

Studying the Structure and Kinetics of Functional Nucleic Acids

by

Kyle Alexander Piccolo

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Doctor of Philosophy

in

Chemistry

Waterloo, Ontario, Canada, 2021

© Kyle Alexander Piccolo 2021

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner: Dr. Pascale Legault
Professor, Department of Biochemistry, University of Montreal

Supervisor(s): Dr. Thorsten Dieckmann
Associate Professor, Department of Chemistry, University of Waterloo

Internal Member: Dr. Juewen Liu
Professor, Department of Chemistry, University of Waterloo

Dr. Richard Manderville
Professor, Department of Chemistry, University of Guelph

Dr. Elizabeth Meiering
Professor, Department of Chemistry, University of Waterloo

Internal-External Member: Dr. Todd Holyoak
Associate Professor, Department of Biology, University of Waterloo

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

LSPR experiments in Chapter 2 (Figures 2.5 and 2.6, Table 2.5) were performed by Brooke McNeil and the electrostatic potential surfaces (Figures 2.22 and 2.23) were calculated by Jeff Crouse and Su Ji Lim. All other experimental work was performed and analyzed by the candidate.

All experimental work in Chapters 3 and 4 was performed and analyzed by the candidate.

In Chapter 5, ThT fluorescence spectroscopy experiments were performed and analyzed by Runjhun Saran, Yanping He, Yongqiang Kang, Po-Jung Jimmy Huang, Chunying Wei and Da Chen. NMR and CD experiments were performed and analyzed by the candidate. All authors contributed to the text in the abstract, introduction and conclusion sections, while the methods and results and discussion sections were written primarily by the author(s) who used those respective methodologies.

Abstract

Aptamers are valuable tools in a variety of biotechnological applications, ranging from biosensing to *in vivo* cell imaging. Studying these molecules can further our understanding of the structure and kinetics of nucleic acid-ligand binding. The sulforhodamine B binding aptamer (SRB-2) is a polyanionic molecule that was selected to bind a ligand with an overall negative charge. It is also a somewhat promiscuous aptamer that can bind ligands that vary markedly in shape, size and charge. These characteristics make its binding mechanisms of interest to characterize.

In Chapter 2, we categorized potential ligands based on their binding mode and structural characteristics required for high affinity and selectivity. Several known and potential ligands of SRB-2 were screened for binding affinity using fluorescence. Promising candidates were subsequently characterized by biophysical assays including fluorescence, ITC, DSC, LSPR, CD and NMR spectroscopy. These studies showed that rhodamine B has the ideal structural and electrostatic properties for selective and high-affinity binding of the SRB-2 aptamer. These are desirable properties in relevant applications of aptamers, indicating that of SRB-2 ligands currently commercially available, rhodamine B would be the most useful in those contexts.

In Chapter 3, we show NMR experiments performed on the SRB-2 aptamer with the goal of elucidating a solution structure. An unlabelled sample, as well as three samples selectively labelled on a single type of nucleotide (A, C and U) were synthesized and analyzed using a variety of pulse sequences. These included multiple variations of HSQC, NOESY, TOCSY and COSY experiments. A significant amount of NMR data were obtained and some tentative and partial assignments were made, but ultimately the data were insufficient for unambiguous assignment. This resulted from several sources, but broad lines caused by structural heterogeneity were the primary obstacle encountered.

In Chapter 4, we opted to pursue alternative NMR approaches which did not require isotopic labelling. This strategy, segmental analysis, involves breaking a molecule down into smaller segments

and studying them individually. The structural information obtained from the segments was then combined to give an overall picture of the full-sized molecule. This method was used when there are multiple stem-loop segments in a sequence that form stable structures independent of the other stem-loops. Otherwise, this method did not yield accurate depictions of the overall structure. In this chapter, strategies for designing representative truncations are outlined. For a 16mer containing a UUCG tetraloop, assignments were completed, and a structure was calculated. This allowed for several resonances in the spectra of full-length SRB-2 to be identified. Results with another stem-loop region indicate that segmental analysis beyond the UUCG tetraloop is not possible for this system due to interactions between the two other stem-loops.

In Chapter 5, a Na^+ -binding aptamer present in several DNAzymes was investigated. These DNAzymes require Na^+ for activity but show no activity in the presence of K^+ or other metal ions. Given that DNA can selectively bind K^+ by forming a G-quadruplex structure, the goal of this work was to determine whether this Na^+ aptamer also uses a G-quadruplex to bind Na^+ . The Na^+ aptamer is embedded in the trivalent lanthanide-dependent DNAzyme Ce13d. Its sequence consists of multiple GG sequences, which is also a prerequisite for the formation of G4 structures. To probe the structure of Ce13d for G4 presence, thioflavin T (ThT) fluorescence spectroscopy, NMR spectroscopy and CD spectroscopy were used. Through comparative ThT fluorescence spectrometry studies, we determined that a control G-quadruplex DNA exhibited fluorescence enhancement in the presence of K^+ . The Ce13d DNAzyme did not show fluorescence enhancement in the presence of K^+ , but displayed fluorescence decrease with low millimolar concentrations of Na^+ . This agrees with NMR experiments that suggest Ce13d adapts a significantly different conformation, or equilibrium of conformations, in the presence of Na^+ versus K^+ . However, it does appear significantly more stable in Na^+ . The absence of characteristic G-quadruplex peaks in 1D ^1H NMR suggest that a G-quadruplex is not responsible for the Na^+ binding and this is supported by the absence of characteristic peaks in the CD spectra of this sequence. Therefore,

we concluded that the aptamer must be selective for Na⁺ and binds using a structural element that does not contain a G-quadruplex structure.

Acknowledgements

First and foremost, I must express my sincerest gratitude to my supervisor Dr. Thorsten Dieckmann. Thank you for opportunity to work with you on this project. Your encouragement, guidance and positive attitude were much appreciated and helped make this an enjoyable and rewarding experience.

Thank you to my advisory committee members: Dr. Juewen Liu, Dr. Richard Manderville and Dr. Elizabeth Meiring for taking the time out of your busy schedules to support me along the way. I appreciate your valuable insights and wisdom. I would also like to thank my examining committee members Dr. Todd Holyoak and Dr. Pascale Legault for reading my thesis and contributing helpful feedback. I also thank our NMR facility manager Jan Venne for her extensive and much appreciated assistance with NMR spectroscopy.

Thank you to all the past and present members of the Dieckmann group who have contributed in endless ways to this research. I must recognize Volition La and Kyla Egerdeen in particular. Volition helped me out a great deal in my first couple of years here, and both Volition and Kyla were valued friends and colleagues throughout my tenure in the Dieckmann Lab. I had the pleasure of working with many undergraduate students over the years, but I would like to single out the following students: Maryam Najib, Damilola Adebajo, Sarah Bickers, Brooke MacNeil and Samantha Cousineau. They all contributed directly to work appearing in this thesis by making samples, helping to run experiments and/or analyzing data.

Thank you to all of my other friends and colleagues in the department, especially in our biochemistry corner, for sharing your expertise and lending your encouragement. Your comradery helped make grad school a positive experience.

Lastly, thank you to my friends and family. This would not have been possible without you.

Dedication

To my parents Mark and Mary, and brother Matt for your endless love and support.

Table of Contents

Examining Committee Membership	ii
Author’s Declaration	iii
Statement of Contributions	iv
Abstract.....	v
Acknowledgements.....	viii
Dedication	ix
List of Figures	xvii
List of Tables	xxi
List of Abbreviations	xxii
Chapter 1: Literature Review	1
1.1 Overview of Nucleic Acids.....	1
1.2 History of Aptamers	2
1.3 SELEX.....	8
1.3.1 Creating a Library	8
1.3.2 Binding and Partitioning.....	10
1.3.3 Amplification and Analysis	11
1.3.4 Counter-selection.....	12
1.3.5 Modified Nucleic Acids.....	13
1.4 Fluorophore-Binding Aptamers	14
1.4.1 Applications of Fluorophore-Binding Aptamers	14

1.4.2 Structural Studies of Fluorophore Binding Aptamers	27
1.5 Catalytic Nucleic Acids	30
1.5.1 Ribozyme Discovery	30
1.5.2 Selection of Catalytic Nucleic Acids	30
1.5.3 DNAzymes	31
1.5.3 Structural and Functional Importance of Metal Ions.....	31
1.5.4 Metal Biosensors.....	33
1.6 Contents of the Thesis	34
Chapter 2: Ligand Specificity and Affinity in the Sulforhodamine B Binding RNA Aptamer	36
2.1: Foreword.....	36
2.2 Chapter Abstract	36
2.3 Introduction	37
2.4 Materials and Methods.....	42
2.4.1 Preparation of RNA Samples and Dyes	42
2.4.2 Emission Scans	44
2.4.3 Fluorescence Anisotropy.....	45
2.4.4 ITC Studies.....	45
2.4.5 LSPR Assay.....	46
2.4.6 NMR Experiments	46
2.4.7 Electrostatic Potential Surfaces	47

2.4.8 Differential Scanning Calorimetry	47
2.4.9 Circular Dichroism Spectroscopy	48
2.5 Results	48
2.5.1 Emission Scans	48
2.5.2 Fluorescence Anisotropy.....	50
2.5.2 Isothermal Titration Calorimetry	52
2.5.3 Surface Plasmon Resonance Studies	54
2.5.4 NMR Spectroscopy.....	56
2.5.5 Differential Scanning Calorimetry.....	61
2.5.6 Circular Dichroism Spectroscopy	61
2.6 Discussion.....	62
2.6.1 Smaller Size and Reduced Negative Charge of Ligands Result in Higher Binding Affinity	62
2.6.3 Negatively Charged Group is Required for Selectivity but is Likely Not Directly Involved in Binding	66
2.6.4 Size and Shape of the Alkyl Groups are Crucial for Binding.....	67
2.6.5 SRB-2 Binding is Salt and Concentration Dependent.....	68
2.6.6 SRB-2 Lacks any Easily Discernable Tertiary Structure Elements.....	69
2.7 Conclusion.....	70
Chapter 3: Characterizing the Solution Structure of the Sulforhodamine B Binding RNA Aptamer by NMR Spectroscopy.....	71
3.1 Chapter Abstract	71

3.2 Introduction	71
3.3 Materials and Methods.....	81
3.3.1 Preparation of RNA Samples and Dyes	81
3.3.2 NMR Experiments	82
3.4 Results.....	83
3.4.1 Homonuclear NMR Experiments	83
3.4.2 Homonuclear NMR Experiments Without Ligand.....	92
3.4.3 Heteronuclear NMR Experiments with Single-nucleotide Labelling.....	93
3.5 Discussion.....	103
3.5.1 Linewidth and Spectral Resolution	103
3.5.2 Structural Heterogeneity	104
3.5.3 HCN Experiments	106
3.5.4 Sample Stability.....	107
3.5.5 Additional Resources for Assignment	108
3.6 Conclusions and Future Work.....	108
Chapter 4: Probing the Structure of the SRB-2 Binding Aptamer.....	111
4.1 Chapter Abstract	111
4.2 Introduction	111
4.3 Materials and Methods.....	115
4.3.1 Preparation of RNA Samples and Dyes	115

4.3.2 Fluorescence Emission Scans	117
4.3.3 Fluorescence Titrations.....	117
4.3.4 NMR Experiments	118
4.3.5 Input Restraints and Structure Calculations	118
4.4 Results and Discussion	120
4.4.1 Sequence Design and Fluorescence Screening.....	120
4.4.2 NMR Structure of the UUCG 16mer.....	126
4.4.3 NMR Studies of the SRB-2 Truncation L3C.....	132
4.5 Conclusions	134
Chapter 5: Thioflavin T Fluorescence and NMR Spectroscopy Suggesting a Non-G-quadruplex Structure for a Sodium Binding Aptamer.....	
	136
5.1 Foreward.....	136
5.2 Chapter Abstract	137
5.3 Introduction	137
5.4 Materials and Methods.....	140
5.4.1 Chemicals	140
5.4.2 ThT Fluorescence Spectroscopy.....	140
5.4.3 Nuclear Magnetic Resonance	141
5.4.4 Circular Dichroism	141
5.5 Results and Discussion	142
5.5.1 The Ce13d DNAzyme.....	142

5.5.2 ThT fluorescence spectroscopy.....	144
5.5.3 Design of a cis-DNAzyme for NMR Spectroscopy	148
5.5.4. Folding of Ce13d in Li ⁺ , Na ⁺ and K ⁺	150
5.5.5 NMR Spectra Suggest the Na ⁺ -binding Structure is not a G-quadruplex.....	153
5.5.6 CD Spectra Confirm the Absence of G-quadruplex Structure.....	155
5.5.7 Potential Structural Information from 2D NMR.....	156
5.5.8 Structural Effects of Lanthanum Titration on Cis13dB	159
5.6 Conclusions	162
Chapter 6: Summary and Future Work.....	164
Copyright Permissions	167
Biochemical and Biophysical Research Communications (Elsevier)	167
Canadian Journal of Chemistry	169
References	171
Appendices.....	191
Appendix A: Supplementary Figures.....	191
Appendix B: NMR Pulse Programs	200
1D 1H with 11-Spin Echo Solvent Suppression.....	200
1D 1H with Presaturation.....	200
2D NOESY with 11-Spin Echo Solvent Suppression	200
2D NOESY with Presaturation	201

2D CITY-TOCSY	202
2D DQF-COSY	203
2D 13C HSQC.....	203
2D 15N HSQC	204
3D HCCH COSY	205
2D NOESY with X-Half Filter in F2	208
2D NOESY with Double X-Half Filters.....	209
3D Edited HSQC-NOESY.....	211
3D HCCNH	212
3D HCCNH-TOCSY.....	215
3D HCN.....	218
Appendix C: UUCG Chemical Shifts and Restraints for Structure Calculations.....	221
UUCG 16mer Assigned Chemical Shifts	221
UUCG 16mer NOE Restraints	223
List of Hydrogen Bond Restraints.....	233
UUCG 16mer Dihedral Angle Restraints	234
UUCG 16mer Planarity Restraints.....	245

List of Figures

Figure 1.1: SELEX schematic outlining Craig Tuerk and Larry Gold's original procedure.	5
Figure 1.2: Schematic outlining Andrew Ellington and Jack Szostak's original protocol for in vitro selection.	7
Figure 1.3: Schematic outlining the steps in a single round of a general aptamer selection.	9
Figure 1.4: Types of nucleic acid modifications used to confer nuclease resistance in aptamers.	14
Figure 1.5: Relevant ligands of fluorophore-binding aptamers.	15
Figure 1.6: Desired redox reaction of the SRB-2 DNA aptamer.	16
Figure 1.7: Site-specific inactivation of mRNA transcripts by chromophore assisted laser inactivation. ...	18
Figure 1.8: Structural basis of light-up aptamers for intracellular imaging.	20
Figure 1.9: Chemical structure of the GFP fluorophore and its RNA aptamer mimics.	23
Figure 1.10: Metabolite sensors based on the malachite green aptamer.	26
Figure 1.11: Structures of selected fluorophore binding aptamers.	29
Figure 1.12: Metal coordination to several nucleic acid structures.	33
Figure 2.1: Secondary structure of SRB-2 and chemical structures of several studied dyes.	41
Figure 2.2: Chemical structures of ligands screened for SRB-2 binding activity.	42
Figure 2.3: Emission scans comparing free dye (blue) and dye bound to SRB-2 (red).	49
Figure 2.4: Fluorescence anisotropy titration curves for SRB-2 and its various ligands.	51
Figure 2.5: Representative ITC data for the binding of SRB-2 to SR.	53
Figure 2.6: LSPR response curve for Biocytin-TMR.	55
Figure 2.7: LSPR response curve for Atto Rho 101.	55
Figure 2.8: 1D ¹ H NMR spectra of free SRB-2 and SRB-2 bound in a 1:1 ratio to several different ligands.	57

Figure 2.9: 1D ^1H NMR spectra of free SRB-2 and SRB-2 bound in a 1:1 ratio to several different ligands.	58
Figure 2.10: TOCSY spectra of 1mM SR101 (red) and 1.1mM SRB-2 fully bound to SR101 (black).....	59
Figure 2.11: NOESY spectrum of 1.8 mM SRB-2 ~1:1 with sulforhodamine B in 90% $\text{H}_2\text{O}/10\%$ D_2O	60
Figure 2.12: Differential scanning calorimetry isotherms.	61
Figure 2.13: CD and UV absorbance spectra of SRB-2.	62
Figure 2.14: Electrostatic potential surfaces for sulforhodamine B (SR), rhodamine B (RB), sulforhodamine 101 (SR101), tetramethylrosamine (TMR), patent blue V (PBV) and fluorescein (FL).	64
Figure 2.15: Electrostatic potential surfaces of 5-biocytn tetramethylrhodamine, 6-biocytn tetramethylrhodamine (B-TMR), Atto rhodamine 101 biotin (AR), 9-aminoacridine (9AA), Acridine orange (AO), Pyronin Y (PY), Atto 495, Rhodamine 6G (R6G), Rhodamine 110 (R110) and Xylene cyanol FF (XC).	65
Figure 3.1: Proton chemical shift ranges of nucleic acids.....	74
Figure 3.2: Important interactions in RNA backbone assignment.....	76
Figure 3.3: Heteronuclear through-bond NMR experiments used for resonance assignment of nucleic acids.	79
Figure 3.4: 1D titration of 1 mM SRB-2 with SR.....	84
Figure 3.5: NOESY of 1.8 mM SRB-2 bound to SR in 90% $\text{H}_2\text{O}/10\%$ D_2O	85
Figure 3.6: TOCSY of 0.6 mM SRB-2 bound to SR.	87
Figure 3.7: DQF COSY of 0.6 mM SRB-2 bound to SR.	88
Figure 3.8: NOESY of 1.8 mM SRB-2 bound to SR in D_2O	89
Figure 3.9: Effect of Mg^{2+} on NOESY of 0.8 mM SRB-2.	90
Figure 3.10: $\text{H1}'\text{-H6/H8}$ region of NOESY in D_2O of 1.3 mM SRB-2 bound to RB.....	91
Figure 3.11: TOCSY spectrum of 1 mM sulforhodamine B.	92

Figure 3.12: NMR spectra acquired on 0.5 mM unbound SRB-2.....	93
Figure 3.13: ^1H - ^{13}C HSQC spectra of 0.7 mM [^{13}C ^{15}N]-Ade SRB-2 bound to SR.....	95
Figure 3.14: ^1H - ^{13}C HSQC spectra of 0.5 mM [^{13}C ^{15}N]-Cyt SRB-2 bound to SR.....	95
Figure 3.15: HSQC spectra of 0.9 mM [^{13}C ^{15}N]-Ura SRB-2 bound to SR.....	97
Figure 3.16: 2D HCCH-COSY spectrum of 0.9 mM [^{13}C ^{15}N]-Ura SRB-2 bound to SR.....	98
Figure 3.17: Double half-x-filtered NOESY of 0.7 mM [^{13}C ^{15}N]-Ade SRB-2 bound to SR.....	99
Figure 3.18: F2 half-x-filtered NOESY of 0.7 mM [^{13}C ^{15}N]-Ade SRB-2 bound to SR.....	100
Figure 3.19: Half x-filtered NOESY of 0.5 mM [^{13}C ^{15}N]-Cyt SRB-2 bound to SR.....	101
Figure 3.20: F1-F3 projections of 3D ^{13}C HSQC-NOESY experiments.....	102
Figure 4.1: Schematic of the cUUCGg tetraloop.....	115
Figure 4.2: Segments of the SRB-2 aptamer designed for NMR studies.....	122
Figure 4.3: SRB-2 variants studied using fluorescence titrations.....	123
Figure 4.4: Fluorescence emission scans of SRB-2 mutants and truncations.....	124
Figure 4.5: Fluorescence titrations of SRB-2 mutants with SR.....	125
Figure 4.6: NOESY of 2.0 mM UUCG 16mer in D_2O	127
Figure 4.7: NOESY of 2.0 mM UUCG 16mer in 90% H_2O /10% D_2O	128
Figure 4.8: TOCSY and COSY spectra of 2.0 mM UUCG 16mer.....	129
Figure 4.9: Stereo view of the superposition of the 20 lowest energy structures of the UUCG 16mer....	131
Figure 4.10: Literature comparison of the UUCG tetraloop.....	132
Figure 4.11: 2D NOESY spectra of L3C.....	133
Figure 4.12: H5-H6 region in the 2D TOCSY of 1.5 mM L3C.....	134
Figure 5.1: Secondary structure of molecules used in ThT fluorescence experiments.....	143
Figure 5.2: ThT fluorescence spectra of Ce13d and G4 control.....	145
Figure 5.3: ThT fluorescence intensities in lower buffer concentration.....	146

Figure 5.4: Fluorescence intensity at different DNA concentrations.	147
Figure 5.5: Ce13d sequence variations used for NMR spectroscopy.	148
Figure 5.6: Imino proton region of 90 % H ₂ O / 10 % D ₂ O 1D ¹ H NMR spectra of Ce13dA at 277K.	152
Figure 5.7: Imino proton region of 90 % H ₂ O / 10 % D ₂ O 1D ¹ H NMR spectra of Ce13dA at 277K.	153
Figure 5.8: CD spectra of Ce13dA in the presence of no added salt, K ⁺ and Na ⁺	156
Figure 5.9: 2D-TOCSY spectrum of 450 μM Ce13dB in 5 mM LiPO ₄ pH 6.8, 200 mM Na ⁺	157
Figure 5.10: H5/H6 proton region of the 2D-TOCSY spectrum of Ce13dB.	158
Figure 5.11: Titration of Cis13dB with lanthanide.	160
Figure 5.12: NOESY of Cis13dB in 90% H ₂ O/10% D ₂ O.....	161
Figure 5.13: H1'-H6/H8 region in the NOESY spectrum of Ce13dA in D ₂ O.....	162
Figure A.1: 1D titration of 1 mM SRB-2 with SR.....	191
Figure A.2: TOCSY spectra of 1mM SR (red) and 1.8mM SRB-2 fully bound to SR (black).	192
Figure A.3: TOCSY spectra of 1mM RB (red) and 1.3mM SRB-2 fully bound to RB (black).....	193
Figure A.4: H ₂ O NOESY spectrum of 1.3 mM SRB-2 bound ~1:1 with rhodamine B.....	194
Figure A.5: H ₂ O NOESY spectrum of 1.1 mM SRB-2 bound ~1:1 with sulforhodamine 101.....	195
Figure A.6: H ₂ O NOESY spectrum of 0.4 mM SRB-2 bound ~1:1 with tetramethylrosamine.	196
Figure A.7: H ₂ O NOESY spectrum of 0.4 mM SRB-2 bound ~1:1 with pyronin Y.	197
Figure A.8: H ₂ O NOESY spectrum of 0.4 mM SRB-2 bound ~1:1 with 9-aminoacridine.....	198
Figure A.9: H ₂ O NOESY spectrum of 0.4 mM SRB-2 bound ~1:1 with Atto 495.	199

List of Tables

Table 2.1: Sample <i>in vitro</i> transcription recipe.....	43
Table 2.2: Wavelength shifts and fluorescence changes determined by emission scans.	50
Table 2.3: Dissociation constants of SRB-2 ligands determined by fluorescence anisotropy.	52
Table 2.4: Thermodynamic parameters of SRB-2 determined by ITC.	54
Table 2.5: Thermodynamic parameters of SRB-2 determined by ITC in 0.5M KCl and no MgCl ₂	54
Table 2.6: Kinetic parameters of SRB-2 binding determined by LSPR.	56
Table 3.1: PDB Statistics by Molecular Type [200].	73
Table 4.1: List of relevant DNA template sequences.....	116
Table 4.2: Dissociation constants of SRB-2 mutants determined by fluorescence titration.	126
Table 4.3: Structure determination statistics for UUCG 16mer.....	130

List of Abbreviations

9AA	9-aminoacridine
AO	Acridine orange
ARB101-B	Atto rhodamine 101 biotin
A/Ade	Adenine
ATP	Adenosine triphosphate
B-TMR	5-(and-6)-tetramethylrhodamine biocytin
CALI	Chromophore-assisted laser inactivation
CARA	Computer aided resonance assignment
CD	Circular dichroism
COSY	Correlation spectroscopy
C/Cyt	Cytidine
CTP	Cytosine triphosphate
DEAE	Diethylaminoethyl
DNA	Deoxyribonucleic acid
DSC	Differential scanning calorimetry
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
<i>E.coli</i>	<i>Escherichia coli</i>
EPS	Electrostatic potential surface
FL	Fluorescein
FMN	Flavin mononucleotide
FPLC	Fast protein liquid chromatography
G4	G-quadruplex

GFP Green fluorescent protein

G/Gua Guanidine

GTP Guanosine triphosphate

HEPES 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

HSQC Heteronuclear single quantum correlation

IPTG Isopropyl- β -D-thiogalactopyranoside

ITC Isothermal titration calorimetry

K_d Dissociation constant

LSPR Light surface plasmon resonance

MG Malachite green

MGA Malachite green aptamer

mRNA Messenger RNA

NMR Nuclear magnetic resonance

NTP Nucleoside triphosphate

NOESY Nuclear Overhauser effect spectroscopy

PBV Patent blue V

PBVF Patent blue VF

PMSF Phenylmethylsulphonylfluoride

PY Pyronin Y

R110 Rhodamine 110

R6G Rhodamine 6G

RB Rhodamine B

RNA Ribonucleic acid

SDS PAGE Sodium dodecyl sulphate polyacrylamide gel electrophoresis

SELEX Systematic evolution of ligands by exponential enrichment

SR Sulforhodamine B

SR101 Sulforhodamine 101

SRB-2 Sulforhodamine B binding aptamer

TBS Tris-buffered saline

T_m Melting temperature

TMR Tetramethylrosamine

TMR3 Tetramethylrhodamine aptamer

TOCSY Total correlation spectroscopy

Tris Tris(hydroxymethyl)aminomethane

Urea PAGE . . . Urea polyacrylamide gel electrophoresis

U/Ura Uridine

UTP Uridine triphosphate

UTR Untranslated region

XC Xylene cyanol FF

Chapter 1: Literature Review

1.1 Overview of Nucleic Acids

Our understanding of the biological importance of nucleic acids has come a long way in the last three-quarters of a century. The significance of DNA and RNA in biology cannot be understated; as a matter of fact, their story can be summarized in Nobel Prizes alone. Modern understanding of nucleic acids all started with James Watson and Francis Crick's (and Maurice Wilkins' and Rosalind Franklin's) model of DNA, and won the 1962 Nobel Prize in Chemistry [1–3]. Shortly thereafter, the mechanism of nucleic acid synthesis was detailed by Ochoa and Kornberg, who discovered and characterized DNA polymerase [4,5]. They were actually recognized by the Nobel committee for their work on nucleic acid synthesis a few years earlier than Watson, Crick and Wilkins, receiving their prize for physiology in 1958. The big picture became much clearer when Robert Holley, Har Khorana, and Marshall Nirenberg were able to classify the functions of biomolecules, specifically the role of DNA in protein synthesis, forming what we know today as the central dogma of molecular biology. Holley discovered and characterized the structure and function of tRNA [6], while Khorana and Nirenberg determined the genetic code, showing how codon triplets code for amino acids [7,8]. Their work won the Nobel Prize for physiology in 1968.

With the fundamentals of nucleic acid structure, chemistry and function now in place, it was time for field of molecular genetics to shine. This was made possible by the discovery of restriction enzymes by Werner Arber [9]. Hamilton Smith and Daniel Nathans discovered additional restriction endonucleases and were then able to show how they work and how they can be used to create genetic maps, respectively [10,11]. These three men received the 1978 Nobel Prize in physiology. Next in line were Paul Berg, Walter Gilbert and Frederick Sanger, who won the Nobel Prize in chemistry just two years later. Berg was recognized for his work understanding the mechanisms of DNA recombination in viruses [12] and Gilbert and Sanger for their development of DNA sequencing methods [13,14]. Aside

from paving the way for the next chapter in studying nucleic acid structure and function, this work also planted the first seeds in the biotechnology field. Now that nucleic acids were becoming better characterized, scientists could begin to think of how to use these biological processes and chemistry in new, innovative ways.

To summarize this section, the biological and evolutionary importance of nucleic acids is immeasurable, but at their core nucleic acids are biomolecules, polymers and chemicals. They are tools that scientists can apply to large variety of applications, just like proteins and small molecules.

1.2 History of Aptamers

In the late '70s and early '80s another revolutionary, Nobel Prize winning discovery was made that would forever change the ways scientists viewed nucleic acids. Sidney Altman and Thomas Cech discovered RNA with self-catalytic capabilities. Altman and Cech independently found that RNA could cut strands of RNA, a discovery suggesting that life could have started as RNA [15,16]. This is now known as the RNA world hypothesis, a phrase eventually coined by the aforementioned Nobel laureate Walter Gilbert. This work was eventually awarded the 1989 Nobel Prize in chemistry. We will come back to catalytic nucleic acids later, but the importance of this discovery with respect to the eventual development of aptamers was that RNA could have functional roles beyond expression of the genetic code. This was expanded upon when research by several groups on HIV was finally able to determine the previously unknown role of viral-associated RNAs, short RNA sequences that had been found in adenoviruses [17–20]. Their role in transcriptional regulation was fascinating in and of itself, but the major discovery here was that short, partially double-stranded nucleic acid sequences can bind targets, in this case proteins, with high affinity and specificity. Today, that is the textbook definition of an aptamer.

Following the establishment of the biological functions of nucleic acids and the existence of functional nucleic acids, researchers began to investigate the diverse structures of single stranded nucleic acids. As aptamer pioneer Larry Gold has said, they are “shapes, not tapes” [21]. Their tertiary structures, particularly those once thought of as ‘unique’ or ‘outliers’ such as rRNA and tRNA, were now recognized as more similar to globular proteins than to linear DNA. It was not lost on biochemists of the time that these RNA molecules somewhat resembled the antigen binding sites of monoclonal antibodies. Of course, by this time, the discovery of monoclonal antibodies and the creation of monoclonal antibody hybridomas had already won Nobel Prizes of their own and their potential to treat diseases was recognized. This inspired the search in the late ‘80s and early ‘90s for so called nucleic acid antibodies. This was done using *in vitro* mutagenesis, a method that would eventually be developed into what we now know as SELEX (the systematic evolution of ligands by exponential enrichment).

The concept of *in vitro* selection originated in the 1960s with work by Sol Spiegelman. His studies on the Q β RNA bacteriophage led to the discovery of its replicase protein, an RNA-dependent RNA polymerase [22]. This enzyme made *in vitro* synthesis of viral RNA possible. Spiegelman’s work with exogenous viral RNA was able to show Darwinian selection when various selective pressures were applied [23]. These experiments resulted in mutants that had eliminated large portions of their nucleotide sequence, retaining only those required for replication. Mutants with other interesting phenotypes such as resistance to ethidium were also obtained and became the basis for studies on selecting mutants with increased replication rates under inhibitory conditions [24]. This was some significant progress, but a couple of limitations existed that would not be lifted until years later. Any sequence of interest required a vector in order to be studied in this capacity due to their swift removal upon any selection pressure. In addition, the scale of amplification, in general, was still minuscule compared to what we have to work with today. In another decade, however, this would all change.

In the 1980s, Kary Mullis revolutionized the future of aptamer research when he came up with the concept of the polymerase chain reaction (PCR) [25]. PCR allowed for the rapid amplification of DNA, making millions to billions of copies in mere hours. It also made the process of amplification sequence independent. This allowed researchers to easily amplify any sequence(s) of interest. This invention won him the Nobel Prize in chemistry in 1993. Reverse transcription PCR was invented shortly thereafter and is a staple of aptamer selection, and many other methodologies, to this day.

In 1989, Arnold Oliphant and Kevin Struhl published a paper in which they had synthesized DNA oligonucleotides 25 base pairs in length with random sequences [26]. They screened those sequences to select for those that bound to a yeast transcription factor, GCN4. However, this work did not use amplification in any sense of the word. They simply used affinity chromatography to screen all of the DNA sequences that they had made. This was really the first time *in vitro* mutagenesis had been used on this scale, but it still lacked the critical components of evolution and amplification.

Finally, by combining the *in vitro* mutagenesis work of Oliphant and Struhl with the PCR technique invented by Mullis, SELEX was established independently by two groups in 1990: first by Larry Gold and his student Craig Tuerk, then by Jack Szostak and his student Andrew Ellington about a month later [27,28]. Craig Tuerk and Larry Gold's work focused on an RNA sequence with a known affinity to bacteriophage T4 DNA polymerase. They took an eight nucleotide stretch of this RNA and completely randomized the sequence, creating a pool of 65 536 (4^8) theoretical mutants. Their initial library was large enough (theoretically 5×10^9 copies of each individual sequence were present) that even at the lowest concentrations they used, millions of copies of each sequence should be present at the start of the experiment, assuming relatively consistent incorporation of each type of nucleotide. This suggests a high statistical probability that all theoretical sequences were sufficiently represented. After initial transcription of the DNA library, selection was performed with the bacteriophage T4 DNA polymerase (gp43). These sequences were isolated and reverse transcribed, followed by PCR amplification. Four

“rounds” of this procedure were performed. In the original publication, they referred to this method as “the systematic evolution of ligands by exponential enrichment” or SELEX for short. From the 65 536 theoretical sequences, they obtained two with high binding affinity for the bacteriophage T4 DNA polymerase, the wild-type sequence and a mutant with four varying nucleotides. A more detailed schematic of this procedure is shown in Figure 1.

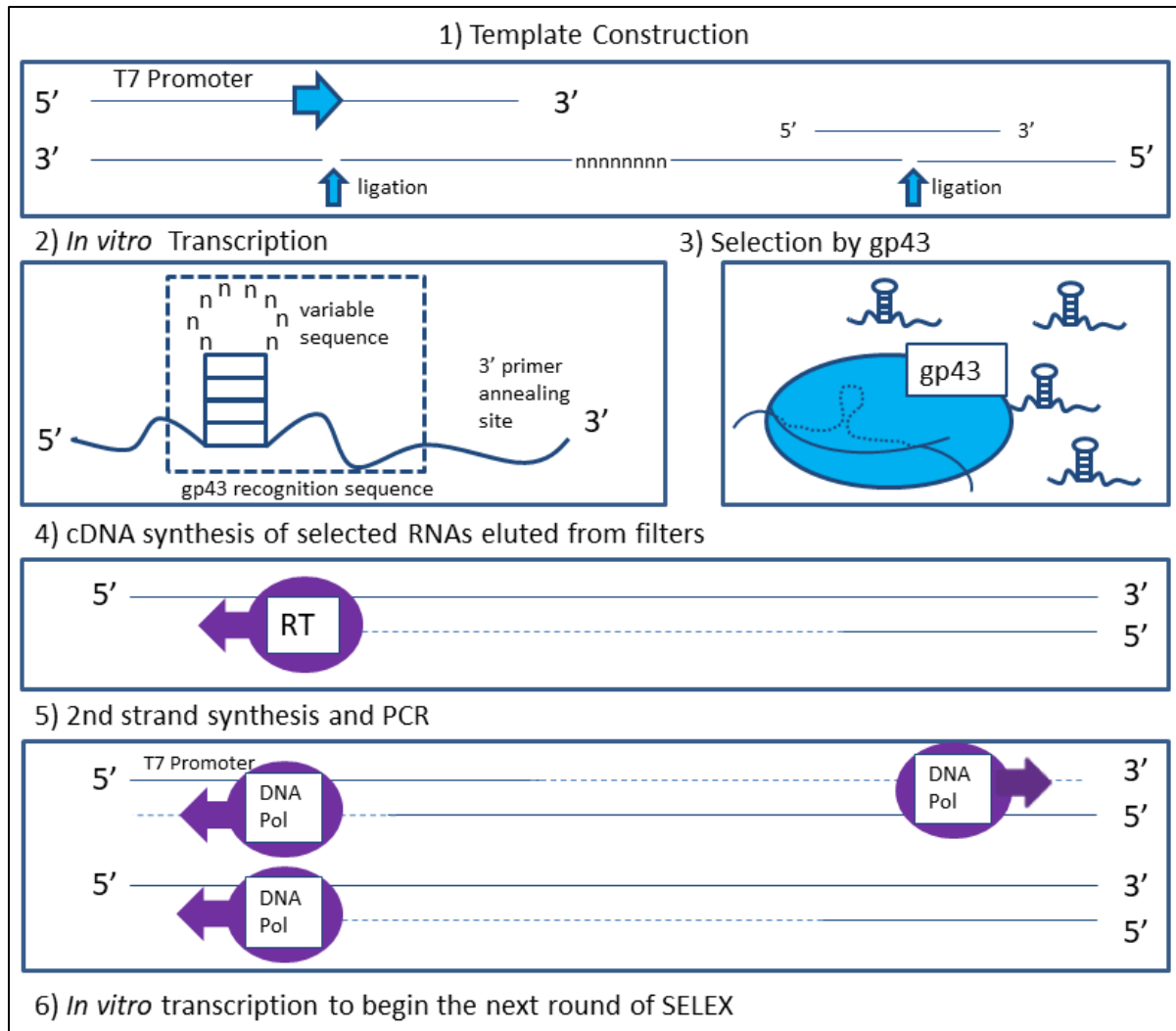


Figure 1.1: SELEX schematic outlining Craig Tuerk and Larry Gold’s original procedure.

The general steps of this procedure include construction of a DNA template for transcription, *in vitro* transcription of the RNA by T7 RNA polymerase, selection on nitrocellulose filters by the target, gp43, reverse transcription, and PCR.

Ellington and Szostak used six rounds of essentially identical methodology (Figure 2) to select sequences that bound to seven different dye columns. They were interested in these particular dyes due to their resemblance to biological cofactors, for example, Cibracon blue, which was known to bind to the NAD binding site of many dehydrogenases and had previously been used for affinity chromatography. They were able to obtain sequences with varying degrees of affinity and selectivity for the various dye columns. Some of the clones obtained varied markedly while others had regions of conservation, suggesting that nucleic acids can fit around their targets in numerous ways but that only certain parts of the sequence would be involved in binding. In fact, Ellington coined the term “aptamer” from this observation as it comes from the Latin *aptus*, which means “to fit.” Ellington and Szostak also concluded that novel ribozymes might be obtained by applying this technique to transition state affinity columns.

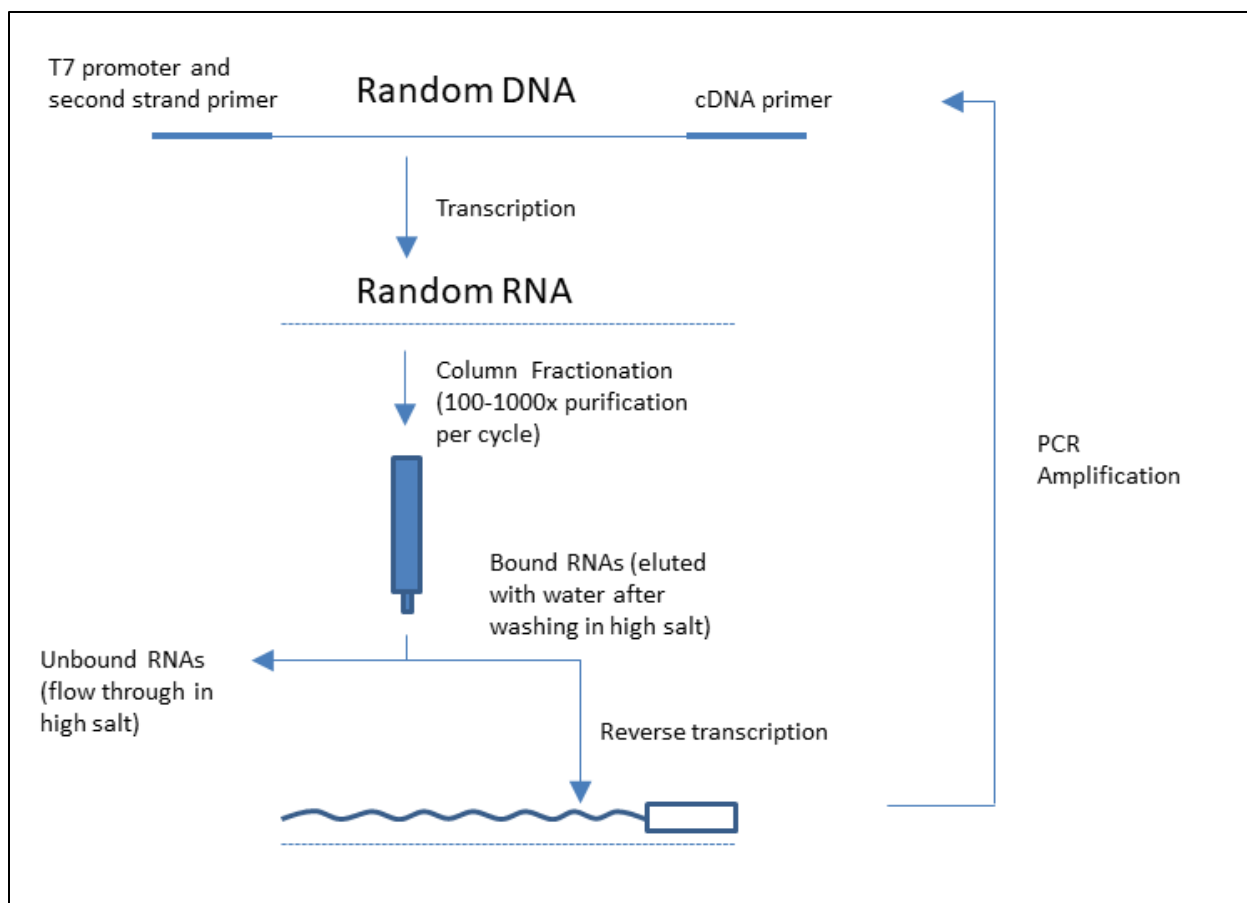


Figure 1.2: Schematic outlining Andrew Ellington and Jack Szostak’s original protocol for in vitro selection.

The general steps of Ellington and Szostak’s procedure mirror those of Tuerk and Gold, however the selection step was done using a column based methodology rather than a nitrocellulose membrane.

Tuerk and Gold’s acronym SELEX stuck, as did Ellington and Szostak’s term “aptamer”. Today you perform SELEX to get aptamers. Ultimately, Craig Tuerk and Larry Gold were awarded the patent for the invention of their SELEX technology [29]. Larry Gold and Gilead Sciences Inc. also held many other broad patents on aptamer selection methodology such as solution SELEX and counter SELEX [30,31]. As a result of these patents and several others, the commercial interest of these molecules was largely subdued until their expiry in the late 2000s and early 2010s. Throughout these years, countless advances in

selection were made, including the improvement of libraries, selection of sequences with complex and dual functions and various strategies for improving the selectivity and nuclease resistance of aptamers. Next we will take a closer look at some of these advances and outline the general methodologies and considerations that go into a contemporary aptamer selection.

1.3 SELEX

Two different schemes for RNA aptamer selection were already shown in Figures 1 and 2, but a generalized scheme applicable to more types of nucleic acid selections is shown in Figure 3. Each step shown in this scheme will be discussed briefly, followed by a few other important factors that are not explicitly shown here.

1.3.1 Creating a Library

The first step of any aptamer selection is to obtain a library of sequences from which to select. A typical selection consists of a sequence of ~30-80 nucleotides partly or entirely randomized, flanked by two fixed sequences of approximately 20 nucleotides in length that are designed to contain appropriate PCR primers [32]. When selecting RNA aptamers, a T7 promoter is also required for transcription prior to each round. Random regions are synthesized by using an equimolar mixture of all nucleotides, resulting in a pool of $\sim 10^{14}$ unique sequences. The probability of finding a particular sequence is $1/4^N$, i.e. the longer the random region, the less likely it is to find a specific sequence due to sampling bias. Therefore, multiple selections may be required in some cases [33]. The oligonucleotide pool is also usually amplified to ensure multiple copies of each sequence are present to limit the potential loss of binding sequences. Sometimes a genomic library can be used instead of a synthetic one, but this of course limits possible aptamer sequences to naturally occurring ones [34].

Another consideration here for selection of DNA aptamers is that the DNA must be single-stranded for selection and for future PCR amplifications. One of the most effective and commonly used methods for separating the strands is by biotinylation of the reverse primer, which allows for the removal of the template strand after PCR amplification by binding to streptavidin. Other more cost efficient methods include asymmetric PCR, size separation by PAGE or enzymatic digestion of the undesired strand [35].

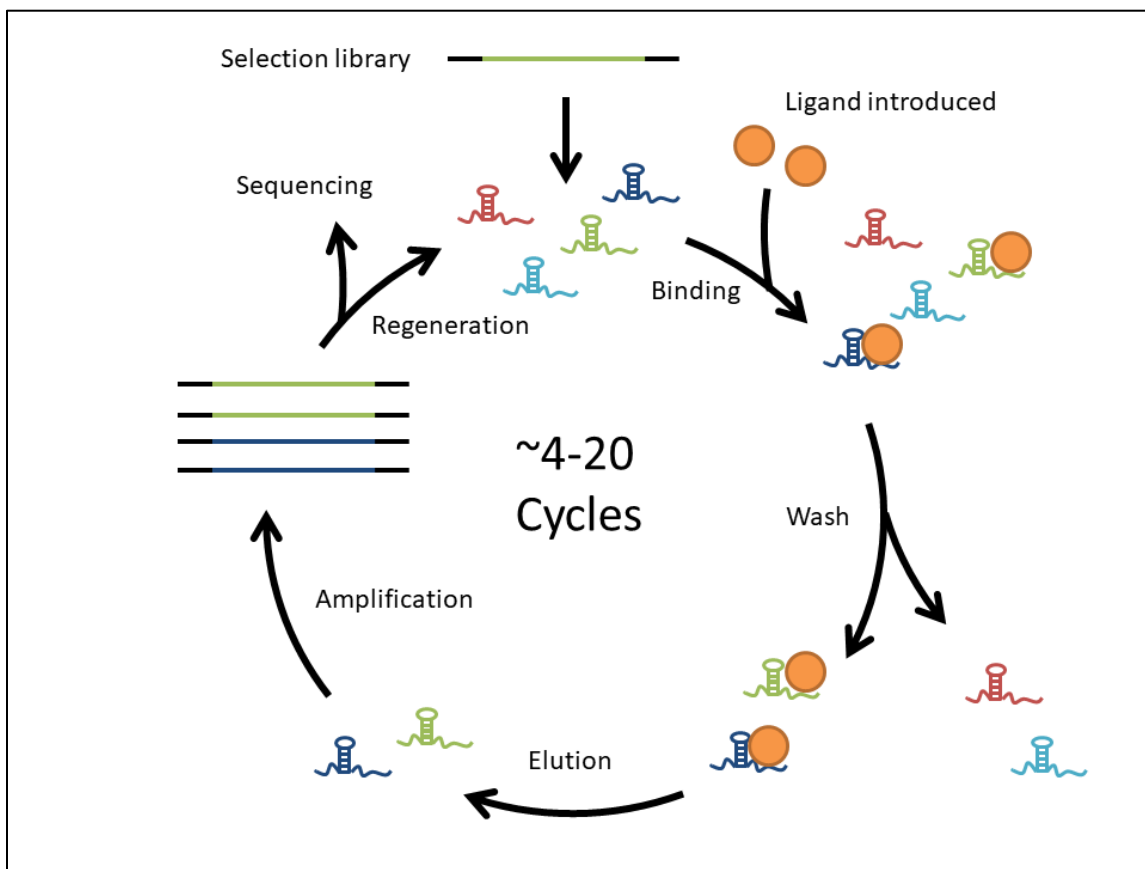


Figure 1.3: Schematic outlining the steps in a single round of a general aptamer selection.

General selection steps are still the same as those used by Szostak and Gold when they independently discovered this process for obtaining RNA aptamers, however, this is a more modern, general schematic that applies to both RNA and DNA selections.

1.3.2 Binding and Partitioning

As shown in Figure 1.3, there are three main steps in a single selection cycle: binding, partitioning and amplification. In practice, ~5-15 rounds of selection are performed. The number of rounds generally depends on the type of target, level of stringency used in the selection and complexity of the initial pool. There are no outright restrictions on the target that is used; in literature, targets include ions, small molecules, biomolecules, and even entire cells [36–39]. The intrinsic properties of nucleic acids do make for some inherent limitations, however. Nucleic acids are small, negatively charged and hydrophilic molecules, which makes targeting a small, positively charged molecule more straightforward, while it may be more difficult to target a large, hydrophobic or negatively charged ligand [40]. The type of matrix used for binding will also depend on the target. Common strategies, particularly for small molecule targets, normally involve immobilization of the target on an affinity column [28] or paramagnetic beads [41]. Nitrocellulose filters are often the matrix of choice when using protein targets [27], and agar plates are used for cell-SELEX [42]. How the target is anchored to the matrix is also a worthwhile consideration as the binding sequences obtained will tend to bind the most accessible parts of the target and may not be able to bind regions close to the linker.

There are several reaction conditions that must be considered as well, including concentrations of the target and library. Buffer conditions must also be optimized to promote binding; for a small molecule selection, this generally consists of neutral pH, high salt and presence of divalent cations which often act as cofactors in target binding [43]. The incubation time and temperature are also crucial parameters as they must be permissive to specific binding but minimize non-specific binding. Several of these conditions, particularly target concentration and incubation time, are also adjusted throughout the cycles to increase the stringency of the selection. This is done to ensure the acquisition of only high-affinity sequences. Stringency is usually not increased until after the first few rounds to prevent the loss of rare sequences and is then increased progressively throughout the rest of the selection. [40]

Partitioning refers to the separation of binding and non-binding sequences. This is normally accomplished by first washing the non-binding and weakly binding sequences away with binding buffer, followed by the elution and isolation of high-affinity sequences using denaturing conditions such as urea, EDTA and heat. In selections using column affinity, stringency can be controlled by increasing the volume of wash buffer used [44]. Once the binding sequences have been isolated, they move on to PCR amplification and sequence analysis.

1.3.3 Amplification and Analysis

The iterative nature of the selection process is what has made SELEX such a powerful tool. It results in high affinity binding sequences dominating the pool, simplifying isolation and sequencing of relevant oligonucleotides. This, of course, would not be possible on such a large scale without Mullis' discovery of PCR amplification [25]. PCR is not a perfect system however and can have inherent biases. For example, the negative strand of DNA can compete with primers for hybridization, resulting in the suppression of rarer sequences [45]. It is therefore wise to optimize the PCR protocol before selection to minimize the number of PCR cycles required to obtain a sufficient quantity of products for the next round of selection. This also holds true for transcription and reverse transcription when selecting RNA aptamers as formation of certain RNA secondary structures can be inhibitory [46].

Another useful facet of amplification is the opportunity it presents to introduce a mutation mechanism if one wishes to do so. Mutant libraries are advantageous because they allow sampling of a larger total population, and in a focused manner. Selections with mutation mechanisms are more likely to produce sequences with higher affinities or enhanced catalytic properties as a result [47,48]. Thermostable DNA polymerases are generally fairly processive; they have an intrinsic error rate ranging from about $10^{-5} - 10^{-7}$ which is not sufficient to maintain a mutant population [49]. Common approaches include error-prone PCR and hypermutagenic PCR, which involve decreasing the fidelity of the

polymerase by using biased dNTP pools and adding Mn^{2+} to promote polymerization after mismatches [50,51]. Recombination methods can also be used if a very large degree of mutation is desired [47].

Progress of a selection is tracked by measuring the proportion of total nucleic acid that is retained by UV absorbance at 260nm. Only a small fraction of the total nucleic acid content will be recovered in the elution of initial rounds and those where stringency is increased. The selection is complete when essentially all nucleic acid is retained in the elution. The pool is amplified one last time and sequenced using next generation sequencing techniques. The sequences are then analyzed using computational techniques to determine one or a few consensus sequences. Minimization and optimization of the sequence is often desirable to highlight the active motif of the aptamer. This is generally done using kinetic-based methodologies. In addition, sequences may be re-selected with a partially random or 'doped' pool used specifically for binding site selection. This may result in an aptamer with improved properties [52].

1.3.4 Counter-selection

Many of the factors of aptamer selection discussed thus far impact primarily the binding affinity of the sequence obtained, but another advantageous property of aptamers is their high selectivity. In fact, aptamers with enantioselectivity for (R)-ibuprofen and (S)-ibuprofen have been isolated [53]. Selectivity can be tailored in the selection process by using counter-selection. This describes a cycle in which an alternate, undesired target is used, and sequences that bind said target are discarded. For example, when using a small molecule target such as sulforhodamine B is used, an analogue such as fluorescein may be used as the counter selection step [37]. A non-target peptide may be used when targeting a specific protein or an unwanted cell type in the case of cell-SELEX [42,54]. Negative selection is also used to ensure that the sequences present are binding the target as intended and eliminate false

positives that survive the selection by interacting with the matrix or binding by other undesirable mechanisms [55].

1.3.5 Modified Nucleic Acids

Nucleic acid aptamers hold many advantages over their protein antibody counterparts, such as ease of storage and modification, low immunogenicity, high-temperature stability and a wider range of potential targets [56,57]. One problem with aptamers, particularly those destined for *in vivo* applications, is that their in-host stability is relatively low due to nuclease susceptibility. RNA aptamers are degraded almost immediately, and DNA aptamers are degraded within a few hours of incubation in human serum [58]. Fortunately, one of the advantages of aptamers is that they are relatively easy to modify, and nuclease resistance can be straightforwardly conferred using various modification strategies before and/or after selection.

The most common modifications are substitutions of the 2' sugar position (Figure 1.4a-b). Initial modification efforts revolved around 2'-amino pyrimidines but they proved difficult to use during solid-phase chemical synthesis (Figure 1.4c) [59]. 2'-fluoro pyrimidines are often used in libraries due to their increased coupling efficiency (Figure 1.4d). The famed Macugen, the only FDA approved aptamer, has 2'-fluoro pyrimidine modifications [60]. 2'-methylhydroxy groups are favoured as post-SELEX modifications due to their cost efficiency and in-cell role as a ubiquitous post-translational modification (Figure 1.4e). Locked nucleic acids are another highly nuclease-resistant modification that consists of a second ring structure being created by connecting the 2' carbon and the 4' carbon of the sugar (Figure 1.4f).

Also popular are 4'-thiol sugars (Figure 1.4g) and various modifications of the inter-nucleotide linkage. Examples include phosphorothioate, methylphosphanate, 5'- α -P-borano and thiazole linkages

(Figure 1.4h-k) [59,61]. Another advantage of modified nucleic acids is that they may diversify aptamer libraries; they may be able to adopt previously unseen conformations or bind novel targets.

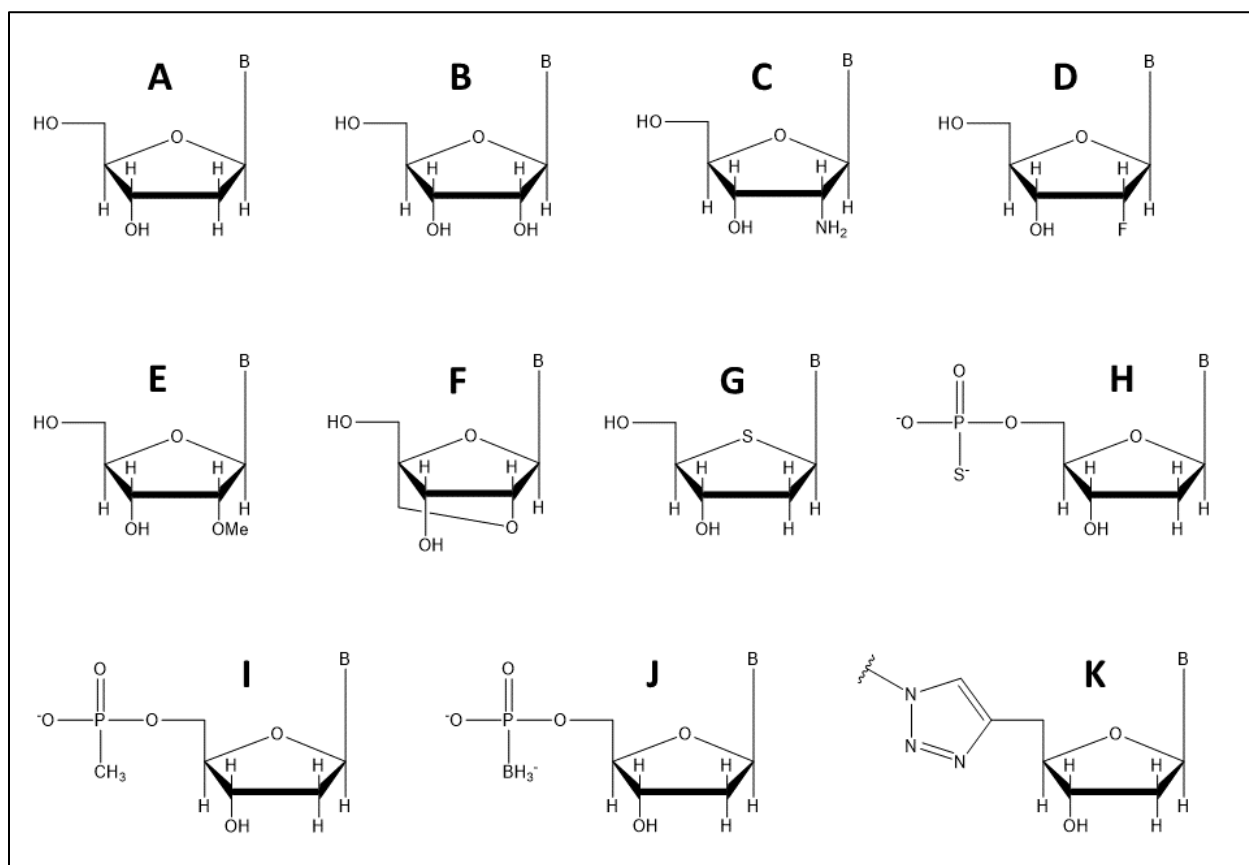


Figure 1.4: Types of nucleic acid modifications used to confer nuclease resistance in aptamers.

A) 2'-H (DNA), B) 2'-OH (RNA), C) 2'-NH₂, D) 2'-F, E) 2'-OMe, F) 4'-thiol, G) locked nucleic acid (LNA), H) phosphorothioate, I) methylphosphanate, J) 5'- α -P-borano, K) thiazole.

1.4 Fluorophore-Binding Aptamers

1.4.1 Applications of Fluorophore-Binding Aptamers

As discussed in section 1.2, some of the first proof of concept RNA aptamers selected by Andrew Ellington and Jack Szostak used organic dyes as targets due to their similarity to biological cofactors [28]. They used the same group of dyes to select some of the first proof of concept DNA

aptamers [62]. The RNA world concept was prominent at the time, and scientists were interested in gauging the catalytic complexity of nucleic acids. In addition, the inherent advantages of having a detectable signal built into the system was not lost on researchers.

Several years later, Jack Szostak and Charles Wilson selected in parallel a DNA aptamer and an RNA aptamer for sulforhodamine B (SR) (Figure 1.5a) [37,63]. With the DNA aptamer, they were interested in finding an aptamer that could catalyze the redox reaction of the colorless compound dihydrotetramethylrosamine to the fluorescent product tetramethylrosamine (Figure 1.6). These compounds oxidize spontaneously, so they instead used the related dye, SR, as the target for their selection. The clones they obtained, however, had only weak redox activity, and this work was abandoned.

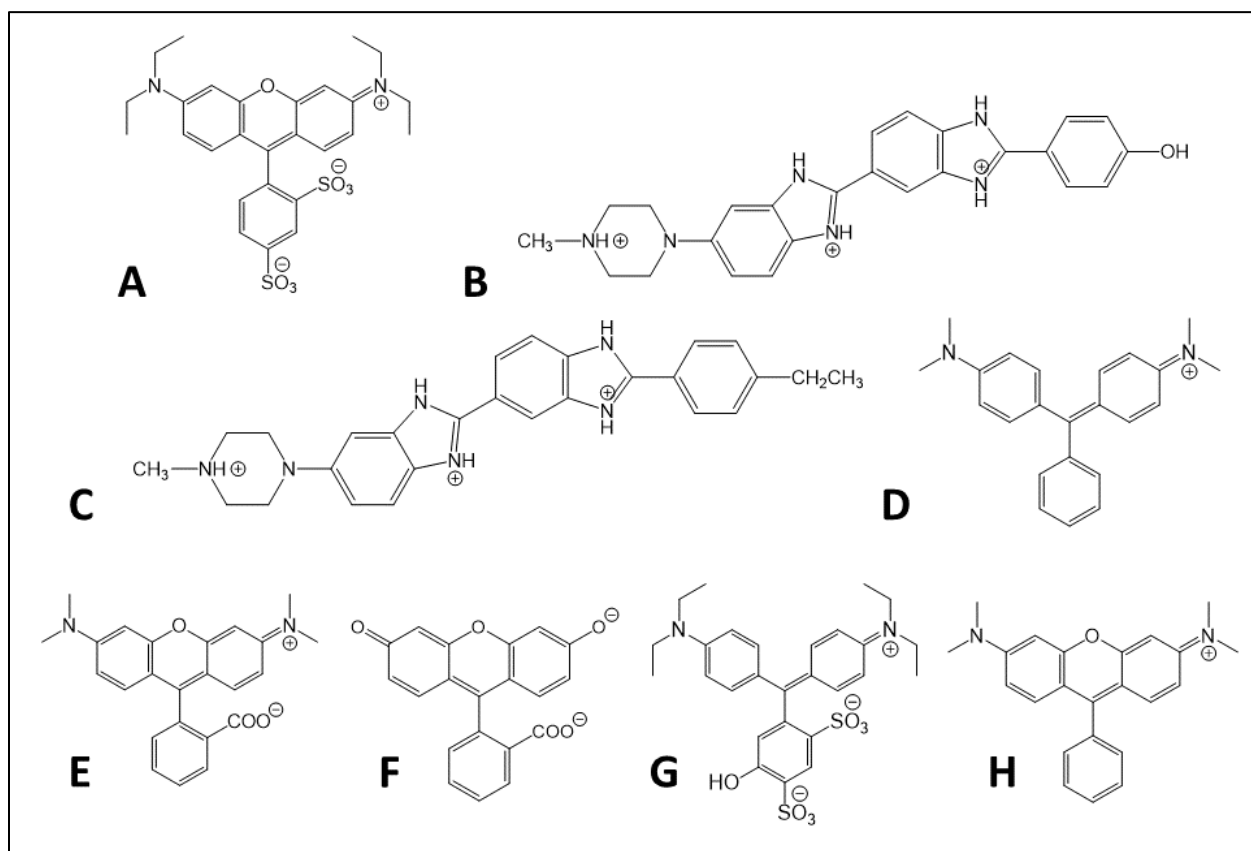


Figure 1.5: Relevant ligands of fluorophore-binding aptamers.

A) sulforhodamine B, B) Hoechst 33258, C) Hoechst 33242, D) malachite green, E) tetramethylrhodamine, F) fluorescein, G) Patent blue V, H) tetramethylrosamine

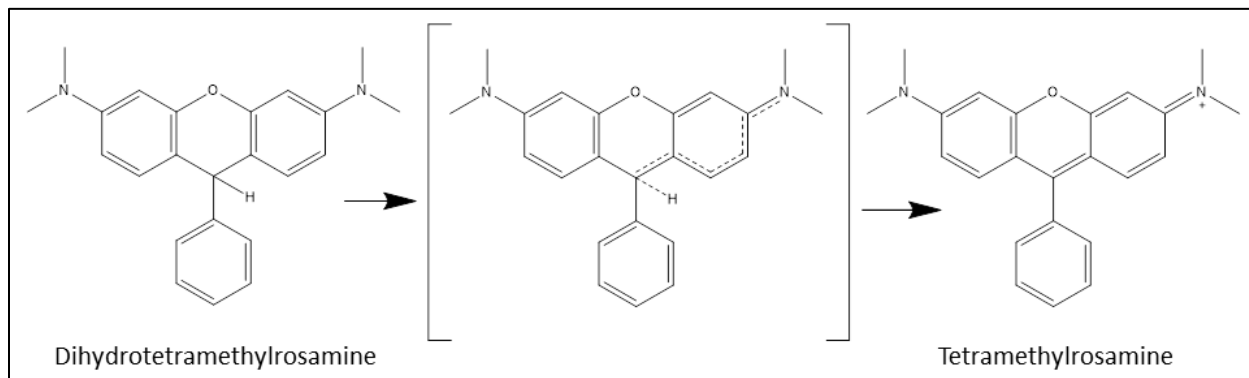


Figure 1.6: Desired redox reaction of the SRB-2 DNA aptamer.

Gold and Wilson attempted to find an aptamer that could catalyze the redox reaction of the colorless compound dihydrotetramethylrosamine to the fluorescent product tetramethylrosamine. This figure shows the general scheme for said reaction.

1.4.1.1 Controlling Gene Expression

Thus far, aptamers have only been discussed in a synthetic capacity. Of course, as it turns out, biological systems have long been using their own version of aptamers, which are generally found in riboswitches. Riboswitches are nucleic acid sequences most commonly found in the 5' UTR region of mRNAs which control gene expression in a *cis* fashion through small molecule binding [64]. Riboswitches consist of two main domains, a gene expression platform and a metabolite-binding (aptamer) domain. Riboswitches function by forming secondary structure elements that impede protein function. The most common mechanisms include rho-independent transcription termination and translation inhibition by sequestering of the ribosome binding site.

The suspicion of the existence of such elements has been around for some time, as proposed by Larry Gold *et al.* in 1984 [65], but substantial proof was lacking until 2002 when riboswitches were finally characterized by several research groups working on several different riboswitches [66–70]. Interest in potential gene regulation by aptamers was heightened in the late '90s after the establishment of SELEX [27,28] and further studies showing modulation of gene expression by small molecules [71–73]. Initial studies using synthetic riboswitches (aptamers) to control gene expression were actually done in 1998, a few years earlier than the 'discovery' of riboswitches in 2002.

In these studies, two aptamers that could specifically bind both the dye Hoechst 33258 (Figure 1.5b) and the related drug Hoechst 33342 (Figure 1.5c) were selected. They were inserted in tandem into a mammalian β -galactosidase expression plasmid which was then transfected into Chinese hamster ovary cells. These studies showed that Hoechst 33342 was able to dose-dependently inhibit β -galactosidase activity by over 90% [74].

The malachite green aptamer (MGA) is among the most extensively studied aptamers and has proven to be dynamic in its applications [75–79]. This aptamer was originally selected to improve upon the chromophore-assisted laser inactivation (CALI) method developed by Jay and Keshishian [80], wherein antibodies are covalently modified with the fluorescent dye malachite green (MG) (Figure 1.5d). Malachite green forms destructive hydroxyl radicals upon irradiation by laser at 630nm, resulting in the inactivation of the antibody and its target. Therefore, gene products can be targeted and inactivated by controlling the laser (Figure 1.7). The malachite green aptamer has a single nucleotide bulge that is cleaved upon exposure to the hydroxyl radical generated by MG. As a result, when MGA is included in the 5' UTR of a gene of interest, the laser can be used in an analogous fashion to downregulate the gene product by reducing transcript stability and translatability [75]. Furthermore, this system was used to demonstrate that cell-cycle control could be rendered dependent on MG in *S.cerevisiae* [81].

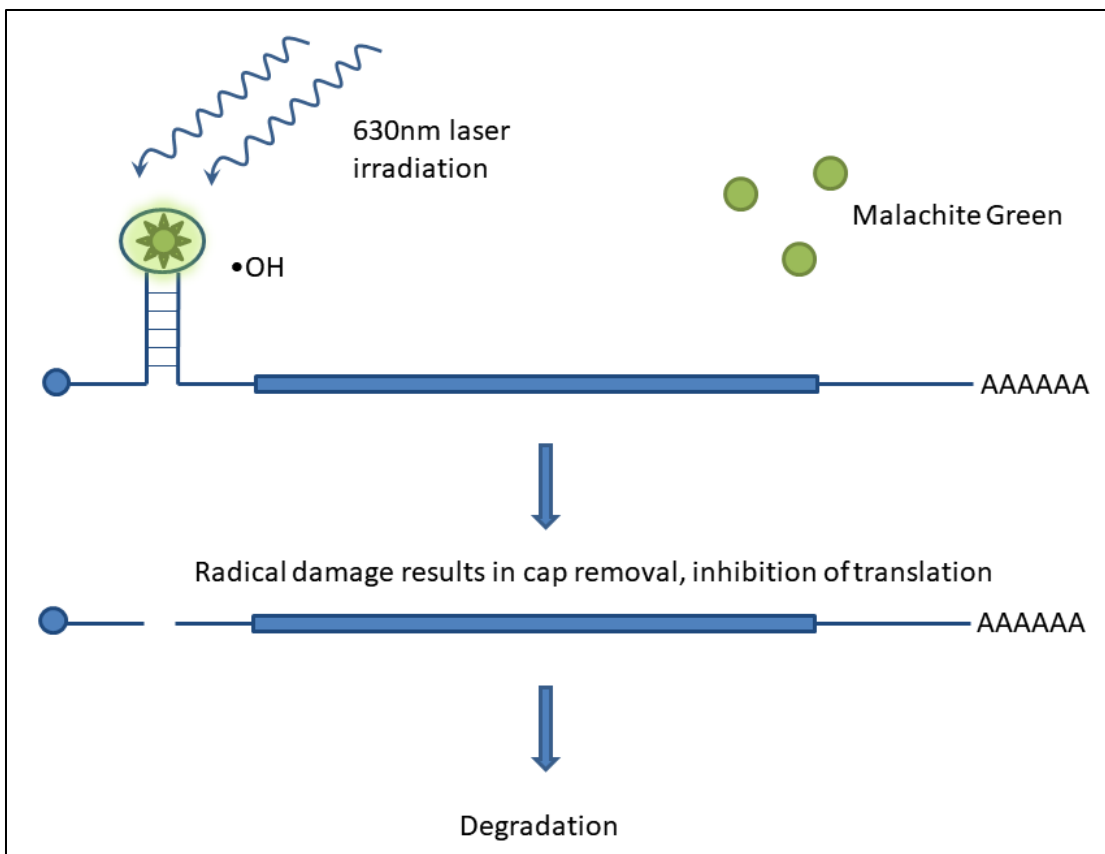


Figure 1.7: Site-specific inactivation of mRNA transcripts by chromophore assisted laser inactivation.

Malachite green forms hydroxyl radicals when exposed to a laser of the appropriate wavelength. Due to close proximity and susceptibility, these radicals attack and cleave a single nucleotide bulge in MGA resulting in cap removal and translation inhibition.

In 2010, aptamers were selected for the dye tetramethylrhodamine (Figure 1.5e) with high-affinity and low magnesium dependence for optimal *in vivo* function, but no further works involving gene expression were conducted. Synthetic riboswitches have remained an area of interest for researchers after the CALI studies of MG, but studies have mostly turned to aptamers with non-fluorogenic ligands that do not require laser mediation such as theophylline, FMN and neomycin [82–84].

1.4.1.2 mRNA Imaging

In 1998, in parallel with their work involving the sulforhodamine B DNA aptamer, Wilson and Szostak also selected an RNA aptamer for sulforhodamine B [37]. The original goal of selecting this RNA aptamer, SRB-2, was to develop a tool for fluorescent RNA labelling analogous to green fluorescent protein (GFP) and its derivatives in proteins. The methodology here involved inserting the DNA template sequence for SRB-2 on the end of a gene of interest, resulting in the fluorescent tagging of that transcript in the presence of the ligand (Figure 1.8). Wilson and Szostak recognized that this ability to fluorescently label nucleic acids *in vivo* or *in vitro* could open up many possible applications. They also used the clones obtained from the SRB-2 selection as a basis to select an aptamer for the similar dye fluorescein (Figure 1.5f), which has a similar structure but a different color than SR. They performed a proof of concept experiment *in vitro* that showed the two aptamers were able to discriminate between the two dyes which allowed localization of the aptamers to be observed when the dyes were anchored to separate beads [37]. This highlighted the potential for future multiple labelling as well as Forster resonance energy transfer (FRET) type experiments; though contemporary FRET sensors generally utilize aptamers that are labelled with fluorophores and/or quenchers rather than aptamers that bind them directly. Despite sulforhodamine B and fluorescein having a significantly greater fluorescence enhancement than biological cofactors such as tryptophan and flavin mononucleotide, background fluorescence was still far too great for biological applications [37,85–87].

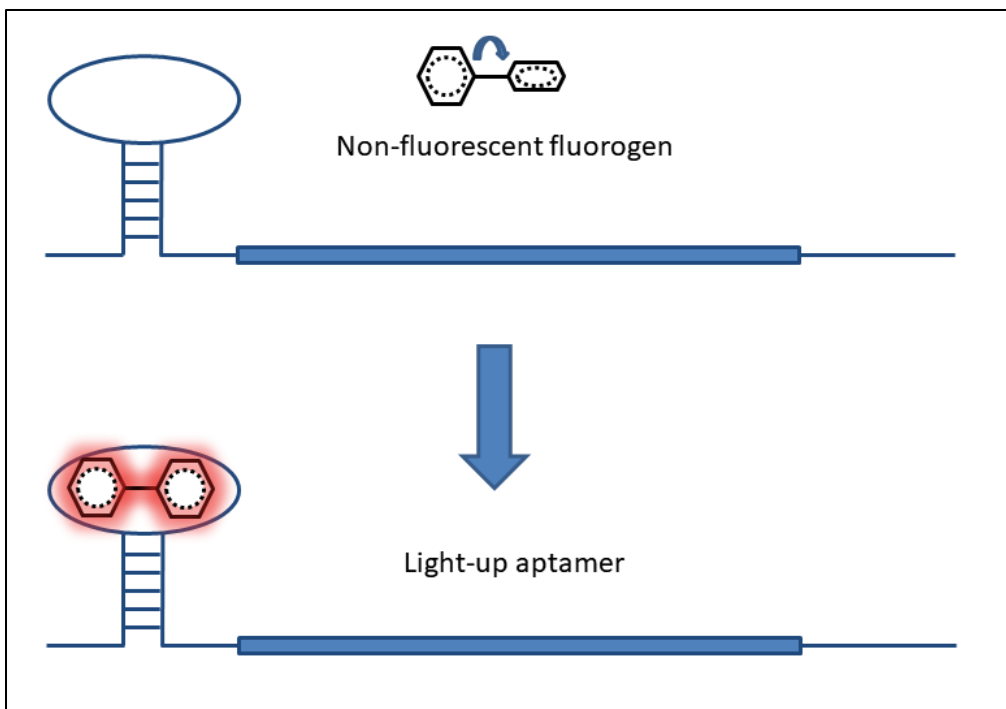


Figure 1.8: Structural basis of light-up aptamers for intracellular imaging.

In solution, fluorogens are non-fluorescent, but when stabilized by stacking interactions with the bases of an aptamer adopt a fluorescent conformation.

More recent studies involving the SRB-2 aptamer have utilized ligand-quencher constructs to increase the fluorescence enhancement of sulforhodamine B upon binding by quenching its intrinsic fluorescence with dinitroaniline [88,89]. This approach yielded success as its brightness surpassed even enhanced GFP (eGFP), and 105-fold enhancement was observed. This is impressive, but other systems with different fluorophores have been characterized since then that have generated even greater fluorescence enhancements.

Studies by Babendure *et al.* with the SRB-2 and MGA aptamers showed that significantly larger fluorescence enhancements are observed with triphenylmethane dyes as opposed to planar xanthene-type dyes [76]. They observed that the triphenylmethane dyes such as patent blue V (PBV) (Figure 1.5g) and MG had a lower intrinsic quantum yield and higher quantum yield when bound to the aptamer,

which they surmised was a result of the locking of the rotationally mobile dyes into planar conformations more similar to the xanthene dyes tetramethylrosamine (TMR) (Figure 1.5h) and sulforhodamine B (SR). Malachite green had a 2360-fold fluorescence enhancement upon binding, significantly surpassing SRB-2 and GFP. This could be further increased by using tandem repeats of MGA [76]. Efforts to optimize the enhancement of aptamer-based RNA imaging has more recently turned to the chromophore of GFP, 4-Hydroxybenzylidene (HBI) (Figure 1.9a) and its structural analogues.

The first works with HBI derivatives involved the selection of an aptamer for DMHBI (Figure 1.9b). It was then determined that the derivative 3,5-difluoro-4-hydroxybenzylidene imidazolinone (DFHBI) (Figure 1.9c) was significantly brighter upon aptamer binding (80% as bright as GFP, DMHBI was only 12% as bright), and this system was termed spinach [90]. Spinach has since spawned several additional light-up aptamer systems with high fluorescence enhancement and of varying color. These include aptamers that bind other GFP mimics, such as broccoli [91] which binds DFHBI-1T (Figure 1.9d), and corn [92] which binds 3,5-difluoro-4-hydroxybenzylidene imidazolinone-2-oxime (DFHO) (Figure 1.9e) and aptamers that bind other fluorogenic dyes such as the mango aptamer [93], which binds thiazole orange and its derivatives. These aptamers have all been sequence optimized and used in various live-cell imaging experiments of different types of RNA [94–102]. In addition, the corn aptamer has been used for quantification of transcription [92] and the mango aptamer has been used to simultaneously fluorescently label and purify biologically relevant RNAs [93].

Although these GFP-based selection targets were able to increase fluorescence enhancement of these systems drastically, there is always room for improvement. These RNA aptamers have been successful in imaging highly abundant, non-translated sequences such as tRNA and rRNA [90,91,103]. However, fluorescence enhancement of these aptamers is insufficient for mRNA imaging owing to its low cellular abundance [104]. One method implemented to improve fluorescence enhancement is the use of fluorescent protein multimers [105,106]. This strategy has also been adapted for the spinach

aptamer, where tandem repeats showed up to a ~17-fold increase in fluorescence enhancement [107]. Tandem repeats of fluorescent proteins can cause some issues, especially when used in organisms with proficient homologous recombination mechanisms [108]. Fortunately, transcription seems to be relatively unaffected by the presence of spinach aptamer repeats [107]. However, the G-quadruplex nature of the spinach aptamer may cause another unforeseen problem with this type of system. Although the ability of the G-quadruplex to stabilize fluorogens is likely responsible for the large fluorescence enhancement of HBI derivatives, it has been suggested that RNA G-quadruplex domains may be kept globally unfolded in eukaryotic cells. This would, of course, limit the use of such aptamers for *in vivo* imaging applications.

This potential hurdle highlights the importance of studying ways to improve the fluorescence enhancement of non- G-quadruplex containing aptamers such as malachite green and SRB-2. In fact, a variation of the tandem strategy was recently theorized using SRB-2 [109]. In these studies, the authors designed a sulforhodamine B dimer called Gemini-561 and selected a linker for an SRB-2 dimer they termed o-Coral. This fluorogenic aptamer showed remarkable fluorescence enhancement, surpassing even the aforementioned HBI derivatives. It also had noticeably higher photostability than the mango, corn and broccoli aptamers which could extend the useful timeframe of mRNA imaging experiments [109].

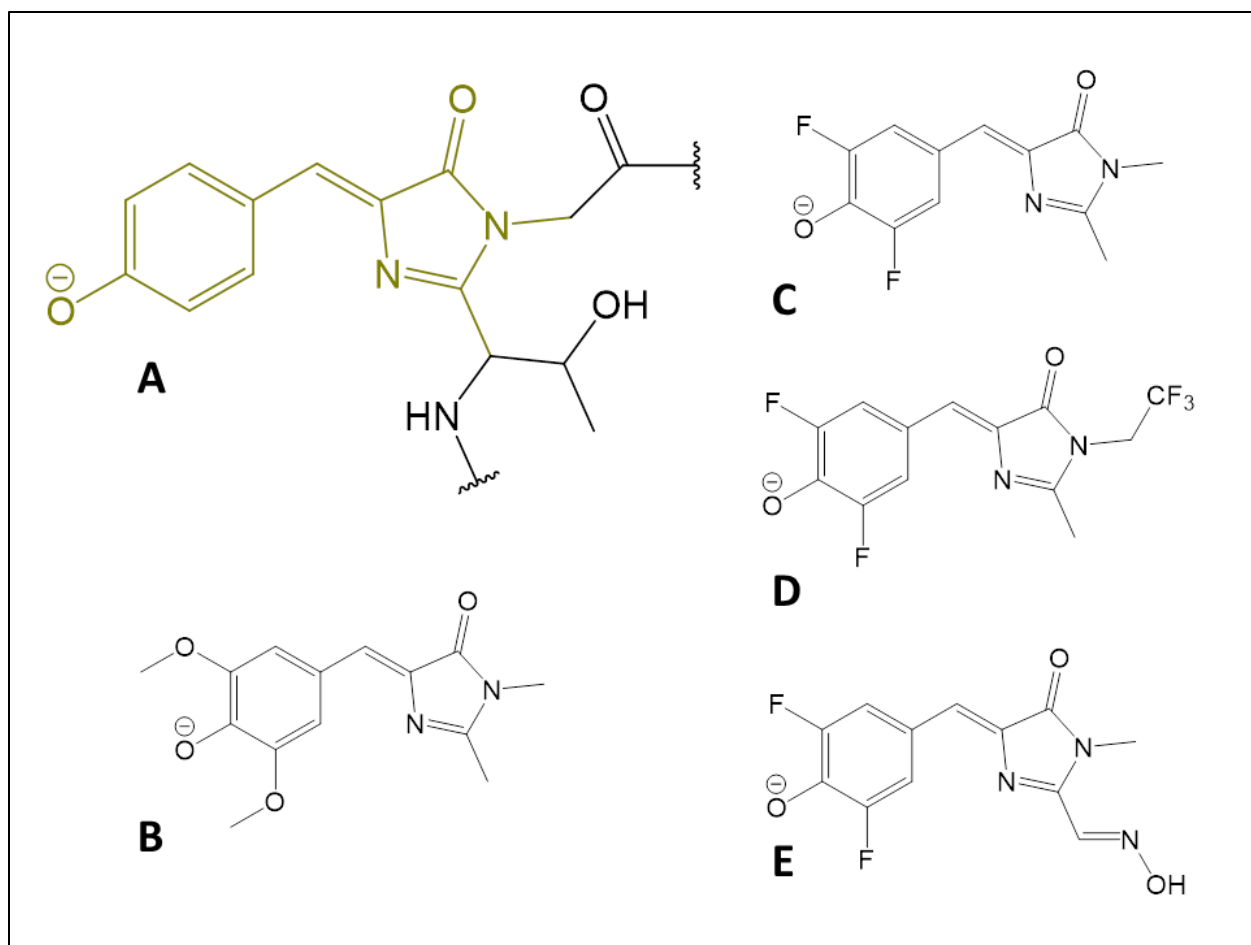


Figure 1.9: Chemical structure of the GFP fluorophore and its RNA aptamer mimics.

A) 4-Hydroxybenzylidene (HBI), GFP's fluorophore. B) 3,5-dimethoxy-4-hydroxybenzylidene imidazolinone (DMHBI), selection target of the spinach aptamer. C) 3,5-difluoro-4-hydroxybenzylidene imidazolinone (DFHBI), a brighter analogue of DMHBI and ligand for the spinach aptamer. D) 3,5-difluoro-4-hydroxybenzylidene imidazolinone 1-trifluoroethyl (DFHBI-1T), an even brighter modification of DFHBI. E) 3,5-difluoro-4-hydroxybenzylidene imidazolinone-2-oxime (DFHO), selection target of the corn aptamer.

1.4.1.3 Metabolite sensing

Another application that fluorophore binding aptamers are well suited for is biosensing. Aptamer-based fluorescent biosensors are most often FRET-based sensors that feature either a signal “turn-on” or “turn-off” mode involving a conformational change that alters the proximity of a fluorophore and a quencher [110–112]. That being said, there are several examples of innovative constructs that directly use fluorophore-binding aptamers.

The first attempt at creating such a construct was work done by Stojanovic and Kolpashchikov with the malachite green aptamer in 2004 [77]. They created modular aptameric sensors using recognition domains consisting of a binding site for ATP, FMN or theophylline and a signalling domain that had a binding site for malachite green (Figure 1.10a). All binding sites were based on existing aptamer sequences. The motivation for this work was to create a sensor for these biological cofactors that could be used for intracellular imaging. In contemporary research, the sensing and imaging of metabolite trafficking has become more popular than imaging RNA trafficking *in vivo* [113].

Shortly after the development of these sensors, Kolpashchikov created a binary malachite green aptamer to be used as a probe for fluorescently reporting the presence of a specific nucleic acid sequence [78]. This sensor consisted of two RNA strands, each consisting of a fragment of the malachite green aptamer and a fragment of the complement RNA sequence to the DNA target sequence (Figure 1.10b). In the presence of the target sequence and MG, the two strands would bind with their respective RNA complement forming the complete malachite green aptamer, which would then bind the dye resulting in a fluorescence enhancement.

In 2010, Lu and Xu combined aspects of these two approaches to make a sensor for adenosine [79]. Their sensor consisted of a construct with aptamers for MG and adenosine fused together, and a bridging strand that was constructed such that it could only be released upon binding of adenosine to

the adenosine aptamer. Once the bridging strand was released, MGA would be free to bind MG, resulting in a fluorescent signal (Figure 1.10c).

One of the problems with using MG for intracellular imaging is that its propensity for forming radicals results in it being relatively cytotoxic. Therefore, more recent studies in this field use aptamers that bind to HBI derivatives such as the spinach aptamer. Interest in this field has mostly pivoted towards the development of sensors for use in intracellular metabolite imaging so many of the constructs created in recent years bear a striking similarity to Stojanovic and Kolpashchikov's original work with the malachite green aptamer. The spinach aptamer alone has since been used to image a variety of metabolites in cells, including adenosine, ADP, guanine, SAM, guanine, GTP, cyclic-di-GMP and c-AMP-GMP [114–116].

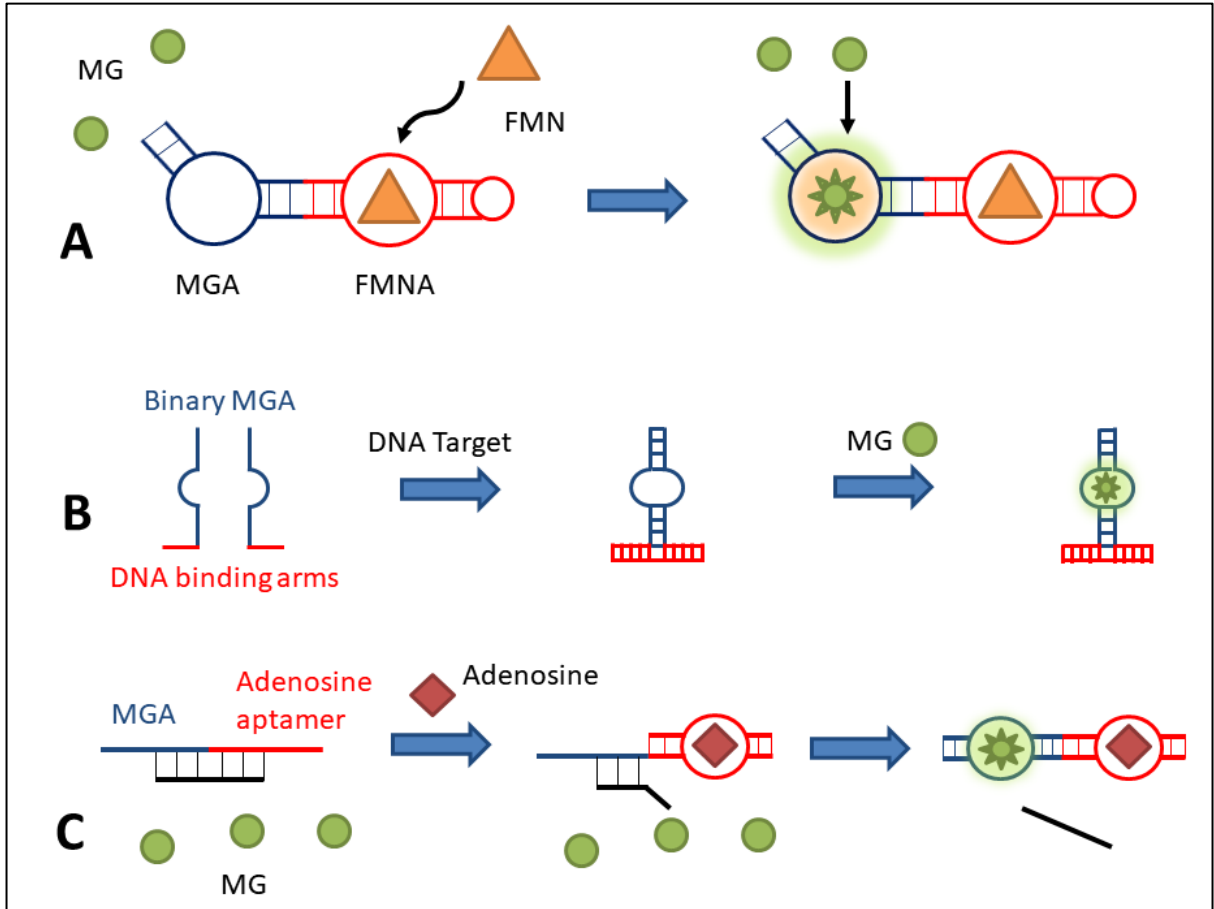


Figure 1.10: Metabolite sensors based on the malachite green aptamer.

