. (2006) Polymorphism of bulged-out residues in HIV-1 RNA DIS kissing complex and structure comparison with solution studies, *Journal of molecular biology* 356, 771-782.
  13. Mujeeb, A., Clever, J. L., Billeci, T. M., James, T. L., and Parslow, T. G. (1998) Structure of the dimer initiation complex of HIV-1 genomic RNA, *Nature structural biology* 5, 432-436.

14. Baba, S., Takahashi, K., Noguchi, S., Takaku, H., Koyanagi, Y., Yamamoto, N., and Kawai, G. (2005) Solution RNA structures of the HIV-1 dimerization initiation site in the kissing-loop and extended-duplex dimers, *Journal of biochemistry* 138, 583-592.
15. Kieken, F., Paquet, F., Brule, F., Paoletti, J., and Lancelot, G. (2006) A new NMR solution structure of the SL1 HIV-1Lai loop-loop dimer, *Nucleic acids research* 34, 343-352.
16. Dardel, F., Marquet, R., Ehresmann, C., Ehresmann, B., and Blanquet, S. (1998) Solution studies of the dimerization initiation site of HIV-1 genomic RNA, *Nucleic acids research* 26, 3567-3571.
17. Takahashi, K. I., Baba, S., Chattopadhyay, P., Koyanagi, Y., Yamamoto, N., Takaku, H., and Kawai, G. (2000) Structural requirement for the two-step dimerization of human immunodeficiency virus type 1 genome, *RNA (New York, N.Y)* 6, 96-102.
18. Spomer, J., and Spackova, N. (2007) Molecular dynamics simulations and their application to four-stranded DNA, *Methods (San Diego, Calif)* 43, 278-290.
19. Noy, A., Meyer, T., Rueda, M., Ferrer, C., Valencia, A., Perez, A., de la Cruz, X., Lopez-Bes, J. M., Pouplana, R., Fernandez-Recio, J., Luque, F. J., and Orozco, M. (2006) Data mining of molecular dynamics trajectories of nucleic acids, *Journal of biomolecular structure & dynamics* 23, 447-456.
20. MacKerell, A. D., Jr., Banavali, N., and Foloppe, N. (2000) Development and current status of the CHARMM force field for nucleic acids, *Biopolymers* 56, 257-265.

21. Sarzynska, J., Reblova, K., Sponer, J., and Kulinski, T. (2008) Conformational transitions of flanking purines in HIV-1 RNA dimerization initiation site kissing complexes studied by CHARMM explicit solvent molecular dynamics, *Biopolymers* 89, 732-746.
22. Hu, W. S., and Temin, H. M. (1990) Genetic consequences of packaging two RNA genomes in one retroviral particle: pseudodiploidy and high rate of genetic recombination, *Proceedings of the National Academy of Sciences of the United States of America* 87, 1556-1560.
23. Goodenow, M., Huet, T., Saurin, W., Kwok, S., Sninsky, J., and Wain-Hobson, S. (1989) HIV-1 isolates are rapidly evolving quasispecies: evidence for viral mixtures and preferred nucleotide substitutions, *Journal of acquired immune deficiency syndromes* 2, 344-352.
24. Laughrea, M., Jette, L., Mak, J., Kleiman, L., Liang, C., and Wainberg, M. A. (1997) Mutations in the kissing-loop hairpin of human immunodeficiency virus type 1 reduce viral infectivity as well as genomic RNA packaging and dimerization, *Journal of virology* 71, 3397-3406.
25. Ian Marshall, S. D. B. J. H. A. M. J. M. W. K. J. F. J. S. (2000) Choice of spectroscopic lineshape model affects metabolite peak areas and area ratios, *Magnetic Resonance in Medicine* 44, 646-649.
26. Berkhout, B. (1996) Structure and function of the human immunodeficiency virus leader RNA, *Progress in nucleic acid research and molecular biology* 54, 1-34.