A) A modular aptamer construct for metabolite detection was devised by fusing binding motifs from the malachite green aptamer (MGA) and several metabolite binding aptamers such as the flavin mononucleotide aptamer. A tangible increase in fluorescence was observed in the presence of FMN [76]. B) Another sensing approach involved a binary aptamer-based on the sequence of MGA. Each strand also had an arm whose sequence was the complement of a target DNA sequence. In the presence of the target sequence, both arms bind to a target sequence and the complete MGA forms, resulting in MG binding and fluorescence enhancement [77]. C) An improvement of the methodology of a modular aptameric sensor that utilizes a bridging strand to increase fluorescence enhancement. This strand

prevents binding of MGA until adenosine has already bound to the adenosine aptamer, which releases the bridging strand and allows MGA to bind MG.

1.4.2 Structural Studies of Fluorophore Binding Aptamers

One characteristic that all of the aforementioned applications have in common is that their rational development relies heavily on fundamental structural knowledge of the aptamer's binding complex. For example, if a researcher wishes to design a riboswitch to control transcription of a gene with a fluorophore binding aptamer, they must ensure that the bound apo-form folds sufficiently to block transcription and the holo-form does not. Fortunately, a distinguishing feature of aptamers is their adaptive binding [117–122]. That is, an aptamer in the absence of its ligand (holo-form) tends to be relatively unstructured but when bound to its ligand (apo-form) forms structured binding motifs. Therefore, aside from characterizing the fundamental binding properties of nucleic acids, obtaining crystal or solution structures for these systems is important because structural knowledge of a system is useful for implementation in various applications. It is no coincidence then that the handful of fluorophore-binding aptamer structures that have been determined mostly coincide with the previously discussed aptamers that have been used in real-world applications.

There are six fluorophore-binding aptamers that have at least one structure available, and each has contributed to the understanding of ligand-induced folding in aptamer-small molecule systems. The first system successfully characterized was the malachite green aptamer, for which a crystal structure was determined when bound to the high-affinity ligand analogue tetramethylrosamine (TMR) (Figure 1.11a) [123]. Shortly thereafter, a solution structure was determined for MGA bound to its original selection target MG (Figure 1.11b) [124]. Comparison of these structures led to insights regarding the ability of the binding pocket to adapt to both planar and non-planar ligands.

The spinach aptamer was successfully crystallized years later in 2014 and revealed for the first time a G-quadruplex motif involved in the binding of a fluorophore (Figure 1.11c) [125]. They observed that this G-quadruplex motif was well suited to bind and induce fluorescence in an organic dye as it was forced into a locked planar conformation upon binding. A crystal structure was also obtained for iSpinach, the result of further selection to optimize the fluorescence and folding properties of the original sequence [126]. This structure showed that mutations in iSpinach resulted in a more stable and compact binding motif than the original sequence and conferred advantageous properties such as thermal stability and reduced salt dependence.

Around the same time as iSpinach, several similarly fluorogenic systems of varying color were also characterized by x-ray crystallography. The mango and corn aptamers were revealed to utilize G-quadruplex motifs as well, cementing the importance of these structures in light-up aptamer systems [96,127]. The fluorescence properties of the mango aptamer were also optimized in a similar fashion to iSpinach and the structure of this mutant was recently obtained [128]. Interestingly, the dimethylindole red (DIR2s) aptamer, which binds the cyanin dyes dimethylindole red and oxazole thiazole blue, was determined to be the first fluorophore binding aptamer that did not include a G-quadruplex motif or a base quadruple such as that observed in MGA. Instead, it relies on a base triple as a stacking platform which diversifies the known range of fluorophore activation by RNA.

Most recently, a structure for a tetramethylrhodamine aptamer (TMR3) was determined by NMR spectroscopy [129]. TMR3 was found to share some structural similarities with the malachite green aptamer, but as shown in Figure 1.11d, it has a base triple similar to DIR2s. Interestingly, the fluorescence of the ligand is quenched upon binding in contrast to all the other systems that have been characterized. The ligand binds with similar base stacking and hydrogen bonding patterns to other fluorogenic RNA aptamers, so it is likely that this low fluorescence conformation is a result of a pKa shift, possibly of one of the carboxyl groups.

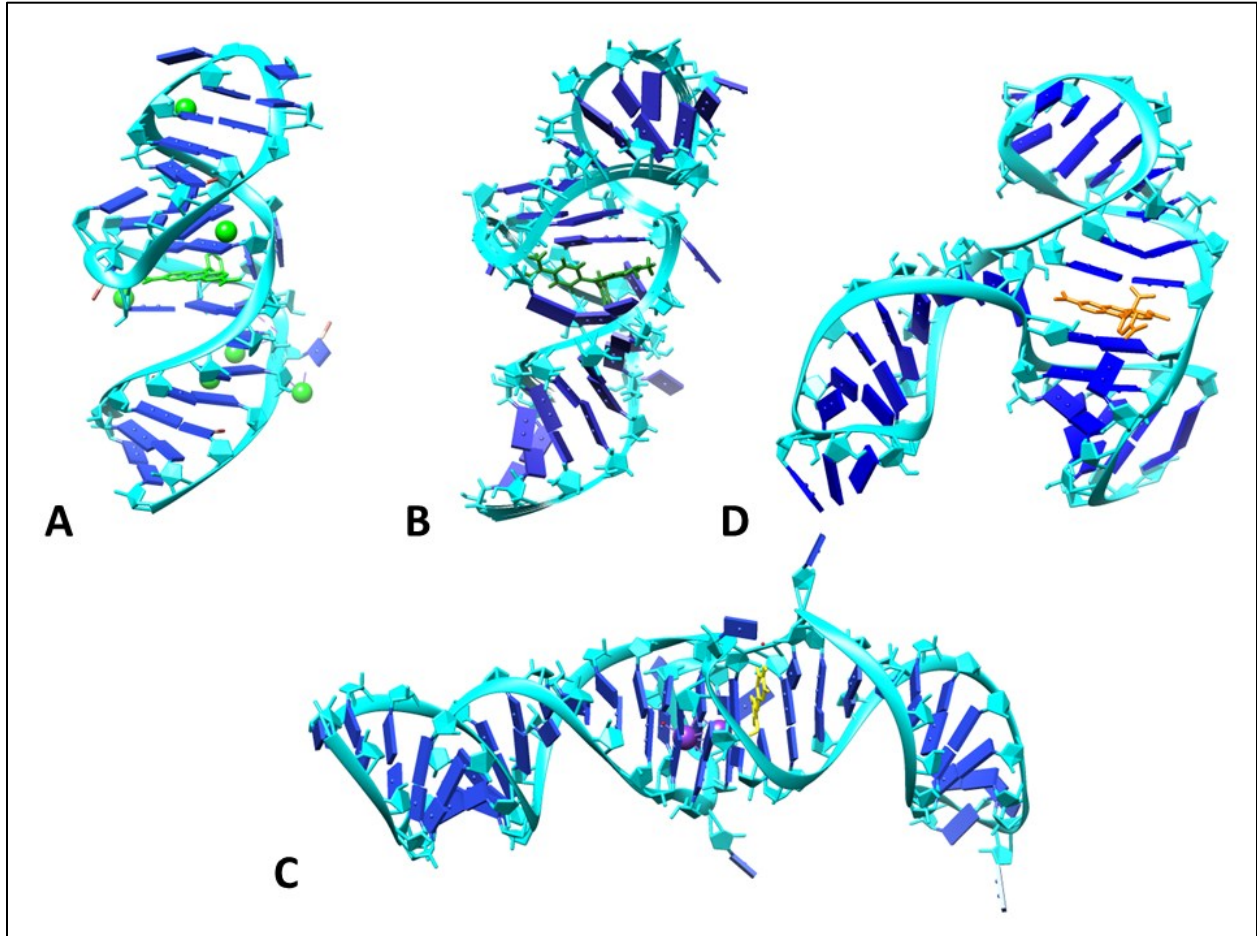


Figure 1.11: Structures of selected fluorophore binding aptamers.

A) Crystal structure of MGA bound to TMR (1F1T) [123]. B) Solution structure of MGA bound to MG (1Q8N) [124]. C) Crystal structure of the iSpinach aptamer bound to DFHBI (5OB3) [126]. D) Solution structure of TMR3 bound to tetramethylrhodamine (6GZK) [129]. Images were created with the modelling software UCSF Chimera, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco [130].

1.5 Catalytic Nucleic Acids

1.5.1 Ribozyme Discovery

In 1982, Thomas Cech was the first to discover RNA with catalytic properties thanks to his work on the *Tetrahymena* intervening sequence RNA (IVS RNA), which is capable of self-splicing [16]. Cech shared the Nobel Prize for Chemistry in 1989 with Sydney Altman, who characterized the catalytic properties of the RNase P ribozyme, which also cleaves RNA, in 1983 [15]. Several biologically relevant ribozymes have since been discovered, including the Varkud satellite (VS) ribozyme [131], the hammerhead ribozyme [132,133], the hairpin ribozyme [134], the ribosome [135] and the spliceosome [136], all of which catalyze cleavage and/or ligation of the nucleic acid backbone. Another interesting ribozyme discovery was GImS, a self-cleaving ribozyme that acts as a riboswitch [137]. More recently, synthetic ribozymes that can perform a variety of chemical reactions have been developed [138–141] as catalytic nucleic acids can be obtained through SELEX, just like aptamers.

1.5.2 Selection of Catalytic Nucleic Acids

Gold and Szostak are usually credited with the invention of SELEX, but the other, often underappreciated, pioneer in this field is Gerald Joyce. Joyce and Robertson also independently came up with an *in vitro* selection method in 1990 [142]. The main difference in this work was that rather than selecting aptamers that bind small molecules or proteins, Joyce was interested in selecting catalytic RNAs. Ellington and Szostak did ponder the possibility of selecting catalytic RNAs in their work earlier in the year [28], but Joyce was the first to actually do it. Using *in vitro* selection, Joyce and Robertson were able to isolate a mutant of the self-splicing *Tetrahymena* ribozyme, which more efficiently cleaved its substrate than the wildtype sequence. Their iteration of *in vitro* selection was again striking similar to Gold's and Szostak's procedures, and contemporary selections of catalytic nucleic acids typically use the same general SELEX procedure as aptamers. Due to the nature of these molecules, there are some

differences such as immobilization not being implemented. In addition, cleaved nucleotides must be separated from intact sequences, which is usually achieved through the use of either polyacrylamide gel electrophoresis (PAGE) or biological tags such as biotin [143].

1.5.3 DNAzymes

Although not found in nature, DNAzymes can also be selected through SELEX. The first DNAzyme, which could catalyze the cleavage of RNA, was isolated by Ronald Breaker and Gerald Joyce in 1994 [144]. DNAzymes have since been created with the capacity to perform many types of catalysis, including nucleic acid cleavage and ligation [144–147], DNA modifications such as phosphorylation [148], adenylation [149] and deglycosylation [150] and even a Diels-Alder reaction [151].

DNAzymes have been used in several applications, including DNAzyme therapeutics, of which a few have participated in clinical trials. These are generally treatments involving the inactivation of mRNA sequences, downregulating the expression of disease-related proteins [152]. The most common class of DNAzymes are ribonucleases, which catalyze the cleavage of an RNA phosphodiester bond via a transesterification reaction. These self-cleaving DNAs generally require specific metal cofactors which has resulted in metal sensing being one of the most studied applications of DNAzymes. For use in such commercial applications, DNAzymes provide a couple of key advantages over ribozymes, including stability, cost efficiency and ease of chemical modification [153].

1.5.3 Structural and Functional Importance of Metal Ions

Due to the polyanionic nature of nucleic acids, it is impossible to discuss their structure and function without also considering metal ions. The negative charges need to be shielded by cations, i.e. metals, to allow the formation of stable secondary structure. Due to their stabilizing effect on nucleic acid structures, metal ions are imperative for the biological functions of DNA and RNA: storage, reading

and homologous recombination of genetic information [154]. Metal-binding also has a significant impact on the stability and reactivity of various functional groups in nucleic acids, especially particular heteroatoms in the bases [155]. They also mediate the ability of nucleic acids to interact with proteins and other ligands [156].

In specific sequences, the presence of metals can lead to several distinct structural motifs, including triplexes (Figure 1.12a), helical junctions (Figure 1.12b), quadruplexes (Figure 1.12c) and the loop E motif (Figure 1.12d) [157–161]. These types of motifs are all dependent on the presence and concentration of common intracellular metals, including Na^+ , K^+ and Mg^{2+} . Some metals can also bind specifically to certain base pairs, such as in thymine- Hg^{2+} -thymine [162] and cytosine- Ag^+ -cytosine [163].

As the diverse roles of metal ions listed may suggest, under correct conditions, DNA binds to certain metals in a highly selective manner. This useful trait is often exploited in biotechnological applications such as DNAzyme-based biosensing.

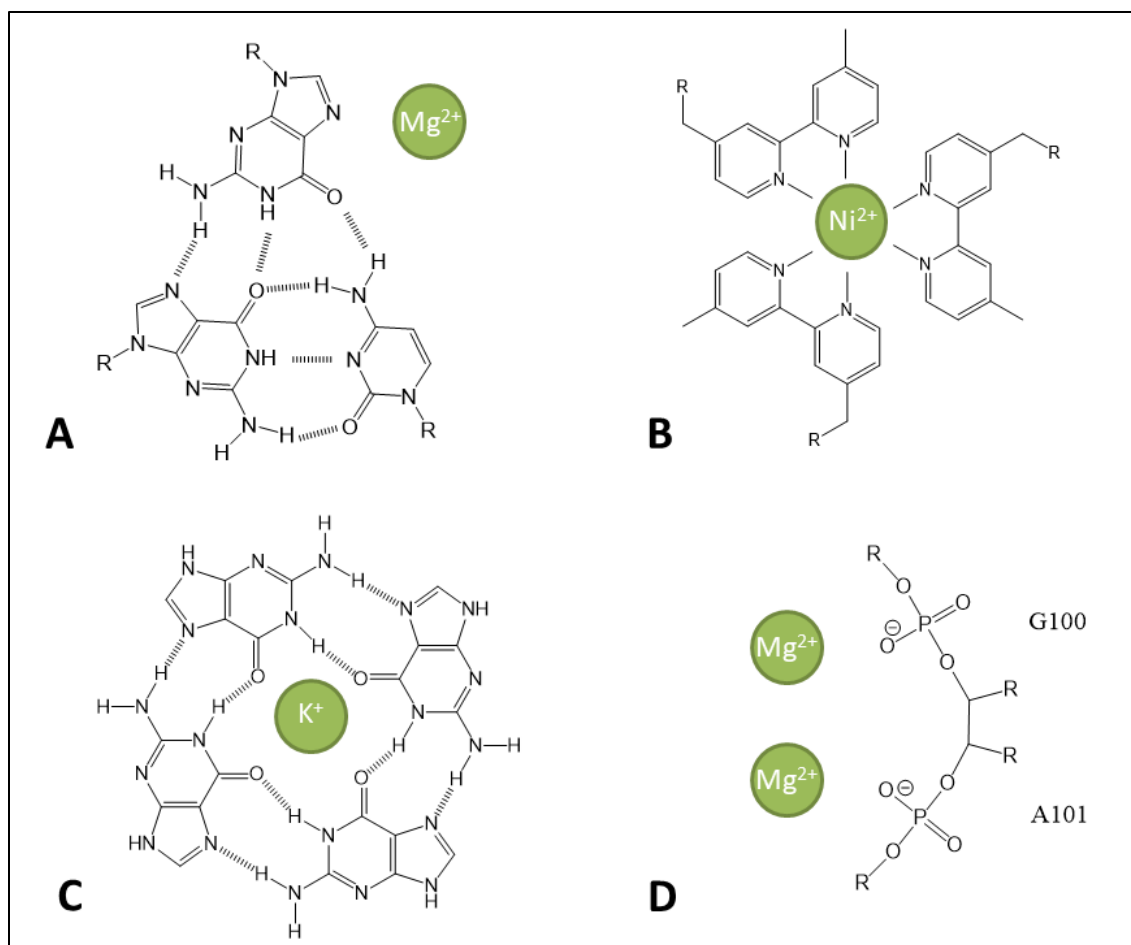


Figure 1.12: Metal coordination to several nucleic acid structures.

A) Penta-hydrated magnesium stabilizes triplex formation by binding to the N7 of guanine in the tertiary strand. B) Nickel (II) stabilizes a three-way junction by complexing with three bipyrimidines. C) Potassium promotes the formation of G-quadruplex in guanine-rich sequences. D) A unique binuclear magnesium cluster observed in the Loop E motif.

1.5.4 Metal Biosensors

The ability to specifically detect metal ions is important for numerous reasons. Metals have great physiological importance, such as their aforementioned roles with respect to nucleic acids. They are crucial for cell signalling and transport [164,165], and metalloproteins are thought to make up about half of all known proteins [166]. Maintaining proper cellular levels of metals is therefore important for

human health, and the ability to reliably test these levels is required for diagnosis of metal deficiencies and other related ailments. Metals are also important in society as anthropological activities such as mining, smelting, electronic waste disposal and domestic and agricultural use of metal containing compounds continuously results in significant pollution of heavy metals [167–170]. Heavy metals are a known hazard to human health and the environment, therefore the ability to test drinking water and other potential sources for contamination is imperative. DNA-based methodologies have been used to detect a large percentage of metals from the periodic table [171]. Different strategies are implemented depending on the metal of interest and the ways it is known to interact with DNA. For example, Na^+ has been detected using an aptamer sequence, K^+ is generally detected using G-quadruplex based sequences and metals such as Pb^{2+} use metal-dependent self-cleaving DNAzymes.

Biosensors generally consist of a biomolecular element, which in this case is DNA and a signal transduction element. Most signal transduction elements in DNAzyme sensors are either fluorescence, colorimetric or electrochemical based. Fluorescent detection is often achieved using fluorophore/quencher pairs [172]. Nanomaterials such as quantum dots (QD), gold nanoparticles (GNPs), graphene oxide and carbon nanotubes have proven useful as vectors for these labelling pairs [173–175]. Additionally, the change in color between dispersed and aggregated GNPs is the basis of several colorimetric biosensors [176–178]. Several unique strategies have been used to develop electrochemical sensors with impressive detection limits for metals such as Pb^{2+} [179–181].

1.6 Contents of the Thesis

In order to develop potential applications such as biosensors and cell imaging, we must have a thorough understanding of how aptamers bind to their targets. This allows chemical modification and sequence optimization of aptamer systems for their use in previously discussed commercial applications. This is where work on characterizing aptamer structure and binding kinetics comes into

play. Chapters 2,3 and 4 feature work on the aptamer SRB-2, which is an interesting one due to its ligand promiscuity and overall negative charge of the ligand. Chapter 2 will present work wherein a variety of biophysical techniques were used to characterize the binding kinetics and affinity of several known and screened SRB-2 ligands. These techniques are also used to categorize the ligands by mechanism of binding. Chapter 3 discusses the attempted determination of a solution structure for SRB-2 bound to SR by NMR spectroscopy. This involved the preparation of several unlabelled and selectivity isotopically labelled SRB-2 samples. Chapter 4 outlines further characterization of the SRB-2 complex using a segmental analysis approach. Several truncated versions of the SRB-2 aptamer were designed and characterized by fluorescence and NMR spectroscopy. Chapter 5 presents work involving a DNAzyme that strongly discriminates between Na^+ and K^+ . This work uses spectroscopy techniques including ThT, NMR and CD to show that, despite high G content, this selective binding ability is not conferred by a G-quadruplex. Chapter 6 presents a summary and review of future work that can be done where applicable.

Chapter 2: Ligand Specificity and Affinity in the Sulforhodamine B Binding RNA Aptamer

2.1: Foreword

The results in this chapter have been published in the journal Biochemical and Biophysical Research Communications: **Kyle A. Piccolo**, Brooke McNeil, Jeff Crouse, Su Ji Lim, Sarah C. Bickers, W. Scott Hopkins & Thorsten Dieckmann. Ligand Specificity and Affinity in the Sulforhodamine B Binding RNA Aptamer. *Biochem. Biophys. Res. Commun.* 2020, 529 (3), pp 666-671. Any permission for further re-use of this material should be requested directly from BBRC with the DOI: <https://doi.org/10.1016/j.bbrc.2020.06.056>.

LSPR experiments in this chapter (Figures 2.5 and 2.6, Table 2.5) were performed by Brooke McNeil and the electrostatic potential surfaces (Figures 2.22 and 2.23) were calculated by Jeff Crouse and Su Ji Lim. All other experimental work was performed and analyzed by the candidate.

All data and figures in this chapter were published in the above article with the following exceptions: DSC experiments shown in (Figure 2.20), CD experiments shown in (Figure 2.21), ITC salt studies shown in Table 2.3 and the following NMR figures (Figures 2.8-12, 17-19). Data and figures listed are unchanged from the original publication, though some of the text was modified for inclusion in this thesis.

2.2 Chapter Abstract

Binding affinity and selectivity are critical properties of aptamers that must be optimized for any application. The sulforhodamine B binding RNA aptamer (SRB-2) is a somewhat promiscuous aptamer that can bind ligands that vary markedly in shape, size and charge. Here we categorize potential ligands

based on their binding mode and structural characteristics required for high affinity and selectivity. Several known and potential ligands of SRB-2 were screened for binding affinity using fluorescence. Promising candidates were subsequently characterized by LSPR, ITC, fluorescence anisotropy, DSC, CD and NMR spectroscopy. These studies show that rhodamine B has the ideal structural and electrostatic properties for selective and high-affinity binding of the SRB-2 aptamer.

2.3 Introduction

The sulforhodamine B binding aptamer (SRB-2) is an RNA aptamer that was selected by the Wilson group to bind the fluorescent dye sulforhodamine B (SR) [37]. The secondary structure of SRB-2 and the structures of six of the studied ligands are shown in Figure 2.1. The structures of several other known and potential ligands studied are shown in Figure 2.2. Some of the very first targets used in the development of SELEX were small, organic, planar fluorophores due to their similarity to biological cofactors and comparatively high quantum yield [28]. Such aptamers, including SRB-2, continue to be widely studied due to their potential applications in bio-imaging and bio-sensing [79,88,89,182].

The SRB-2 aptamer was selected for use in so-called “light-up” RNA aptamer systems [37]. This is a promising method wherein the complement of a fluorophore binding aptamer sequence is added to a gene of interest. The target fluorophores have low intrinsic fluorescence but become highly fluorescent upon binding the aptamer [113,182]. Several other fluorophore binding aptamers, particularly the class based on derivatives of fluorescent protein fluorophores such as spinach [90], broccoli [91], corn [92] and mango [93] have been selected and studied to create a rainbow of RNA reporter molecules for nucleic acid researchers analogous to the arsenal of fluorescent proteins used by protein researchers [113,183]. SRB-2 has been used successfully in mRNA imaging experiments with SR and with SR-quencher constructs designed to reduce background noise in fluorescence measurements [88,89].

Aptamers have several characteristics that make them attractive receptors to use in biosensors compared to antibodies. These include similar high affinity, but with higher stability, ability to be produced synthetically, easy modification and wider range of potential targets. There are a few literature examples of fluorophore binding aptamers like SRB-2 being used in this type of application. One prime example is the label-free sensor discussed in Chapter 1 that incorporated the RNA aptamer MGA (malachite green aptamer) [79]. This sensor consisted of a fusion of MGA and the ATP aptamer, and a bridging strand connecting the two. The bridging strand is released upon ATP binding, allowing MGA to bind MG and significantly increase the fluorescence. In order to develop such systems, the binding mode of the aptamer needs to be well characterized, which we intend to do with the SRB-2 aptamer.

One of the defining characteristics of SRB-2 is its ligand promiscuity. SRB-2 can bind a variety of planar dyes which contain a similar xanthene ring structure, which is characteristic of rhodamine derivatives. These dyes are diverse in color, size, charge and selectivity. Much of the motivation of this work is to determine what structural characteristics of the dyes make for the best SRB-2 ligand. The dyes used in this work were selected on a number of factors which include structural similarity to SR, commercial availability and presence in prior SRB-2 studies [37,89]. The authors who originally selected SRB-2 used tetramethylrosamine and xylene cyanol among others in a competitive elution experiment with SR agarose [37]. There was also work done by Sunbul and Jaschke, where the K_d values of various SRB-2 ligands were determined and we took a representative sample of those including pyronin Y (PY), acridine orange (AO), 9-aminoacridine (9AA) and Atto 495 [89]. The problem with some of these ligands is that despite higher fluorescence enhancement, they are known nucleic acid stains. This means they bind sequence independently in the RNA/DNA backbone, which could limit their practical use [184–186]. All other dyes studied, including rhodamine B (RB), sulforhodamine 101 (SR101), rhodamine 6G (R6G), rhodamine 110 (R110), Atto rhodamine 101 biotin (AR101-B) and 5/6-biocyin tetramethylrhodamine (B-

TMR) were chosen solely for their structural similarity to SR and commercial availability and to our knowledge have not been studied as SRB-2 ligands in any prior literature.

It is worth noting that SRB-2 is also capable of binding the dyes Patent Blue V (PBV) and Patent Blue VF (PBVF) [76]. PBV is very similar to SR but lacks the bridging oxygen in the central ring, resulting in a propeller-like shape. PBVF differs from PBV by only the addition of a hydroxyl group in the meta position opposite the ortho sulfonate. However, these dyes bind with K_d values of just 23 μM for PBV and 86 μM for PBVF. These binding constants are significantly higher than the literature value of 310 nM for SR. This may limit the practicality of the aptamer in imaging and sensing applications. Therefore, we opted not to study these ligands any further.

Another part of what makes SRB-2 an intriguing system to study is the overall negative charge of its ligand, SR (Figure 2.1). Negatively charged aptamer ligands are generally less common than neutral and positively charged ones due to potentially repulsive backbone interactions. There are few structures available of aptamers bound to negatively charged ligands available. One prominent example is the ATP aptamer. However, negative charges are not explicitly required for binding and based on the structure when bound the AMP, the negative charges of the phosphate do not interact directly with the RNA. Given the consistent presence of the xanthene-type group in ligands studied, this may well be the case for SRB-2 as well.

In these studies, we use a variety of biophysical techniques to characterize and compare the kinetics and binding mode for several aptamer-ligand complexes of SRB-2. Fluorescence scans are used initially to screen ligands for binding activity. Fluorescence anisotropy, isothermal titration calorimetry (ITC), and light surface plasmon resonance (LSPR) are used to probe binding affinity and selectivity of known and potential ligands. Thermodynamic parameters of selective binders are also assessed via ITC and differential scanning calorimetry (DSC). Aptamer-ligand interactions are examined using

homonuclear nuclear magnetic resonance (NMR) spectroscopy, including 1D ^1H NMR titrations and 2D NOESY experiments.

Characterizing aptamer systems such as SRB-2 is essential to rationalize the motifs and intermolecular contacts that are critical for binding affinity and specificity. This information can then be used to implement aptamers in several types of applications including cell imaging and biosensing. In these studies, we screen several known potential SRB-2 ligands and use them to probe the structure of SRB-2. We determined the binding affinity, selectivity and mode of each of these ligands with the goal of determining what an ideal SRB-2 ligand would look like. Of the ligands studied, rhodamine B possesses the ideal characteristics for use in applications including selective binding, high binding affinity and reasonable fluorescence compared to other selective binders.

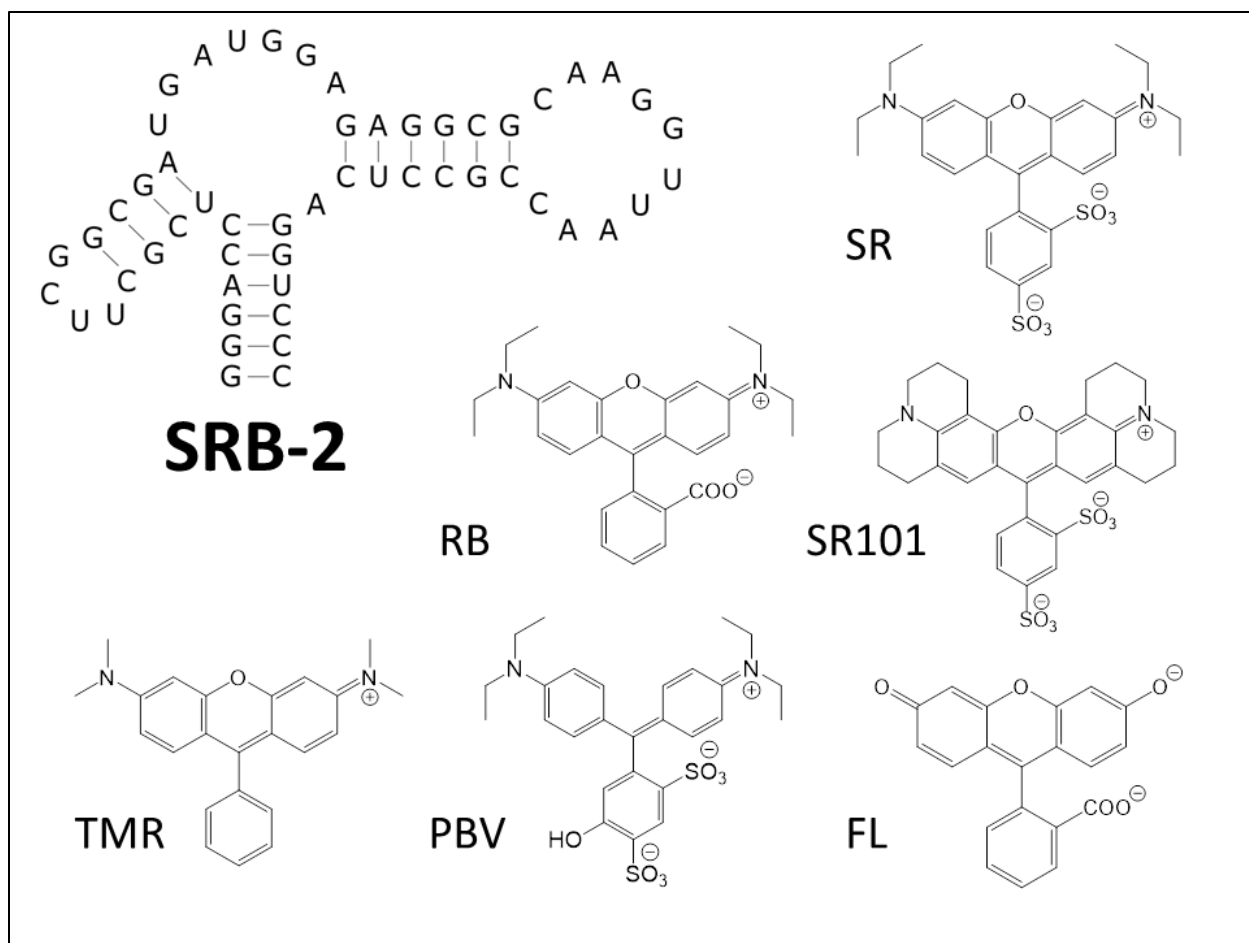


Figure 2.1: Secondary structure of SRB-2 and chemical structures of several studied dyes.

These include sulforhodamine B (SR), rhodamine B (RB), sulforhodamine 101 (SR101), tetramethylrhodamine (TMR), Patent blue V (PBV) and fluorescein (FL). Chemical structures of all other dyes discussed can be found in Figure 2.2.

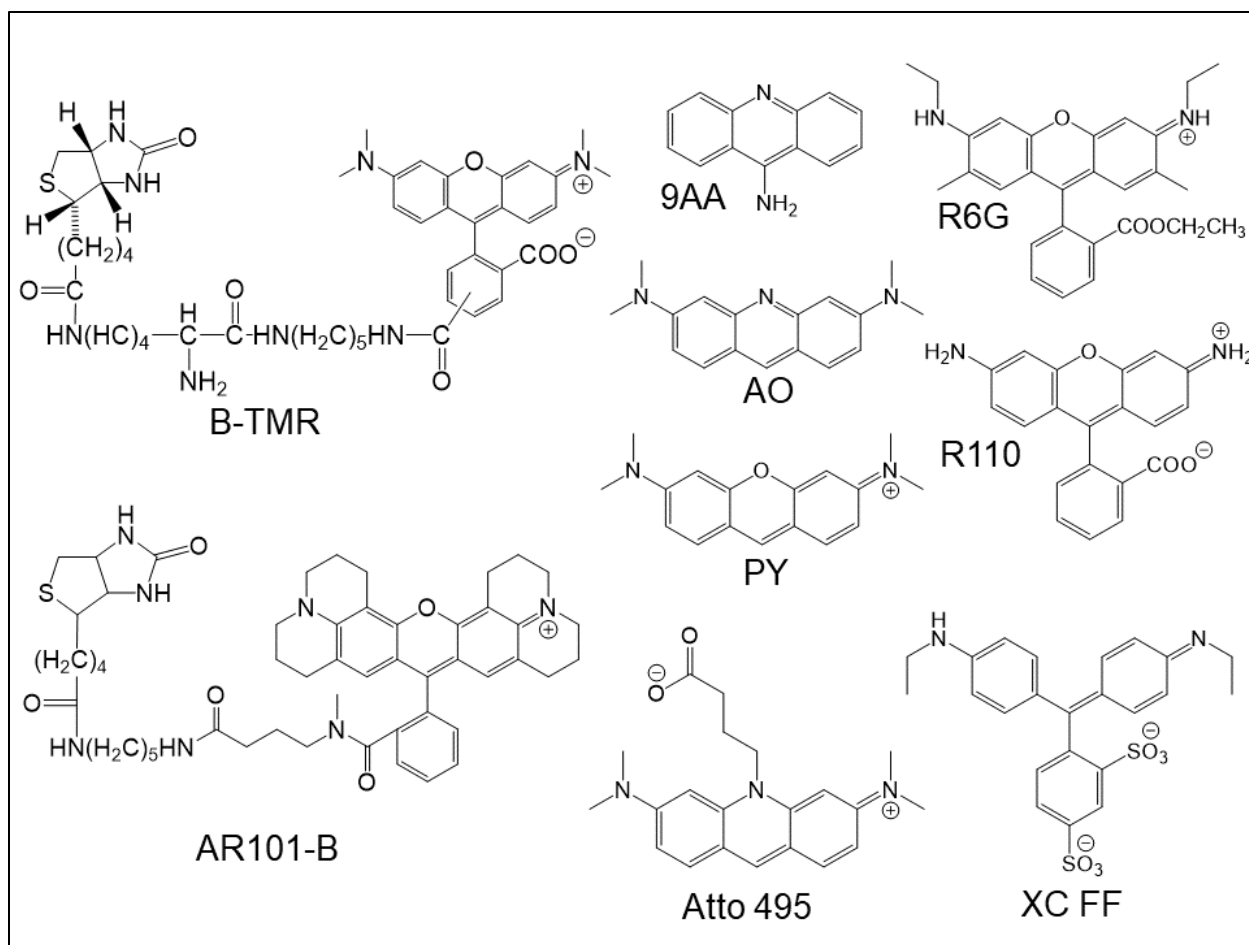


Figure 2.2: Chemical structures of ligands screened for SRB-2 binding activity.

These include 5/6-biotin tetramethylrhodamine (B-TMR), Atto rhodamine 101 biotin (AR101-B), 9-aminoacridine (9AA), Acridine orange (AO), Pyronin Y (PY), Atto 495, Rhodamine 6G (R6G), Rhodamine 110 (R110) and Xylene cyanol FF (XC).

2.4 Materials and Methods

2.4.1 Preparation of RNA Samples and Dyes

The SRB-2 RNA was synthesized enzymatically using a T7 RNA polymerase and a double-stranded synthetic DNA template (Integrated DNA Technologies, Inc., Coralville, Iowa). The sequences used were as follows:

SRB-2 template DNA:

5'-GGGACCTGAGGCGGTTAACCTTGCGCCTCTCCATCATCGCCGAAGCGAGGTCCCTATAGTGAGTCGTATTA-3'

SRB-2 coding DNA:

5'-TAATACGACTCACTATAGGGACCTCGCTTCGGCGATGGAGAGGCGCAAGGTTAACCGCCTCAGGTCCC-3'

A single base-pair substitution in the original sequence of SRB-2 (A3 → G3 and U52 → C52) was introduced to increase transcriptional yield, as the presence of three consecutive G residues at the beginning of the sequence is thought to increase initiation of transcription [187,188]. RNA was transcribed using the recipe listed in Table 2.1. The transcription buffer was tris-buffered saline (TBS), which contains 40 μM Tris and 100μM NaCl.

Table 2.1: Sample recipe for a 10mL transcription.

Reagent	Volume	Stock concentration	Final Concentration
DMSO	2 mL	100 %	20 %
MQ	400 μL	N/A	N/A
TBS (pH 8.3)	1 mL	10X	1X
MgCl ₂	450 μL	1 M	40 mM
DTT	500 μL	100 mM	5 mM
NTPs (A,C,G,U)	750 μL	100 mM	7.5 mM
PEG 8000	400 μL	40% w/v	1.6% w/v
DNA Template	800 μL	20 μM	1.6 μM
T7 RNAP (in 50% glycerol)	1 mL	~3 mg/mL	~0.3 mg/mL

The RNA was purified on a 10% Urea PAGE gel, and the band containing the aptamer was cut out. The RNA was eluted from the gel by crush & soak in 300mM NaCl or by electroelution. This was followed by clean-up on a HiPrep 16/10 DEAE FF anion-exchange column (GE Healthcare, Uppsala, Sweden) and desalting on a HiPrep 26/10 Desalting column (GE Healthcare, Uppsala, Sweden). The RNA was also precipitated with 70% ethanol prior to running each column. Pure samples obtained from the desalting column were then lyophilized and dissolved in the appropriate buffer. For the fluorescence assay, MGA and Mango RNA aptamers were also used. In this case, a single-stranded template annealed to the T7 promoter was used in place of a double-stranded template sequence. These aptamers were otherwise transcribed and purified using methodologies analogous to SRB-2. DNA sequences used are as follows:

MGA DNA: 5'-GGATCCATTCGTTACCTGGCTCTCGCCAGTCGGGATCCTATAGTGAGTCGTATTA-3'

Mango DNA: 5'-CTGCTCTCCTCTCCGCACCGTCCCTTCGCTCCCTATAGTGAGTCGTATTA-3'

T7 promoter: 5'-TAATACGACTCACTATAG-3'

Sulforhodamine B (Life Technologies, Eugene, OR), sulforhodamine 101, acridine orange, 9-aminoacridine, Atto 495, rhodamine 110 chloride, fluorescein (Sigma, St. Louis, MO), rhodamine B, rhodamine 6G (Alfa Aesar, Ward Hill, MA), tetramethylrosamine (Invitrogen Corporation, Carlsbad, CA), pyronin Y (Acros Organics, Geel, Belgium) and xylene cyanol FF (EMD, Burlington, MA) were used without further purification. Dyes were prepared in assay buffer (10mM HEPES, 10mM KCl and 5mM MgCl₂ at pH 7.4) or NMR buffer (10 mM potassium phosphate buffer and 10 mM KCl) and concentrations were verified by UV/vis spectrophotometry.

2.4.2 Emission Scans

Spectra were obtained using a Spectramax M3 Multi-Mode Microplate Reader using a 100 nm wide scan with 1nm intervals. SR, RB and TMR were measured at an λ_{ex} of 520 nm and an λ_{em} range of

540-640 nm. SR101 was measured using an λ_{ex} of 550 nm and an λ_{em} range of 570 nm to 670 nm. PY was measured using an λ_{ex} of 510 nm and an λ_{em} range of 530 nm to 630 nm. AO and Atto 495 were measured using an λ_{ex} of 460 nm and an λ_{em} range of 480 nm to 580 nm. 9AA was measured using an λ_{ex} of 390 nm and an λ_{em} range of 410 nm to 510 nm. Readings were taken in Greiner 96-well black microplates (Kremsmünster, Austria). Samples were prepared with 10 μ M dye and 50 μ M RNA in assay buffer (10mM HEPES, 10mM KCl and 5mM MgCl₂ at pH 7.4).

2.4.3 Fluorescence Anisotropy

Fluorescence anisotropy measurements were acquired on a Photon Technology International (Edison, NJ) LS-100 spectrofluorimeter equipped with a continuous Ushio (Cypress, CA) UXL-75Xe xenon arc lamp and a PTI 814 photomultiplier detection system (Photon Technology International). Measurements were taken using a λ_{ex} of 563 nm and a λ_{em} of 585 nm, which were determined experimentally using excitation and emission scans. Samples were prepared with varying concentrations of RNA (0.1-50 μ M) and 1 μ M SR in assay buffer (10mM HEPES, 10mM KCl and 5mM MgCl₂ at pH 7.4). Samples were incubated at 25°C for at least 10 minutes before measurement and then read in a Hellma high-performance quartz glass cuvette with a 3.00 mm optical path length and 45 μ L volume (Müllheim, Germany). The resulting binding curves were fit using a non-linear, least squares method to the simple hyperbolic function: $\% \text{FA}_{\text{max}} = \text{FA}_{\text{max}} \times [\text{SRB-2}] / (K_d + [\text{SRB-2}])$ where FA_{max} is the maximum fluorescence anisotropy and K_d is the dissociation constant. Anisotropy measurements were normalized to a percentage of FA_{max} for each respective ligand.

2.4.4 ITC Studies

Experiments were performed on a MicroCal ITC 200 microcalorimeter (MicroCal Inc., Northampton, MA). The samples of 30-35 μ M RNA and dye solutions of 0.7-1.1 mM were each prepared

by dissolving in assay buffer. All experiments were carried out at 25°C and performed in triplicate. For each dye or experimental condition, a blank run with assay buffer in the cell and fluorophore dissolved in assay buffer in the syringe was performed. This run was subtracted from the experimental run to account for the heat of dilution of the fluorophore solutions. All data sets were analyzed and fitted using the Origin 7 software package provided by MicroCal. All data was fit to a single-site binding model using a non-linear least-squares approach.

2.4.5 LSPR Assay

Localized surface plasmon resonance (LSPR) binding assays were carried out on an OpenSPR™ (Nicoya, Waterloo, Canada) at room temperature. For the analysis of SRB-2, 10µM biocytin TMR (Invitrogen Corporation, Carlsbad, CA) or biotinylated Atto Rho101 (Sigma, St. Louis, MO) were immobilized onto the streptavidin sensor chip. SRB-2 was then injected at a flow rate of 20 µL/min at concentrations of 0.9 – 7.57 µM in HEPES (10 mM HEPES, 200 mM KCl, 10 mM MgCl₂) running buffer. The binding time was 240 seconds, and the disassociation time was 150 seconds. The sensor was regenerated with 10 mM glycine-HCl (pH 2.5). The data was retrieved and analyzed with TraceDrawer software (Ridgeview Instruments AB, Sweden). A 1:1 Langmuir interaction was fit to the data for each compound, and the fit delivered the on rate (k_a), off rate (k_d), and the equilibrium constant (K_d).

2.4.6 NMR Experiments

NMR samples were prepared by dissolving an appropriate weight of lyophilized RNA in 500µL of 10 mM potassium phosphate buffer and 10 mM KCl (90% H₂O/10% D₂O). Lyophilized aliquots of dye were added to titrate the RNA to a ~1:1 ratio unless otherwise specified. The samples were dried by lyophilization and re-dissolved in 500 µL of 99.996% D₂O (Cambridge Isotopes) to perform experiments on non-exchangeable resonances. All spectra were collected on a Bruker DRX-600 spectrometer

equipped with an HCN triple-resonance, triple-axis PFG probe. Quadrature detection for the indirect dimensions in multidimensional experiments was achieved using the States-TPPI method [189]. Samples in 90% H₂O/10% D₂O that were used to observe exchangeable protons were run using $1\bar{1}$ -spin echo solvent suppression [190]. Two-dimensional NOESY spectra [191] in 90% H₂O/10% D₂O were acquired at 277 K with a mixing time of 150 ms. Samples in D₂O that were used to observe non-exchangeable protons were run with presaturation solvent suppression [192]. 2D C1Y-TOCSY [193] with a mixing time of 50 ms, and NOESY with a mixing time of 150 ms in 100% D₂O were acquired at 298 K. All NMR samples had a volume of 500 μ L and were read in standard 5 mm NMR tubes. Relevant pulse programs can be found in Appendix B.

2.4.7 Electrostatic Potential Surfaces

All electrostatic potential surface (EPS) calculations were performed using the Gaussian 16 program package for computational chemistry. The molecules were optimized at B3LYP/def2-TZVPP level of theory, besides Atto Rhodamine 101, 6-Biocytin TMR, and Patent Blue V, which were optimized at B3LYP/def2-SVP level. Frequency calculations were performed to ensure that structures correspond to minima and not transition states.

2.4.8 Differential Scanning Calorimetry

DSC experiments were performed using an LLP cap DSC (MicroCal Inc., Malvern Instruments Ltd.). 50 μ M SRB-2 in assay buffer (10mM HEPES, 10mM KCl and 5mM MgCl₂ at pH 7.4) was scanned at a rate of 1 $^{\circ}$ C per minute over a range of 20 $^{\circ}$ C to 100 $^{\circ}$ C. Where applicable, 100 μ M of either SR or TMR were also present (2x the SRB-2 concentration). When plotted, each RNA scan has a reference scan of buffer (or dye + buffer where applicable) subtracted from it.

2.4.9 Circular Dichroism Spectroscopy

CD experiments were performed on a Jasco J-715 spectropolarimeter (Jasco Inc., Easton, MD). CD scanning experiments were run from 320 nm to 220 nm with a path length of 0.1 cm, data interval of 0.5 nm, band width of 0.5 nm, response of 1 second, scanning speed of 200 nm minute⁻¹ and a total of four accumulated scans. Samples contained 40 μM RNA in 10mM HEPES, 10mM KCl and 5mM MgCl₂ at pH 7.4. Where applicable, samples also contained 50 μM of either SR or TMR. The samples were heated to 85 °C for 5min, cooled to 4 °C and incubated for at least 24hrs before acquisition at 25 °C. A blank containing buffer or buffer + ligand was subtracted from each sample and the resulting data was smoothed using a Savitzky-Golay smoothing function [194].

2.5 Results

2.5.1 Emission Scans

Fluorescence emission spectra were obtained for each dye both on their own and bound to SRB-2 in a 5:1 ligand:RNA ratio (Figure 2.3). Fluorescence results are collected in Table 2.2. The specificity of these dyes was tested using additional RNA and dsDNA sequences. Dyes that could bind SRB-2 only included sulforhodamine B (SR), rhodamine B (RB) and sulforhodamine 101 (SR101), as well as 5/6-biotin tetramethylrhodamine (B-TMR) and Atto rhodamine 101 biotin (AR101-B), which are functionalized analogues of RB and SR101, respectively. All other dyes listed were capable of binding all sequences tested. Most dyes experienced a red shift in their maximum emission wavelength and an increase in fluorescence upon binding. As shown in Table 2.2, 9-aminoacridine (9AA) and acridine orange (AO) are exceptions, likely a result of a pKa change in the amines, which alters their protonation state and results in a non-fluorescent conformation in the case of 9AA and a higher energy complex in AO.

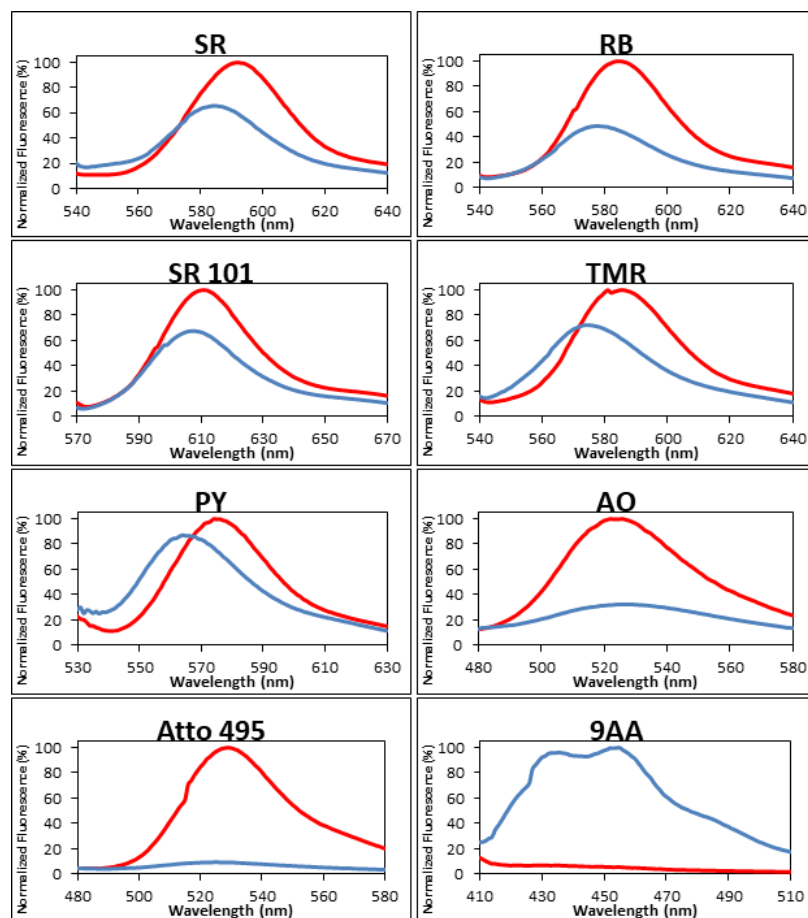


Figure 2.3: Emission scans comparing free dye (blue) and dye bound to SRB-2 (red).

The are examples of some of the emission scans performed, comparing the fluorescence of free ligands to fluorescence when bound to SRB-2. As seen in most of these graphs, a fluorescence increase and a red shift are generally observed upon aptamer binding of a fluorogenic ligand. 9AA is an exception due to the presence of an amine group. Upon binding, the protonation state of this amine is changed, resulting in a non-fluorescent conformation. This plots are normalized to the fluorescence of the more fluorescent state (this would be the bound state in all cases except 9AA, as previously mentioned).

Table 2.2: Wavelength shifts and fluorescence changes determined by emission scans.

Oligonucleotide		SR	RB	SR101	TMR	PyrY	AO	Atto 495	9AA	B-TMR	AR101-B
Assay buffer	RF	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	$\lambda_{\max}(\text{nm})$	580	576	604	573	563	526	516	435	578	605
SRB-2	RF	1.52	2.05	1.48	1.38	1.15	3.08	10.82	0.06	1.11	1.11
	$\Delta\lambda_{\max}(\text{nm})$	8	6	4	10	9	-2	11	0	8	1
dsDNA template	RF	0.91	0.80	0.74	0.35	0.28	2.39	2.53	0.19	0.92	0.84
	$\Delta\lambda_{\max}(\text{nm})$	-2	0	0	9	9	-1	7	0	1	2
MG aptamer	RF	0.87	0.78	1.04	0.37	0.28	2.90	2.50	0.26	0.92	0.92
	$\Delta\lambda_{\max}(\text{nm})$	-1	0	0	6	9	-3	10	0	2	1
Mango aptamer	RF	0.92	0.79	0.86	0.30	0.94	1.94	3.29	0.37	0.81	0.78
	$\Delta\lambda_{\max}(\text{nm})$	-2	0	1	9	9	1	12	0	1	3

The λ_{\max} for each ligand is determined by inspection of the raw emission data of ligand in assay buffer. The $\Delta\lambda_{\max}$ for each ligand-nucleic mixture is determined by taking the difference between the λ_{\max} of the ligand in assay buffer and the respective ligand-nucleic acid mixture. For purposes of identifying binding, the ligand-nucleic acid mixture must have a fluorescence increase of at least 10% and/or a λ_{\max} that increases by at least 5 nm.

2.5.2 Fluorescence Anisotropy

Fluorescence anisotropy was used to determine the binding affinity of the ligands tested as this was the method used by the authors of the original SRB-2 selection [37]. 9AA was not tested due to the unusual behaviour observed in the fluorescence emission scans. Through the dissociation constant, K_d , we find the order of binding affinities to SRB-2 is $\text{RB} > \text{SR} > \text{SR101}$, as shown in Table 2.3. This trend matches up with the ITC data listed in Table 2.4, however, affinities are reported as being lower in general. This may be due to the concentration of ligands not being limiting enough in these assays. TMR,

PY, AO and Atto 495 were also tested, but they are suspected to bind in multiple sites at SRB-2. The simple hyperbolic function used to fit these data is meant to represent single-site binding. Therefore, this methodology may be inappropriate for such ligands. The K_d for TMR is slightly higher than SR. Like RB and SR101, this value was also previously unreported, but due to its presumed intercalation, we were unable to substantiate this result with ITC. The K_d values determined for PY and AO are both lower than SR, which agrees with previous studies [89]. Atto 495 was determined to have the lowest binding affinity, though it was expected to be similar to SR based on previous studies [89].

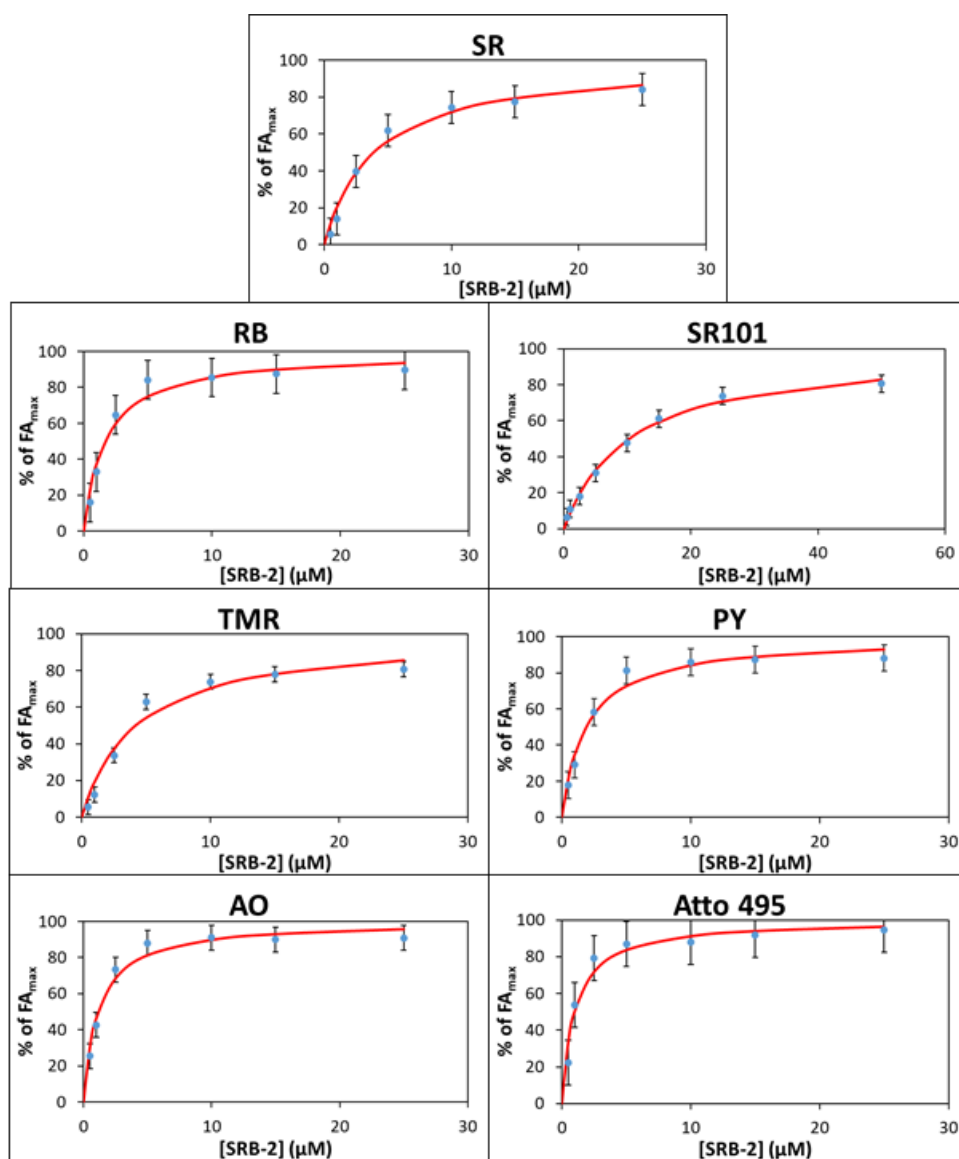


Figure 2.4: Fluorescence anisotropy titration curves for SRB-2 and its various ligands.

Experimental data is shown as blue points and fit data is shown as a red line. Error bars are one standard deviation from the mean of the replicates acquired.

Table 2.3: Dissociation constants of SRB-2 ligands determined by fluorescence anisotropy.

Ligand	K_d (μM)
SR	3.94 ± 0.34
RB	1.67 ± 0.18
SR 101	10.3 ± 0.5
TMR	4.24 ± 0.17
PY	1.89 ± 0.14
AO	1.16 ± 0.08
Atto 495	0.989 ± 0.12

2.5.2 Isothermal Titration Calorimetry

Isothermal titration calorimetry (ITC) was also used to determine the binding affinity of the ligands due to the extra thermodynamic information it can provide. It has the inherent advantage over fluorescence anisotropy of being able to directly measure the molar ratio, n and the enthalpy, ΔH . The entropy, ΔS , and Gibbs free energy, ΔG , can also be determined indirectly. A sample ITC isotherm and its integrated plot are given in Figure 2.5. Through the dissociation constant, K_d , we find the order of binding affinities to SRB-2 is $\text{RB} > \text{SR} > \text{SR101}$, as shown in Table 2.4. This agrees with K_d values obtained from the fluorescence anisotropy experiments.

Several other dyes, including tetramethyl rosamine (TMR), pyronin Y (PY), Atto 495, acridine orange (AO) and 9-aminoacridine (9AA) were tested by ITC. However, these small, planar and electron-poor dyes intercalate between base pairs in double-stranded nucleic acid sequences. The resulting isotherms adhered poorly to a single-site binding model. Therefore, reasonable binding affinities and thermodynamic parameters were not determined by this method. PY, AO, 9AA and Atto 495 also mirror the trend observed in previous studies using fluorescence titrations [89].

The binding of ligands to aptamers is commonly dependent on divalent counter-ions to stabilize the negatively charged backbone. To probe this relationship, the ITC experiments were repeated in the absence of Mg^{2+} . No binding was observed under these conditions for SR, SR101 or RB. Weak binding was observed for TMR, which carries an overall positive charge. However, due to its non-specific binding nature, an appropriate fit could not be determined. ITC experiments at 500mM KCl in the absence of Mg^{2+} showed weak binding of the dye ligands to SRB-2, as shown in Table 2.5. SR101 is not shown as binding was too weak to determine a reliable fit.

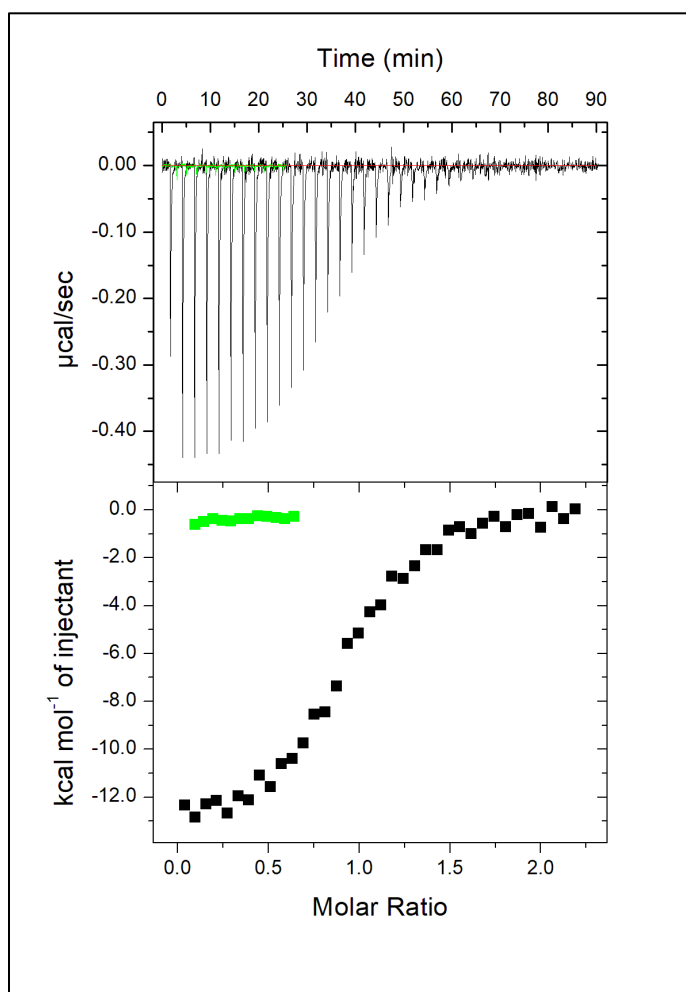


Figure 2.5: Representative ITC data for the binding of SRB-2 to SR.

A titration curve of SR injected into SRB-2 is shown in black in the top panel. SR into buffer control is shown in green. The binding isotherm is shown in black in the bottom frame, with the SR into buffer control again being shown in green.

Table 2.4: Thermodynamic parameters of SRB-2 determined by ITC in 5 mM MgCl₂ and 10 mM KCl.

Ligand	<i>n</i>	K _d (μM)	ΔH (kcal mol ⁻¹)	ΔS (cal mol ⁻¹ K ⁻¹)
SR	0.885 ± 0.012	1.45 ± 0.09	-13.1 ± 0.3	-17.2 ± 0.4
SR101	0.882 ± 0.031	4.7 ± 0.70	-10.1 ± 0.5	-9.27 ± 0.8
RB	0.922 ± 0.010	0.447 ± 0.08	-14.2 ± 0.2	-18.5 ± 1.2

Table 2.5: Thermodynamic parameters of SRB-2 determined by ITC in 0.5 M KCl and no MgCl₂

Ligand	<i>n</i>	K _d (μM)	ΔH (kcal mol ⁻¹)	ΔS (cal mol ⁻¹ K ⁻¹)
SR	1.09 ± 0.008	11.9 ± 0.2	-36.6 ± 0.4	-99.6 ± 6.0
RB	1.08 ± 0.013	6.02 ± 0.3	-44.1 ± 0.8	-124 ± 2.6

2.5.3 Surface Plasmon Resonance Studies

LSPR anisotropy experiments were performed in order to obtain kinetic data for the structurally relevant rhodamine derivatives biocytin-tetramethyl rhodamine (B-TMR) and Atto rhodamine 101 biotin (AR101-B). These dyes were selected based on commercial availability and ability to bind the surface of SPR chips. Dissociation curves are shown in Figures 2.6 and 2.7, and binding data obtained for these dyes is listed in Table 2.6.

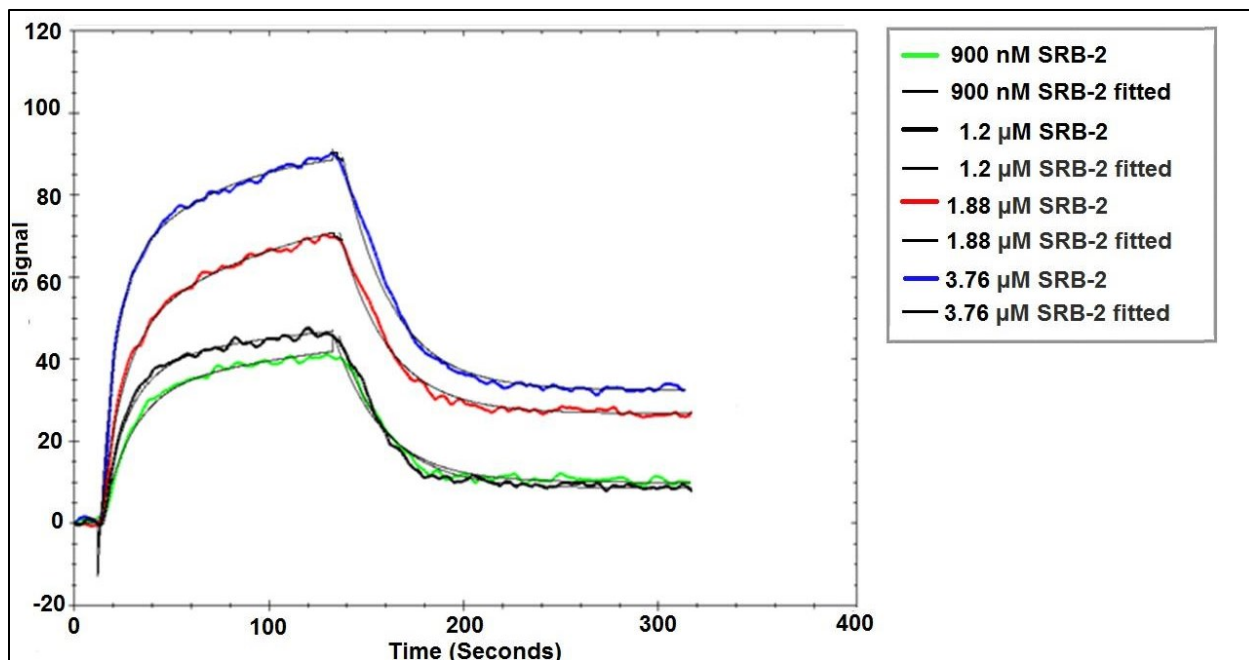


Figure 2.6: LSPR response curve for Biocytin-TMR.

LSPR response curve and 1:1 stoichiometry kinetic fit models for B-TMR. B-TMR was immobilized onto the streptavidin sensor chip and SRB-2 was injected at a flow rate of 20 $\mu\text{L}/\text{min}$ at concentrations of 0.9 – 3.76 μM .

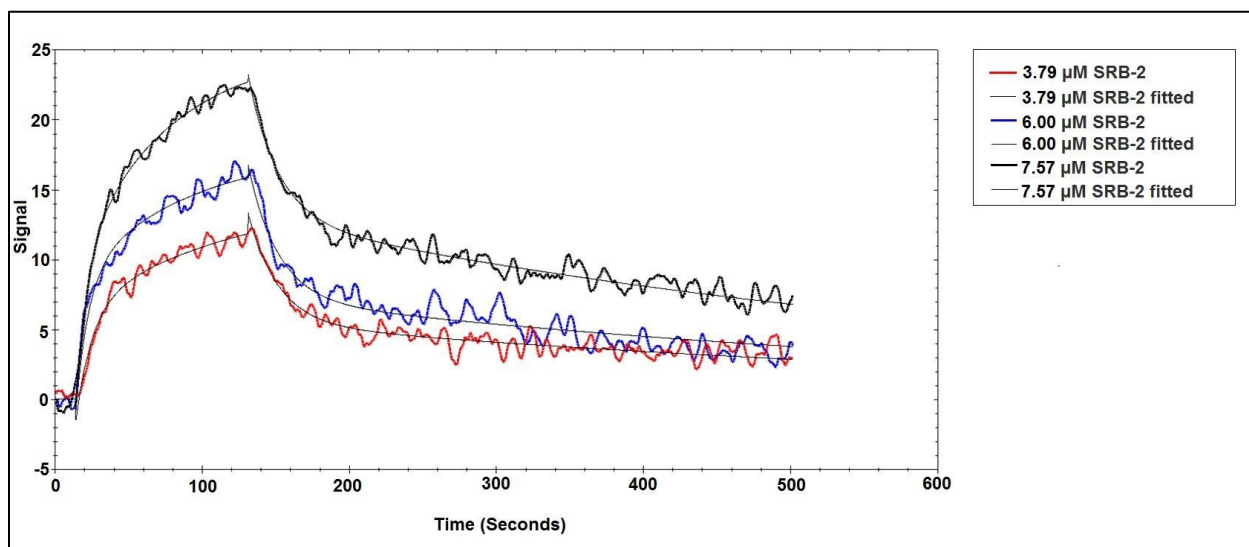


Figure 2.7: LSPR response curve for Atto Rho 101.

LSPR response curve and 1:1 stoichiometry kinetic fit models for AR101-B. AR101-B was immobilized onto the streptavidin sensor chip and SRB-2 was injected at a flow rate of 20 $\mu\text{L}/\text{min}$ at concentrations of 3.79 – 7.57 μM .

Table 2.6: Kinetic parameters of SRB-2 binding determined by LSPR.

Ligand	k_{on} ($\mu\text{M}^{-1}\text{s}^{-1}$)	k_{off} (s^{-1})	K_{d} (μM)
B-TMR	207 ± 78.0	0.0292 ± 0.0156	1.37 ± 0.23
AR101-B	735 ± 559	0.0246 ± 0.0346	4.66 ± 0.47

2.5.4 NMR Spectroscopy

The binding behavior of SRB-2 with each ligand was studied by NMR spectroscopy. SRB-2 was titrated with each ligand in 3-5 steps. 1D ^1H -NMR spectra in 90% $\text{H}_2\text{O}/10\% \text{D}_2\text{O}$ were used to follow the changes during the titrations. The spectrum of SRB-2 in the absence of dye shows relatively few and generally broad peaks (Figure 2.8 A). This is typical for the adaptive binding mode found in many aptamers [119,122], where the binding pocket in the absence of ligand is mostly unstructured. When the aptamer has been titrated to approximately 1:1 with SR, RB and SR101 (Figure 2.8 B-D), an NMR spectrum consistent with a single conformation is observed. This is indicated by the presence of additional peaks compared to the spectrum of the free SRB-2 and linewidths that are noticeably narrower. Several of the new peaks are in the 9-11 ppm region, which is typical for imino protons in non-canonical base pairs. Narrower line shapes in the NMR spectrum and additional peaks in the 9-11 ppm region are consistent with a more highly structured bound conformation with more contacts forming between the imino protons of guanine or uridine residues and other atoms in the aptamer or ligand.

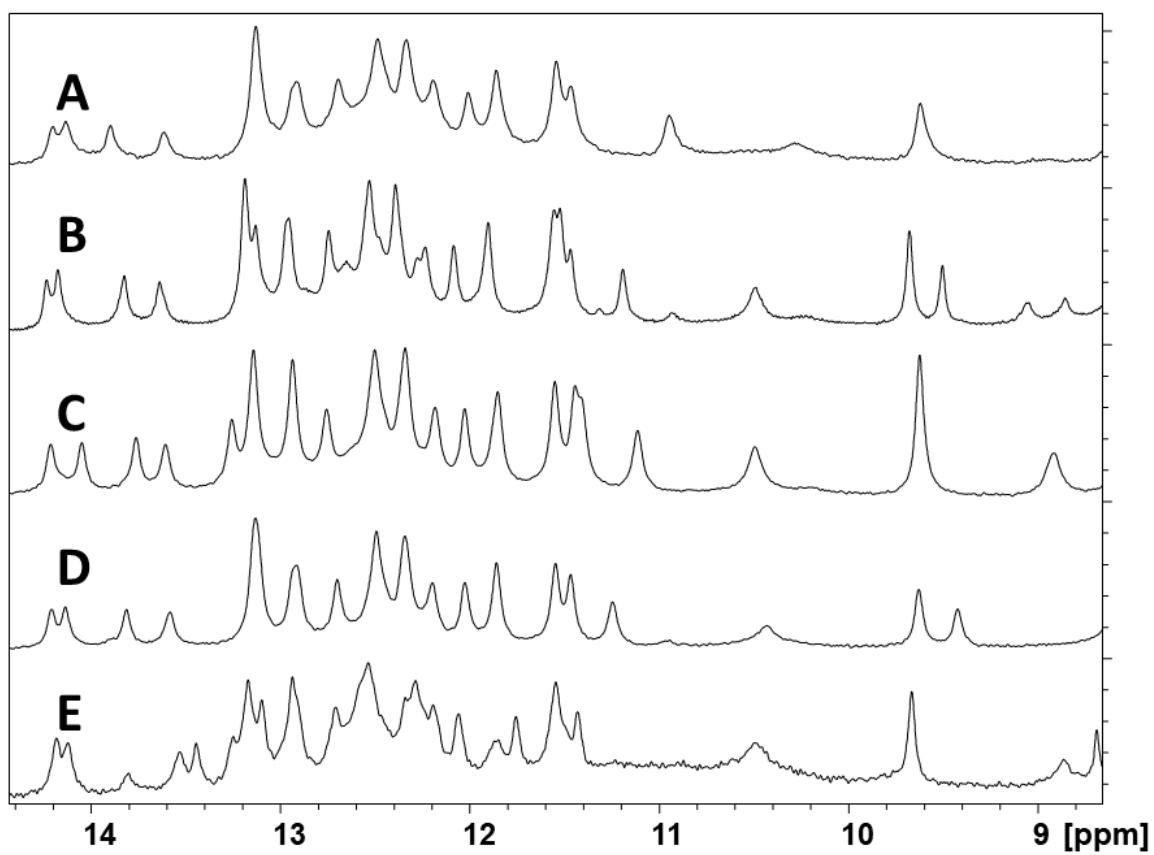


Figure 2.8: 1D ¹H NMR spectra of free SRB-2 and SRB-2 bound in a 1:1 ratio to several different ligands.

A) 0.9 mM SRB-2 only B) 1.8 mM SR, C) 1.3 mM RB, D) 1.1 mM SR101, E) 0.4 mM TMR.

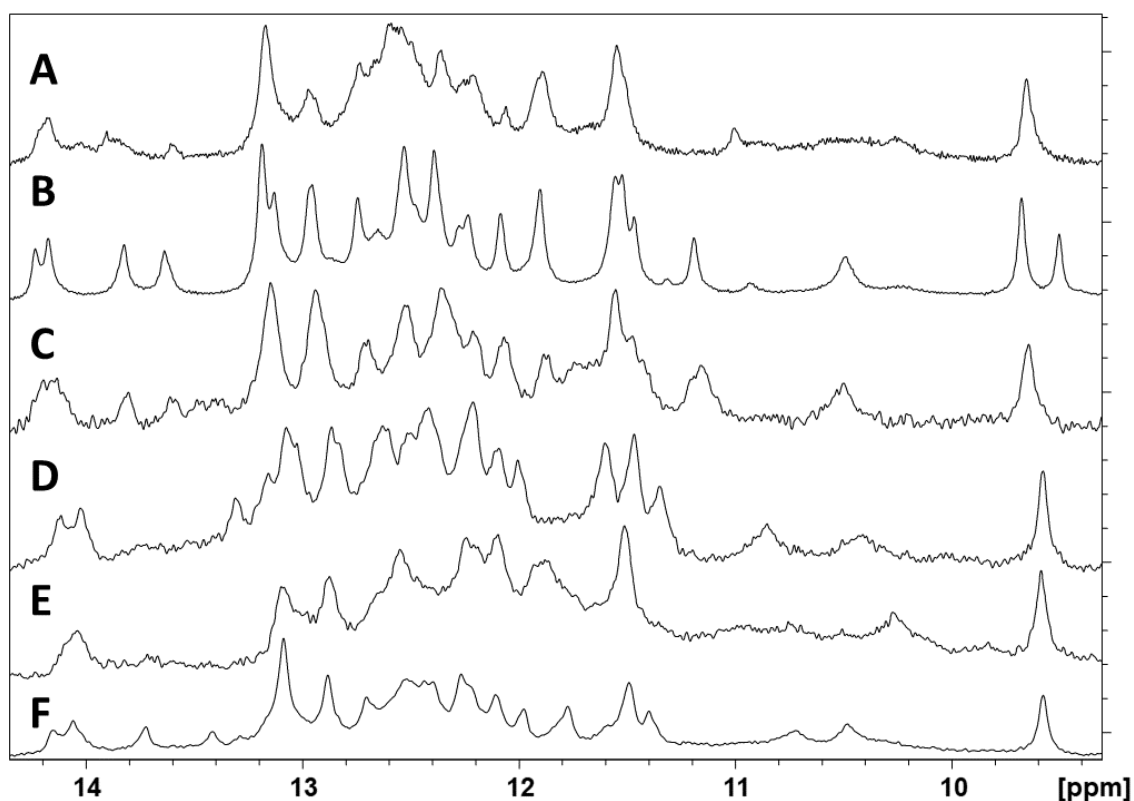


Figure 2.9: 1D ^1H NMR spectra of free SRB-2 and SRB-2 bound in a 1:1 ratio to several different ligands.

A) 0.9 mM SRB-2 only B) 1.8 mM SR, C) 0.4 mM PY, D) 0.2 mM AO, E) 0.4 mM 9AA, F) 0.4 mM Atto 495.

TOCSY experiments were run on SRB-2 complexed with SR, RB and SR101. The aromatic protons of the ligands, particularly those on the single rings, are shifted downfield when bound to SRB-2 as a result of diamagnetic anisotropy experienced by these aromatic protons when they are involved in π - π stacking interactions with bases of the aptamer. It was qualitatively observed in these spectra that RB protons have a larger change in the chemical shift between free and bound forms ($\Delta\delta_{\text{free-bound}} = \delta_{\text{free}} - \delta_{\text{bound}}$) as compared with SR, which shows larger shifts as compared with SR101. The spectrum for SR on its own is shown in Figure 3.11 and the spectra for SR and RB bound to SRB-2 are listed in Appendix A.

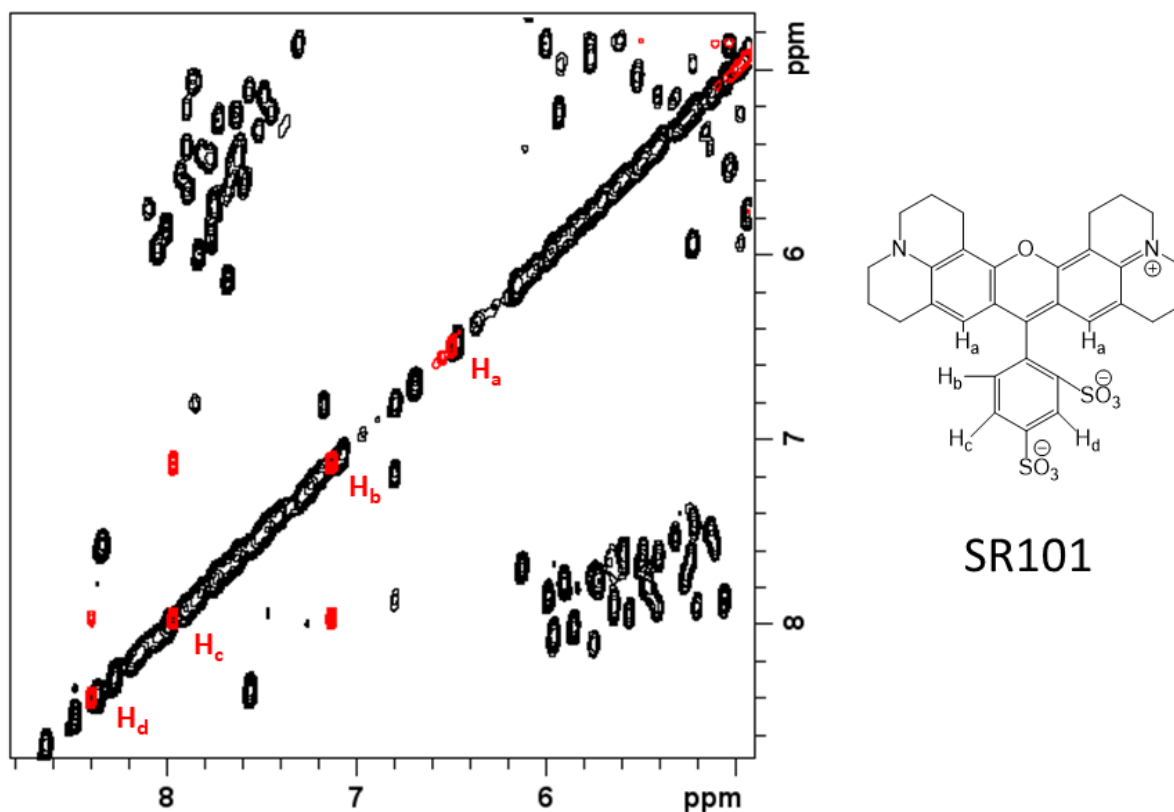


Figure 2.10: TOCSY spectra of 1mM SR101 (red) and 1.1mM SRB-2 fully bound to SR101 (black).

TOCSY was acquired with a mixing time of 50 ms. Aromatic protons of SR101 are labelled in red. Clear changes in the chemical shifts of these protons can be observed in free (red) vs. bound (black) spectra.

NOESY experiments in 90% H₂O/10% D₂O are used to map interactions between imino-protons. The NOEs between exchangeable imino protons in adjacent base-pairs allow a comparison of the patterns formed in the unfolded and folded states of the aptamer. SR has the greatest number of NOEs between exchangeable protons (Figure 2.14). This is expected as SR was the ligand SRB-2 was initially selected to bind, so the binding site is optimized for this molecule. The NOESY spectra for RB (Figure 2.15) and SR101 are almost identical to SR, except for some minor chemical shift changes. This indicates that they likely bind in the same location. In the NOESY spectrum of TMR (Figure 2.17), fewer stacking NOEs are present when compared to SR. In addition, multiple conformations were clearly observed

based on the excess number of peaks compared to SR, RB and SR101 in the 1D ^1H -NMR spectrum (Figure 2.8E). PY, 9AA and Atto 495 all have similar NOESY and 1D spectra to TMR (Figure 2.9 B-F and Figures 2.18-2.20), indicating that they are also binding in multiple different locations. In addition, lack of NOEs present in these spectra suggest that each of these dyes intercalate in the base-paired stems, rather than participating in adaptive aptamer binding. A NOESY spectrum was not acquired for AO due to limited solubility of the low-purity dye. NOESY spectra for all dyes except SR (Figure 2.11) are listed in Appendix A.

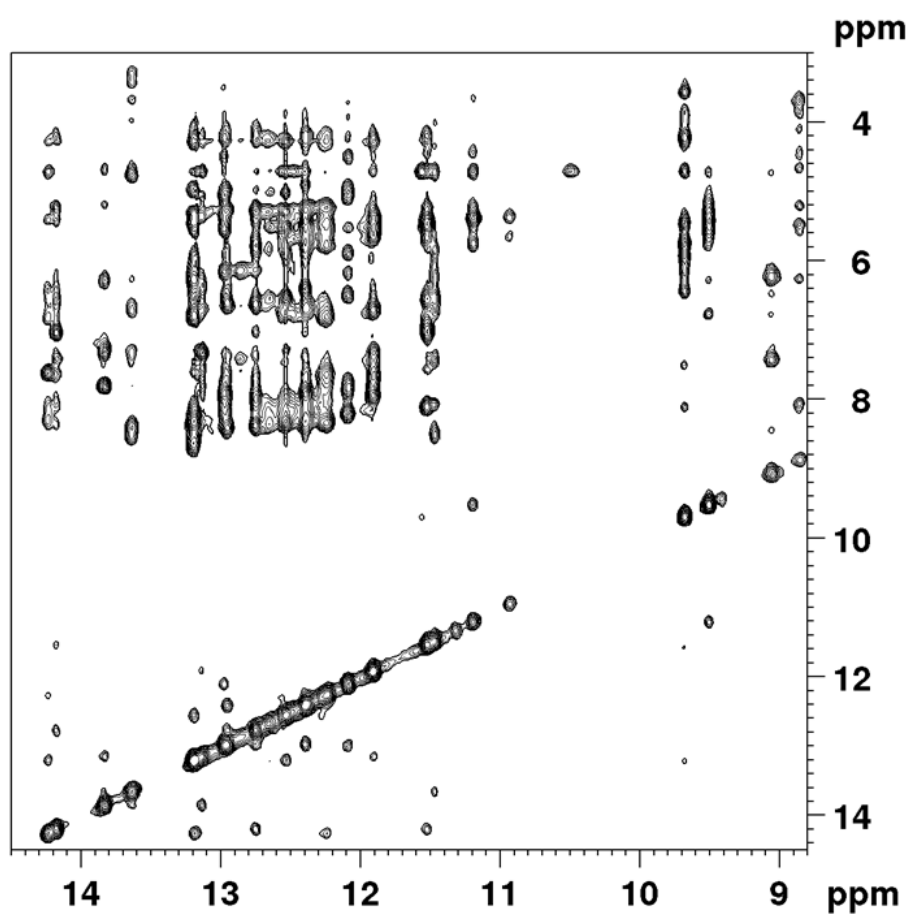


Figure 2.11: NOESY spectrum of 1.8 mM SRB-2 ~1:1 with sulforhodamine B in 90% H_2O /10% D_2O .

Many imino-imino stacking NOEs are observed between 11-14.5 ppm, indicating the presence of several base paired regions, consistent with the predicted secondary structure of SRB-2. This NOESY spectrum was acquired with a mixing time of 150ms.

2.5.5 Differential Scanning Calorimetry

Several differential scanning calorimetry (DSC) isotherms were obtained for SRB-2 in order to compare its melting behaviour when bound to a ligand known to bind specifically with a ligand that is known to bind sequence-independently. Based on observations from the emission scans, SR and TMR were chosen to represent each of those categories, respectively. A comparison of the melting curves of SRB-2 and SRB-2 + SR is shown in Figure 2.20 A and a comparison of SRB-2 + SR and SRB-2 + TMR is shown in Figure 2.20 B. There is melting observed around 70°C in both bound forms that is not observed in SRB-2 alone. This suggests that there are tertiary structures that only form when the aptamer is bound. The secondary structure melts at 90°C in all samples.

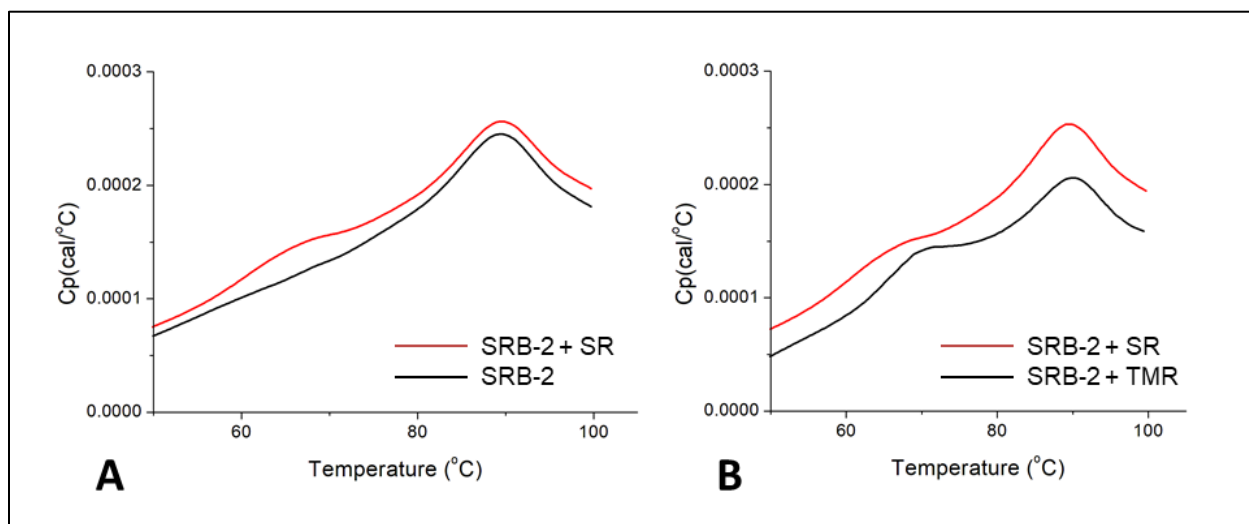


Figure 2.12: Differential scanning calorimetry isotherms.

A) Comparison of unbound SRB-2 to SRB-2 bound to SR and B) comparison of SRB-2 when bound to SR versus when bound to TMR.

2.5.6 Circular Dichroism Spectroscopy

Circular dichroism (CD) spectra were acquired to assess the possibility of any distinct tertiary structures being present in SRB-2. Similar to DSC, CD and UV absorbance spectra were acquired for SRB-2 on its own, SRB + SR and SRB-2 + TMR (Figure 2.21). These dyes were again chosen as a general

representative of their respective binding behaviors. No distinct tertiary structure features such as G-quadruplex or triplex RNA are observed in any of the spectra.

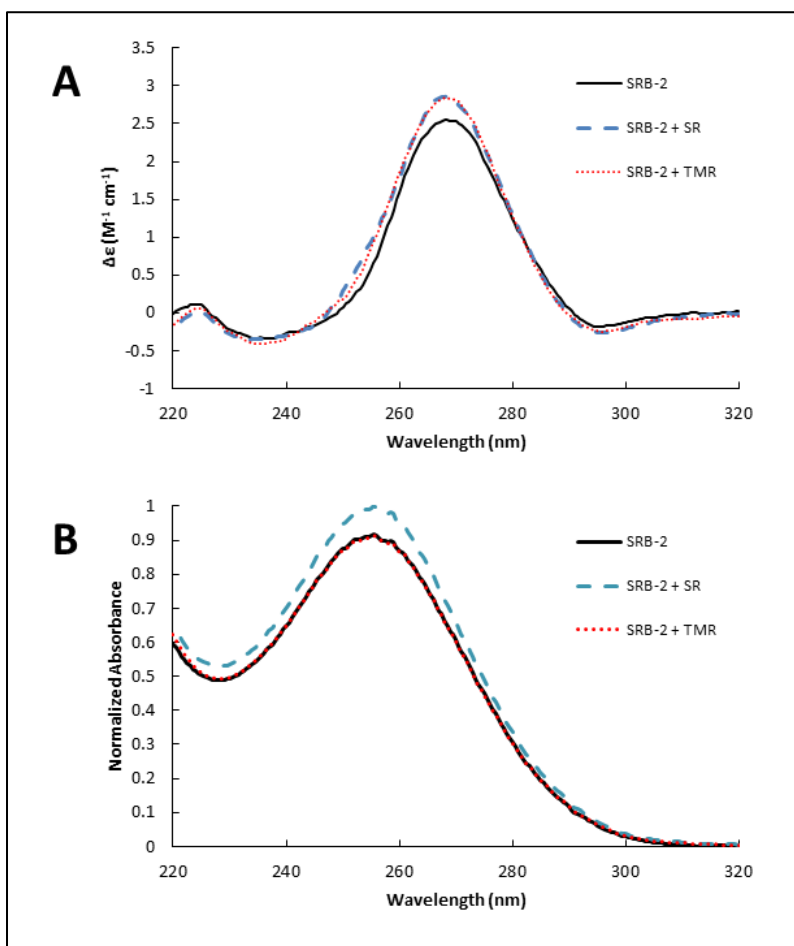


Figure 2.13: CD and UV absorbance spectra of SRB-2.