27. McBride, M. S., and Panganiban, A. T. (1997) Position dependence of functional hairpins important for human immunodeficiency virus type 1 RNA encapsidation in vivo, *Journal of virology* 71, 2050-2058.
28. Amarasinghe, G. K., De Guzman, R. N., Turner, R. B., and Summers, M. F. (2000) NMR structure of stem-loop SL2 of the HIV-1 psi RNA packaging signal reveals a novel A-U-A base-triple platform, *Journal of molecular biology* 299, 145-156.
29. Freed, E. O. (1998) HIV-1 gag proteins: diverse functions in the virus life cycle, *Virology* 251, 1-15.
30. Freed, E. O. (2003) The HIV-TSG101 interface: recent advances in a budding field, *Trends in microbiology* 11, 56-59.
31. Rein, A., Henderson, L. E., and Levin, J. G. (1998) Nucleic-acid-chaperone activity of retroviral nucleocapsid proteins: significance for viral replication, *Trends in biochemical sciences* 23, 297-301.
32. Chan, B., Weidemaier, K., Yip, W. T., Barbara, P. F., and Musier-Forsyth, K. (1999) Intra-tRNA distance measurements for nucleocapsid protein-independent tRNA unwinding during priming of HIV reverse transcription, *Proceedings of the National Academy of Sciences of the United States of America* 96, 459-464.
33. Tisne, C., Roques, B. P., and Dardel, F. (2001) Heteronuclear NMR studies of the interaction of tRNA(Lys)<sub>3</sub> with HIV-1 nucleocapsid protein, *Journal of molecular biology* 306, 443-454.

34. De Guzman, R. N., Wu, Z. R., Stalling, C. C., Pappalardo, L., Borer, P. N., and Summers, M. F. (1998) Structure of the HIV-1 nucleocapsid protein bound to the SL3 psi-RNA recognition element, *Science (New York, N.Y)* 279, 384-388.
35. Remy, E., de Rocquigny, H., Petitjean, P., Muriaux, D., Theilleux, V., Paoletti, J., and Roques, B. P. (1998) The annealing of tRNA<sup>3</sup>Lys to human immunodeficiency virus type 1 primer binding site is critically dependent on the NCp7 zinc fingers structure, *The Journal of biological chemistry* 273, 4819-4822.
36. Laughrea, M., and Jette, L. (1996) Kissing-loop model of HIV-1 genome dimerization: HIV-1 RNAs can assume alternative dimeric forms, and all sequences upstream or downstream of hairpin 248-271 are dispensable for dimer formation, *Biochemistry* 35, 1589-1598.
37. Marquet, R., Baudin, F., Gabus, C., Darlix, J. L., Mougel, M., Ehresmann, C., and Ehresmann, B. (1991) Dimerization of human immunodeficiency virus (type 1) RNA: stimulation by cations and possible mechanism, *Nucleic acids research* 19, 2349-2357.
38. Rist, M. J., and Marino, J. P. (2002) Mechanism of nucleocapsid protein catalyzed structural isomerization of the dimerization initiation site of HIV-1, *Biochemistry* 41, 14762-14770.
39. Pyle, A. M. (2002) Metal ions in the structure and function of RNA, *J Biol Inorg Chem* 7, 679-690.
40. DeRose, V. J. (2003) Metal ion binding to catalytic RNA molecules, *Current opinion in structural biology* 13, 317-324.

41. Shan, S., Kravchuk, A. V., Piccirilli, J. A., and Herschlag, D. (2001) Defining the catalytic metal ion interactions in the Tetrahymena ribozyme reaction, *Biochemistry* 40, 5161-5171.
42. Shiman, R., and Draper, D. E. (2000) Stabilization of RNA tertiary structure by monovalent cations, *Journal of molecular biology* 302, 79-91.
43. Tinoco, I., Jr., and Bustamante, C. (1999) How RNA folds, *Journal of molecular biology* 293, 271-281.
44. Cole, P. E., Yang, S. K., and Crothers, D. M. (1972) Conformational changes of transfer ribonucleic acid. Equilibrium phase diagrams, *Biochemistry* 11, 4358-4368.
45. Heilman-Miller, S. L., Pan, J., Thirumalai, D., and Woodson, S. A. (2001) Role of counterion condensation in folding of the Tetrahymena ribozyme. II. Counterion-dependence of folding kinetics, *Journal of molecular biology* 309, 57-68.
46. Brown, T. S., Chadalavada, D. M., and Bevilacqua, P. C. (2004) Design of a highly reactive HDV ribozyme sequence uncovers facilitation of RNA folding by alternative pairings and physiological ionic strength, *Journal of molecular biology* 341, 695-712.
47. Bock, C. W., Katz, A. K., Markham, G. D., and Glusker, J. P. (1999) Manganese as a Replacement for Magnesium and Zinc: Functional Comparison of the Divalent Ions, *Journal of the American Chemical Society* 121, 7360-7372.
48. Weber, E., and Aoyama, Y. (1998) Design of Organic Solids, in *Design of Organic Solids* (E., W., Ed.).

49. Conn, G. L., Gittis, A. G., Lattman, E. E., Misra, V. K., and Draper, D. E. (2002) A Compact RNA Tertiary Structure Contains a Buried Backbone-K<sup>+</sup> Complex, *Journal of molecular biology* 318, 963-973.
50. Felg, A. L., and Uhlenbeck, O. C. (1999) *The RNA World*, Second Edition ed.
51. Anderson, C. F., and Record, M. T., Jr. (1995) Salt-nucleic acid interactions, *Annual review of physical chemistry* 46, 657-700.
52. Gonzalez, R. L., Jr., and Tinoco, I., Jr. (2001) Identification and characterization of metal ion binding sites in RNA, *Methods in enzymology* 338, 421-443.
53. Horton, T. E., Clardy, D. R., and DeRose, V. J. (1998) Electron paramagnetic resonance spectroscopic measurement of Mn<sup>2+</sup> binding affinities to the hammerhead ribozyme and correlation with cleavage activity, *Biochemistry* 37, 18094-18101.
54. Yamada, A., Akasaka, K., and Hatano, H. (1976) Proton and phosphorus-31 magnetic relaxation studies on the interaction of polyriboadenylic acid with Mn<sup>2+</sup>, *Biopolymers* 15, 1315-1331.
55. Tanaka, Y., and Taira, K. (2005) Detection of RNA nucleobase metalation by NMR spectroscopy, *Chemical communications (Cambridge, England)*, 2069-2079.
56. Klein, D. J., Moore, P. B., and Steitz, T. A. (2004) The contribution of metal ions to the structural stability of the large ribosomal subunit, *RNA (New York, N.Y)* 10, 1366-1379.
57. Selvin, P. R. (2000) The renaissance of fluorescence resonance energy transfer, *Nature structural biology* 7, 730-734.