A) CD spectra comparing unbound SRB-2, SRB-2 bound to SR and SRB-2 bound to TMR.

B) Corresponding UV absorbance spectra for SRB-2, SRB-2 bound to SR and SRB-2 bound to TMR.

2.6 Discussion

2.6.1 Smaller Size and Reduced Negative Charge of Ligands Result in Higher Binding Affinity

From the fluorescence emission scans, we show that RB has a similar red shift but a greater fluorescence intensity increase than SR (Table 2.2) when binding to SRB-2. A comparable red shift is

expected due to their conjugation systems being very similar. The fluorescence intensity increase for RB binding is likely higher as compared with SR due to differences in the polarity of the bound dyes. RB is missing a negatively charged group in the para-position of the single ring, which results in the ligand's conformation being locked more tightly in the binding site. Figure 2.22 shows that the negatively charged sulfonates in SR draw electron density from the three-ring system, creating a significant concentration of negative partial charge on the bottom ring (electrostatic potential surfaces of all dyes referenced are shown in Figures 2.22 and 2.23). RB has a single carboxyl group rather than two sulfonate groups. The carboxyl group of RB is smaller and less electron-dense than the sulfonates on SR, resulting in a dye with less overall electron density. The reduction in repulsive interactions with the RNA backbone and the size of groups on the extended ring would account for the small relative increase in enthalpy and entropy, respectively. It also makes sense that RB would have a lower K_d for SRB-2 because the original selection target had a PEG-anchor at the para position rather than a negative charge and therefore possessed an overall neutral charge like RB. The NMR spectra support this as only minor chemical shift changes are observed between the two complexes.

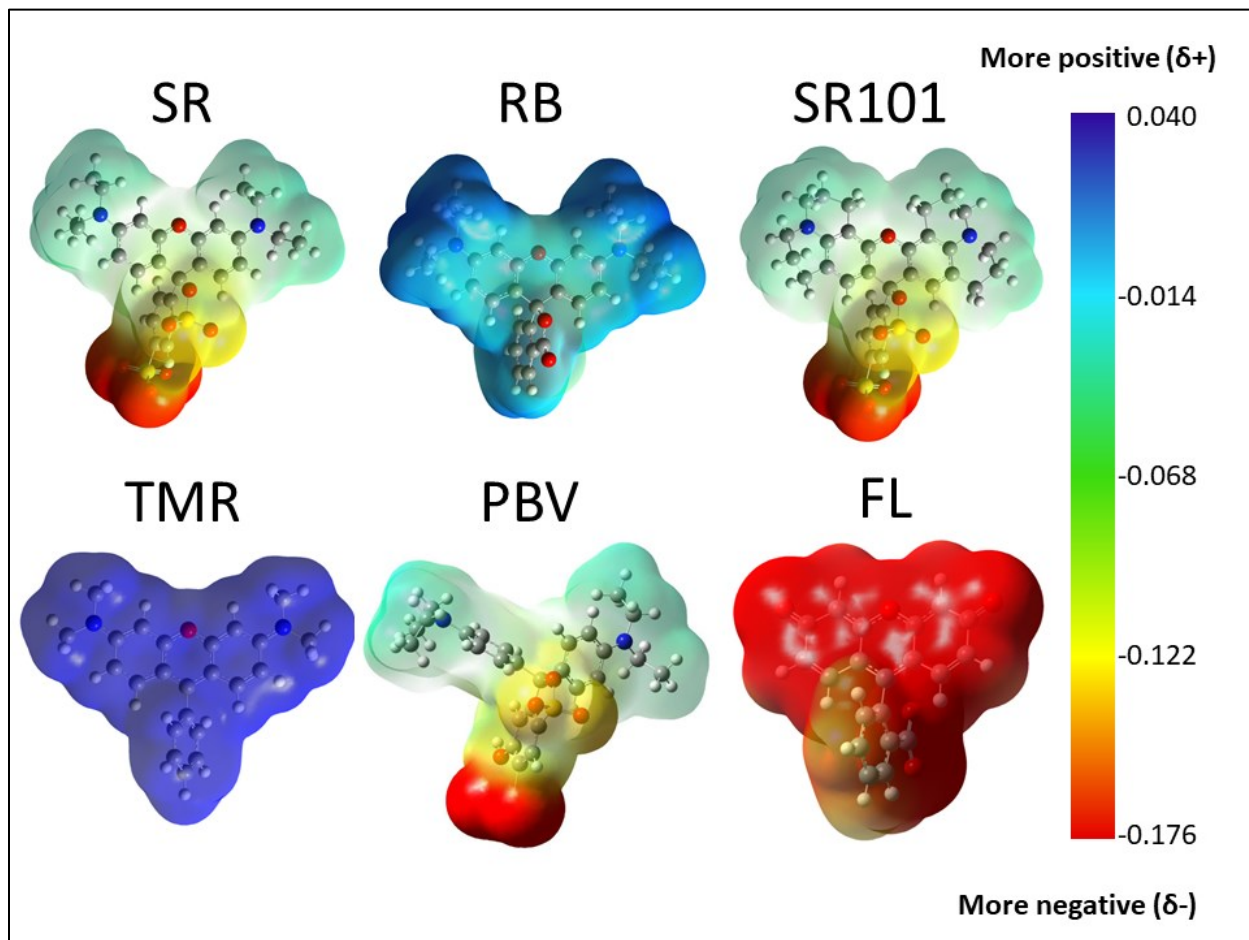


Figure 2.14: Electrostatic potential surfaces for sulforhodamine B (SR), rhodamine B (RB), sulforhodamine 101 (SR101), tetramethylrosamine (TMR), patent blue V (PBV) and fluorescein (FL). The upper charge limit was set by averaging the maximum positive charges in all ligands calculated. The lower charge limit was set by averaging the minimum negative charges in all ligands calculated and subtracting one standard deviation.

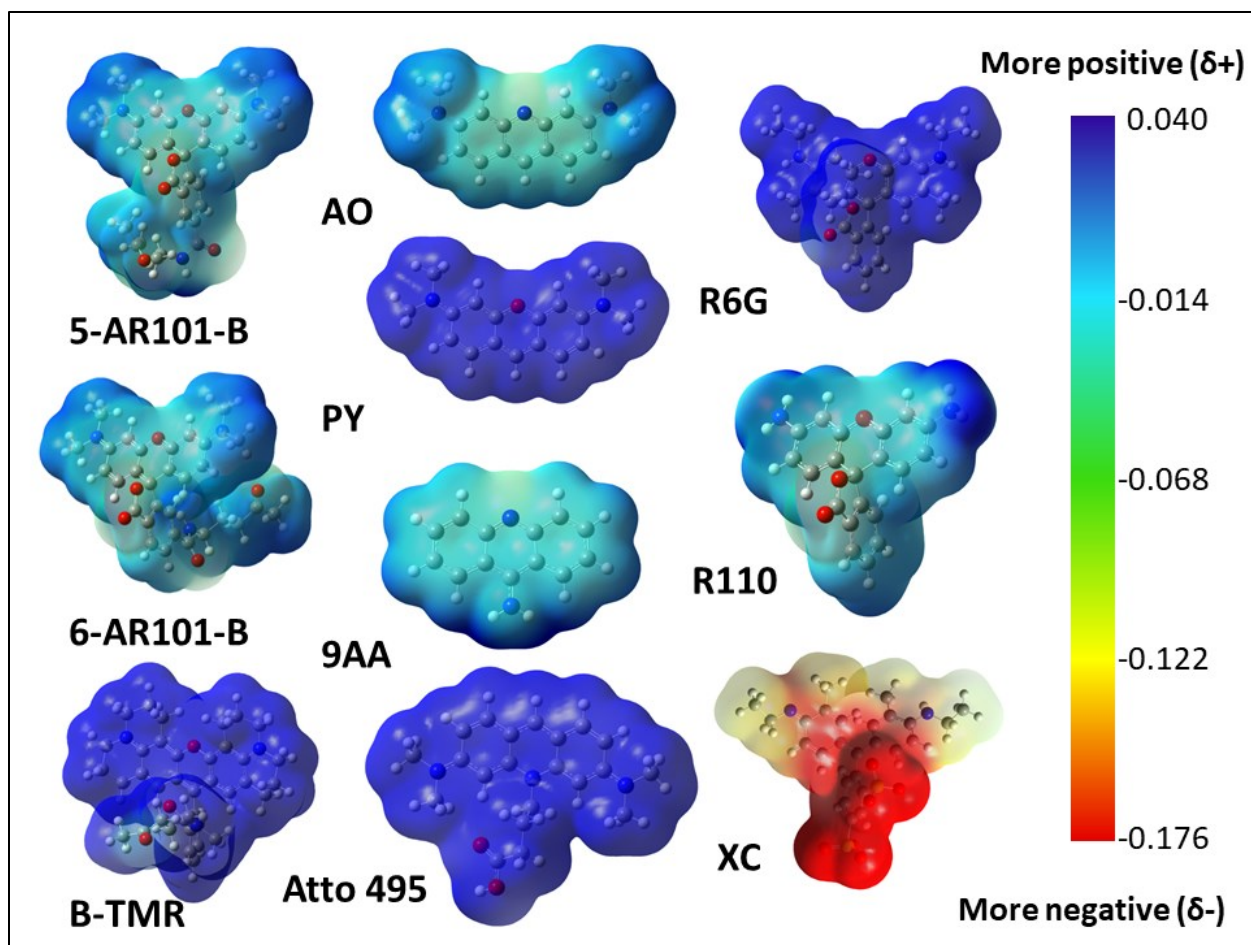


Figure 2.15: Electrostatic potential surfaces of 5-biocytin tetramethylrhodamine, 6-biocytin tetramethylrhodamine (B-TMR), Atto rhodamine 101 biotin (AR), 9-aminoacridine (9AA), Acridine orange (AO), Pyronin Y (PY), Atto 495, Rhodamine 6G (R6G), Rhodamine 110 (R110) and Xylene cyanol FF (XC).

The upper charge limit was set by averaging the maximum positive charges in all ligands calculated. The lower charge limit was set by averaging the minimum negative charges in all ligands calculated and subtracting one standard deviation.

It can also be observed from the fluorescence data in Table 2.2 that SR101 has a similar fluorescence increase but a smaller red shift compared to SR. The smaller red shift results from SR101

being more rigid in its free conformation. This would fit with the ITC results in Table 2.3, which indicate that the entropy loss upon binding is relatively small for SR 101. This is in agreement with the observed K_d values as SR101 is expectedly higher than SR as the tighter the binding pocket, the stronger the π - π stacking interactions will be, resulting in a larger fluorescence redshift. There is also a significant decrease in the entropy loss of SR101 compared to SR. This decrease in entropy loss can be attributed to SR101 having a more rigid ring structure in place of the freely rotating ethyl groups in SR, the more ordered structure results in a smaller entropy loss upon binding. Solvation entropies may also differ between these two ligands. The enthalpy change for SR101 binding is also smaller compared to SR, which can also be explained by the steric interactions between the bulkier rings, resulting in less ideal π - π stacking with the xanthene ring structure of the dye. NOESY experiments indicate that the aptamer is unable to form some contacts with SR101 that are observed with SR and RB, indicating that SR101 doesn't fit as well in the aptamer binding pocket.

2.6.3 Negatively Charged Group is Required for Selectivity but is Likely Not Directly Involved in Binding

TMR binding to SRB-2 resulted in the most significant chemical shift changes in the NOESY spectrum as compared to SR binding. However, even with significant over-titration, multiple conformations appear to be present based on the number of H5-H6 crosspeaks observed. There are also fewer NOEs observed in the SRB-2/TMR NOESY spectrum than in the SRB-2/SR spectrum, indicating that TMR is not inducing the formation of a specific binding pocket at all. TMR is known, from studies with MGA, to intercalate between base pairs of double-stranded RNA [124]. Similar observations were made for the NOESY spectra of PY, 9AA and Atto 495 (Figures A.7, A.8, A.9). Overall, this suggests that all of these small, positively charged dyes intercalate rather than participate in specific adaptive aptamer

binding. These observations also indicate that selective binding of SRB-2 is only possible in the presence of a negatively charged ring well.

SR101 and AR101-B differ slightly in that the para sulfonate is replaced by biotin, and the other is missing entirely. Little difference is observed in the K_d for complexation with SRB-2 between these two dyes, despite a marked change in the electrostatic potential map. This suggests that the negative charge often found in the ortho position of these dyes plays no significant role in binding and likely interacts mostly with solvent, similar to the phosphate groups of ATP when bound to the ATP aptamer [117]. Aside from the biocytin group, B-TMR and SR differ in that B-TMR has a carboxyl group instead of a sulfonate on the single ring and diaminomethyl groups in place of diaminoethyl groups. Based on the SRB-2 binding observations with AR101-B and SR101, the difference in charged groups is expected to have little impact. However, despite B-TMR being more structurally similar to RB, the K_d for complexation with SRB-2 is more similar to that found for SR using ITC. This could be due to the presence of the smaller methyl groups in B-TMR versus ethyl groups in SR, however this cannot be concluded for certain as these K_d values were determined using different methodologies (SPR vs. ITC).

2.6.4 Size and Shape of the Alkyl Groups are Crucial for Binding

Concerning the nature of these alkyl groups, the SRB-2 aptamer has specific requirements. Fluorescein was used in a counter selection step during SELEX [37], resulting in the removal of sequences capable of binding to a dye with electron-withdrawing oxygens rather than the tertiary amines found in SR. The resulting consensus sequence would be expected to have electrostatic interactions with the RNA backbone preventing the correct folding of the aptamer in this case. Dyes such as Xylene cyanol FF (XC) and Patent Blue V (PBV) are missing the bridging oxygen in the middle ring, resulting in a propeller-like structure rather than a planar structure. It is probable then that stacking interactions are much less ideal in this case due to its freedom of mobility and steric hindrance by the

additional methyl groups adjacent to the aminoethyl groups. Interestingly, PBV has been shown to bind SRB-2 weakly [76], indicating a molecule of this shape can fit in the SRB-2 binding site. However, XC cannot bind, and the only difference is that PBV has the traditional diaminoethyl groups while XC has aminoethyl groups, which are secondary amines. Rhodamine 6G has the same aminoethyl groups as XC and also cannot bind, despite being planar like SR. Rhodamine 110 has primary amines and likewise is unable to bind. This suggests that, despite the similarities in electron density, SRB-2 is selective for tertiary amine groups. This fact indicates that with respect to SR101, the steric hindrance of the bulky ring groups is the limiting factor in binding, likely preventing the aptamer from coming in close enough contact to the negative charges for them to have an effect on binding affinity. Supporting this idea are the association and dissociation rates also obtained in LSPR experiments. As seen in Table 2.5, the difference in binding constant between TMR-B and AR101-B clearly comes from the association rate. The aptamer's structure may have to adapt and strain further to allow AR101-B to bind compared to TMR-B, resulting in a slower rate of binding. The ligands are released by the aptamer at a comparable rate due to their relative similarity in charge distribution.

2.6.5 SRB-2 Binding is Salt and Concentration Dependent

Adaptive binding of aptamers is commonly dependent on counter-ions for stabilization of the negatively charged backbone of RNA. In order to investigate this further, ITC experiments were conducted in the absence of divalent cation. Binding was not observed under these conditions for SR, SR101 or RB. Some weak binding was observed for TMR, which is likely due to its positive charge interacting considerably more favourably with the RNA backbone than the other dyes, all of which contain negative charges. However, NMR titrations show ~1:1 binding in the absence of magnesium, indicating that a ~10 to 50 fold higher concentration of aptamer is enough to push the system to bind.

ITC experiments were also run at 500mM KCl (Table 2.4) to observe the ability of excess monovalent cations to assume the role of divalent cations. Binding was observed in the presence of high KCl for all ligands but with lower affinity than with magnesium present. This indicates that a high enough concentration of monovalent cations can result in enough positive charges being present in appropriate positions to assume the role of divalent cations in ligand binding, albeit less effectively.

2.6.6 SRB-2 Lacks any Easily Discernable Tertiary Structure Elements

In attempts to further characterize the structure of SRB-2 and solidify the categorization of its ligands into two distinct groups, DSC and CD experiments were performed.

The rationale behind the DSC experiments was to see if the tertiary structure formed upon ligand binding was able to increase the melting temperature (T_m) of SRB-2 when bound to ligands involved in adaptive binding versus intercalation. Results from isotherms of bound and unbound SRB-2 (Figure 2.20 A) clearly show that the secondary structure of the aptamer melts at 90 °C. This is relatively high for nucleic acids and can certainly be attributed to the presence of the highly stable UUCG tetraloop. In comparing the isotherms of SRB-2 bound the SR and TMR (Figure 2.20 B), their tertiary structures unfold at approximately 67 °C and 69 °C, respectively. They also have similar heat capacities, with TMR's being slightly higher. Based on knowledge from other methodologies, we concluded that the slightly higher T_m and heat capacity of TMR is very likely due to electrostatics rather than formation of stable tertiary structures, and therefore, this line of investigation was abandoned due to the lack of obvious structural differences.

CD was similarly used to assess tertiary structure formation upon SRB-2 binding of the ligands SR and TMR. An ellipticity change is observed in Figure 2.22 A that expectedly indicates stabilization of the aptamer structure when bound to either SR or TMR [195]. However, the general topology of unbound SRB-2 and the two bound conformations are quite similar to each other and are analogous to literature

CD spectra for double-stranded nucleic acids [196,197]. This indicates that no significant tertiary structure elements such as a G-quadruplex exist in this sequence before or after ligand binding. UV absorbance spectra are also nearly identical for the three samples tested (Figure 2.22 B). Small differences in SR compared to TMR and unbound can likely be attributed to minute differences in the aptamer structure or to the ligand itself.

It can be concluded from these methodologies that notable tertiary structure elements such as G-quadruplex are not responsible for ligand binding of SRB-2.

2.7 Conclusion

In summary, we have shown that an ideal ligand for SRB-2 has several requirements. For optimal binding affinity, the amine groups must have small alkyl constituents that are able to participate in hydrophobic interactions but are not so large as to interfere with the aptamer's binding conformation. The absence of a negative charge in the para position on the single ring also contributes positively to binding affinity. A negatively charged group on the ortho position has little impact on binding affinity but is required for selective binding of SRB-2 as repulsive interactions with the RNA backbone prevent these ligands from intercalating in base paired regions of the aptamer. As observed with AR101-B, bulky amine constituents can also confer selective binding of SRB-2 in the absence of this negative charge but at a considerable loss of binding affinity. Taking these observations into account, rhodamine B was shown to be the SRB-2 ligand with the best properties among those tested. We also show that SRB-2 binding is strongly dependent on divalent cations at low concentration and that its binding mechanism does not involve a G-quadruplex or formation of helical regions.

Chapter 3: Characterizing the Solution Structure of the Sulforhodamine B Binding RNA Aptamer by NMR Spectroscopy

All experiments outlined in this chapter were performed and analyzed by the candidate.

3.1 Chapter Abstract

Aptamers are valuable tools in a variety of biotechnological applications, ranging from biosensing to *in vivo* cell imaging. Studying the structure of these molecules can further our understanding of the structure and kinetics of nucleic acid-ligand binding. The SRB-2 aptamer is a polyanionic molecule that binds a ligand with an overall negative charge, making this an interesting binding mechanism to characterize. Here, we show NMR experiments performed on the SRB-2 aptamer with the goal of working towards a solution structure. An unlabelled sample, as well as three samples selectively labelled on a single type of nucleotide (A, C and U) were synthesized and analyzed using a variety of pulse programs. These included multiple variations of HSQC, NOESY, TOCSY and COSY experiments. Some partial assignments were made, but ultimately the data obtained were not sufficient for unambiguous assignment. This structure remains an interesting one to study, and initial NMR studies show some promise, especially if more optimal experimental conditions are found.

3.2 Introduction

Structure determination of biomolecules is invaluable to researchers as it can reveal the intricacies of how these molecules move, fold, function and interact with other molecules. With respect

to aptamers, for example, this information can then be used to elucidate unknown binding mechanisms and to rationally design mutations and/or ligands. This may result in optimization of the binding properties discussed in Chapter 2 for use in desirable applications such as biosensing or cell imaging. As discussed in Chapter 2, the SRB-2 aptamer is of particular interest due to both its ligand promiscuity and ability to bind ligands with an overall negative charge. For these reasons, we set out to investigate the structure of this aptamer using NMR spectroscopy.

NMR spectroscopy is a powerful tool to study biomolecules in solution. As shown in Chapter 2, it is quite useful for mapping structural changes in varying solution conditions. It will also be shown in Chapter 5 that it is a helpful technique to use, particularly in combinations with others like CD, to assess the presence of specific sequence motifs such as G-quadruplex or triplex DNA [198,199]. But perhaps most valuable is its ability to determine high-resolution 3D structures, offering an alternative, and often complementary, method to x-ray crystallography [200,201]. NMR is particularly useful for studying molecules that are poor candidates for crystallization, such as those that are flexible or have unorganized intermolecular interactions. Proteins have a diverse set of functional groups that can result in charge complementarity between molecules, while nucleic acids are primarily covered in negative charges which results in intermolecular repulsion. This makes it difficult to pack them into an appropriate crystal lattice structure for x-ray analysis [202]. For this reason, a much higher percentage of nucleic acid structures (37.8%) are determined by NMR compared to proteins (7.5%) and protein-nucleic acid complexes (18.0%) as shown in Table 3.1 [203]. Cryo-EM is another methodology for structure determination that is currently on the rise [204]. It has proven quite powerful, particularly in studying membrane-bound proteins, which previously lacked a reliable methodology for structure determination. It is a promising technique for studying nucleic acids as well, though only a handful of structures are available thus far [203]. NMR is the technique of interest in this case, so some assignment strategies will

be reviewed below. The general scheme for this overview is loosely based on volume 338 of the Methods in Enzymology book series [205].

Table 3.1: PDB Statistics by Molecular Type as of September 21st, 2021 [200].

Molecular Type	X-ray	NMR	EM	Multiple	Neutron	Other	Total
Protein (only)	141397	11638	5718	176	70	32	159165
Protein/Oligosaccharide	8370	31	913	5	0	0	9319
Protein/NA	7468	271	1943	3	0	0	9685
Nucleic acid (only)	2362	1369	53	8	2	1	3802
Other	149	31	3	0	0	0	183
Oligosaccharide (only)	11	6	0	1	0	4	22

A critical type of experiment in the characterization of 3D molecule structures is nuclear Overhauser effect spectroscopy (NOESY). Through-bond couplings (J-coupling) are studied in most NMR experiments, but NOESY takes advantage of ¹H-¹H dipolar couplings. This type of coupling provides a means for cross-relaxation to occur, resulting in the detection of resonances that are close in space, regardless of connectivity [206,207]. Peak intensity, *I*, can then be correlated to inter-proton distance, *r*, by the sixth inverse power, ($I \propto 1/r^6$), which in turn can be used to set distance restraints for structure calculations [208]. NOESY experiments are extremely useful for studying tertiary structures of biomolecules or receptor-ligand interactions because it allows resonances that are very far apart in the molecule or in different molecules entirely to be correlated.

With respect to structure determination of nucleic acid-ligand complexes, the first step is to prepare a sample of the sequence fully bound to ligand. To determine a high-resolution structure, there must be only a single conformation present to prevent spectral overlap and ambiguity in the resonance assignment. This is done by titrating of the sequence of interest with aliquots of lyophilized ligand. ¹H NMR spectra are acquired for each aliquot and the titration continues until changes in the imino region

of the spectrum are no longer observed upon addition. A chart illustrating the chemical shift ranges of nucleic acids is shown in Figure 3.1.

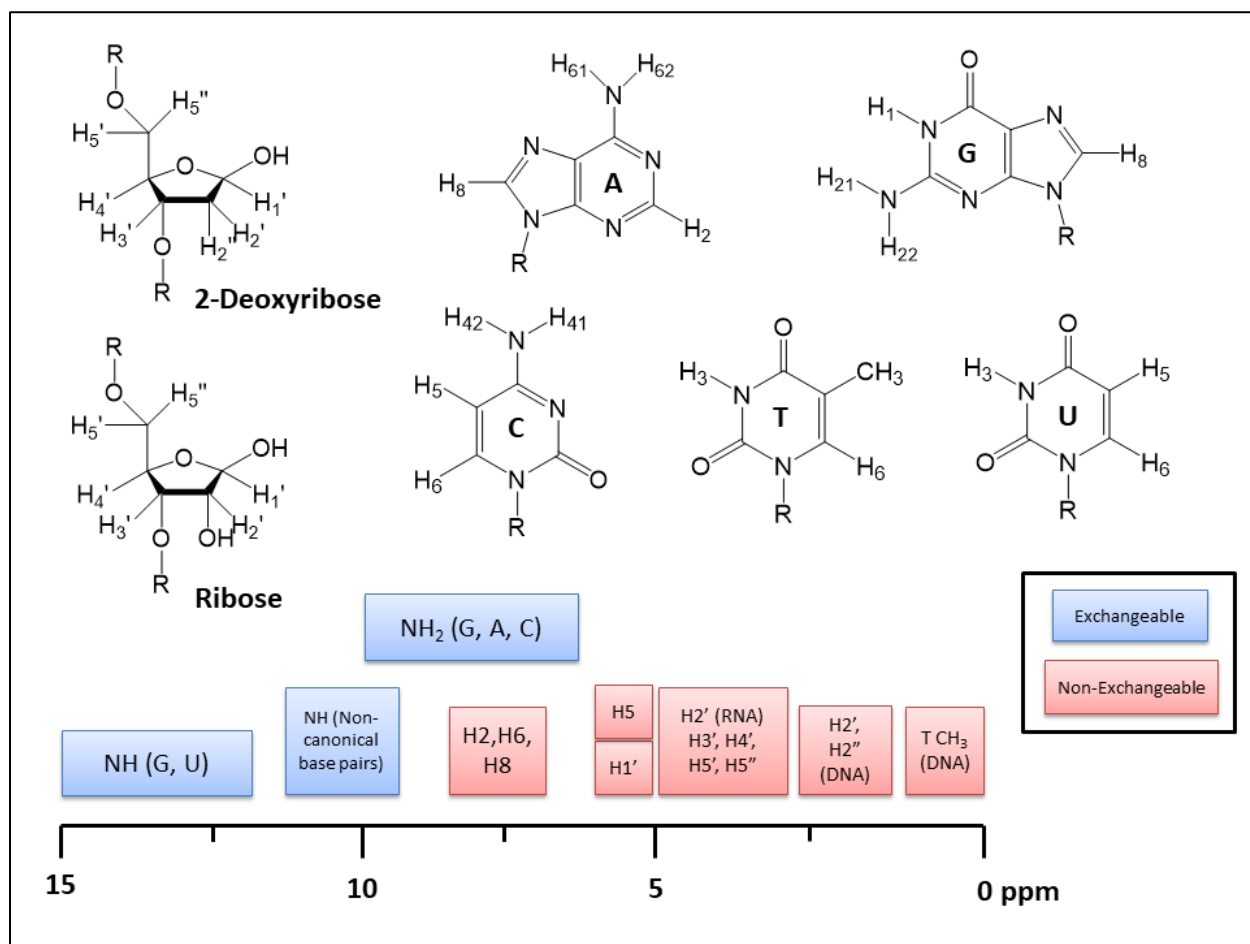


Figure 3.1: Proton chemical shift ranges of nucleic acids.

Blue and red boxes show the approximate literature chemical shift ranges for each type of proton in RNA and DNA. A- Adenine, G- guanine, C- cytosine, T- thymine, U-uracil.

Resonance assignment in unlabelled nucleic acids is possible but generally limited to smaller molecules. This type of assignment involves the contemporary through bond experiments DQF COSY [209] and TOCSY [193], as well as through-space NOESY [191] experiments. One generally begins by acquiring a NOESY in 90% H₂O/10% D₂O [210], since the sample is titrated in these conditions to observe changes of the exchangeable protons. This NOESY spectrum is used to establish base-pairing by

assigning imino-imino and amino-amino NOEs. It is also used to assign imino to base and imino to sugar correlations. The sample is then lyophilized, resuspended in 100 % D₂O and NOESY, TOCSY and DQF COSY experiments are acquired. The D₂O experiment is the primary source of sequential assignments. This is done by 'walking' through the H6/H8-H1' region in the NOESY, as described in Figure 3.2. The H6/H8-H2' region is also useful for this purpose, particularly in DNA where H2' and H2'' are well resolved from the other sugar protons. Methyl groups from thymine residues may also aid in the sequential assignment of DNA. TOCSY and/or COSY experiments allow the identification of C/U base spin systems (H5/H6) that also appear in this region. TOCSY and COSY are also used to identify sugar spin systems and provide information about sugar puckering. Lastly, ³¹P experiments may be performed on unlabelled samples since this isotope, which has the highest natural abundance, is an NMR-active nucleus. These experiments, usually heteronuclear correlation sequences (HETCOR) [211], are used to confirm or extend sugar spin systems and assign phosphorus resonances. In short sequences, the experiments listed may be sufficient for structural determination. However, in longer sequences, the spectral overlap becomes too great and additional data is required.

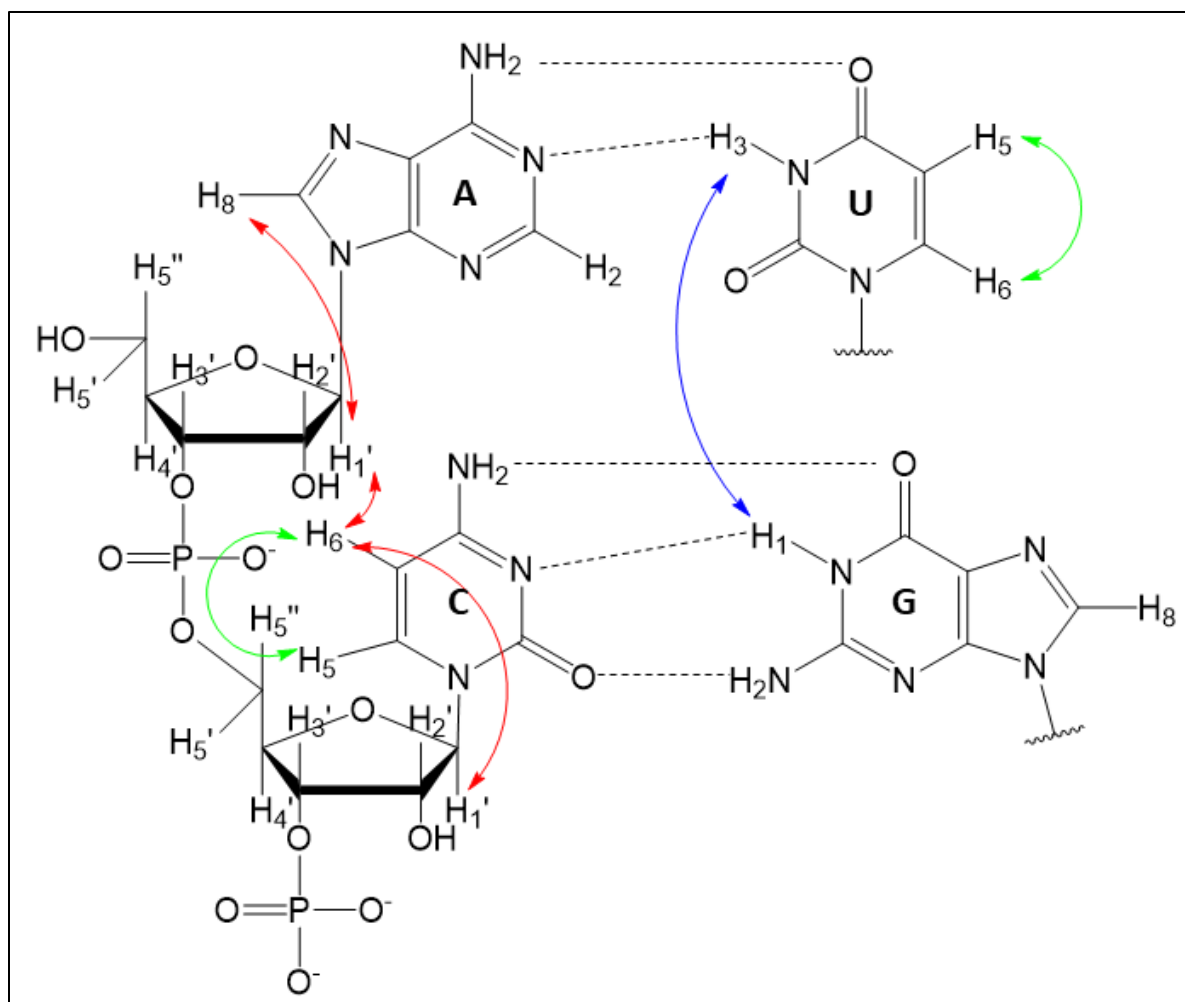


Figure 3.2: Important interactions in RNA backbone assignment.

The NOEs used to 'walk' between sugars and bases (H1'-H6/H8) are shown with red arrows. The H5-H6 TOCSY cross-peaks are shown with green arrows. These signals are used to differentiate pyrimidines from purines in the NOESY walk. Based-paired regions of RNA can be sequentially assigned using imino-imino interactions between stacked bases. An example of this type of interaction is shown with a blue arrow.

Isotopic labelling with ^{13}C and ^{15}N is an extremely powerful method for deconvolution of more complex spectra, particularly for RNA. The enzymatic synthesis of RNA by T7 RNA polymerase allows for simple incorporation of isotopically labelled nucleotide triphosphates which can be isolated from *E.coli*

cells grown on ^{15}N ammonium chloride and ^{13}C glucose, for example. This allows for simple, uniform labelling of single or multiple nucleotide types. Labelling of DNA is not so straightforward due to the lack of viable precursors readily available for solid-phase phosphoramidite chemistry, though it does have the advantage of labelling by position rather than being limited to the type of nucleotide. SRB-2 is an RNA aptamer, so we will focus on RNA methodologies here.

As indicated above, isotopically labelled nucleoside triphosphates are incorporated into RNA sequences via *in vitro* transcription. As a result, any combination of the four bases may be labelled in a given synthesis. Because the ligand titration is performed in water, it is generally most efficient to run all experiments that are required to be in H_2O first, then lyophilize and switch over to D_2O . For this reason, exchangeable proton/nitrogen correlations are generally acquired first.

2D ^1H - ^{15}N -HSQC experiments [212,213] are used to correlate imino (G N1H, U N3H) and amino (A N6H₂, C N4H₂, G N2H₂) protons to their respective nitrogen atoms. A 2D HNN COSY [214] or variation thereof may be used to directly identify any hydrogen bonds involving two nitrogen atoms. Sequential assignment of exchangeable proton/nitrogen resonances in base paired regions of RNA may be accomplished using 3D ^1H - ^{15}N -NOESY-HSQC [213] type experiments. Edited versions of 3D ^1H - ^{15}N -NOESY-HSQC may be used to correlate imino and amino resonances to non-exchangeable protons. This can also be done using a through bond experiment such as HCCNH or HCCNH-TOCSY [215].

Once information about exchangeable resonances is acquired, the next step is the sequential assignment of sugar-base correlations. There are two main strategies for doing this, NOE-based assignment and through-bond coherence transfer assignment. NOE-based assignment involves filtered and edited type NOESY experiments, where protons attached to ^{13}C are either filtered out or selected for, respectively. Filters can be applied in both F1 and F2, resulting in experiments selecting for protons that originate on a specified nucleus, end up on a specified nucleus or both. These filters are most often applied to 2D NOESY [191] or 3D ^1H - ^{13}C -HSQC-NOESY pulse sequences. With a fully labelled sample, a 3D

^1H - ^{13}C -HSQC-NOESY can also be used to distinguish NOEs originating from the (unlabelled) ligand. The through-bond approach involves triple resonance experiments that allow correlation between the base and sugar of each nucleotide, including HCN [216], HCNH [217], HCNCH [218] experiments. HMBC [219] may also be used to correlate H2/H8 protons in purines. Sequential sugar-phosphate assignments may be completed using ^1H , ^{13}C , ^{31}P experiments such as HCP, PCCH-TOCSY and HPHCH [220–224].

After sequential assignment, the next step is to assign carbon resonances and the remaining sugar protons. 2D ^1H - ^{13}C -HSQC [212,213] experiments are used to identify carbon-proton spin systems, and remaining carbon and proton resonances, particularly those in the sugars, are assigned using either HCCH-TOCSY [225] or HCCH-COSY [226] experiments. Finally, ^{31}P resonances are assigned via HETCOR if ^1H , ^{13}C , ^{31}P experiments were not already acquired.

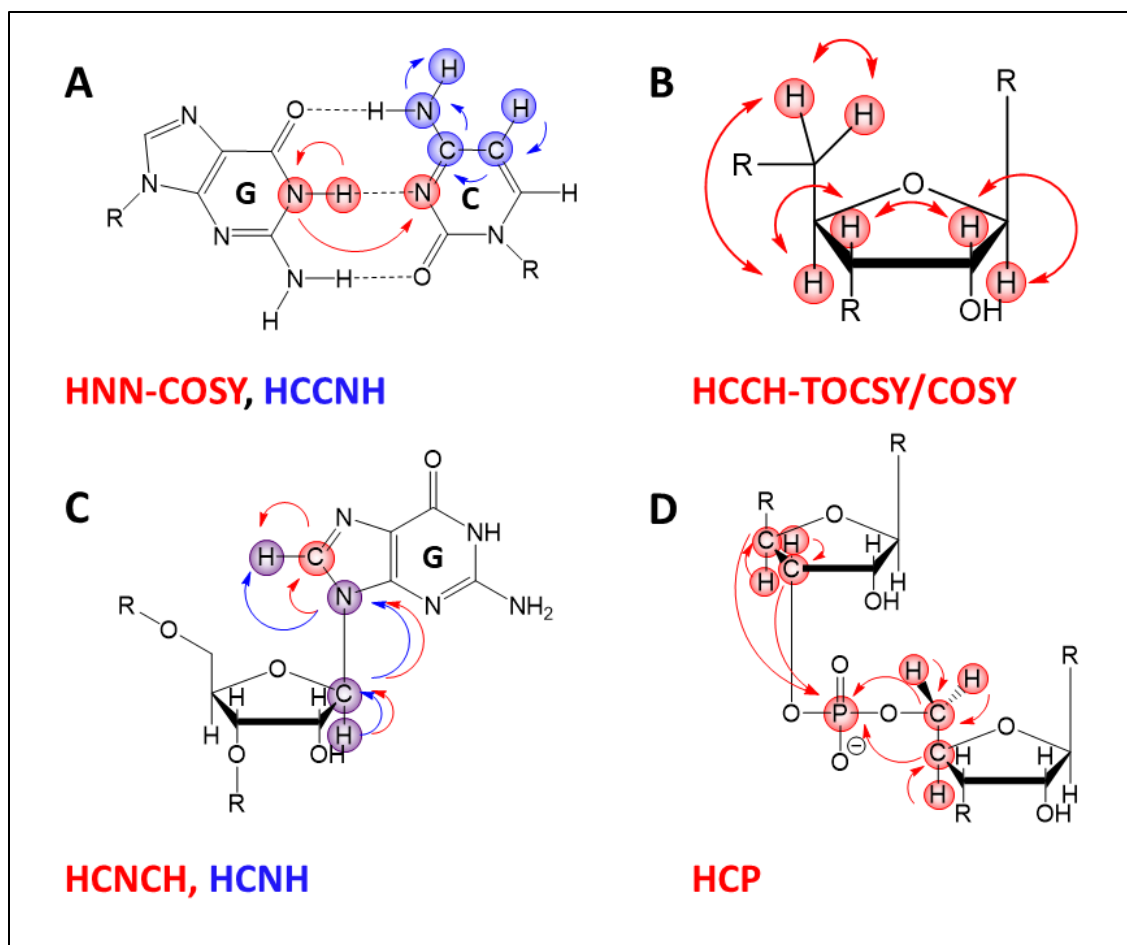


Figure 3.3: Heteronuclear through-bond NMR experiments used for resonance assignment of nucleic acids.

The transfers of magnetization in each type of experiment are shown by arrows. Relevant nuclei are colored to match the name of their pulse sequence. A) HNN-COSY [214], which correlates the hydrogen bond donor imino group in either G or U to the hydrogen bond acceptor nitrogen in C or A (shown in red for G-C base pair). Exchangeable protons can also be correlated to non-exchangeable protons in HCCNH or HCCNH-TOCSY [215] experiments. An example of an H5-H3 correlation in cytidine is shown in blue. B) HCCH-TOCSY [225] (shown in red) and COSY [226] experiments are used to assign sugar spin systems. Pulse programs for detection of ^{13}C are also available. C) Correlation of sugar to base resonances may be achieved using triple resonance experiments such as HCNCH [218] (shown in red) and HCNH [217]

(shown in blue, overlap with HCNCH is shown in purple). It is also possible to obtain the same information from two separate HCN [216] experiments, one optimized for the ribose and the other optimized for the base. D) HCP [221–223] experiments correlate H3'/C3', H4'/C4' on the 5' side and H4'/C4' and H5',H5''/C5' resonances on the 3' side of a given phosphorus atom. An HCCP-TOCSY [220,221,224] experiment may be used to extend assignments to the rest of the ribose.

This is by no means a step-by-step guide, as the experiments required to assign RNA will vary significantly with sequence complexity. There are also many other pulse sequences available that may be useful depending on gaps of information that need to be filled for a particular project. Also, it is important to note that the types of labelled sequences synthesized will affect the relevant pulse programs that one may use. A common strategy is to acquire spectra with fully labelled ^{13}C , ^{15}N and/or ^{15}N samples and either single or dual nucleotide ^{13}C , ^{15}N labelled samples to allow unambiguous assignment of resonances belonging to each type of base. That being said, isotopic labelling is rather expensive, so it is generally the goal of a researcher to assign the sequence with as few labelled samples as possible.

Another valuable labelling technique not covered above is deuterium (^2H) labelling. For nucleic acids, this generally refers to either site-specific ribose (D3', D4', D5' and D5'') or base deuteration/enrichment of perdeuterated bases. Partial deuteration of the ribose simplifies a NOESY spectrum significantly, as removing spins unimportant for sequential assignment drastically reduces the amount of dipolar relaxation experienced by the other protons, decreasing signal linewidth [227]. Due to the established lability of H8 protons, particularly under alkaline conditions, it is possible to either selectively deuterate fully protonated purine bases or selectively proton-enrich perdeuterated bases [228]. Procedures have also been reported for H5 or H6 deuteration with pyrimidine 5'-monophosphates [229,230]. This type of labelling results in relevant NOESY regions, including only

signals from particular bases, thereby simplifying assignment. Deuterating the H5 or H6 position has an added linewidth reduction benefit, similar to site-specific ribose labelling. This is due to a combination of large dipolar and scalar couplings between these positions [231].

In our work with the SRB-2 aptamer, we acquired the standard homonuclear spectra listed above, as well as some of the heteronuclear spectra discussed with three different single-nucleotide ¹³C, ¹⁵N samples (A, C, U). Some regions of these spectra had unusually broad peaks, which is generally the result of chemical exchange. In fast chemical exchange, only one signal that represents the average of the contributions from each environment is observed. In slow chemical exchange, multiple sharp signals for each environment are observed. In intermediate exchange, one or multiple broad peaks may be observed. Because this is an aptamer-ligand complex, the chemical exchange and hence broad lines being observed is likely due to the presence of at least one additional aptamer conformation. This, among other reasons, resulted in the spectra listed above being of insufficient quality to assign unambiguously.

3.3 Materials and Methods

3.3.1 Preparation of RNA Samples and Dyes

The SRB-2 RNA was synthesized enzymatically using a T7 RNA polymerase and a double-stranded synthetic DNA template (Integrated DNA Technologies, Inc., Coralville, Iowa). The sequences used were as follows:

SRB-2 template DNA:

5'-GGGACCTGAGGCGGTTAACCTTGC GCCTCTCCATCATCGCCGAAGCGAGGTCCCTATAGTGAGTCGTATTA-3'

SRB-2 coding DNA:

5'-TAATACGACTCACTATAGGGACCTCGCTTCGGCGATGGAGAGGCGCAAGGTTAACCGCCTCAGGTCCC-3'

A single base-pair substitution in the original sequence of SRB-2 (A3 → G3 and U52 → C52) was introduced to increase transcriptional yield, as the presence of three consecutive G residues at the beginning of sequence is thought to increase initiation of transcription [187,188]. RNA was transcribed using the recipe listed in Table 2.1, and labelled samples were synthesized by incorporating the respective labelled nucleotide triphosphates in place of their unlabelled version. Single nucleotide labelled samples were synthesized with A, C and U. The RNA was purified on a 10% Urea PAGE gel and the band containing the aptamer was cut out. The RNA was eluted from the gel by crush & soak in 300mM NaCl or by electroelution. This was followed by clean-up on a HiPrep 16/10 DEAE FF anion-exchange column (GE Healthcare, Uppsala, Sweden) and desalting on a HiPrep 26/10 Desalting column (GE Healthcare, Uppsala, Sweden). The RNA was precipitated with 70% ethanol prior to running each column. Pure samples obtained from the desalting column were then lyophilized and dissolved in 500µL of 10 mM potassium phosphate and 10 mM KCl (90% H₂O/10% D₂O).

3.3.2 NMR Experiments

Lyophilized aliquots of Sulforhodamine B (Life Technologies, Eugene, OR) were added to the 90% H₂O/10% D₂O samples to titrate the RNA to slightly above a 1:1 ratio. Once experiments examining exchangeable resonances were performed, the samples were dried by lyophilization and re-dissolved in 500 µL of 99.996% D₂O (Cambridge Isotopes) to perform experiments on non-exchangeable resonances. All spectra were collected on a Bruker DRX-600 spectrometer equipped with an HCN triple-resonance, triple-axis PFG probe. Quadrature detection for the indirect dimensions in multidimensional experiments was achieved using the States-TPPI method [189]. Samples in 90% H₂O/10% D₂O that were used to observe exchangeable protons were acquired at 277 K, and water suppression was achieved using a 1 $\bar{1}$ -spin echo [190]. As for the experiments in D₂O that were used to observe non-exchangeable

protons, 2D NOESY [191], 2D CITY-TOCSY [193] and DQF COSY [209] were run with presaturation solvent suppression [192]. Double half x-filtered and F2 half x-filtered NOESY experiments [232,233] used the WATERGATE suppression sequence [234]. All other heteronuclear experiments had echo-antiecho coherence selection where solvent signals are suppressed using the inherent dephasing effects of pulsed field gradients [235]. All 2D NOESY spectra [191] were acquired with a mixing time of 150 ms, and TOCSY experiments were acquired with a mixing time of 50 ms. All NMR samples had a volume of 500 μ L and were read in standard 5 mm NMR tubes. The labelled C sample was an exception with a volume of 200 μ L and was read in a 5mm Shigemi tube with 8mm bottom matched to D₂O. Relevant pulse programs can be found in Appendix B.

3.4 Results

3.4.1 Homonuclear NMR Experiments

As discussed in Chapter 1, aptamers are known to use an adaptive binding mechanism [117–122]. This means that they undergo a significant conformational change upon binding their ligand. For this reason, the first step of the NMR studies was to titrate an SRB-2 sample with SR. This was done by preparing three lyophilized aliquots of ligand that each contained enough ligand to bind one-third of the RNA sample. ¹H NMR spectra were acquired for the RNA sample on its own (Figure 3.4 D) and after each aliquot of SR was added (Figure 3.4 A-C). The imino range of these spectra is shown in Figure 3.4 because these protons have the highest degree of variation in NMR when tertiary structure is formed or altered. These resonances can exchange with the solvent and have inherently broad lines as a result. They are also mostly unobservable in NMR in the absence of hydrogen bonding. In adaptive binding, the aptamer folds tightly around the ligand. This results in more opportunities for imino protons in single-stranded regions to hydrogen bond with other parts of the molecule and protects imino protons in general from being exposed to solvent. Therefore, observing new peaks forming and/or existing peaks

sharpening in this region is a good way to observe changes in nucleic acid structures upon binding ligands. For SRB-2, it was observed that the linewidths of the peaks decreased throughout the titration, and clear changes in chemical shift and intensity of peaks could be identified. This suggests a change in conformation is occurring upon ligand binding.

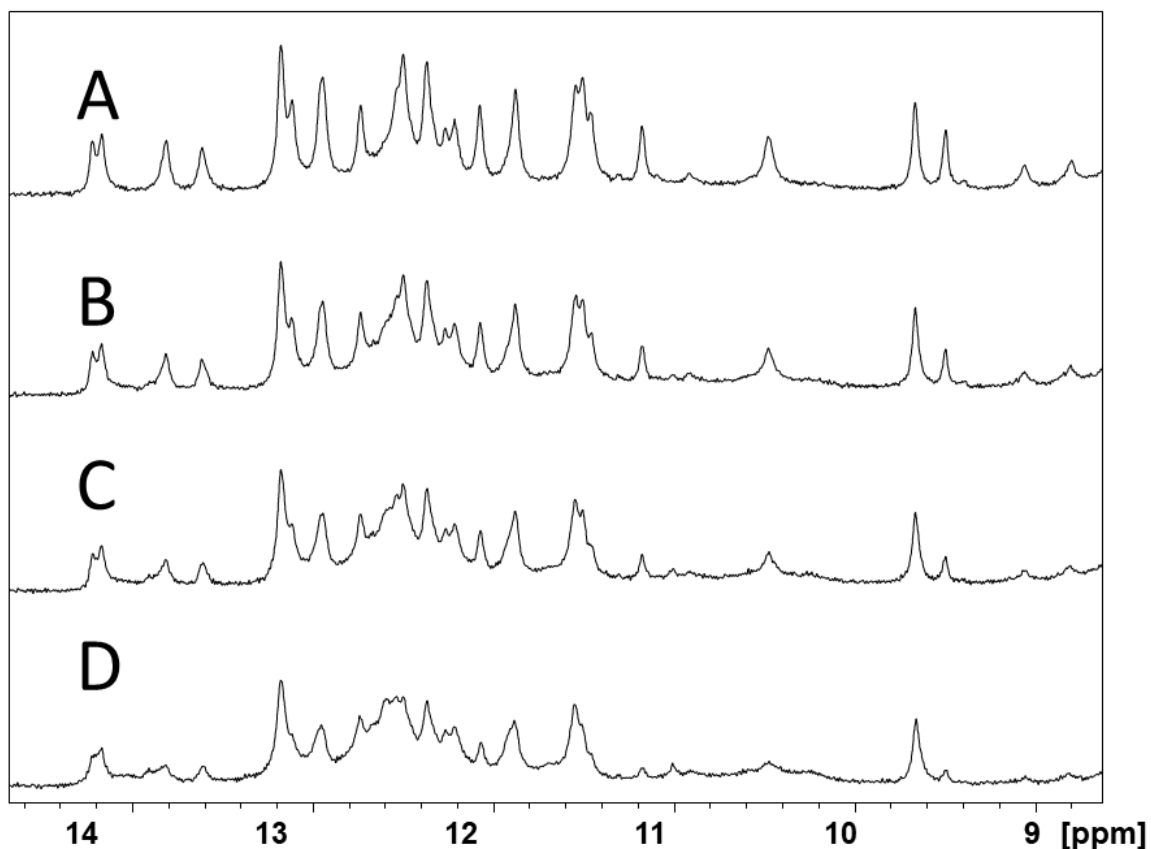


Figure 3.4: 1D titration of 1 mM SRB-2 with SR.

Imino proton range showing the change in structure upon binding of the ligand in three aliquots. A) SRB-2 fully bound, B) SRB-2 two-thirds bound, C) SRB-2 one-third bound, and D) SRB-2 completely unbound.

While the sample was still in 90% H₂O/10% D₂O, a 2D NOESY was acquired (Figure 3.5). Initially, NOESY experiments may be useful in assessing the likelihood of assignment based on the number of NOEs that are present and their resolution. Additionally, folding simulators such as UNAFold [236] are

often used to predict secondary structure formation in nucleic acids and looking at the number of G and U imino resonances present may give an indication of whether or not predicted stems are forming.

Ultimately, the NOESY in 90% H₂O/10% D₂O is used to assign NOEs involving imino and amino protons.

With respect to SRB-2, a significant number of NOEs from each imino proton were observed, even at lower RNA concentrations. Partial assignments of the stem regions were made from these NOEs, including unambiguous assignment of the UUCG tetraloop as shown in Figure 3.5 B. After the acquisition of this experiment, the sample was lyophilized and transferred into D₂O.

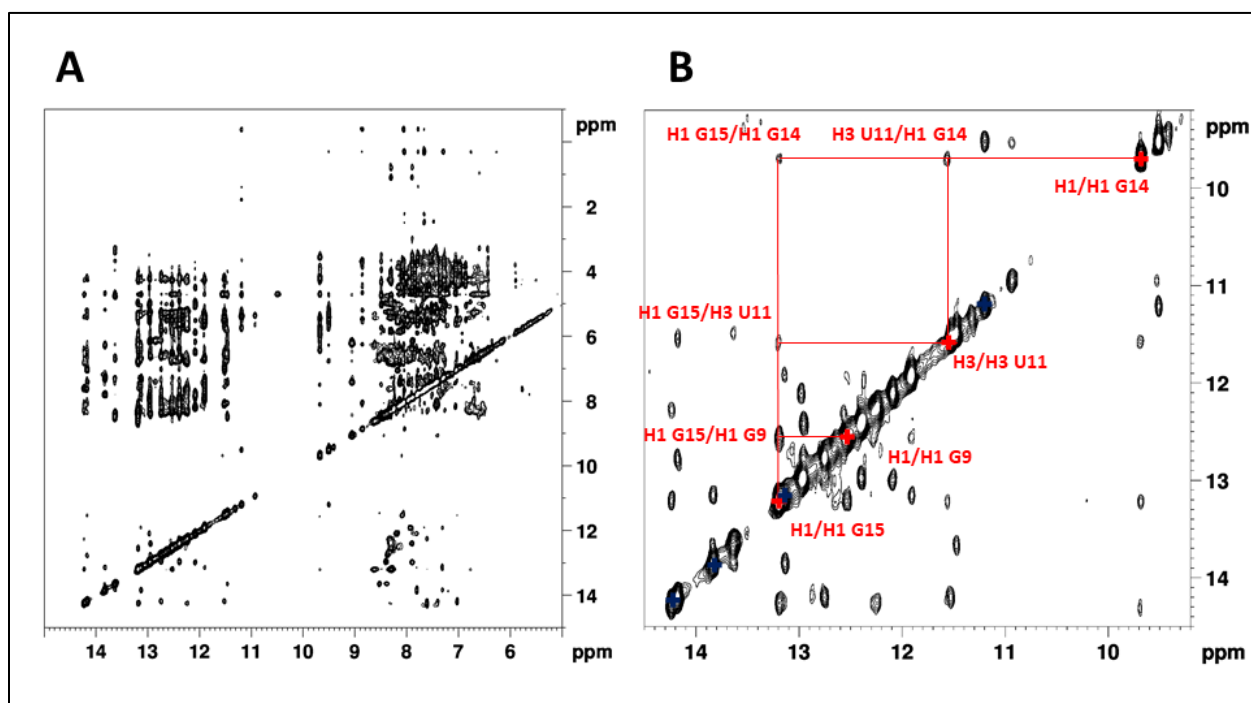


Figure 3.5: NOESY of 1.8 mM SRB-2 bound to SR in 90% H₂O/10% D₂O.

A) Full spectrum. B) Imino-to-imino region. Assignments for the UUCG tetraloop determined in Chapter 4 are shown in red, and U H3 protons are labelled blue based on chemical shifts observed in the ¹H-¹⁵N HSQC spectrum.

Recalling the initial titration, a clear change in conformation was observed. However, we must be sure that this is a single conformation and that an equilibrium of multiple conformations is not what was observed. When multiple conformers are present, the resulting overlap renders multidimensional spectra unassignable. In order to determine whether or not SRB-2 was in a single bound conformation, a TOCSY in 100% D₂O was acquired (Figure 3.6). Cytosine and uridine bases have H5 and H6 protons, giving rise to a strong cross-peak in TOCSY experiments due to the 3-bond j-coupling. Therefore, the number of H5-H6 cross-peaks observed in a TOCSY is correlated to the number of U and C residues in the sequence. If significantly more peaks are observed, then there are likely multiple conformations present. For example, in the SRB-2 spectrum, approximately 32 peaks were counted in this region. There are 16 cytosine and 9 uridine residues. Therefore, the number of peaks observed exceeds the expected number, which indicates that multiple conformations are likely present. Having a single conformation present is generally a required condition for assignment. Due to several of these peaks being substantially weaker than most, we carried on with NMR experiments and classified these peaks as being from alternative conformations as demonstrated in Figure 3.6.

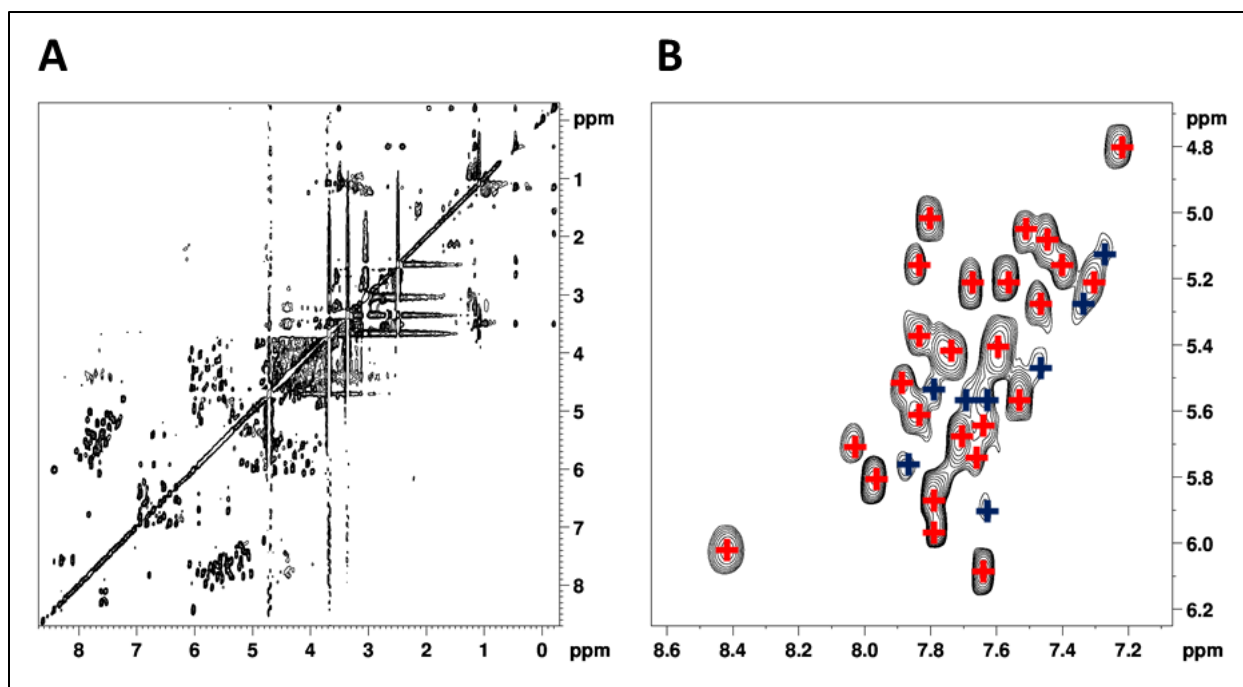


Figure 3.6: TOCSY of 0.6 mM SRB-2 bound to SR.

A) Full spectrum. B) Close-up of the H5-H6 region. Strong peaks that are likely from the dominant conformation are marked as red, and weaker ones from other conformations are marked as blue.

A DQF COSY was acquired as well (Figure 3.7), which provides similar information to the TOCSY with respect to H5-H6 cross-peaks. The TOCSY, however, tends to be easier to resolve in this respect due to the peaks being in-phase. Because of the antiphase multiplet structure observed in COSY [237], destruction of overlapped, opposite phase signals is possible, resulting in more difficult assignment. However, COSY experiments are extremely valuable when doing structure calculations because of the information they provide regarding J-coupling constants. For example, the presence and magnitude of $^3J_{H1'-H2'}$ and $^3J_{H3'-H4'}$ couplings indicates the type of sugar pucker that a particular residue has adopted (C3'-endo, C2'-endo or an intermediate) [238]. In this spectrum, 22 H5-H6 cross-peaks were counted. This is less than the expected 25, but due to the antiphase multiplet structure, it is possible that some of the signals in close proximity were cancelling each other.

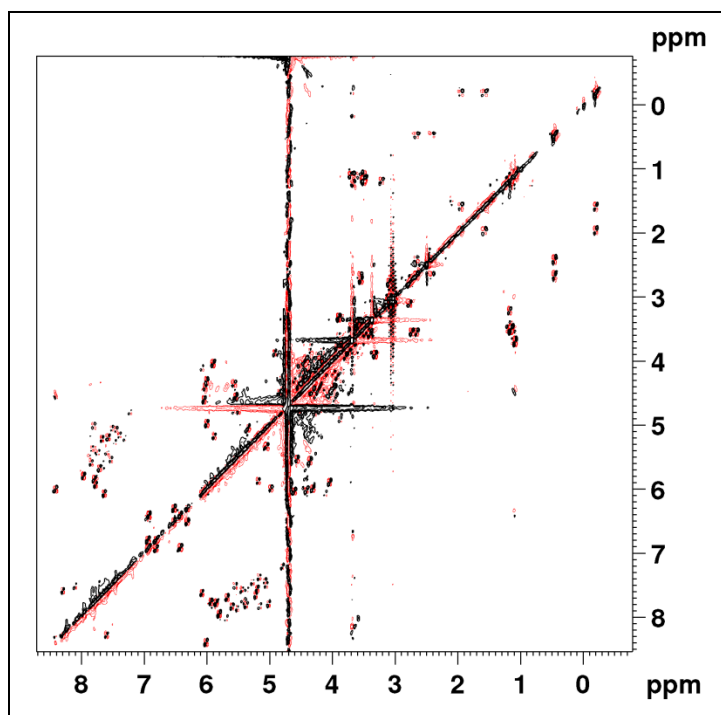


Figure 3.7: DQF COSY of 0.6 mM SRB-2 bound to SR.

Positive peaks are shown in black and negative peaks are shown in red.

Now that the sample was in D_2O , a second NOESY was acquired, this time optimized for the chemical shift range of non-exchangeable protons in the RNA and ligand (Figure 3.8). Non-exchangeable protons are plentiful in nucleic acids, resulting in the NOESY in D_2O containing a vast amount of information. In short sequences, residues are sequentially assigned using a sequential walk in the H6/H8-H1' region of the spectrum. The location and orientation of the sequence in this assignment process is determined with the help of the TOCSY as it highlights C and U residues. Once the backbone sequence has been determined, assignments are extended throughout the ribose, resulting in a large number of NOEs to be used in structure calculations. Therefore, with regards to short oligonucleotide sequences, the NOESY can be responsible for the overwhelming bulk of assignments. This highlights the significance of acquiring a high-resolution NOESY spectrum in D_2O . In larger sequences such as SRB-2, the overlap is significant, and unambiguous sequential assignment using the base-sugar walk is not

possible. However, the NOESY is still critical as other experiments such as filtered/edited NOESYs are often overlaid on it for assignment. The spectrum acquired for SRB-2 is of decent quality and many strong NOEs are present, indicating that the SR-SRB-2 complex is highly structured. However, there are regions in this spectrum with broad peaks. Some of these may be a result of overlap, but this also suggests, like the TOCSY, that there is some chemical exchange happening here.

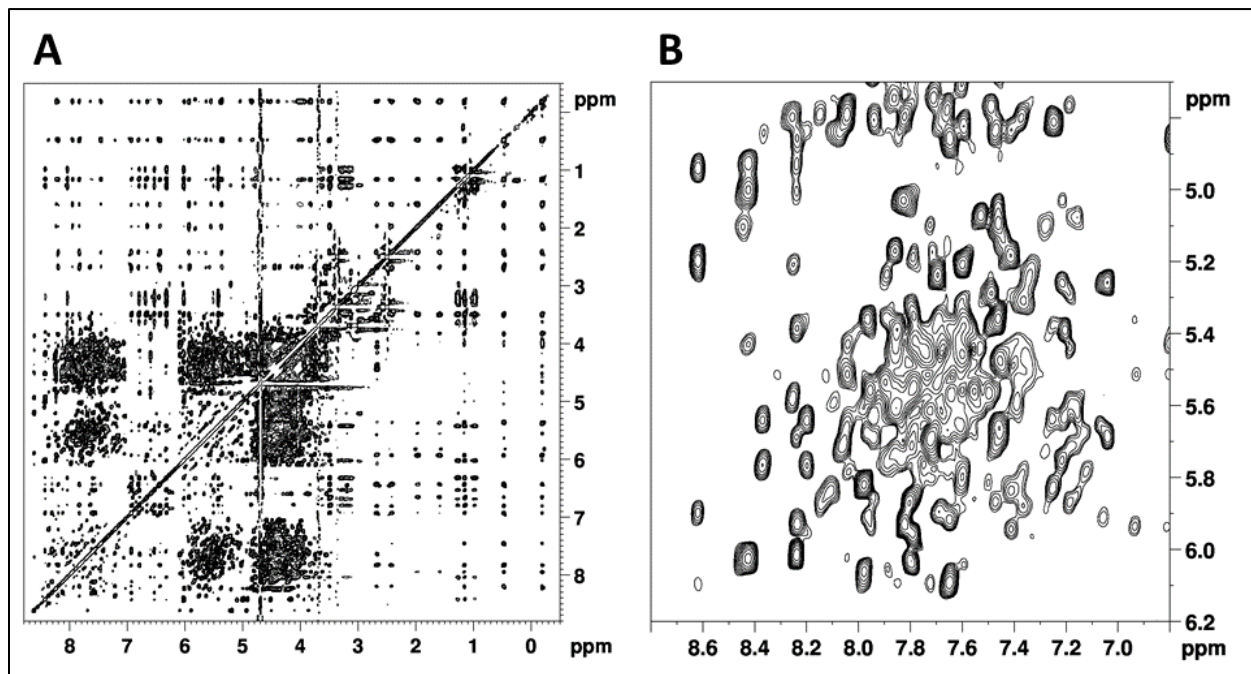


Figure 3.8: NOESY of 1.8 mM SRB-2 bound to SR in D₂O.

A) Full spectrum. B) Close-up of the H1' to H6/H8 region. Mixing time: 150ms.

SRB-2 was originally selected using a buffer that contained MgCl₂, so in an attempt to reduce the chemical exchange observed in Figure 3.8, we acquired a NOESY in D₂O that contained 2 mM MgCl₂ (Figure 3.9). Some minor changes in the spectrum were observed when compared to the spectrum acquired in the absence of MgCl₂, but this did not help significantly with the problem of spectral overlap. MgCl₂ was not used in future experiments because despite its role as an important cofactor, Mg²⁺ is also known to catalyze RNA degradation [239].

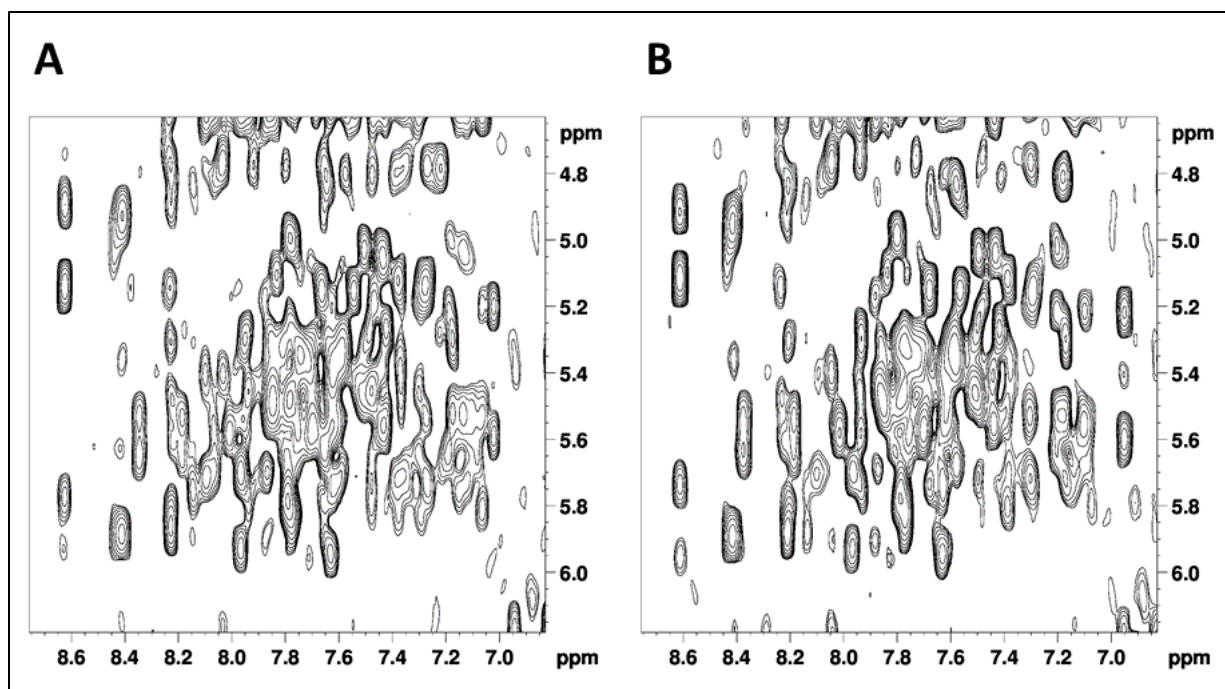


Figure 3.9: Effect of Mg^{2+} on NOESY of 0.8 mM SRB-2.

A) H1'-H6/H8 region of NOESY in the absence of Mg^{2+} B) H1'-H6/H8 region of NOESY with 2 mM MgCl.

Mixing times: 150 ms.

In another attempt to reduce the amount of chemical exchange observed, we took a cue from Chapter 2 and acquired a NOESY in D_2O of SRB-2 bound to the best ligand, RB (Figure 3.10). This spectrum looked near identical to the one with SRB-2 bound to SR in Figure 3.8, albeit with a slightly lower concentration. Little to no difference was observed in the amount of structural heterogeneity.

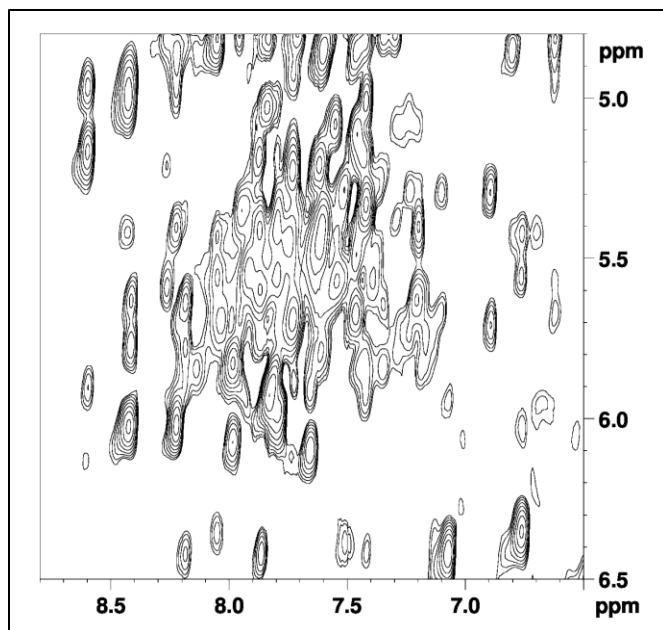


Figure 3.10: H1'-H6/H8 region of NOESY in D₂O of 1.3 mM SRB-2 bound to RB.

A significant degree of spectral overlap is observed even when a ligand with higher affinity, RB, is used in place of SR. Spectrum was acquired with 150 ms mixing time.

In addition to NOEs from the RNA, the intra- and inter- molecular NOES from the ligand are also clearly observed in Figure 3.8. In order to determine which signals in the NOESY are coming from the ligand and which cross-peaks involve interaction with the RNA, a TOCSY was acquired for sulforhodamine B (Figure 3.10). There appear to be a number of NOEs between the ligand and the RNA, and this is especially obvious with the ethyl groups from the ligand, as there are no resonances in RNA that show up at the same chemical shift (~0-3 ppm) as the ligand's ethyl groups.

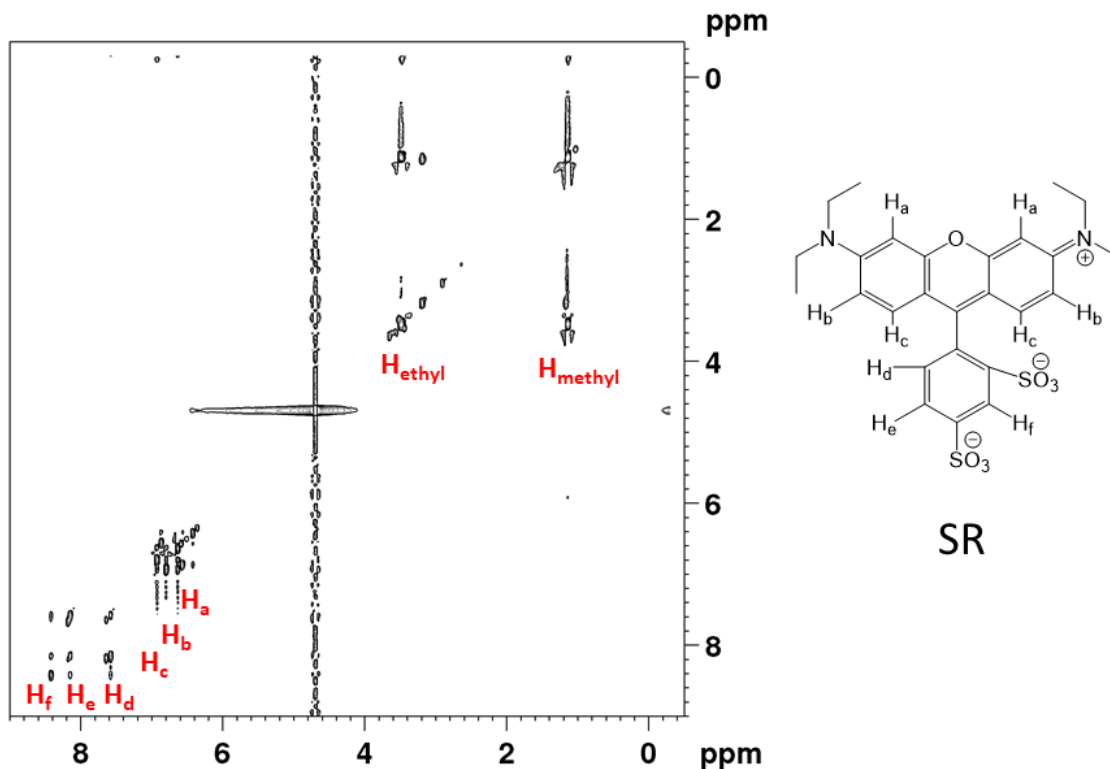


Figure 3.11: TOCSY spectrum of 1 mM sulforhodamine B.

This spectrum was acquired with a mixing time of 50 ms. Aromatic protons of SR are labelled in red and their locations on the ligand are indicated on the structure to the right.

3.4.2 Homonuclear NMR Experiments Without Ligand

In an effort to probe the native secondary structure of SRB-2, a NOESY in D₂O (Figure 3.11 A) and a TOCSY (Figure 3.11 B) of SRB-2 in the absence of ligand were acquired. Due to the adaptive binding nature of aptamers, they tend to lack a significant amount of tertiary structure in the absence of ligand. It is therefore likely that only protons in structured (double-stranded) regions of the aptamer are observable under these conditions. This could potentially be useful in assignments by helping to distinguish residues that are in the stem regions of the RNA. These experiments are also very convenient to run as they are simply acquired using a sample that will later be titrated with ligand for further study.

Unfortunately, the number of TOCSY H5-H6 peaks observed in the SRB-2 spectrum did not match the number of C and U residues in base paired regions and the variance in the intensity of signals was quite large, which significantly limits the practicality of this experiment. In the NOESY spectrum acquired in D₂O, similarly underwhelming results were obtained. An obvious reduction in the quantity of peaks was observed as expected. However, the signals remained somewhat overlapped and many are clearly TOCSY peaks. For a few peaks, it may be inferred that the residue they belong to is in a base-paired region, but ultimately there is not a lot of information provided here.

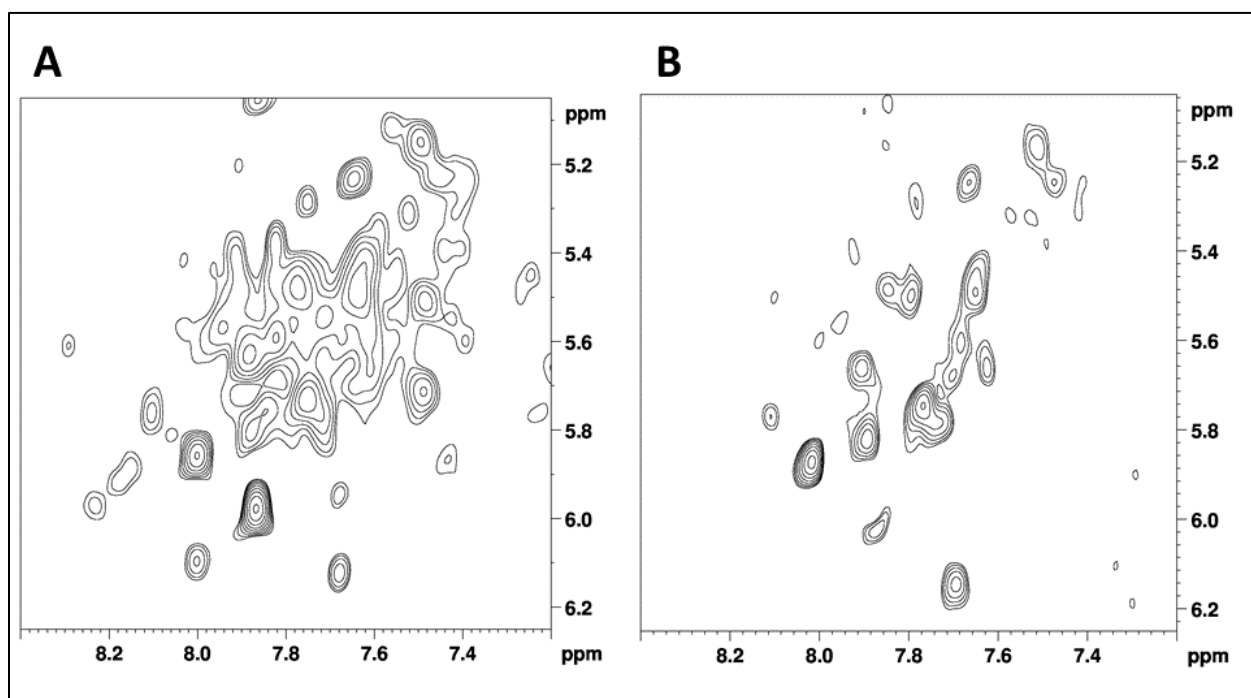


Figure 3.12: NMR spectra acquired on 0.5 mM unbound SRB-2.

A) Close-up of the H1'-H6/H8 region of the NOESY spectrum in D₂O. B) Close-up of the H5-H6 region of the 2D CITY TOCSY.

3.4.3 Heteronuclear NMR Experiments with Single-nucleotide Labelling

As discussed above, spectral overlap is great in sequences as large as SRB-2 and homonuclear experiments are not sufficient for structural determination. Therefore, we prepared isotopically labelled samples to gain additional information. Three samples were prepared, and each had a single type of nucleotide fully ^{13}C and ^{15}N labelled (i.e. all nitrogen and carbons atoms in the base and in the intra-residue sugar). These are henceforth denoted as [^{13}C ^{15}N]-Ade, [^{13}C ^{15}N]-Cyt and [^{13}C ^{15}N]-Ura to represent the samples where adenine, cytidine and uridine are labelled, respectively.

3.4.3.1 HSQC Experiments

The first experiments acquired with each sample were two ^{13}C HSQC experiments, one optimized for H1' and the other for aromatic protons. These spectra are used to identify sugar carbon spin systems and proton/carbon aromatic spin systems, respectively. These spin systems are assigned using 3D HCCH-TOCSY/COSY experiments. Proton chemical shift information, particularly with respect to the aromatic protons, can also be used to aid in NOESY assignment. Spectra for the [^{13}C ^{15}N]-Ade sample are shown in Figure 3.12. The experiment optimized for the sugar is shown in panel A. H2 and H8 aromatic proton regions from the aromatic optimized experiment are shown in panels B and C, respectively. There are 10 adenine residues in this sequence, but despite ~10 C2-H2 peaks being observed, there are 16 observed for both C1'-H1' and C8-H8. This indicates a high degree of heterogeneity in these residues.

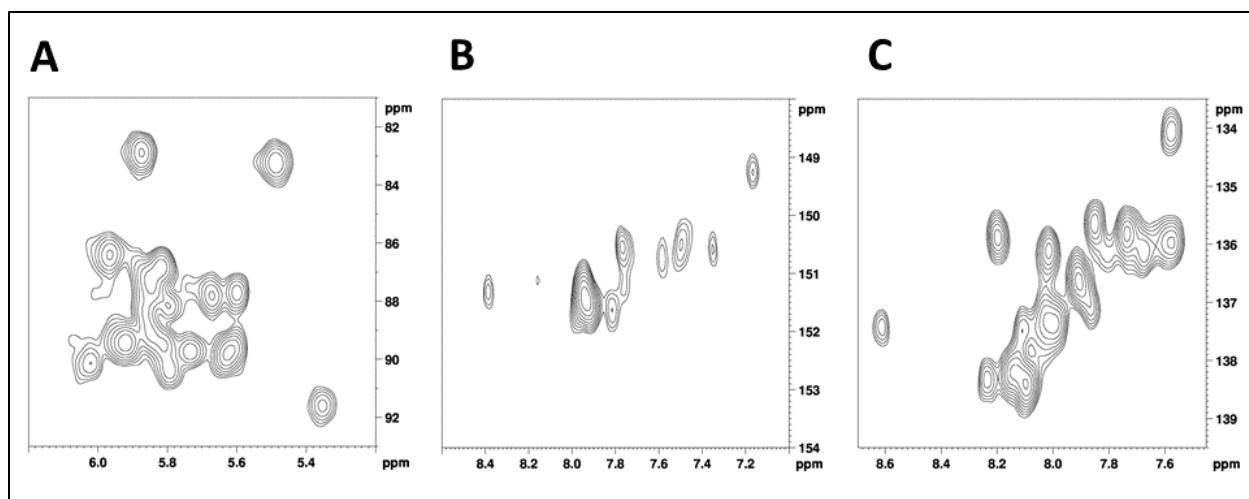


Figure 3.13: ^1H - ^{13}C HSQC spectra of 0.7 mM [^{13}C ^{15}N]-Ade SRB-2 bound to SR.

Optimized for A) C1'-H1' B) C2-H2 and C) C8-H8.

Spectra for the [^{13}C ^{15}N]-Cyt sample are shown in Figure 3.13. The experiment optimized for the sugar is shown in panel A. H5 and H6 aromatic proton regions from the aromatic optimized experiment are shown in panels B and C, respectively. There are 16 cytidine residues in SRB-2, and similar to adenine, exchange is observed. This is not evident for the C1'-H1' or C5-H5 regions, but at least 18 peaks are observed in the C6-H6 region.

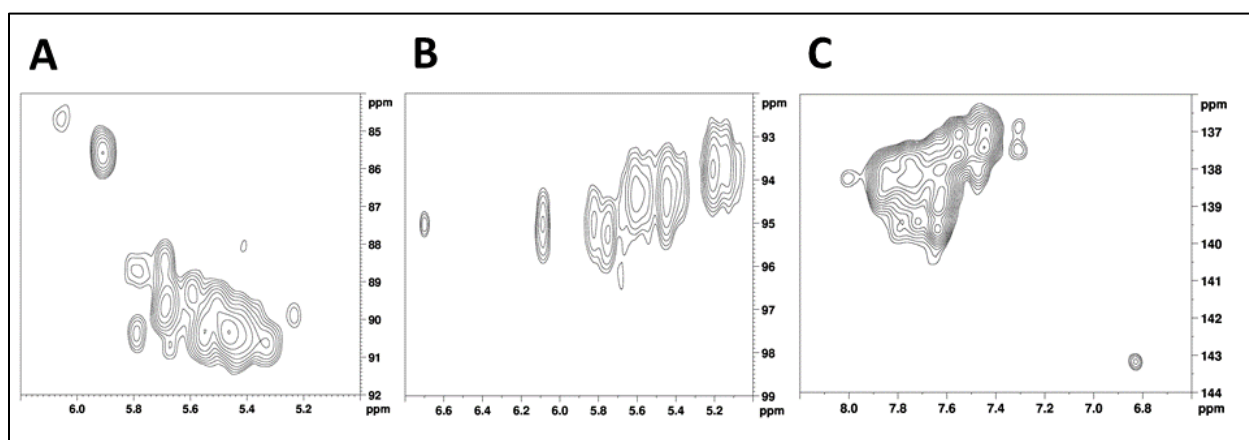


Figure 3.14: ^1H - ^{13}C HSQC spectra of 0.5 mM [^{13}C ^{15}N]-Cyt SRB-2 bound to SR.

Optimized for A) C1'-H1' B) C5-H5 and C) C6-H6.

Spectra for the [^{13}C ^{15}N]-Ura sample are shown in Figure 3.14. The experiment optimized for the sugar and in this case, the H5, are shown in panels A and B, respectively. The experiment optimized for H6 is shown in panel C. For the [^{13}C ^{15}N]-Ura sample, a ^{15}N HSQC in 90% H_2O /10% D_2O (Figure 3.13 D) was acquired in addition to the ^{13}C HSQC experiments. This HSQC not only provided the exchangeable nitrogen-proton correlations, but the proton chemical shifts also allowed for partial assignment of imino-imino NOEs in the NOESY in 90% H_2O /10% D_2O , as shown in Figure 3.5. A significant contribution from structural heterogeneity is again observed in these spectra. There are 9 uridine residues in SRB-2, but at least 12 peaks were counted in all regions of the ^1H - ^{13}C HSQCs. Six were observed in the N3-H3 range of the ^1H - ^{15}N HSQC, which suggests that three U H3 protons are not involved in hydrogen bonding.

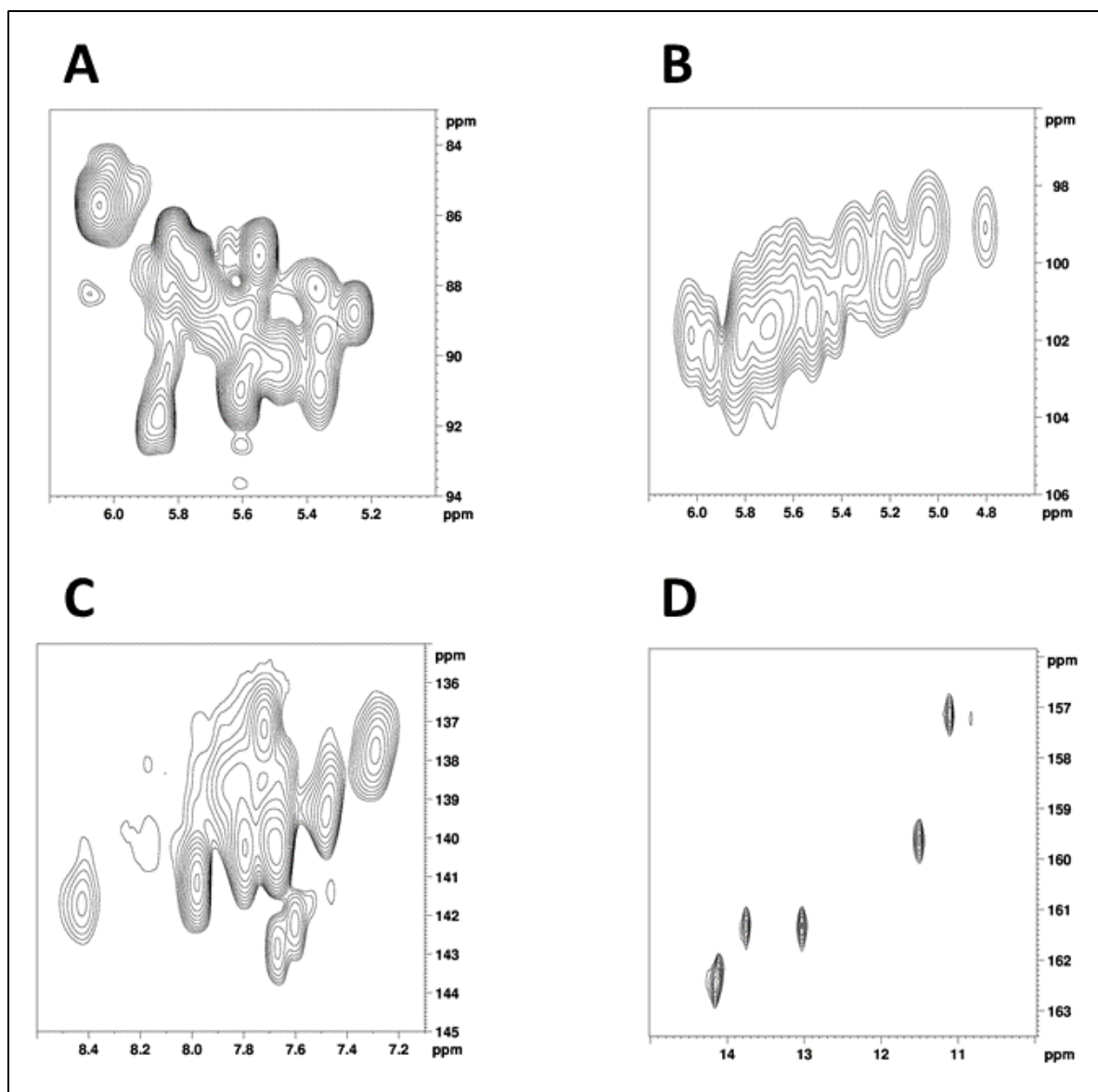


Figure 3.15: HSQC spectra of 0.9 mM [$^{13}\text{C} \ ^{15}\text{N}$]-Ura SRB-2 bound to SR.

A) ^1H - ^{13}C HSQC optimized for C1'-H1' and B) C5-H5. C) ^{13}C HSQC optimized for C6-H6. D) ^1H - ^{15}N HSQC optimized for N3-H3.

3.4.3.2 2D HCCH-COSY

Also acquired for the [$^{13}\text{C} \ ^{15}\text{N}$]-Ura sample was a 2D HCCH-COSY (Figure 3.15). This experiment was acquired primarily to distinguish between uridine and cytidine residues in the 2D CITY-TOCSY

experiment. It also contains pertinent chemical shift information about sugars in U residues, however, 3D TOCSY/COSY experiments will be required for a full assignment.

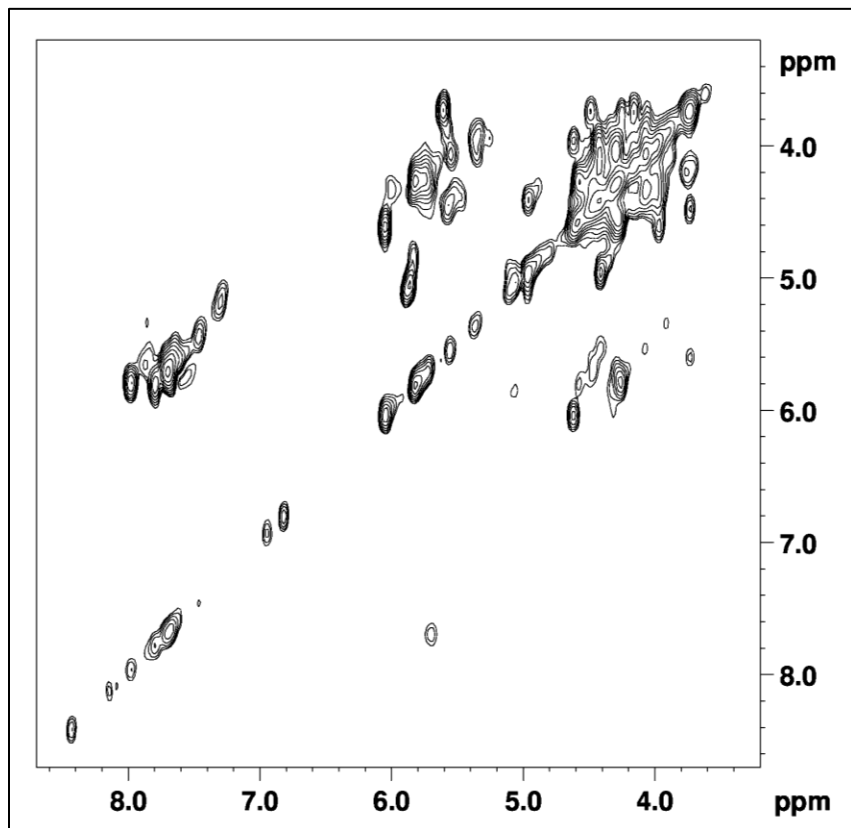


Figure 3.16: 2D HCCH-COSY spectrum of 0.9 mM [^{13}C ^{15}N]-Ura SRB-2 bound to SR.

3.4.3.3 Filtered and Edited NOESY Experiments

One of the primary goals with these labelled samples was to help reduce overlap in the NOESY in D_2O so that sequential assignments could be made. For this reason, several filtered and edited NOESY experiments were performed with each sample. These types of pulse programs allow for the selection or elimination of protons bound to isotopically labelled heteroatoms. The resulting spectra can be overlaid on the unlabelled NOESY spectrum in D_2O , which allows for the classification of residue type. Two filtered NOESY spectra were acquired with the [^{13}C ^{15}N]-Ade sample. The first spectrum, shown in

Figure 3.16, is a 2D double half x-filtered NOESY, which has two separate filters, one in each acquisition dimension. This effectively filters out magnetization that originates and/or ends up on a proton bound to ^{13}C .

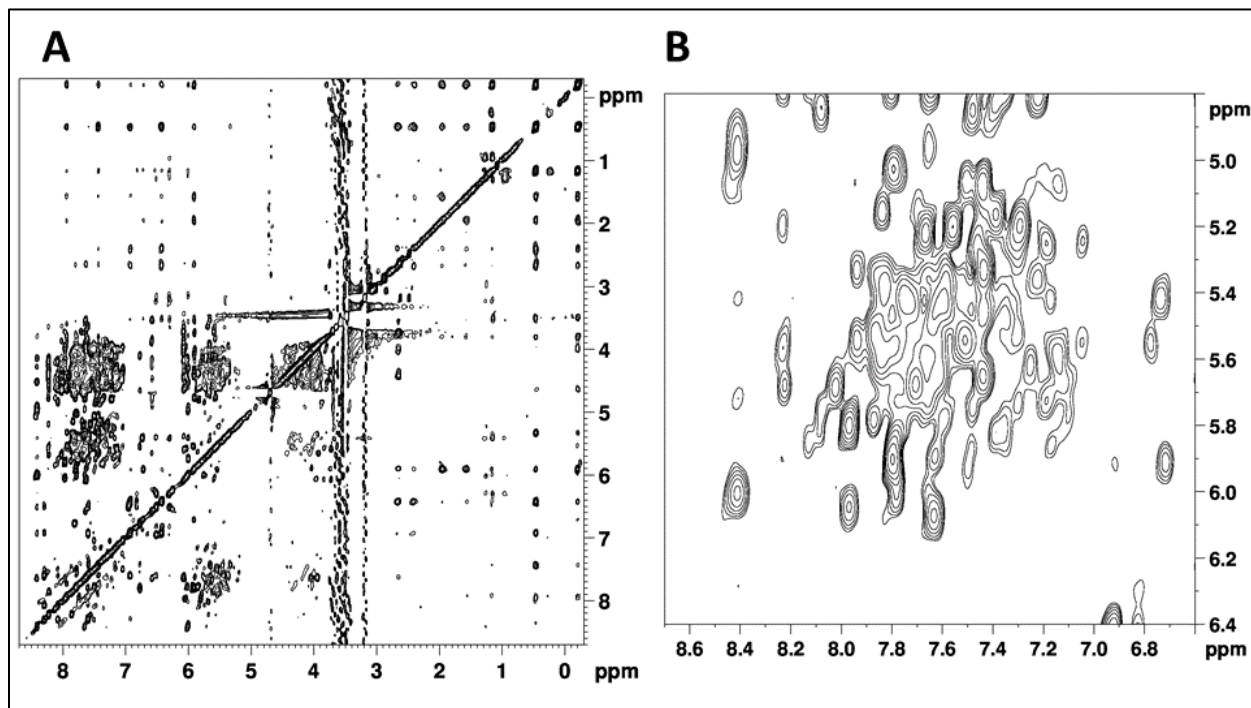


Figure 3.17: Double half-x-filtered NOESY of 0.7 mM $[^{13}\text{C } ^{15}\text{N}]$ -Ade SRB-2 bound to SR.

A) Full spectrum. B) Close-up of the H1'-H6/H8 region. Spectrum acquired with mixing time of 150 ms.

The second filtered NOESY acquired for the $[^{13}\text{C } ^{15}\text{N}]$ -Ade sample is a 2D half F2 x-filtered NOESY (Figure 3.17). This experiment only has a filter in F2, resulting in signals that originate on an adenine being filtered out of the spectrum. In combination with the double half x-filtered NOESY, this may distinguish signals originating on an A from those ending up on an A. The spectra suffer from the same heterogeneity issues as described for the unlabelled sample.

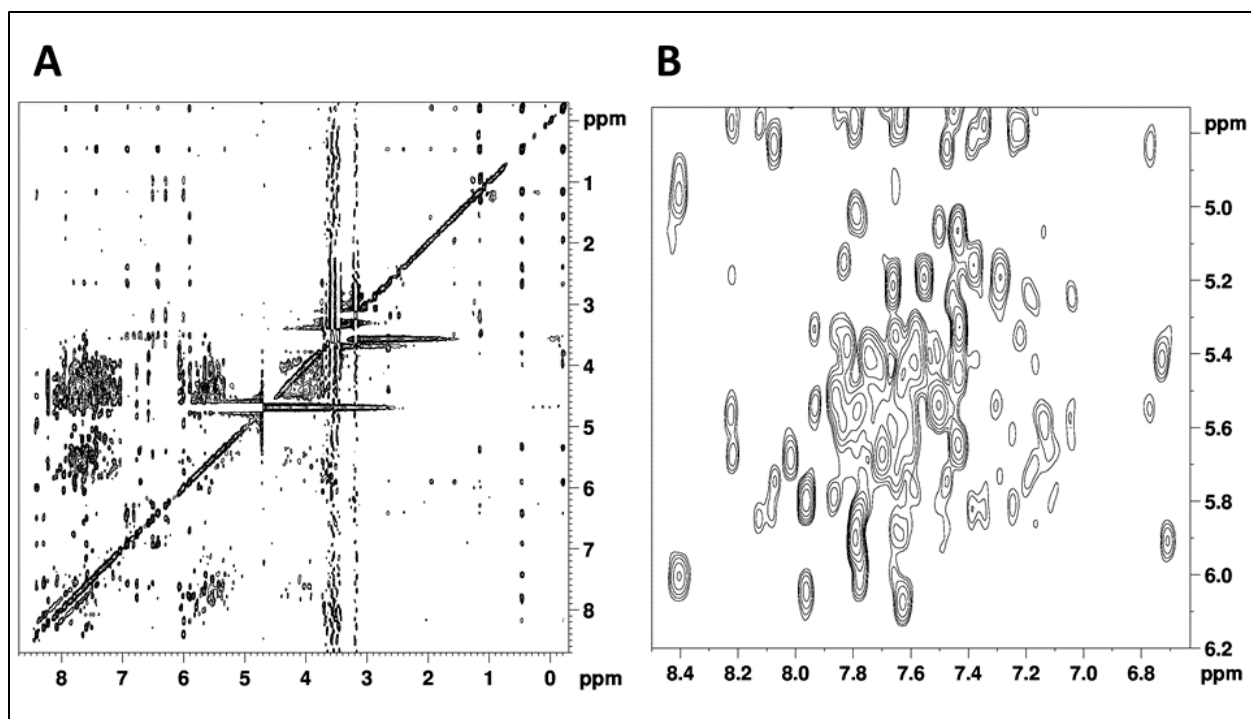


Figure 3.18: F2 half-x-filtered NOESY of 0.7 mM [^{13}C ^{15}N]-Ade SRB-2 bound to SR.

A) Full spectrum. B) Close-up of the H1'-H6/H8 region. Spectrum acquired with mixing time of 150 ms.

A 2D half x-filtered NOESY was also obtained for the [^{13}C ^{15}N]-Cyt sample (Figure 3.18).

Unfortunately, the concentration was lower, and therefore, the signal-to-noise ratio was somewhat lower than in the spectra acquired on [^{13}C ^{15}N]-Ade.

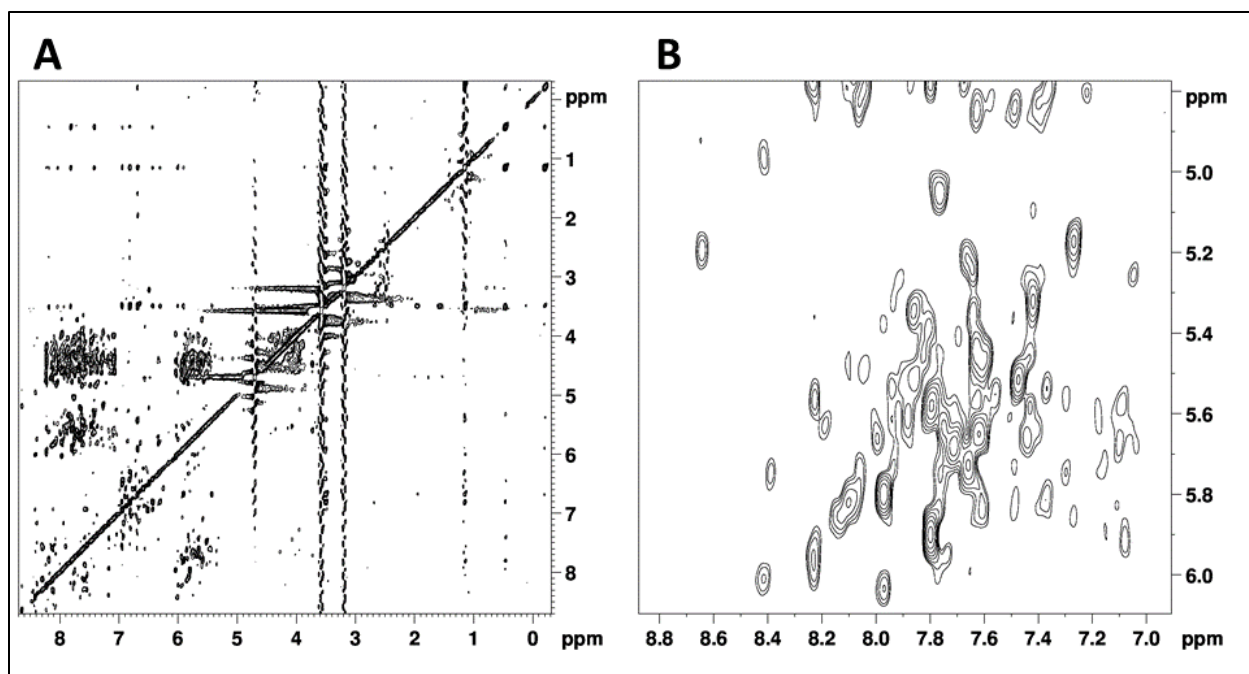


Figure 3.19: Half x-filtered NOESY of 0.5 mM [^{13}C ^{15}N]-Cyt SRB-2 bound to SR.

A) Full spectrum. B) Close-up of the H1'-H6/H8 region. Spectrum acquired with mixing time of 150 ms.

In addition to the 2D NOESYs, 3D edited ^{13}C HSQC-NOESY experiments were acquired for the [^{13}C ^{15}N]-Ade and [^{13}C ^{15}N]-Ura samples (Figure 3.18). They were both optimized for the sugar in the carbon dimension. These spectra can be used to identify which peaks in the NOESY in D_2O are from protons in adenine or uridine residues which is critical in the sequential assignment of a rather large RNA molecule like SRB-2. The F1-F3 projections for [^{13}C ^{15}N]-Ade and [^{13}C ^{15}N]-Ura are shown in Figure 3.18 A and B, respectively. The signal-to-noise ratio of the [^{13}C ^{15}N]-Ura spectrum is significantly higher despite only having a slightly higher RNA concentration and identical processing parameters to the [^{13}C ^{15}N]-Ade spectrum.

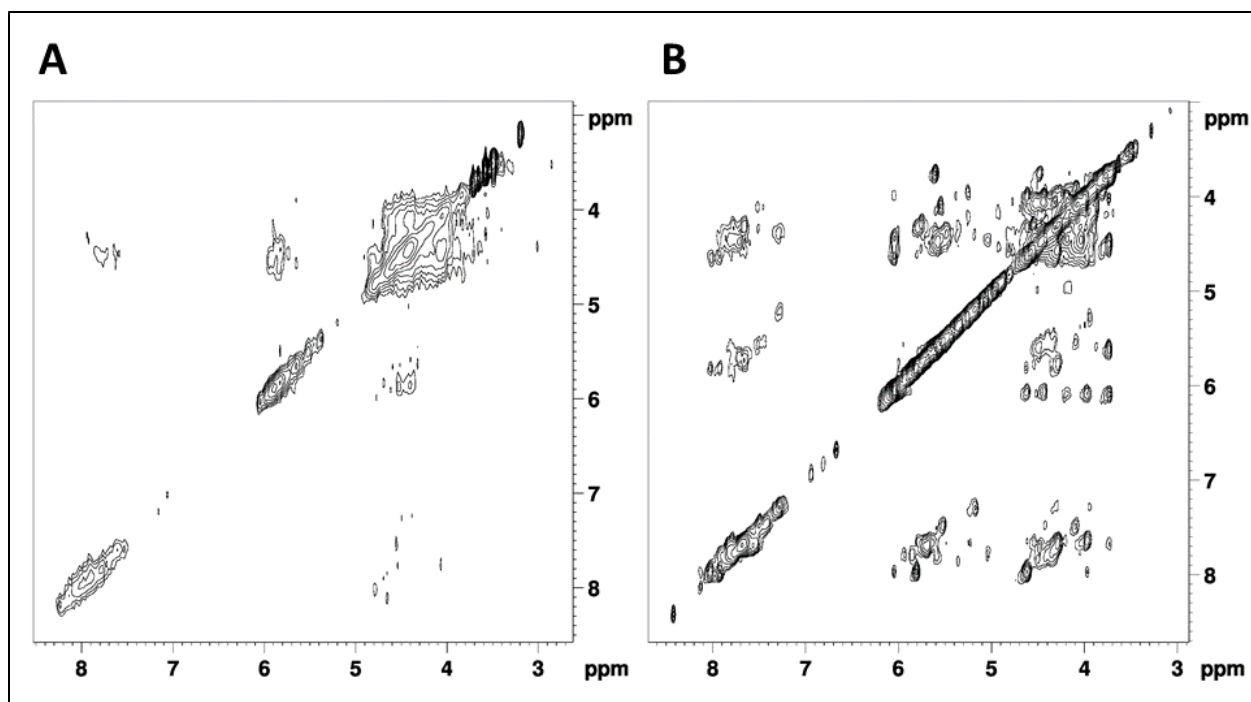


Figure 3.20: F1-F3 projections of 3D ^{13}C HSQC-NOESY experiments.

A) 0.7 mM $^{[13}\text{C } ^{15}\text{N]}$ -Ade SRB-2 bound to SR. B) 0.9mM $^{[13}\text{C } ^{15}\text{N]}$ -Ura SRB-2 bound to SR. Spectra were acquired with 150 ms mixing times.

3.4.3.4 HCN Experiments

In order to correlate imino and amino resonances to non-exchangeable base protons, we also attempted to acquire 3D HCCNH and 3D HCCNH-TOCSY experiments with the labelled-U sample. Despite meticulous optimization, no accumulation of signal was observed in the FID of these experiments, so we were unable to acquire a spectrum with either pulse program. We also ran a 3D HCN experiment to see if it would be possible to correlate the N1 to H1' and/or H6 but obtained the same unsuccessful results as the HCCNH experiments.

3.5 Discussion

As evidenced in section 3.4, a large quantity of NMR data was obtained on SRB-2. Some nice quality spectra were acquired, such as the HSQC experiments, filtered 2D NOESYs with [^{13}C ^{15}N]-Ade and the 2D HCCH-COSY and 3D HSQC NOESY with [^{13}C ^{15}N]-Ura. From these spectra, some assignments could be made, however, this was far from sufficient to fully assign and calculate a structure for SRB-2. There were a number of obstacles encountered that prevented the completion of this project, including broad lines, low-resolution spectra, structural heterogeneity, low sensitivity of pulse programs, degradation of samples, and lack of required samples. These barriers and possible resolutions will be discussed in detail below.

3.5.1 Linewidth and Spectral Resolution

As discussed in 3.2, many of the spectra acquired were of decent quality. However, they were plagued with broad lines caused by conformational heterogeneity. It is likely that heterogenous sequences were being observed in these spectra. This is evidenced by the presence of two distinct types of peaks in many of the spectrum acquired. Some peaks appeared to be strong and sharp, likely from the dominant binding conformation, while some regions contained weak, broad peaks, which appear to be the result of structural heterogeneity. We tried to separate these peaks when attempting to assign the various spectra, but the amount of heterogeneity coupled with significant overlap from the large number of signals in a 54nt sequence made assignment difficult. It is possible that conditions may exist that shift the equilibrium of conformations. For example, we tried adding MgCl_2 , but this did not solve the problem as discussed in 3.4.1.

There were also several spectra, including the 3D HSQC NOESY with [^{13}C ^{15}N]-Ade and the filtered 2D NOESY with [^{13}C ^{15}N]-Ura, which were of overall poor quality. Only one attempt at the 3D HSQC NOESY was made before the sample degraded, so we were unfortunately not able to further

optimize the experiment at that time. The same pulse program was run with [^{13}C ^{15}N]-Ura and the spectral resolution was much higher, so I believe that this was at least partially a result of parameter adjustment. It should also be noted that the concentrations of the three labelled samples were all different. [^{13}C ^{15}N]-Ade was 0.7 mM, [^{13}C ^{15}N]-Cyt was 0.5 mM and [^{13}C ^{15}N]-Ura was 0.9 mM. The fact that the concentration of [^{13}C ^{15}N]-Ura was higher may have also contributed to the difference in spectral resolution. This was also likely a factor with the filtered 2D NOESY with [^{13}C ^{15}N]-Cyt. This sample was the lowest concentration and was also acquired in a Shigemi tube with a sample volume of only 200 μL . The concentrations of all these samples are significantly lower than the 1.8 mM unlabelled sample. This is an inherent disadvantage of a filtered NOESY compared to an edited one, as peaks disappearing may be a result of lower sensitivity due to concentration rather than being filtered out due to the NMR-active ^{13}C . For this reason, it is difficult to confidently extract information from the filtered 2D NOESY with [^{13}C ^{15}N]-Cyt. To resolve this problem, more RNA would likely need to be produced. This is quite cost-prohibitive, especially when labelling is involved. Another possibility, which will be discussed more in Chapter 4, is the optimization of transcription initiation by changing the first 5-8 nucleotides of the sequence. At the very least, acquisition of edited NOESYs should be prioritized over filtered NOESYs.

3.5.2 Structural Heterogeneity

There are quite a few possibilities for the origin of the structural heterogeneity, several having to do with *in vitro* transcription by T7 RNA polymerase. Despite its usefulness in molecular biology, T7 RNA polymerase is known to produce a number of undesirable side products that result in careful analysis and purification being required. Side reactions can result in aborted products [240,241], use of alternative template initiation sites [242,243], polymerase slippage [244], non-templated nucleotide additions at the 3' end of RNA and [187,245] and 5' sequence heterogeneity [240,242,246].

Aborted sequences and sequences varying significantly in length resulting from alternative initiation sites are easily removed from a sample by gel purification. However, products with 3' and 5' heterogeneity are the same length or within a couple of nucleotides. Even with high resolution gels, it is not always possible to separate these kinds of products. There are several possible methods that could be used to reduce or eliminate these side products. For example, it has been shown that using a DNA template with C2'-methoxyl modifications at the 5' end can be effective in reducing these additional nucleotides [247]. For both 3' and 5' heterogeneity, it has been shown that using self-cleaving or cleavable tags can be highly effective in reducing side products [246]. It may also be prudent to reduce 3' phosphate cyclization and ensure homogeneity by incubation with T4 polynucleotide kinase–phosphatase [248]. Of particular concern to us, it has been shown that sequences beginning with GGG are particularly susceptible to 5' heterogeneity [240,242,246]. However, we had issues with low transcription yields with other 5' sequences. Therefore, it would certainly be of benefit to spend more time optimizing the 5' sequence for a combination of reduced heterogeneity and sufficient product yield.

Another source of heterogeneity sometimes encountered in nucleic acid samples, especially at concentrations required for NMR spectroscopy, is dimerization. Dimerization often results when self-complementary regions exist in a nucleic acid sequence. Based on the SRB-2 sequence, it is possible that one of the regions in what we will later refer to as loop 3 could do this. However, we did not observe any conclusive NMR evidence, such as additional NOEs consistent with this type of base pairing, that would support this notion. In addition, a native gel run by a former student on various SRB-2-ligand complexes also showed no indication of dimerization or oligomerization.

Finally, one source of heterogeneity possible with any receptor-ligand complex is if saturation has not been reached. We did not include $MgCl_2$ in the NMR despite ITC studies indicating that it may be required. This is due to initial NMR titrations we did in the absence of Mg^{2+} , wherein changes in the

spectrum were not observed after approximately a 1:1 titration with ligand. We titrated multiple equivalents of the ligand initially, but this was not enough to induce any observable differences in the spectrum. As a result, we generally aimed for ~1.2:1 ligand:RNA in our NMR samples. It is possible that the absence of peak observed in ITC has to do with thermodynamics of binding in the presence of $MgCl_2$ rather than an actual absence of binding. A control SPR experiment without any magnesium may be useful to confirm this theory. Another possibility here is that other residual metal ions present may have taken on the role of Mg^{2+} . However, despite not de-metalating our solutions, there was a significant amount of EDTA present both in the gel and anion exchange buffers, giving little reason to believe that this was the case. For these reasons, we believe that the 5' and 3' heterogeneity discussed were more likely to be the source of heterogeneity.

3.5.3 HCN Experiments

One of the biggest disappointments of this project was our inability to acquire any triple resonance experiments with our labelled samples. These experiments are desirable as they can reduce ambiguity in NOESY experiments, but they are not without their challenges. Spectral resolution and sensitivity of RNA are known to decrease significantly with increasing molecular weight [227,249], and these types of pulse programs involve several transfers of magnetization, resulting in significant signal losses. Because SRB-2 is relatively large (54 nt), and these pulse programs used several resonance transfers, we postulate that the sensitivity of these programs was too low to detect signal in these multiple-transfer experiments. If this project were to continue, I would suggest synthesizing a very short [^{13}C ^{15}N] oligonucleotide, which could be used as positive control to confirm this hypothesis and rule out any issues with the pulse programs themselves or relevant parameter settings. It is also worth noting that the console of the Bruker DRX-600 spectrometer used was replaced and the software used to run it was updated during the period in which these experiments were run. This may have resulted in spectra

of varying quality before and after the change. It may have also resulted in compatibility issues with some pulse programs.

3.5.4 Sample Stability

In addition to the low resolution of some spectra, there were additional spectra that we intended to acquire but were unable to due to degradation of samples. Hence the missing NOESY experiments for some samples in section 3.4. RNA is particularly susceptible to degradation due to the ubiquitous presence of RNases on surfaces and in reagents that are not certified RNase-free. RNase activity has even been observed in RNA samples frozen at -20 °C [250]. Even if much care is taken to avoid RNase contamination, RNA is often only stable for up to a year when frozen at -20 °C or -80 °C and mere days at room temperature [251–254]. Despite -80 °C storage, certified nuclease-free labware and reagents, and re-suspension with ultra-pure Milli-Q water, all three of the labelled samples degraded within weeks or months of synthesis. Unfortunately, two of these samples ($[^{13}\text{C } ^{15}\text{N}]$ -Cyt and $[^{13}\text{C } ^{15}\text{N}]$ -Ura) degraded during the months of storage resulting from the provincial lockdown in the spring of 2020.

Although RNA degradation is sometimes out of a researcher's control, there are a few things I would do differently in the future. It has been suggested in the literature, that although freezing RNA in an aqueous solution at -80 °C is commonplace, storage as either a lyophilized powder or ethanol precipitate may be more effective long-term storage techniques [251,254]. Storage in an air-tight container should be ensured as atmospheric humidity is known to play a role in RNA degradation [251]. I also took time to analyze each spectrum before running more, and I believe that more meticulous planning and efficient scheduling of experiments could help guarantee the acquisition of all desired spectra by decreasing the amount of time that samples need to stay viable.

3.5.5 Additional Resources for Assignment

The reality is that even if a full complement of high-resolution spectra were acquired for the three samples studied here, it is nearly certain that more data would still be required. This would likely involve the uniformly labelled samples [U-¹⁵N] and [U-¹³C ¹⁵N]. It may be useful to have another single (or dual) nucleotide labelled sample that has guanine, since it has not been labelled in any of our previous samples. It is also possible that selectively deuterated samples would be required to assign the severely overlapped H1'-H6/H8 region of the NOESY in D₂O. Unfortunately, the isotopically labelled nucleotide triphosphates that are required to synthesize these samples are extremely cost prohibitive. There are alternative paths that may somewhat decrease the price tag of these samples, such as preparing the nucleotides ourselves. However, this requires significant time and expertise, and the starting materials and enzyme cocktail required are still very expensive. In addition to the quality of spectra obtained, cost was a factor in the stalling of this project after the three existing samples had degraded.

3.6 Conclusions and Future Work

In summary, the solution structure of SRB-2 was probed using homonuclear NMR techniques as well as with the three single-nucleotide labelled samples [¹³C ¹⁵N]-Ade, [¹³C ¹⁵N]-Cyt and [¹³C ¹⁵N]-Ura. The major obstacle encountered was the detrimental effects of multiple conformations on the spectra. This resulted in an insufficient amount of quality data for the unambiguous assignment of SRB-2. There are a few things one could do to potentially improve quality of these spectra. There were some minor changes observed with low Mg²⁺ concentrations, so investigations with more salt concentrations may be useful. Spectra obtained using other ligands were not promising, however, we could also consider looking at SRB-2 mutants that may have a stronger equilibrium of preferred to alternative conformations. This includes sequences with optimized 5' and 3' sequences, and perhaps other

modifications such as C2'-methoxylation of the DNA template, or cleavable RNA tags discussed in 3.5.2. In terms of NMR spectroscopy, there are also methodologies available that look at molecular dynamics, which may provide some insight into what is happening in this system [255].

If we found a solution for these structural heterogeneity issues, we would likely start by synthesizing [U- ^{15}N], [U- ^{13}C ^{15}N] and either a single or dual nucleotide labelled sample involving G. With these samples, we would consider replicating some of the strategies used by Duchardt-Ferner *et al.* who recently characterized the solution structure of the TMR3 aptamer bound to tetramethylrhodamine [129]. This aptamer is a similar size (48 nt) as SRB-2 (54 nt) and binds tetramethylrhodamine (TMR), an analogue of SR. The binding affinity and kinetics of TMR with respect to SRB-2 were discussed back in Chapter 2. With the [U- ^{15}N] sample, we would acquire standard and long-range 2D HNN-COSY experiments and a ^{15}N HSQC-NOESY. With respect to NOESYs, Duchardt-Ferner *et al.* similarly acquired 3D ^{13}C HSQC NOESYs with their selectively labelled samples [^{13}C ^{15}N]-Ade, [^{13}C ^{15}N]-Ade, Cyt and [^{13}C ^{15}N]- Gua, Ura. These largely provide the same information as our single-nucleotide labelled samples. The major difference is we never synthesized a sample with labelled G residues, which would be useful to have. Another sample that would be quite useful to have is [U- ^{13}C ^{15}N]. With this sample, a double x-half filtered 2D NOESY could be acquired to isolate intra-ligand NOEs. An F1-filtered 3D ^{13}C HSQC could also be acquired to specifically identify RNA-ligand NOEs. Similarly, a sample bound to ^{13}C -labelled ligand could also prove valuable.

As discussed in 3.5.3, HCN experiments did not work out, and there is little motivation to pursue them further. If we needed to further reduce ambiguity in the NOESYs, I would suggest using at least one partially deuterated sample, likely with one or more nucleotides deuterated in the H3', H4', H5' and H5'' positions. Removing some of these spins would drastically reduce the amount of dipolar relaxation experienced by the other protons, narrowing the linewidths of their cross-peaks and therefore reducing overlap in the spectrum [227].

As will be described in Chapter 4, we opted to pursue a 'divide and conquer' approach as an alternative to isotopic labelling. This involves designing and synthesizing various truncations of the SRB-2 aptamer which are more financially sustainable to synthesize than labelled samples.

Chapter 4: Probing the Structure of the SRB-2 Binding Aptamer

All experiments in this chapter were performed and analyzed by the candidate.

4.1 Chapter Abstract

The SRB-2 aptamer is a polyanionic RNA that binds a ligand with an overall negative charge, making this an interesting system to characterize. There are various NMR strategies that can be used to acquire structural information, many of which involve isotopic labelling, as described in Chapter 3. Due to the lack of success for this sequence, we opted to pursue alternative approaches which do not, at least initially, require isotopic labelling. Segmental analysis, or the “divide and conquer” approach, involves resolving the structure of smaller pieces of a sequence then combining the information to produce a single, overall structure. This is a particularly useful method if there are multiple stem-loop segments in a sequence that may form stable structures independent of the other stem-loops. In this chapter, strategies for designing representative truncations are outlined. For a 16mer containing a UUCG tetraloop, assignments are completed, and a structure is calculated. This allowed for several resonances in the spectra of full-length SRB-2 to be identified. A similar study was attempted with another stem-loop structure (loop 3) from SRB-2. However, results obtained suggested that the structure of loop 3 alone is heterogenous, and in SRB-2 the structure of loop 3 is likely co-dependent with loop 2. These results indicate that segmental analysis beyond the UUCG tetraloop is not possible for this system.

4.2 Introduction

As established in previous chapters, SRB-2 is a desirable system to study due to its intrinsic properties including ligand promiscuity and overall negative charge. The kinetics and binding affinity of

SRB-2 ligands are outlined in Chapter 2. Structural NMR studies of SRB-2 bound to SR are presented in Chapter 3. These NMR studies involved unlabelled samples as well as the single nucleotide labelled samples [^{13}C ^{15}N]-Ade, [^{13}C ^{15}N]-Cyt and [^{13}C ^{15}N]-Ura. The spectra acquired with these samples were insufficient for unambiguous assignment due to structural heterogeneity, and further labelling was cost prohibitive, so we sought other avenues to pursue the goal of elucidating a solution structure of SRB-2.

One such strategy that has been used to characterize numerous large RNAs by NMR is the so-called divide and conquer approach [256–259]. This approach involves characterizing smaller pieces of the sequence, often individual stem-loop structures from a multi stem-loop complex. These shorter sequences are more amenable to assignment and are relatively inexpensive to synthesize. These smaller sequences potentially provide partial assignments for the overall sequence, and when all the data is pieced together, the result is significant deconvolution in the NMR spectra for the full-length sequence. This also provided an opportunity to become familiar with how structure calculations are done, as it became clear in Chapter 3 that a full structure determination was unlikely.

Of course, there are several assumptions that must hold in order for this approach to work. First, the individual segments must be folded. Otherwise, unambiguous assignment will not be possible. Secondly, because we are talking about an aptamer, the segments must also retain the same structure in the presence and absence of ligand. It is highly unlikely that such severe truncations would retain any binding activity, so these experiments would be performed in the absence of ligand and if their structure is significantly different than when bound, assignments of the segment will not translate to the full-length sequence. Finally, the structure of each stem-loop segment must not be dependent on the presence of other such segments from the full sequence. For example, if there is base-pairing between two distinct loop regions, then segmental analysis of the independent stem-loops will not be an effective method of assignment.

With regards to segmental analysis of the SRB-2 aptamer, the most logical place to start is the UUCG tetraloop. Tetraloops are highly stable secondary structure elements that often cap stems in RNA. They were first discovered by Woese *et al.* in 1983 when they noted a disproportionate number of hairpin loops consisting of four nucleotides in 16S rRNA [260]. The vast majority of these sequences fell into two main families, GNRA and UNCG. Further studies indicated that the UUCG tetraloop is the most stable tetraloop as a result of its high thermodynamic and structural favorability [261]. It was later noted by Tuerk *et al.* that a probable function of these biologically abundant hairpins is to organize the correct folding of large, complex RNA sequences [262]. Some tetraloops, including GAAA [263] and AUUA [264] are known to interact with receptors or proteins. However, the UUCG tetraloop lacks any known propensity for protein-RNA or RNA-RNA interactions [265]. For these reasons, UUCG tetraloops are often incorporated in aptamer sequences where possible to promote proper folding while avoiding unwanted interaction with the ligand or other parts of the RNA sequence. High stability and low interactivity are also properties that make the UUCG tetraloop of SRB-2 an ideal candidate for segmental NMR analysis.

Due to its unusually high stability and experimental practicality, the structure of the UUCG tetraloop is well characterized. Several structures have been determined for UUCG tetraloops with varying short stems [266–268] and in larger RNA molecules such as a self-splicing intron [269], an influenza virus promoter [270], and the ATP-binding aptamer [77]. A schematic of the UUCG tetraloop is shown in Figure 4.1. The first solution structure was a 12-mer elucidated in 1991 by Varani *et al.* and this structure revealed many of the unusual contacts that result in the high stability of the UUCG tetraloop [266]. These include a U1-G4 reverse wobble base pair where G4 is in a *syn* conformation, a C3 base stack on U1 and a hydrogen bond between the N4 of C3 and a phosphate oxygen of U2 [266]. A few years later, the first crystal structure of the UUCG tetraloop was determined as part of a 57nt segment of the 16s rRNA. This structure revealed additional hydrogen bonds between O2' of U2 and O6 of G4,

and between O2' of C3 and O2 of C3, which were not observed in the solution structure. In the years since several NMR structures of higher quality have been reported for a UUCG 14mer as a result of improvements in both NMR and computational methodologies [268,271–273]. In recent studies on the same 14mer, the existence of a low-populated, non-native conformation where U1 and G4 do not pair was shown using exact NOE (eNOE) measurements and molecular dynamics simulations [273].

Because of the high stability and well-defined structure of the UUCG tetraloop, this segment should be an ideal candidate for our segmental analysis. It provided a suitable model for learning how to do structure calculations and proving the efficacy of our assignment and calculation processes. We also designed several versions of what we termed loop 3, the other stem-loop structure in SRB-2. The goal for designing these segments was to optimize transcriptional yield while maintaining the structural integrity of the aptamer. For this reason, the binding affinity of some modified versions of SRB-2 were tested to ensure that the sequence changes implemented would be representative of the original SRB-2 sequence. In these studies, a complete NMR structure of a UUCG-containing 16mer was elucidated. Initial studies with loop 3 truncations indicate that elucidation of a structure is unlikely.

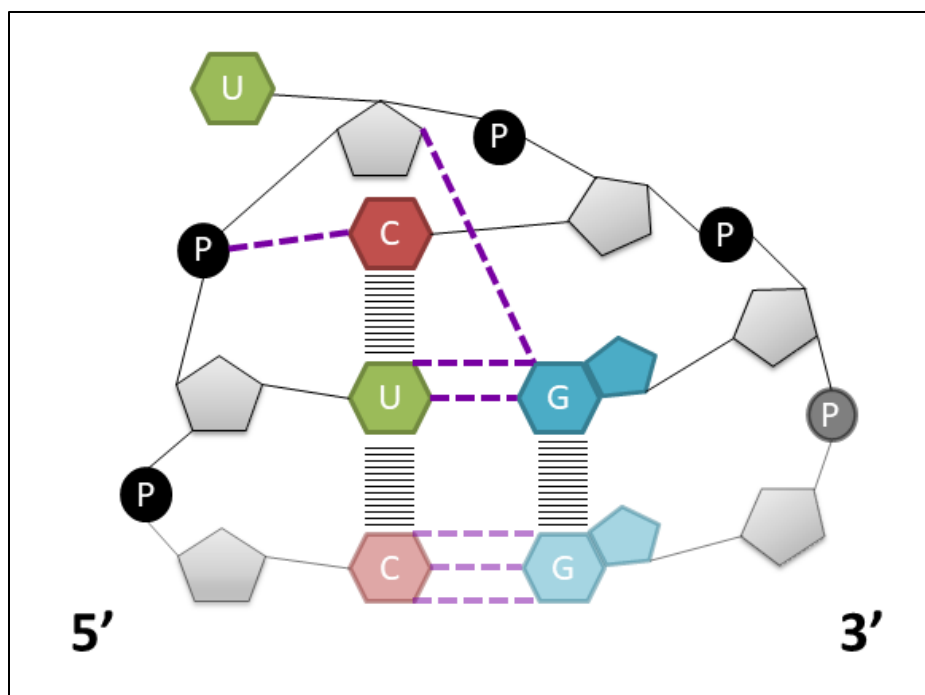


Figure 4.1: Schematic of the cUUCGg tetraloop.

Solid black lines indicate bonds, horizontal black lines indicate stacking interactions and purple dashed lines indicate hydrogen bonds. This tetraloop was inserted into the SRB-2 sequence (Figure 2.1) by the original authors to promote proper folding [37].

4.3 Materials and Methods

4.3.1 Preparation of RNA Samples and Dyes

RNA was synthesized enzymatically using a T7 RNA polymerase and either a double-stranded synthetic DNA template or a single-stranded DNA template with T7 promoter. DNA was purchased from Integrated DNA Technologies, Inc. (Coralville, Iowa) and Eurofins (Huntsville, Alabama). The sequences used are summarized in Table 4.1.

Table 4.1: List of relevant DNA template sequences

Aptamer	Aptamer Template Sequence
SRB-2	5'-GGG ACC TGA GGC GGT TAA CCT TGC GCC TCT CCA TCA TCG CCG AAG CGA GGT CCC TAT AGT GAG TCG TAT TA-3'
SRB-2 reverse	5'-TAA TAC GAC TCA CTA TAG GGA CCT CGC TTC GGC GAT GGA GAG GCG CAA GGT TAA CCG CCT CAG GTC CC-3'
SRB-2 min	5'-GGG ACT GAG GCG GTT AAC CTT GCG CCT CTC CAT CAT CGC ACC CTA TAG TGA GTC GTA TTA-3'
L3	5'-GGG GCG GTT AAC CTT GCG CCC CTA TAG TGA GTC GTA TTA-3'
L3 reverse	5'-TAA TAC GAC TCA CTA TAG GGG CGC AAG GTT AAC CGC CCC-3'
L3A	5'-GGG ACG GTT AAC CTT GCG TCC CTA TAG TGA GTC GTA TTA-3'
L3B	5'-GGG CCG GTT AAC CTT GCG GCC CTA TAG TGA GTC GTA TTA-3'
L3C	5'-GGG AAA GGT TAA CCT TGC TTT CCC TAT AGT GAG TCG TAT TA-3'
SRB-2C	5'-GGG ACC TGG GAA AGG TTA ACC TTG CTT TCC CTC CAT CAT CGC CGA AGC GAG GTC CCT ATA GTG AGT CGT ATT A-3'
SRB-2D	5'-GGG ACC TGG GAA GGT TAA CCT TGC TTC CCT CCA TCA TCG CCG AAG CGA GGT CCC TAT AGT GAG TCG TAT TA-3'
UUCG	5'-GGG CGC CGA AGC GCC CTA TAG TGA GTC GTA TTA-3'
T7 promoter	5'-TAA TAC GAC TCA CTA TA-3'

RNA was transcribed using the recipe listed in Table 2.1. The RNA was purified on a 10% Urea PAGE gel and the band containing the aptamer was cut out. The RNA was eluted from the gel by crush & soak in 300mM NaCl or by electroelution. This was followed by clean-up on a HiPrep 16/10 DEAE FF

anion-exchange column (GE Healthcare, Uppsala, Sweden) and desalting on a HiPrep 26/10 Desalting column (GE Healthcare, Uppsala, Sweden). The RNA was also precipitated with 70% ethanol prior to running each column. Pure samples obtained from the desalting column were then lyophilized and dissolved in 500 μ L of 10 mM potassium phosphate buffer and 10 mM KCl (90% H₂O/10% D₂O).

4.3.2 Fluorescence Emission Scans

Spectra were obtained using a Spectramax M5 Multi-Mode Microplate Reader using a 100 nm wide scan with 1nm intervals. Samples were measured at an λ_{ex} of 535 nm and an λ_{em} range of 560-660 nm. Readings were taken in Greiner 96-well black microplates (Kremsmünster, Austria). Samples were prepared with 10 μ M dye and 40 μ M RNA in assay buffer (10mM HEPES, 10mM KCl and 5mM MgCl₂ at pH 7.4). Samples were incubated at 25°C for at least 10 minutes before measurement. All data was normalized as a percentage of the fluorescence recorded at λ_{max} (563nm) of SRB-2 bound to SR.

4.3.3 Fluorescence Titrations

Titration were performed on a Spectramax M5 Multi-Mode Microplate Reader with a Greiner 384-well black microplate (Kremsmünster, Austria). Measurements were taken using a λ_{ex} of 563 nm and a λ_{em} of 585 nm, which were determined experimentally using excitation and emission scans. Samples were prepared with varying concentrations of RNA (0.4-20 μ M) and 1 μ M SR in assay buffer (10mM HEPES, 10mM KCl and 5mM MgCl₂ at pH 7.4). Samples were incubated at 25°C for at least 10 minutes before measurement. The resulting binding curves were fit using a non-linear, least squares method to the simple hyperbolic function: % of $F_{\text{max}} = F_{\text{max}} \times [\text{SR}] / (K_{\text{d}} + [\text{SR}])$ where F_{max} is the maximum fluorescence and K_{d} is the dissociation constant. Raw fluorescence values were normalized to a percentage of their sequence's respective F_{max} .

4.3.4 NMR Experiments

Once experiments in 90% H₂O/10% D₂O were acquired, the samples were dried by lyophilization and re-dissolved in 500 μ L of 99.996% D₂O (Cambridge Isotopes) to perform experiments on non-exchangeable resonances. Experiments in 90% H₂O/10% D₂O that were used to observe exchangeable protons were run using $1\bar{1}$ -spin echo solvent suppression [190]. Two-dimensional NOESY spectra [191] in 90% H₂O/10% D₂O were acquired at 277 K with a mixing time of 150 ms for both the UUCG and L3C samples. Experiments in D₂O that were used to observe non-exchangeable protons were run with presaturation solvent suppression [192]. For the UUCG sample, DQF-COSY [209], a 2D CITY-TOCSY [193] with a mixing time of 50 ms, and NOESYs with mixing times of 80, 150 and 200 ms in 100% D₂O were acquired at 298 K. For the L3C sample, a 2D CITY-TOCSY and a NOESY with a 240ms mixing time were acquired. All NMR samples had a volume of 500 μ L and were read in standard 5 mm NMR tubes. All spectra were collected on a Bruker DRX-600 spectrometer equipped with an HCN triple-resonance, triple-axis PFG probe. Quadrature detection for the indirect dimensions in multidimensional experiments was achieved using the States-TPPI method [189]. Relevant pulse programs can be found in Appendix B.

4.3.5 Input Restraints and Structure Calculations

Interproton distance restraints for hydrogens were obtained from NOESY spectra of the sample in D₂O acquired at 298 K with 80, 150 and 240 ms mixing times. Exchangeable NOE distances were obtained from a 2D NOESY in 90% H₂O/10% D₂O at 277 K with a 150 ms mixing time. Based on resonance assignments from these spectra, NOE peak intensities were examined and initially grouped into three distinct categories: strong (~ 2.5 Å), medium (~ 3.5 Å) and weak (~ 5 Å). These were based on an internal standard of 2.5 Å for H5-H6 NOESY cross-peaks. Upper bounds were set at a very conservative 2.0 Å above the proposed distance and lower bounds were set at the Van der Waals distance. The

structure of the UUCG tetraloop is known from previous studies [266,267], and the remainder of the 16mer UUCG sequence is a stem similar to that of the 12mer used in those studies. After a few rounds of calculation, it was apparent that the 16mer consisted of a standard UUCG tetraloop and a stem of RNA in A-form as expected. Once this was established, further refinement of the distance restraints was performed based on the literature interproton distances of the UUCG tetraloop and of A-form RNA, particularly for intra-residue distances. Upper and lower restraints were set at 0.5 Å for strong NOES, 1.0 Å for medium intensity NOEs and 1.5 Å for weaker NOEs. Planarity restraints were included for each of the six Watson-Crick base pairs in the sequence to prevent tilting of the bases. Standard hydrogen bond restraints were also included for these base pairs. A few hydrogen bond restraints were also included for well-established contacts in the UUCG tetraloop, including U7 O2 to G10 N2 and H22, U7 O1P to C9 H41 and U8 O2' to G10 O6. The dihedral angles of all residues in Watson-Crick base pairs as well as U7 were restrained to a standard A-form helix and an N-type sugar pucker. The ribose angles ν_1 and ν_2 in U8 and C9 were restrained to an S-type sugar pucker based on coupling constants observed in the DQF COSY, while all other sugar and backbone angles in these residues were left unconstrained. G10 is known to adopt a *syn* conformation [266,267] and is evidenced by the very strong H8-H1' correlation observed in the NOESY in D₂O. For simplicity, all dihedral angles for this residue were left unconstrained. Lists of all chemical shifts, distance restraints and dihedral angles used in structure calculations can be found in Appendix C.

All structure calculations were done with CNSsolve version 1.3 [274,275]. A total of 20 structures were calculated with 8200 simulated annealing steps per conformer. This consisted of 4000 high temperature steps of torsion angle molecular dynamics followed by two slow-cool annealing stages with 1000 and 3000 steps, respectively. Finally, a minimization stage consisting of 200 steps was performed. Structures were visualized and analyzed using the software packages RasMol [276] and UCSF Chimera [130].

4.4 Results and Discussion

4.4.1 Sequence Design and Fluorescence Screening

The two main domains of SRB-2 we investigated for segmental analysis were the UUCG tetraloop (loop 1) and loop 3. One of the first considerations that must be made in the design of these sequences is the effect of the sequence on T7 RNA polymerase activity. Optimizing activity is particularly important for NMR studies because relatively large (mg) quantities of RNA are required for an NMR sample of adequate concentration. It has been established in previous literature that having G residues in the +1 to +3 positions of the sequence is often optimal for *in vitro* transcription [277]. The +1 G base has been shown to have several specific interactions with residues in the active site of T7 RNA polymerase, which explain its strong preference for this type of nucleotide [278,279]. During transcription initiation, the T7 polymerase binds the DNA promotor from nucleotide position -17 to -5. The DNA template is then melted from position -4 to +3 to prime RNA synthesis, followed by expansion of this initiation bubble to at least the +7 position. It has been suggested that the triplet of G residues may help prevent slippage between the RNA-DNA complex and the active site during translocation, which in turn reduces the frequency of aborted sequences [280]. Expansion of the bubble becomes much less favourable after the +8 position, highlighting the possible significance of the +4 to +8 region. It has also been demonstrated in more recent studies that altering the nucleotides in positions +4 to +8 can affect transcription activity by up to 5-fold [280]. A major observation from these studies was sequences that are AU-rich in the +4 to +8 region generally had the highest activity. This is likely due to the facilitation of strand melting upon transcription initiation [280]. These observations were taken into account when designing truncations of the SRB-2 aptamer.

As seen in Figure 4.2 A, the original sequence of the UUCG stem-loop in SRB-2 is a 12mer with the sequence 5'-UCGCUUCGGCGA-3'. In order to make this sequence more suitable for transcription

while maintaining maximum sequence analogy, the A-U base pair was removed, and three consecutive G-C base pairs were added (Figure 4.2 B). The middle ten bases are identical to SRB-2. Due to the A-U base pair being situated at the end of the stem in SRB-2, the chemical environment of these two bases would likely be significantly different with the addition of a G triplet, and therefore, they do not warrant inclusion here.

In our initial attempt to design a truncation for loop 3, the A-U base pair in the stem was simply changed to a G-C base pair, resulting in a run of four G-C base pairs to begin the sequence (Figure 4.2 C). The transcriptional yield of this sequence, titled L3, was quite poor, even when a double-stranded DNA template was used. For this reason, several modified versions of this sequence were designed and screened in order to optimize transcription (Figure 4.2 D-F). In the sequence L3A, the fourth G-C base pair was substituted with an A-U base pair to increase the A-U content as described above. The presence of an A-U base pair in the stem may also help to simplify resonance assignment. In L3B, the fourth G-C base pair was again replaced but this time by a C-G base pair. This was done to replicate the start of the full SRB-2 aptamer sequence, where transcription activity was already sufficient. In the third and final version, L3C, the fourth G-C base pair and the following C-G base pair were removed and replaced with three A-U base pairs. This was done to substantially increase the AU content and attempt to replicate the results of the literature referenced above. These sequences were transcribed, run on a urea PAGE gel and qualitatively assessed by UV shadowing. The transcription yield observed was $L3 \ll L3A < L3B < L3C$. It was therefore decided that L3C would be the sequence used for NMR studies.

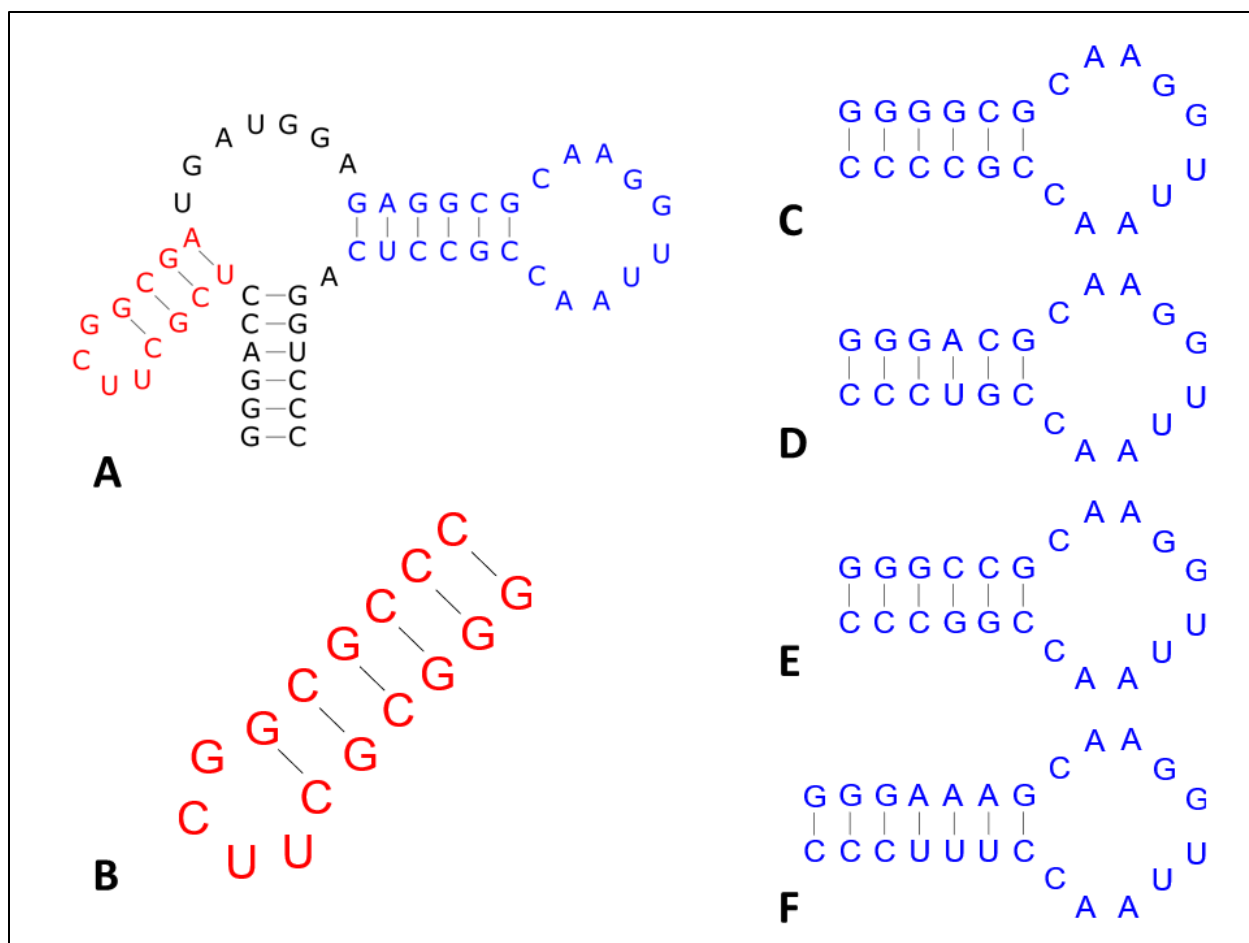


Figure 4.2: Segments of the SRB-2 aptamer designed for NMR studies.

A) SRB-2, B) UUCG, C) L3, D) L3A, E) L3B and F) L3C. The UUCG tetraloop (loop 1) is shown in red and loop 3 is shown in blue. The remainder of the sequence is shown in black.

Before an NMR sample was synthesized, we wanted to ensure that the structure of L3C was representative of the original sequence. To do this, we designed two versions of the full SRB-2 sequence with alterations to the stem of loop 3. SRB-2C has a stem identical to L3C, and SRB-2D has a stem similar to L3B. SRB-2C was designed as a direct comparison for L3C, but this sequence involves an additional base pair in the stem. That means that differences in binding affinity observed may be a result of either a change in sequence or a change in stem length. In order to distinguish these effects, the SRB-2D sequence was created, as it carries some of the same alterations as SRB-2C, but its stem is the same

length as the original SRB-2 sequence. These sequences were examined using fluorescence techniques. They are shown in Figure 4.3.

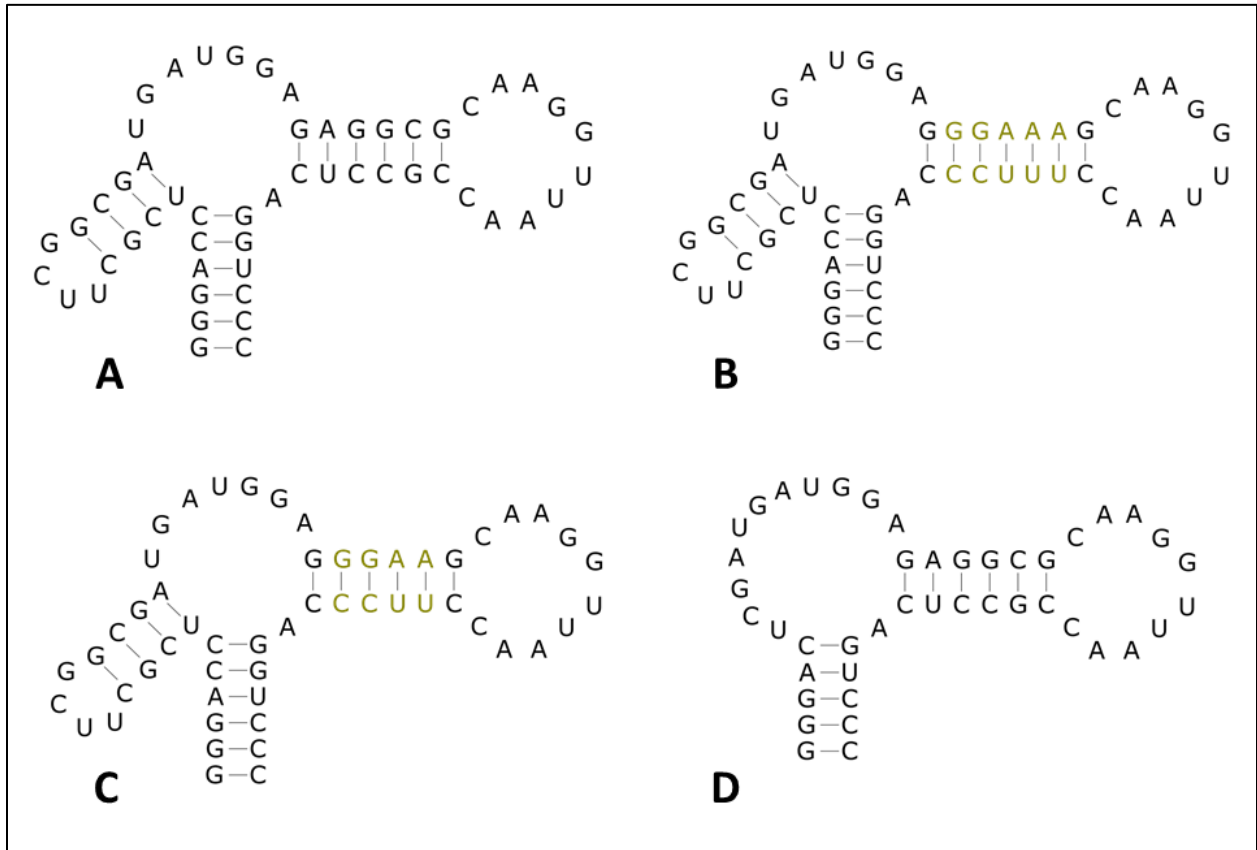


Figure 4.3: SRB-2 variants studied using fluorescence titrations.

A) SRB-2, B) SRB-2C, C) SRB-2D and D) SRB-2 min. For B and C, region that varies is shown in green.

Some of the sequences discussed above were screened for SR binding using fluorescence emissions scans (Figure 4.4). UUCG and L3C by themselves showed no discernable binding activity, which is unsurprising due to the large amount of sequence that was removed. It was also observed that the binding activity of both SRB-2C and SRB-2D were comparable to SRB-2. The final sequence included in the emission scans was SRB-2 min. This was a truncation created by the authors who originally selected SRB-2, and we tested it to see if it was worth studying further. However, this sequence showed

only a slight shift in maximum emission wavelength and no change in fluorescence intensity. This indicates that at best, this sequence binds extremely weakly, which is unsuitable for NMR studies.

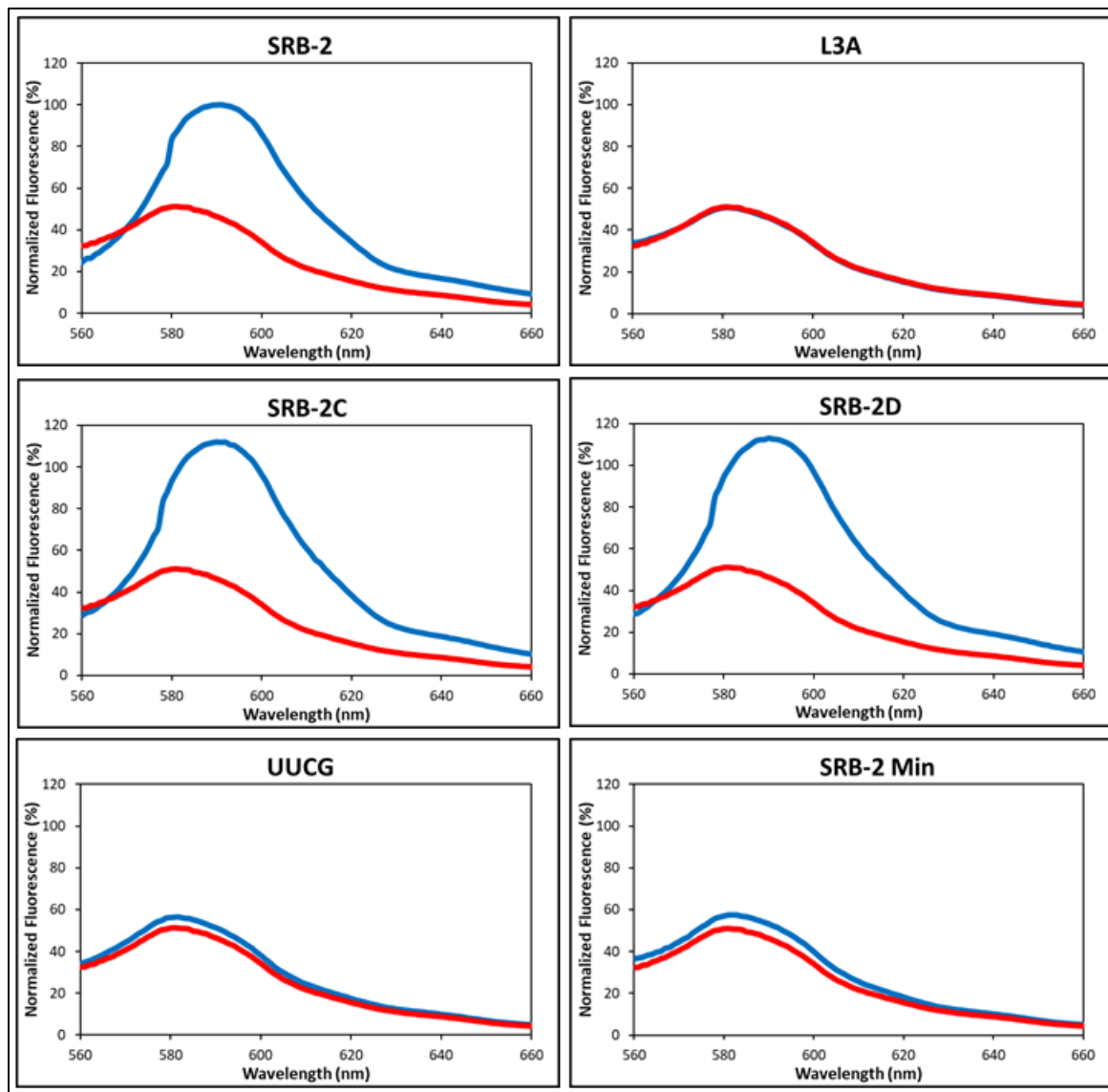


Figure 4.4: Fluorescence emission scans of SRB-2 mutants and truncations.

Emission spectra of ligand on its own are shown in red and emission spectra of ligand in the presence of excess SRB-2 are shown in blue. SRB-2, SRB-2C and SRB-2D all clearly bind SR, while L3A, UUCG and SRB-2 Min do not.

In order to compare the binding affinities of SRB-2, SRB-2C and SRB-2D, fluorescence titrations were performed (Figure 4.5). As shown in Table 4.2, it was determined from these experiments that SRB-2C and SRB-2D had only slightly weaker binding affinity than SRB-2. This observation shows that the same general structure is likely being formed in the presence of ligand and that the sequence alterations have minimal effect on this structure. L3C should therefore be a reasonable representation of loop 3 in NMR studies.

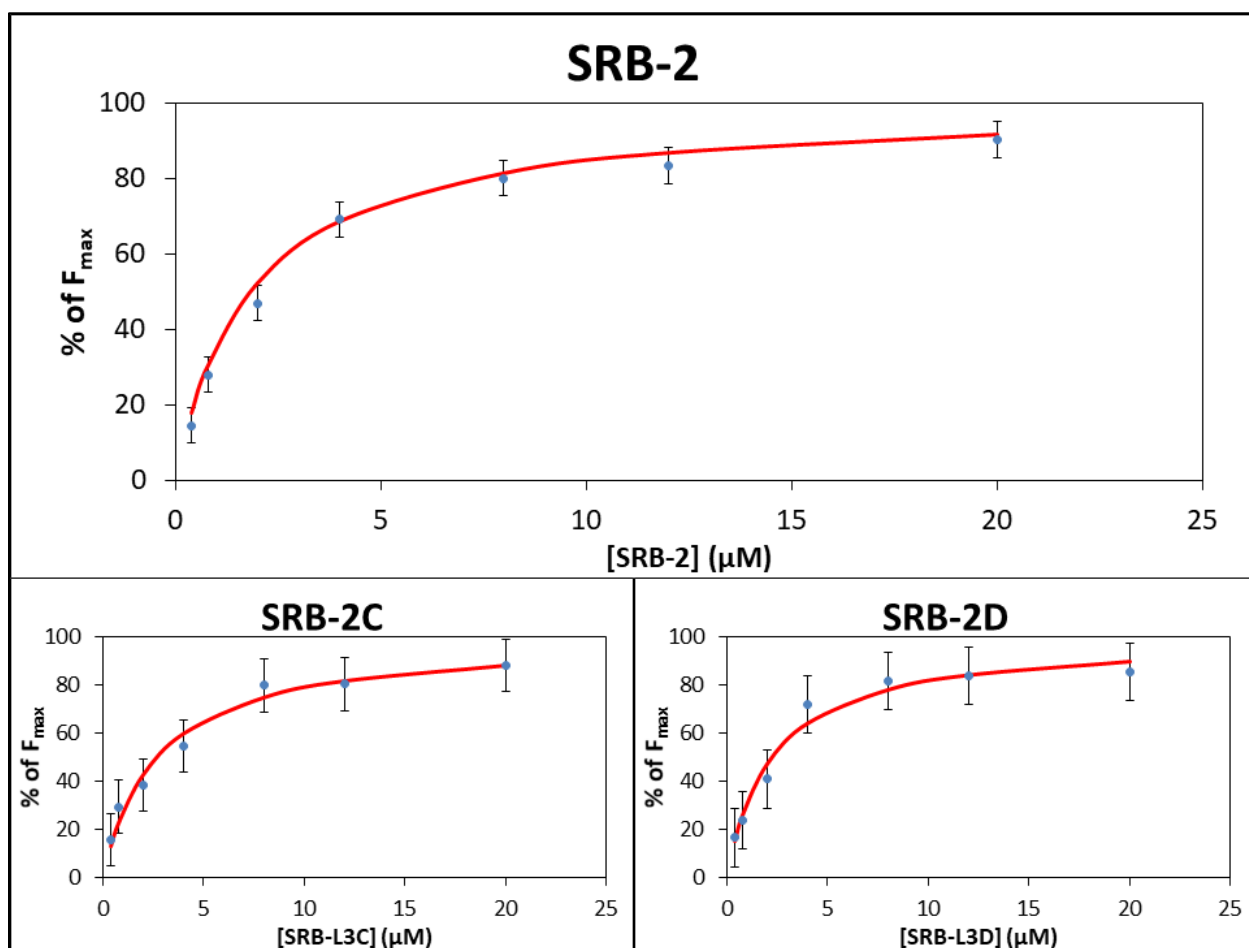


Figure 4.5: Fluorescence titrations of SRB-2 mutants with SR.

Experimental data is shown as blue points and fit data is shown as a red line. Concentration of SR was 5 μM and SRB-2 ranged in concentration from 0.5-25 μM . Error bars are one standard deviation from the mean of the replicates acquired.

Table 4.2: Dissociation constants of SRB-2 mutants determined by fluorescence titration.

Sequence	K_d (μM)
SRB-2	2.09 ± 0.12
SRB-2C	2.69 ± 0.30
SRB-2D	2.23 ± 0.27

4.4.2 NMR Structure of the UUCG 16mer

Due to its small size and stable structure, resonance assignment of the UUCG 16mer was completed using only homonuclear experiments, including a NOESY in 90% H₂O/10% D₂O and NOESY, TOCSY and COSY in D₂O. Inter-residue connectivity was first established by assigning the cross-peaks in the H6/H8 to H1' region of the NOESY in D₂O. This NOESY is shown in Figure 4.6 A, and a close up of the H6/H8 to H1' region is shown in Figure 4.6 B. After the chemical shifts of the H6/H8 proton from each base were determined, assignments were extended from each base to the H2', H3', H4', H5' and H5'' protons.

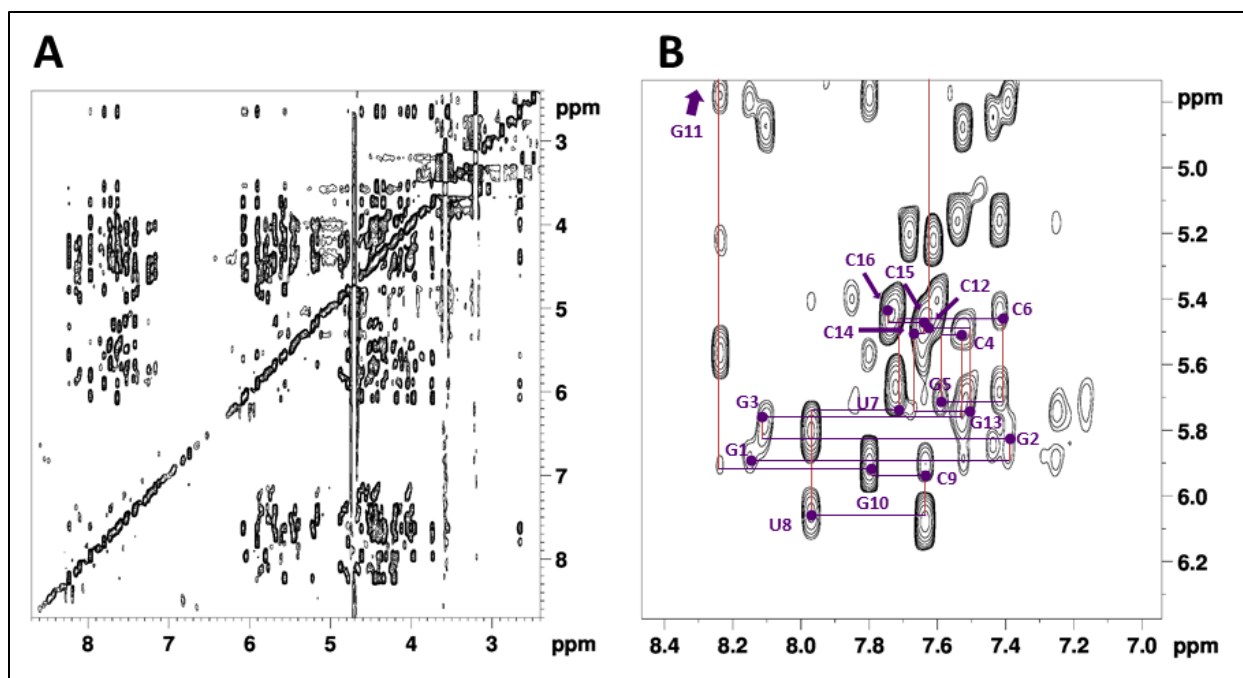


Figure 4.6: NOESY of 2.0 mM UUCG 16mer in D₂O.

A) Full spectrum. B) Close-up of the H1' to H6/H8 region. A sequential 'walk' is labelled on this spectrum, where cross-peaks connected by horizontal purple lines are two different bases that see the same sugar, and cross-peaks connected by vertical red lines are two different sugar protons that see the same base. Blue dots indicate intra-nucleotide cross-peaks. G11 has an H1' with an uncharacteristically low chemical shift and appears outside of this figure's chemical shift range. This is represented by a purple arrow. Spectrum was acquired with mixing time of 150 ms.

Upon completing resonance assignments for the non-exchangeable protons, the NOESY in 90% H₂O/10% D₂O was used to assign the exchangeable imino and amino protons on the bases where possible. The full spectrum is shown in Figure 4.7 A, and a close up of the imino-imino region of the spectrum is shown in Figure 4.7 B. Assignment was relatively straightforward for imino protons belonging to nucleotides in the tetraloop, but unambiguously assignments were not determined for the amino and imino protons in the stem due to significant overlap. The absence of any A-U content in the

stem contributed to this problem, resulting in a lack of chemical shift variation. Stacking NOEs between adjacent guanines in the stem were not clearly observed in this spectrum. This limited the usefulness of the imino-imino region shown in Figure 4.7 B. Assignment of the UUCG tetraloop resonances in the full-length SRB-2 sequence was made possible by this work as shown in Figure 3.5.

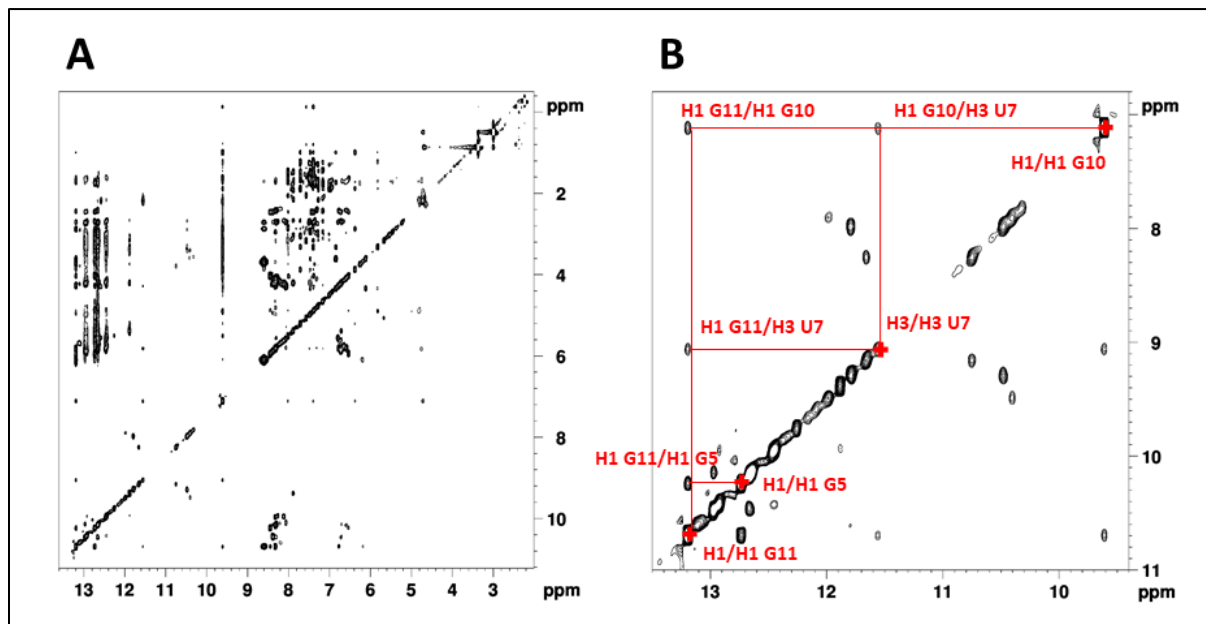


Figure 4.7: NOESY of 2.0 mM UUCG 16mer in 90% H₂O/10% D₂O.

A) Full spectrum. B) Imino-to-imino region. Relevant assignments from the tetraloop are shown in red. Spectrum was acquired with mixing time of 150 ms.

A TOCSY experiment was also used to help identify the H5-H6 cross-peaks from C and U in the NOESY in D₂O and to distinguish base-sugar cross-peaks involving the pyrimidine residues from those involving purine residues (Figure 4.8 A). There are nine C and U residues in the UUCG 16mer and ten peaks observed in the TOCSY. However, it appears likely that the two weakest peaks belong to the same resonance, allowing unambiguous assignment of the spectrum. A DQF COSY was also obtained and was

used to confirm the known S-type sugar puckers in the tetraloop, U8 and C9. This spectrum is shown in Figure 4.8 B.

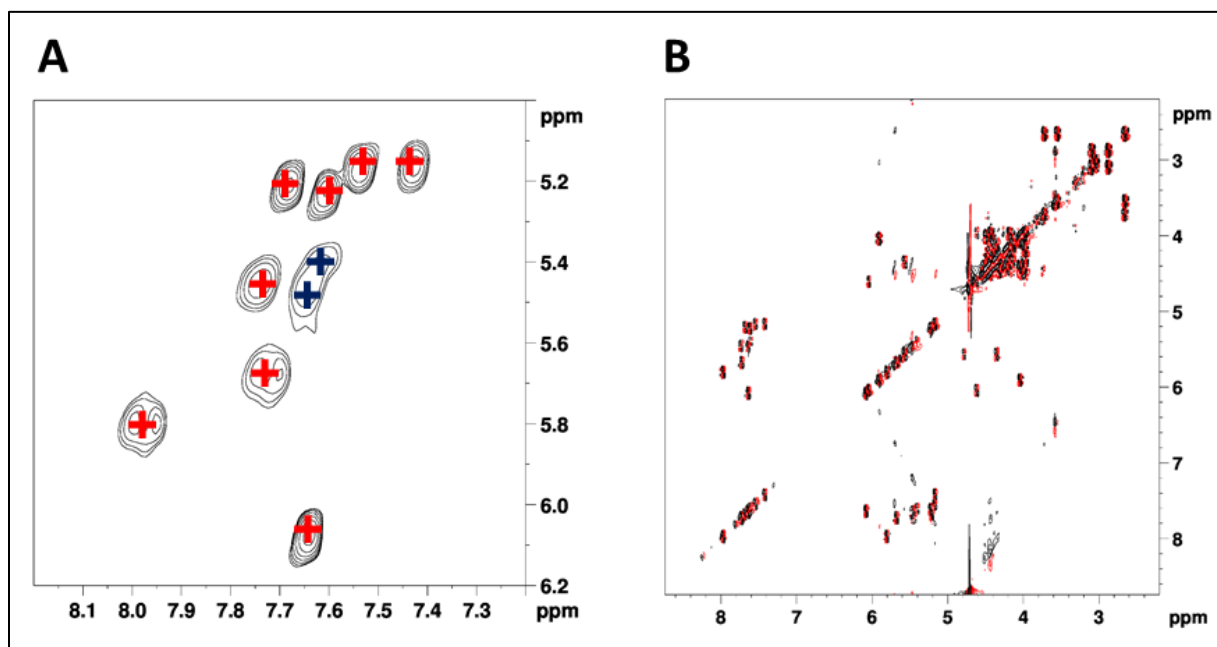


Figure 4.8: TOCSY and COSY spectra of 2.0 mM UUCG 16mer.

A) H5-H6 region of the TOCSY spectrum. Strong signals from the dominant conformation are marked in red, and weaker signals experiencing chemical exchange are marked in blue. B) DQF COSY where positive contours are shown in black and negative contours are shown in red. Spectrum was acquired with mixing time of 50 ms.

Distance constraints and dihedral angles were obtained as outlined in 4.3.5. A total of 540 NOE distance constraints and 135 dihedral angles were used, which corresponds to an average of 33.8 distance and 8.4 dihedral angle restraints per nucleotide as listed in Table 4.3. The distance constraints include a total of 40 hydrogen bonds that were added for the six base-pairs in the stem of the molecule as well as four in the tetraloop since previous studies indicate that these exist and are observed in experimental data. Within the tetraloop portion of the molecule, 145 NOEs were obtained, 47 of which

were inter-residue NOEs. Therefore, the tetraloop structure was defined by an average of 36.3 NOE constraints per nucleotide.

Table 4.3: Structure determination statistics for UUCG 16mer

Distance Restraints	
Tetraloop (7-10) intra-nucleotide	98
Tetraloop (7-10) inter-nucleotide	47
Total intra-nucleotide	363
Total inter-nucleotide	137
Hydrogen Bonds	40
Total distance restraints	540
Dihedral Restraints	
Ribose pucker	67
Backbone	68
Total dihedral restraints	135
Structural Statistics	
NOE violations	None > 0.5 Å
Angle violations	None > 5°
Loop (6-11) RMSD	0.629 Å
Stem (1-5, 12-16) RMSD	1.808 Å

For visualization and RMSD calculation, the structure was separated into two segments. Due to the lack of long-range restraints included in the calculation, large variations in the bending of the stem were observed and aligning the two halves separately resulted in the structure being displayed more clearly. The first segment, shown in Figure 4.9 A, includes the tetraloop and the first base pair of the stem (residues 6-11). The second segment, shown in Figure 4.9 B, is the rest of the stem (residues 1-5, 12-16).

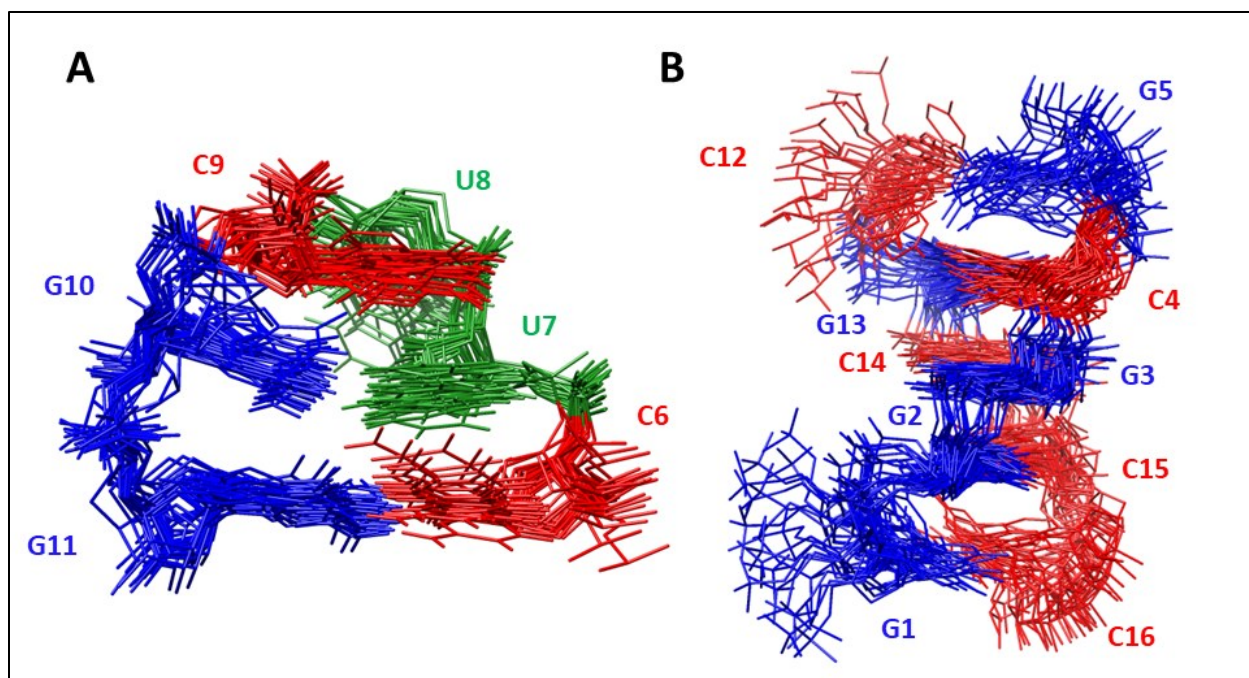


Figure 4.9: Stereo view of the superposition of the 20 lowest energy structures of the UUCG 16mer.

Separated into A) UUCG tetraloop and first base of the stem (nucleotides 6-11) and B) The stem (nucleotides 1-5 and 12-16). G residues are shown in blue, C's in red and U's in green.

Figure 4.10 shows a comparison between one of the 20 structures obtained in previous studies [268] and one of the 20 obtained in this work. As expected, there is little discernable difference between the two sets of structures. This is partly due to the inclusion of hydrogen bond restraints based on these previous structures. These include several of the defining interactions in the tetraloop: the *trans*-wobble U7-G10 base pair, the C9 amine to the U8 phosphate and the O6 of G10 to the O2' of U8 [266,268]. These were all confirmed experimentally. As shown in Table 4.3 and Figure 4.9, a reasonably small deviation between structures was obtained. Additionally, this structure was already well defined and our calculation compared to literature versions relatively well. As such, we were not justified in spending any additional time refining the structure.

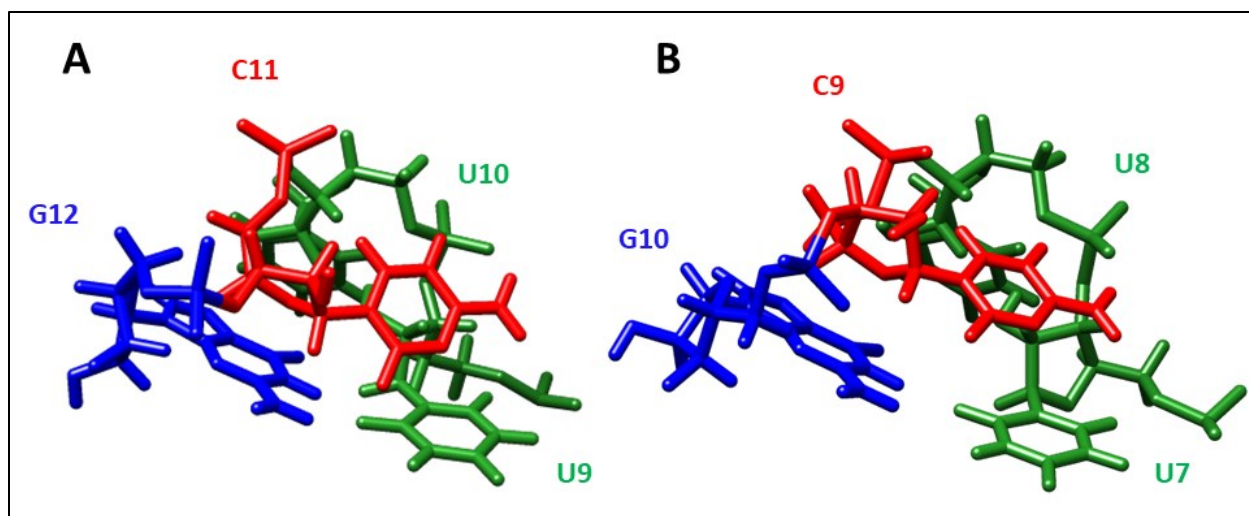


Figure 4.10: Literature comparison of the UUCG tetraloop.