58. Murchie, A. I., Clegg, R. M., von Kitzing, E., Duckett, D. R., Diekmann, S., and Lilley, D. M. (1989) Fluorescence energy transfer shows that the four-way DNA junction is a right-handed cross of antiparallel molecules, *Nature* 341, 763-766.
59. Hardman, S. J., and Thompson, K. C. (2006) Influence of base stacking and hydrogen bonding on the fluorescence of 2-aminopurine and pyrrolocytosine in nucleic acids, *Biochemistry* 45, 9145-9155.
60. Jean, J. M., and Hall, K. B. (2001) 2-Aminopurine fluorescence quenching and lifetimes: role of base stacking, *Proceedings of the National Academy of Sciences of the United States of America* 98, 37-41.
61. Sowers, L. C., Fazakerley, G. V., Eritja, R., Kaplan, B. E., and Goodman, M. F. (1986) Base pairing and mutagenesis: observation of a protonated base pair between 2-aminopurine and cytosine in an oligonucleotide by proton NMR, *Proceedings of the National Academy of Sciences of the United States of America* 83, 5434-5438.
62. Fagan, P. A., Fabrega, C., Eritja, R., Goodman, M. F., and Wemmer, D. E. (1996) NMR Study of the Conformation of the 2-Aminopurine:Cytosine Mismatch in DNA, *Biochemistry* 35, 4026-4033.
63. Nordlund, T. M., Andersson, S., Nilsson, L., Rigler, R., Graeslund, A., and McLaughlin, L. W. (1989) Structure and dynamics of a fluorescent DNA oligomer containing the EcoRI recognition sequence: fluorescence, molecular dynamics, and NMR studies, *Biochemistry* 28, 9095-9103.

64. Law, S. M., Eritja, R., Goodman, M. F., and Breslauer, K. J. (1996) Spectroscopic and Calorimetric Characterizations of DNA Duplexes Containing 2-Aminopurine, *Biochemistry* 35, 12329-12337.
65. Borjesson, K., Sandin, P., and Wilhelmsson, L. M. (2009) Nucleic acid structure and sequence probing using fluorescent base analogue tC(O), *Biophysical chemistry* 139, 24-28.
66. Xu, D., Evans, K. O., and Nordlund, T. M. (1994) Melting and Premelting Transitions of an Oligomer Measured by DNA Base Fluorescence and Absorption, *Biochemistry* 33, 9592-9599.
67. Rist, M. J., and Marino, J. P. (2002) Fluorescent Nucleotide Base Analogs as Probes of Nucleic Acid Structure, Dynamics and Interactions, *Current Organic Chemistry* 6, 775.
68. Zhao, C., and Marino, J. P. (2007) Synthesis of HIV-1 Psi-site RNA sequences with site specific incorporation of the fluorescent base analog 2-aminopurine, *Tetrahedron* 63, 3575-3584.
69. Kubo Ryogo, K. T. (1954) A General Theory of Magnetic Resonance Absorption, *Journal of the Physical Society of Japan* 9, 888.
70. Caruthers, M. H., Beaucage, S. L., Efcavitch, J. W., Fisher, E. F., Matteucci, M. D., and Stabinsky, Y. (1980) New chemical methods for synthesizing polynucleotides, *Nucleic acids symposium series*, 215-223.
71. Mihailescu, M. R., and Marino, J. P. (2004) A proton-coupled dynamic conformational switch in the HIV-1 dimerization initiation site kissing complex,

*Proceedings of the National Academy of Sciences of the United States of America* 101, 1189-1194.

72. Lee, H. W., Briggs, K. T., and Marino, J. P. (2009) Dissecting structural transitions in the HIV-1 dimerization initiation site RNA using 2-aminopurine fluorescence, *Methods (San Diego, Calif.*
73. Milligan, J. F., Groebe, D. R., Witherell, G. W., and Uhlenbeck, O. C. (1987) Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates, *Nucleic acids research* 15, 8783-8798.
74. Milligan, J. F., and Uhlenbeck, O. C. (1989) Synthesis of small RNAs using T7 RNA polymerase, *Methods in enzymology* 180, 51-62.
75. Xu, J., Lapham, J., and Crothers, D. M. (1996) Determining RNA solution structure by segmental isotopic labeling and NMR: application to *Caenorhabditis elegans* spliced leader RNA 1, *Proceedings of the National Academy of Sciences of the United States of America* 93, 44-48.
76. D'Souza, V., and Summers, M. F. (2004) Structural basis for packaging the dimeric genome of Moloney murine leukaemia virus, *Nature* 431, 586-590.
77. Furtig, B., Richter, C., Wohnert, J., and Schwalbe, H. (2003) NMR spectroscopy of RNA, *Chembiochem* 4, 936-962.
78. Varani, G., and Tinoco, I., Jr. (1991) RNA Structure and NMR Spectroscopy, *Q. Rev. Biophys.* 24, 479-532.
79. Pardi, A. (1995) Multidimensional heteronuclear NMR experiments for structure determination of isotopically labeled RNA, *Methods in enzymology* 261, 350-380.