Lowest energy structure of A) the UUCG tetraloop as determined by Allain and Varani (1HLX) [269] and B) the UUCG tetraloop as determined in these studies.

4.4.3 NMR Studies of the SRB-2 Truncation L3C

Following the successful resonance assignment and structure calculation of the UUCG 16mer, the next step was to design and study another segment of the SRB-2 aptamer. Several variations of the ‘loop 3’ stem-loop were designed, and we proceeded with NMR studies of L3C, for reasons discussed above. As with the UUCG 16mer, NOESYs in both D₂O and 90% H₂O/10% D₂O were collected for this sample, as well as a TOCSY. From the NOESY spectra, it is evident that some of the heterogeneity issues seen with the full SRB-2 sequence were significantly worse with this sequence. The peaks in the D₂O spectrum are very broad, resulting in poor resolution. In the NOESY in 90% H₂O/10% D₂O (Figure 4.11), only two imino resonances are observed, likely from the middle of the stem region. There are no peaks in the region where non-Watson-Crick type pairing is typically seen. This suggests that L3C is highly dynamic in solution, forming little tangible structure outside of the stem.

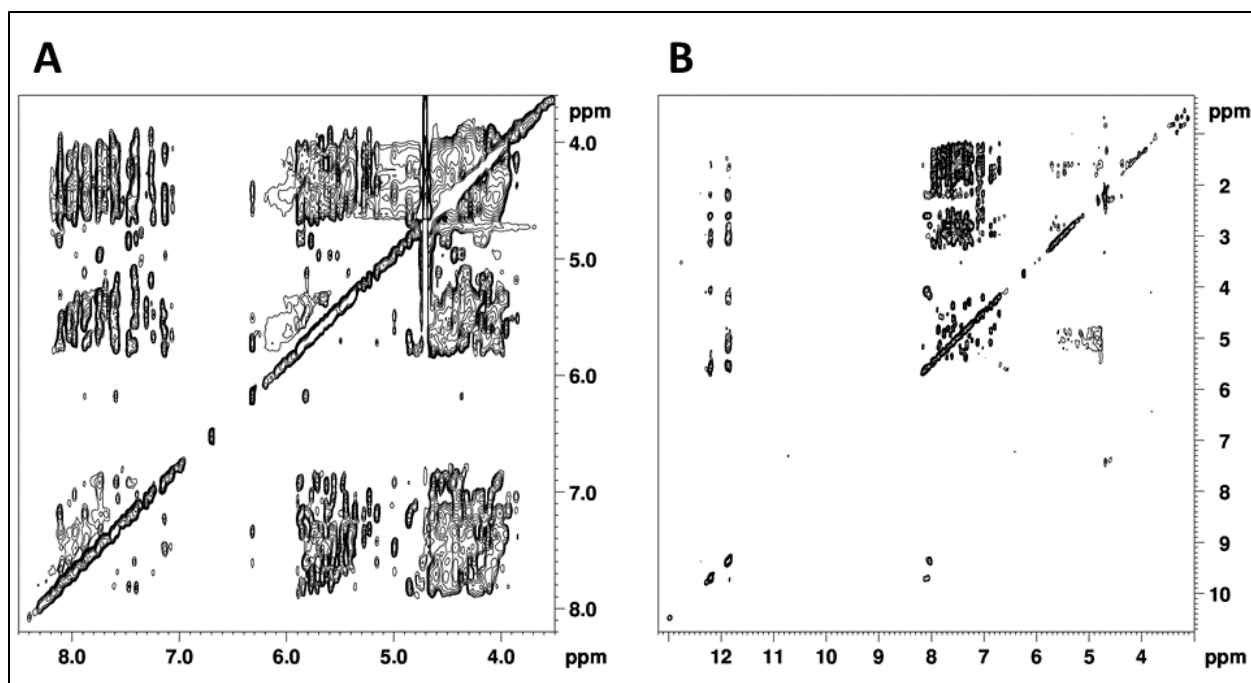


Figure 4.11: 2D NOESY spectra of L3C.

1.5 mM RNA in A) D_2O and B) 90% H_2O /10% D_2O . Very little structure is observed in these spectra.

Spectra were both acquired with a mixing time of 150 ms.

This is supported by the TOCSY (Figure 4.12), where it is clear that multiple conformations exist. There are eleven peaks expected in the H5-H6 region of the TOCSY (six C and five U in L3C), but at least 17 peaks can be counted in this spectrum. At this point, we concluded that resonance assignment was not possible for this segment of SRB-2. Based on the results obtained, it is highly likely that nucleotides in the loop 3 region of the sequence interact with some of those in the center loop to form the binding site. This, unfortunately, results in segmental analysis being impractical for SRB-2.

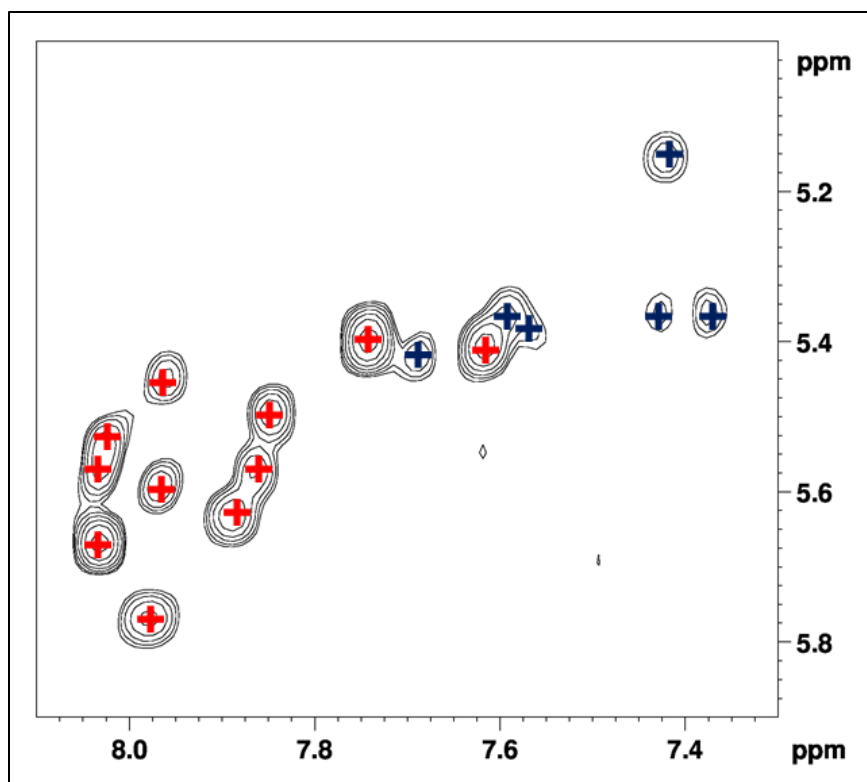


Figure 4.12: H5-H6 region in the 2D TOCSY of 1.5 mM L3C.

Peaks are labelled to show how they were counted. Peaks labelled in red are strong, and those labelled in blue are weak. Blue peaks are likely a result of structural homogeneity.

4.5 Conclusions

Overall, the segmental analysis of SRB-2 yielded mixed results. The UUCG 16mer was successfully assigned, and 20 structures of satisfactory energy and residual violations were acquired. As shown in Figure 4.10, no discernible differences were observed between the experimental and literature structures. When compared to SRB-2 spectra, cross-peaks in the tetraloop can be easily identified, particularly those belonging to exchangeable resonances in the NOESY spectrum in 90% H₂O/10% D₂O (Figure 3.5). Unfortunately, studies with the L3C sequence were unsuccessful due to the lack of tertiary structure present. This was not an unexpected result considering that the ligand is unable to bind this sequence. There is also a palindromic sequence in the loop that may cause dimerization and therefore

structural heterogeneity in the NMR spectra. Based on this significantly lower degree of structure observed in the NMR spectra for L3C compared to SRB-2, it is highly likely that there is significant interaction between loop 3 and loop 2 in the folded structure SRB-2. Therefore, segmental analysis will not be useful beyond the UUCG tetraloop. A minimal version of SRB-2 aptamer, SRB-2 Min, was proposed by the authors who originally selected SRB-2 [37]. The UUCG tetraloop is absent in this structure, while loop 2 and loop 3 are intact. However, fluorescence studies shown in Figure 4.4 indicate that the binding affinity of this sequence is quite low at best, which does not make NMR studies a realistic option. If future NMR studies are done on this system, finding conditions to limit exchange and hence peak broadening, would be the primary goal. Then, the isotope labelling strategies discussed in Chapter 3 may be more effective.

Chapter 5: Thioflavin T Fluorescence and NMR Spectroscopy Suggesting a Non-G-quadruplex Structure for a Sodium Binding Aptamer

5.1 Foreward

The results in this chapter have been accepted for publication in the Canadian Journal of Chemistry: Runjhun Saran, **Kyle A. Piccolo**, Yanping He, Yongqiang Kang, Po-Jung Jimmy Huang, Chunying Wei, Da Chen, Thorsten Dieckmann and Juewen Liu. "Thioflavin T fluorescence and NMR spectroscopy suggesting a non-G-quadruplex structure for a sodium binding aptamer embedded in DNAzymes." Any permission for further re-use of this material should be requested directly from the Canadian Journal of Chemistry.

*R.S., K.A.P. and Y.H. contributed equally to this work.

ThT fluorescence spectroscopy experiments were performed and analyzed by Runjhun Saran, Yanping He, Yongqiang Kang, Po-Jung Jimmy Huang, Chunying Wei and Da Chen. NMR and CD experiments were performed and analyzed by the candidate. All authors contributed to the text in the abstract, introduction, and conclusion sections, while the methods and results and discussion sections were written primarily by the author(s) who used those respective methodologies.

All data and figures in this chapter were published in the above article with the following exceptions: Figures 5.11-13 and analysis in section 5.5.8.

5.2 Chapter Abstract

Recently, a Na^+ -binding aptamer was reported to be embedded in a few RNA-cleaving DNAzymes, including NaA43, Ce13d and NaH1. These DNAzymes require Na^+ for activity but show no activity in the presence of K^+ or other metal ions. Given that DNA can selectively bind K^+ by forming a G-quadruplex structure, this work aims to answer whether this Na^+ aptamer also uses a G-quadruplex to bind Na^+ . The Na^+ aptamer embedded in Ce13d consists of multiple GG sequences, which is also a prerequisite for the formation of G4 structures. To delineate the structural differences and similarities between Ce13d and G-quadruplex in terms of metal binding, thioflavin T (ThT) fluorescence spectroscopy, NMR spectroscopy and CD spectroscopy were used. Through comparative ThT fluorescence spectrometry studies, we deciphered that while a control G-quadruplex DNA exhibited notable fluorescence enhancement up to 5 mM K^+ with a K_d of 0.52 mM, the Ce13d DNAzyme fluorescence was negligibly perturbed with similar concentrations of K^+ . Opposed to this, Ce13d displayed specific remarkable fluorescence decrease with low millimolar concentrations of Na^+ . NMR experiments at two different pH values suggest that Ce13d adopts a significantly different conformation or equilibrium of conformations in the presence of Na^+ versus K^+ and has a more stable structure in the presence of Na^+ . Additionally, absence of characteristic G-quadruplex peaks in 1D ^1H NMR suggest that G4 is not responsible for the Na^+ binding. This theory is confirmed by absence of characteristic peaks in the CD spectra of this sequence. Therefore, we concluded that the aptamer must be selective for Na^+ and binds using a structural element that does not contain G4.

5.3 Introduction

Understanding metal-binding to DNA is important not only for studying the biological functions of DNA, but also for biosensor development [171,281], drug development [282], and nanotechnology [283]. In biological studies, Na^+ and K^+ are among the most abundant physiological metal ions. They can

control the ionic strength of buffers and solutions and screen the negative charges on DNA, resulting in more stable DNA duplexes [284]. In addition, they can also have specific binding interactions with certain single-stranded DNA sequences [171]. The most famous example is of the stabilization of G-quadruplex (G4) DNA [285]. Normally K^+ is much more effective than Na^+ in stabilizing G4 structures [286,287]. Na^+ is less effective, often attributed to its smaller size and also because thermodynamically it has a higher energy of dehydration [287].

Recently, a Na^+ -binding DNA aptamer has been reported, [288,289] which was derived from the conserved sequences of the DNAzymes NaA43, NaH1, and Ce13d all originally discovered through *in-vitro* selections [290–293]. The NaA43 DNAzyme was reported by Lu and co-workers [290], and it specifically requires Na^+ for cleaving an RNA containing substrate. NaA43 shares its conserved sequence with the Ce13d DNAzyme, which was selected by our group in a lanthanide-dependent selection [291]. The conserved sequence is the main part of a Na^+ -binding aptamer [294–297]. The identification of this Na^+ -aptamer proved instrumental in understanding the reason for the specificity of NaA43 and Ce13d DNAzymes for Na^+ , although the mechanism underlying specific Na^+ binding by the DNA still remains intriguing [289,296,298,299]. Our knowledge on specific Na^+ binding by DNA is limited and from the literature known, and a possible mechanism may rely on G4 structures. In such a case, the G4 structure would require a superior Na^+ -induced stabilization than K^+ , as the aptamer is known to show a higher affinity to Na^+ in comparison to K^+ , especially at room temperature [288,295,298,300]. Outside the G4 context, Na^+ binds more strongly to DNA than K^+ since it can better increase the melting temperature (T_m) of DNA [301]. With respect to G4 structures, so far only a few specialized examples are known where Na^+ can stabilize G4 more than K^+ does. Alberti and coworkers reported a structure containing two contiguous G4 units with a greater stabilization by Na^+ [302]. Other examples of Na^+ being a better stabilizer were all from mutated human telomeric sequences, but the advantage of Na^+ was extremely small. For example, by replacing a certain guanine with a O^6 -methylguanines, the T_m was enhanced by

just 1 °C with Na⁺, while the T_m of the original DNA was 8 °C higher with K⁺ [303]. Moderate advantages were also observed by replacing certain guanines by abasic sites [304], or adenines [305]. Overall, such mutations significantly decreased the overall stability of the G4 structures. Sun *et al.* reported an unmodified G4 sequence that showed different folding in the presence of Na⁺ and K⁺, and Na⁺ appeared to have an even tighter binding [306]. For unmodified simple G4 sequences, no examples are known for Na⁺ being a better stabilizer. Therefore, it would be extremely intriguing to probe whether the mechanism underlying Na⁺-binding to the aptamer derived from NaA43 and Ce13d DNazymes involves Na⁺-G4 interactions.

Another interesting facet regarding Ce13d is its ability to act as a detection beacon for trivalent lanthanides [291]. Lanthanides are used in commercial applications such as magnets, batteries, lasers and contrasting agents, so the ability to detect these metal ions is important to identify and prevent pollution [307–310]. Their similarity in physical properties makes them difficult to separate without the use of complex instrumentation methods [311]. Previous studies with Ce13d indicate that several trivalent lanthanides can be detected in the low nanomolar concentration range [291]. This DNazyme was the product of the first selection to contain lanthanides as the sole metal cofactor, though lanthanides have been noted to have activating or inhibitory effects on nucleic acid catalysis in the past [312–316]. From a practical perspective, lanthanides are a promising choice of cofactor for nucleic acid cleavage since their ability to act as strong Lewis acids make them adept at performing non-specific nucleic acid hydrolysis [309,317,318]. Here we examine by NMR the structural effects that lanthanide presence has on Ce13d.

Thioflavin T is a popular dye that becomes fluorescent upon binding to G4 DNA, and it has been extensively used to probe G4 [319]. In addition, NMR is a powerful method for studying G4 structures [199,320] and for studying biomolecule-ligand interactions in general. In this work, we used ThT to study Na⁺ binding by the Ce13d DNazyme and a comparison was made with a G4 structure. In addition, NMR

spectroscopy was performed to further analyze the Ce13d DNAzyme structure. The results argued against the presence of a G4 structure to be responsible for the specific Na⁺ binding by the aptamer. NMR results also show that sodium and a lanthanide together are not sufficient to observe Cis13d in a single rigid conformation.

5.4 Materials and Methods

5.4.1 Chemicals

The DNA sequences were obtained from Integrated DNA Technologies (Coralville, IA) and Eurofins (Huntsville, AL). Metal salts including lithium chloride (LiCl), sodium chloride (NaCl), lanthanum (III) nitrate (La(NO₃)₃), KH₂PO₄, K₂HPO₄, Na₂HPO₄, and NaH₂PO₄ were obtained from Sigma-Aldrich, VWR, and Fischer Scientific Canada at the highest purity available. ThT was from Sigma-Aldrich. 99.996% D₂O was from Cambridge Isotope Laboratories.

5.4.2 ThT Fluorescence Spectroscopy

For ThT fluorescence spectroscopy, the Ce13d DNAzyme or G4 complexes were annealed at a final concentration of 20 μM in buffer A (25 mM LiCl, 50 mM HEPES, pH 7.5) by heating the samples to 85 °C for 5 min and then gradually cooling to 4 °C over 30 min. For the experiments, final concentration of 0.6 μM DNA complexes were added to a final concentration of 3 μM ThT solution in buffer B at room temperature (500 mM tris-acetate, pH 8). After 15 min reaction at 4°C, the sample was recovered to room temperature. Then fluorescence readings were collected on a Cary Eclipse fluorometer in a 1x1 cm quartz fluorescence cuvette with the excitation wavelength (λ_{exc}) as 442 nm and the scanning emission wavelength (λ_{emm}) range from 455 to 650 nm at room temperature.

5.4.3 Nuclear Magnetic Resonance

DNA for NMR experiments was purified by 10% denaturing polyacrylamide gel electrophoresis (dPAGE). The DNA was eluted from the dPAGE using 300 mM LiCl. This was followed by purification on a HiPrep 16/10 DEAE FF anion-exchange column (GE Healthcare, Uppsala, Sweden), and desalting on a HiPrep 26/10 Desalting column (GE Healthcare, Uppsala, Sweden). Buffers containing only Li⁺ cations (no Na⁺ or K⁺) were used throughout purification. NMR samples were prepared by dissolving an appropriate weight of lyophilized powder in 400 μL of either water (no salt samples), 5 mM sodium phosphate buffer and 5 mM NaCl, 5 mM potassium phosphate buffer and 5 mM KCl, 80 mM NaCl or 80 mM KCl. The pH was adjusted to 5.8 or 6.8 with ammonia, NaOH or KOH depending on the cation already present. The samples were dried by lyophilization and re-dissolved in 500 μL of 90% H₂O/10% D₂O or 99.996% D₂O. Samples were heated to 85°C for 5 min and cooled to 4°C before spectra were acquired. All spectra were collected on a Bruker DRX-600 spectrometer equipped with a HCN triple-resonance, triple-axis PFG probe (Bruker, Billerica, MA). NMR experiments were carried out at 277 K in 90% H₂O/10% D₂O or 298 K in D₂O. Solvent suppression was achieved using 1 $\bar{1}$ -spin echo pulse sequences [190] for 90% H₂O/10% D₂O or presaturation [192] for D₂O samples. All 1D spectra were processed using identical parameters and window function. The 2D CITY TOCSY experiments [193] were run with a mixing time of 50 ms and NOESY experiments [191] were run with a mixing time of 150 ms. Quadrature detection for the indirect dimension was achieved using the States-TPPI method [189]. All NMR samples had a volume of 500 μL and were read in standard 5 mm NMR tubes. Relevant pulse programs can be found in Appendix B.

5.4.4 Circular Dichroism

CD experiments were performed on a Jasco J-815 spectropolarimeter (Jasco Inc., Easton, MD). CD scanning experiments were run from 330 nm to 200 nm with a path length of 0.1 cm, data interval of 0.5 nm, bandwidth of 0.5 nm, response of 1 second, scanning speed of 200 nm minute⁻¹ and a total of

four accumulated scans. Samples contained 5 μM DNA at pH 6.8 and either H_2O , 80 mM KCl or 80 mM NaCl. The samples were also heated to 85 $^\circ\text{C}$ for 5min, cooled to 4 $^\circ\text{C}$ and incubated for at least 24hrs before acquisition at 25 $^\circ\text{C}$.

5.5 Results and Discussion

5.5.1 The Ce13d DNAzyme

The secondary structure of the Ce13d DNAzyme is shown in Figure 5.1 A [291]. Its substrate strand (shown in green) contains a single RNA linkage (rA in red for ribo-adenine) that serves as the cleavage site. For most of the studies in this work, this RNA linkage was replaced by its DNA analog to avoid cleavage (dA in red for deoxyribo-adenine). Previous assays have shown that such a change does not perturb Na^+ binding [288,300]. The two ends of the enzyme strand (shown in blue) bind the substrate via two stems (shown as blue/green duplexes in Figure 5.1 A) respectively, and between these two stems the enzyme contains a hairpin (shown as blue) followed by a large loop (shown as red and purple), which is the main part of the Na^+ aptamer. Stretches of G bases present in the catalytic loop of Ce13d DNAzyme as well as the substrate strand are highlighted in purple (Figure 1A). G4 structures are composed of stacked G-quartet, where each quartet consists of 4 guanines Hoogsteen base-paired in a square planar array (Figure 5.1 D). G4s may form by one to four nucleic acid strands that bear continuous runs of guanines or G-tracts in presence of metal ions such as K^+ [321,322].

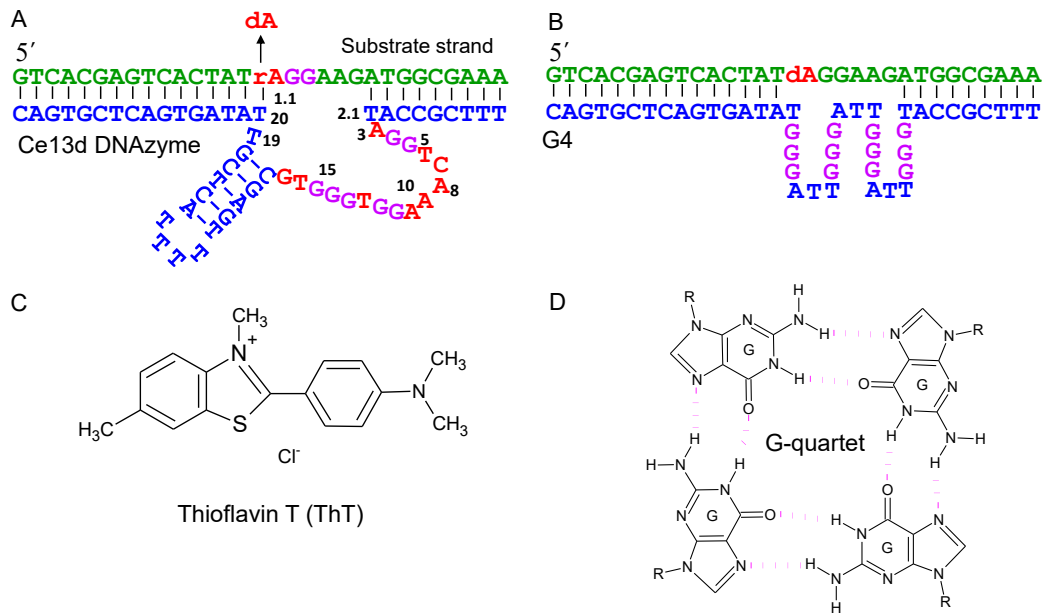


Figure 5.1: Secondary structure of molecules used in ThT fluorescence experiments.

(A) the Ce13d DNAzyme and (B) G4 construct, designed by replacing the Ce13d catalytic loop by a G4 DNA. The guanine stretches are marked in purple. (C) The structure of ThT. (D) Structural representation of a G-quartet, where the hydrogen bonds are shown in pink color, G stands for guanine, and R depicts the rest of the nucleic acid chain attached to G.

From the secondary structure of Ce13d, we can find four GG or GGG stretches (purple, Figure 5.1 A) in its catalytic loop, and thus it has the chemical components to form a G4. From the previously published DMS foot-printing experiment, most of these guanines in the enzyme strand were protected in presence of Na^+ , indicating that these guanines are involved in the Na^+ -binding pocket [288]. However, this DNAzyme is known to be inactive with K^+ [298]. In addition, upon replacing one of the critical guanines in the enzyme catalytic loop with base hypoxanthine, the modified Ce13d DNAzyme still retained the same Na^+ -induced activity [323]. This modification must disrupt G4 structures, however it did not hamper the Ce13d activity. Therefore, whether Ce13d uses G4 to bind Na^+ remains elusive. To

address this problem through comparative studies, we designed a G4 construct as a positive control, in which we replaced the Ce13d catalytic loop with a G4 sequence (Figure 5.1 B).

5.5.2 ThT fluorescence spectroscopy

We started by using ThT to probe for the presence of G4 structures in the Ce13d DNAzyme and the G4 control sequence. The structure of ThT is shown in Figure 5.1 C, and it is commonly used for staining G4 DNA [319,324–326] although ThT also has its limitations such as it prefers to bind parallel G-quadruplex over anti-parallel ones [327,328]. Before studying our Ce13d DNAzyme, we first did a control experiment using the G4 construct in Figure 5.1 B. We mixed ThT with this G4 structure, and an emission peak at 488 nm was observed with 442 nm excitation (Figure 5.2 A, black spectrum). Upon adding 10 mM K^+ , an increase in the fluorescence was observed, suggesting formation of a G4 structure (Figure 5.2 A, red spectrum). For quantitative understanding, we gradually titrated K^+ (Figure 5.2 B, green trace) to see a concentration-dependent effect. A sharp increase in fluorescence occurred between 0 and 5 mM K^+ and then the fluorescence saturated. A K_d of 0.52 mM K^+ was obtained by fitting the curve. With more than 10 mM K^+ , the fluorescence started to drop, which might be attributed to the general effect of salt in screening the interaction between ThT and the DNA. While the increase in fluorescence in Figure 5.2 B was sharp, it was relatively small in terms of fold-enhancement i.e. ~ 2 -fold. This could be attributed to the long DNA structure in which only a small fraction of the nucleotides makes the G4 structure. The non-guanine nucleotides may non-specifically bind ThT and thus may have contributed to a high background fluorescence [326]. In addition, this G4 DNA might fold into an anti-parallel structure, which would also limit the amount of fluorescence increase (see discussion on its CD spectra later). When Li^+ was titrated, no fluorescence increase was observed and it even dropped slightly (Figure 5.2 B, black trace). When Na^+ was titrated, the drop in fluorescence was even more (Figure 5.2 B, red trace). Overall,

the control G4 experiment indicated that ThT can stain the G4 structure in our two-strand system (Figure 5.1 B), and only K^+ promoted formation of the G4 structure.

We then titrated the metal ions to the Ce13d DNAzyme containing the non-cleavable substrate (Figure 5.2 C). Interestingly, we observed decreased fluorescence intensity upon addition of Na^+ , while K^+ almost had no influence on the signal, similar to the response to Li^+ . We reason that Na^+ can fold the DNAzyme into a tight binding structure, releasing previously associated ThT to decrease its fluorescence. Such a binding structure was unlikely to be a G4 sequence due to the drastic fluorescence quenching by Na^+ and the lack of response to K^+ . To ensure that the data is representative, we also performed the metal titration in the presence of a lower buffer concentration (Figure 5.3). Still, Na^+ showed the largest ThT fluorescence decrease, confirming specific Na^+ -binding but likely to a non-G4 structure.

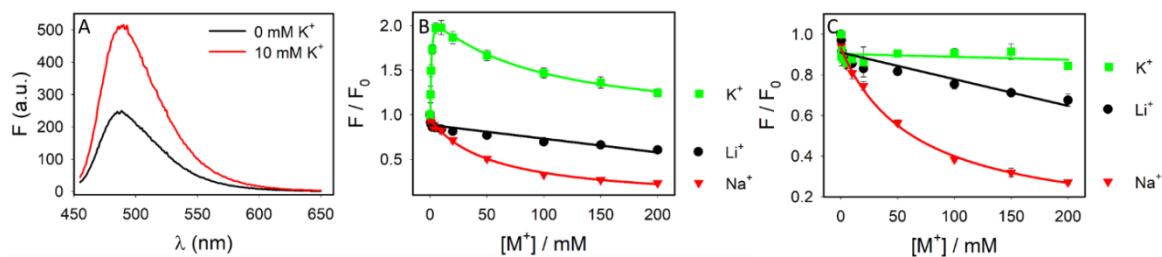


Figure 5.2: ThT fluorescence spectra of Ce13d and G4 control.

(A) Fluorescence spectra of the G4/ThT mixture without and with 10 mM K^+ . Fluorescence titration of the (B) G4/ThT, and (C) Ce13d/ThT mixture with various monovalent metal ions in 500 mM Tris-acetate buffer, pH 8.0.

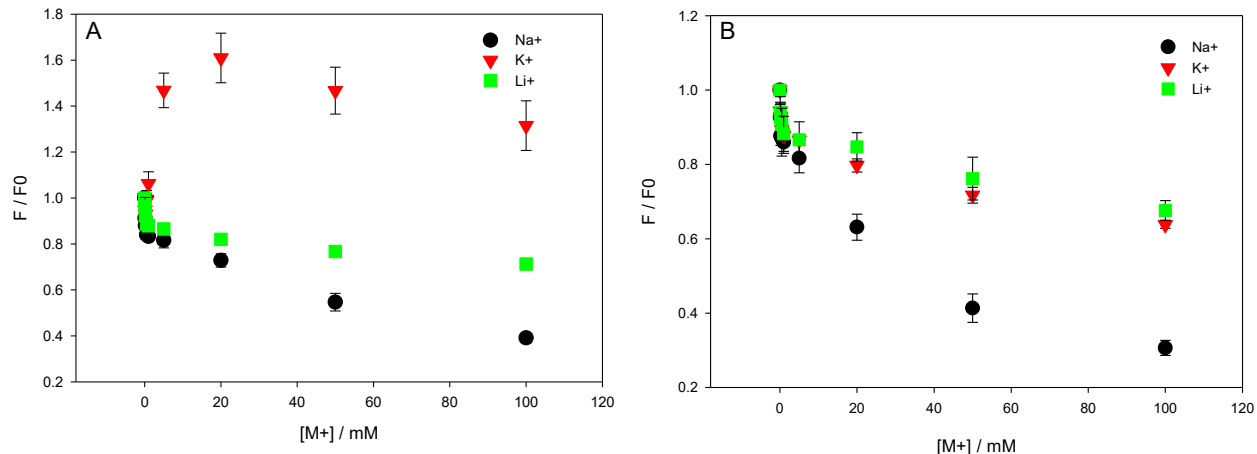


Figure 5.3: ThT fluorescence intensities in lower buffer concentration.

(A) G4/ThT, and (B) Ce13dA/ThT mixture with various monovalent metal ions in 50 mM Tris-acetate buffer, pH 8.0. The ratio of the G4/Ce13dA to ThT was 0.6 μ M : 3 μ M.

Another possibility is the formation of inter-molecular G4 complexes by multiple DNAzymes specifically interacting with each other. To test this, we varied the concentration of the Ce13d DNAzyme (keeping the ThT concentration the same). As we increased the concentration of DNAzyme, the initial fluorescence increased, which is consistent with formation of increasing DNA/ThT complexes. However, this response to Na⁺ was observed to be independent of DNAzyme concentration, upon plotting the relative fluorescence change (Figure 5.4). This data advocate that the effect of the Na⁺-binding is conferred upon individual DNAzyme molecules rather than the formation of inter-molecular complexes.

An important aspect of ThT staining to be considered is the possibility of G4-induced fluorescence reduction. It has been previously reported that using ThT to stain G4 DNA followed by addition of metal ions may not always accompany fluorescence increase, and sometimes fluorescence decrease may also be observed [319]. Based on the available literature, in most common cases with unmodified DNA we expect K⁺ to be better than Na⁺ to stabilize G4 structures, although exceptions were also reported [306,329]. The fact that only Na⁺ had a strong response of decreasing fluorescence with

negligible fluorescence perturbation in presence of K^+ (Figure 5.2 C) did not provide a strong support for a G4 structure in Ce13d with Na^+ . The insights from previous 2-aminopurine spectroscopy studies [298], in addition to the data fished out in our study herein, strengthen the notion of Ce13d DNAzyme to fold differently than G4 structures in presence of Na^+ . Due to the limitations of ThT as a G4 probe, the ThT data alone cannot conclude the structure of the Ce13d DNAzyme in the presence of Na^+ . Therefore, we then used spectroscopic methods that do not require labelling or staining of the DNA.

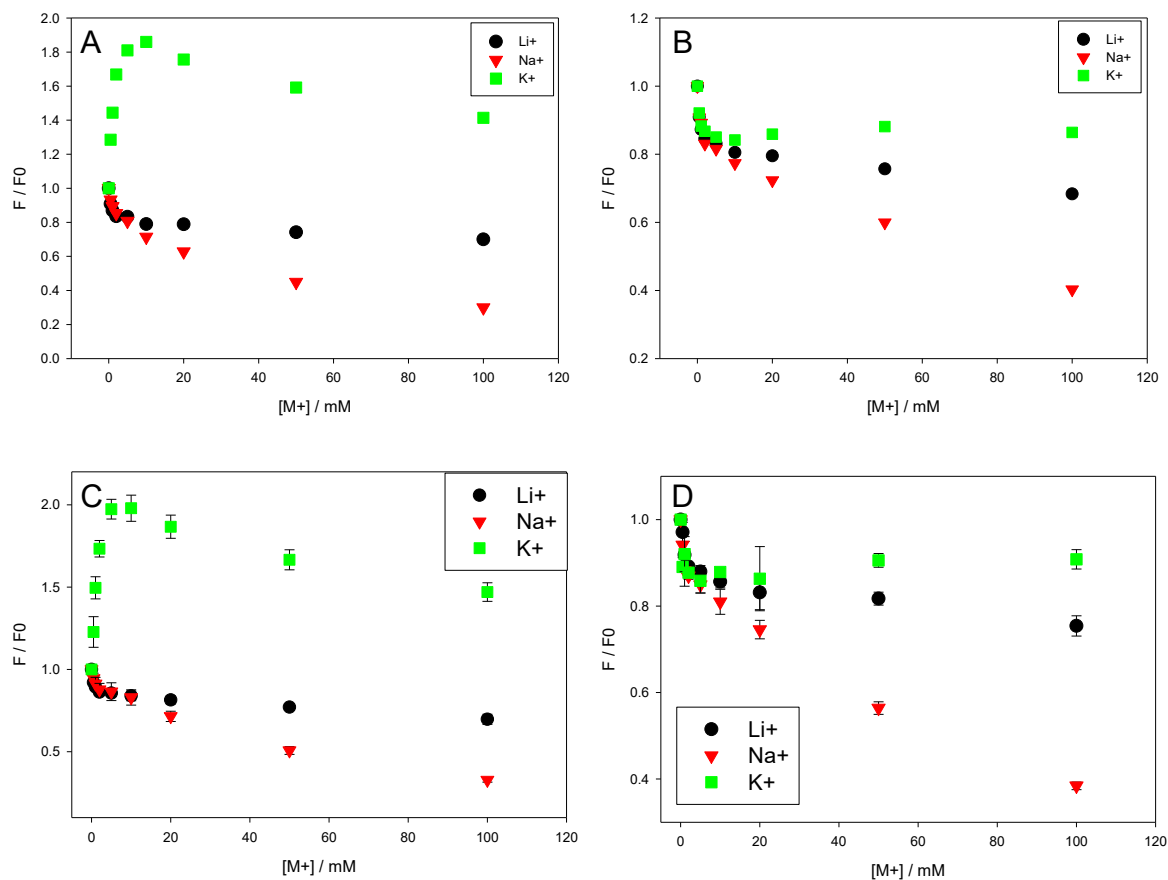


Figure 5.4: Fluorescence intensity at different DNA concentrations.

(A) G4/ThT and (B) Ce13dA/ThT mixture in the G4/Ce13dA : ThT ratio of 1.2 μ M : 3 μ M, with various monovalent metal ions. Fluorescence spectra of the (C) G4/ ThT and (D) Ce13dA/ ThT mixture in the G4/Ce13dA : ThT ratio of 0.6 μ M : 3 μ M, with various monovalent metal ions. This data is collected in 500 mM Tris-Acetate buffer pH 8.0.

5.5.3 Design of a cis-DNAzyme for NMR Spectroscopy

To further test the structure of the DNAzymes, we then performed NMR spectroscopy. One of the main bottlenecks in obtaining information from nucleic acid NMR is the length of the sequence under study. The chemical diversity of the nucleotide monomers (i.e. adenine, thymine/uracil, cytosine, and guanine) present in naturally occurring nucleic acids is very low. Due to this there is high spectral overlap in their NMR peaks [227]. This problem becomes more and more significant as the number of nucleic acid polymers or the number of nucleotides increase [330]. The DNAzyme version used for ThT experiments (Figure 5.1 A) contains two separate strands, and the full Ce13d DNAzyme used for previous studies had nearly 90 nucleotides. It is difficult to prepare a homogenous NMR sample with the two-strand system since it is hard to control the presence of any unhybridized strand by having exactly the same ratio of the two strands. Such heterogeneity adds spectral overlap of NMR peaks as well, making NMR analysis even more difficult. Therefore, to lessen the probability of spectral overlap, short cis versions of Ce13d were designed for NMR studies.

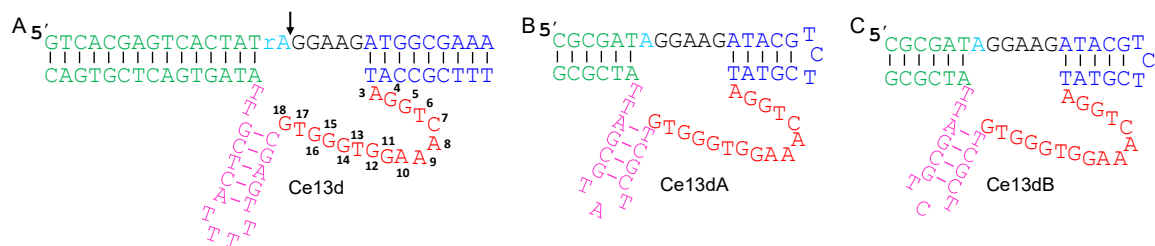


Figure 5.5: Ce13d sequence variations used for NMR spectroscopy.

(A) trans-cleaving DNAzyme Ce13d with the conserved nucleotides (red) numbered from 3-18, and its non-cleavable analogues (B) Ce13dA, and (C) Ce13dB.

The secondary structure of the trans-cleaving Ce13d previously used for biochemical characterizations [323], and the two short cis versions: Ce13dA and Ce13dB used for NMR studies are shown in Figure 3. The two substrate-binding arms of these two cis DNAzymes are 6 base pair (shown in green color) and 5 base pairs (shown in blue color) long, much shorter than those in Ce13d. Previous studies showed that the hairpin size and composition can be changed as long as a hairpin structure is retained [291,295,298]. In the catalytic loop, the length of the hairpin was also shortened. The only difference between Ce13dA and Ce13dB is that the adenine in the tip of the hairpin loop was changed to a cytosine. Shortened cis-DNAzymes were used to solve the DNA length and substrate/enzyme ratio problems. The region shown in dark red is the same for all three versions shown. These conserved nucleotides present in the enzyme loop of Ce13d are most important for Na⁺-binding as well as catalytic activity (nucleotides numbered 3-18 in Figure 5.5 A). A systematic mutation study of the conserved enzyme loop, in which each nucleotide was mutated to the other three has revealed interesting insights [295]. It was found that most of the mutants except for A3G, A8G, G14A, and G14T, were incapable of specific Na⁺-binding. In terms of catalytic activity, the nucleotides A3, G14 and T17 exhibited tolerance to mutations, and mutants C7A, A8G, and T13C were found active. Except for these, all the other mutants remarkably hampered the Ce13d catalysis. These data present a good correlation between Na⁺-binding and catalytic activity, showing that Na⁺-binding is a key factor for catalysis to take place. These data also validate the usage of Ce13dA and Ce13dB for NMR, as these have the conserved set of nucleotides preserved. In the trans-cleaving Ce13d DNAzyme (Figure 5.5 A), the cleavage site is denoted with a black arrow, and the cleavage site ribonucleotide 'rA' is colored in cyan. The cis-versions of Ce13d are designed to be non-cleavable by replacing the cleavage site 'rA' to deoxy-ribonucleotide 'A' (colored in cyan in Figure 5.5 B and 5.5 C).

5.5.4. Folding of Ce13d in Li⁺, Na⁺ and K⁺.

To see if we could gain a deeper understanding of the folding of Ce13d in the presence of various monovalent ions we probed the 1D ¹H spectrum of Ce13dA in the presence of various monovalent ions at two different pH values (Figures 5.6 and 5.7). The imino proton regions of 90% H₂O/10% D₂O 1D ¹H NMR spectra of Ce13dA were collected with no salt added (only trace amounts of Li⁺ present) at pH 6.8 (Figure 5.6 A), 10 mM K⁺ at pH 6.8 (Figure 5.6 B), and 10 mM Na⁺ at pH 6.8 (Figure 5.6 C) respectively. The imino region of the 1D ¹H spectrum contains peaks for the exchangeable imino protons of guanine (H1) and thymine (H3) [331]. More specifically, the region of 12-14 ppm represents signals from imino (NH) protons which are strongly hydrogen bonded in Watson-Crick base pairs, while the signals in the region around 9-12 ppm belong to imino protons that are typically involved in non-canonical base pairs which are useful for characterizing the secondary structures formed by complexed DNA [332,333]. A comparison of the three spectra in the region of 12-14 ppm in Figure 4 suggests that there are similar number of peaks and several shared chemical shifts between each of the three spectra, suggesting that the structure of the base paired regions shown in Figure 3B was relatively rigid and stable in the presence of traces of Li⁺, or 10 mM Na⁺, and K⁺ at pH 6.8. However, the region of 9-12 ppm is quite different with respect to the number of peaks and chemical shifts of the peaks for each of the three spectra, indicating that Ce13dA adopted a different conformation and/or equilibrium of conformations in the presence of no salt (Li⁺ traces), Na⁺ and K⁺.

In Figure 5.7, similar spectra were acquired but at a lower pH of 5.8 and with higher salt concentrations of 80 mM K⁺ (Figure 5B) and 80 mM Na⁺ (Figure 5.7 C) to drive the binding of the cations. Under these conditions, the spectra for K⁺, and to a lesser extent, Li⁺, had much broader linewidths and more spectral overlap, resulting in poorly defined peaks. This is indicative of the presence of multiple conformations, which is unsurprising at a lower pH where exchange occurs more readily due to higher H⁺ concentration. On the other hand, it can be observed in Figure 5.7 C that there are shifts in the Na⁺

spectrum from higher salt concentration and lower pH, but in general it retains its structured conformation. From Figure 5.6, it is evident that many of the peaks affected in presence of Na^+ are different from those affected with K^+ , and at the lower pH of 5.8 where the exchange rate is higher, Ce13dA visibly retains much more structure in the presence of Na^+ than the free DNA or in the presence of K^+ . This emphasizes that Ce13dA adopts a different conformation and/or equilibrium of conformations in the presence of Na^+ versus K^+ . It is also worth noting that in Figure 5.6, there are fewer peaks with a narrower distribution present in the absence of salt than there are in the presence of Na^+ or K^+ which implies that some features of the folded structures are unable to form without cation stabilization. Most of these differences are observed in the lower ppm range of the imino region, which is characteristic of non-Watson Crick base pairing. These interpretations support the conclusions of previous results, where using intrinsic fluorescence changes of 2-aminopurine labelled at the cleavage site, it was shown that the folding pattern with Na^+ -binding was completely different from K^+ -binding, where K^+ is considered to induce misfolding of Ce13d [298,300].

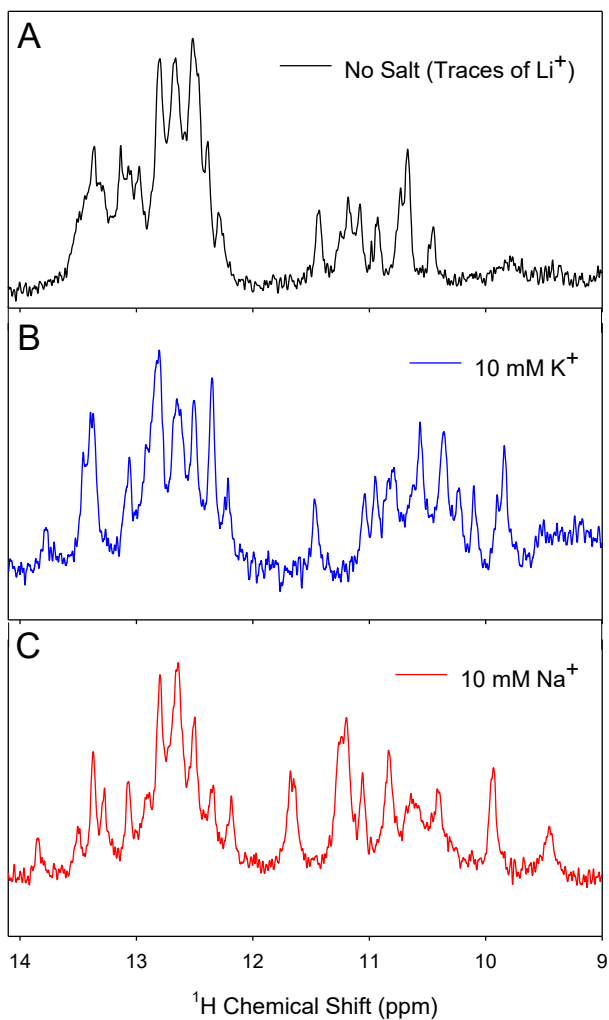


Figure 5.6: Imino proton region of 90 % H_2O / 10 % D_2O 1D ^1H NMR spectra of Ce13dA at 277K.

(A) 600 μM Ce13dA with no salt added (only trace amounts of Li^+ from purification present), pH 6.8, (B)

150 μM Ce13dA in 10mM K^+ , pH 6.8, (C) 150 μM Ce13dA in 10mM Na^+ , pH 6.8.

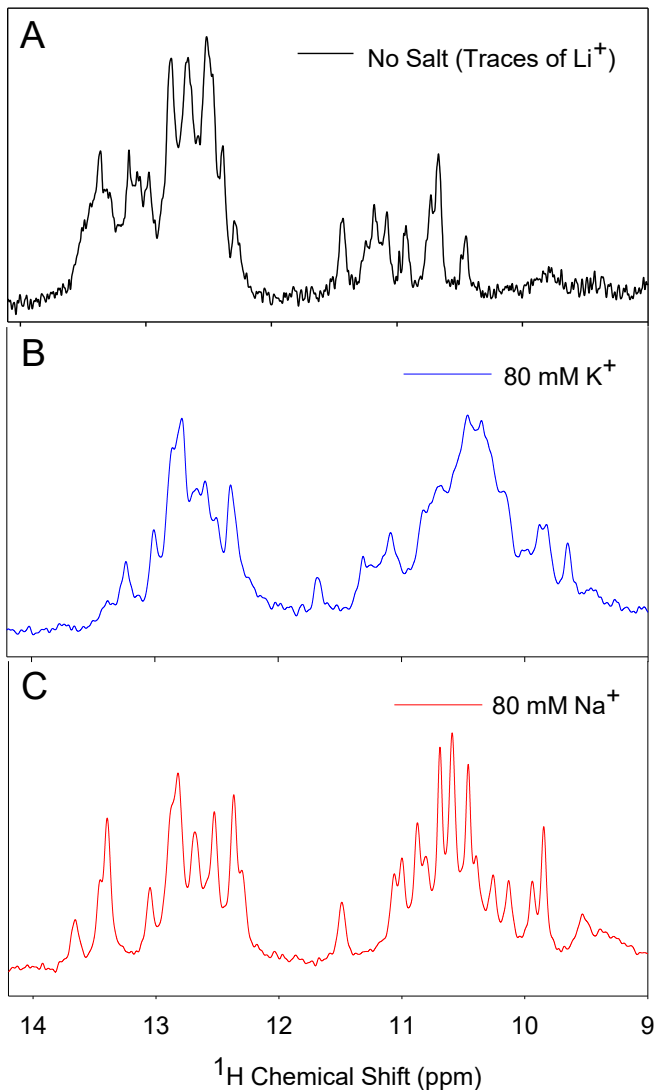


Figure 5.7: Imino proton region of 90 % H₂O / 10 % D₂O 1D ¹H NMR spectra of Ce13dA at 277K.

(A) 150 μM Ce13dA with no salt added (only trace amounts of Li⁺ from purification present), pH 5.8, (B) 150 μM Ce13dA in 80mM K⁺, pH 5.8, (C) 150μM Ce13dA in 80mM Na⁺, pH 5.8.

5.5.5 NMR Spectra Suggest the Na⁺-binding Structure is not a G-quadruplex

Many G-rich DNA aptamers contain G-quadruplex structures for molecular recognition, and these structures have fairly well defined guanine imino ¹H NMR shifts between 10.5-12.5 ppm [198,334–336]. G-quadruplex DNA is a highly stable structure and therefore these peaks are typically defined by

high intensity and narrow linewidth. Due to the Na^+ dependence of the Ce13d DNAzyme and its sequence containing sufficient G-rich regions, NMR was also used to qualitatively assess the presence of G-quadruplex DNA. This needed the investigation of Ce13d in presence of Na^+ due to its functional role, and also in presence of K^+ because of the well-established preference of G-tetrads for K^+ [337,338]. The spectra in Figures 4 and 5 were analyzed for this purpose. However, no compelling evidence supported the presence of a G-quadruplex in Ce13dA in the presence of Na^+ or K^+ . There are some peaks between 10.5 ppm and 12.5 ppm at both pH ranges, but this is not atypical of DNA, and based on the linewidths, G-quadruplex is not conclusively present in any of the spectra. At pH 5.8, it is highly likely that a G-quadruplex would be stable and retain its characteristic, narrow imino peaks between 10.5-12.5 ppm and it is clear that this is not the case for free DNA or in the presence of K^+ . In the presence of Na^+ , the peaks in this region have narrower linewidths but this simply indicates that conformational stability is conferred but not necessarily due to a G-quadruplex. Analysis of Figure 5.6 shows that there are not significantly more peaks in the presence of Na^+ than K^+ . Based on these observations and the fact K^+ is known to have a higher propensity for G-quadruplex formation than Na^+ , it is unlikely that Ce13dA forms a G-quadruplex. However, despite the lack of evidence, it is possible that different G-quadruplexes are formed in presence of different cations and it cannot be explicitly ruled out by these spectra alone that G-quadruplex is formed. Therefore, we sought additional evidence to support this claim. In order to collect this evidence, we lyophilized the three samples from Figure 5.7 after the previous spectra were obtained and resuspended them in 100% D_2O for acquisition of further 1D ^1H NMR spectra. Under these conditions, signals from exchangeable imino and amino resonances from G-quadruplex G residues may survive for up to two or more weeks in D_2O [339]. There were no residual imino or amino peaks observed in any of these spectra, despite being observed within 30 minutes of resuspension. These findings, in combination with those from spectra in Figures 5.6 and 5.7, support the conclusion that there is no G-quadruplex formation in the presence of Na^+ or K^+ . To summarize, these ^1H NMR spectra

show that Ce13d is unique in its ability to discriminate between monovalent cations and supports the idea that there is an aptamer specific for Na^+ within the catalytic loop of Ce13d. This NMR data also indicates that this aptamer is not based on a G-quadruplex structure.

5.5.6 CD Spectra Confirm the Absence of G-quadruplex Structure.

CD spectra were then obtained for Ce13dA under the same three salt conditions used for NMR experiments (no salt added, 80 mM Na^+ and 80 mM K^+). We chose the cis-cleaving Ce13dA to better match the results of the NMR experiments. All three spectra had maxima at approximately 280 nm, minima at 250 nm and a cross-over point from positive to negative intensity around 260 nm which is typical of duplex DNA (Figure 5.8) [196,340]. G-quadruplex DNA can have different forms, all with characteristic CD signatures, such as parallel (~264 nm max, 245 nm min), antiparallel (~ 295 max, 260 min) or hybrid (~ 295 max, 260 max, 245 min) [341,342]. These peaks are clearly not present in any of the CD spectra obtained. In addition to this, all three salt conditions give nearly identical CD spectra, which is not consistent with the presence of a G-quadruplex. Since G-quadruplex formation is dependent on salt, a sequence containing G-quadruplex would experience significant shifts in wavelengths and increases in peak magnitudes in the presence of K^+ compared to the absence of K^+ [343]. We previously measured the CD spectra of the trans-cleaving Ce13d DNAzyme and also the G-quadruplex control shown in Figure 5.1 A and 5.1 B, respectively [299]. The trans-cleaving Ce13d spectra were very similar to that of the cis-cleaving Ce13dA presented in Figure 5.8, suggesting that they had a similar overall folding. The G4 control, on the other hand, had the peaks shifted to 290 nm and 250 nm in the presence of K^+ , suggesting its folding into an anti-parallel G-quadruplex. The peaks did not perfectly match with the ideal values since a portion of the DNA was in duplex. This evidence indicates that Ce13dA does not form a G-quadruplex, in agreement with 1D NMR data.

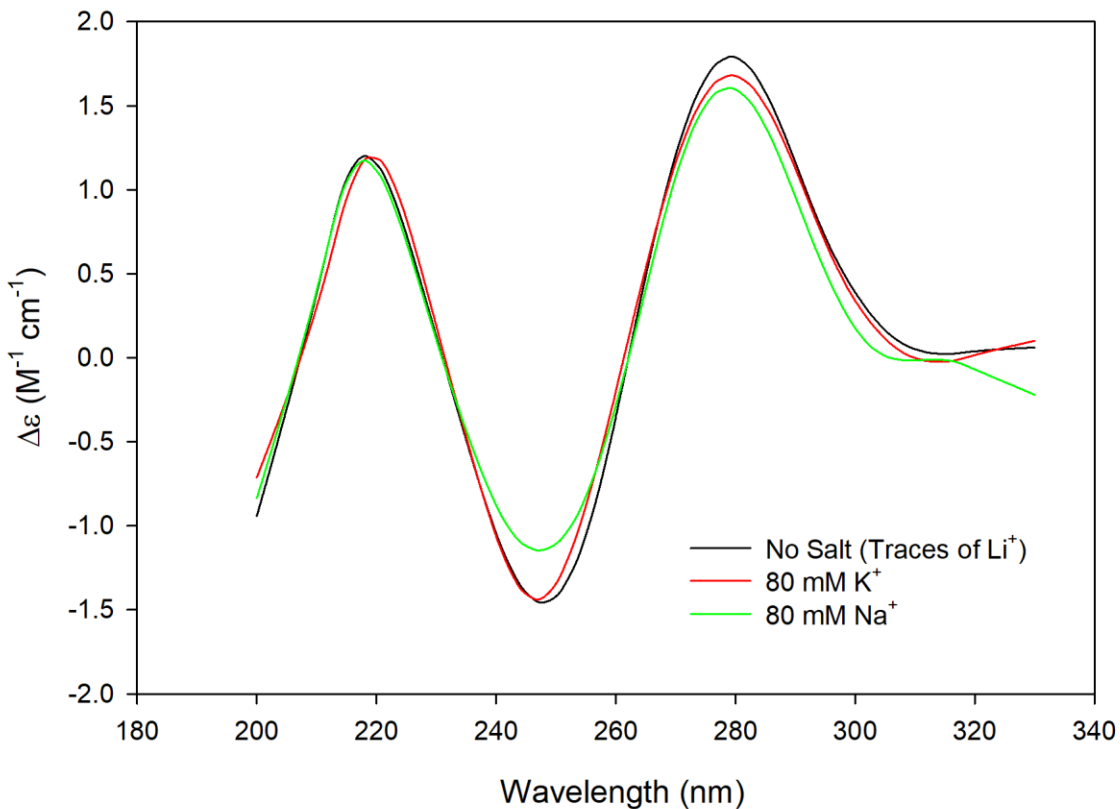


Figure 5.8: CD spectra of Ce13dA in the presence of no added salt, K^+ and Na^+ .

No significant changes are observed in CD spectra when different ions are present and characteristic G-quadruplex peaks are absent, indicating a lack of G-quadruplex structure in this sequence.

5.5.7 Potential Structural Information from 2D NMR

In addition to the 1D 1H NMR, we probed the structure of Ce13d with 2D NMR. For this, we used the Ce13dB construct. The Ce13dB differs from Ce13dA by a cytosine residue in its hairpin-loop (shown in pink in Figure 5.5 B and 5.5 C). This change could be afforded as this position is known to be insignificant in Na^+ binding and catalysis of Ce13d [291,323]. This was done to increase the number of cytosine residues as it proves beneficial for spectral assignment of peaks, and therefore in determining the homogeneity of the sample. Typically for cytosines, the H5 and H6 protons show up peaks between 5-6 ppm and 6.9-7.9 ppm, respectively. The through-bond interaction between H5 and H6 protons is

unique to cytosines, and the number of peaks coming from this interaction directly correlates to the number of cytosines in the structure. To determine if Ce13dB is present in a single homogeneous conformation, we probed the structure of Ce13dB with a 2D TOCSY experiment (Figure 5.9) and looked at the peaks generated by the through-bond interactions of H5/H6 protons in the cytosine nucleotides (Figure 5.10). The number of cytosines in Ce13dB is 12 (Figure 5.5 C), while the number of peaks showing up in the 100 % D₂O ¹H5/¹H6 2D TOCSY is 18 (Figure 5.9). This clearly indicated that Ce13dB is present in multiple three-dimensional conformations. Since conformational homogeneity is a prerequisite for structure determination through NMR, any further spectra for structure determination were not acquired in this study.

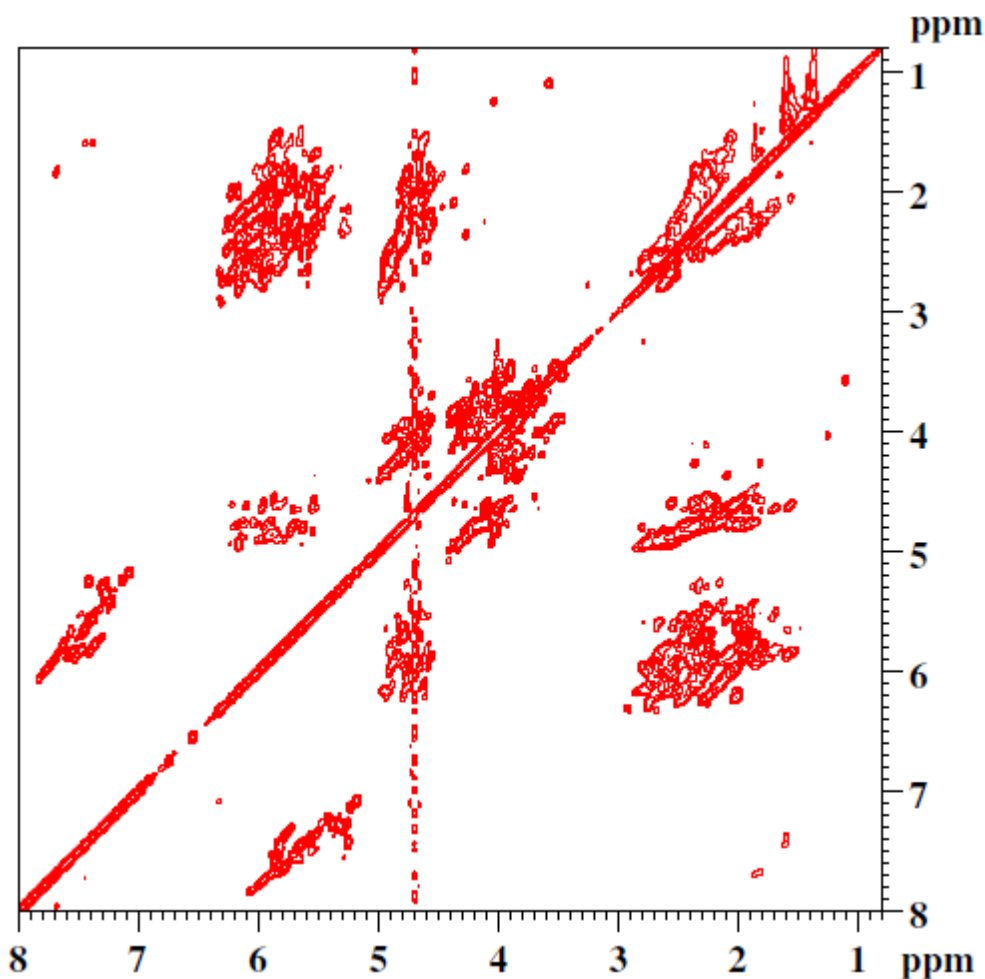


Figure 5.9: 2D-TOCSY spectrum of 450 μ M Ce13dB in 5 mM LiPO₄ pH 6.8, 200 mM Na⁺.

This is the full TOCSY spectrum of Ce13dB in the presence of high Na⁺. Spectrum was acquired with a mixing time of 50 ms.

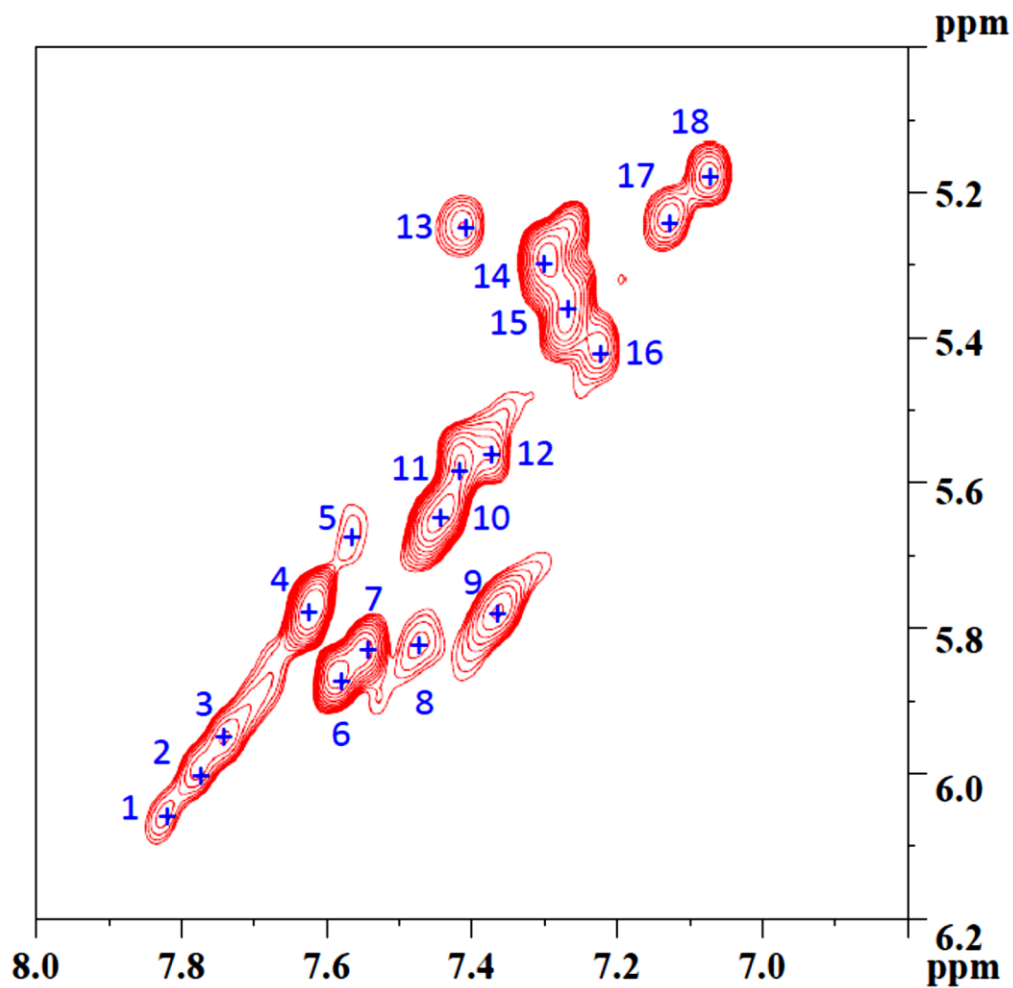


Figure 5.10: H5/H6 proton region of the 2D-TOCSY spectrum of Ce13dB.

The spectrum shows cross-peaks generated by through-bond interactions of H5/H6 hydrogens of the cytosine residues of Ce13dB. These peaks are conservatively counted in blue text. The chemical shift range of H5 protons is 5-6.2 ppm (y-axis) and the chemical shift range of H6 protons is 6.8-8.0 ppm (x-axis). This spectrum was acquired with a 450 μ M sample in 5 mM LiPO₄ pH 6.0, 200 mM Na⁺.

5.5.8 Structural Effects of Lanthanum Titration on Cis13dB

It has been noted in previous work on Cis13d that cleavage activity is catalyzed by several trivalent lanthanides [291]. There was a distinct lack of structural homogeneity observed in TOCSY experiments with the non-cleaving Cis13dB, even in the presence of sodium. Because of Cis13d's ability to interact with lanthanides at the cleavage site, we decided to probe Cis13dB with lanthanum to see if any structure-stabilizing effects were observed in its presence. To do this, we acquired TOCSY experiments after the addition of several aliquots of lanthanum, as shown in Figure 5.11. A clear change in conformation and/or equilibrium of conformations was observed over the course of this titration. After the addition of the final aliquot (4:1 lanthanum to DNA), shown in red in Figure 5.11, four fewer peaks are counted than at the beginning of the titration (no lanthanum, 80mM NaCl), shown in blue in Figure 5.11. This is still more than expected for a single conformation but is an obvious improvement over the spectra obtained with sodium alone. An additional aliquot (6:1 lanthanum to DNA) was obtained, but little difference was observed compared to the 4:1 aliquot so the titration was concluded at that point. Though the TOCSY titration suggests that multiple conformations are still present, NOESY experiments were performed to probe the tertiary structure and confirm the supposition of the TOCSYs.

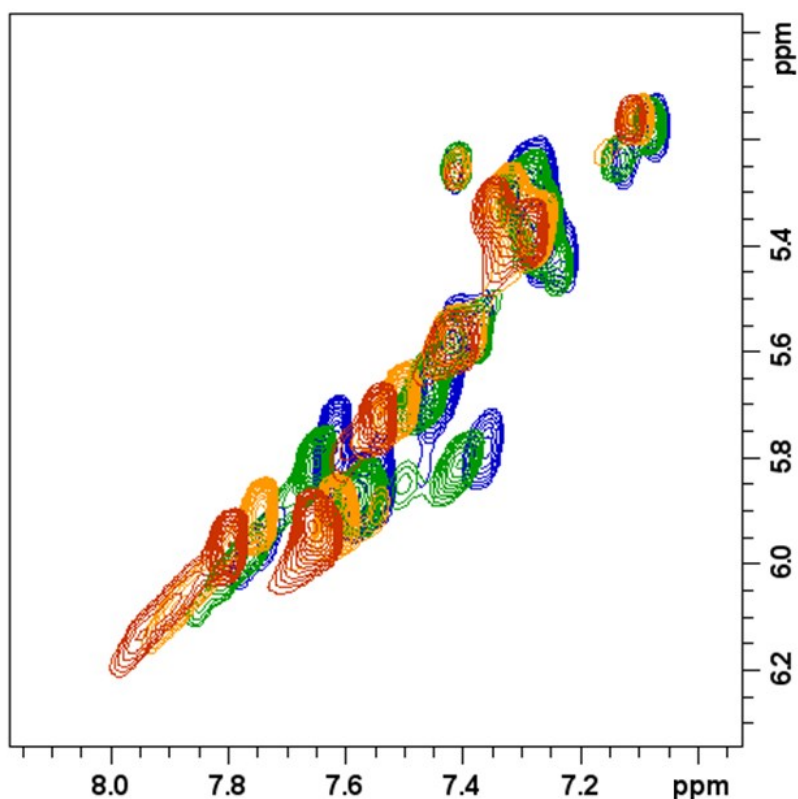


Figure 5.11: Titration of Cis13dB with lanthanide.

Blue- 600 μM DNA, 80 mM NaCl, Green- 600 μM DNA, 80 mM NaCl, 600 μM $\text{La}(\text{NO}_3)_3$, Orange- 600 μM DNA, 80 mM NaCl, 1200 μM $\text{La}(\text{NO}_3)_3$, Red- 600 μM DNA, 80 mM NaCl, 2400 μM $\text{La}(\text{NO}_3)_3$. All spectra were acquired with mixing times of 50 ms.

A NOESY in 90% $\text{H}_2\text{O}/10\%$ D_2O (Figure 5.12) and a NOESY in D_2O (Figure 5.13) were acquired on the same sample used for the TOCSY experiments in Figure 5.11 (6:1 lanthanum to DNA and 80mM NaCl). Though spectra of reasonable resolution were acquired, the significant lack of NOEs is evident, particularly when comparing to high-quality spectra of a sequence of similar length (Figure 3.8). This is indicative of a lack of tertiary structure forming. At this point, no further spectra were acquired as it became clear that assignments were impossible.

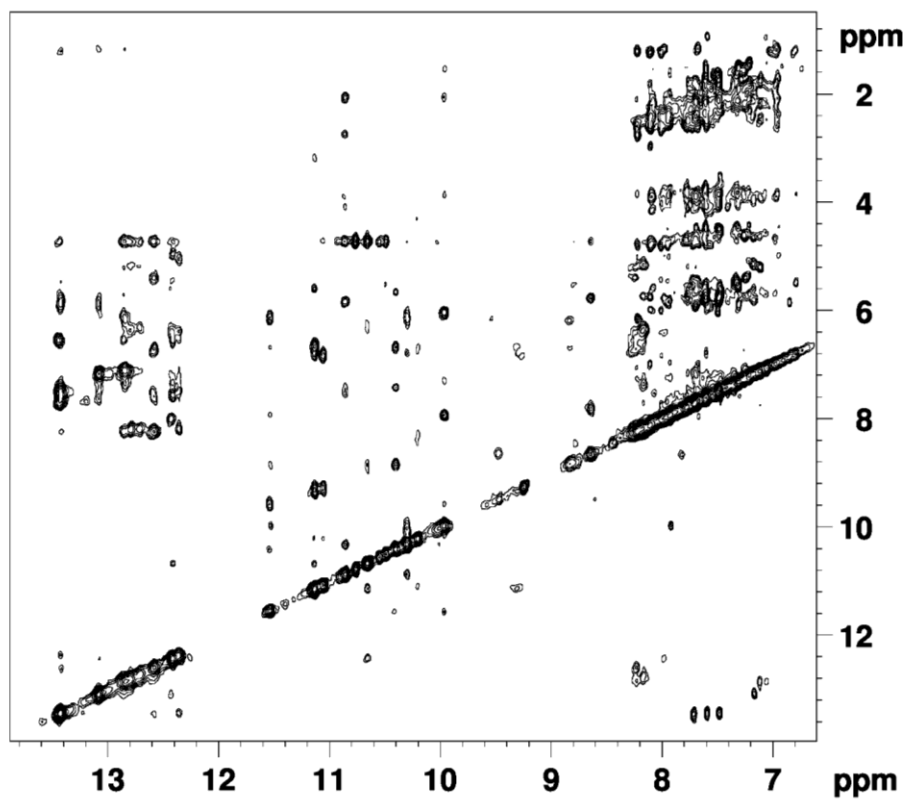


Figure 5.12: NOESY of Cis13dB in 90% H₂O/10% D₂O.

Sample contained 600 μM DNA, 80 mM NaCl, 3600 μM La(NO₃)₃ and was run with a 150 ms mixing time.

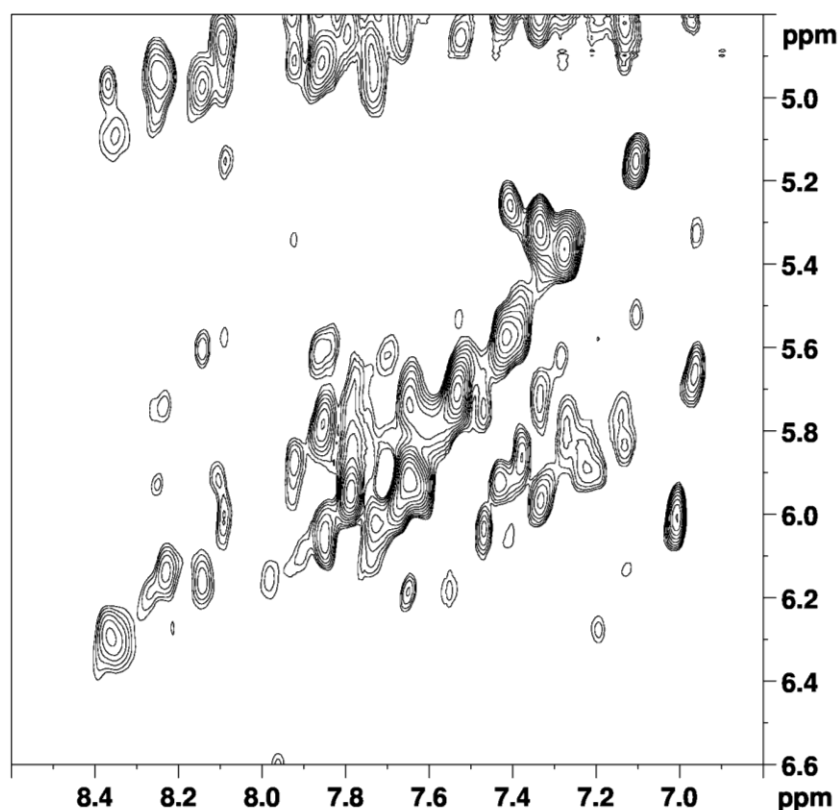


Figure 5.13: H1'-H6/H8 region in the NOESY spectrum of Ce13dA in D₂O.

Sample contained 600 μM DNA, 80 mM NaCl, 3600 μM $\text{La}(\text{NO}_3)_3$ and was run with a 150 ms mixing time.

5.6 Conclusions

In this study, ThT staining, NMR spectroscopy and CD spectroscopy were employed to study Na^+ binding by its aptamer, which is embedded in the Ce13d and NaA43 DNAzymes. By accomplishing comparative analysis between Ce13d Na^+ -aptamer versus a G4 construct, it was observed that both show a distinct fluorescence change in the presence of Li^+ , Na^+ and K^+ . In the case of Ce13d, while most of the binding was observed with Na^+ , no evidence supported that formation of a G4 structure makes the basis of Na^+ -binding, and thus this aptamer likely uses other mechanisms to bind Na^+ . NMR provided a similar conclusion arguing against a G4 structure in the presence of Na^+ . This is further supported by

the lack of G4 observed in CD. This report not only explicitly demonstrates the presence of a uniquely folding novel Na⁺-aptamer in Ce13d, but also substantiates the fact that isolation of novel aptamer containing DNAzymes or Aptazymes is a prudent way of discovering novel distinctly folding metal-binding aptamers. Additionally, this study highlights the possibility of utilizing monovalent metal ions to play novel and unique roles in DNA scaffolding and DNA nanotechnology in general, other than just nucleic acid duplex stabilization.

Chapter 6: Summary and Future Work

In Chapter 2, we screened and investigated SRB-2 ligands for optimal binding properties, namely affinity and selectivity. We determined that the amine groups on the ligand must have small alkyl constituents that are able to participate in hydrophobic interactions. They must not, however, be so large as to interfere with the aptamer's binding conformation. Binding affinity is also increased when the total number of negative charges is decreased. A single negatively charged group was shown to have little impact on binding affinity but is required for selective binding of SRB-2. This is due to repulsive interactions with the RNA backbone preventing these ligands from non-specifically binding to A-form regions of the aptamer. As observed with AR101-B, bulky amine constituents can also confer selective binding of SRB-2 in the absence of a negative charge but at a considerable loss of binding affinity. Considering all these observations, we determined rhodamine B was the best ligand out of those we tested. We also showed via ITC that SRB-2 binding is dependent on divalent cations at low concentration and determined using CD that its binding mechanism does not involve a G-quadruplex and does not form any new helical regions upon binding.

In Chapter 3, the solution structure of SRB-2 was probed using homonuclear NMR techniques as well as with the three single-nucleotide labelled samples [^{13}C ^{15}N]-Ade, [^{13}C ^{15}N]-Cyt and [^{13}C ^{15}N]-Ura. The presence of structural heterogeneity ultimately resulted in insufficient quality of data to pursue structure calculations. There were some minor changes observed with low Mg^{2+} concentrations so investigations with more salt concentrations may be useful. Initial spectra obtained using RB as the ligand were not successful. However, sequence modification of SRB-2 to address 3' and 5' heterogeneity may be more successful. If such changes were impactful, one could then consider re-acquiring some of the NMR experiments performed and perhaps synthesize additional samples that may be of use, such as a uniform ^{13}C ^{15}N sample. Another potentially useful strategy to address spectral overlap would be to

use at least one partially deuterated sample, likely with one or more nucleotides deuterated in the H3', H4', H5' and H5'' positions. Removing some of these spins would drastically reduce the amount of dipolar relaxation experienced by the other protons, resulting in a significantly sharpened spectrum that may reduce some of the remaining peak overlap.

In Chapter 4, we pursued a “divide and conquer” strategy to acquire more structural information about SRB-2. Resonance assignments for a UUCG 16mer were acquired using homonuclear spectra, and a family of 20 structures was calculated. These structures were in agreement with literature structures, and characteristic cross-peaks from the tetraloop could be easily identified in the NOESY spectra of SRB-2. NMR studies with the L3C sequence were unsuccessful due to the high degree of structural heterogeneity present. Using this truncation clearly did not solve the overlap or heterogeneity problems discussed in Chapter 3. Based on the NMR spectra for L3C, it is likely that there is significant interaction between loop 3 and loop 2 in the folded structure SRB-2. Therefore, this “divide and conquer” approach will not be useful beyond the UUCG tetraloop. A minimal version of SRB-2 aptamer, SRB-2 min, was proposed by the authors who originally selected SRB-2. The UUCG tetraloop is absent in this structure, while loop 2 and loop 3 are intact. However, NMR studies are not possible due to the extremely low binding affinity observed in fluorescence experiments. If future NMR studies are done on this system, finding conditions to limit exchange and hence peak broadening would be the primary goal. Then, the isotope labelling strategies discussed in Chapter 3 may be more effective.

In Chapter 5, ThT staining, NMR spectroscopy and CD spectroscopy were used to study Na⁺ binding by its aptamer, which is embedded in the Ce13d and NaA43 DNAzymes. By accomplishing comparative analysis between Ce13d Na⁺-aptamer versus a G4 construct, it was observed that both show a distinct fluorescence change in the presence of Li⁺, Na⁺ and K⁺. In the case of Ce13d, most of the binding was observed with Na⁺, but no evidence supported that formation of a G4 structure makes the basis of Na⁺-binding. NMR and CD provided similar conclusions regarding the absence of G4 structure in

the presence of Na^+ . However, based on the metal selectivity observed in this system, we were able to demonstrate the presence of a uniquely folding novel Na^+ -aptamer in Ce13d, but also suggests that isolation of novel aptamer containing DNAzymes or Aptazymes may be a practical way of discovering novel metal-binding aptamers. Additionally, this work highlighted the possibility of utilizing monovalent metal ions to serve novel and unique roles in DNA scaffolding and DNA nanotechnology in general.

To conclude this thesis, I would like to briefly state the most significant impacts of this research. The binding kinetics investigated in Chapter 2 were able to show, with respect to SRB-2, that a ligand that was not the selection target, rhodamine B, actually had the most ideal binding properties of those studied, including a higher binding affinity than the original ligand, sulforhodamine B. Advances in aptamer selection methodologies have come a long way since SRB-2 was selected in 1997, so perhaps this indicates more optimal sequences could exist for some of these ligands. In literature, several SRB-2 ligands showed promise due to their high fluorescence enhancement in previous studies, but we showed here that they were in fact binding to the aptamer non-specifically, questioning the usefulness of these ligands in potential applications. In Chapter 5, we studied a DNAzyme containing a Na^+ aptamer that clearly discriminates between Na^+ and K^+ without the use of a G-quadruplex, as would generally be expected. This existence of a selective Na^+ aptamer is also encouraging with respect to biosensor development for metal ions. Distinguishing ions as similar in properties to Li^+ , Na^+ and K^+ is no small feat and indicates that DNAzymes and/or aptazymes may be a useful method for sensing unique metal ions.

Copyright Permissions

Biochemical and Biophysical Research Communications (Elsevier)

[About Elsevier](#) [Products & Solutions](#) [Services](#) [Shop & Discover](#) [Search](#)

[Overview](#) [Author rights](#) [Institution rights](#) [Government rights](#) [Find out more](#)

Overview

In order for Elsevier to publish and disseminate research articles, we need certain publishing rights from authors, which are determined by a publishing agreement between the author and Elsevier.

For articles published open access, the authors license exclusive rights in their article to Elsevier.

For articles published under the subscription model, the authors transfer copyright to Elsevier.

Regardless of whether they choose to publish open access or subscription with Elsevier, authors have many of the same rights under our publishing agreement, which support their need to share, disseminate and maximize the impact of their research.

For open access articles, authors will also have additional rights, depending on the Creative Commons end user license that they select. This Creative Commons license sets out the rights that readers (as well as the authors) have to re-use and share the article: please see [here](#) for more information on how articles can be re-used and shared under these licenses.

This page aims to summarise authors' rights when publishing with Elsevier; these are explained in more detail in the [↓ publishing agreement](#) between the author and Elsevier.

Irrespective of how an article is published, Elsevier is committed to protect and defend authors' works and their reputation. We take allegations of infringement, plagiarism, ethical disputes, and fraud very seriously.

Author rights




The below table explains the rights that authors have when they publish with Elsevier, for authors who choose to publish either open access or subscription. These apply to the corresponding author and all co-authors.

Author rights

The below table explains the rights that authors have when they publish with Elsevier, for authors who choose to

Overview	Author rights	Institution rights	Government rights	Find out more
Author rights in Elsevier's proprietary journals		Published open access	Published subscription	
	Retain patent and trademark rights	✓	✓	
	Retain the rights to use their research data freely without any restriction	✓	✓	
	Receive proper attribution and credit for their published work	✓	✓	
	Re-use their own material in new works without permission or payment (with full acknowledgement of the original article): 1. Extend an article to book length 2. Include an article in a subsequent compilation of their own work 3. Re-use portions, excerpts, and their own figures or tables in other works.	✓	✓	
	Use and share their works for scholarly purposes (with full acknowledgement of the original article): 1. In their own classroom teaching. Electronic and physical distribution of copies is permitted 2. If an author is speaking at a conference, they can present the article and distribute copies to the attendees 3. Distribute the article, including by email, to their students and to research colleagues who they know for their personal use 4. Share and publicize the article via Share Links, which offers 50 days' free access for anyone, without signup or registration 5. Include in a thesis or dissertation (provided this is not published commercially) 6. Share copies of their article privately as part of an invitation-only work group on commercial sites with which the publisher has a hosting agreement	✓	✓	
	Publicly share the preprint on any website or repository at any time.	✓	✓	
	Publicly share the accepted manuscript on non-commercial sites	✓	✓ using a CC BY-NC-ND license and usually only after an embargo period (see Sharing Policy for more information)	
	Publicly share the final published article	✓ In line with the author's choice of end user license	x	
	Retain copyright	✓	x	





Publishing Policy

Cette politique est également disponible [en français](#).

Canadian Science Publishing (CSP) is a not-for-profit federally incorporated publishing company that owns and publishes a suite of scientific journals.

At CSP our mission is to publish high-quality, international scientific journals that provide researchers in Canada and around the world with a means to communicate their findings. We are champions of scientific knowledge exchange, committed to strengthening the integrity, relevance, and reach of science. We ensure that scientific knowledge is easy to discover, use, and share.

MORE ABOUT CSP

- Leadership
- Policies
- Advertising
- Community
- Partners
- Careers
- News
- Contact

Preprints and postprints

To help extend the reach of science, we support the sharing of new research findings before and after publication. Before publication, authors may:

- post an un-refereed, preliminary version of the manuscript to a preprint server, institutional repository, personal website, or funder archive;
- present the work at a conference as an abstract, poster, lecture, or paper, so long as the manuscript submitted to the journal presents substantial additional content; and
- submit their thesis or dissertation (in full or in part) as a manuscript, and may include their published manuscript (in full or in part) in their thesis or dissertation.

Upon submission, authors should:

- disclose whether the work was presented at a conference and, if so, include conference details;
- disclose the availability of any published, preliminary work in the covering letter;
- confirm that they either retain copyright of any published conference material or acquired permission from the copyright holder, in the covering letter; and
- cite the preliminary version in the manuscript.

We will not consider manuscripts that have been previously published, or are under consideration for publication, in another journal.

After publication, authors retain the right to:

- post the submitted or accepted version of the manuscript to a preprint server, institutional repository, personal website, or funder archive, providing a link back to the published version of record once available;
- reproduce the final published article for the purpose of teaching; and
- ➔ • reuse the final published article (in full or in part) in other works created by them for noncommercial purposes, provided the published version of record is acknowledged and cited.

References

- [1] J.D. Watson, F.H.C. Crick, Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid, *Nature*. 171 (1953) 737–738. <https://doi.org/10.1038/171737a0>.
- [2] R.E. Franklin, R.G. Gosling, Evidence for 2-chain Helix in crystalline structure of sodium deoxyribonucleate, *Nature*. 172 (1953) 156–157. <https://doi.org/10.1038/172156a0>.
- [3] M.H.F. Wilkins, A.R. Stokes, H.R. Wilson, Molecular structure of nucleic acids: Molecular structure of deoxypentose nucleic acids, *Nature*. 171 (1953) 738–740. <https://doi.org/10.1038/171738a0>.
- [4] M. Grunberg-Manago, S. Ochoa, Enzymatic synthesis and breakdown of polynucleotides; polynucleotide phosphorylase, *J. Am. Chem. Soc.* 77 (1955) 3165–3166. <https://doi.org/10.1021/ja01616a093>.
- [5] B. Arthur Kornberg, V. Euler, ENZYMATIC SYNTHESIS OF TRIPHOSPHOPYRIDINE NUCLEOTIDE Downloaded from, n.d. <http://www.jbc.org/> (accessed March 27, 2020).
- [6] R.W. Holley, J. Apgar, G.A. Everett, J.T. Madison, M. Marquisee, S.H. Merrill, J.R. Penswick, A. Zamir, Structure of a ribonucleic acid, *Science* (80-.). 147 (1965) 1462–1465. <https://doi.org/10.1126/science.147.3664.1462>.
- [7] H.G. Khorana, H. Büchi, H. Ghosh, N. Gupta, T.M. Jacob, H. Kössel, R. Morgan, S.A. Narang, E. Ohtsuka, R.D. Wells, Polynucleotide synthesis and the genetic code., *Cold Spring Harb. Symp. Quant. Biol.* 31 (1966) 39–49. <https://doi.org/10.1101/SQB.1966.031.01.010>.
- [8] M.R. Bernfield, M.W. Nirenberg, RNA codewords and protein synthesis, *Science* (80-.). 147 (1965) 479–484. <https://doi.org/10.1126/science.147.3657.479>.
- [9] W. Arber, D. Dussoix, Host specificity of DNA produced by *Escherichia coli*: I. Host controlled modification of bacteriophage λ , *J. Mol. Biol.* 5 (1962) 18–36. [https://doi.org/10.1016/S0022-2836\(62\)80058-8](https://doi.org/10.1016/S0022-2836(62)80058-8).
- [10] H.O. Smith, K.W. Welcox, A Restriction enzyme from *Hemophilus influenzae*. I. Purification and general properties, *J. Mol. Biol.* 51 (1970) 379–391. [https://doi.org/10.1016/0022-2836\(70\)90149-X](https://doi.org/10.1016/0022-2836(70)90149-X).
- [11] K. Danna, D. Nathans, Specific cleavage of simian virus 40 DNA by restriction endonuclease of *Hemophilus influenzae*., *Proc. Natl. Acad. Sci. U. S. A.* 68 (1971) 2913–2917. <https://doi.org/10.1073/pnas.68.12.2913>.
- [12] P.W.J. Rigby, M. Dieckmann, C. Rhodes, P. Berg, Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I, *J. Mol. Biol.* 113 (1977) 237–251. [https://doi.org/10.1016/0022-2836\(77\)90052-3](https://doi.org/10.1016/0022-2836(77)90052-3).
- [13] A.M. Maxam, W. Gilbert, A new method for sequencing DNA, *Proc. Natl. Acad. Sci. U. S. A.* 74 (1977) 560–564. <https://doi.org/10.1073/pnas.74.2.560>.
- [14] F. Sanger, S. Nicklen, A.R. Coulson, DNA sequencing with chain-terminating inhibitors., *Proc. Natl. Acad. Sci. U. S. A.* 74 (1977) 5463–5467. <https://doi.org/10.1073/pnas.74.12.5463>.
- [15] C. Guerrier-Takada, K. Gardiner, T. Marsh, N. Pace, S. Altman, The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme, *Cell*. 35 (1983) 849–857. [https://doi.org/10.1016/0092-8674\(83\)90117-4](https://doi.org/10.1016/0092-8674(83)90117-4).
- [16] K. Kruger, P.J. Grabowski, A.J. Zaug, J. Sands, D.E. Gottschling, T.R. Cech, Self-splicing RNA: Autoexcision and autocyclization of the ribosomal RNA intervening sequence of tetrahymena, *Cell*. 31 (1982) 147–157. [https://doi.org/10.1016/0092-8674\(82\)90414-7](https://doi.org/10.1016/0092-8674(82)90414-7).
- [17] B. Thimmappaya, C. Weinberger, R.J. Schneider, T. Shenk, Adenovirus VAI RNA is required for efficient translation of viral mRNAs at late times after infection, *Cell*. 31 (1982) 543–551. [https://doi.org/10.1016/0092-8674\(82\)90310-5](https://doi.org/10.1016/0092-8674(82)90310-5).

- [18] C. Svensson, G. Akusjärvi, Adenovirus VA RNAI: a positive regulator of mRNA translation., *Mol. Cell. Biol.* 4 (1984) 736–42. <https://doi.org/10.1128/mcb.4.4.736>.
- [19] R.P. O'Malley, T.M. Mariano, J. Siekierka, M.B. Mathews, A mechanism for the control of protein synthesis by adenovirus VA RNAI, *Cell.* 44 (1986) 391–400. [https://doi.org/10.1016/0092-8674\(86\)90460-5](https://doi.org/10.1016/0092-8674(86)90460-5).
- [20] R. Strijker, D.T. Fritz, A.D. Levinson, Adenovirus VAI-RNA regulates gene expression by controlling stability of ribosome-bound RNAs., *EMBO J.* 8 (1989) 2669–75. <http://www.ncbi.nlm.nih.gov/pubmed/2583126> (accessed March 25, 2020).
- [21] L. Gold, SELEX: How It Happened and Where It will Go, *J. Mol. Evol.* 81 (2015) 140–143. <https://doi.org/10.1007/s00239-015-9705-9>.
- [22] D. Gillespie, S. Spiegelman, A quantitative assay for DNA-RNA hybrids with DNA immobilized on a membrane, *J. Mol. Biol.* 12 (1965) 829–842. [https://doi.org/10.1016/S0022-2836\(65\)80331-X](https://doi.org/10.1016/S0022-2836(65)80331-X).
- [23] D.R. Mills, R.L. Peterson, S. Spiegelman, An extracellular Darwinian experiment with a self-duplicating nucleic acid molecule., *Proc. Natl. Acad. Sci. U. S. A.* 58 (1967) 217–224. <https://doi.org/10.1073/pnas.58.1.217>.
- [24] F.R. Kramer, D.R. Mills, P.E. Cole, T. Nishihara, S. Spiegelman, Evolution in vitro: Sequence and phenotype of a mutant RNA resistant to ethidium bromide, *J. Mol. Biol.* 89 (1974) 719–736. [https://doi.org/10.1016/0022-2836\(74\)90047-3](https://doi.org/10.1016/0022-2836(74)90047-3).
- [25] K.B. MULLIS, F.A. FALONNA, Specific Synthesis of DNA in Vitro via a Polymerase-Catalyzed Chain Reaction, in: *Recomb. DNA Methodol.*, Elsevier, 1989: pp. 189–204. <https://doi.org/10.1016/b978-0-12-765560-4.50015-0>.
- [26] A.R. Oliphant, K. Struhl, An efficient method for generating proteins with altered enzymatic properties: Application to β -lactamase, *Proc. Natl. Acad. Sci. U. S. A.* 86 (1989) 9094–9098. <https://doi.org/10.1073/pnas.86.23.9094>.
- [27] C. Tuerk, L. Gold, Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase, *Science* (80-.). 249 (1990) 505–510. <https://doi.org/10.1126/science.2200121>.
- [28] A.D. Ellington, J.W. Szostak, In vitro selection of RNA molecules that bind specific ligands, *Nature.* 346 (1990) 818–822. <https://doi.org/10.1038/346818a0>.
- [29] L. Gold, C. Tuerk, *Methods for identifying nucleic acid ligands*, US5270163A, 1993.
- [30] L. Gold, S. Ringquist, *Systematic evolution of ligands by exponential enrichment: Solution SELEX*, US5567588A, 1996.
- [31] L. Gold, R.D. Jenison, B. Polisky, *High-affinity nucleic acid ligands that discriminate between theophylline and caffeine*, US5580737A, 1996.
- [32] M.C. Cowperthwaite, A.D. Ellington, Bioinformatic analysis of the contribution of primer sequences to aptamer structures, *J. Mol. Evol.* 67 (2008) 95–102. <https://doi.org/10.1007/s00239-008-9130-4>.
- [33] J. Ciesiolka, M. Illangasekare, I. Majerfeld, T. Nickles, M. Welch, M. Yarus, S. Zinnen, Affinity selection-amplification from randomized ribooligonucleotide pools., *Methods Enzymol.* 267 (1996) 315–35. [https://doi.org/10.1016/s0076-6879\(96\)67021-9](https://doi.org/10.1016/s0076-6879(96)67021-9).
- [34] B. Zimmermann, I. Bilusic, C. Lorenz, R. Schroeder, Genomic SELEX: A discovery tool for genomic aptamers, *Methods.* 52 (2010) 125–132. <https://doi.org/10.1016/j.ymeth.2010.06.004>.
- [35] C. Marimuthu, T.H. Tang, J. Tominaga, S.C. Tan, S.C.B. Gopinath, Single-stranded DNA (ssDNA) production in DNA aptamer generation, *Analyst.* 137 (2012) 1307–1315. <https://doi.org/10.1039/c2an15905h>.
- [36] H. Qu, A.T. Csordas, J. Wang, S.S. Oh, M.S. Eisenstein, H.T. Soh, Rapid and Label-Free Strategy to Isolate Aptamers for Metal Ions, *ACS Nano.* 10 (2016) 7558–7565. <https://doi.org/10.1021/acsnano.6b02558>.
- [37] L.A. Holeman, S.L. Robinson, J.W. Szostak, C. Wilson, Isolation and characterization of fluorophore-binding RNA

- aptamers, *Fold. Des.* 3 (1998) 423–431. [https://doi.org/10.1016/S1359-0278\(98\)00059-5](https://doi.org/10.1016/S1359-0278(98)00059-5).
- [38] M. Famulok, J.S. Hartig, G. Mayer, Functional aptamers and aptazymes in biotechnology, diagnostics, and therapy, *Chem. Rev.* 107 (2007) 3715–3743. <https://doi.org/10.1021/cr0306743>.
- [39] D.H.J. Bunka, P.G. Stockley, Aptamers come of age - At last, *Nat. Rev. Microbiol.* 4 (2006) 588–596. <https://doi.org/10.1038/nrmicro1458>.
- [40] D.S. Wilson, J.W. Szostak, In Vitro Selection of Functional Nucleic Acids, *Annu. Rev. Biochem.* 68 (1999) 611–647. <https://doi.org/10.1146/annurev.biochem.68.1.611>.
- [41] R. Stoltenburg, T. Schubert, B. Strehlitz, In vitro Selection and Interaction Studies of a DNA Aptamer Targeting Protein A, (2015). <https://doi.org/10.1371/journal.pone.0134403>.
- [42] T. Iwagawa, S.P. Ohuchi, S. Watanabe, Y. Nakamura, Selection of RNA aptamers against mouse embryonic stem cells, *Biochimie.* 94 (2012) 250–257. <https://doi.org/10.1016/j.biochi.2011.10.017>.
- [43] M. Yarus, How many catalytic RNAs? Ions and the Cheshire cat conjecture., *FASEB J.* 7 (1993) 31–39. <https://doi.org/10.1096/fasebj.7.1.8422972>.
- [44] G.J. Connell, M. Yarus, RNAs with dual specificity and dual RNAs with similar specificity, *Science* (80-.). 264 (1994) 1137–1141. <https://doi.org/10.1126/science.7513905>.
- [45] F. Mathieu-Daudé, J. Welsh, T. Vogt, M. McClelland, DNA rehybridization during PCR: the “Cot effect” and its consequences., *Nucleic Acids Res.* 24 (1996) 2080–6. <https://doi.org/10.1093/nar/24.11.2080>.
- [46] D.B. Olsen, S.S. Carroll, J.C. Culbertson, J.A. Shafer, L.C. Kuo, Effect of template secondary structure on the inhibition of HIV-1 reverse transcriptase by a pyridinone non-nucleoside inhibitor, 1994.
- [47] J.A. Bittker, B. V. Le, D.R. Liu, Nucleic acid evolution and minimization by nonhomologous random recombination, *Nat. Biotechnol.* 20 (2002) 1024–1029. <https://doi.org/10.1038/nbt736>.
- [48] J. Tsang, G.F. Joyce, Specialization of the DNA-cleaving activity of a group I ribozyme through in vitro evolution, *J. Mol. Biol.* 262 (1996) 31–42. <https://doi.org/10.1006/jmbi.1996.0496>.
- [49] J. Cline, J.C. Braman, H.H. Hogrefe, PCR fidelity of Pfu DNA polymerase and other thermostable DNA polymerases, 1996.
- [50] J.-P. Vartanian, M. Henry, S. Wain-Hobson, Hypermutagenic PCR involving all four transitions and a sizeable proportion of transversions, 1996.
- [51] D.S. Wilson, A.D. Keefe, Random Mutagenesis by PCR, in: *Curr. Protoc. Mol. Biol.*, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2001. <https://doi.org/10.1002/0471142727.mb0803s51>.
- [52] R. Knight, M. Yarus, Analyzing partially randomized nucleic acid pools: straight dope on doping, *Nucleic Acids Res.* 31 (2003) e30. <https://doi.org/10.1093/NAR/GNG030>.
- [53] Y.S. Kim, C.J. Hyun, I.A. Kim, M.B. Gu, Isolation and characterization of enantioselective DNA aptamers for ibuprofen, *Bioorganic Med. Chem.* 18 (2010) 3467–3473. <https://doi.org/10.1016/j.bmc.2010.03.074>.
- [54] M. Blank, T. Weinschenk, M. Priemer, H. Schluesener, Systematic evolution of a DNA aptamer binding to rat brain tumor microvessels: Selective targeting of endothelial regulatory protein pigpen, *J. Biol. Chem.* 276 (2001) 16464–16468. <https://doi.org/10.1074/jbc.M100347200>.
- [55] Z. Zhuo, Y. Yu, M. Wang, J. Li, Z. Zhang, J. Liu, X. Wu, A. Lu, G. Zhang, B. Zhang, Recent Advances in SELEX Technology and Aptamer Applications in Biomedicine, *Int. J. Mol. Sci.* 18 (2017) 2142. <https://doi.org/10.3390/ijms18102142>.

- [56] J. Zhou, J. Rossi, Aptamers as targeted therapeutics: Current potential and challenges, *Nat. Rev. Drug Discov.* 16 (2017) 181–202. <https://doi.org/10.1038/nrd.2016.199>.
- [57] K.M. Song, S. Lee, C. Ban, Aptamers and their biological applications, *Sensors.* 12 (2012) 612–631. <https://doi.org/10.3390/s120100612>.
- [58] C. Kratschmer, M. Levy, Effect of Chemical Modifications on Aptamer Stability in Serum, *Nucleic Acid Ther.* 27 (2017) 335–344. <https://doi.org/10.1089/nat.2017.0680>.
- [59] A.D. Keefe, S.T. Cload, SELEX with modified nucleotides, *Curr. Opin. Chem. Biol.* 12 (2008) 448–456. <https://doi.org/10.1016/j.cbpa.2008.06.028>.
- [60] E.W.M. Ng, D.T. Shima, P. Calias, E.T. Cunningham, D.R. Guyer, A.P. Adamis, Pegaptanib, a targeted anti-VEGF aptamer for ocular vascular disease, *Nat. Rev. Drug Discov.* 5 (2006) 123–132. <https://doi.org/10.1038/nrd1955>.
- [61] S. Ni, H. Yao, L. Wang, J. Lu, F. Jiang, A. Lu, G. Zhang, Chemical modifications of nucleic acid aptamers for therapeutic purposes, *Int. J. Mol. Sci.* 18 (2017). <https://doi.org/10.3390/ijms18081683>.
- [62] A.D. Ellington, J.W. Szostak, Selection in vitro of single-stranded DNA molecules that fold into specific ligand-binding structures, *Nature.* 355 (1992) 850–852. <https://doi.org/10.1038/355850a0>.
- [63] C. Wilson, J.W. Szostak, Isolation of a fluorophore-specific DNA aptamer with weak redox activity, *Chem. Biol.* 5 (1998) 609–617. [https://doi.org/10.1016/S1074-5521\(98\)90289-7](https://doi.org/10.1016/S1074-5521(98)90289-7).
- [64] A.D. Garst, A.L. Edwards, R.T. Batey, Riboswitches: Structures and mechanisms, *Cold Spring Harb. Perspect. Biol.* 3 (2011) 1–13. <https://doi.org/10.1101/cshperspect.a003533>.
- [65] L. Gold, G. Stormo, R. Saunders, Escherichia coli translational initiation factor IF3: A unique case of translational regulation, *Proc. Natl. Acad. Sci. U. S. A.* 81 (1984) 7061–7065. <https://doi.org/10.1073/pnas.81.22.7061>.
- [66] B.A. Murphy McDaniel, F.J. Grundy, I. Artsimovitch, T.M. Henkin, Transcription termination control of the S box system: Direct measurement of S-adenosylmethionine by the leader RNA, *Proc. Natl. Acad. Sci. U. S. A.* 100 (2003) 3083–3088. <https://doi.org/10.1073/pnas.0630422100>.
- [67] A.S. Mironov, I. Gusarov, R. Rafikov, L.E. Lopez, K. Shatalin, R.A. Kreneva, D.A. Perumov, E. Nudler, Sensing small molecules by nascent RNA: A mechanism to control transcription in bacteria, *Cell.* 111 (2002) 747–756. [https://doi.org/10.1016/S0092-8674\(02\)01134-0](https://doi.org/10.1016/S0092-8674(02)01134-0).
- [68] W.C. Winkler, S. Cohen-Chalamish, R.R. Breaker, An mRNA structure that controls gene expression by binding FMN, *Proc. Natl. Acad. Sci. U. S. A.* 99 (2002) 15908–15913. <https://doi.org/10.1073/pnas.212628899>.
- [69] W. Winkler, A. Nahvi, R.R. Breaker, Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression, *Nature.* 419 (2002) 952–956. <https://doi.org/10.1038/nature01145>.
- [70] A. Nahvi, N. Sudarsan, M.S. Ebert, X. Zou, K.L. Brown, R.R. Breaker, Genetic control by a metabolite binding mRNA, *Chem. Biol.* 9 (2002) 1043–1049. [https://doi.org/10.1016/S1074-5521\(02\)00224-7](https://doi.org/10.1016/S1074-5521(02)00224-7).
- [71] M. Gossen, H. Bujard, Tight control of gene expression in mammalian cells by tetracycline-responsive promoters, *Proc. Natl. Acad. Sci. U. S. A.* 89 (1992) 5547–5551. <https://doi.org/10.1073/pnas.89.12.5547>.
- [72] D.M. Spencer, T.J. Wandless, S.L. Schreiber, G.R. Crabtree, Controlling signal transduction with synthetic ligands, *Science (80-.)*. 262 (1993) 1019–1024. <https://doi.org/10.1126/science.7694365>.
- [73] L. Gold, D. Brown, Y.Y. He, T. Shtatland, B.S. Singer, Y. Wu, From oligonucleotide shapes to genomic SELEX: Novel biological regulatory loops, *Proc. Natl. Acad. Sci. U. S. A.* 94 (1997) 59–64. <https://doi.org/10.1073/pnas.94.1.59>.
- [74] G. Werstuck, M.R. Green, Controlling gene expression in living cells through small molecule-RNA interactions, *Science*

- (80-). 282 (1998) 296–298. <https://doi.org/10.1126/science.282.5387.296>.
- [75] D. Grate, C. Wilson, Laser-mediated, site-specific inactivation of RNA transcripts, *Proc. Natl. Acad. Sci. U. S. A.* 96 (1999) 6131–6136. <https://doi.org/10.1073/pnas.96.11.6131>.
- [76] J.R. Babendure, S.R. Adams, R.Y. Tsien, Aptamers Switch on Fluorescence of Triphenylmethane Dyes, *J. Am. Chem. Soc.* 125 (2003) 14716–14717. <https://doi.org/10.1021/ja037994o>.
- [77] M.N. Stojanovic, D.M. Kolpashchikov, Modular aptameric sensors, *J. Am. Chem. Soc.* 126 (2004) 9266–9270. <https://doi.org/10.1021/ja032013t>.
- [78] D.M. Kolpashchikov, Binary malachite green aptamer for fluorescent detection of nucleic acids, *J. Am. Chem. Soc.* 127 (2005) 12442–12443. <https://doi.org/10.1021/ja0529788>.
- [79] W. Xu, Y. Lu, Label-free fluorescent aptamer sensor based on regulation of malachite green fluorescence, *Anal. Chem.* 82 (2010) 574–578. <https://doi.org/10.1021/ac9018473>.
- [80] D.G. Jay, H. Keshishian, Laser inactivation of fasciclin I disrupts axon adhesion of grasshopper pioneer neurons, *Nature.* 348 (1990) 548–550. <https://doi.org/10.1038/348548a0>.
- [81] D. Grate, C. Wilson, Inducible regulation of the *S. cerevisiae* cell cycle mediated by an RNA aptamer-ligand complex, in: *Bioorganic Med. Chem.*, 2001: pp. 2565–2570. [https://doi.org/10.1016/S0968-0896\(01\)00031-1](https://doi.org/10.1016/S0968-0896(01)00031-1).
- [82] B. Suess, B. Fink, C. Berens, R. Stentz, W. Hillen, A theophylline responsive riboswitch based on helix slipping controls gene expression in vivo, *Nucleic Acids Res.* 32 (2004) 1610–1614. <https://doi.org/10.1093/nar/gkh321>.
- [83] J.K. Wickiser, W.C. Winkler, R.R. Breaker, D.M. Crothers, The speed of RNA transcription and metabolite binding kinetics operate an FMN riboswitch, *Mol. Cell.* 18 (2005) 49–60. <https://doi.org/10.1016/j.molcel.2005.02.032>.
- [84] J.E. Weigand, M. Sanchez, E.B. Gunnesch, S. Zeiher, R. Schroeder, B. Suess, Screening for engineered neomycin riboswitches that control translation initiation, *RNA.* 14 (2008) 89–97. <https://doi.org/10.1261/rna.772408>.
- [85] P. Burgstaller, M. Famulok, Isolation of RNA Aptamers for Biological Cofactors by In Vitro Selection, *Angew. Chemie Int. Ed. English.* 33 (1994) 1084–1087. <https://doi.org/10.1002/anie.199410841>.
- [86] M. Famulok, J.W. Szostak, Stereospecific Recognition of Tryptophan Agarose by in Vitro Selected RNA, *J. Am. Chem. Soc.* 114 (1992) 3990–3991. <https://doi.org/10.1021/ja00036a065>.
- [87] C.T. Lauhon, J.W. Szostak, RNA Aptamers that Bind Flavin and Nicotinamide Redox Cofactors, *J. Am. Chem. Soc.* 117 (1995) 1246–1257. <https://doi.org/10.1021/ja00109a008>.
- [88] A. Arora, M. Sunbul, A.J. Aschke, Dual-colour imaging of RNAs using quencher-and fluorophore-binding aptamers, *Nucleic Acids Res.* 43 (2015). <https://doi.org/10.1093/nar/gkv718>.
- [89] M. Sunbul, A.J. Aschke, SRB-2: a promiscuous rainbow aptamer for live-cell RNA imaging, *Nucleic Acids Res.* 46 (2018). <https://doi.org/10.1093/nar/gky543>.
- [90] J.S. Paige, K.Y. Wu, S.R. Jaffrey, RNA mimics of green fluorescent protein, *Science* (80-). 333 (2011) 642–646. <https://doi.org/10.1126/science.1207339>.
- [91] G.S. Filonov, J.D. Moon, N. Svensen, S.R. Jaffrey, Broccoli: Rapid selection of an RNA mimic of green fluorescent protein by fluorescence-based selection and directed evolution, *J. Am. Chem. Soc.* 136 (2014) 16299–16308. <https://doi.org/10.1021/ja508478x>.
- [92] W. Song, G.S. Filonov, H. Kim, M. Hirsch, X. Li, J.D. Moon, S.R. Jaffrey, Imaging RNA polymerase III transcription using a photostable RNA-fluorophore complex, *Nat. Chem. Biol.* 13 (2017) 1187–1194. <https://doi.org/10.1038/nchembio.2477>.

- [93] E. V. Dolgosheina, S.C.Y. Jeng, S.S.S. Panchapakesan, R. Cojocar, P.S.K. Chen, P.D. Wilson, N. Hawkins, P.A. Wiggins, P.J. Unrau, RNA Mango aptamer-fluorophore: A bright, high-affinity complex for RNA labeling and tracking, *ACS Chem. Biol.* 9 (2014) 2412–2420. <https://doi.org/10.1021/cb500499x>.
- [94] W.Q. Ong, Y.R. Citron, S. Sekine, B. Huang, Live Cell imaging of endogenous mRNA Using RNA-Based fluorescence “turn-on” probe, *ACS Chem. Biol.* 12 (2017) 200–205. <https://doi.org/10.1021/acscchembio.6b00586>.
- [95] V.B. O’Leary, S. Hain, D. Maugg, J. Smida, O. Azimzadeh, S. Tapio, S.V. Ovsepian, M.J. Atkinson, Long non-coding RNA PARTICLE bridges histone and DNA methylation, *Sci. Rep.* 7 (2017). <https://doi.org/10.1038/s41598-017-01875-1>.
- [96] K.D. Warner, L. Sjekloa, W. Song, G.S. Filonov, S.R. Jaffrey, A.R. Ferré-D’Amaré, A homodimer interface without base pairs in an RNA mimic of red fluorescent protein, *Nat. Chem. Biol.* 13 (2017) 1195–1201. <https://doi.org/10.1038/nchembio.2475>.
- [97] I. Masuda, T. Igarashi, R. Sakaguchi, R.G. Nitharwal, R. Takase, K.Y. Han, B.J. Leslie, C. Liu, H. Gamper, T. Ha, S. Sanyal, Y.M. Hou, A genetically encoded fluorescent tRNA is active in live-cell protein synthesis, *Nucleic Acids Res.* (2016).
- [98] A. Autour, S.C.Y. Jeng, A.D. Cawte, A. Abdolhazadeh, A. Galli, S.S.S. Panchapakesan, D. Rueda, M. Ryckelynck, P.J. Unrau, Fluorogenic RNA Mango aptamers for imaging small non-coding RNAs in mammalian cells, *Nat. Commun.* 9 (2018) 1–12. <https://doi.org/10.1038/s41467-018-02993-8>.
- [99] Z.M. Ying, B. Tu, L. Liu, H. Tang, L.J. Tang, J.H. Jiang, Spinach-based fluorescent light-up biosensors for multiplexed and label-free detection of microRNAs, *Chem. Commun.* 54 (2018) 3010–3013. <https://doi.org/10.1039/c8cc00123e>.
- [100] Z. JA, R. M, J. C, M. KA, B. MF, Live-cell imaging of small nucleolar RNA tagged with the broccoli aptamer in yeast., *FEMS Yeast Res.* 18 (2018). <https://doi.org/10.1093/FEMSYR/FOY093>.
- [101] M. Zhou, X. Teng, Y. Li, R. Deng, J. Li, Cascade Transcription Amplification of RNA Aptamer for Ultrasensitive MicroRNA Detection, *Anal. Chem.* 91 (2019) 5295–5302. <https://doi.org/10.1021/acs.analchem.9b00124>.
- [102] X. Chen, D. Zhang, N. Su, B. Bao, X. Xie, F. Zuo, L. Yang, H. Wang, L. Jiang, Q. Lin, M. Fang, N. Li, X. Hua, Z. Chen, C. Bao, J. Xu, W. Du, L. Zhang, Y. Zhao, L. Zhu, J. Loscalzo, Y. Yang, Visualizing RNA dynamics in live cells with bright and stable fluorescent RNAs, *Nat. Biotechnol.* 37 (2019) 1287–1293. <https://doi.org/10.1038/s41587-019-0249-1>.
- [103] R.L. Strack, M.D. Disney, S.R. Jaffrey, A superfolding Spinach2 reveals the dynamic nature of trinucleotide repeat-containing RNA, *Nat. Methods.* 10 (2013) 1219–1224. <https://doi.org/10.1038/nmeth.2701>.
- [104] G. Pothoulakis, F. Ceroni, B. Reeve, T. Ellis, The Spinach RNA aptamer as a characterization tool for synthetic biology, *ACS Synth. Biol.* 3 (2014) 182–187. <https://doi.org/10.1021/sb400089c>.
- [105] J. Chloe Bulinski, D. Gruber, K. Faire, P. Prasad, W. Chang, GFP Chimeras of E-MAP-115 (ensconsin) Domains Mimic Behavior of the Endogenous Protein in vitro and in vivo., *Cell Struct. Funct.* 24 (1999) 313–320. <https://doi.org/10.1247/csf.24.313>.
- [106] G. Genové, B.S. Glick, A.L. Barth, Brighter reporter genes from multimerized fluorescent proteins, *Biotechniques.* 39 (2005) 814–818. <https://doi.org/10.2144/000112056>.
- [107] J. Zhang, J. Fei, B.J. Leslie, K.Y. Han, T.E. Kuhlman, T. Ha, Tandem spinach array for mRNA imaging in living bacterial cells, *Sci. Rep.* 5 (2015) 17295. <https://doi.org/10.1038/srep17295>.
- [108] P. Rugbjerg, C. Knuf, J. Förster, M.O.A. Sommer, Recombination-stable multimeric green fluorescent protein for characterization of weak promoter outputs in *Saccharomyces cerevisiae*, *FEMS Yeast Res.* 15 (2015) fov085. <https://doi.org/10.1093/femsyr/fov085>.
- [109] F. Bouhedda, K.T. Fam, M. Collot, A. Autour, S. Marzi, A. Klymchenko, M. Ryckelynck, A dimerization-based fluorogenic dye-aptamer module for RNA imaging in live cells, *Nat. Chem. Biol.* 16 (2020) 69–76. <https://doi.org/10.1038/s41589-019-0381-8>.

- [110] E. Heyduk, T. Heyduk, Nucleic acid-based fluorescence sensors for detecting proteins, *Anal. Chem.* 77 (2005) 1147–1156. <https://doi.org/10.1021/ac0487449>.
- [111] H. Chang, L. Tang, Y. Wang, J. Jiang, J. Li, Graphene fluorescence resonance energy transfer aptasensor for the thrombin detection, *Anal. Chem.* 82 (2010) 2341–2346. <https://doi.org/10.1021/ac9025384>.
- [112] M.N. Stojanovic, P. de Prada, D.W. Landry, Aptamer-based folding fluorescent sensor for cocaine, *J. Am. Chem. Soc.* 123 (2001) 4928–4931. <https://doi.org/10.1021/ja0038171>.
- [113] J. Ouellet, RNA fluorescence with light-Up aptamers, *Front. Chem.* 4 (2016). <https://doi.org/10.3389/fchem.2016.00029>.
- [114] J.S. Paige, T. Nguyen-Duc, W. Song, S.R. Jaffrey, Fluorescence imaging of cellular metabolites with RNA, *Science* (80-.). 335 (2012) 1194. <https://doi.org/10.1126/science.1218298>.
- [115] S. Nakayama, Y. Luo, J. Zhou, T.K. Dayie, H.O. Sintim, Nanomolar fluorescent detection of c-di-GMP using a modular aptamer strategy, *Chem. Commun.* 48 (2012) 9059–9061. <https://doi.org/10.1039/c2cc34379g>.
- [116] C.A. Kellenberger, S.C. Wilson, J. Sales-Lee, M.C. Hammond, RNA-based fluorescent biosensors for live cell imaging of second messengers cyclic di-GMP and cyclic AMP-GMP, *J. Am. Chem. Soc.* 135 (2013) 4906–4909. <https://doi.org/10.1021/ja311960g>.
- [117] T. Dieckmann, E. Suzuki, G.K. Nakamura, J. Feigon, Solution structure of an ATP-binding RNA aptamer reveals a novel fold, *RNA*. 2 (1996) 628–640. <https://doi.org/10.2210/pdb1raw/pdb>.
- [118] T. Dieckmann, S.E. Butcher, M. Sassanfar, J.W. Szostak, J. Feigon, Mutant ATP-binding RNA aptamers reveal the structural basis for ligand binding, *J. Mol. Biol.* 273 (1997) 467–478. <https://doi.org/10.1006/jmbi.1997.1329>.
- [119] D.J. Patel, A.K. Suri, F. Jiang, L. Jiang, P. Fan, R.A. Kumar, S. Nonin, [1] D.J. Patel, A.K. Suri, F. Jiang, L. Jiang, P. Fan, R.A. Kumar, S. Nonin, Structure, recognition and adaptive binding in RNA aptamer complexes, *J. Mol. Biol.* 272 (1997) 645–664. <https://doi.org/10.1006/jmbi.1997.1281>. Structure, recognition and adaptive, *J. Mol. Biol.* 272 (1997) 645–664. <https://doi.org/10.1006/jmbi.1997.1281>.
- [120] T. Hermann, D.J. Patel, Adaptive recognition by nucleic acid aptamers, *Science* (80-.). 287 (2000) 820–825. <https://doi.org/10.1126/science.287.5454.820>.
- [121] S.D. Gilbert, F.E. Reyes, A.L. Edwards, R.T. Batey, Adaptive Ligand Binding by the Purine Riboswitch in the Recognition of Guanine and Adenine Analogs, *Structure*. 17 (2009) 857–868. <https://doi.org/10.1016/j.str.2009.04.009>.
- [122] J.B. Da Costa, A.I. Andreiev, T. Dieckmann, Thermodynamics and Kinetics of Adaptive Binding in the Malachite Green RNA Aptamer, *Biochemistry*. 52 (2013) 6575–6583. <https://doi.org/10.1021/bi400549s>.
- [123] C. Baugh, D. Grate, C. Wilson, 2.8 Å Crystal structure of the malachite green aptamer, *J. Mol. Biol.* 301 (2000) 117–128. <https://doi.org/10.1006/jmbi.2000.3951>.
- [124] J. Flinders, S.C. DeFina, D.M. Brackett, C. Baugh, C. Wilson, T. Dieckmann, Recognition of Planar and Nonplanar Ligands in the Malachite Green-RNA Aptamer Complex, *ChemBioChem*. 5 (2004) 62–72. <https://doi.org/10.1002/cbic.200300701>.
- [125] K.D. Warner, M.C. Chen, W. Song, R.L. Strack, A. Thorn, S.R. Jaffrey, A.R. Ferré-D'Amaré, Structural basis for activity of highly efficient RNA mimics of green fluorescent protein, *Nat. Struct. Mol. Biol.* 21 (2014) 658–663. <https://doi.org/10.1038/nsmb.2865>.
- [126] P. Fernandez-Millan, A. Autour, E. Ennifar, E. Westhof, M. Ryckelynck, Crystal structure and fluorescence properties of the iSpinach aptamer in complex with DFHBI, *RNA*. 23 (2017) 1788–1795. <https://doi.org/10.1261/rna.063008.117>.
- [127] R.J. Trachman, A. Abdolazadeh, A. Andreoni, R. Cojocar, J.R. Knutson, M. Ryckelynck, P.J. Unrau, A.R. Ferré-D'Amaré,

Crystal Structures of the Mango-II RNA Aptamer Reveal Heterogeneous Fluorophore Binding and Guide Engineering of Variants with Improved Selectivity and Brightness, *Biochemistry*. 57 (2018) 3544–3548. <https://doi.org/10.1021/acs.biochem.8b00399>.

- [128] R.J. Trachman, A. Autour, S.C.Y. Jeng, A. Abdolazadeh, A. Andreoni, R. Cojocar, R. Garipov, E. V. Dolgosheina, J.R. Knutson, M. Ryckelynck, P.J. Unrau, A.R. Ferré-D'Amaré, Structure and functional reselection of the Mango-III fluorogenic RNA aptamer, *Nat. Chem. Biol.* 15 (2019) 472–479. <https://doi.org/10.1038/s41589-019-0267-9>.
- [129] E. Duchardt-Ferner, M. Juen, B. Bourgeois, T. Madl, C. Kreutz, O. Ohlenschlager, J. Wohnert, Structure of an RNA aptamer in complex with the fluorophore tetramethylrhodamine, *Nucleic Acids Res.* 48 (2020) 949–961.
- [130] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera - A visualization system for exploratory research and analysis, *J. Comput. Chem.* 25 (2004) 1605–1612. <https://doi.org/10.1002/jcc.20084>.
- [131] B.J. Saville, R.A. Collins, A site-specific self-cleavage reaction performed by a novel RNA in neurospora mitochondria, *Cell*. 61 (1990) 685–696. [https://doi.org/10.1016/0092-8674\(90\)90480-3](https://doi.org/10.1016/0092-8674(90)90480-3).
- [132] G.A. Prody, J.T. Bakos, J.M. Buzayan, I.R. Schneider, G. Bruening, Autolytic processing of dimeric plant virus satellite RNA, *Science* (80-.). 231 (1986) 1577–1580. <https://doi.org/10.1126/science.231.4745.1577>.
- [133] A.C. Forster, R.H. Symons, Self-cleavage of plus and minus RNAs of a virusoid and a structural model for the active sites, *Cell*. 49 (1987) 211–220. [https://doi.org/10.1016/0092-8674\(87\)90562-9](https://doi.org/10.1016/0092-8674(87)90562-9).
- [134] J.M. Kaper, M.E. Tousignant, G. Steger, Nucleotide sequence predicts circularity and self-cleavage of 300-ribonucleotide satellite of arabis mosaic virus, *Biochem. Biophys. Res. Commun.* 154 (1988) 318–325. [https://doi.org/10.1016/0006-291X\(88\)90687-0](https://doi.org/10.1016/0006-291X(88)90687-0).
- [135] T.R. Cech, The ribosome is a ribozyme, *Science* (80-.). 289 (2000) 878–879. <https://doi.org/10.1126/SCIENCE.289.5481.878>.
- [136] C.A. Collins, C. Guthrie, The question remains: Is the spliceosome a ribozyme?, *Nat. Struct. Biol.* 2000 710. 7 (2000) 850–854. <https://doi.org/10.1038/79598>.
- [137] W.C. Winkler, A. Nahvi, A. Roth, J.A. Collins, R.R. Breaker, Control of gene expression by a natural metabolite-responsive ribozyme, *Nature*. 428 (2004) 281–286. <https://doi.org/10.1038/nature02362>.
- [138] B. Seelig, A. Jaschke, A small catalytic RNA motif with Diels-Alderase activity, *Chem. Biol.* 6 (1999) 167–176. <http://biomednet.com/elecref/1074552100600167> (accessed May 26, 2020).
- [139] S. Tsukiji, S.B. Pattnaik, H. Suga, An alcohol dehydrogenase ribozyme, *Nat. Struct. Biol.* 10 (2003) 713–717. <https://doi.org/10.1038/nsb964>.
- [140] S. Tsukiji, S.B. Pattnaik, H. Suga, Reduction of an Aldehyde by a NADH/Zn²⁺-Dependent Redox Active Ribozyme, *J. Am. Chem. Soc.* 126 (2004) 5044–5045. <https://doi.org/10.1021/ja0495213>.
- [141] M. Popović, P.S. Fliss, M.A. Ditzler, In vitro evolution of distinct self-cleaving ribozymes in diverse environments, *Nucleic Acids Res.* 43 (2015) 7070–7082. <https://doi.org/10.1093/nar/gkv648>.
- [142] D.L. Robertson, G.F. Joyce, Selection in vitro of an RNA enzyme that specifically cleaves single-stranded DNA, *Nature*. 344 (1990) 467–468. <https://doi.org/10.1038/344467a0>.
- [143] G.F. Joyce, Directed Evolution of Nucleic Acid Enzymes, *Annu. Rev. Biochem.* 73 (2004) 791–836. <https://doi.org/10.1146/annurev.biochem.73.011303.073717>.
- [144] R.R. Breaker, G.F. Joyce, A DNA enzyme that cleaves RNA, *Chem. Biol.* 1 (1994) 223–229. [https://doi.org/10.1016/1074-5521\(94\)90014-0](https://doi.org/10.1016/1074-5521(94)90014-0).

- [145] N. Carmi, S.R. Balkhi, R.R. Breaker, Cleaving DNA with DNA, *Proc. Natl. Acad. Sci. U. S. A.* 95 (1998) 2233–2237. <https://doi.org/10.1073/pnas.95.5.2233>.
- [146] W.E. Purtha, R.L. Coppins, M.K. Smalley, S.K. Silverman, General deoxyribozyme-catalyzed synthesis of native 3'-5' RNA linkages, *J. Am. Chem. Soc.* 127 (2005) 13124–13125. <https://doi.org/10.1021/ja0533702>.
- [147] B. Cuenoud, J.W. Szostak, A DNA metalloenzyme with DNA ligase activity, *Nature*. 375 (1995) 611–614. <https://doi.org/10.1038/375611a0>.
- [148] W. Wang, L.P. Billen, Y. Li, Sequence diversity, metal specificity, and catalytic proficiency of metal-dependent phosphorylating DNA enzymes, *Chem. Biol.* 9 (2002) 507–517. [https://doi.org/10.1016/S1074-5521\(02\)00127-8](https://doi.org/10.1016/S1074-5521(02)00127-8).
- [149] Y. Li, Y. Liu, R.R. Breaker, Capping DNA with DNA, *Biochemistry*. 39 (2000) 3106–3114. <https://doi.org/10.1021/bi992710r>.
- [150] T.L. Sheppard, P. Ordoukhanian, G.F. Joyce, A DNA enzyme with N-glycosylase activity, *Proc. Natl. Acad. Sci. U. S. A.* 97 (2000) 7802–7807. <https://doi.org/10.1073/pnas.97.14.7802>.
- [151] M. Chandra, S.K. Silverman, DNA and RNA can be equally efficient catalysts for carbon-carbon bond formation, *J. Am. Chem. Soc.* 130 (2008) 2936–2937. <https://doi.org/10.1021/ja7111965>.
- [152] N. Krug, J.M. Hohlfeld, A.-M. Kirsten, O. Kornmann, K.M. Beeh, D. Kappeler, S. Korn, S. Ignatenko, W. Timmer, C. Rogon, J. Zeitvogel, N. Zhang, J. Bille, U. Homburg, A. Turowska, C. Bachert, T. Werfel, R. Buhl, J. Renz, H. Garn, H. Renz, Allergen-Induced Asthmatic Responses Modified by a GATA3-Specific DNAzyme, *N. Engl. J. Med.* 372 (2015) 1987–1995. <https://doi.org/10.1056/NEJMoa1411776>.
- [153] K. Saito, N. Shimada, A. Maruyama, Cooperative enhancement of deoxyribozyme activity by chemical modification and added cationic copolymer, *Sci. Technol. Adv. Mater.* 17 (2016) 437–442. <https://doi.org/10.1080/14686996.2016.1208627>.
- [154] G.L. Eichhorn, J.J. Butzow, Y.A. Shin, Some effects of metal ions on DNA structure and genetic information transfer*, 1985.
- [155] M. Morell Cerdà, D. Amantia, B. Costisella, A. Houlton, B. Lippert, Multiple metal binding to 6-oxopurine nucleobases as a source of deprotonation. The role of metal ions at N7 and N3, *Dalt. Trans.* (2006) 3894–3899. <https://doi.org/10.1039/b603650c>.
- [156] K.A. Sharp, B. Honig, Salt effects on nucleic acids, *Curr. Opin. Struct. Biol.* 5 (1995) 323–328. [https://doi.org/10.1016/0959-440X\(95\)80093-X](https://doi.org/10.1016/0959-440X(95)80093-X).
- [157] J. Muñoz, J.L. Gelpí, M. Soler-López, J.A. Subirana, M. Orozco, F.J. Luque, Can divalent metal cations stabilize the triplex motif? Theoretical study of the interaction of the hydrated Mg²⁺ cation with the G-G-C triplet, *J. Phys. Chem. B.* 106 (2002) 8849–8857. <https://doi.org/10.1021/jp026096w>.
- [158] D.M.J. Lilley, R.M. Clegg, The Structure of the Four-Way Junction in DNA, *Annu. Rev. Biophys. Biomol. Struct.* 22 (1993) 299–328. <https://doi.org/10.1146/annurev.bb.22.060193.001503>.
- [159] A. Wong, G. Wu, Selective Binding of Monovalent Cations to the Stacking G-Quartet Structure Formed by Guanosine 5'-Monophosphate: A Solid-State NMR Study, *J. Am. Chem. Soc.* 125 (2003) 13895–13905. <https://doi.org/10.1021/ja0302174>.
- [160] N.B. Leontis, P. Ghosh, P.B. Moore, Effect of Magnesium Ion on the Structure of the 5S RNA from *Escherichia coli*. An Imino Proton Magnetic Resonance Study of the Helix I, IV, and V Regions of the Molecule, *Biochemistry*. 25 (1986) 7386–7392. <https://doi.org/10.1021/bi00371a021>.
- [161] Y. Takezawa, M. Shionoya, Supramolecular DNA Three-Way Junction Motifs With a Bridging Metal Center, *Front. Chem.* 7 (2020) 925. <https://doi.org/10.3389/fchem.2019.00925>.

- [162] J. Wang, B. Liu, Highly sensitive and selective detection of Hg²⁺ in aqueous solution with mercury-specific DNA and Sybr Green I, *Chem. Commun.* (2008) 4759–4761. <https://doi.org/10.1039/b806885b>.
- [163] A. Ono, S. Cao, H. Togashi, M. Tashiro, T. Fujimoto, T. MacHinami, S. Oda, Y. Miyake, I. Okamoto, Y. Tanaka, Specific interactions between silver(i) ions and cytosine-cytosine pairs in DNA duplexes, *Chem. Commun.* 0 (2008) 4825–4827. <https://doi.org/10.1039/b808686a>.
- [164] J.M. DeMoor, D.J. Koropatnick, Metals and cellular signaling in mammalian cells., *Cell. Mol. Biol. (Noisy-Le-Grand)*. 46 (2000) 367–381.
- [165] Y. Pirahanchi, N.R. Aeddula, Physiology, Sodium Potassium Pump (Na⁺ K⁺ Pump), StatPearls Publishing, 2019. <http://www.ncbi.nlm.nih.gov/pubmed/30725773> (accessed January 28, 2021).
- [166] K. Degtyarenko, Metalloproteins, in: *Encycl. Genet. Genomics, Proteomics Bioinforma.*, John Wiley & Sons, Ltd, 2005. <https://doi.org/10.1002/047001153x.g306204>.
- [167] Z.L. He, X.E. Yang, P.J. Stoffella, Trace elements in agroecosystems and impacts on the environment, *J. Trace Elem. Med. Biol.* 19 (2005) 125–140. <https://doi.org/10.1016/j.jtemb.2005.02.010>.
- [168] N. Herawati, S. Suzuki, K. Hayashi, I.F. Rivai, H. Koyama, Cadmium, copper, and zinc levels in rice and soil of Japan, Indonesia, and China by soil type, *Bull. Environ. Contam. Toxicol.* 64 (2000) 33–39. <https://doi.org/10.1007/s001289910006>.
- [169] J. Li, H. Duan, P. Shi, Heavy metal contamination of surface soil in electronic waste dismantling area: Site investigation and source-apportionment analysis, *Waste Manag. Res.* 29 (2011) 727–738. <https://doi.org/10.1177/0734242X10397580>.
- [170] S. Shallari, C. Schwartz, A. Haskob, J.L. Morelat, S. Shalan, Heavy metals in soils and plants of serpentine and industrial sites of Albania, 1998.
- [171] W. Zhou, R. Saran, J. Liu, Metal Sensing by DNA, *Chem. Rev.* 117 (2017) 8272–8325. <https://doi.org/10.1021/acs.chemrev.7b00063>.
- [172] W. Chiuman, Y. Li, Efficient signaling platforms built from a small catalytic DNA and doubly labeled fluorogenic substrates, *Nucleic Acids Res.* 35 (2007) 401–405. <https://doi.org/10.1093/nar/gkl1056>.
- [173] X.H. Zhao, R.M. Kong, X.B. Zhang, H.M. Meng, W.N. Liu, W. Tan, G.L. Shen, R.Q. Yu, Graphene-DNAzyme based biosensor for amplified fluorescence “turn-On” detection of Pb²⁺ with a high selectivity, *Anal. Chem.* 83 (2011) 5062–5066. <https://doi.org/10.1021/ac200843x>.
- [174] C.S. Wu, M.K. Khaing Oo, X. Fan, Highly sensitive multiplexed heavy metal detection using quantum-dot-labeled DNAzymes, *ACS Nano.* 4 (2010) 5897–5904. <https://doi.org/10.1021/nn1021988>.
- [175] X.B. Zhang, Z. Wang, H. Xing, Y. Xiang, Y. Lu, Catalytic and molecular beacons for amplified detection of metal ions and organic molecules with high sensitivity, *Anal. Chem.* 82 (2010) 5005–5011. <https://doi.org/10.1021/ac1009047>.
- [176] R.A. Reynolds, C.A. Mirkin, R.L. Letsinger, Homogeneous, nanoparticle-based quantitative colorimetric detection of oligonucleotides [13], *J. Am. Chem. Soc.* 122 (2000) 3795–3796. <https://doi.org/10.1021/ja000133k>.
- [177] Z. Wang, J.H. Lee, Y. Lu, Label-Free Colorimetric Detection of Lead Ions with a Nanomolar Detection Limit and Tunable Dynamic Range by using Gold Nanoparticles and DNAzyme, *Adv. Mater.* 20 (2008) 3263–3267. <https://doi.org/10.1002/adma.200703181>.
- [178] N. Kanayama, T. Takarada, M. Maeda, Rapid naked-eye detection of mercury ions based on non-crosslinking aggregation of double-stranded DNA-carrying gold nanoparticles, *Chem. Commun.* 47 (2011) 2077–2079. <https://doi.org/10.1039/c0cc05171c>.

- [179] Y. Xiao, A.A. Rowe, K.W. Plaxco, Electrochemical detection of parts-per-billion lead via an electrode-bound DNAzyme assembly, *J. Am. Chem. Soc.* 129 (2007) 262–263. <https://doi.org/10.1021/ja067278x>.
- [180] L. Shen, Z. Chen, Y. Li, S. He, S. Xie, X. Xu, Z. Liang, X. Meng, Q. Li, Z. Zhu, M. Li, X.C. Le, Y. Shao, Electrochemical DNAzyme sensor for lead based on amplification of DNA-Au bio-bar codes, *Anal. Chem.* 80 (2008) 6323–6328. <https://doi.org/10.1021/ac800601y>.
- [181] J. Zhuang, L. Fu, M. Xu, Q. Zhou, G. Chen, D. Tang, DNAzyme-based magneto-controlled electronic switch for picomolar detection of lead (II) coupling with DNA-based hybridization chain reaction, *Biosens. Bioelectron.* 45 (2013) 52–57. <https://doi.org/10.1016/j.bios.2013.01.039>.
- [182] F. Bouhedda, A. Autour, M. Ryckelynck, Light-up RNA aptamers and their cognate fluorogens: From their development to their applications, *Int. J. Mol. Sci.* 19 (2018). <https://doi.org/10.3390/ijms19010044>.
- [183] S. Neubacher, S. Hennig, RNA Structure and Cellular Applications of Fluorescent Light-Up Aptamers, *Angew. Chemie - Int. Ed.* 58 (2019) 1266–1279. <https://doi.org/10.1002/anie.201806482>.
- [184] Z. Darzynkiewicz, J. Kapuscinski, F. Traganos, H.A. Crissman, Application of pyronin Y(G) in cytochemistry of nucleic acids, *Cytometry.* 8 (1987) 138–145. <https://doi.org/10.1002/cyto.990080206>.
- [185] D.M. Damas-Souza, R. Nunes, H.F. Carvalho, An improved acridine orange staining of DNA/RNA, *Acta Histochem.* 121 (2019) 450–454. <https://doi.org/10.1016/j.acthis.2019.03.010>.
- [186] S. Chatterjee, G.S. Kumar, Binding of fluorescent acridine dyes acridine orange and 9-aminoacridine to hemoglobin: Elucidation of their molecular recognition by spectroscopy, calorimetry and molecular modeling techniques, *J. Photochem. Photobiol. B Biol.* 159 (2016) 169–178. <https://doi.org/10.1016/j.jphotobiol.2016.03.045>.
- [187] J.F. Milligan, D.R. Groebe, G.W. Witherell, O.C. Uhlenbeck, Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates, *Nucleic Acids Res.* 15 (1987) 8783–8798. <https://doi.org/10.1093/nar/15.21.8783>.
- [188] J.F. Milligan, O.C. Uhlenbeck, Synthesis of small RNAs using T7 RNA polymerase, *Methods Enzymol.* 180 (1989) 51–62. [https://doi.org/10.1016/0076-6879\(89\)80091-6](https://doi.org/10.1016/0076-6879(89)80091-6).
- [189] D. Marion, P.C. Driscoll, L.E. Kay, P.T. Wingfield, A. Bax, A.M. Gronenborn, G.M. Clore, Overcoming the Overlap Problem in the Assignment of ¹H NMR Spectra of Larger Proteins by Use of Three-Dimensional Heteronuclear ¹H-¹⁵N Hartmann-Hahn-Multiple Quantum Coherence and Nuclear Overhauser-Multiple Quantum Coherence Spectroscopy: Application to Interleukin 1 β , *Biochemistry.* 28 (1989) 6150–6156. <https://doi.org/10.1021/bi00441a004>.
- [190] V. Sklenář, R. Tschudin, A. Bax, Water suppression using a combination of hard and soft pulses, *J. Magn. Reson.* 75 (1987) 352–357. [https://doi.org/10.1016/0022-2364\(87\)90041-2](https://doi.org/10.1016/0022-2364(87)90041-2).
- [191] A. Kumar, R.R. Ernst, K. Wüthrich, A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules, *Top. Catal.* 95 (1980) 1–6. [https://doi.org/10.1016/0006-291X\(80\)90695-6](https://doi.org/10.1016/0006-291X(80)90695-6).
- [192] D.I. Hoult, Solvent peak saturation with single phase and quadrature fourier transformation, *J. Magn. Reson.* 21 (1976) 337–347. [https://doi.org/10.1016/0022-2364\(76\)90081-0](https://doi.org/10.1016/0022-2364(76)90081-0).
- [193] J. Briand, R.R. Ernst, Computer-optimized homonuclear TOCSY experiments with suppression of cross relaxation, *Chem. Phys. Lett.* 185 (1991) 276–285. [https://doi.org/10.1016/S0009-2614\(91\)85060-A](https://doi.org/10.1016/S0009-2614(91)85060-A).
- [194] A. Savitzky, M.J.E. Golay, Smoothing and Differentiation of Data by Simplified Least Squares Procedures, *Anal. Chem.* 36 (1964) 1627–1639. <https://doi.org/10.1021/ac60214a047>.
- [195] J. Carvalho, A. Paiva, M.P. Cabral Campello, A. Paulo, J.L. Mergny, G.F. Salgado, J.A. Queiroz, C. Cruz, Aptamer-based Targeted Delivery of a G-quadruplex Ligand in Cervical Cancer Cells, *Sci. Rep.* 9 (2019) 7945. <https://doi.org/10.1038/s41598-019-44388-9>.

- [196] G.R. Bishop, J.B. Chaires, Characterization of DNA Structures by Circular Dichroism, in: *Curr. Protoc. Nucleic Acid Chem.*, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2002: pp. 7.11.1-7.11.8. <https://doi.org/10.1002/0471142700.nc0711s11>.
- [197] M. Kejnovska, I. Renciuik, D. Palacky, J. Vorlickova, G-Quadruplex Nucleic Acids: Methods and Protocols, in: C. Yang, D. Lin (Ed.), *Humana Press*, New York, 2019: pp. 25–44.
- [198] J. Feigon, K.M. Koshlap, F.W. Smith, ¹H NMR spectroscopy of DNA triplexes and quadruplexes, *Methods Enzymol.* 261 (1995) 225–255. [https://doi.org/10.1016/S0076-6879\(95\)61012-X](https://doi.org/10.1016/S0076-6879(95)61012-X).
- [199] M. Adrian, B. Heddi, A.T. Phan, NMR spectroscopy of G-quadruplexes, *Methods.* 57 (2012) 11–24. <https://doi.org/10.1016/j.ymeth.2012.05.003>.
- [200] A.T. Brünger, X-ray crystallography and NMR reveal complementary views of structure and dynamics, *Nat. Struct. Biol.* 4 (1997) 862–865. <https://europepmc.org/article/med/9377160> (accessed March 24, 2021).
- [201] A. Schirò, A. Carlon, G. Parigi, G. Murshudov, V. Calderone, E. Ravera, C. Luchinat, On the complementarity of X-ray and NMR data, *J. Struct. Biol. X.* 4 (2020) 100019. <https://doi.org/10.1016/j.yjsbx.2020.100019>.
- [202] W. Zhang, J.W. Szostak, Z. Huang, Nucleic acid crystallization and X-ray crystallography facilitated by single selenium atom, (n.d.). <https://doi.org/10.1007/s11705-016-1565-3>.
- [203] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The Protein Data Bank, *Nucleic Acids Res.* 28 (2000) 235–242. <https://doi.org/10.1093/nar/28.1.235>.
- [204] E. Callaway, Revolutionary cryo-EM is taking over structural biology, *Nature.* 578 (2020) 201. <https://doi.org/10.1038/d41586-020-00341-9>.
- [205] T.L. James, V. Dotsch, U. Schmitz, *Methods in Enzymology: Nuclear Magnetic Resonance of Biological Macromolecules Part A*, Volume 338, Academic Press, San Diego, California, 2001.
- [206] A.W. Overhauser, Polarization of nuclei in metals, *Phys. Rev.* 92 (1953) 411–415. <https://doi.org/10.1103/PhysRev.92.411>.
- [207] P. Lundström, Nuclear Overhauser Effect, in: *Encycl. Biophys.*, Springer Berlin Heidelberg, 2013: pp. 1753–1759. https://doi.org/10.1007/978-3-642-16712-6_308.
- [208] Kurt Wuthrich, *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, Inc., New York, NY, 1986.
- [209] A.E. Derome, M.P. Williamson, Rapid-pulsing artifacts in double-quantum-filtered COSY, *J. Magn. Reson.* 88 (1990) 177–185. [https://doi.org/10.1016/0022-2364\(90\)90123-Q](https://doi.org/10.1016/0022-2364(90)90123-Q).
- [210] V. Sklenář, A. Bax, Spin-echo water suppression for the generation of pure-phase two-dimensional NMR spectra, *J. Magn. Reson.* 74 (1987) 469–479. [https://doi.org/10.1016/0022-2364\(87\)90269-1](https://doi.org/10.1016/0022-2364(87)90269-1).
- [211] V. Sklenář, H. Miyashiro, G. Zon, H. Todd Miles, A. Bax, Assignment of the ³¹P and ¹H resonances in oligonucleotides by two-dimensional NMR spectroscopy, *FEBS Lett.* 208 (1986) 94–98. [https://doi.org/10.1016/0014-5793\(86\)81539-3](https://doi.org/10.1016/0014-5793(86)81539-3).
- [212] A.G. Palmer, J. Cavanagh, P.E. Wright, M. Rance, Sensitivity improvement in proton-detected two-dimensional heteronuclear correlation NMR spectroscopy, *J. Magn. Reson.* 93 (1991) 151–170. [https://doi.org/10.1016/0022-2364\(91\)90036-S](https://doi.org/10.1016/0022-2364(91)90036-S).
- [213] A.L. Davis, J. Keeler, E.D. Laue, D. Moskau, Experiments for recording pure-absorption heteronuclear correlation spectra using pulsed field gradients, *J. Magn. Reson.* 98 (1992) 207–216. [https://doi.org/10.1016/0022-2364\(92\)90126-R](https://doi.org/10.1016/0022-2364(92)90126-R).
- [214] A.J. Dingley, S. Grzesiek, Direct observation of hydrogen bonds in nucleic acid base pairs by internucleotide ²J(NN) couplings, *J. Am. Chem. Soc.* 120 (1998) 8293–8297. <https://doi.org/10.1021/ja981513x>.

- [215] J. Wöhnert, R. Ramachandran, M. Görlach, L.R. Brown, Triple-Resonance Experiments for Correlation of H5 and Exchangeable Pyrimidine Base Hydrogens in ¹³C, ¹⁵N-Labeled RNA, *J. Magn. Reson.* 139 (1999) 430–433. <https://doi.org/10.1006/jmre.1999.1797>.
- [216] B.T. Farmer Ip', L. Mtiler, E.P. Nikonowicz, A. Pardi, Unambiguous through-bond sugar-to-base correlations for purines in ¹³C,¹⁵N-labeled nucleic acids" The HsCsNb, HsCs(N)bCb, and HbNbC b experiments, 1994.
- [217] S.I. Tate, A. Ono, M. Kainosho, An Alternative Triple-Resonance Method for the Through-Bond Correlation of Intranucleotide H1' and H8 NMR Signals of Purine Nucleotides. Application to a DNA Dodecamer with Fully ¹³C/¹⁵N-Labeled Deoxyadenosine Residues, *J. Am. Chem. Soc.* 116 (1994) 5977–5978. <https://doi.org/10.1021/ja00092a062>.
- [218] W. Hu, Y.Q. Gosser, W. Xu, D.J. Patel, Novel 2D and 3D multiple-quantum bi-directional HCNC experiments for the correlation of ribose and base protons/carbons in ¹³C/¹⁵N labeled RNA, *J. Biomol. NMR.* 20 (2001) 167–172. <https://doi.org/10.1023/A:1011226415824>.
- [219] M.J.P. Van Dongen, S.S. Wijmenga, R. Eritja, F. Azorin, C.W. Hilbers, Through-bond correlation of adenine H2 and H8 protons in unlabeled DNA fragments by HMBC spectroscopy, *J. Biomol. NMR.* 8 (1996) 207–212. <https://doi.org/10.1007/BF00211166>.
- [220] S.S. Wijmenga A', H.A. Heus, H.A.E. Leeuw, H. Hoppe, M. Van Der G, C.W. Hilbers, Sequential backbone assignment of uniformly ¹³C-labeled RNAs by a two-dimensional P(CC)H-TOCSY triple resonance NMR experiment, Bax and Ikura, 1995.
- [221] H.A. Heus, F.J.M. van de Ven, S.S. Wijmenga, C.W. Hilbers, Sequential Backbone Assignment in ¹³C-Labeled RNA via Through-Bond Coherence Transfer Using Three-Dimensional Triple Resonance Spectroscopy (¹H, ¹³C, ³¹P) and Two-Dimensional Hetero TOCSY, *J. Am. Chem. Soc.* 116 (1994) 4983–4984. <https://doi.org/10.1021/ja00090a052>.
- [222] G. Varani, F. Aboul-ela, F. Allain, C.C. Gubser, Novel three-dimensional ¹H-¹³C-³¹P triple resonance experiments for sequential backbone correlations in nucleic acids, *J. Biomol. NMR.* 5 (1995) 315–320. <https://doi.org/10.1007/BF00211759>.
- [223] J.P. Marino, H. Schwalbe, C. Anklin, W. Bermel, D.M. Crothers, C. Griesinger, A Three-Dimensional Triple-Resonance *H,¹³C,³¹P Experiment: Sequential Through-Bond Correlation of Ribose Protons and Intervening Phosphorus along the RNA Oligonucleotide Backbone Two consecutive •H-,³C and ¹³C-³¹P INEPT and reverse INEPT transfer steps17, 1994. <https://pubs.acs.org/sharingguidelines> (accessed April 4, 2021).
- [224] J.P. Marino, J.H. Prestegard, D.M. Crothers, Uniformly ¹³C Labeled RNAs by 2D HCCH-TOCSY: A New Tool for Assignment, Gronenborn, A. M. Prog. NMR Spectrosc, 1994. <https://pubs.acs.org/sharingguidelines> (accessed April 4, 2021).
- [225] L.E. Kay, G.Y. Xu, A.U. Singer, D.R. Muhandiram, J.D. Formankay, A Gradient-Enhanced HCCH-TOCSY Experiment for Recording Side-Chain ¹H and ¹³C Correlations in H₂O Samples of Proteins, *J. Magn. Reson. Ser. B.* 101 (1993) 333–337. <https://doi.org/10.1006/jmrb.1993.1053>.
- [226] L.E. Kay, M. Ikura, A. Bax, Proton-Proton Correlation via Carbon-Carbon Couplings: A Three-Dimensional NMR Approach for the Assignment of Aliphatic Resonances in Proteins Labeled with Carbon-13, *J. Am. Chem. Soc.* 112 (1990) 888–889. <https://doi.org/10.1021/ja00158a070>.
- [227] R.P. Barnwal, F. Yang, G. Varani, Applications of NMR to structure determination of RNAs large and small, *Arch. Biochem. Biophys.* 628 (2017) 42–56. <https://doi.org/10.1016/j.abb.2017.06.003>.
- [228] J.A. Rabi, J.J. Fox, Nucleosides. LXXIX. Facile Base-Catalyzed Hydrogen Isotope Labeling at Position 6 of Pyrimidine Nucleosides, *J. Am. Chem. Soc.* 95 (1973) 1628–1632. <https://doi.org/10.1021/ja00786a044>.
- [229] P.M. Cullis, M.E. Malone, I.D. Podmore, M.C.R. Symons, Site of protonation of one-electron-reduced cytosine and its derivatives in aqueous methanol glasses, *J. Phys. Chem.* 99 (1995) 9293–9298. <https://doi.org/10.1021/j100022a051>.

- [230] X. Huang, P. Yu, E. Leproust, X. Gao, An efficient and economic site-specific deuteration strategy for NMR studies of homologous oligonucleotide repeat sequences, *Nucleic Acids Res.* 25 (1997) 4758–4763. <https://doi.org/10.1093/nar/25.23.4758>.
- [231] K. Lu, Y. Miyazaki, M.F. Summers, Isotope labeling strategies for NMR studies of RNA, *J. Biomol. NMR.* 46 (2010) 113–125. <https://doi.org/10.1007/s10858-009-9375-2>.
- [232] M. Ikura, A. Bax, Isotope-Filtered 2D NMR of a Protein-Peptide Complex: Study of a Skeletal Muscle Myosin Light Chain Kinase Fragment Bound to Calmodulin, *J. Am. Chem. Soc.* 114 (1992) 2433–2440. <https://doi.org/10.1021/ja00033a019>.
- [233] K. Ogura, H. Terasawa, F. Inagaki, An improved double-tuned and isotope-filtered pulse scheme based on a pulsed field gradient and a wide-band inversion shaped pulse, *J. Biomol. NMR.* 8 (1996) 492–498. <https://doi.org/10.1007/BF00228150>.
- [234] M. Piotto, V. Saudek, V. Sklenář, Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions, *J. Biomol. NMR.* 2 (1992) 661–665. <https://doi.org/10.1007/BF02192855>.
- [235] T. Parella, Pulsed field gradients: A new tool for routine NMR, *Magn. Reson. Chem.* 36 (1998) 467–495. [https://doi.org/10.1002/\(SICI\)1097-458X\(199807\)36:7<467::AID-OMR325>3.0.CO;2-S](https://doi.org/10.1002/(SICI)1097-458X(199807)36:7<467::AID-OMR325>3.0.CO;2-S).
- [236] M. Zuker, Mfold web server for nucleic acid folding and hybridization prediction, *Nucleic Acids Res.* 31 (2003) 3406–3415. <https://doi.org/10.1093/nar/gkg595>.
- [237] D.E. Wemmer, Homonuclear correlated spectroscopy (COSY): The basics of two-dimensional NMR, *Concepts Magn. Reson.* 1 (1989) 59–72. <https://doi.org/10.1002/cmr.1820010204>.
- [238] U. Nagaswamy, X. Gao, S.A. Martinis, G.E. Fox, NMR structure of a ribosomal RNA hairpin containing a conserved CUCAA pentaloop, *Nucleic Acids Res.* 29 (2001) 5129–5139. <https://doi.org/10.1093/nar/29.24.5129>.
- [239] M.G. AbouHaidar, I.G. Ivanov, Non-enzymatic RNA hydrolysis promoted by the combined catalytic activity of buffers and magnesium ions, *Zeitschrift Fur Naturforsch. - Sect. C J. Biosci.* 54 (1999) 542–548. <https://doi.org/10.1515/znc-1999-7-813>.
- [240] C.T. Martin, D.K. Muller, J.E. Coleman, Processivity in early stages of transcription by T7 RNA polymerase, *Biochemistry.* 27 (1988) 3966–3974. <https://doi.org/10.1021/BI00411A012>.
- [241] S. Moroney, J. Piccirilli, Abortive products as initiating nucleotides during transcription by T7 RNA polymerase, *Biochemistry.* 30 (1991) 10343–10349. <https://doi.org/10.1021/BI00106A036>.
- [242] J.A. Pleiss, M.L. Derrick, O.C. Uhlenbeck, T7 RNA polymerase produces 5' end heterogeneity during in vitro transcription from certain templates., *RNA.* 4 (1998) 1313. <https://doi.org/10.1017/S135583829800106X>.
- [243] M. Helm, H. Brulé, R. Giegé, C. Florentz, More mistakes by T7 RNA polymerase at the 5' ends of in vitro-transcribed RNAs., *RNA.* 5 (1999) 618. <https://doi.org/10.1017/S1355838299982328>.
- [244] L.E. Macdonald, Y. Zhou, W.T. McAllister, Termination and slippage by bacteriophage T7 RNA polymerase, *J. Mol. Biol.* 232 (1993) 1030–1047. <https://doi.org/10.1006/JMBI.1993.1458>.
- [245] G. Krupp, RNA synthesis: strategies for the use of bacteriophage RNA polymerases, *Gene.* 72 (1988) 75–89. [https://doi.org/10.1016/0378-1119\(88\)90129-1](https://doi.org/10.1016/0378-1119(88)90129-1).
- [246] A. Salvail-Lacoste, G. Di Tomasso, B.L. Piette, P. Legault, Affinity purification of T7 RNA transcripts with homogeneous ends using ARIBo and CRISPR tags, *RNA.* 19 (2013) 1003. <https://doi.org/10.1261/RNA.037432.112>.
- [247] C. Kao, M. Zheng, S. Rüdiger, A simple and efficient method to reduce nontemplated nucleotide addition at the 3' terminus of RNAs transcribed by T7 RNA polymerase., *RNA.* 5 (1999) 1268.

<https://doi.org/10.1017/S1355838299991033>.

- [248] U. Das, S. Shuman, Mechanism of RNA 2',3'-cyclic phosphate end healing by T4 polynucleotide kinase-phosphatase, *Nucleic Acids Res.* 41 (2013) 355–365. <https://doi.org/10.1093/NAR/GKS977>.
- [249] O.B. Becette, G. Zong, B. Chen, K.M. Taiwo, D.A. Case, T.K. Dayie, Solution NMR readily reveals distinct structural folds and interactions in doubly ¹³C- And ¹⁹F-labeled RNAs, *Sci. Adv.* 6 (2020) eabc6572. <https://doi.org/10.1126/sciadv.abc6572>.
- [250] S. Ma, Y. Huang, R.B. Van Huystee, Improved plant RNA stability in storage, *Anal. Biochem.* 326 (2004) 122–124. <https://doi.org/10.1016/j.ab.2003.10.026>.
- [251] A.L. Fabre, M. Colotte, A. Luis, S. Tuffet, J. Bonnet, An efficient method for long-term room temperature storage of RNA, *Eur. J. Hum. Genet.* 22 (2014) 379–385. <https://doi.org/10.1038/ejhg.2013.145>.
- [252] R.E. Farrell, Creating a Ribonuclease-Free Environment, in: *RNA Methodol.*, Elsevier, 2017: pp. 55–74. <https://doi.org/10.1016/b978-0-12-804678-4.00002-6>.
- [253] G.E. Hernandez, T.S. Mondala, S.R. Head, Assessing a novel room-temperature RNA storage medium for compatibility in microarray gene expression analysis, *Biotechniques.* 47 (2009) 667–670. <https://doi.org/10.2144/000113209>.
- [254] H. Nielsen, Working with RNA., *Methods Mol. Biol.* 703 (2011) 15–28. https://doi.org/10.1007/978-1-59745-248-9_2.
- [255] L.E. Kay, Protein dynamics from NMR, *Biochem. Cell Biol.* 76 (1998) 145–152. <https://doi.org/10.1139/bcb-76-2-3-145>.
- [256] M. Ziegeler, M. Cevec, C. Richter, H. Schwalbe, NMR Studies of HAR1 RNA Secondary Structures Reveal Conformational Dynamics in the Human RNA, *ChemBioChem.* 13 (2012) 2100–2112. <https://doi.org/10.1002/cbic.201200401>.
- [257] S.C. Keane, X. Heng, K. Lu, S. Kharytonchyk, V. Ramakrishnan, G. Carter, S. Barton, A. Hoscic, A. Florwick, J. Santos, N.C. Bolden, S. McCowin, D.A. Case, B.A. Johnson, M. Salemi, A. Telesnitsky, M.F. Summers, Structure of the HIV-1 RNA packaging signal, *Science* (80-.). 348 (2015) 917–921. <https://doi.org/10.1126/science.aaa9266>.
- [258] S. Imai, P. Kumar, C.U.T. Hellen, V.M. D'Souza, G. Wagner, An accurately preorganized IRES RNA structure enables eIF4G capture for initiation of viral translation, *Nat. Struct. Mol. Biol.* 23 (2016) 859–864. <https://doi.org/10.1038/nsmb.3280>.
- [259] P. Dagenais, N. Girard, E. Bonneau, P. Legault, Insights into RNA structure and dynamics from recent NMR and X-ray studies of the *Neurospora Varkud* satellite ribozyme, *Wiley Interdiscip. Rev. RNA.* 8 (2017) 1421. <https://doi.org/10.1002/wrna.1421>.
- [260] C.R. Woese, R. Gutell, R. Gupta, H.F. Noller, Detailed analysis of the higher-order structure of 16S-like ribosomal ribonucleic acids, *Microbiol. Rev.* 47 (1983) 621–669. <https://doi.org/10.1128/mbr.47.4.621-669.1983>.
- [261] V.P. Antao, S.Y. Lai, I. Tinoco, A thermodynamic study of unusually stable RNA and DNA hairpins, n.d.
- [262] C. Tuerk, P. Gauss, C. Thermes, D.R. Groebe, M. Gayle, N. Guild, G. Stormo, Y. d'Aubenton-Carafa, O.C. Uhlenbeck, I. Tinoco, E.N. Brody, L. Gold, CUUCGG hairpins: extraordinarily stable RNA secondary structures associated with various biochemical processes., *Proc. Natl. Acad. Sci. U. S. A.* 85 (1988) 1364–1368. <https://doi.org/10.1073/pnas.85.5.1364>.
- [263] J.H. Davis, M. Tonelli, L.G. Scott, L. Jaeger, J.R. Williamson, S.E. Butcher, RNA helical packing in solution: NMR structure of a 30 kDa GAAA tetraloop-receptor complex, *J. Mol. Biol.* 351 (2005) 371–382. <https://doi.org/10.1016/j.jmb.2005.05.069>.
- [264] K. Valegård, J.B. Murray, N.J. Stonehouse, S. Van Den Worm, P.G. Stockley, L. Liljas, The three-dimensional structures of two complexes between recombinant MS2 capsids and RNA operator fragments reveal sequence-specific protein-RNA interactions, *J. Mol. Biol.* 270 (1997) 724–738. <https://doi.org/10.1006/jmbi.1997.1144>.

- [265] K.B. Hall, Mighty tiny, *RNA*. 21 (2015) 630–631. <https://doi.org/10.1261/rna.050567.115>.
- [266] G. Varani, C. Cheong, I. Tinoco, Structure of an Unusually Stable RNA Hairpin, *Biochemistry*. 30 (1991) 3280–3289. <https://doi.org/10.1021/bi00227a016>.
- [267] E. Ennifar, A. Nikulin, S. Tishchenko, A. Serganov, N. Nevskaya, M. Garber, B. Ehresmann, C. Ehresmann, S. Nikonov, P. Dumas, The crystal structure of UUCG tetraloop, *J. Mol. Biol.* 304 (2000) 35–42. <https://doi.org/10.1006/jmbi.2000.4204>.
- [268] S. Nozinovic, B. Fürtig, H.R.A. Jonker, C. Richter, H. Schwalbe, High-resolution NMR structure of an RNA model system: The 14-mer cUUCGg tetraloop hairpin RNA, *Nucleic Acids Res.* 38 (2009) 683–694. <https://doi.org/10.1093/nar/gkp956>.
- [269] F.H.T. Allain, G. Varani, Structure of the P1 helix from group I self-splicing introns, *J. Mol. Biol.* 250 (1995) 333–353. <https://doi.org/10.1006/jmbi.1995.0381>.
- [270] S.H. Bae, H.K. Cheong, J.H. Lee, C. Cheong, M. Kainosho, B.S. Choi, Structural features of an influenza virus promoter and their implications for viral RNA synthesis, *Proc. Natl. Acad. Sci. U. S. A.* 98 (2001) 10602–10607. <https://doi.org/10.1073/pnas.191268798>.
- [271] B. Fürtig, C. Richter, W. Bermel, H. Schwalbe, New NMR experiments for RNA nucleobase resonance assignment and chemical shift analysis of an RNA UUCG tetraloop, *J. Biomol. NMR.* 28 (2004) 69–79. <https://doi.org/10.1023/B:JNMR.0000012863.63522.1f>.
- [272] P.J. Nichols, M.A. Henen, A. Born, D. Strotz, P. Güntert, B. Vögeli, High-resolution small RNA structures from exact nuclear Overhauser enhancement measurements without additional restraints, *Commun. Biol.* 1 (2018). <https://doi.org/10.1038/s42003-018-0067-x>.
- [273] S. Bottaro, P.J. Nichols, B. Vögeli, M. Parrinello, K. Lindorff-Larsen, Integrating NMR and simulations reveals motions in the UUCG tetraloop, *Nucleic Acids Res.* 48 (2020) 5839–5848. <https://doi.org/10.1093/NAR/GKAA399>.
- [274] A.T. Brünger, P.D. Adams, G.M. Clore, W.L. Delano, P. Gros, R.W. Grossekunstleve, J.S. Jiang, J. Kuszewski, M. Nilges, N.S. Pannu, R.J. Read, L.M. Rice, T. Simonson, G.L. Warren, Crystallography & NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr. Sect. D Biol. Crystallogr.* 54 (1998) 905–921. <https://doi.org/10.1107/S0907444998003254>.
- [275] A.T. Brunger, Version 1.2 of the crystallography and nmr system, *Nat. Protoc.* 2 (2007) 2728–2733. <https://doi.org/10.1038/nprot.2007.406>.
- [276] R.A. Sayle, E.J. Milner-White, RASMOL: biomolecular graphics for all, *Trends Biochem. Sci.* 20 (1995) 374–376. [https://doi.org/10.1016/S0968-0004\(00\)89080-5](https://doi.org/10.1016/S0968-0004(00)89080-5).
- [277] D. Imburgio, M. Rong, K. Ma, W.T. McAllister, Studies of promoter recognition and start site selection by T7 RNA polymerase using a comprehensive collection of promoter variants, *Biochemistry*. 39 (2000) 10419–10430. <https://doi.org/10.1021/bi000365w>.
- [278] L.G. Briebe, R. Padilla, R. Sousa, Role of T7 RNA polymerase His784 in start site selection and initial transcription, *Biochemistry*. 41 (2002) 5144–5149. <https://doi.org/10.1021/bi016057v>.
- [279] I. Kuzmine, P.A. Gottlieb, C.T. Martin, Binding of the priming nucleotide in the initiation of transcription by T7 RNA polymerase, *J. Biol. Chem.* 278 (2003) 2819–2823. <https://doi.org/10.1074/jbc.M208405200>.
- [280] T. Conrad, I. Plumbom, M. Alcobendas, R. Vidal, S. Sauer, Maximizing transcription of nucleic acids with efficient T7 promoters, *Commun. Biol.* 3 (2020) 1–8. <https://doi.org/10.1038/s42003-020-01167-x>.
- [281] X.B. Zhang, R.M. Kong, Y. Lu, Metal ion sensors based on DNAzymes and related DNA molecules, *Annu. Rev. Anal. Chem.* 4 (2011) 105–128. <https://doi.org/10.1146/annurev.anchem.111808.073617>.

- [282] E.R. Jamieson, S.J. Lippard, Structure, recognition, and processing of cisplatin-DNA adducts, *Chem. Rev.* 99 (1999) 2467–2498. <https://doi.org/10.1021/cr980421n>.
- [283] T. Oh, S.S. Park, C.A. Mirkin, Stabilization of Colloidal Crystals Engineered with DNA, *Adv. Mater.* 31 (2019) 1805480. <https://doi.org/10.1002/adma.201805480>.
- [284] R.K.O. Sigel, Intimate Relationships between Metal Ions and Nucleic Acids, *Angew. Chemie Int. Ed.* 46 (2007) 654–656. <https://doi.org/10.1002/anie.200605065>.
- [285] E. Largy, J.L. Mergny, V. Gabelica, Role of Alkali Metal Ions in G-Quadruplex Nucleic Acid Structure and Stability, in: *Met. Ions Life Sci.*, Walter de Gruyter GmbH, 2016: pp. 203–258. https://doi.org/10.1007/978-3-319-21756-7_7.
- [286] B. Saccà, L. Lacroix, J.L. Mergny, The effect of chemical modifications on the thermal stability of different G-quadruplex-forming oligonucleotides, *Nucleic Acids Res.* 33 (2005) 1182–1192. <https://doi.org/10.1093/nar/gki257>.
- [287] N. V. Hud, F.W. Smith, F.A.L. Anet, J. Feigon, The selectivity for K⁺ versus Na⁺ in DNA quadruplexes is dominated by relative free energies of hydration: A thermodynamic analysis by 1H NMR, *Biochemistry.* 35 (1996) 15383–15390. <https://doi.org/10.1021/bi9620565>.
- [288] W. Zhou, Y. Zhang, P.J.J. Huang, J. Ding, J. Liu, A DNAzyme requiring two different metal ions at two distinct sites, *Nucleic Acids Res.* 44 (2016) 354–363. <https://doi.org/10.1093/nar/gkv1346>.
- [289] S.F. Torabi, Y. Lu, Identification of the Same Na⁺-Specific DNAzyme Motif from Two In Vitro Selections Under Different Conditions, *J. Mol. Evol.* 81 (2015) 225–234. <https://doi.org/10.1007/s00239-015-9715-7>.
- [290] S.F. Torabi, P. Wu, C.E. McGhee, L. Chen, K. Hwang, N. Zheng, J. Cheng, Y. Lu, In vitro selection of a sodium-specific DNAzyme and its application in intracellular sensing, *Proc. Natl. Acad. Sci. U. S. A.* 112 (2015) 5903–5908. <https://doi.org/10.1073/pnas.1420361112>.
- [291] P.J.J. Huang, J. Lin, J. Cao, M. Vazin, J. Liu, Ultrasensitive DNAzyme beacon for lanthanides and metal speciation, *Anal. Chem.* 86 (2014) 1816–1821. <https://doi.org/10.1021/ac403762s>.
- [292] L. Ma, S. Kartik, B. Liu, J. Liu, From general base to general acid catalysis in a sodium-specific DNAzyme by a guanine-to-adenine mutation, *Nucleic Acids Res.* 47 (2019) 8154–8162. <https://doi.org/10.1093/nar/gkz578>.
- [293] L. Ma, J. Liu, An in Vitro-Selected DNAzyme Mutant Highly Specific for Na⁺ under Slightly Acidic Conditions, *ChemBioChem.* 20 (2019) 537–542. <https://doi.org/10.1002/cbic.201800322>.
- [294] W. Zhou, R. Saran, J. Ding, J. Liu, Two Completely Different Mechanisms for Highly Specific Na⁺ Recognition by DNAzymes, *ChemBioChem.* 18 (2017) 1828–1835. <https://doi.org/10.1002/cbic.201700184>.
- [295] W. Zhou, J. Ding, J. Liu, A Selective Na⁺ Aptamer Dissected by Sensitized Tb³⁺ Luminescence, *ChemBioChem.* 17 (2016) 1563–1570. <https://doi.org/10.1002/cbic.201600174>.
- [296] Y. He, Y. Zhou, D. Chen, J. Liu, Global Folding of a Na⁺-Specific DNAzyme Studied by FRET, *ChemBioChem.* 20 (2019) 385–393. <https://doi.org/10.1002/cbic.201800548>.
- [297] S.J. Admiraal, D. Herschlag, Mapping the transition state for ATP hydrolysis: implications for enzymatic catalysis, *Chem. Biol.* 2 (1995) 729–739. [https://doi.org/10.1016/1074-5521\(95\)90101-9](https://doi.org/10.1016/1074-5521(95)90101-9).
- [298] Y. He, D. Chen, P.J.J. Huang, Y. Zhou, L. Ma, K. Xu, R. Yang, J. Liu, Misfolding of a DNAzyme for ultrahigh sodium selectivity over potassium, *Nucleic Acids Res.* 46 (2018) 10262–10271. <https://doi.org/10.1093/nar/gky807>.
- [299] Y. He, Y. Chang, D. Chen, J. Liu, Probing Local Folding Allows Robust Metal Sensing Based on a Na⁺-Specific DNAzyme, *ChemBioChem.* 20 (2019) 2241–2247. <https://doi.org/10.1002/cbic.201900143>.
- [300] W. Zhou, J. Ding, J. Liu, A highly specific sodium aptamer probed by 2-Aminopurine for robust Na⁺ sensing, *Nucleic*

- Acids Res. 44 (2016) 10377–10385. <https://doi.org/10.1093/nar/gkw845>.
- [301] B. Liu, E.Y. Kelly, J. Liu, Cation-size-dependent DNA adsorption kinetics and packing density on gold nanoparticles: An opposite trend, *Langmuir*. 30 (2014) 13228–13234. <https://doi.org/10.1021/la503188h>.
- [302] C. Saintomé, S. Amrane, J.L. Mergny, P. Alberti, The exception that confirms the rule: A higher-order telomeric G-quadruplex structure more stable in sodium than in potassium, *Nucleic Acids Res.* 44 (2016) 2926–2935. <https://doi.org/10.1093/nar/gkw003>.
- [303] C.S. Mekmaysy, L. Petraccone, N.C. Garbett, P.A. Ragazzon, R. Gray, J.O. Trent, J.B. Chaires, Effect of O6-methylguanine on the stability of G-quadruplex DNA, *J. Am. Chem. Soc.* 130 (2008) 6710–6711. <https://doi.org/10.1021/ja801976h>.
- [304] P. Školáková, K. Bednářová, M. Vorlíčková, J. Sagi, Quadruplexes of human telomere dG3(TTAG3)3 sequences containing guanine abasic sites, *Biochem. Biophys. Res. Commun.* 399 (2010) 203–208. <https://doi.org/10.1016/j.bbrc.2010.07.055>.
- [305] J. Sagi, D. Renčiuk, M. Tomaško, M. Vorlíčková, Quadruplexes of human telomere DNA analogs designed to contain G:A:G:A, G:G:A:A, and A:A:A:A tetrads, *Biopolymers*. 93 (2010) 880–886. <https://doi.org/10.1002/bip.21481>.
- [306] H. Sun, J. Xiang, W. Gai, Y. Liu, A. Guan, Q. Yang, Q. Li, Q. Shang, H. Su, Y. Tang, G. Xu, Quantification of the Na⁺/K⁺ ratio based on the different response of a newly identified G-quadruplex to Na⁺ and K⁺, *Chem. Commun.* 49 (2013) 4510–4512. <https://doi.org/10.1039/c3cc39020a>.
- [307] F. Wang, Y. Han, C.S. Lim, Y. Lu, J. Wang, J. Xu, H. Chen, C. Zhang, M. Hong, X. Liu, Simultaneous phase and size control of upconversion nanocrystals through lanthanide doping, *Nature*. 463 (2010) 1061–1065. <https://doi.org/10.1038/nature08777>.
- [308] G. Haxel, *Rare Earth Elements: Critical Resources for High Technology*, 2002. https://books.google.ca/books?hl=en&lr=&id=sd5zZG3EqAsC&oi=fnd&pg=PP3&ots=rD7tXDUkT&sig=tHb-IUDGsIs3-mOrZQO0gZwyNTU&redir_esc=y#v=onepage&q&f=false (accessed March 16, 2021).
- [309] S. Kobayashi, M. Sugiura, H. Kitagawa, W.W.L. Lam, Rare-earth metal triflates in organic synthesis, *Chem. Rev.* 102 (2002) 2227–2302. <https://doi.org/10.1021/cr010289i>.
- [310] T.I. Kostelnik, C. Orvig, Radioactive Main Group and Rare Earth Metals for Imaging and Therapy, *Chem. Rev.* 119 (2019) 902–956. <https://doi.org/10.1021/acs.chemrev.8b00294>.
- [311] C. Evans, *Biochemistry of the Lanthanides*, Springer Science+Business Media, LLC, 1990. https://books.google.ca/books?hl=en&lr=&id=Jj6eBgAAQBAJ&oi=fnd&pg=PA1&ots=6uE1eMETk9&sig=_VyDAvoG72k74ndHixG2SDItHwO&redir_esc=y#v=onepage&q&f=false (accessed March 16, 2021).
- [312] H.-K. Kim, J. Li, N. Nagraj, Y. Lu, Probing Metal Binding in the 8-17 DNAzyme by Tb (III) Luminescence Spectroscopy, *Chem. - A Eur. J.* 14 (2008) 8696–8703. <https://doi.org/10.1002/chem.200701789>.
- [313] N. Sugimoto, T. Ohmichi, Site-specific cleavage reaction catalyzed by leadzyme is enhanced by combined effect of lead and rare earth ions, *FEBS Lett.* 393 (1996) 97–100. [https://doi.org/10.1016/0014-5793\(96\)00860-5](https://doi.org/10.1016/0014-5793(96)00860-5).
- [314] Y. Wang, S.K. Silverman, Deoxyribozymes that synthesize branched and lariat RNA, *J. Am. Chem. Soc.* 125 (2003) 6880–6881. <https://doi.org/10.1021/ja035150z>.
- [315] F. Javadi-Zarnaghi, C. Höbartner, Lanthanide cofactors accelerate DNA-catalyzed synthesis of branched RNA, *J. Am. Chem. Soc.* 135 (2013) 12839–12848. <https://doi.org/10.1021/ja406162z>.
- [316] V. Dokukin, S.K. Silverman, Lanthanide ions as required cofactors for DNA catalysts, *Chem. Sci.* 3 (2012) 1707–1714. <https://doi.org/10.1039/c2sc01067d>.
- [317] M. Komiyama, N. Takeda, H. Shigekawa, Hydrolysis of DNA and RNA by lanthanide ions: Mechanistic studies leading to

- new applications, *Chem. Commun.* 0 (1999) 1443–1451. <https://doi.org/10.1039/a901621j>.
- [318] F. Mancin, P. Scrimin, P. Tecilla, U. Tonellato, Artificial metallonucleases, *Chem. Commun.* (2005) 2540–2548. <https://doi.org/10.1039/b418164f>.
- [319] F. Yeasmin Khusbu, X. Zhou, H. Chen, C. Ma, K. Wang, Thioflavin T as a fluorescence probe for biosensing applications, *TrAC - Trends Anal. Chem.* 109 (2018) 1–18. <https://doi.org/10.1016/j.trac.2018.09.013>.
- [320] M. Webba da Silva, NMR methods for studying quadruplex nucleic acids, *Methods.* 43 (2007) 264–277. <https://doi.org/10.1016/j.ymeth.2007.05.007>.
- [321] J.L. Mergny, D. Sen, DNA quadruple helices in nanotechnology, *Chem. Rev.* 119 (2019) 6290–6325. <https://doi.org/10.1021/acs.chemrev.8b00629>.
- [322] H. Ueyama, M. Takagi, S. Takenaka, A novel potassium sensing in aqueous media with a synthetic oligonucleotide derivative. Fluorescence resonance energy transfer associated with guanine quartet-potassium ion complex formation, *J. Am. Chem. Soc.* 124 (2002) 14286–14287. <https://doi.org/10.1021/ja026892f>.
- [323] M. Vazin, P.J.J. Huang, Z. Matuszek, J. Liu, Biochemical Characterization of a Lanthanide-Dependent DNAzyme with Normal and Phosphorothioate-Modified Substrates, *Biochemistry.* 54 (2015) 6132–6138. <https://doi.org/10.1021/acs.biochem.5b00691>.
- [324] S. Liu, P. Peng, H. Wang, L. Shi, T. Li, Thioflavin T binds dimeric parallel-stranded GA-containing non-G-quadruplex DNAs: A general approach to lighting up double-stranded scaffolds, *Nucleic Acids Res.* 45 (2017) 12080–12089. <https://doi.org/10.1093/nar/gkx942>.
- [325] A.J. Guan, X.F. Zhang, X. Sun, Q. Li, J.F. Xiang, L.X. Wang, L. Lan, F.M. Yang, S.J. Xu, X.M. Guo, Y.L. Tang, Ethyl-substitutive Thioflavin T as a highly-specific fluorescence probe for detecting G-quadruplex structure, *Sci. Rep.* 8 (2018) 1–12. <https://doi.org/10.1038/s41598-018-20960-7>.
- [326] A.R. De La Faverie, A. Guédin, A. Bedrat, L.A. Yatsunyk, J.L. Mergny, Thioflavin T as a fluorescence light-up probe for G4 formation, *Nucleic Acids Res.* 42 (2014) e65. <https://doi.org/10.1093/nar/gku111>.
- [327] D. Zhao, X. Dong, N. Jiang, D. Zhang, C. Liu, Selective recognition of parallel and anti-parallel thrombin-binding aptamer G-quadruplexes by different fluorescent dyes, *Nucleic Acids Res.* 42 (2014) 11612–11621. <https://doi.org/10.1093/nar/gku833>.
- [328] R.R. Gao, T.M. Yao, X.Y. Lv, Y.Y. Zhu, Y.W. Zhang, S. Shi, Integration of G-quadruplex and DNA-templated Ag NCs for nonarithmetic information processing, *Chem. Sci.* 8 (2017) 4211–4222. <https://doi.org/10.1039/c7sc00361g>.
- [329] K.N. Luu, A.T. Phan, V. Kuryavyi, L. Lacroix, D.J. Patel, Structure of the human telomere in K⁺ solution: An intramolecular (3 + 1) G-quadruplex scaffold, *J. Am. Chem. Soc.* 128 (2006) 9963–9970. <https://doi.org/10.1021/ja062791w>.
- [330] B. Fürtig, C. Richter, J. Wöhnert, H. Schwalbe, NMR Spectroscopy of RNA, *ChemBioChem.* 4 (2003) 936–962. <https://doi.org/10.1002/cbic.200300700>.
- [331] D.J. Patel, L. Shapiro, D. Hare, DNA and RNA: NMR studies of conformations and dynamics in solution, *Q. Rev. Biophys.* 20 (1987) 35–112. <https://doi.org/10.1017/S0033583500004224>.
- [332] Y.-Y. Tseng, S.-H. Chou, Systematic NMR Assignments of DNA Exchangeable Protons, *J. Chinese Chem. Soc.* 46 (1999) 699–706. <https://doi.org/10.1002/jccs.199900096>.
- [333] P. Blancafort, S. V. Steinberg, B. Paquin, R. Klinck, J.K. Scott, R. Cedergren, The recognition of a noncanonical RNA base pair by a zinc finger protein, *Chem. Biol.* 6 (1999) 585–597. [https://doi.org/10.1016/S1074-5521\(99\)80091-X](https://doi.org/10.1016/S1074-5521(99)80091-X).
- [334] S. Amrane, M. Adrian, B. Heddi, A. Serero, A. Nicolas, J.L. Mergny, A.T. Phan, Formation of pearl-necklace

- monomeric G-quadruplexes in the human CEB25 minisatellite, *J. Am. Chem. Soc.* 134 (2012) 5807–5816. <https://doi.org/10.1021/ja208993r>.
- [335] A.T. Phan, Y.S. Modi, D.J. Patel, Two-repeat Tetrahymena telomeric d(TGGGGTTGGGGT) sequence interconverts between asymmetric dimeric G-quadruplexes in solution, *J. Mol. Biol.* 338 (2004) 93–102. <https://doi.org/10.1016/j.jmb.2004.02.042>.
- [336] X.A. Mao, L.A. Marky, W.H. Gmeiner, Nmr structure of the thrombin-binding dna aptamer stabilized by sr²⁺, *J. Biomol. Struct. Dyn.* 22 (2004) 25–33. <https://doi.org/10.1080/07391102.2004.10506977>.
- [337] M. Meyer, A. Hocquet, J. Sühnel, Interaction of sodium and potassium ions with sandwiched cytosine-, guanine-, thymine-, and uracil-base tetrads, *J. Comput. Chem.* 26 (2005) 352–364. <https://doi.org/10.1002/jcc.20176>.
- [338] C.C. Hardin, T. Watson, M. Corregan, C. Bailey, Cation-Dependent Transition between the Quadruplex and Watson-Crick Hairpin Forms of d(CGCG3GCG), *Biochemistry.* 31 (1992) 833–841. <https://doi.org/10.1021/bi00118a028>.
- [339] F.W. Smith, J. Feigon, Strand Orientation in the DNA Quadruplex Formed from the Oxytricha Telomere Repeat Oligonucleotide d(G4T4G4) in Solution, *Biochemistry.* 32 (1993) 8682–8692. <https://doi.org/10.1021/bi00084a040>.
- [340] M. Vorlíčková, I. Kejnovská, K. Bednářová, D. Renčíuk, J. Kypr, Circular dichroism spectroscopy of DNA: From duplexes to quadruplexes, in: *Chirality, Chirality*, 2012: pp. 691–698. <https://doi.org/10.1002/chir.22064>.
- [341] I. Kejnovská, D. Renčíuk, J. Palacký, M. Vorlíčková, CD Study of the G-Quadruplex Conformation, in: *Methods Mol. Biol., Humana Press Inc.*, 2019: pp. 25–44. https://doi.org/10.1007/978-1-4939-9666-7_2.
- [342] R. del Villar-Guerra, J.O. Trent, J.B. Chaires, G-Quadruplex Secondary Structure Obtained from Circular Dichroism Spectroscopy, *Angew. Chemie - Int. Ed.* 57 (2018) 7171–7175. <https://doi.org/10.1002/anie.201709184>.
- [343] A. Mammana, G.T. Carroll, B.L. Feringa, Circular Dichroism of Dynamic Systems: Switching Molecular and Supramolecular Chirality, in: *Compr. Chiroptical Spectrosc.*, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2012: pp. 289–316. <https://doi.org/10.1002/9781118120392.ch8>.