80. Aboul-ela, F., and Varani, G. (1995) Novel techniques in nuclear magnetic resonance for nucleic acids, *Current opinion in biotechnology* 6, 89-95.
81. Varani, G., Aboul-ela, F., and Allain, F. H. T. (1996) NMR investigation of RNA structure, *Progress in Nuclear Magnetic Resonance Spectroscopy* 29, 51-127.
82. Boris Furtig, C. R. J. W. H. S. (2003) NMR Spectroscopy of RNA, *ChemBioChem* 4, 936-962.
83. Scott, L. G., and Hennig, M. (2008) RNA Structure Determination by NMR, pp 29-61.
84. Rance, M., Sorensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., and Wuthrich, K. (1983) Improved spectral resolution in cosy  $^1\text{H}$  NMR spectra of proteins via double quantum filtering, *Biochemical and biophysical research communications* 117, 479-485.
85. Claridge, T. D. W. (2004) *High-Resolution NMR Techniques in Organic Chemistry*, Vol. 19, Reprinted ed., Elsevier, Oxford, UK.
86. Marino, J. P., Prestegard, J. H., and Crothers, D. M. (1994) Correlation of Adenine H2/H8 Resonances in Uniformly  $^{13}\text{C}$  Labeled RNAs by 2D HCCH-TOCSY: A New Tool for  $^1\text{H}$  Assignment, *Journal of the American Chemical Society* 116, 2205-2206.
87. Sklenář, V., Peterson, R. D., Rejante, M. R., and Feigon, J. (1993) Two-and three-dimensional HCN experiments for correlating base and sugar resonances in  $^{15}\text{N}$ ,  $^{13}\text{C}$ -labeled RNA oligonucleotides, *Journal of biomolecular NMR* 3, 721-727.
88. Heus, H. A., Wijmenga, S. S., van de Ven, F. J. M., and Hilbers, C. W. (1994) Sequential Backbone Assignment in  $^{13}\text{C}$ -Labeled RNA via Through-Bond



- Coherence Transfer Using Three-Dimensional Triple Resonance Spectroscopy ( $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{31}\text{P}$ ) and Two-Dimensional Hetero TOCSY, *Journal of the American Chemical Society* 116, 4983-4984.
89. Marino, J. P., Schwalbe, H., Anklin, C., Bermel, W., Crothers, D. M., and Griesinger, C. (1995) Sequential correlation of anomeric ribose protons and intervening phosphorus in RNA oligonucleotides by a  $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{31}\text{P}$  triple resonance experiment: HCP-CCH-TOCSY, *Journal of biomolecular NMR* 5, 87-92.
  90. Peterson, R. D., Theimer, C. A., Wu, H., and Feigon, J. (2004) New applications of 2D filtered/edited NOESY for assignment and structure elucidation of RNA and RNA-protein complexes, *Journal of biomolecular NMR* 28, 59-67.
  91. Duchardt, E., Richter, C., Reif, B., Glaser, S. J., Engels, J. W., Griesinger, C., and Schwalbe, H. (2001) Measurement of  $2J(\text{H,C})$ - and  $3J(\text{H,C})$ -coupling constants by alpha/beta selective HC(C)H-TOCSY, *Journal of biomolecular NMR* 21, 117-126.
  92. Jucker, F. M., and Pardi, A. (1995) Solution structure of the CUUG hairpin loop: a novel RNA tetraloop motif, *Biochemistry* 34, 14416-14427.
  93. Schwalbe, H., Marino, J. P., King, G. C., Wechselberger, R., Bermel, W., and Griesinger, C. (1994) Determination of a complete set of coupling constants in  $^{13}\text{C}$ -labeled oligonucleotides, *Journal of biomolecular NMR* 4, 631-644.
  94. Dingley, A. J., and Grzesiek, S. (1998) Direct Observation of Hydrogen Bonds in Nucleic Acid Base Pairs by Internucleotide  $2J(\text{NN})$  Couplings, *Journal of the American Chemical Society* 120, 8293-8297.

95. Kime, M. J., and Moore, P. B. (1983) Nuclear Overhauser experiments at 500 MHz on the downfield proton spectra of 5S ribonucleic acid and its complex with ribosomal protein L25, *Biochemistry* 22, 2622-2629.
96. Dingley, A. J., Masse, J. E., Feigon, J., and Grzesiek, S. (2000) Characterization of the hydrogen bond network in guanosine quartets by internucleotide  $3hJ(NC)'$  and  $2hJ(NN)$  scalar couplings, *Journal of biomolecular NMR* 16, 279-289.
97. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, New York, Wiley-Interscience.
98. Craven, C. J., Baxter, N. J., Murray, E. H., Hill, N. J., Martin, J. R., Ylinenjarvi, K., Bjork, I., Waltho, J. P., and Murray, I. A. (2000) Wild-type and Met-65→Leu variants of human cystatin A are functionally and structurally identical, *Biochemistry* 39, 15783-15790.
99. Hansen, M. R., Mueller, L., and Pardi, A. (1998) Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions, *Nature structural biology* 5, 1065-1074.
100. Sklenar, V., Peterson, R. D., Rejante, M. R., and Feigon, J. (1994) Correlation of nucleotide base and sugar protons in a  $^{15}\text{N}$ -labeled HIV-1 RNA oligonucleotide by  $1\text{H}$ - $^{15}\text{N}$  HSQC experiments, *Journal of biomolecular NMR* 4, 117-122.
101. Mori, S., Abeygunawardana, C., Johnson, M. O., and van Zijl, P. C. (1995) Improved sensitivity of HSQC spectra of exchanging protons at short interscan delays using a new fast HSQC (FHSQC) detection scheme that avoids water saturation, *Journal of magnetic resonance* 108, 94-98.

102. Santoro, J., and King, G. (1992) A Constant-Time 2D Overboderhausen Experiment for Inverse Correlation of Isotopically Enriched Species, *J Magn Reson* 97, 202-207.
103. Griesinger, C., and Eggenberger, U. (1992) Determination of Proton-Proton Coupling Constants in <sup>13</sup>C-Labeled Molecules, *J. Magn. Reson.* 97, 426-434.
104. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes, *Journal of biomolecular NMR* 6, 277-293.
105. Goddard, T. D., and Kneller, D. G. SPARKY 3, SPARKY 3 ed., University of California, San Francisco, San Francisco.
106. Bax, A., and Freeman, R. (1981) *J. Magn. Reson.* 44, 542-561.
107. Moore, P. B. (1995) Determination of RNA conformation by nuclear magnetic resonance. , *Acc. Chem.Res.* 28, 251-256.
108. Sigel, R. K., Vaidya, A., and Pyle, A. M. (2000) Metal ion binding sites in a group II intron core, *Nature structural biology* 7, 1111-1116.
109. Holbrook, S. R., and Kim, S. H. (1997) RNA crystallography, *Biopolymers* 44, 3-21.
110. Hansen, M. R., Simorre, J. P., Hanson, P., Mokler, V., Bellon, L., Beigelman, L., and Pardi, A. (1999) Identification and characterization of a novel high affinity metal-binding site in the hammerhead ribozyme, *RNA (New York, N.Y)* 5, 1099-1104.
111. Bertini, I., and Luchinat, C. (1986) *NMR of Paramagnetic Molecules in Biological Systems*, Menlo Park, CA.