Appendices

Appendix A: Supplementary Figures

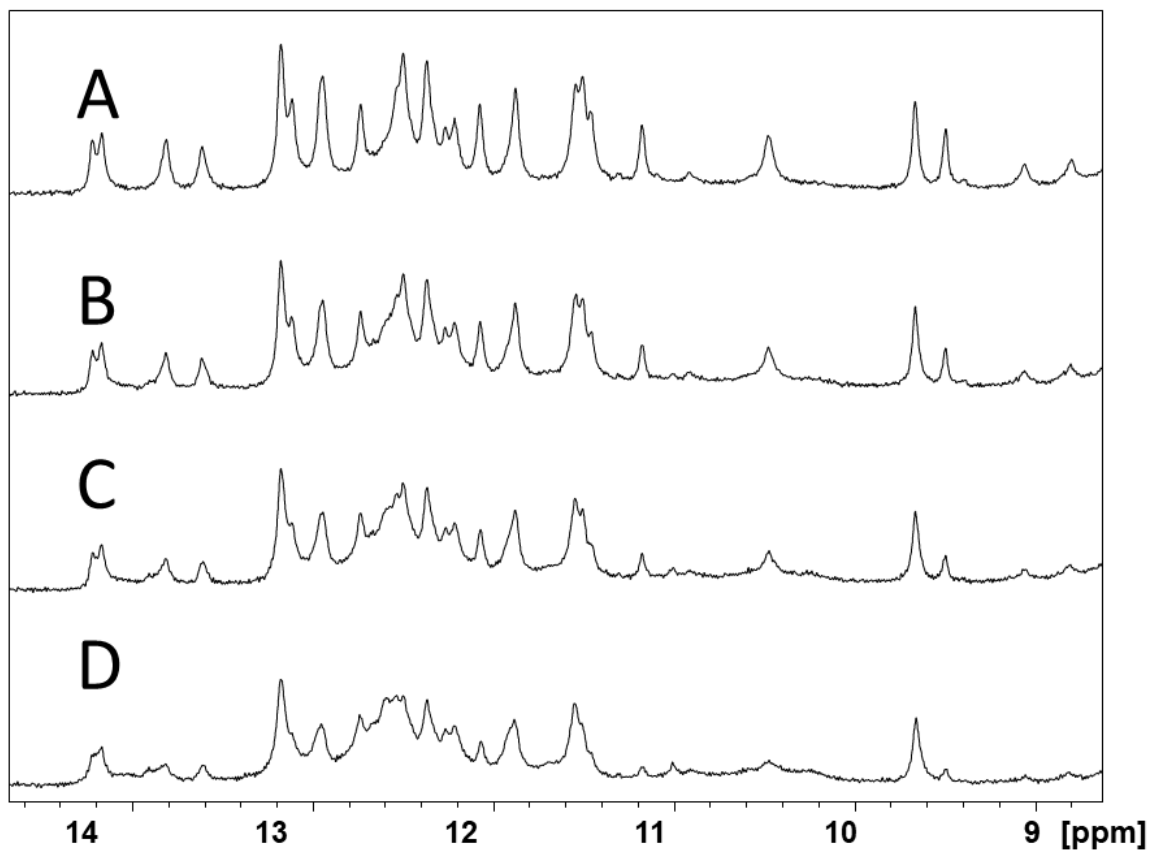


Figure A.1: 1D titration of 1 mM SRB-2 with SR.

Imino proton range showing change in structure upon binding of the ligand in three aliquots. A) SRB-2 fully bound, B) SRB-2 two-thirds bound, C) SRB-2 one-third bound, and D) SRB-2 completely unbound.

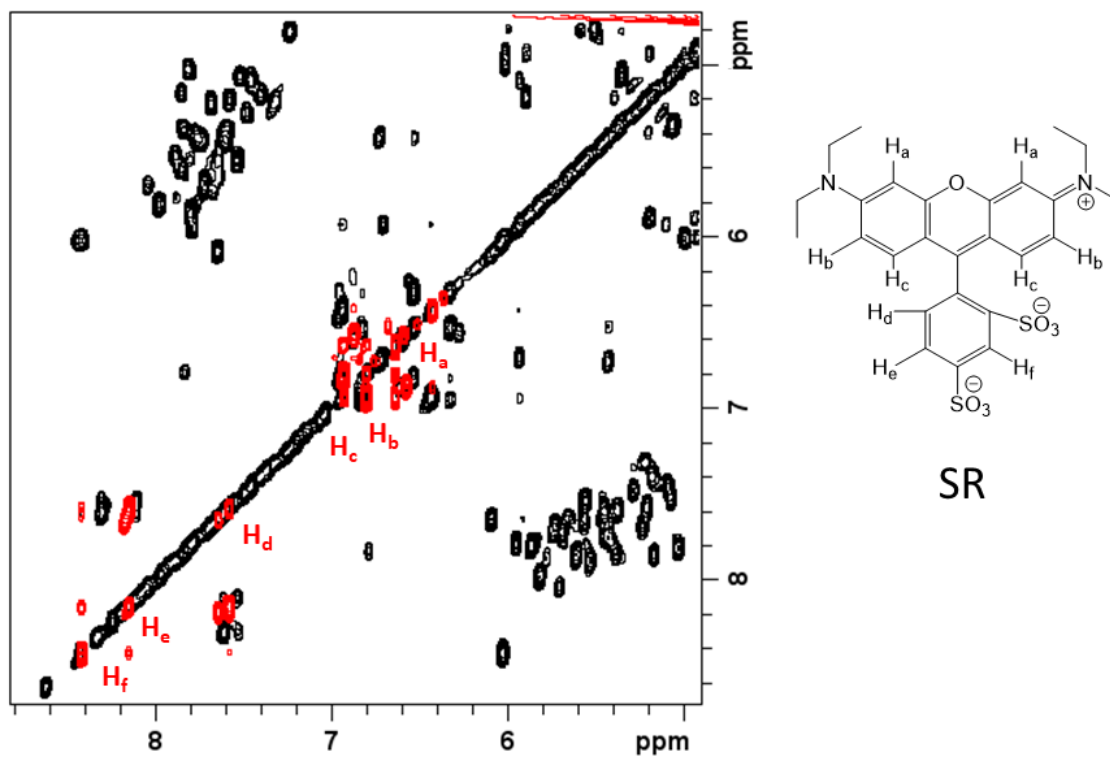


Figure A.2: TOCSY spectra of 1mM SR (red) and 1.8mM SRB-2 fully bound to SR (black).

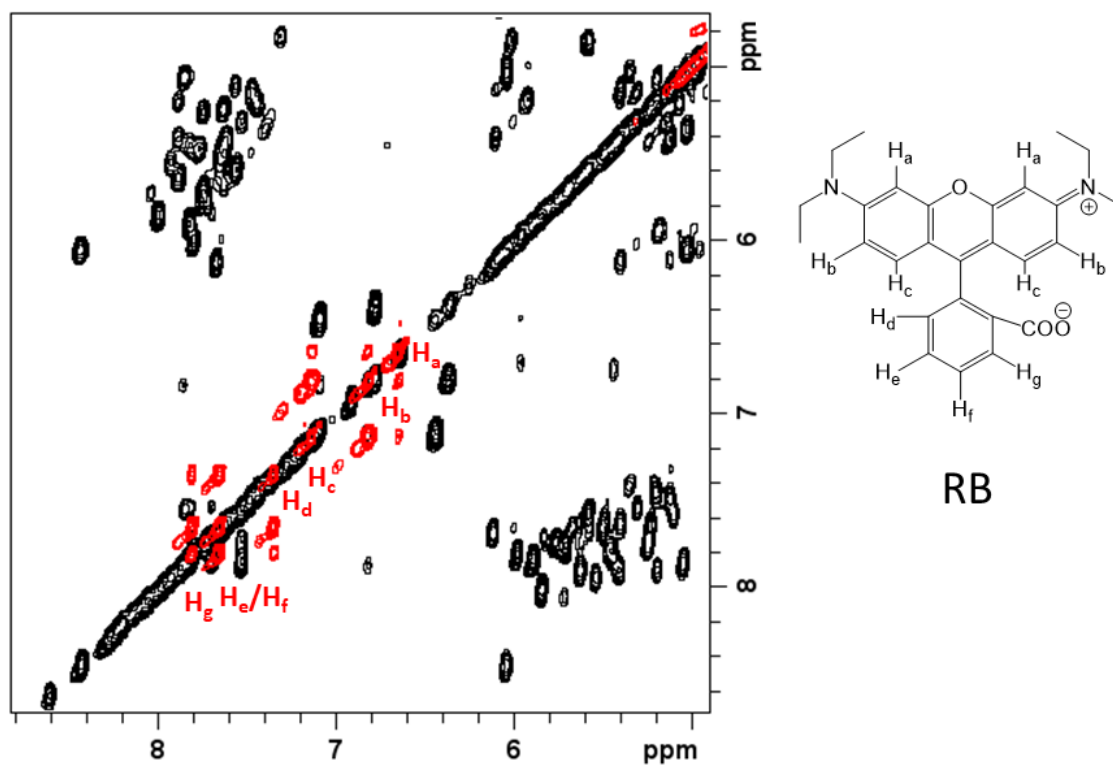


Figure A.3: TOCSY spectra of 1mM RB (red) and 1.3mM SRB-2 fully bound to RB (black).

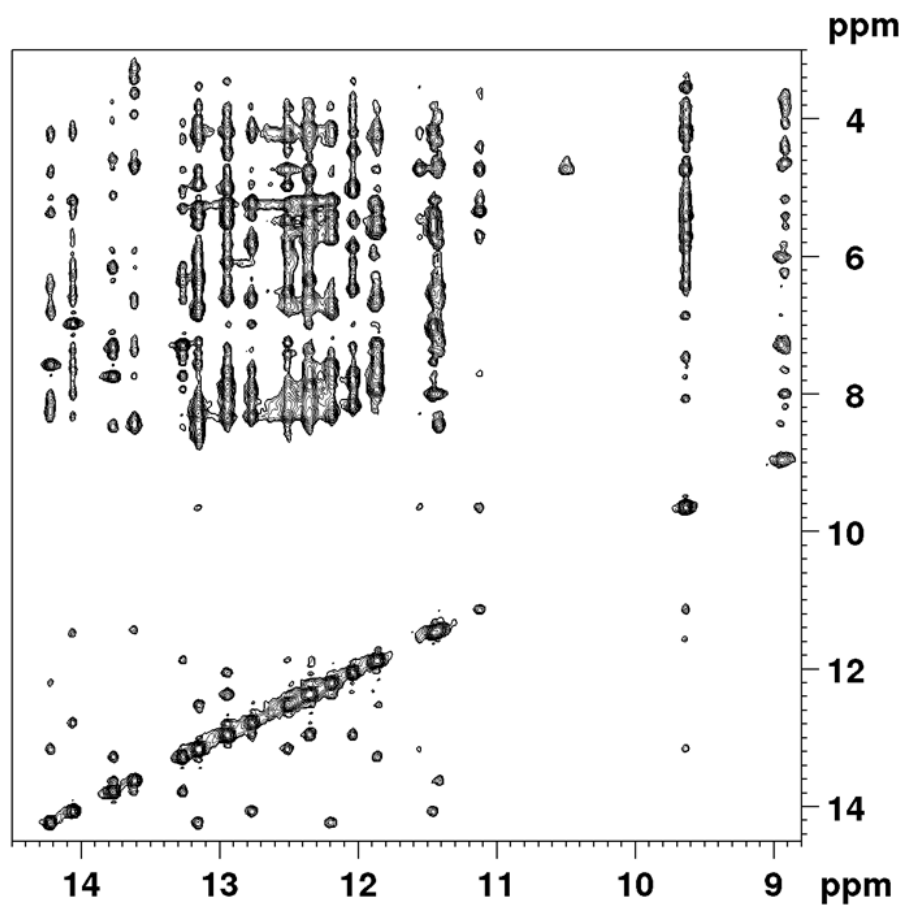


Figure A.4: NOESY spectrum of 1.3 mM SRB-2 bound ~1:1 with rhodamine B in 90% H₂O/10% D₂O.

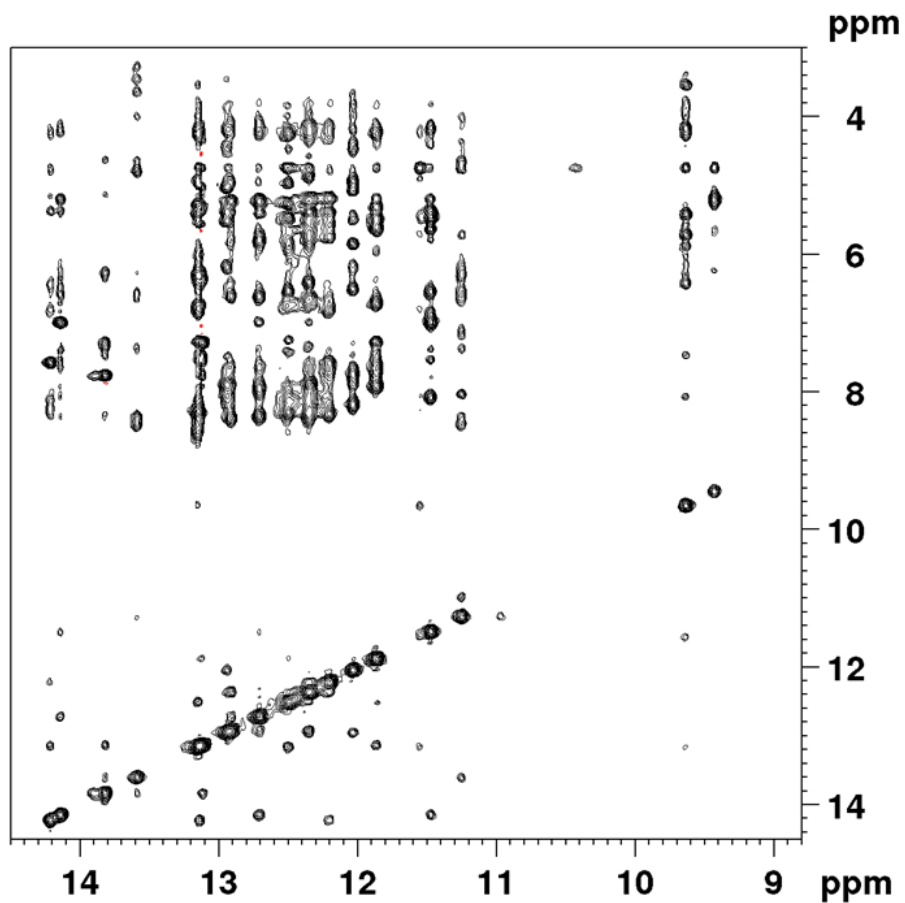


Figure A.5: NOESY spectrum of 1.1 mM SRB-2 bound ~1:1 with sulforhodamine 101 in 90% H₂O/10% D₂O.

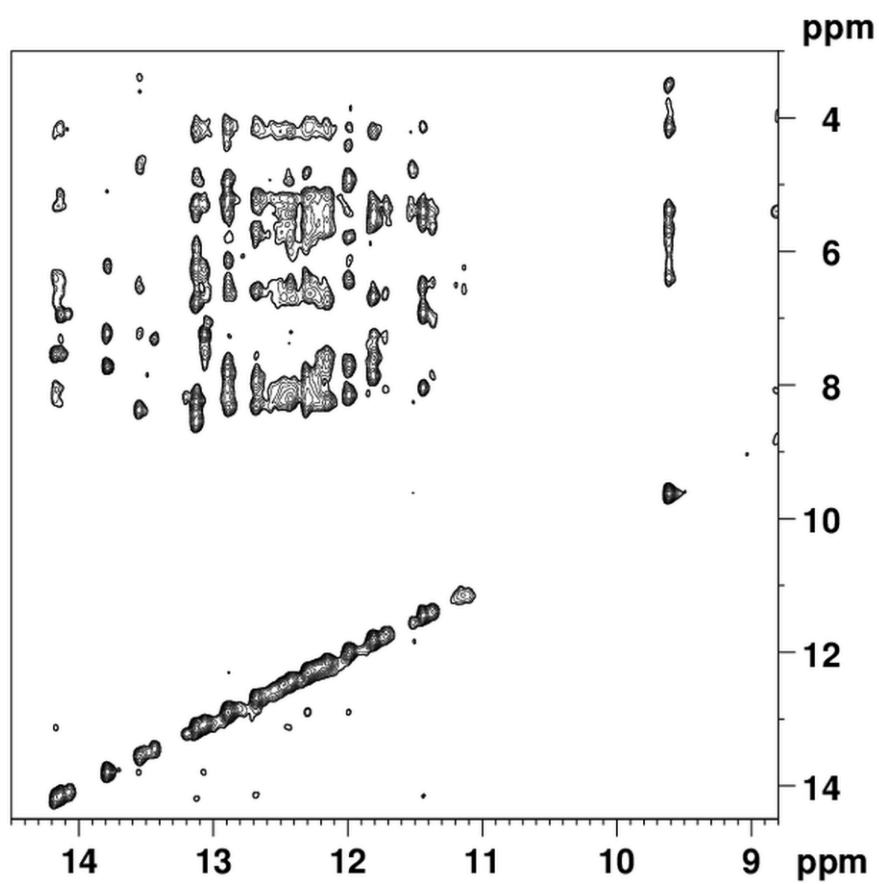


Figure A.6: NOESY spectrum of 0.4 mM SRB-2 bound ~1:1 with tetramethylrosamine in 90% H₂O/10% D₂O.

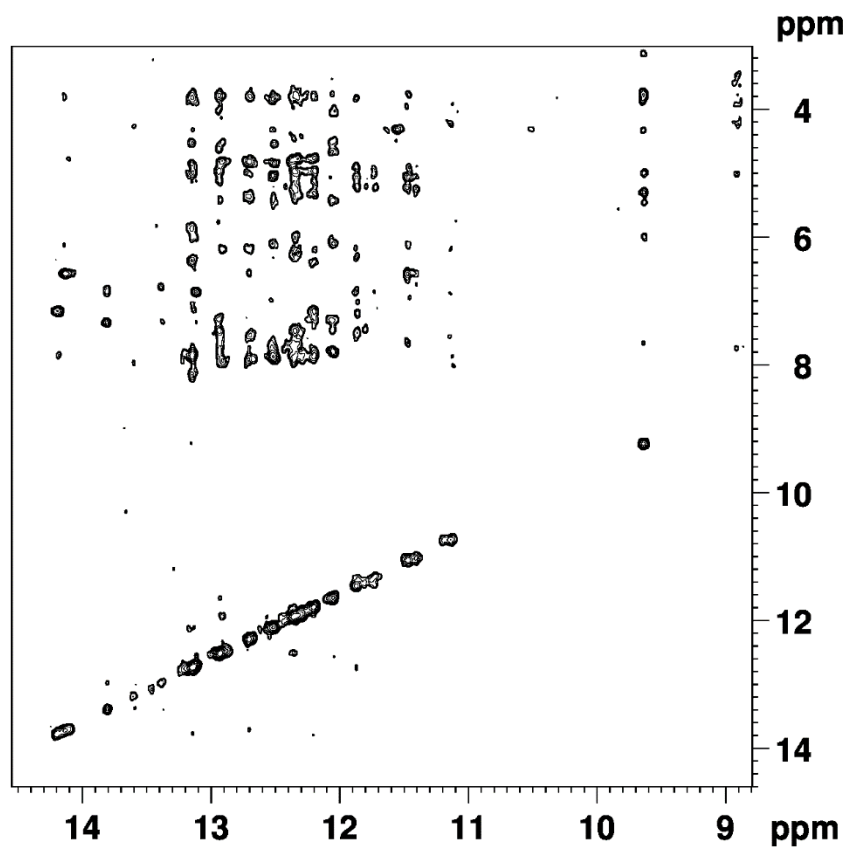


Figure A.7: NOESY spectrum of 0.4 mM SRB-2 bound ~1:1 with pyronin Y in 90% H₂O/10% D₂O.

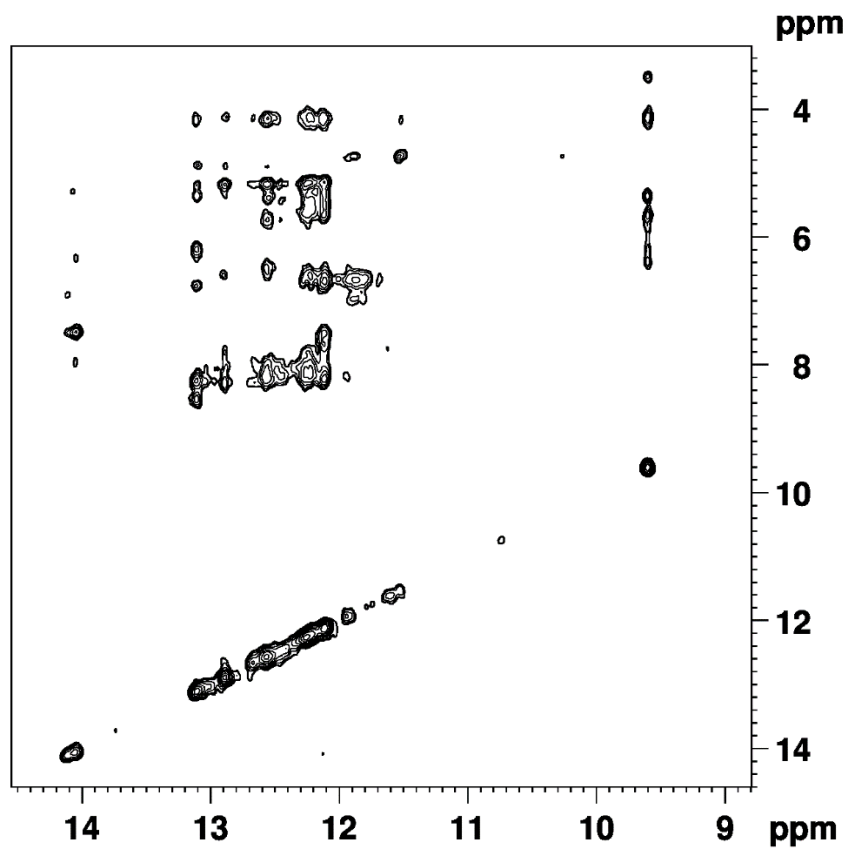


Figure A.8: NOESY spectrum of 0.4 mM SRB-2 bound ~1:1 with 9-aminoacridine in 90% H₂O/10% D₂O.

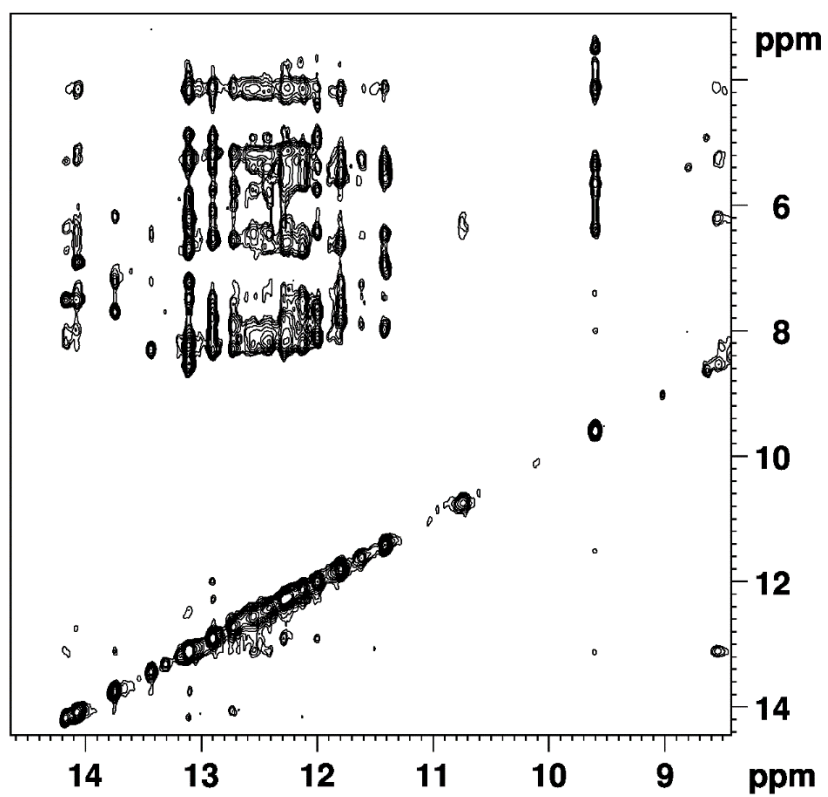


Figure A.9: H₂O NOESY spectrum of 0.4 mM SRB-2 bound ~1:1 with Atto 495.

Appendix B: NMR Pulse Programs

1D 1H with $1\bar{1}$ -Spin Echo Solvent Suppression

```
;SD_11echozg
;1D sequence for 11 spin echo water suppression with
gradients
;low power mode
;4.8.94 vladi sklenar
;modified 3-26-97, td
```

```
#include <Avance.incl>
#include <Grad.incl>
```

```
"d11=30m"
"d16=200u"
"d13=20u"
"d2=1s/(cnst2*4)-(p1*0.667)"
```

```
1 ze
d11
2 d13
d13
LOCKH_OFF
d1
LOCKH_ON
p1 ph1
d2
p1 ph2
p16:gp1
d16
p1 ph3
d2
d2
p1 ph4
p16:gp1
d16
go=2 ph31
wr #0
LOCKH_OFF
exit
```

```
ph1 = 0 0 0 0 1 1 1 1 2 2 2 2 3 3 3 3
ph2 = 2 2 2 2 3 3 3 3 0 0 0 0 1 1 1 1
ph3 = 0 1 2 3 1 2 3 0 2 3 0 1 3 0 1 2
ph4 = 2 3 0 1 3 0 1 2 0 1 2 3 1 2 3 0
ph30= 0
ph31= 0 2 0 2 1 3 1 3 2 0 2 0 3 1 3 1
```

```
;p1: 1H high power(~0 dB)
;p1 : 90 degree ecoupler pulse
;d1 : relaxation delay; 1-5 * T1
;d2 : 1/4*d2 excitation maximum calculated based on
; CNST2 = distance in Hz to first excitation max
```

1D 1H with Presaturation

```
;SD_1Dpresat
;avance-version
;1D sequence with f1 presaturation
```

```
#include <Avance.incl>
```

```
"d12=20u"
"d13=3u"
```

```
1 ze
2 d12 pl9:f1
p18*0.6 ph28
d13
p18*0.4 ph29
d12 pl1:f1
p1 ph1
go=2 ph31
wr #0
exit
```

```
ph1=0 2 2 0 1 3 3 1
ph28= 0
ph29= 1
ph31=0 2 2 0 1 3 3 1
```

```
;p1 : f1 channel - power level for pulse (default)
;p19 : f1 channel - power level for presaturation
;p1 : f1 channel - 90 degree high power pulse
;d1 : relaxation delay; 1-5 * T1
;d12: delay for power switching [20 usec]
;d13: short delay [3 usec]
```

2D NOESY with $1\bar{1}$ -Spin Echo Solvent Suppression

```
;SD_noe11ezg
;2D NOE with 11echo and gradients
;phase sensitive using States-TPPI
;with presaturation during relaxation delay and mixing
time
;the shortest mixing time ~50 ms!!!! otherwise problems
with radiation
;damping
;added decoupling, 9/15/94; thorsten
;changed for DRX, 05/01/96; thorsten
```

```
;Allows you to center spectrum anywhere.
;Uses a frequency list:
; 1) set the first value to the desired center of
spectrum in f1
; 2) set the second value to water
```



```

;      3) set the third value to the desired center of
spectrum in f2
;      4) Set o1 to the desired center of spectrum in f2
(needed for DQD)
;changed 1/21/97 - ELW

```

```

;d2 calculation added 3-22-97, td

```

```

#include <Avance.incl>
#include <Grad.incl>

```

```

"d11=30m"
"d16=200u"
"l3=td1/2"
"d0=(in0/2)-(p1*4/3.14159)"
"d8=d9-p16-d16"
"d2=1s/(cnst2*4)-(p1*0.667)"

```

```

1 ze
  d11
2 d11
3 d11 pl1:f1
4 d11
  LOCKH_OFF
  d1 fq1:f1           ;Center
  LOCKH_ON
  p1 ph1
  d0
  p1 ph2
  d8
  p16:gp1
  d16 fq1:f1         ;Water
  p1 ph3
  d2
  p1 ph6
  p16:gp2
  d16
  p1 ph7
  d2
  d2
  p1 ph8
  p16:gp2
  d16 fq1:f1         ;Center
  go=2 ph31
  d11 wr #0 if #0 ip1 zd
  lo to 3 times 2
  d11 id0
  lo to 4 times l3
  LOCKH_OFF
exit

ph1= 0 2 0 2 1 3 1 3 2 0 2 0 3 1 3 1
    2 0 2 0 3 1 3 1 0 2 0 2 1 3 1 3
ph2= 0 0 2 2 1 1 3 3 2 2 0 0 3 3 1 1
    2 2 0 0 3 3 1 1 0 0 2 2 1 1 3 3
ph3= 0 0 0 0 1 1 1 1 2 2 2 2 3 3 3 3

```

```

ph6= 2 2 2 2 3 3 3 3 0 0 0 0 1 1 1 1
ph7= 0 0 0 0 1 1 1 1 2 2 2 2 3 3 3 3
ph8= 2 2 2 2 3 3 3 3 0 0 0 0 1 1 1 1
ph20=0 0 0 0 0 0 0 0 2 2 2 2 2 2 2 2
ph31=0 2 2 0 1 3 3 1 2 0 0 2 3 1 1 3

```

```

;p1: Proton power level
;p1 : 90 degree proton pulse
;d9 : the mixing time
;d0 : incremented delay (2D) = in0/2-(4/pi*p1)
;d2 : delay for excitation maximum, calculated based
; on cnst2 = offset of first max in Hz
;d11: delay for disk I/O           [30 msec]
;gp1: Homospoil during mixing time
;gp2: Watergate gradient
;in0: 1/(SW)
;nd0: 1
;NS: 2,4 8,16 or 32 * n
;DS: 4
;td1: number of experiments
;MC2: States-TPPI

```

2D NOESY with Presaturation

```

;SD_noesypr
;avance-version
;2D homonuclear correlation via dipolar coupling
;dipolar coupling may be due to noe or chemical exchange.
;phase sensitive using States-TPPI method
;with presaturation during relaxation delay

; changed presat to 'vladi-trick' 03-24-96, td

```

```

#include <Avance.incl>
#include <Grad.incl>

```

```

"d11=30m"
"d12=20u"
"d13=5u"
"d16=200u"

"l3=(td1/2)"
"d0=((in0/2)-(p1*4/3.14159))"
"p19=d9-d12-p16-d16"

```

```

1 ze
2 d11
3 d11
4 d12 pl9:f1
  LOCKH_OFF
  p18*0.8 ph29
  d13
  p18*0.2 ph30
  LOCKH_ON
  d12 pl1:f1
  p1 ph1
  d0
  p1 ph2

```

```

d12 pl9:f1
p19 ph18
p16:gp1
d16 pl1:f1
p1 ph3
go=2 ph31
d11 wr #0 if #0 ip1 zd
lo to 3 times 2
d11 id0
lo to 4 times l3
LOCKH_OFF
exit

ph1 =0 2
ph2 =0 0 0 0 0 0 0 0 2 2 2 2 2 2 2 2
ph3 =0 0 2 2 1 1 3 3
ph18=0
ph29=0
ph30=1
ph31=0 2 2 0 1 3 3 1 2 0 0 2 3 1 1 3

;pl1 : f1 channel - power level for pulse (default)
;pl9 : f1 channel - power level for presaturation
;p1 : f1 channel - 90 degree high power pulse
;d0 : incremented delay (2D) [3 usec]
;d1 : relaxation delay; 1-5 * T1
;d9 : mixing time
;d11: delay for disk I/O [30 msec]
;d12: delay for power switching [20 usec]
;d13: short delay [3 usec]
;l3: loop for phase sensitive 2D using States-TPPI method:
l3 = td1/2
;in0: 1/(1 * SW) = 2 * DW
;nd0: 1
;NS: 8 * n
;DS: 16
;td1: number of experiments
;MC2: States-TPPI

```

2D CITY-TOCSY

```

;SD_cityprst
;2D clean CITY with water presat
;States-TPPI t1-detection
;DRX version with gradient purging
;4.7.1995 vladi sklenar
;03-28-97, td drx version

#include <Avance.incl>

define delay mix

"d0=3u"
"d11=30m"
"d12=20u"
"d16=200u"
"d20=(p6/90)*103"
"l3=td1/2"

```

```

"mix=(p6*48.532+d20*32)*l1"

1 ze
2 d11
3 d11
4 d12 pl9:f1
d13 LOCKH_OFF
p18*0.8 ph18
d13
p18*0.2 ph19
d13 LOCKH_ON
d12 pl1:f1
p1 ph1
d0
p1 ph2
p16:gp1
d16
d12 pl11:f1
7 (d20 p6*2 ph22 d20*2 p6*2 ph22 d20)
(p6*0.533 ph23 p6*3.067 ph21 p6*0.533 ph23)
(d20 p6*2 ph22 d20*2 p6*2 ph22 d20)
(d20 p6*2 ph22 d20*2 p6*2 ph22 d20)
(p6*0.533 ph23 p6*3.067 ph21 p6*0.533 ph23)
(d20 p6*2 ph22 d20*2 p6*2 ph22 d20)
(d20 p6*2 ph24 d20*2 p6*2 ph24 d20)
(p6*0.533 ph21 p6*3.067 ph23 p6*0.533 ph21)
(d20 p6*2 ph24 d20*2 p6*2 ph24 d20)
(d20 p6*2 ph24 d20*2 p6*2 ph24 d20)
(p6*0.533 ph21 p6*3.067 ph23 p6*0.533 ph21)
(d20 p6*2 ph24 d20*2 p6*2 ph24 d20)
lo to 7 times l1
p16:gp2
d16
d12 pl1:f1
p1 ph3
go=2 ph31
d11 wr #0 if #0 ip1 zd
lo to 3 times 2
d11 id0
lo to 4 times l3
mix
3u LOCKH_OFF
exit

ph1 = 0 2 0 2 0 2 0 2 1 3 1 3 1 3 1 3
ph2 = 0 0 0 0 2 2 2 2 1 1 1 1 3 3 3 3
ph3 = 0 0 2 2 0 0 2 2 1 1 3 3 1 1 3 3
ph4 = 0
ph9 = 0
ph10 = 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1
ph11 = 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2
ph21 = 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1
ph22 = 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2
ph23 = 2 2 2 2 2 2 2 2 3 3 3 3 3 3 3 3
ph24 = 3 3 3 3 3 3 3 3 0 0 0 0 0 0 0 0
ph18 = 0
ph19 = 1
ph31 = 0 2 2 0 2 0 0 2 1 3 3 1 3 1 1 3

```

```

;p19 : transmitt water presat level
;p11 : transmitter excitation level
;p111 : transmitter TOCSY mixing level
;p1 : 90 transmitter pulse at p1 level
;p6 : 90 transmitter pulse at p2 level
;p18 : presaturation pulse with p19
;d0 : incremented delay
;d11 : disc I/O delay
;d20 : delay for NOE-ROE compensation (1.1*p6? to be
optimized)
;l1 : loop for TOCSY mixing (48.532*p6+32*d20)
;NS=8 minimal phase cycle for pure phase spectra
including ax.peak suppr.
;NS=16 + short CYCLOPS
;MC2 : States-TPPI (90/-180 phase correction in F1)

```

2D DQF-COSY

```

;SD_cosydfprst
;avance-version
;2D homonuclear shift correlation
;with presaturation during relaxation delay
;phase sensitive using States-TPPI method
;with double quantum filter
;phasecycle: A. Derome & M. Williamson, J. Magn. Reson.
88,
; 177 - 185 (1990)

```

```

#include <Avance.incl>

```

```

"d0=3u"
"d11=30m"
"d12=20u"
"d13=3u"

```

```

"l3=(td1/2)"

```

```

1 ze
2 d11
3 d11
4 d12 p19:f1
p18*0.8 ph29
d13
p18*0.2 ph30
d12 p11:f1
p1 ph1
d0
p1 ph2
d13
p1 ph3
go=2 ph31
d11 wr #0 if #0 ip1 zd
lo to 3 times 2
d11 id0
lo to 4 times l3
exit

```

```

ph1=1 1 1 1 0 0 0 0
ph2=0 0 0 0 1 1 1 1
ph3=1 2 3 0 2 3 0 1
ph29=0
ph30=1
ph31=0 3 2 1 3 2 1 0

```

```

;p11 : f1 channel - power level for pulse (default)
;p19 : f1 channel - power level for presaturation
;p1 : f1 channel - 90 degree high power pulse
;d0 : incremented delay (2D) [3 usec]
;d1 : relaxation delay; 1-5 * T1
;d11: delay for disk I/O [30 msec]
;d12: delay for power switching [20 usec]
;d13: short delay [3 usec]
;l3: loop for phase sensitive 2D using States-TPPI method:
l3 = td1/2
;in0: 1/(1 * SW) = 2 * DW
;nd0: 1
;NS: 8 * n
;DS: 16
;td1: number of experiments
;MC2: States-TPPI

```

2D 13C HSQC

```

;SD_hsqchcgsse
;avance-version
;2D H-1/13C correlation via double inept transfer
; using sensitivity improvement
;phase sensitive using Echo/Antiecho gradient selection
;with decoupling during acquisition
;A.G. Palmer III, J. Cavanagh, P.E. Wright & M. Rance, J.
Magn.
; Reson. 93, 151-170 (1991)
;L.E. Kay, P. Keifer & T. Saarinen, J. Am. Chem. Soc. 114,
; 10663-5 (1992)
;J. Schleucher et al., Angew. Chem. 114(10), 1518 (1993)

```

```

;!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
;!! !!
;!! Attention: DS has to be a multiple of 2*NS !!
;!! !!
;!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!

```

```

#include <Avance.incl>

```

```

"p2=p1*2"
"p4=p3*2"
"d4=1s/(cnst2*4)"
"d24=1s/(cnst2*6)"

```

```

"l3=(td1/2)"

```

```

"d21=d13+p16+d16+4u"
"d0=3u" ;(in0/2-p3*1.273)"
"d20=p16+d16+50u+p2+d0*2"

```

```
#include <Grad.incl>
```

```
1 ze
  d11 pl12:f2
2 d1 do:f2
  6m
3 d11
4 (p1 ph1)
  d4 pl2:f2
  (p2 ph1) (p4 ph1):f2
  d4
  (p1 ph2) (p3 ph3):f2
  d0
  p2 ph1
  d0
  50u
  GRADIENT(const21)
  d16
  (p4 ph4):f2
  d20
  (p1 ph1) (p3 ph4):f2
  d24
  (p2 ph1) (p4 ph1):f2
  d24
  (p1 ph2) (p3 ph5):f2
  d4
  (p2 ph1) (p4 ph1):f2
  d4
  (p1 ph1)
  d21
  (p2 ph1)
  d13
  GRADIENT(const22)
  d16 pl12:f2
  4u
  go=2 ph31 cpd2:f2
  d1 do:f2 wr #0 if #0 zd
  3m ip5
  3m ip5
  lo to 3 times 2
  d11 id0
  lo to 4 times l3
exit
```

```
ph1=0
ph2=1
ph3=0 2
ph4=0 0 2 2
ph5=1 1 3 3
ph31=0 2 2 0
```

```
;pl1 : f1 channel - power level for pulse (default)
;pl2 : f2 channel - power level for pulse (default)
;pl12: f2 channel - power level for CPD/BB decoupling
;p1 : f1 channel - 90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p3 : f2 channel - 90 degree high power pulse
;p4 : f2 channel - 180 degree high power pulse
```

```
;p16: homospoil/gradient pulse
;p28: f1 channel - trim pulse
;p31: - 90 degree pulse for decoupling sequence
;d0 : incremented delay (2D) [3 usec]
;d1 : relaxation delay; 1-5 * T1
;d4 : 1/(4J)XH
;d11: delay for disk I/O [30 msec]
;d13: short delay (e.g. to compensate delay line) [3 usec]
;d16: delay for homospoil/gradient recovery
;d20: = p16+d16+50u+p2+d0*2
;d21: = d13+p16+d16+4u
;d24: 1/(4J)XH for XH
; 1/(6J)XH for all multiplicities
;l3: loop for phase sensitive 2D using E/A method : l3 =
td1/2
;in0: 1/(2 * SW(X)) = DW(X)
;nd0: 2
;NS: 1 * n
;DS: >= 16, but 2 * ns * m
;td1: number of experiments
;MC2: echo-antiecho
;cpd2: decoupling according to sequence defined by
cpdprg2
```

```
;use gradient program (GRDPROG) : SD_2sineea3
```

```
;use gradient ratio: const21 : const22 : const23 : const24 :
const25
```

```
; 4 : 1 : 4 : 1 : -1
for C-13
```

2D 15N HSQC

```
;hsqcetf3gp
;avance-version (02/07/15)
;HSQC
;2D H-1/X correlation via double inept transfer
;phase sensitive using Echo/Antiecho-TPPI gradient
selection
;with decoupling during acquisition
;using trim pulses in inept transfer
;using f3 - channel
;
;Davis et al., J. Magn. Reson. 98, 207 - (1992)
;
;,$CLASS=HighRes
;,$DIM=2D
;,$TYPE=
;,$SUBTYPE=
;,$COMMENT=
```

```
#include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>
```

```
"p2=p1*2"
"p22=p21*2"
"d0=3u"
```

```

"d11=30m"
"d13=4u"
"d26=1s/(cnst4*4)"

"DELTA1=d26-p16-d13-4u"

# ifdef LABEL_CN
"DELTA=p16+d16+larger(p2,p14)+d0*2"
# else
"DELTA=p16+d16+p2+d0*2"
# endif /*LABEL_CN*/

1 ze
  d11 pl16:f3
2 d1 do:f3
3 (p1 ph1)
  d26 pl3:f3
  (center (p2 ph1) (p22 ph6):f3 )
  d26 UNBLKGRAD
  p28 ph1
  d13
  (p1 ph2)
  3u
  p16:gp1
  d16
  (p21 ph3):f3
  d0

# ifdef LABEL_CN
  (center (p2 ph5) (p14:sp3 ph1):f2 )
# else
  (p2 ph5)
# endif /*LABEL_CN*/

d0
p16:gp2*EA
d16
(p22 ph4):f3
DELTA
(ralign (p1 ph1) (p21 ph4):f3 )
d26
(center (p2 ph1) (p22 ph1):f3 )
d13
p16:gp3
DELTA1 pl16:f3
4u BLKGRAD
go=2 ph31 cpd3:f3
d1 do:f3 mc #0 to 2
  F1EA(igrad EA, id0 & ip3*2 & ip6*2 & ip31*2)
exit

ph1=0
ph2=1
ph3=0 2
ph4=0 0 0 2 2 2 2
ph5=0 0 2 2

```

```

ph6=0
ph31=0 2 0 2 2 0 2 0

;p1 : f1 channel - power level for pulse (default)
;p3 : f3 channel - power level for pulse (default)
;p16: f3 channel - power level for CPD/BB decoupling
;sp3: f2 channel - shaped pulse 180 degree (adiabatic)
;p1 : f1 channel - 90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p14: f2 channel - 180 degree shaped pulse for inversion
(adiabatic)
;p16: homospoil/gradient pulse [1 msec]
;p21: f3 channel - 90 degree high power pulse
;p22: f3 channel - 180 degree high power pulse
;p28: f1 channel - trim pulse [1 msec]
;d0 : incremented delay (2D) [3 usec]
;d1 : relaxation delay; 1-5 * T1
;d11: delay for disk I/O [30 msec]
;d13: short delay [4 usec]
;d16: delay for homospoil/gradient recovery
;d26: 1/(4J(YH))
;cnst4: = J(YH)
;in0: 1/(2 * SW(X)) = DW(X)
;nd0: 2
;NS: 1 * n
;DS: >= 16
;td1: number of experiments
;FnMODE: echo-antiecho
;cpd3: decoupling according to sequence defined by
cpdprg3
;pcpd3: f3 channel - 90 degree pulse for decoupling
sequence

;use gradient ratio: gp 1 : gp 2 : gp 3
;50 : 80 : 20.1 for C-13
;50 : 80 : 8.1 for N-15

;for z-only gradients:
;gpz1: 50%
;gpz2: 80%
;gpz3: 20.1% for C-13, 8.1% for N-15

;use gradient files:
;gpnam1: SINE.100
;gpnam2: SINE.100
;gpnam3: SINE.100

;preprocessor-flags-start
;LABEL_CN: for C-13 and N-15 labeled samples start
experiment with
; option -DLABEL_CN (eda: ZGOPTNS)
;preprocessor-flags-end

;$Id: hsqctf3gp,v 1.2.8.1 2005/11/10 13:18:57 ber Exp $

```

3D HCCH COSY

```

;hcchcogp3d
;avance-version
;HCCH-COSY
;3D sequence with
; inverse correlation using multiple inept transfer and
;
; F1(H,t1) -> F2(C,t2) -> F2(C') -> F1(H',t3)
;
;off resonance C=O pulse using shaped pulse
;phase sensitive using States-TPPI
;(L.E. Kay, G.Y. Xu, A.U. Singer, D.R. Muhandiram & J. D.
Forman-Kay
; J. Magn. Reson. B 101, 333 - 337 (1993))

#include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>

"d0=3u"
"d10=3u"
"d11=30m"
"d12=20u"
"d4=1.6m"
"d21=1.1m"
"d23=475u"
"d24=3.6m"

"DELTA1=d4-p16-d16-4u"
"DELTA2=d4-p16-d16-4u+d0*2+p4"
"DELTA3=d23-p29-d16"
"DELTA4=p22+p2+d10*2+4u"
"DELTA5=d24-p29-d16-4u"
"DELTA6=d21-p16-d16-4u"
"DELTA7=d4-p16-d16-p3*2-7u+p1"

"TAU=(p3*2+3u)-p1"

"CEN_HC1=(p3-p1)/2"
"CEN_HC2=(p4-p2)/2"

"l3=(td1/2)"
"l13=(td2/2)"

"d22=3.6m-d23*2-(in10*l13/2)"

aqseq 312

1 d11 ze
d11 pl12:f2
2 d11 do:f2
d11
3 d11
4 d11
5 d11
6 d1
50u UNBLKGRAD
d12 pl1:f1

```

```

(p1 ph3)
4u
p16:gp1
d16
DELTA1 pl2:f2
d0
(p4 ph1):f2
d0
(p2 ph1)
4u
p16:gp1
d16
DELTA2 pl3:f3
(p1 ph2)

p19:gp3
d16

(p3 ph4):f2
d10
; (p22 ph1):f3
4u
p29:gp2
d16
DELTA3 pl0:f2
; (p14:sp5 ph1):f2
4u
p29:gp2
d16
DELTA3 pl2:f2
p2 ph1
d10
d22
(p4 ph1):f2
DELTA4
p29:gp2
d16
DELTA3
; (p14:sp5 ph1):f2
4u
p29:gp2
d16
DELTA3 pl2:f2
d22

(p3 ph1):f2
4u
p29:gp2
d16
DELTA5
(p4 ph1):f2
4u
p29:gp2
d16
DELTA5
(p3 ph2):f2

(p17 ph1)

```

```

(p17*2 ph2)
4u
p30:gp4
d16
(p1 ph1)
4u
p31:gp4
d16

(p3 ph2):f2
4u
p16:gp1
d16
DELTA6 pl2:f2
(CEN_HC2 p2 ph1) (p4 ph1):f2
4u
p16:gp1
d16
DELTA6
(CEN_HC1 p1 ph1) (p3 ph1):f2

4u
p16:gp1
d16
DELTA1
(CEN_HC2 p2 ph1) (p4 ph1):f2
4u
p16:gp1
d16
DELTA7
(TAU p1 ph1) (p3 ph1 3u p3 ph5):f2
4u pl12:f2
4u BLKGRAD
go=2 ph31 cpd2:f2
d11 do:f2 wr #0 if #0 zd

d11 ip3
lo to 3 times 2
d11 id0
lo to 4 times l3
d11 rd0 ip4
lo to 5 times 2
d11 id10
lo to 6 times l13

exit

ph1=0
ph2=1
ph3=0 2
ph4=0 0 2 2
ph5=0 0 0 2 2 2 2
ph31=0 2 2 0

;p10 : 120dB
;p1 : f1 channel - power level for pulse (default)
;p2 : f2 channel - power level for pulse (default)
;p3 : f3 channel - power level for pulse (default)

;p12: f2 channel - power level for CPD/BB decoupling
;sp5: f2 channel - shaped pulse 180 degree (C=O off
resonance)
;p1 : f1 channel - 90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p3 : f2 channel - 90 degree high power pulse
;p4 : f2 channel - 180 degree high power pulse
;p14: f2 channel - 180 degree shaped pulse
;p16: homospoil/gradient pulse [500 usec]
;p17: f1 channel - trim pulse [1 msec]
;p19: gradient pulse 2 [2 msec]
;p22: f3 channel - 180 degree high power pulse
;p27: f2 channel - trim pulse [2 msec]
;p29: gradient pulse 3 [300 usec]
;p30: gradient pulse 4 [5 msec]
;p31: gradient pulse 5 [4.4 msec]
;d0 : incremented delay (F1 in 3D) [3 usec]
;d4 : 1/(4J(CH)) - tau a [1.6 msec]
;d10: incremented delay (F2 in 3D) [3 usec]
;d11: delay for disk I/O [30 msec]
;d12: delay for power switching [20 usec]
;d16: delay for homospoil/gradient recovery
;d21: 1/(6J'(CH)) - tau c [1.1 msec]
;d22: d22+d23*2+t2max*0.5 = 1/(4J(CC)) : 3.6m-d23*2-
(in10*113/2)
;d23: tau b [475 usec]
;d24: 1/(4J(CC)) [3.6 msec]
;l3: loop for phase sensitive 3D using States-TPPI : l3 =
td1/2
;l13: loop for phase sensitive 3D using States-TPPI : l13 =
td2/2
;in0: 1/(2 * SW(H)) = DW(H)
;nd0: 2
;in10: 1/(2 * SW(C)) = DW(C)
;nd10: 2
;NS: 16 * n
;DS: 32
;td1: number of experiments in F1
;td2: number of experiments in F2
;MC2: States-TPPI in F1
;MC2: States-TPPI in F2
;cpd2: decoupling according to sequence defined by
cpdprg2
;pcpd2: f2 channel - 90 degree pulse for decoupling
sequence

;use gradient ratio: gp 1 : gp 2 : gp 3 : gp 4
; 16 : 16 : 30 : 60

;for z-only gradients:
;gpz1: 16%
;gpz2: 16%
;gpz3: 30%
;gpz4: 60%

;use gradient files:
;gpnam1: SINE.100
;gpnam2: SINE.50

```

;gpnam3: SINE.100
;gpnam4: SINE.100

2D NOESY with X-Half Filter in F2

;noesygpphwx2
;avance-version (05/02/17)
;2D homonuclear correlation via dipolar coupling
; with C-13 and N-15 filter in F2
;
; H (t1) -> H[C12,N14] (t2)
;
;dipolar coupling may be due to noe or chemical exchange.
;phase sensitive
;with gradient pulses in mixing time
;water suppression using watergate sequence
;
;A.L. Breeze, Prog. NMR Spectrosc. 36, 323-372 (2000)
;C. Zwahlen, P. Legault, S.J.F. Vincent, J. Greenblatt, R.
Konrat &
; L.E. Kay, J. Am. Chem. Soc. 119 6711-6721 (1997)
;K. Ogura, H. Terasawa & F. Inagaki, J. Biomol. NMR 8, 492-
498 (1996)
;J. Iwahara, J.M. Wojciak & R.T. Clubb, J. Biomol. NMR 19,
231-241 (2001)
;
;\$CLASS=HighRes
;\$DIM=2D
;\$TYPE=
;\$SUBTYPE=
;\$COMMENT=

prosol relations=<triple>

#include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>

"p2=p1*2"
"p22=p21*2"
"d0=3u"
"d11=30m"
"d26=1s/(cnst4*4)"
"d28=1s/(cnst6*4)"
"d29=1s/(cnst7*4)"

"DELTA=larger(p8,p22)+d0*2"
"DELTA1=d28-p19-d16-p15/2"
"DELTA2=d26-d28-p15/2"
"DELTA3=d26-p19-d16"
"DELTA4=d29-p19-d16-p15/2"
"DELTA5=d26-d29-p15/2"

"TAU=d8-p16-d16"

1 ze
2 d11
3 d1 pl1:f1

(p1 ph4)
DELTA
(p2 ph5)
d0
(center (p8:sp13 ph1):f2 (p22 ph1):f3)
d0
(p1 ph6)
TAU UNBLKGRAD
p16:gp1
d16 pl2:f2

(p1 ph1)
p19:gp2
d16
DELTA1
(p15:sp18 ph1):f2
DELTA2
(center (p2 ph2) (p22 ph1):f3)
DELTA3
p19:gp2
d16 pl2:f2
(p1 ph3)

(p3 ph8):f2
(p21 ph8):f3
p16:gp3
d16

(p1 ph1)
p19:gp4
d16
DELTA4
(p15:sp18 ph1):f2
DELTA5
(center (p2 ph2) (p22 ph1):f3)
DELTA3
p19:gp4
d16 pl2:f2
(p1 ph3)

(p3 ph8):f2
(p21 ph8):f3
p16:gp5
d16

(p1 ph7)
4u
p16:gp6
d16 pl0:f1
(p11:sp1 ph9:r):f1
4u
4u pl1:f1
(p2 ph10)
4u pl0:f1
(p11:sp1 ph9:r):f1
4u
p16:gp6
d16

4u BLKGRAD

go=2 ph31
d11 mc #0 to 2 F1PH(ip4 & ip5, id0)
exit

ph1=0
ph2=1
ph3=2
ph4=0 0 0 0 2 2 2 2
ph5=1 1 1 1 3 3 3 3
ph6=0 0 0 0 0 0 2 2 2 2 2 2 2
ph7=0 1 2 3
ph8=0 1 2 3
ph9=2 2 2 2 3 3 3 3 0 0 0 1 1 1 1
ph10=0 0 0 0 1 1 1 1 2 2 2 2 3 3 3 3
ph31=0 3 2 1 0 3 2 1 2 1 0 3 2 1 0 3

;pl0 : 120dB
;pl1 : f1 channel - power level for pulse (default)
;pl2 : f2 channel - power level for pulse (default)
;pl3 : f3 channel - power level for pulse (default)
;sp1: f1 channel - shaped pulse 90 degree
;sp13: f2 channel - shaped pulse 180 degree (adiabatic)
;spnam13: Crp60,0.5,20.1
;sp18: f2 channel - shaped pulse 180 degree (adiabatic)
;p1 : f1 channel - 90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p3 : f2 channel - 90 degree high power pulse
;p8 : f2 channel - 180 degree shaped pulse for inversion (adiabatic)
;p11: f1 channel - 90 degree shaped pulse
;p15: f2 channel - 180 degree shaped pulse for inversion (adiabatic)
;p16: homospoil/gradient pulse [1 msec]
;p19: gradient pulse 2 [500 usec]
;p21: f3 channel - 90 degree high power pulse
;p22: f3 channel - 180 degree high power pulse
;d0 : incremented delay (2D) [3 usec]
;d1 : relaxation delay; 1-5 * T1
;d11: delay for disk I/O [30 msec]
;d8 : mixing time
;d16: delay for homospoil/gradient recovery
;d26: 1/(4*J(NH))
;d28: 1/(4*J(CH)min)
;d29: 1/(4*J(CH)max)
;cnst4: = J(NH)
;cnst6: = 1J(CH)min
;cnst7: = 1J(CH)max
;in0: 1/(2 * SW) = DW
;nd0: 2
;NS: 8 * n
;DS: 32
;td1: number of experiments
;FnMODE: States-TPPI, TPPI, States or QSEQ

;use gradient ratio: gp 1 : gp 2 : gp 3 : gp 4 : gp 5 : gp 6
; 50 : 7 : 40 : 5 : 3 : 30

;for z-only gradients:

;gpz1: 50%
;gpz2: 7%
;gpz3: 40%
;gpz4: 5%
;gpz5: 3%
;gpz6: 30%

;use gradient files:

;gpnam1: SINE.100
;gpnam2: SINE.100
;gpnam3: SINE.100
;gpnam4: SINE.100
;gpnam5: SINE.100
;gpnam6: SINE.100

;for p15:sp18 use either "Crp60,0.5,20.1" or adiabatic pulse

; with matched sweep

;for matched sweep use

; low to high field sweep,

; carrier shifted to low field end via spoofs18

; d28 = d29 = 1/(4*J(CH)aro)

;\$Id: noesygpwhgx2,v 1.2.6.2 2006/04/03 14:25:11 ber Exp \$

2D NOESY with Double X-Half Filters

;noesygpwhgx2

;avance-version (05/09/23)

;2D homonuclear correlation via dipolar coupling

;dipolar coupling may be due to noe or chemical exchange

;phase sensitive

;selecting C-12 or N-14 bound protons in F1 and F2

;water suppression using wtergate sequence

;

;M. Ikura & A. Bax, J. Am. Chem. Soc. 114, 2433-2440 (1992)

;M. Piotto, V. Saudek & V. Sklenar, J. Biomol. NMR 2, 661 - 666 (1992)

;V. Sklenar, M. Piotto, R. Leppik & V. Saudek, J. Magn. Reson.,

; Series A 102, 241 -245 (1993)

;

;\$CLASS=HighRes

;\$DIM=2D

;\$TYPE=

;\$SUBTYPE=

;\$COMMENT=

prosol relations=<triple>

#include <Avance.incl>

#include <Grad.incl>

#include <Delay.incl>

```

"p2=p1*2"
"p22=p21*2"
"d0=3u"
"d2=1s/(cnst2*2)"
"d11=30m"
"d12=20u"
"d21=1s/(cnst4*2)"

"p16=1m"
"p19=3m"

"DELTA=d21-d2-larger(p2,p22)-p3"
"DELTA1=d2*2-d21-p21+larger(p2,p22)+p3-d0"
"DELTA2=d24-p16-p11-d12-8u"
"DELTA3=d21-d24-larger(p2,p22)-p3-p11-p16-d12-8u"
"DELTA4=d24*2-d21-p21+larger(p2,p22)+p3-4u"

"TAU=d8/2--p3/2-p21/2"
"TAU1=d8/2-p21/2-p19-4u"

1 ze
2 d11
3 d1 pl1:f1
  (p1 ph1)
  d2
  (p3 ph6):f2
  (center (p2 ph2) (p22 ph7):f3 )
  DELTA
  (p21 ph6):f3
  DELTA1
  (p3 ph7):f2
  d0
  (p1 ph3)
  TAU UNBLKGRAD
  (p3 ph1):f2
  (p21 ph1):f3
  4u
  p19:gp1
  TAU1
  (p1 ph4):f1
  4u
  p16:gp2
  DELTA2 pl0:f1
  (p11:sp1 ph5:r):f1
  4u
  d12 pl1:f1
  (p3 ph6):f2
  (center (p2 ph1) (p22 ph7):f3 )
  4u
  d12 pl0:f1
  (p11:sp1 ph5:r):f1
  4u
  p16:gp3
  DELTA3
  (p21 ph9):f3
  DELTA4
  (p3 ph8):f2
  4u BLKGRAD

go=2 ph31
d11 mc #0 to 2 F1PH(dp3, id0)
exit

ph1=0
ph2=0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1
  2 3 2 3 2 3 2 3 2 3 2 3 2 3 2 3
ph3=0
ph4=0 0 1 1 2 2 3 3
ph5=2
ph6=0 0 0 0 0 0 0 2 2 2 2 2 2 2 2
ph7=0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
  2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
ph8=0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
  2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
  2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
  2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
  2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
ph9=0 0 0 0 2 2 2 2
ph31=0 2 3 1 2 0 1 3

;p0 : 120dB
;p1 : f1 channel - power level for pulse (default)
;p2 : f2 channel - power level for pulse (default)
;p3 : f3 channel - power level for pulse (default)
;sp1 : f1 channel - shaped pulse 90 degree
;p1 : f1 channel - 90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p3 : f2 channel - 90 degree high power pulse
;p11: f1 channel - 90 degree shaped pulse [1 msec]
;p16: homospoil/gradient pulse [1 msec]
;p19: gradient pulse 2 [3 msec]
;p21: f3 channel - 90 degree high power pulse
;p22: f3 channel - 180 degree high power pulse
;d0 : incremented delay (2D) [3 usec]
;d1 : relaxation delay; 1-5 * T1
;d2 : 1/(2J(CH))
;d8 : mixing time
;d11: delay for disk I/O [30 msec]
;d12: delay for power switching [20 usec]
;d16: delay for homospoil/gradient recovery
;d21: 1/(2J(NH))
;d24: 1/(2J'(CH))
;in0: 1/(1 * SW) = 2 * DW
;nd0: 1
;NS: 8 * n
;DS: 16
;td1: number of experiments
;FnMODE: States-TPPI, TPPI, States or QSEQ

;use gradient ratio: gp 1 : gp 2 : gp 3
; 50 : 30 : 30

```

```

;for z-only gradients:
;gpz1: 50%
;gpz2: 30%
;gpz3: 30%

;use gradient files:
;gpnam1: SINE.100
;gpnam2: SINE.100
;gpnam3: SINE.100

;$Id: noesygpqhwxgxf,v 1.5.10.2 2006/04/03 14:25:11 ber
Exp $

```

3D Edited HSQC-NOESY

```

;noesyhsqcetgp3d
;avance-version (05/10/28)
;NOESY-HSQC
;3D sequence with
; homonuclear correlation via dipolar coupling
; dipolar coupling may be due to noe or chemical
exchange.
; H-1/X correlation via double inept transfer
;phase sensitive (t1)
;phase sensitive using Echo/Antiecho-TPPI gradient
selection (t2)
;using trim pulses in inept transfer
;with decoupling during acquisition
;using shaped pulses for inversion on f2 - channel
;(use parameterset NOESYHSQCETGP3D)
;
;A.L. Davis, J. Keeler, E.D. Laue & D. Moskau, J. Magn.
Reson. 98,
; 207-216 (1992)
;
;$CLASS=HighRes
;$DIM=3D
;$TYPE=
;$SUBTYPE=
;$COMMENT=

#include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>

"p2=p1*2"
"p4=p3*2"
"d0=3u"
"d4=1s/(cnst2*4)"
"d10=3u"
"d11=30m"
"d12=20u"

# ifdef LABEL_CN
"p22=p21*2"
# else
# endif /*LABEL_CN*/

```

```

"DELTA2=d4-larger(p2,p14)/2-p16-8u"
"DELTA3=d4-larger(p2,p14)/2"

# ifdef LABEL_CN
"DELTA=larger(p14,p22)+d0*2"
"DELTA1=p16+d16+larger(p2,p22)+d10*2"
# else
"DELTA=p14+d0*2"
"DELTA1=p16+d16+p2+d10*2"
# endif /*LABEL_CN*/

aqseq 321

1 ze
d11 pl12:f2
2 d1 do:f2
3 d12 pl0:f2
(p1 ph7)
DELTA
(p2 ph8)
d0

# ifdef LABEL_CN
(center (p14:sp3 ph1):f2 (p22 ph1):f3 )
# else
(p14:sp3 ph1):f2
# endif /*LABEL_CN*/

d0
(p1 ph1):f1
d8
(p1 ph1):f1
DELTA3 pl0:f2
4u
(center (p2 ph1) (p14:sp3 ph6):f2 )
4u
DELTA3 pl2:f2 UNBLKGRAD
p28 ph1
4u
(p1 ph2) (p3 ph3):f2
d10

# ifdef LABEL_CN
(center (p2 ph5) (p22 ph1):f3 )
# else
(p2 ph5)
# endif /*LABEL_CN*/

d10
p16:gp1*EA
d16
(p4 ph4):f2
DELTA1
(ralign (p1 ph1) (p3 ph4):f2 )
DELTA3 pl0:f2
(center (p2 ph1) (p14:sp3 ph1):f2 )
4u
p16:gp2

```

```

DELTA2 pl12:f2
4u BLKGRAD
go=2 ph31 cpd2:f2
d1 do:f2 mc #0 to 2
  F1PH(rd10 & rp3 & rp6 & rp31 & ip7 & ip8, id0)
  F2EA(igrad EA, id10 & ip3*2 & ip6*2 & ip31*2)
exit

ph1=0
ph2=1
ph3=0 2
ph4=0 0 0 0 0 0 0 2 2 2 2 2 2 2 2
ph5=0 0 2 2
ph6=0
ph7=0 0 0 0 2 2 2 2
ph8=1 1 1 1 3 3 3 3
ph31=0 2 0 2 2 0 2 0 2 0 2 0 0 2 0 2

;pl0 : 120dB
;pl1 : f1 channel - power level for pulse (default)
;pl2 : f2 channel - power level for pulse (default)
;pl3 : f3 channel - power level for pulse (default)
;pl12: f2 channel - power level for CPD/BB decoupling
;sp3: f2 channel - shaped pulse 180 degree
;p1 : f1 channel - 90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p3 : f2 channel - 90 degree high power pulse
;p4 : f2 channel - 180 degree high power pulse
;p14: f2 channel - 180 degree shaped pulse for inversion
;p16: homospoil/gradient pulse
;p22: f3 channel - 180 degree high power pulse
;p28: f1 channel - trim pulse
;d0 : incremented delay (F1 in 3D)      [3 usec]
;d1 : relaxation delay; 1-5 * T1
;d4 : 1/(4)XH
;d8 : mixing time
;d10: incremented delay (F2 in 3D)      [3 usec]
;d11: delay for disk I/O                [30 msec]
;d12: delay for power switching         [20 usec]
;d16: delay for homospoil/gradient recovery
;cnst2: = J(XH)
;in0: 1/(2 * SW(H)) = DW(H)
;nd0: 2
;in10: 1/(2 * SW(X)) = DW(X)
;nd10: 2
;NS: 8 * n
;DS: >= 16
;td1: number of experiments in F1
;td2: number of experiments in F2
;FnMODE: States-TPPI (or TPPI) in F1
;FnMODE: echo-antiecho in F2
;cpd2: decoupling according to sequence defined by
cpdprg2
;pcpd2: f2 channel - 90 degree pulse for decoupling
sequence

;use gradient ratio:  gp 1 : gp 2
;                      80 : 20.1   for C-13

```

```

;                      80 : 8.1   for N-15

;for z-only gradients:
;gpz1: 80%
;gpz2: 20.1% for C-13, 8.1% for N-15

;use gradient files:
;gpnam1: SINE.100
;gpnam2: SINE.100

;preprocessor-flags-start
;LABEL_CN: for C-13 and N-15 labeled samples start
experiment with
;          option -DLABEL_CN (eda: ZGOPTNS)
;preprocessor-flags-end

;$Id: noesyhsqcetgp3d,v 1.3.2.1 2005/11/10 13:18:59 ber
Exp $

```

3D HCCNH

```

;na_h56c56c4n3h_3d
;avance-version (04/11/04)
;H6/5(C5C4)NH
;3D sequence with
; inverse correlation for triple resonance
;   via inept transfer steps
;   and DIPSI3 CC spinlock
;
; in U:
;   F1(H6/5,t1) -> F2(C6/5 -> C4) -> F3(N3,t2) -> F1(H3,t3)
;
; in C:
;   F1(H6/5,t1) -> F2(C6/5 -> C4) -> F3(N4,t2) -> F1(H4,t3)
;
;on/off resonance C pulses using shaped pulse
;phase sensitive (t1)
;phase sensitive (t2)
;using constant time in t2
;water suppression using watergate sequence
;(use parameterset NA_H56C56C4N3H_3D)
;
;J. Woehnert, M. Goerlach & H. Schwalbe,
; J. Biomol. NMR 26, 79-83 (2003)
;
;CLASS=HighRes
;$DIM=3D
;$TYPE=
;$SUBTYPE=
;$COMMENT=

prosol relations=<triple_na>

#include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>

```

"p2=p1*2"	DELTA3
"p22=p21*2"	4u
"d11=30m"	(center (p2 ph1) (p8:sp13 ph1):f2)
	4u
"d4=1.25m"	DELTA3 pl15:f2
"d25=10m"	
"d26=2.5m"	
	;begin DIPSI3
"in30=in10"	4 (p9*2.722 ph11):f2
	(p9*4.389 ph12):f2
"d0=3u"	(p9*2.778 ph11):f2
"d10=3u"	(p9*3.056 ph12):f2
"d30=d25"	(p9*0.333 ph11):f2
	(p9*2.556 ph12):f2
"l1=(d15/(p9*217.30))+0.5"	(p9*4.000 ph11):f2
	(p9*2.722 ph12):f2
"DELTA1=d4-d0-p8/2"	(p9*4.111 ph11):f2
"DELTA2=d4+d0+p8/2"	(p9*3.778 ph12):f2
"DELTA3=d4-larger(p2,p8)/2"	(p9*3.889 ph11):f2
"DELTA4=d25-d10-d24-p2"	(p9*2.889 ph12):f2
"DELTA5=d26-p16-d16-p11-8u"	(p9*3.000 ph11):f2
	(p9*0.333 ph12):f2
"spoff2=0"	(p9*2.500 ph11):f2
"spoff3=0"	(p9*4.050 ph12):f2
"spoff8=0"	(p9*2.830 ph11):f2
"spoff9=0"	(p9*4.389 ph12):f2
"spoff13=0"	(p9*2.722 ph12):f2
	(p9*4.389 ph11):f2
"cnst0=(cnst28+cnst24)/2"	(p9*2.778 ph12):f2
	(p9*3.056 ph11):f2
aqseq 312	(p9*0.333 ph12):f2
	(p9*2.556 ph11):f2
1 ze	(p9*4.000 ph12):f2
d11 pl16:f3	(p9*2.722 ph11):f2
2 d11 do:f3	(p9*4.111 ph12):f2
	(p9*3.778 ph11):f2
3 d1 pl1:f1 pl3:f3	(p9*3.889 ph12):f2
30u fq=cnst0(bf ppm):f2	(p9*2.889 ph11):f2
50u UNBLKGRAD	(p9*3.000 ph12):f2
	(p9*0.333 ph11):f2
(p3 ph1):f2	(p9*2.500 ph12):f2
(p21 ph1):f3	(p9*4.050 ph11):f2
p16:gp3	(p9*2.830 ph12):f2
d16	(p9*4.389 ph11):f2
(p1 ph3)	(p9*2.722 ph12):f2
DELTA1	(p9*4.389 ph11):f2
d0	(p9*2.778 ph12):f2
(p8:sp13 ph1):f2	(p9*3.056 ph11):f2
d0	(p9*0.333 ph12):f2
(p2 ph1)	(p9*2.556 ph11):f2
DELTA2	(p9*4.000 ph12):f2
(p1 ph2)	(p9*2.722 ph11):f2
	(p9*4.111 ph12):f2
p16:gp4	(p9*3.778 ph11):f2
d16	(p9*3.889 ph12):f2
	(p9*2.889 ph11):f2
(p3 ph4):f2	(p9*3.000 ph12):f2

```

(p9*0.333 ph11):f2
(p9*2.500 ph12):f2
(p9*4.050 ph11):f2
(p9*2.830 ph12):f2
(p9*4.389 ph11):f2

(p9*2.722 ph11):f2
(p9*4.389 ph12):f2
(p9*2.778 ph11):f2
(p9*3.056 ph12):f2
(p9*0.333 ph11):f2
(p9*2.556 ph12):f2
(p9*4.000 ph11):f2
(p9*2.722 ph12):f2
(p9*4.111 ph11):f2
(p9*3.778 ph12):f2
(p9*3.889 ph11):f2
(p9*2.889 ph12):f2
(p9*3.000 ph11):f2
(p9*0.333 ph12):f2
(p9*2.500 ph11):f2
(p9*4.050 ph12):f2
(p9*2.830 ph11):f2
(p9*4.389 ph12):f2
lo to 4 times l1

;end DIPS13
4u pl2:f2
(p3 ph2):f2

4u
30u fq=cnst24(bf ppm):f2

(p13:sp2 ph5):f2
d25
(center (p14:sp3 ph1):f2 (p30:sp9 ph1):f3 )
d25
(p13:sp8 ph2):f2

(p21 ph6):f3
d10
d24
(p2 ph1)
DELTA4
(center (p14:sp3 ph1):f2 (p30:sp9 ph1):f3 )
d30
(p21 ph1):f3

p16:gp4
d16

(p1 ph1)
p16:gp5
d16
DELTA5 pl0:f1
(p11:sp1 ph7):f1
4u
4u pl1:f1

(center (p2 ph1) (p22 ph1):f3 )
4u pl0:f1
(p11:sp1 ph7):f1
DELTA5
p16:gp5
d16 pl16:f3
4u BLKGRAD

go=2 ph31 cpd3:f3
d11 do:f3 mc #0 to 2
F1PH(ip3, id0)
F2PH(rd0 & rp3 & ip6, id10 & dd30)
exit

ph1=0
ph2=1
ph3=0 0 0 0 2 2 2 2
ph4=0 0 0 0 0 0 0 2 2 2 2 2 2 2 2
ph5=0 0 2 2
ph6=0 2
ph7=2
ph11=0
ph12=2
ph31=0 2 2 0 2 0 0 2 2 0 0 2 0 2 2 0

;pl0 : 120db
;p1 : f1 channel - power level for pulse (default)
;p2 : f2 channel - power level for pulse (default)
;p3 : f3 channel - power level for pulse (default)
;p15: f2 channel - power level for TOCSY-spinlock
;p16: f3 channel - power level for CPD/BB decoupling
;sp1: f1 channel - shaped pulse 90 degree (H2O on
resonance)
;sp2 : f2 channel - shaped pulse 90 degree (on resonance)
;sp3 : f2 channel - shaped pulse 180 degree (on resonance)
;sp8 : f2 channel - shaped pulse 90 degree (on resonance)
; for time reversed pulse
;sp9 : f3 channel - shaped pulse 180 degree (N on
resonance)
;sp13: f2 channel - shaped pulse 180 degree (adiabatic)
;spnam13: Crp60,0.5,20.1
;p1 : f1 channel - 90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p3 : f2 channel - 90 degree high power pulse
;p8 : f2 channel - 180 degree shaped pulse for inversion
(adiabatic)
; = 500usec for Crp60,0.5,20.1
;p9 : f2 channel - 90 degree low power pulse
;p11: f1 channel - 90 degree shaped pulse [1 msec]
;p13: f2 channel - 90 degree shaped pulse
;p14: f2 channel - 180 degree shaped pulse
;p16: homospoil/gradient pulse [1 msec]
;p21: f3 channel - 90 degree high power pulse
;p22: f3 channel - 180 degree high power pulse
;p30: f3 channel - 180 degree shaped pulse
;d0 : incremented delay (F1 in 3D) [3 usec]
;d1 : relaxation delay: 1-5 * T1
;d4 : 1/(4J(C5H5)) [1.25 msec]

```

```

;d10: incremented delay (F2 in 3D)          [3 usec]
;d11: delay for disk I/O                    [30 msec]
;d15: TOCSY mixing time (CC)                [19.5 msec]
;d16: delay for homospoil/gradient recovery
;d24: for U: 1/((4J)NH))                   [2.7 msec]
;   for C: 1/((8J)NH))                     [1.4 msec]
;d25: 1/(4J(C4N))                          [10 msec]
;d26: 1/(4J(NH))                           [2.5 msec]
;d30: decremented delay (F2 in 3D) = d25
;cnst0 : (cnst28 + cnst24)/2
;cnst24: C4 (C/U) chemical shift (offset, in ppm) [169 ppm]
;cnst28: C5 (C/U) chemical shift (offset, in ppm) [105 ppm]
;o3p: for U: 160 ppm, for C: 100 ppm
;l1: loop for DIPSI cycle:
;   mixing time = ((p9*217.3) * l1)        [19.5 msec]
;in0: 1/(2 * SW(H)) = DW(H)
;nd0: 2
;in10: 1/(2 * SW(N)) = DW(N)
;nd10: 2
;in30: = in10
;NS: 16 * n
;DS: >= 32
;td1: number of experiments in F1
;td2: number of experiments in F2   td2 max = 2 * d30 / in30
;FnMODE: States-TPPI (or TPPI) in F1
;FnMODE: States-TPPI (or TPPI) in F2
;cpd3: decoupling according to sequence defined by cpdprg3
;pcpd3: f3 channel - 90 degree pulse for decoupling sequence

;use gradient ratio: gp 3 : gp 4 : gp 5
;                               50 : 80 : 30

;for z-only gradients:
;gpz3: 50%
;gpz4: 80%
;gpz5: 30%

;use gradient files:
;gpnam3: SINE.100
;gpnam4: SINE.100
;gpnam5: SINE.100

;calculate pulselength according to:
;
; (DeltaOmega * DeltaT) / (width of region[ppm] * SFOn{MHz})
;
;for p13 (sp2/sp8) use Q5/Q5tr pulse (DeltaOmega * DeltaT = 6.180)
;   to cover 41.0ppm on resonance
;   (pulselength: 1.0ms at 600.13 MHz)
;for p14 (sp3) use Q3 pulse (DeltaOmega * DeltaT = 3.448)
;   to cover 22.8ppm on or off resonance

```

```

; (pulselength: 1.0ms at 600.13 MHz)
;for p30 (sp9) use Q3 pulse (DeltaOmega * DeltaT = 3.448)
; to cover 81ppm (N1/N9)
; (pulselength: 700us at 600.13 MHz)

```

```

;$Id: na_h56c56c4n3h_3d,v 1.1.2.2 2006/01/26 08:40:39
ber Exp $

```

3D HCCNH-TOCSY

```

;na_hccnhdigpwg3d
;avance-version (05/10/20)
;HCCNH-TOCSY: H6/8(CCC)NH
;3D sequence with
; inverse correlation for triple resonance
;   via multiple inept transfer steps
;   and DIPSI3 CC spinlock and CN hetero TOCSY
;
; in U:
;   F1(H6,t1) -> F2(C) -> F3(N3,t2) -> F1(NH,t3)
;
; in G:
;   F1(H8,t1) -> F2(C) -> F3(N1,t2) -> F1(NH,t3)
;
;phase sensitive (t1)
;phase sensitive (t2)
;water suppression using watergate sequence
;(use parameterset NA_HCCNHDIGPWG3D)
;
;V. Sklenar, T. Dieckmann, S. E. Butcher & J. Feigon,
; J. Biomol. NMR 7, 83 - 87 (1996)
;
;$CLASS=HighRes
;$DIM=3D
;$TYPE=
;$SUBTYPE=
;$COMMENT=

prosol relations=<triple_na>

#include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>

"p2=p1*2"
"p4=p3*2"
"p22=p21*2"
"d0=3u"
"d4=1s/(cnst2*4)"
"d10=6u"
"d11=30m"
"d12=20u"
"d13=4u"
"d26=1s/(cnst4*4)"

"DELTA=d4-p4-d0*2"
"DELTA1=d26-p16-d16-p11-d12*2-d13"
"DELTA2=d26-d10*2-4u-d12-d13-larger(p2,p4)"

```

aqseq 312

1 ze

d11 pl16:f3

2 d11 do:f3

3 d12 pl9:f1 pl2:f2

d1 cw:f1 ph29

4u do:f1

d12 pl1:f1

(p1 ph3):f1

DELTA UNBLKGRAD

d0

(p4 ph6):f2

d0

(p2 ph1)

d4

(p1 ph2)

d13

p16:gp1

d16

(p3 ph4):f2

d4

(center (p2 ph1) (p4 ph1):f2)

d4 pl26:f2 pl23:f3

(p20 ph7):f2

;begin

DIPSI3:f2

4 (p25*2.722 ph7):f2

(p25*4.389 ph9):f2

(p25*2.778 ph7):f2

(p25*3.056 ph9):f2

(p25*0.333 ph7):f2

(p25*2.556 ph9):f2

(p25*4.000 ph7):f2

(p25*2.722 ph9):f2

(p25*4.111 ph7):f2

(p25*3.778 ph9):f2

(p25*3.889 ph7):f2

(p25*2.889 ph9):f2

(p25*3.000 ph7):f2

(p25*0.333 ph9):f2

(p25*2.500 ph7):f2

(p25*4.050 ph9):f2

(p25*2.830 ph7):f2

(p25*4.389 ph9):f2

(p25*2.722 ph9):f2

(p25*4.389 ph7):f2

(p25*2.778 ph9):f2

(p25*3.056 ph7):f2

(p25*0.333 ph9):f2

(p25*2.556 ph7):f2

(p25*4.000 ph9):f2

(p25*2.722 ph7):f2

(p25*4.111 ph9):f2

(p25*3.778 ph7):f2

(p25*3.889 ph9):f2

(p25*2.889 ph7):f2

(p25*3.000 ph9):f2

(p25*0.333 ph7):f2

(p25*2.500 ph9):f2

(p25*4.050 ph7):f2

(p25*2.830 ph9):f2

(p25*4.389 ph7):f2

(p25*2.722 ph9):f2

(p25*4.389 ph7):f2

(p25*2.778 ph9):f2

(p25*3.056 ph7):f2

(p25*0.333 ph9):f2

(p25*2.556 ph7):f2

(p25*4.000 ph9):f2

(p25*2.722 ph7):f2

(p25*4.111 ph9):f2

(p25*3.778 ph7):f2

(p25*3.889 ph9):f2

(p25*2.889 ph7):f2

(p25*3.000 ph9):f2

(p25*0.333 ph7):f2

(p25*2.500 ph9):f2

(p25*4.050 ph7):f2

(p25*2.830 ph9):f2

(p25*4.389 ph7):f2

(p25*2.722 ph7):f2

(p25*4.389 ph9):f2

(p25*2.778 ph7):f2

(p25*3.056 ph9):f2

(p25*0.333 ph7):f2

(p25*2.556 ph9):f2

(p25*4.000 ph7):f2

(p25*2.722 ph9):f2

(p25*4.111 ph7):f2

(p25*3.778 ph9):f2

(p25*3.889 ph7):f2

(p25*2.889 ph9):f2

(p25*3.000 ph7):f2

(p25*0.333 ph9):f2

(p25*2.500 ph7):f2

(p25*4.050 ph9):f2

(p25*2.830 ph7):f2

(p25*4.389 ph9):f2

lo to 4 times l1

;begin

DIPSI3:f3

5 (p25*2.722 ph7):f2 (p25*2.722 ph7):f3

(p25*4.389 ph9):f2 (p25*4.389 ph9):f3

(p25*2.778 ph7):f2 (p25*2.778 ph7):f3

(p25*3.056 ph9):f2 (p25*3.056 ph9):f3

(p25*0.333 ph7):f2 (p25*0.333 ph7):f3

(p25*2.556 ph9):f2 (p25*2.556 ph9):f3

(p25*4.000 ph7):f2 (p25*4.000 ph7):f3

(p25*2.722 ph9):f2 (p25*2.722 ph9):f3

(p25*4.111 ph7):f2 (p25*4.111 ph7):f3
 (p25*3.778 ph9):f2 (p25*3.778 ph9):f3
 (p25*3.889 ph7):f2 (p25*3.889 ph7):f3
 (p25*2.889 ph9):f2 (p25*2.889 ph9):f3
 (p25*3.000 ph7):f2 (p25*3.000 ph7):f3
 (p25*0.333 ph9):f2 (p25*0.333 ph9):f3
 (p25*2.500 ph7):f2 (p25*2.500 ph7):f3
 (p25*4.050 ph9):f2 (p25*4.050 ph9):f3
 (p25*2.830 ph7):f2 (p25*2.830 ph7):f3
 (p25*4.389 ph9):f2 (p25*4.389 ph9):f3
 (p25*2.722 ph9):f2 (p25*2.722 ph9):f3
 (p25*4.389 ph7):f2 (p25*4.389 ph7):f3
 (p25*2.778 ph9):f2 (p25*2.778 ph9):f3
 (p25*3.056 ph7):f2 (p25*3.056 ph7):f3
 (p25*0.333 ph9):f2 (p25*0.333 ph9):f3
 (p25*2.556 ph7):f2 (p25*2.556 ph7):f3
 (p25*4.000 ph9):f2 (p25*4.000 ph9):f3
 (p25*2.722 ph7):f2 (p25*2.722 ph7):f3
 (p25*4.111 ph9):f2 (p25*4.111 ph9):f3
 (p25*3.778 ph7):f2 (p25*3.778 ph7):f3
 (p25*3.889 ph9):f2 (p25*3.889 ph9):f3
 (p25*2.889 ph7):f2 (p25*2.889 ph7):f3
 (p25*3.000 ph9):f2 (p25*3.000 ph9):f3
 (p25*0.333 ph7):f2 (p25*0.333 ph7):f3
 (p25*2.500 ph9):f2 (p25*2.500 ph9):f3
 (p25*4.050 ph7):f2 (p25*4.050 ph7):f3
 (p25*2.830 ph9):f2 (p25*2.830 ph9):f3
 (p25*4.389 ph7):f2 (p25*4.389 ph7):f3
 (p25*2.722 ph9):f2 (p25*2.722 ph9):f3
 (p25*4.389 ph7):f2 (p25*4.389 ph7):f3
 (p25*2.778 ph9):f2 (p25*2.778 ph9):f3
 (p25*3.056 ph7):f2 (p25*3.056 ph7):f3
 (p25*0.333 ph9):f2 (p25*0.333 ph9):f3
 (p25*2.556 ph7):f2 (p25*2.556 ph7):f3
 (p25*4.000 ph9):f2 (p25*4.000 ph9):f3
 (p25*2.722 ph7):f2 (p25*2.722 ph7):f3
 (p25*4.111 ph9):f2 (p25*4.111 ph9):f3
 (p25*3.778 ph7):f2 (p25*3.778 ph7):f3
 (p25*3.889 ph9):f2 (p25*3.889 ph9):f3
 (p25*2.889 ph7):f2 (p25*2.889 ph7):f3
 (p25*3.000 ph9):f2 (p25*3.000 ph9):f3
 (p25*0.333 ph7):f2 (p25*0.333 ph7):f3
 (p25*2.500 ph9):f2 (p25*2.500 ph9):f3
 (p25*4.050 ph7):f2 (p25*4.050 ph7):f3
 (p25*2.830 ph9):f2 (p25*2.830 ph9):f3
 (p25*4.389 ph7):f2 (p25*4.389 ph7):f3
 (p25*2.722 ph9):f2 (p25*2.722 ph9):f3
 (p25*4.389 ph9):f2 (p25*4.389 ph9):f3
 (p25*2.778 ph7):f2 (p25*2.778 ph7):f3
 (p25*3.056 ph9):f2 (p25*3.056 ph9):f3
 (p25*0.333 ph7):f2 (p25*0.333 ph7):f3
 (p25*2.556 ph9):f2 (p25*2.556 ph9):f3
 (p25*4.000 ph7):f2 (p25*4.000 ph7):f3
 (p25*2.722 ph9):f2 (p25*2.722 ph9):f3
 (p25*4.111 ph7):f2 (p25*4.111 ph7):f3
 (p25*3.778 ph9):f2 (p25*3.778 ph9):f3
 (p25*3.889 ph7):f2 (p25*3.889 ph7):f3
 (p25*2.889 ph9):f2 (p25*2.889 ph9):f3