112. Bloembergen, N. (1957) Proton Relaxation Times in Paramagnetic Solutions, *The Journal of Chemical Physics* 27, 572-573.
113. Bertini, I., Donaire, A., Luchinat, C., and Rosato, A. (1997) Paramagnetic relaxation as a tool for solution structure determination: Clostridium pasteurianum ferredoxin as an example, *Proteins* 29, 348-358.
114. Ubbink, M., Worrall, J. A., Canters, G. W., Groenen, E. J., and Huber, M. (2002) Paramagnetic resonance of biological metal centers, *Annual review of biophysics and biomolecular structure* 31, 393-422.
115. Gonzalez, R. L., Jr., and Tinoco, I., Jr. (1999) Solution structure and thermodynamics of a divalent metal ion binding site in an RNA pseudoknot, *Journal of molecular biology* 289, 1267-1282.
116. Gdaniec, Z., Sierzputowska-Gracz, H., and Theil, E. C. (1998) Iron regulatory element and internal loop/bulge structure for ferritin mRNA studied by cobalt(III) hexammine binding, molecular modeling, and NMR spectroscopy, *Biochemistry* 37, 1505-1512.
117. Kieft, J. S., and Tinoco, I., Jr. (1997) Solution structure of a metal-binding site in the major groove of RNA complexed with cobalt (III) hexammine, *Structure* 5, 713-721.
118. Erat, M. C., Zerbe, O., Fox, T., and O, S. R. K. (2007) Cover Picture: 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* 8, 253.

119. Ott, G., Arnold, L., and Limmer, S. (1993) Proton NMR studies of manganese ion binding to tRNA-derived acceptor arm duplexes, *Nucleic acids research* 21, 5859-5864.
120. Frederick, A. F., Kay, L. E., and Prestegard, J. H. (1988) Location of divalent ion sites in acyl carrier protein using relaxation perturbed 2D NMR, *FEBS letters* 238, 43-48.
121. Allain, F. H., and Varani, G. (1995) Divalent metal ion binding to a conserved wobble pair defining the upstream site of cleavage of group I self-splicing introns, *Nucleic acids research* 23, 341-350.
122. Colmenarejo, G., and Tinoco, I., Jr. (1999) Structure and thermodynamics of metal binding in the P5 helix of a group I intron ribozyme, *Journal of molecular biology* 290, 119-135.
123. Haddrick, M., Han-Liu, Z., Lau, A., Heaphy, S., and Cann, A. J. (1996) Production of non-infectious human immunodeficiency virus-like particles which package specifically viral RNA, *Journal of virological methods* 61, 89-93.
124. Moore, M. D., Fu, W., Nikolaitchik, O., Chen, J., Ptak, R. G., and Hu, W. S. (2007) Dimer initiation signal of human immunodeficiency virus type 1: its role in partner selection during RNA copackaging and its effects on recombination, *Journal of virology* 81, 4002-4011.
125. Ennifar, E., Paillart, J. C., Bodlenner, A., Walter, P., Weibel, J. M., Aubertin, A. M., Pale, P., Dumas, P., and Marquet, R. (2006) Targeting the dimerization initiation site of HIV-1 RNA with aminoglycosides: from crystal to cell, *Nucleic acids research* 34, 2328-2339.

126. Darlix, J. L., Lastra, M. L., Mely, Y., and Roques, B. (2002) Nucleocapsid Protein Chaperoning of Nucleic Acids at the Heart of HIV Structure, Assembly and cDNA Synthesis., *HIV Sequence Compendium*, 69-88.
127. Vo, M. N., Barany, G., Rouzina, I., and Musier-Forsyth, K. (2009) Effect of  $Mg^{2+}$  and  $Na^{+}$  on the nucleic acid chaperone activity of HIV-1 nucleocapsid protein: implications for reverse transcription, *Journal of molecular biology* 386, 773-788.