(p25*3.000 ph7):f2 (p25*3.000 ph7):f3
 (p25*0.333 ph9):f2 (p25*0.333 ph9):f3
 (p25*2.500 ph7):f2 (p25*2.500 ph7):f3
 (p25*4.050 ph9):f2 (p25*4.050 ph9):f3
 (p25*2.830 ph7):f2 (p25*2.830 ph7):f3
 (p25*4.389 ph9):f2 (p25*4.389 ph9):f3
 lo to 5 times l4

;end

DIPS13:f2, DIPS13:f3

d13
 d12 pl2:f2 pl3:f3
 DELTA2
 d10 gron0
 2u groff
 (center (p2 ph1):f1 (p4 ph1):f2)
 d10 gron0*-1
 2u groff
 (p22 ph1):f3
 d26
 (p21 ph5):f3

d13
 p16:gp2
 d16 pl0:f1
 (p29:sp11 ph10:r):f1
 d13
 d12 pl1:f1

(p1 ph1):f1
 DELTA1
 p16:gp3
 d16
 d12 pl0:f1
 (p11:sp1 ph11:r):f1
 d13
 d12 pl1:f1
 (center (p2 ph1):f1 (p22 ph8):f3)

d13
 d12 pl0:f1
 (p11:sp1 ph11:r):f1
 d12

p16:gp3
 d16
 DELTA1 pl16:f3
 4u BLKGRAD
 go=2 ph31 cpd3:f3
 d11 do:f3 mc #0 to 2
 F1PH(ip3 & ip29, id0)
 F2PH(rd0 & rp3 & ip5 & ip8, id10)
 exit

ph1 = 0
 ph2 = 1
 ph3 = 0
 ph4 = 1 3
 ph5 = 1 1 3 3
 ph6 = 0 0 0 2 2 2 2

```

ph7 = 1
ph8 = 0
ph9 = 3
ph10= 2
ph11= 2
ph29=0
ph31= 0 2 2 0

;p0 : 120db
;p1 : f1 channel - power level for pulse (default)
;p2 : f2 channel - power level for pulse (default)
;p3 : f3 channel - power level for pulse (default)
;p9 : f1 channel - power level for presaturation
;p16: f3 channel - power level for CPD/BB decoupling
;p23: f3 channel - power level for TOCSY-spinlock
;p26: f2 channel - power level for TOCSY-spinlock (higher
sel. II)
;   to match p25/pl23
;sp1 : f1 channel - shaped pulse 90 degree (H2O on
resonance)
;sp11: f1 channel - shaped pulse 90 degree (H2O on
resonance)
;p1 : f1 channel - 90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p3 : f2 channel - 90 degree high power pulse
;p4 : f2 channel - 180 degree high power pulse
;p11: f1 channel - 90 degree shaped pulse      [1 msec]
;p16: homospoil/gradient pulse 1              [1 msec]
;p20: f2 channel - trim pulse                  [2 msec]
;p21: f3 channel - 90 degree high power pulse
;p22: f3 channel - 180 degree high power pulse
;p25: f3 channel - 90 degree pulse at pl23
;p29: f1 channel - 90 degree shaped pulse
      [1 msec]
;d0 : incremented delay (F1 in 3D)             [3 usec]
;d1 : relaxation delay; 1-5 * T1
;d4 : 1/(4J(HC))                              [1.3 msec]
;d10: incremented delay (F2 in 3D)            [6 usec]
;d11: delay for disk I/O                      [30 msec]
;d12: delay for power switching                [20 usec]
;d13: short delay                             [4 usec]
;d16: delay for homospoil/gradient recovery
;d26: 1/(4J(HN))                              [2.5 msec]
;cnst2 = J(HC)
;cnst4 = J(HN)
;o2p: 149 ppm
;o3p: 160 ppm
;l1: loop for DIPSI cycle: mixing time C-13 only
;   mixing time = ((p25*217.3) * I1) + (p20) [19.4
msec]
;l4: loop for DIPSI cycle: mixing time C-13/N-15
;   mixing time = ((p25*217.3) * I4)        [58 msec]
;in0: 1/(2 * SW(H)) = DW(H)
;nd0: 2
;in10: 1/(2 * SW(C)) = DW(C)
;nd10: 2
;NS: 16 * n

```

```

;DS: 32
;td1: number of experiments in F1
;td2: number of experiments in F2
;FnMODE: States-TPPI (or TPPI) in F1
;FnMODE: States-TPPI (or TPPI) in F2
;cpd3: decoupling according to sequence defined by
cpdprg3
;pcpd3: f3 channel - 90 degree pulse for decoupling
sequence

;use gradient ratio:  gp 0 : gp 1 : gp 2 : gp 3
;                   3 : 35 : 25 : 60

;for z-only gradients:
;gpz0: 3%
;gpz1: 50%
;gpz2: 60%
;gpz3: 70%

;use gradient files:
;gpname1: SINE.100
;gpname2: SINE.100
;gpname3: SINE.100

;set pl9 to 120dB when presaturation is not required
; use 75 - 80dB to reduce radiation damping

;$Id: na_hccnhdigpwg3d,v 1.1.2.2 2006/03/10 18:08:32
ber Exp $

```

3D HCN

```

;na_hcnetgpsisp3d
;avance-version (05/10/20)
;HCN
;3D sequence with
; inverse correlation for triple resonance using multiple
; inept transfer steps
;
; F1(H6/8) -> F2(C6/8) -> F3(N1/9,t1)
;           -> F2(C6/8,t2) -> F1(H6/8,t3)
; and/or
; F1(H1') -> F2(C1') -> F3(N1/9,t1)
;           -> F2(C1',t2) -> F1(H1',t3)
;
;using shaped pulses on f2 and f3
;phase sensitive (t1)
;phase sensitive using Echo/Antiecho gradient selection
(t2)
;using constant time in t2
;with decoupling during acquisition
;(use parameterset NA_HCNETGPSISP3D)
;
;V. Sklenar, R.D. Peterson, M.R. Rejante & J. Feigon,
; J. Biomol. NMR 3, 721 - 727 (1993)
;

```

```

; $CLASS=HighRes
; $DIM=3D
; $TYPE=
; $SUBTYPE=
; $COMMENT=

prosol relations=<triple_na>

#include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>

"p2=p1*2"
"p4=p3*2"
"p22=p21*2"
"d0=3u"
"d10=3u"
"d11=30m"
"d13=4u"

"in30=in10"

"d30=d22-d13-p16-d16-larger(p14,p30)/2"

"DELTA=d0*2+larger(p2,p4)"
"DELTA1=d22-d13-p16-d16-larger(p14,p30)/2"
"DELTA2=d22-larger(p2,p22)-d16-p16-larger(p14,p30)/2-d10"
"DELTA3=d13+p16+d16+4u

"spoff3=0"
"spoff9=0"

aqseq 321

1 ze
  d11 pl12:f2 pl16:f3
2 d11 do:f2 do:f3
3 d1
  (p1 ph1)
  d4 pl2:f2 pl3:f3
  (center (p2 ph1) (p4 ph1):f2 )
  d4 UNBLKGRAD
  p28 ph1
  d13
  (p1 ph2)

  d13
  p16:gp1
  d16

  (p3 ph3):f2
  DELTA1
  d13
  p16:gp2
  d16 pl0:f2 pl0:f3
  (center (p14:sp3 ph1):f2 (p30:sp9 ph1):f3 )
  d13

p16:gp2
d16 pl2:f2 pl3:f3
DELTA1
(p3 ph2):f2

(p21 ph5):f3
d0
(center (p2 ph1) (p4 ph6):f2 )
d0
(p22 ph1):f3
DELTA
(p21 ph1):f3

d13
p16:gp5
d16

(p3 ph7):f2
d10
(center (p2 ph1) (p22 ph1):f3 )
DELTA2
p16:gp3*EA*-1
d16 pl0:f2 pl0:f3
(center (p14:sp3 ph1):f2 (p30:sp9 ph1):f3 )
d13
p16:gp3*EA
d16 pl2:f2
d30

(center (p1 ph1) (p3 ph8):f2 )
d4
(center (p2 ph1) (p4 ph1):f2 )
d4
(center (p1 ph2) (p3 ph9):f2 )
d4
(center (p2 ph1) (p4 ph1):f2 )
d4
(p1 ph1)
DELTA3
(p2 ph1)
d13
p16:gp4
d16 pl12:f2 pl16:f3
4u BLKGRAD
go=2 ph31 cpd2:f2 cpd3:f3
d11 do:f2 do:f3 mc #0 to 2
  F1PH(rd10 & rd30 & rp6 & rp7 & rp31 & ip4 & ip5, id0)
  F2EA(igrad EA & ip9*2, id10 & dd30 & ip6*2 & ip7*2 &
ip31*2)
exit

ph1=0
ph2=1
ph3=0 0 0 0 2 2 2 2
ph4=0
ph5=0 2
ph6=0
ph7=1

```

```

ph8=0 0 2 2
ph9=3 3 1 1
ph31=0 2 2 0 2 0 2

;p10 : 120db
;p11 : f1 channel - power level for pulse (default)
;p12 : f2 channel - power level for pulse (default)
;p13 : f3 channel - power level for pulse (default)
;p12: f2 channel - power level for CPD/BB decoupling
;p16: f3 channel - power level for CPD/BB decoupling
;sp3: f2 channel - shaped pulse 180 degree (on resonance)
;sp9: f3 channel - shaped pulse 180 degree (on resonance)
;p1 : f1 channel - 90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p3 : f2 channel - 90 degree high power pulse
;p4 : f2 channel - 180 degree high power pulse
;p14: f2 channel - 180 degree shaped pulse
;p16: homospoil/gradient pulse
      [1 msec]
;p21: f3 channel - 90 degree high power pulse
;p22: f3 channel - 180 degree high power pulse
;p28: f1 channel - trim pulse           [1 msec]
;p30: f3 channel - 180 degree shaped pulse for inversion
;d0 : incremented delay (F1 in 3D)
      [3 usec]
;d1 : relaxation delay; 1-5 * T1
;d4 : 1/(4J(HC)) for both:             [1.4 msec]
;      for H1':           [1.56 msec]
;      for H8':           [1.25 msec]
;d10: incremented delay (F2 in 3D)
      [3 usec]
;d11: delay for disk I/O                [30 msec]
;d13: short delay                       [4 usec]
;d16: delay for homospoil/gradient recovery
;d22: 1/(4J(CN))
;      C1':                [19 msec]
;      C6/C8:              [16 msec]
;      both                 [17.5 msec]
;      and constant time delay T(C) = 1/(J(CC)) with J(CC)=40
Hz
;d30: decremented delay (F2 in 3D) = d22-d13-p16-d16-
p14/2
;o2p: 90(C1') or 137(C6/C8) or 113.5 ppm(both)
;o3p: 157 ppm
;in0: 1/(2 * SW(N)) = DW(N)
;nd0: 2
;in10: 1/(2 * SW(C)) = DW(C)
;nd10: 2
;in30: = in10
;NS: 8 * n
;DS: >= 32
;td1: number of experiments in F1
;td2: number of experiments in F2   td2 max = 2 * d30 /
in30
;FnMODE: States-TPPI (or TPPI) in F1
;FnMODE: echo-antiecho in F2
;cpd2: decoupling according to sequence defined by
cpdprg2

```

```

;cpd3: decoupling according to sequence defined by
cpdprg3
;pcpd2: f2 channel - 90 degree pulse for decoupling
sequence
;pcpd3: f3 channel - 90 degree pulse for decoupling
sequence

;use gradient ratio:      gp 1 : gp 2 : gp 3 : gp 4 : gp 5
;                          -40 : 25 : 40 : 20.1 : -50

;for z-only gradients:
;gpz1: -40%
;gpz2: 25%
;gpz3: 40%
;gpz4: 20.1%
;gpz5: -50%

;use gradient files:
;gpname1: SINE.100
;gpname2: SINE.100
;gpname3: SINE.100
;gpname4: SINE.100
;gpname5: SINE.100

;Processing

;F2 reverse: true

;calculate pulselength according to:
;
; (DeltaOmega * DeltaT) / (width of region[ppm] *
SFOn{MHz})
;
;for p14 (sp3) use Q3 pulse (DeltaOmega * DeltaT = 3.448)
; either one pulse to cover 22.8ppm (C1' or C6/C8 in
separate
; experiments)
; (pulselength: 1ms at 600.13 MHz)
; or a twofold modulated pulse to cover 8.8ppm at
90ppm (C1')
; and 137ppm (C6/C8) (both in one experiment)
; (pulselength: 2.6ms at 600.13 MHz)
;for p30 (sp9) use Q3 pulse (DeltaOmega * DeltaT = 3.448)
; to cover 81ppm (N1/N9)
; (pulselength: 700us at 600.13 MHz)

;$Id: na_hcnetgpcisp3d,v 1.1.2.2 2006/03/10 18:08:32 ber
Exp $

```

Appendix C: UUCG Chemical Shifts and Restraints for Structure Calculations

UUCG 16mer Assigned Chemical Shifts

1	7.633	0.000	H6	9
2	2.634	0.000	H5pp	9
3	3.533	0.000	H5p	9
4	3.727	0.000	H4p	9
5	4.418	0.000	H3p	9
6	6.074	0.000	H5	9
7	5.893	0.000	H1p	9
8	7.966	0.000	H6	8
9	5.799	0.000	H5	8
10	6.047	0.000	H1p	8
11	4.606	0.000	H2p	8
12	3.982	0.000	H5pp	8
13	4.196	0.000	H5p	8
14	3.962	0.000	H3p	8
15	4.438	0.000	H4p	8
16	7.719	0.000	H6	7
17	5.671	0.000	H5	7
18	5.607	0.000	H1p	7
19	3.727	0.000	H2p	7
20	4.022	0.000	H5pp	7
21	4.324	0.000	H4p	7
22	4.438	0.000	H5p	7
23	4.458	0.000	H3p	7
24	7.795	0.000	H8	10
25	5.893	0.000	H1p	10
26	5.558	0.000	H3p	10
27	4.773	0.000	H2p	10
28	4.129	0.000	H5pp	10
29	8.235	0.000	H8	11
30	4.207	0.000	H3p	11
34	7.524	0.000	H6	4
35	5.158	0.000	H5	4
36	7.419	0.000	H6	6
37	5.162	0.000	H5	6
38	8.097	0.000	H8	3
39	5.767	0.000	H1p	3
40	8.148	0.000	H8	1
41	5.872	0.000	H1p	1
44	7.634	0.000	H6	15
45	5.452	0.000	H1p	15
46	7.391	0.000	H8	2
47	5.783	0.000	H1p	2
48	7.594	0.000	H8	5
49	5.688	0.000	H1p	5
50	7.610	0.000	H6	12
51	5.214	0.000	H5	12

52 7.682 0.000 H6 14
53 5.188 0.000 H5 14
54 7.513 0.000 H8 13
55 5.700 0.000 H1p 13
56 5.423 0.000 H1p 6
58 5.487 0.000 H1p 4
59 4.654 0.000 H2p 2
60 4.798 0.000 H2p 1
61 4.507 0.000 H2p 4
62 4.874 0.000 H2p 3
64 5.463 0.000 H1p 14
65 5.469 0.000 H1p 12
66 4.355 0.000 H1p 11
67 4.038 0.000 H2p 9
68 4.334 0.000 H5p 10
69 4.343 0.000 H4p 10
70 4.443 0.000 H5p 11
71 4.369 0.000 H2p 11
72 4.340 0.000 H4p 11
73 3.979 0.000 H5pp 6
74 4.179 0.000 H3p 6
75 4.495 0.000 H4p 6
76 4.478 0.000 H5p 6
77 4.437 0.000 H2p 6
78 4.011 0.000 H5pp 5
79 4.044 0.000 H5p 5
80 4.504 0.000 H2p 5
81 4.478 0.000 H3p 5
82 4.378 0.000 H4p 5
83 4.648 0.000 H3p 1
84 4.525 0.000 H4p 1
85 4.466 0.000 H5p 1
86 4.243 0.000 H5pp 1
87 4.645 0.000 H3p 2
88 4.628 0.000 H4p 2
89 4.126 0.000 H5p 2
90 4.117 0.000 H5pp 2
91 4.663 0.000 H3p 3
92 4.190 0.000 H5pp 3
93 4.217 0.000 H5p 3
94 4.387 0.000 H4p 3
95 4.425 0.000 H3p 4
96 4.196 0.000 H4p 4
97 4.035 0.000 H5p 4
98 4.003 0.000 H5pp 4
99 4.437 0.000 H3p 12
100 4.428 0.000 H4p 12
101 4.416 0.000 H5p 12
102 4.407 0.000 H5pp 12
103 4.512 0.000 H2p 13
104 4.438 0.000 H3p 13

105 4.407 0.000 H4p 13
 106 4.049 0.000 H5p 13
 107 4.036 0.000 H5pp 13
 108 4.455 0.000 H2p 14
 109 4.429 0.000 H3p 14
 110 4.394 0.000 H4p 14
 111 4.018 0.000 H5p 14
 112 4.004 0.000 H5pp 14
 113 4.431 0.000 H2p 15
 114 4.398 0.000 H3p 15
 115 4.209 0.000 H4p 15
 116 3.974 0.000 H5pp 15
 117 3.986 0.000 H5p 15
 119 7.738 0.000 H6 16
 120 5.456 0.000 H1p 16
 121 5.462 0.000 H5 16
 122 4.378 0.000 H2p 16
 123 4.335 0.000 H3p 16
 124 4.215 0.000 H4p 16
 125 4.004 0.000 H5p 16
 126 3.974 0.000 H5pp 16
 127 4.445 0.000 H2p 12
 128 9.843 0.000 H1 10
 129 13.425 0.000 H1 11
 130 11.789 0.000 H3 7
 131 12.116 0.000 H3 8
 132 7.075 0.000 H42 6
 133 6.607 0.000 H22 11

UUCG 16mer NOE Restraints

assi (resid 10 and name H2')(resid 10 and name H4') 3.8 1.0 1.0
 assi (resid 2 and name H2')(resid 2 and name H5'') 5.0 1.5 1.5
 assi (resid 13 and name H8)(resid 12 and name H4') 5.8 1.5 1.5
 assi (resid 8 and name H2')(resid 8 and name H4') 3.9 1.0 1.0
 assi (resid 10 and name H1')(resid 10 and name H3') 3.9 1.0 1.0
 assi (resid 3 and name H2')(resid 3 and name H3') 2.5 1.0 1.0
 assi (resid 9 and name H1')(resid 10 and name H3') 5.0 1.0 1.0
 assi (resid 11 and name H1')(resid 11 and name H4') 3.5 1.0 1.0
 assi (resid 7 and name H5')(resid 7 and name H5'') 1.8 0.5 0.5
 assi (resid 9 and name H3')(resid 9 and name H2') 2.5 0.5 0.5
 assi (resid 13 and name H3')(resid 13 and name H5') 3.8 1.0 1.0
 assi (resid 9 and name H5)(resid 9 and name H5') 6.2 1.0 1.0
 assi (resid 11 and name H2')(resid 11 and name H1') 2.8 1.0 1.0
 assi (resid 8 and name H5)(resid 8 and name H5') 5.9 1.0 1.0
 assi (resid 3 and name H3')(resid 3 and name H4') 3.0 1.0 1.0
 assi (resid 8 and name H5)(resid 8 and name H5'') 6.6 1.5 1.5
 assi (resid 9 and name H6)(resid 8 and name H5') 5.0 1.5 1.5
 assi (resid 3 and name H2')(resid 3 and name H4') 3.8 1.0 1.0
 assi (resid 8 and name H5)(resid 8 and name H4') 6.8 1.0 1.0

assi (resid 6 and name H5')(resid 6 and name H3') 3.8 1.0 1.0
assi (resid 9 and name H6)(resid 8 and name H2') 4.0 1.5 1.5
assi (resid 8 and name H5)(resid 8 and name H2') 4.5 1.0 1.0
assi (resid 8 and name H1')(resid 8 and name H3') 3.9 1.0 1.0
assi (resid 12 and name H6)(resid 11 and name H3') 3.2 1.5 1.5
assi (resid 3 and name H3')(resid 3 and name H5'') 2.5 1.0 1.0
assi (resid 9 and name H6)(resid 8 and name H4') 4.6 1.5 1.5
assi (resid 2 and name H3')(resid 2 and name H5') 3.8 1.5 1.5
assi (resid 5 and name H3')(resid 6 and name H3') 5.8 1.5 1.5
assi (resid 13 and name H4')(resid 13 and name H5') 2.3 1.0 1.0
assi (resid 2 and name H1')(resid 2 and name H2') 2.8 1.0 1.0
assi (resid 2 and name H1')(resid 2 and name H4') 3.5 1.0 1.0
assi (resid 11 and name H2')(resid 11 and name H4') 3.8 1.0 1.0
assi (resid 13 and name H8)(resid 12 and name H3') 2.4 1.5 1.5
assi (resid 13 and name H8)(resid 12 and name H2') 3.5 1.5 1.5
assi (resid 8 and name H5)(resid 7 and name H5') 5.2 1.0 1.0
assi (resid 12 and name H6)(resid 11 and name H2') 2.2 1.5 1.5
assi (resid 14 and name H4')(resid 14 and name H5'') 4.0 2.5 2.5
assi (resid 1 and name H2')(resid 1 and name H5') 5.3 1.5 1.5
assi (resid 3 and name H1')(resid 3 and name H2') 2.8 1.0 1.0
assi (resid 3 and name H1')(resid 3 and name H5'') 5.0 1.5 1.5
assi (resid 2 and name H2')(resid 2 and name H5') 5.3 1.5 1.5
assi (resid 3 and name H1')(resid 3 and name H5') 3.5 1.5 1.5
assi (resid 5 and name H4')(resid 5 and name H5') 4.0 1.5 1.5
assi (resid 3 and name H1')(resid 3 and name H3') 4.0 1.5 1.5
assi (resid 4 and name H3')(resid 4 and name H5') 3.8 1.0 1.0
assi (resid 1 and name H2')(resid 1 and name H4') 3.8 1.0 1.0
assi (resid 13 and name H3')(resid 13 and name H5'') 2.5 1.0 1.0
assi (resid 8 and name H5)(resid 8 and name H3') 6.4 1.0 1.0
assi (resid 9 and name H1')(resid 9 and name H3') 3.7 0.5 0.5
assi (resid 4 and name H3')(resid 4 and name H4') 3.0 1.0 1.0
assi (resid 10 and name H1')(resid 10 and name H5'') 4.7 1.0 1.0
assi (resid 12 and name H6)(resid 11 and name H5') 7.0 1.5 1.5
assi (resid 7 and name H6)(resid 6 and name H5') 6.4 1.5 1.5
assi (resid 8 and name H6)(resid 9 and name H5') 7.3 1.5 1.5
assi (resid 2 and name H1')(resid 2 and name H5') 5.0 1.5 1.5
assi (resid 11 and name H8)(resid 9 and name H4') 7.0 1.5 1.5
assi (resid 1 and name H2')(resid 1 and name H3') 2.5 1.0 1.0
assi (resid 14 and name H4')(resid 14 and name H5') 2.2 1.0 1.0
assi (resid 3 and name H1')(resid 3 and name H4') 3.5 1.0 1.0
assi (resid 16 and name H2')(resid 16 and name H3') 2.4 1.0 1.0
assi (resid 13 and name H8)(resid 12 and name H5') 6.0 1.5 1.5
assi (resid 8 and name H4')(resid 8 and name H5'') 2.4 1.0 1.0
assi (resid 14 and name H2')(resid 14 and name H5') 5.4 1.5 1.5
assi (resid 9 and name H5)(resid 9 and name H2') 4.5 1.0 1.0
assi (resid 10 and name H1')(resid 9 and name H4') 4.0 1.0 1.0
assi (resid 8 and name H1')(resid 8 and name H5') 4.6 1.0 1.0
assi (resid 10 and name H8)(resid 9 and name H5') 3.3 1.0 1.0
assi (resid 14 and name H2')(resid 14 and name H5'') 4.7 1.0 1.0
assi (resid 7 and name H3')(resid 7 and name H5'') 2.9 1.0 1.0
assi (resid 8 and name H1')(resid 8 and name H5'') 5.0 1.5 1.5

assi (resid 9 and name H1')(resid 9 and name H2') 3.0 0.5 0.5
assi (resid 9 and name H1')(resid 9 and name H4') 3.2 1.0 1.0
assi (resid 15 and name H2')(resid 15 and name H5'') 4.7 1.5 1.5
assi (resid 1 and name H1')(resid 1 and name H3') 3.9 1.0 1.0
assi (resid 8 and name H1')(resid 8 and name H2') 3.0 0.5 0.5
assi (resid 1 and name H1')(resid 1 and name H2') 2.7 1.0 1.0
assi (resid 8 and name H1')(resid 8 and name H4') 3.2 0.5 0.5
assi (resid 9 and name H6)(resid 9 and name H5') 3.9 1.5 1.5
assi (resid 4 and name H6)(resid 4 and name H4') 4.2 1.0 1.0
assi (resid 6 and name H6)(resid 6 and name H4') 4.2 1.0 1.0
assi (resid 12 and name H6)(resid 12 and name H5') 4.0 1.5 1.5
assi (resid 11 and name H8)(resid 10 and name H2') 4.0 1.0 1.0
assi (resid 9 and name H5)(resid 9 and name H1') 5.4 1.0 1.0
assi (resid 8 and name H1')(resid 8 and name H5) 5.5 1.0 1.0
assi (resid 11 and name H8)(resid 10 and name H3') 3.8 0.5 0.5
assi (resid 10 and name H1')(resid 10 and name H2') 2.9 0.5 0.5
assi (resid 6 and name H6)(resid 6 and name H2') 3.7 2.5 2.5
assi (resid 13 and name H8)(resid 13 and name H5'') 3.8 1.0 1.0
assi (resid 9 and name H6)(resid 9 and name H5'') 4.2 1.5 1.5
assi (resid 15 and name H6)(resid 15 and name H5'') 4.0 1.0 1.0
assi (resid 5 and name H8)(resid 5 and name H5'') 3.8 1.0 1.0
assi (resid 5 and name H3')(resid 5 and name H5'') 2.5 1.0 1.0
assi (resid 15 and name H6)(resid 15 and name H2') 3.7 1.0 1.0
assi (resid 14 and name H6)(resid 14 and name H5) 2.5 1.0 1.0
assi (resid 5 and name H2')(resid 5 and name H5') 5.4 1.5 1.5
assi (resid 16 and name H6)(resid 16 and name H1') 3.5 1.0 1.0
assi (resid 10 and name H8)(resid 10 and name H2') 4.0 1.0 1.0
assi (resid 4 and name H6)(resid 4 and name H3') 3.2 1.0 1.0
assi (resid 5 and name H8)(resid 5 and name H4') 4.3 1.0 1.0
assi (resid 6 and name H6)(resid 6 and name H3') 3.2 1.5 1.5
assi (resid 14 and name H6)(resid 14 and name H5') 4.0 1.0 1.0
assi (resid 5 and name H8)(resid 5 and name H3') 3.0 1.0 1.0
assi (resid 15 and name H6)(resid 15 and name H5') 4.0 1.0 1.0
assi (resid 15 and name H6)(resid 15 and name H4') 4.2 1.0 1.0
assi (resid 12 and name H6)(resid 11 and name H1') 5.0 1.5 1.5
assi (resid 12 and name H6)(resid 12 and name H3') 3.2 1.5 1.5
assi (resid 15 and name H6)(resid 15 and name H3') 3.2 1.5 1.5
assi (resid 5 and name H8)(resid 5 and name H1') 3.8 1.0 1.0
assi (resid 1 and name H8)(resid 1 and name H3') 3.0 1.0 1.0
assi (resid 11 and name H8)(resid 11 and name H5') 4.2 1.0 1.0
assi (resid 4 and name H6)(resid 3 and name H1') 5.2 1.0 1.0
assi (resid 4 and name H2')(resid 4 and name H5'') 4.7 1.0 1.0
assi (resid 10 and name H8)(resid 9 and name H1') 7.1 1.5 1.5
assi (resid 10 and name H8)(resid 10 and name H1') 2.4 1.0 1.0
assi (resid 6 and name H6)(resid 6 and name H5') 4.0 1.0 1.0
assi (resid 9 and name H6)(resid 10 and name H3') 7.3 1.5 1.5
assi (resid 9 and name H6)(resid 9 and name H2') 2.7 1.0 1.0
assi (resid 5 and name H2')(resid 5 and name H5'') 4.7 1.0 1.0
assi (resid 1 and name H8)(resid 1 and name H4') 4.3 1.0 1.0
assi (resid 3 and name H8)(resid 2 and name H1') 5.0 1.0 1.0
assi (resid 9 and name H6)(resid 8 and name H1') 6.0 1.5 1.5

assi (resid 13 and name H8)(resid 13 and name H3') 3.0 1.0 1.0
assi (resid 4 and name H6)(resid 4 and name H2') 3.7 1.5 1.5
assi (resid 15 and name H6)(resid 14 and name H1') 5.0 1.5 1.5
assi (resid 13 and name H8)(resid 13 and name H2') 3.8 1.0 1.0
assi (resid 13 and name H8)(resid 12 and name H1') 5.0 1.5 1.5
assi (resid 10 and name H1')(resid 10 and name H4') 3.1 0.5 0.5
assi (resid 3 and name H8)(resid 3 and name H5') 4.2 1.5 1.5
assi (resid 8 and name H6)(resid 7 and name H1') 5.5 1.0 1.0
assi (resid 9 and name H6)(resid 9 and name H4') 4.3 1.5 1.5
assi (resid 12 and name H6)(resid 12 and name H2') 4.0 1.5 1.5
assi (resid 5 and name H3')(resid 6 and name H5'') 5.2 1.5 1.5
assi (resid 12 and name H6)(resid 12 and name H4') 4.2 1.0 1.0
assi (resid 6 and name H6)(resid 6 and name H5'') 4.0 1.5 1.5
assi (resid 3 and name H8)(resid 3 and name H3') 3.0 1.5 1.5
assi (resid 12 and name H6)(resid 12 and name H5'') 4.0 1.5 1.5
assi (resid 11 and name H8)(resid 12 and name H5) 4.0 1.5 1.5
assi (resid 7 and name H6)(resid 7 and name H5) 2.4 0.5 0.5
assi (resid 6 and name H5')(resid 5 and name H5'') 6.2 1.5 1.5
assi (resid 6 and name H6)(resid 6 and name H5) 2.5 1.0 1.0
assi (resid 13 and name H8)(resid 13 and name H1') 3.8 1.0 1.0
assi (resid 6 and name H5')(resid 5 and name H5') 7.0 1.5 1.5
assi (resid 1 and name H8)(resid 1 and name H5') 4.0 1.5 1.5
assi (resid 7 and name H6)(resid 7 and name H2') 3.9 1.0 1.0
assi (resid 10 and name H8)(resid 10 and name H3') 5.4 1.5 1.5
assi (resid 6 and name H6)(resid 6 and name H1') 3.5 1.0 1.0
assi (resid 13 and name H8)(resid 13 and name H5') 4.2 1.5 1.5
assi (resid 4 and name H6)(resid 4 and name H5') 4.3 1.5 1.5
assi (resid 13 and name H8)(resid 13 and name H4') 4.3 1.0 1.0
assi (resid 5 and name H8)(resid 5 and name H2') 3.8 1.0 1.0
assi (resid 14 and name H6)(resid 14 and name H5'') 4.0 1.0 1.0
assi (resid 5 and name H8)(resid 5 and name H5') 4.2 1.0 1.0
assi (resid 14 and name H6)(resid 14 and name H3') 3.2 1.0 1.0
assi (resid 11 and name H5')(resid 11 and name H1') 4.9 1.0 1.0
assi (resid 13 and name H2')(resid 13 and name H5'') 4.7 1.0 1.0
assi (resid 4 and name H2')(resid 4 and name H5') 5.4 1.5 1.5
assi (resid 14 and name H6)(resid 14 and name H2') 3.7 1.0 1.0
assi (resid 7 and name H3')(resid 7 and name H4') 3.0 1.0 1.0
assi (resid 13 and name H1')(resid 13 and name H3') 4.0 1.0 1.0
assi (resid 11 and name H1)(resid 10 and name H1) 5.3 1.0 1.0
assi (resid 13 and name H2')(resid 13 and name H5') 5.3 1.5 1.5
assi (resid 11 and name H8)(resid 10 and name H1') 6.8 1.0 1.0
assi (resid 11 and name H8)(resid 11 and name H2') 3.8 1.0 1.0
assi (resid 16 and name H6)(resid 16 and name H5'') 4.0 1.0 1.0
assi (resid 14 and name H6)(resid 14 and name H4') 4.2 1.0 1.0
assi (resid 7 and name H6)(resid 7 and name H1') 3.5 1.0 1.0
assi (resid 4 and name H6)(resid 4 and name H5'') 4.0 1.0 1.0
assi (resid 11 and name H8)(resid 11 and name H3') 3.0 1.0 1.0
assi (resid 1 and name H1')(resid 1 and name H4') 3.5 1.0 1.0
assi (resid 7 and name H6)(resid 7 and name H5'') 4.0 1.0 1.0
assi (resid 11 and name H8)(resid 11 and name H1') 3.8 1.5 1.5
assi (resid 7 and name H6)(resid 7 and name H4') 4.1 1.0 1.0

assi (resid 16 and name H6)(resid 15 and name H1') 5.0 1.5 1.5
assi (resid 9 and name H6)(resid 9 and name H1') 3.5 1.0 1.0
assi (resid 4 and name H6)(resid 4 and name H5) 2.5 1.0 1.0
assi (resid 8 and name H6)(resid 8 and name H5) 2.4 0.5 0.5
assi (resid 8 and name H6)(resid 8 and name H5') 3.8 0.5 0.5
assi (resid 4 and name H6)(resid 4 and name H1') 3.5 1.0 1.0
assi (resid 14 and name H6)(resid 13 and name H1') 5.0 1.5 1.5
assi (resid 3 and name H8)(resid 3 and name H1') 3.8 1.0 1.0
assi (resid 1 and name H8)(resid 1 and name H1') 3.8 1.0 1.0
assi (resid 7 and name H6)(resid 7 and name H3') 3.0 1.0 1.0
assi (resid 3 and name H8)(resid 3 and name H4') 4.3 1.5 1.5
assi (resid 10 and name H8)(resid 10 and name H4') 5.4 1.5 1.5
assi (resid 11 and name H5')(resid 11 and name H4') 2.3 1.0 1.0
assi (resid 7 and name H6)(resid 7 and name H5') 3.7 1.0 1.0
assi (resid 12 and name H6)(resid 12 and name H1') 3.5 1.0 1.0
assi (resid 7 and name H6)(resid 6 and name H1') 5.2 1.0 1.0
assi (resid 15 and name H6)(resid 15 and name H1') 3.5 1.0 1.0
assi (resid 8 and name H6)(resid 8 and name H3') 4.1 1.0 1.0
assi (resid 8 and name H6)(resid 8 and name H2') 2.5 0.5 0.5
assi (resid 8 and name H6)(resid 8 and name H5'') 4.3 1.0 1.0
assi (resid 8 and name H6)(resid 8 and name H1') 3.7 0.5 0.5
assi (resid 16 and name H6)(resid 16 and name H4') 4.2 1.0 1.0
assi (resid 16 and name H6)(resid 16 and name H2') 3.7 1.0 1.0
assi (resid 9 and name H6)(resid 9 and name H5) 2.4 1.5 1.5
assi (resid 3 and name H8)(resid 3 and name H2') 3.8 1.0 1.0
assi (resid 9 and name H6)(resid 9 and name H3') 4.5 1.5 1.5
assi (resid 10 and name H8)(resid 10 and name H5'') 6.3 1.5 1.5
assi (resid 1 and name H8)(resid 1 and name H2') 3.8 1.0 1.0
assi (resid 12 and name H6)(resid 12 and name H5) 2.5 0.5 0.5
assi (resid 4 and name H6)(resid 3 and name H4') 5.8 1.5 1.5
assi (resid 5 and name H8)(resid 4 and name H1') 5.0 1.0 1.0
assi (resid 1 and name H4')(resid 1 and name H5') 2.3 1.0 1.0
assi (resid 3 and name H8)(resid 2 and name H2') 2.0 0.5 0.5
assi (resid 12 and name H2')(resid 12 and name H5'') 4.7 1.0 1.0
assi (resid 14 and name H2')(resid 14 and name H4') 3.8 1.0 1.0
assi (resid 6 and name H6)(resid 5 and name H4') 5.8 1.5 1.5
assi (resid 16 and name H6)(resid 16 and name H3') 3.2 1.0 1.0
assi (resid 6 and name H6)(resid 5 and name H1') 5.0 1.5 1.5
assi (resid 6 and name H6)(resid 5 and name H2') 2.2 1.0 1.0
assi (resid 12 and name H2')(resid 13 and name H4') 4.2 1.5 1.5
assi (resid 13 and name H3')(resid 13 and name H4') 3.0 1.5 1.5
assi (resid 2 and name H8)(resid 2 and name H2') 3.8 1.0 1.0
assi (resid 3 and name H8)(resid 3 and name H5'') 3.8 1.0 1.0
assi (resid 14 and name H6)(resid 13 and name H3') 3.3 1.5 1.5
assi (resid 16 and name H6)(resid 16 and name H5) 2.5 1.0 1.0
assi (resid 15 and name H6)(resid 14 and name H5'') 5.7 1.5 1.5
assi (resid 6 and name H6)(resid 5 and name H3') 3.3 1.5 1.5
assi (resid 16 and name H6)(resid 16 and name H5') 4.0 1.0 1.0
assi (resid 11 and name H8)(resid 11 and name H4') 4.3 1.0 1.0
assi (resid 6 and name H5')(resid 5 and name H4') 4.8 1.5 1.5
assi (resid 14 and name H6)(resid 13 and name H5'') 5.7 1.5 1.5

assi (resid 13 and name H2')(resid 13 and name H4') 3.8 1.0 1.0
assi (resid 14 and name H6)(resid 14 and name H1') 3.5 1.0 1.0
assi (resid 12 and name H2')(resid 12 and name H5') 5.3 1.5 1.5
assi (resid 13 and name H2')(resid 13 and name H3') 2.4 1.0 1.0
assi (resid 4 and name H6)(resid 3 and name H3') 3.3 1.0 1.0
assi (resid 4 and name H6)(resid 3 and name H5') 7.0 1.5 1.5
assi (resid 7 and name H5')(resid 7 and name H4') 2.4 1.0 1.0
assi (resid 9 and name H5)(resid 9 and name H4') 7.6 1.5 1.5
assi (resid 1 and name H1')(resid 1 and name H5') 5.0 1.5 1.5
assi (resid 9 and name H5)(resid 9 and name H3') 6.7 1.5 1.5
assi (resid 14 and name H3')(resid 14 and name H4') 3.0 1.0 1.0
assi (resid 4 and name H6)(resid 3 and name H5'') 5.7 1.5 1.5
assi (resid 6 and name H4')(resid 6 and name H2') 3.8 1.0 1.0
assi (resid 15 and name H2')(resid 15 and name H5') 5.3 1.5 1.5
assi (resid 6 and name H2')(resid 6 and name H5'') 4.7 1.5 1.5
assi (resid 11 and name H5')(resid 11 and name H3') 3.7 1.0 1.0
assi (resid 10 and name H1')(resid 9 and name H5') 4.0 1.0 1.0
assi (resid 15 and name H2')(resid 16 and name H5'') 4.5 1.0 1.0
assi (resid 5 and name H2')(resid 5 and name H4') 3.8 1.0 1.0
assi (resid 3 and name H8)(resid 2 and name H3') 3.0 1.5 1.5
assi (resid 4 and name H2')(resid 4 and name H3') 2.4 1.0 1.0
assi (resid 6 and name H6)(resid 5 and name H5') 7.0 1.5 1.5
assi (resid 9 and name H1')(resid 9 and name H5') 4.7 1.0 1.0
assi (resid 6 and name H5')(resid 6 and name H5'') 1.7 0.5 0.5
assi (resid 7 and name H6)(resid 6 and name H4') 5.4 1.0 1.0
assi (resid 2 and name H3')(resid 2 and name H5'') 2.5 1.0 1.0
assi (resid 16 and name H2')(resid 16 and name H5') 5.4 1.5 1.5
assi (resid 5 and name H1')(resid 5 and name H5') 5.0 1.5 1.5
assi (resid 8 and name H2')(resid 8 and name H5') 4.3 1.0 1.0
assi (resid 15 and name H3')(resid 16 and name H5'') 5.2 1.5 1.5
assi (resid 10 and name H4')(resid 10 and name H5') 2.3 0.5 0.5
assi (resid 8 and name H6)(resid 7 and name H5') 6.2 1.0 1.0
assi (resid 13 and name H1')(resid 13 and name H5'') 5.1 1.5 1.5
assi (resid 14 and name H6)(resid 13 and name H2') 2.2 1.0 1.0
assi (resid 13 and name H8)(resid 12 and name H5'') 6.0 1.5 1.5
assi (resid 8 and name H6)(resid 7 and name H4') 4.0 1.0 1.0
assi (resid 10 and name H2')(resid 10 and name H5'') 4.5 1.0 1.0
assi (resid 6 and name H4')(resid 6 and name H3') 3.0 1.0 1.0
assi (resid 14 and name H6)(resid 13 and name H4') 5.8 1.5 1.5
assi (resid 7 and name H5)(resid 7 and name H5'') 6.2 1.5 1.5
assi (resid 16 and name H2')(resid 16 and name H5'') 4.7 1.5 1.5
assi (resid 13 and name H1')(resid 13 and name H5') 5.0 1.5 1.5
assi (resid 5 and name H4')(resid 5 and name H5'') 2.6 1.0 1.0
assi (resid 10 and name H2')(resid 10 and name H5') 5.0 1.0 1.0
assi (resid 2 and name H3')(resid 2 and name H2') 2.4 1.0 1.0
assi (resid 12 and name H5'')(resid 13 and name H5') 6.2 1.5 1.5
assi (resid 8 and name H6)(resid 7 and name H5'') 5.3 1.0 1.0
assi (resid 4 and name H3')(resid 4 and name H5'') 2.5 1.0 1.0
assi (resid 6 and name H2')(resid 6 and name H3') 2.4 1.0 1.0
assi (resid 8 and name H6)(resid 7 and name H3') 4.7 1.5 1.5
assi (resid 1 and name H3')(resid 1 and name H5') 3.7 1.0 1.0

assi (resid 3 and name H2')(resid 3 and name H5') 5.3 1.5 1.5
assi (resid 7 and name H5)(resid 7 and name H4') 6.5 1.0 1.0
assi (resid 7 and name H5)(resid 7 and name H2') 5.3 1.0 1.0
assi (resid 6 and name H5)(resid 6 and name H4') 6.5 1.5 1.5
assi (resid 1 and name H3')(resid 1 and name H4') 3.0 1.0 1.0
assi (resid 6 and name H5)(resid 5 and name H3') 3.8 1.5 1.5
assi (resid 11 and name H8)(resid 10 and name H5') 5.0 1.5 1.5
assi (resid 2 and name H2')(resid 2 and name H4') 3.8 1.0 1.0
assi (resid 5 and name H1')(resid 5 and name H4') 3.4 1.0 1.0
assi (resid 6 and name H5)(resid 6 and name H5') 6.3 1.5 1.5
assi (resid 4 and name H5)(resid 4 and name H2') 5.2 1.5 1.5
assi (resid 9 and name H6)(resid 8 and name H3') 2.2 1.5 1.5
assi (resid 9 and name H6)(resid 8 and name H5'') 3.4 1.5 1.5
assi (resid 3 and name H4')(resid 3 and name H5') 2.3 0.5 0.5
assi (resid 8 and name H4')(resid 8 and name H3') 2.7 1.0 1.0
assi (resid 15 and name H3')(resid 15 and name H4') 3.0 1.0 1.0
assi (resid 3 and name H2')(resid 3 and name H5'') 4.7 1.5 1.5
assi (resid 2 and name H3')(resid 2 and name H4') 3.0 1.0 1.0
assi (resid 12 and name H5'')(resid 13 and name H5'') 6.8 1.5 1.5
assi (resid 13 and name H4')(resid 13 and name H5'') 2.6 0.5 0.5
assi (resid 15 and name H3')(resid 15 and name H5'') 2.5 1.0 1.0
assi (resid 7 and name H5)(resid 7 and name H5') 6.0 1.0 1.0
assi (resid 10 and name H8)(resid 9 and name H4') 4.5 1.5 1.5
assi (resid 5 and name H1')(resid 5 and name H5'') 5.0 1.5 1.5
assi (resid 13 and name H1')(resid 13 and name H4') 3.5 1.0 1.0
assi (resid 4 and name H4')(resid 4 and name H5') 2.3 1.0 1.0
assi (resid 6 and name H1')(resid 6 and name H4') 3.4 1.0 1.0
assi (resid 2 and name H4')(resid 2 and name H5'') 2.6 1.0 1.0
assi (resid 5 and name H1')(resid 5 and name H2') 2.8 1.0 1.0
assi (resid 8 and name H2')(resid 8 and name H3') 2.4 1.0 1.0
assi (resid 16 and name H5)(resid 16 and name H3') 5.2 1.5 1.5
assi (resid 13 and name H1')(resid 13 and name H2') 2.8 1.0 1.0
assi (resid 2 and name H4')(resid 2 and name H5') 2.3 0.5 0.5
assi (resid 12 and name H5)(resid 12 and name H4') 6.6 1.5 1.5
assi (resid 4 and name H4')(resid 4 and name H5'') 2.6 1.0 1.0
assi (resid 16 and name H3')(resid 16 and name H4') 3.0 1.0 1.0
assi (resid 9 and name H3')(resid 9 and name H4') 2.7 1.0 1.0
assi (resid 16 and name H5)(resid 16 and name H5') 6.3 1.5 1.5
assi (resid 16 and name H5)(resid 16 and name H4') 6.6 1.5 1.5
assi (resid 5 and name H1')(resid 5 and name H3') 3.9 1.0 1.0
assi (resid 8 and name H4')(resid 8 and name H5') 2.4 1.0 1.0
assi (resid 5 and name H1')(resid 6 and name H5') 6.0 1.5 1.5
assi (resid 12 and name H5)(resid 12 and name H3') 5.2 1.5 1.5
assi (resid 15 and name H1')(resid 16 and name H5'') 7.0 1.5 1.5
assi (resid 3 and name H3')(resid 3 and name H5') 3.8 1.0 1.0
assi (resid 4 and name H1')(resid 4 and name H4') 4.0 1.0 1.0
assi (resid 4 and name H1')(resid 4 and name H3') 3.5 1.0 1.0
assi (resid 11 and name H8)(resid 10 and name H4') 6.5 1.5 1.5
assi (resid 16 and name H1')(resid 16 and name H5') 4.9 1.0 1.0
assi (resid 4 and name H5)(resid 4 and name H5'') 6.3 1.5 1.5
assi (resid 16 and name H3')(resid 16 and name H5') 3.8 1.0 1.0

assi (resid 12 and name H5)(resid 12 and name H2') 5.2 1.5 1.5
assi (resid 16 and name H4')(resid 16 and name H5') 2.3 1.0 1.0
assi (resid 15 and name H3')(resid 15 and name H5') 3.8 1.0 1.0
assi (resid 14 and name H1')(resid 14 and name H4') 3.5 1.0 1.0
assi (resid 16 and name H4')(resid 16 and name H5'') 2.6 1.0 1.0
assi (resid 8 and name H2')(resid 8 and name H5'') 4.0 1.0 1.0
assi (resid 9 and name H4')(resid 9 and name H5'') 2.6 0.5 0.5
assi (resid 7 and name H1')(resid 7 and name H5'') 5.2 1.0 1.0
assi (resid 7 and name H5')(resid 7 and name H2') 4.0 1.5 1.5
assi (resid 12 and name H5)(resid 12 and name H5') 6.3 1.5 1.5
assi (resid 15 and name H4')(resid 15 and name H5'') 2.6 1.0 1.0
assi (resid 14 and name H1')(resid 14 and name H5'') 5.0 1.5 1.5
assi (resid 16 and name H1')(resid 16 and name H5'') 5.0 1.5 1.5
assi (resid 3 and name H4')(resid 3 and name H5'') 2.6 1.0 1.0
assi (resid 16 and name H1')(resid 16 and name H2') 2.8 1.0 1.0
assi (resid 11 and name H2')(resid 11 and name H3') 2.4 1.0 1.0
assi (resid 14 and name H1')(resid 14 and name H5') 5.0 1.0 1.0
assi (resid 10 and name H5')(resid 10 and name H5'') 1.7 0.5 0.5
assi (resid 6 and name H1')(resid 6 and name H5') 5.0 1.0 1.0
assi (resid 6 and name H1')(resid 6 and name H2') 2.8 1.0 1.0
assi (resid 16 and name H1')(resid 16 and name H4') 3.5 1.0 1.0
assi (resid 9 and name H2')(resid 9 and name H5'') 3.5 0.5 0.5
assi (resid 15 and name H4')(resid 15 and name H5') 2.3 1.0 1.0
assi (resid 9 and name H3')(resid 9 and name H5'') 2.3 0.5 0.5
assi (resid 15 and name H1')(resid 15 and name H3') 3.9 1.0 1.0
assi (resid 6 and name H1')(resid 6 and name H3') 4.0 1.0 1.0
assi (resid 11 and name H4')(resid 11 and name H3') 3.0 1.0 1.0
assi (resid 16 and name H1')(resid 16 and name H3') 3.5 1.0 1.0
assi (resid 16 and name H5)(resid 16 and name H2') 5.2 1.5 1.5
assi (resid 4 and name H5)(resid 4 and name H4') 6.6 1.5 1.5
assi (resid 4 and name H1')(resid 4 and name H5') 5.0 1.0 1.0
assi (resid 4 and name H5)(resid 4 and name H5') 6.3 1.5 1.5
assi (resid 9 and name H5')(resid 9 and name H5'') 1.8 0.5 0.5
assi (resid 7 and name H1')(resid 7 and name H2') 2.8 1.0 1.0
assi (resid 7 and name H4')(resid 7 and name H5'') 2.4 1.0 1.0
assi (resid 6 and name H1')(resid 5 and name H3') 6.0 1.5 1.5
assi (resid 7 and name H1')(resid 7 and name H5') 4.6 1.0 1.0
assi (resid 12 and name H1')(resid 12 and name H2') 2.8 1.0 1.0
assi (resid 6 and name H5)(resid 6 and name H3') 5.2 1.5 1.5
assi (resid 12 and name H1')(resid 12 and name H4') 3.5 1.5 1.5
assi (resid 4 and name H2')(resid 4 and name H4') 3.8 1.0 1.0
assi (resid 14 and name H5)(resid 14 and name H5') 6.3 1.5 1.5
assi (resid 4 and name H1')(resid 4 and name H5'') 5.0 1.0 1.0
assi (resid 9 and name H2')(resid 9 and name H4') 3.8 0.5 0.5
assi (resid 7 and name H1')(resid 7 and name H3') 3.9 1.0 1.0
assi (resid 12 and name H1')(resid 12 and name H3') 4.0 1.0 1.0
assi (resid 14 and name H1')(resid 14 and name H2') 2.8 1.0 1.0
assi (resid 8 and name H4')(resid 7 and name H4') 8.0 1.5 1.5
assi (resid 6 and name H1')(resid 6 and name H5'') 5.0 1.5 1.5
assi (resid 10 and name H3')(resid 10 and name H5'') 2.4 0.5 0.5
assi (resid 14 and name H1')(resid 14 and name H3') 4.0 1.0 1.0

assi (resid 11 and name H1')(resid 11 and name H3') 4.0 1.0 1.0
assi (resid 10 and name H3')(resid 10 and name H4') 3.0 0.5 0.5
assi (resid 15 and name H1')(resid 15 and name H2') 2.8 1.0 1.0
assi (resid 11 and name H5')(resid 11 and name H2') 5.3 1.5 1.5
assi (resid 9 and name H4')(resid 9 and name H5') 2.3 0.5 0.5
assi (resid 15 and name H1')(resid 15 and name H5') 5.0 1.0 1.0
assi (resid 7 and name H1')(resid 7 and name H4') 3.3 1.0 1.0
assi (resid 14 and name H5)(resid 14 and name H2') 5.2 1.5 1.5
assi (resid 6 and name H5)(resid 6 and name H2') 5.2 1.5 1.5
assi (resid 12 and name H5)(resid 12 and name H5'') 6.3 1.5 1.5
assi (resid 15 and name H2')(resid 15 and name H4') 3.8 1.0 1.0
assi (resid 5 and name H3')(resid 5 and name H4') 3.0 1.0 1.0
assi (resid 8 and name H4')(resid 7 and name H2') 6.2 1.5 1.5
assi (resid 10 and name H3')(resid 10 and name H5') 3.2 0.5 0.5
assi (resid 10 and name H4')(resid 10 and name H5'') 2.4 0.5 0.5
assi (resid 4 and name H5)(resid 4 and name H3') 5.2 1.5 1.5
assi (resid 7 and name H4')(resid 7 and name H2') 3.8 1.0 1.0
assi (resid 16 and name H5)(resid 15 and name H5'') 6.3 1.5 1.5
assi (resid 9 and name H3')(resid 9 and name H5') 3.6 0.5 0.5
assi (resid 1 and name H5'')(resid 1 and name H2') 4.7 1.5 1.5
assi (resid 7 and name H5'')(resid 7 and name H2') 4.9 1.0 1.0
assi (resid 14 and name H5)(resid 14 and name H4') 6.6 1.5 1.5
assi (resid 8 and name H5')(resid 8 and name H3') 3.7 1.0 1.0
assi (resid 12 and name H1')(resid 12 and name H5) 5.2 1.5 1.5
assi (resid 8 and name H4')(resid 9 and name H4') 7.6 1.5 1.5
assi (resid 4 and name H6)(resid 3 and name H2') 2.2 1.0 1.0
assi (resid 15 and name H1')(resid 15 and name H5'') 5.1 1.0 1.0
assi (resid 10 and name H3')(resid 10 and name H2') 2.4 0.5 0.5
assi (resid 6 and name H1')(resid 6 and name H5) 5.2 1.5 1.5
assi (resid 9 and name H2')(resid 9 and name H5') 4.2 0.5 0.5
assi (resid 14 and name H5)(resid 14 and name H3') 5.2 1.5 1.5
assi (resid 4 and name H1')(resid 4 and name H5) 5.2 1.5 1.5
assi (resid 4 and name H1')(resid 4 and name H2') 2.8 1.0 1.0
assi (resid 14 and name H1')(resid 14 and name H5) 5.2 1.5 1.5
assi (resid 16 and name H2')(resid 16 and name H4') 3.8 1.0 1.0
assi (resid 8 and name H5')(resid 8 and name H5'') 1.7 1.0 1.0
assi (resid 7 and name H3')(resid 7 and name H2') 2.4 1.0 1.0
assi (resid 6 and name H5)(resid 6 and name H5'') 6.3 1.5 1.5
assi (resid 6 and name H3')(resid 6 and name H5'') 2.5 1.0 1.0
assi (resid 16 and name H5)(resid 16 and name H5'') 6.3 1.5 1.5
assi (resid 14 and name H5)(resid 14 and name H5'') 6.3 1.5 1.5
assi (resid 2 and name H8)(resid 1 and name H2') 2.0 1.0 1.0
assi (resid 12 and name H1')(resid 12 and name H5') 5.0 1.0 1.0
assi (resid 14 and name H2')(resid 14 and name H3') 2.4 1.0 1.0
assi (resid 12 and name H1')(resid 12 and name H5'') 5.0 1.0 1.0
assi (resid 16 and name H3')(resid 16 and name H5'') 2.5 1.0 1.0
assi (resid 6 and name H5')(resid 6 and name H2') 5.3 1.5 1.5
assi (resid 5 and name H8)(resid 4 and name H5') 7.5 1.5 1.5
assi (resid 1 and name H5'')(resid 1 and name H3') 2.5 1.0 1.0
assi (resid 6 and name H6)(resid 5 and name H5'') 5.8 1.5 1.5
assi (resid 7 and name H6)(resid 6 and name H3') 2.7 1.0 1.0

assi (resid 1 and name H5'')(resid 1 and name H5') 1.7 0.5 0.5
assi (resid 12 and name H2')(resid 12 and name H3') 2.4 0.5 0.5
assi (resid 15 and name H2')(resid 15 and name H3') 2.4 0.5 0.5
assi (resid 1 and name H5'')(resid 1 and name H1') 5.0 1.5 1.5
assi (resid 1 and name H5'')(resid 1 and name H8) 3.8 1.0 1.0
assi (resid 2 and name H8)(resid 2 and name H5') 4.2 1.0 1.0
assi (resid 1 and name H5'')(resid 1 and name H4') 2.6 1.0 1.0
assi (resid 12 and name H5')(resid 12 and name H5'') 1.7 0.5 0.5
assi (resid 2 and name H8)(resid 2 and name H1') 3.8 1.0 1.0
assi (resid 2 and name H8)(resid 1 and name H4') 5.8 1.5 1.5
assi (resid 2 and name H8)(resid 2 and name H5'') 6.3 1.5 1.5
assi (resid 9 and name H6)(resid 7 and name H5) 8.0 2.5 2.5
assi (resid 2 and name H8)(resid 1 and name H1') 5.0 1.5 1.5
assi (resid 7 and name H3)(resid 10 and name H1) 3.7 0.5 0.5
assi (resid 2 and name H8)(resid 2 and name H3') 3.0 1.5 1.5
assi (resid 16 and name H6)(resid 15 and name H5'') 5.7 1.5 1.5
assi (resid 5 and name H3')(resid 6 and name H2') 6.8 1.5 1.5
assi (resid 6 and name H4')(resid 5 and name H3') 5.8 1.5 1.5
assi (resid 16 and name H6)(resid 15 and name H2') 2.2 1.5 1.5
assi (resid 16 and name H6)(resid 15 and name H4') 5.8 1.5 1.5
assi (resid 12 and name H4')(resid 13 and name H4') 6.5 1.5 1.5
assi (resid 2 and name H8)(resid 2 and name H4') 4.3 1.0 1.0
assi (resid 7 and name H6)(resid 6 and name H5'') 5.0 1.0 1.0
assi (resid 12 and name H3')(resid 12 and name H5') 3.8 1.0 1.0
assi (resid 2 and name H8)(resid 1 and name H3') 3.0 1.5 1.5
assi (resid 12 and name H2')(resid 12 and name H4') 3.8 1.0 1.0
assi (resid 16 and name H6)(resid 15 and name H3') 3.3 1.5 1.5
assi (resid 5 and name H8)(resid 4 and name H3') 3.0 1.5 1.5
assi (resid 1 and name H8)(resid 2 and name H8) 5.0 1.5 1.5
assi (resid 3 and name H8)(resid 4 and name H6) 5.5 1.5 1.5
assi (resid 5 and name H8)(resid 4 and name H2') 2.0 1.5 1.5
assi (resid 11 and name H8)(resid 12 and name H6) 5.5 1.5 1.5
assi (resid 16 and name H6)(resid 15 and name H5') 7.0 1.5 1.5
assi (resid 12 and name H4')(resid 12 and name H5'') 2.6 1.0 1.0
assi (resid 6 and name H4')(resid 6 and name H5') 2.3 0.5 0.5
assi (resid 15 and name H6)(resid 14 and name H4') 5.8 1.5 1.5
assi (resid 10 and name H8)(resid 9 and name H5'') 4.9 1.5 1.5
assi (resid 7 and name H3')(resid 7 and name H5') 3.8 1.0 1.0
assi (resid 7 and name H3')(resid 8 and name H4') 7.0 1.5 1.5
assi (resid 15 and name H6)(resid 14 and name H3') 3.3 1.5 1.5
assi (resid 15 and name H6)(resid 14 and name H5') 7.0 1.5 1.5
assi (resid 10 and name H1)(resid 11 and name H8) 5.8 1.0 1.0
assi (resid 8 and name H6)(resid 9 and name H6) 5.8 1.5 1.5
assi (resid 8 and name H6)(resid 7 and name H2') 4.4 1.0 1.0
assi (resid 7 and name H6)(resid 6 and name H6) 4.8 1.0 1.0
assi (resid 5 and name H8)(resid 4 and name H5'') 6.3 1.5 1.5
assi (resid 5 and name H2')(resid 6 and name H5') 4.0 1.5 1.5
assi (resid 12 and name H3')(resid 12 and name H4') 3.0 1.0 1.0
assi (resid 5 and name H8)(resid 4 and name H4') 5.8 1.5 1.5
assi (resid 12 and name H4')(resid 12 and name H5') 2.3 1.0 1.0
assi (resid 7 and name H6)(resid 6 and name H2') 2.6 1.0 1.0

assi (resid 5 and name H2')(resid 5 and name H3') 2.4 1.0 1.0
 assi (resid 12 and name H3')(resid 12 and name H5'') 2.5 1.0 1.0
 assi (resid 11 and name H1)(resid 7 and name H3) 3.0 1.5 1.5
 assi (resid 8 and name H5)(resid 7 and name H5'') 4.8 1.0 1.0
 assi (resid 8 and name H5)(resid 7 and name H1') 5.7 1.0 1.0
 assi (resid 8 and name H5)(resid 7 and name H2') 5.6 1.0 1.0
 assi (resid 8 and name H5)(resid 7 and name H3') 5.4 1.0 1.0
 assi (resid 8 and name H5)(resid 7 and name H4') 3.3 1.0 1.0
 assi (resid 8 and name H2')(resid 7 and name H4') 5.4 1.0 1.0
 assi (resid 8 and name H2')(resid 7 and name H1') 5.4 1.0 1.0
 assi (resid 8 and name H3')(resid 8 and name H5'') 2.5 1.0 1.0
 assi (resid 8 and name H3')(resid 9 and name H5) 3.0 1.0 1.0
 assi (resid 1 and name H1')(resid 2 and name H8) 5.0 1.5 1.5
 assi (resid 2 and name H8)(resid 3 and name H8) 5.0 1.5 1.5
 assi (resid 3 and name H1')(resid 4 and name H6) 5.0 1.5 1.5
 assi (resid 4 and name H1')(resid 5 and name H8) 5.0 1.5 1.5
 assi (resid 12 and name H6)(resid 13 and name H8) 5.5 1.5 1.5
 assi (resid 13 and name H1')(resid 14 and name H6) 5.0 1.5 1.5
 assi (resid 14 and name H1')(resid 15 and name H6) 5.0 1.5 1.5
 assi (resid 14 and name H6)(resid 15 and name H6) 5.0 1.5 1.5
 assi (resid 15 and name H1')(resid 16 and name H6) 5.0 1.5 1.5
 assi (resid 15 and name H6)(resid 16 and name H6) 5.0 1.5 1.5

List of Hydrogen Bond Restraints

! for G1/ C16 base pair

assign (resid 1 and name N1) (resid 16 and name N3) 2.91 0.5 0.5
 assign (resid 1 and name O6) (resid 16 and name N4) 2.71 0.5 0.5
 assign (resid 1 and name N2) (resid 16 and name O2) 3.08 0.5 0.5
 assign (resid 1 and name H1) (resid 16 and name N3) 1.89 0.5 0.5
 assign (resid 1 and name O6) (resid 16 and name H42) 1.71 0.5 0.5
 assign (resid 1 and name H22) (resid 16 and name O2) 2.08 0.5 0.5

! for G2/ C15 base pair

assign (resid 2 and name N1) (resid 15 and name N3) 2.91 0.5 0.5
 assign (resid 2 and name O6) (resid 15 and name N4) 2.71 0.5 0.5
 assign (resid 2 and name N2) (resid 15 and name O2) 3.08 0.5 0.5
 assign (resid 2 and name H1) (resid 15 and name N3) 1.89 0.5 0.5
 assign (resid 2 and name O6) (resid 15 and name H42) 1.71 0.5 0.5
 assign (resid 2 and name H22) (resid 15 and name O2) 2.08 0.5 0.5

! for G3/ C14 base pair

assign (resid 3 and name N1) (resid 14 and name N3) 2.91 0.5 0.5
 assign (resid 3 and name O6) (resid 14 and name N4) 2.71 0.5 0.5
 assign (resid 3 and name N2) (resid 14 and name O2) 3.08 0.5 0.5
 assign (resid 3 and name H1) (resid 14 and name N3) 1.89 0.5 0.5
 assign (resid 3 and name O6) (resid 14 and name H42) 1.71 0.5 0.5
 assign (resid 3 and name H22) (resid 14 and name O2) 2.08 0.5 0.5

! for C4/ G13 base pair

assign (resid 13 and name N1) (resid 4 and name N3) 2.91 0.5 0.5

assign (resid 13 and name O6) (resid 4 and name N4) 2.71 0.5 0.5
assign (resid 13 and name N2) (resid 4 and name O2) 3.08 0.5 0.5
assign (resid 13 and name H1) (resid 4 and name N3) 1.89 0.5 0.5
assign (resid 13 and name O6) (resid 4 and name H42) 1.71 0.5 0.5
assign (resid 13 and name H22) (resid 4 and name O2) 2.08 0.5 0.5

! for G5/ C12 base pair

assign (resid 5 and name N1) (resid 12 and name N3) 2.91 0.5 0.5
assign (resid 5 and name O6) (resid 12 and name N4) 2.71 0.5 0.5
assign (resid 5 and name N2) (resid 12 and name O2) 3.08 0.5 0.5
assign (resid 5 and name H1) (resid 12 and name N3) 1.89 0.5 0.5
assign (resid 5 and name O6) (resid 12 and name H42) 1.71 0.5 0.5
assign (resid 5 and name H22) (resid 12 and name O2) 2.08 0.5 0.5

! for C6/ G11 base pair

assign (resid 11 and name N1) (resid 6 and name N3) 2.91 0.5 0.5
assign (resid 11 and name O6) (resid 6 and name N4) 2.71 0.5 0.5
assign (resid 11 and name N2) (resid 6 and name O2) 3.08 0.5 0.5
assign (resid 11 and name H1) (resid 6 and name N3) 1.89 0.5 0.5
assign (resid 11 and name O6) (resid 6 and name H42) 1.71 0.5 0.5
assign (resid 11 and name H22) (resid 6 and name O2) 2.08 0.5 0.5

! for U7/ G10 base pair

assign (resid 7 and name O2) (resid 10 and name N2) 3.08 0.5 0.5
assign (resid 7 and name O2) (resid 10 and name H22) 2.50 1.0 1.0

! for U8/ C9 base pair

assign (resid 8 and name O1P) (resid 9 and name H41) 2.50 1.0 1.0

! for U8/ G10 base pair

assign (resid 8 and name O2') (resid 10 and name O6) 2.50 1.0 1.0

UUCG 16mer Dihedral Angle Restraints

! G1 chi

assign (resid 1 and name O4') (resid 1 and name C1')
(resid 1 and name N9) (resid 1 and name C4)
1 -158 30 2

! G1 nu0

assign (resid 1 and name C4') (resid 1 and name O4')
(resid 1 and name C1') (resid 1 and name C2')
1 6 15 2

! G1 nu1

assign (resid 1 and name O4') (resid 1 and name C1')
(resid 1 and name C2') (resid 1 and name C3')
1 -25 15 2

! G1 nu2

assign (resid 1 and name C1') (resid 1 and name C2')
(resid 1 and name C3') (resid 1 and name C4')

1 37 15 2
 ! G1 nu3
 assign (resid 1 and name C2') (resid 1 and name C3')
 (resid 1 and name C4') (resid 1 and name O4')
 1 -33 15 2
 ! G1 nu4
 assign (resid 1 and name C3') (resid 1 and name C4')
 (resid 1 and name O4') (resid 1 and name C1')
 1 17 15 2
 ! G2 alpha
 assign (resid 1 and name O3') (resid 2 and name P)
 (resid 2 and name O5') (resid 2 and name C5')
 1 -65 15 2
 ! G2 beta
 assign (resid 2 and name P) (resid 2 and name O5')
 (resid 2 and name C5') (resid 2 and name C4')
 1 178 25 2
 ! G2 gamma
 assign (resid 2 and name O5') (resid 2 and name C5')
 (resid 2 and name C4') (resid 2 and name C3')
 1 54 15 2
 ! G2 epsilon
 assign (resid 2 and name C4') (resid 2 and name C3')
 (resid 2 and name O3') (resid 3 and name P)
 1 200 40 2
 ! G2 zeta
 assign (resid 2 and name C3') (resid 2 and name O3')
 (resid 3 and name P) (resid 3 and name O5')
 1 -71 15 2
 ! G2 chi
 assign (resid 2 and name O4') (resid 2 and name C1')
 (resid 2 and name N9) (resid 2 and name C4)
 1 -158 20 2
 ! G2 nu0
 assign (resid 2 and name C4') (resid 2 and name O4')
 (resid 2 and name C1') (resid 2 and name C2')
 1 6 15 2
 ! G2 nu1
 assign (resid 2 and name O4') (resid 2 and name C1')
 (resid 2 and name C2') (resid 2 and name C3')
 1 -25 15 2
 ! G2 nu2
 assign (resid 2 and name C1') (resid 2 and name C2')
 (resid 2 and name C3') (resid 2 and name C4')
 1 37 15 2
 ! G2 nu3
 assign (resid 2 and name C2') (resid 2 and name C3')
 (resid 2 and name C4') (resid 2 and name O4')
 1 -33 15 2
 ! G2 nu4
 assign (resid 2 and name C3') (resid 2 and name C4')

```

(resid 2 and name O4') (resid 2 and name C1')
  1 17 15 2
! G3 alpha
assign (resid 2 and name O3') (resid 3 and name P )
  (resid 3 and name O5') (resid 3 and name C5')
  1 -68 15 2
! G3 beta
assign (resid 3 and name P ) (resid 3 and name O5')
  (resid 3 and name C5') (resid 3 and name C4')
  1 178 25 2
! G3 gamma
assign (resid 3 and name O5') (resid 3 and name C5')
  (resid 3 and name C4') (resid 3 and name C3')
  1 54 15 2
! G3 epsilon
assign (resid 3 and name C4') (resid 3 and name C3')
  (resid 3 and name O3') (resid 4 and name P )
  1 200 40 2
! G3 zeta
assign (resid 3 and name C3') (resid 3 and name O3')
  (resid 4 and name P ) (resid 4 and name O5')
  1 -71 15 2
! G3 chi
assign (resid 3 and name O4') (resid 3 and name C1')
  (resid 3 and name N9 ) (resid 3 and name C4 )
  1 -158 20 2
! G3 nu0
assign (resid 3 and name C4') (resid 3 and name O4')
  (resid 3 and name C1') (resid 3 and name C2')
  1 6 15 2
! G3 nu1
assign (resid 3 and name O4') (resid 3 and name C1')
  (resid 3 and name C2') (resid 3 and name C3')
  1 -25 15 2
! G3 nu2
assign (resid 3 and name C1') (resid 3 and name C2')
  (resid 3 and name C3') (resid 3 and name C4')
  1 37 15 2
! G3 nu3
assign (resid 3 and name C2') (resid 3 and name C3')
  (resid 3 and name C4') (resid 3 and name O4')
  1 -33 15 2
! G3 nu4
assign (resid 3 and name C3') (resid 3 and name C4')
  (resid 3 and name O4') (resid 3 and name C1')
  1 17 15 2
! C4 alpha
assign (resid 3 and name O3') (resid 4 and name P )
  (resid 4 and name O5') (resid 4 and name C5')
  1 -65 15 2
! C4 beta

```

```

assign (resid 4 and name P ) (resid 4 and name O5')
      (resid 4 and name C5') (resid 4 and name C4')
      1 160 25 2
! C4 gamma
assign (resid 4 and name O5') (resid 4 and name C5')
      (resid 4 and name C4') (resid 4 and name C3')
      1 70 15 2
! C4 epsilon
assign (resid 4 and name C4') (resid 4 and name C3')
      (resid 4 and name O3') (resid 5 and name P )
      1 200 30 2
! C4 zeta
assign (resid 4 and name C3') (resid 4 and name O3')
      (resid 5 and name P ) (resid 5 and name O5')
      1 -85 15 2
! C4 chi
assign (resid 4 and name O4') (resid 4 and name C1')
      (resid 4 and name N1 ) (resid 4 and name C2 )
      1 -158 20 2
! C4 nu0
assign (resid 4 and name C4') (resid 4 and name O4')
      (resid 4 and name C1') (resid 4 and name C2')
      1 6 15 2
! C4 nu1
assign (resid 4 and name O4') (resid 4 and name C1')
      (resid 4 and name C2') (resid 4 and name C3')
      1 -25 15 2
! C4 nu2
assign (resid 4 and name C1') (resid 4 and name C2')
      (resid 4 and name C3') (resid 4 and name C4')
      1 37 15 2
! C4 nu3
assign (resid 4 and name C2') (resid 4 and name C3')
      (resid 4 and name C4') (resid 4 and name O4')
      1 -33 15 2
! C4 nu4
assign (resid 4 and name C3') (resid 4 and name C4')
      (resid 4 and name O4') (resid 4 and name C1')
      1 17 15 2
! G5 alpha
assign (resid 4 and name O3') (resid 5 and name P )
      (resid 5 and name O5') (resid 5 and name C5')
      1 -50 15 2
! G5 beta
assign (resid 5 and name P ) (resid 5 and name O5')
      (resid 5 and name C5') (resid 5 and name C4')
      1 165 15 2
! G5 gamma
assign (resid 5 and name O5') (resid 5 and name C5')
      (resid 5 and name C4') (resid 5 and name C3')
      1 54 15 2

```

! G5 epsilon
assign (resid 5 and name C4') (resid 5 and name C3')
(resid 5 and name O3') (resid 6 and name P)
1 220 30 2

! G5 zeta
assign (resid 5 and name C3') (resid 5 and name O3')
(resid 6 and name P) (resid 6 and name O5')
1 -80 15 2

! G5 chi
assign (resid 5 and name O4') (resid 5 and name C1')
(resid 5 and name N9) (resid 5 and name C4)
1 -158 20 2

! G5 nu0
assign (resid 5 and name C4') (resid 5 and name O4')
(resid 5 and name C1') (resid 5 and name C2')
1 6 15 2

! G5 nu1
assign (resid 5 and name O4') (resid 5 and name C1')
(resid 5 and name C2') (resid 5 and name C3')
1 -25 15 2

! G5 nu2
assign (resid 5 and name C1') (resid 5 and name C2')
(resid 5 and name C3') (resid 5 and name C4')
1 37 15 2

! G5 nu3
assign (resid 5 and name C2') (resid 5 and name C3')
(resid 5 and name C4') (resid 5 and name O4')
1 -33 15 2

! G5 nu4
assign (resid 5 and name C3') (resid 5 and name C4')
(resid 5 and name O4') (resid 5 and name C1')
1 17 15 2

! C6 alpha
assign (resid 5 and name O3') (resid 6 and name P)
(resid 6 and name O5') (resid 6 and name C5')
1 -40 30 2

! C6 beta
assign (resid 6 and name P) (resid 6 and name O5')
(resid 6 and name C5') (resid 6 and name C4')
1 160 20 2

! C6 gamma
assign (resid 6 and name O5') (resid 6 and name C5')
(resid 6 and name C4') (resid 6 and name C3')
1 54 20 2

! C6 epsilon
assign (resid 6 and name C4') (resid 6 and name C3')
(resid 6 and name O3') (resid 7 and name P)
1 220 40 2

! C6 zeta
assign (resid 6 and name C3') (resid 6 and name O3')
(resid 7 and name P) (resid 7 and name O5')

1 -80 20 2
 ! C6 chi
 assign (resid 6 and name O4') (resid 6 and name C1')
 (resid 6 and name N1) (resid 6 and name C2)
 1 -158 20 2
 ! C6 nu0
 assign (resid 6 and name C4') (resid 6 and name O4')
 (resid 6 and name C1') (resid 6 and name C2')
 1 6 15 2
 ! C6 nu1
 assign (resid 6 and name O4') (resid 6 and name C1')
 (resid 6 and name C2') (resid 6 and name C3')
 1 -25 15 2
 ! C6 nu2
 assign (resid 6 and name C1') (resid 6 and name C2')
 (resid 6 and name C3') (resid 6 and name C4')
 1 37 15 2
 ! C6 nu3
 assign (resid 6 and name C2') (resid 6 and name C3')
 (resid 6 and name C4') (resid 6 and name O4')
 1 -33 15 2
 ! C6 nu4
 assign (resid 6 and name C3') (resid 6 and name C4')
 (resid 6 and name O4') (resid 6 and name C1')
 1 17 15 2
 ! U7 alpha
 assign (resid 6 and name O3') (resid 7 and name P)
 (resid 7 and name O5') (resid 7 and name C5')
 1 -68 30 2
 ! U7 beta
 assign (resid 7 and name P) (resid 7 and name O5')
 (resid 7 and name C5') (resid 7 and name C4')
 1 178 30 2
 ! U7 gamma
 assign (resid 7 and name O5') (resid 7 and name C5')
 (resid 7 and name C4') (resid 7 and name C3')
 1 54 25 2
 ! U7 epsilon
 assign (resid 7 and name C4') (resid 7 and name C3')
 (resid 7 and name O3') (resid 8 and name P)
 1 215 50 2
 ! U7 zeta
 assign (resid 7 and name C3') (resid 7 and name O3')
 (resid 8 and name P) (resid 8 and name O5')
 1 -71 30 2
 ! U7 chi
 assign (resid 7 and name O4') (resid 7 and name C1')
 (resid 7 and name N1) (resid 7 and name C2)
 1 -158 20 2
 ! U7 nu0
 assign (resid 7 and name C4') (resid 7 and name O4')

```

(resid 7 and name C1') (resid 7 and name C2')
1 6 15 2
! U7 nu1
assign (resid 7 and name O4') (resid 7 and name C1')
(resid 7 and name C2') (resid 7 and name C3')
1 -25 15 2
! U7 nu2
assign (resid 7 and name C1') (resid 7 and name C2')
(resid 7 and name C3') (resid 7 and name C4')
1 37 15 2
! U7 nu3
assign (resid 7 and name C2') (resid 7 and name C3')
(resid 7 and name C4') (resid 7 and name O4')
1 -33 15 2
! U7 nu4
assign (resid 7 and name C3') (resid 7 and name C4')
(resid 7 and name O4') (resid 7 and name C1')
1 17 15 2
! U8 chi
assign (resid 8 and name O4') (resid 8 and name C1')
(resid 8 and name N1 ) (resid 8 and name C2 )
1 -150 40 2
! U8 nu1
assign (resid 8 and name O4') (resid 8 and name C1')
(resid 8 and name C2') (resid 8 and name C3')
1 25 15 2
! U8 nu2
assign (resid 8 and name C1') (resid 8 and name C2')
(resid 8 and name C3') (resid 8 and name C4')
1 -25 15 2
! C9 chi
assign (resid 9 and name O4') (resid 9 and name C1')
(resid 9 and name N1 ) (resid 9 and name C2 )
1 -140 50 2
! C9 nu1
assign (resid 9 and name O4') (resid 9 and name C1')
(resid 9 and name C2') (resid 9 and name C3')
1 30 25 2
! C9 nu2
assign (resid 9 and name C1') (resid 9 and name C2')
(resid 9 and name C3') (resid 9 and name C4')
1 -35 25 2
! G11 alpha
assign (resid 10 and name O3') (resid 11 and name P )
(resid 11 and name O5') (resid 11 and name C5')
1 -135 50 2
! G11 beta
assign (resid 11 and name P ) (resid 11 and name O5')
(resid 11 and name C5') (resid 11 and name C4')
1 165 30 2
! G11 gamma

```



```

assign (resid 11 and name O5') (resid 11 and name C5')
  (resid 11 and name C4') (resid 11 and name C3')
  1 110 50 2
! G11 epsilon
assign (resid 11 and name C4') (resid 11 and name C3')
  (resid 11 and name O3') (resid 12 and name P )
  1 200 30 2
! G11 zeta
assign (resid 11 and name C3') (resid 11 and name O3')
  (resid 12 and name P ) (resid 12 and name O5')
  1 -100 30 2
! G11 chi
assign (resid 11 and name O4') (resid 11 and name C1')
  (resid 11 and name N9 ) (resid 11 and name C4 )
  1 -150 20 2
! G11 nu0
assign (resid 11 and name C4') (resid 11 and name O4')
  (resid 11 and name C1') (resid 11 and name C2')
  1 6 15 2
! G11 nu1
assign (resid 11 and name O4') (resid 11 and name C1')
  (resid 11 and name C2') (resid 11 and name C3')
  1 -25 15 2
! G11 nu2
assign (resid 11 and name C1') (resid 11 and name C2')
  (resid 11 and name C3') (resid 11 and name C4')
  1 37 15 2
! G11 nu3
assign (resid 11 and name C2') (resid 11 and name C3')
  (resid 11 and name C4') (resid 11 and name O4')
  1 -33 15 2
! G11 nu4
assign (resid 11 and name C3') (resid 11 and name C4')
  (resid 11 and name O4') (resid 11 and name C1')
  1 17 15 2
! C12 alpha
assign (resid 11 and name O3') (resid 12 and name P )
  (resid 12 and name O5') (resid 12 and name C5')
  1 -65 25 2
! C12 beta
assign (resid 12 and name P ) (resid 12 and name O5')
  (resid 12 and name C5') (resid 12 and name C4')
  1 178 25 2
! C12 gamma
assign (resid 12 and name O5') (resid 12 and name C5')
  (resid 12 and name C4') (resid 12 and name C3')
  1 54 25 2
! C12 epsilon
assign (resid 12 and name C4') (resid 12 and name C3')
  (resid 12 and name O3') (resid 13 and name P )
  1 200 40 2

```

```

! C12 zeta
assign (resid 12 and name C3') (resid 12 and name O3')
      (resid 13 and name P ) (resid 13 and name O5')
      1 -71 25 2
! C12 chi
assign (resid 12 and name O4') (resid 12 and name C1')
      (resid 12 and name N1 ) (resid 12 and name C2 )
      1 -158 25 2
! C12 nu0
assign (resid 12 and name C4') (resid 12 and name O4')
      (resid 12 and name C1') (resid 12 and name C2')
      1 6 15 2
! C12 nu1
assign (resid 12 and name O4') (resid 12 and name C1')
      (resid 12 and name C2') (resid 12 and name C3')
      1 -25 15 2
! C12 nu2
assign (resid 12 and name C1') (resid 12 and name C2')
      (resid 12 and name C3') (resid 12 and name C4')
      1 37 15 2
! C12 nu3
assign (resid 12 and name C2') (resid 12 and name C3')
      (resid 12 and name C4') (resid 12 and name O4')
      1 -33 15 2
! C12 nu4
assign (resid 12 and name C3') (resid 12 and name C4')
      (resid 12 and name O4') (resid 12 and name C1')
      1 17 15 2
! G13 alpha
assign (resid 12 and name O3') (resid 13 and name P )
      (resid 13 and name O5') (resid 13 and name C5')
      1 -65 20 2
! G13 beta
assign (resid 13 and name P ) (resid 13 and name O5')
      (resid 13 and name C5') (resid 13 and name C4')
      1 178 20 2
! G13 gamma
assign (resid 13 and name O5') (resid 13 and name C5')
      (resid 13 and name C4') (resid 13 and name C3')
      1 54 20 2
! G13 epsilon
assign (resid 13 and name C4') (resid 13 and name C3')
      (resid 13 and name O3') (resid 14 and name P )
      1 200 40 2
! G13 zeta
assign (resid 13 and name C3') (resid 13 and name O3')
      (resid 14 and name P ) (resid 14 and name O5')
      1 -71 20 2
! G13 chi
assign (resid 13 and name O4') (resid 13 and name C1')
      (resid 13 and name N9 ) (resid 13 and name C4 )

```

```

1 -158 20 2
! G13 nu0
assign (resid 13 and name C4') (resid 13 and name O4')
      (resid 13 and name C1') (resid 13 and name C2')
      1 6 15 2
! G13 nu1
assign (resid 13 and name O4') (resid 13 and name C1')
      (resid 13 and name C2') (resid 13 and name C3')
      1 -25 15 2
! G13 nu2
assign (resid 13 and name C1') (resid 13 and name C2')
      (resid 13 and name C3') (resid 13 and name C4')
      1 37 15 2
! G13 nu3
assign (resid 13 and name C2') (resid 13 and name C3')
      (resid 13 and name C4') (resid 13 and name O4')
      1 -33 15 2
! G13 nu4
assign (resid 13 and name C3') (resid 13 and name C4')
      (resid 13 and name O4') (resid 13 and name C1')
      1 17 15 2
! C14 alpha
assign (resid 13 and name O3') (resid 14 and name P )
      (resid 14 and name O5') (resid 14 and name C5')
      1 -65 15 2
! C14 beta
assign (resid 14 and name P ) (resid 14 and name O5')
      (resid 14 and name C5') (resid 14 and name C4')
      1 178 25 2
! C14 gamma
assign (resid 14 and name O5') (resid 14 and name C5')
      (resid 14 and name C4') (resid 14 and name C3')
      1 54 15 2
! C14 epsilon
assign (resid 14 and name C4') (resid 14 and name C3')
      (resid 14 and name O3') (resid 15 and name P )
      1 200 40 2
! C14 zeta
assign (resid 14 and name C3') (resid 14 and name O3')
      (resid 15 and name P ) (resid 15 and name O5')
      1 -71 15 2
! C14 chi
assign (resid 14 and name O4') (resid 14 and name C1')
      (resid 14 and name N1 ) (resid 14 and name C2 )
      1 -158 20 2
! C14 nu0
assign (resid 14 and name C4') (resid 14 and name O4')
      (resid 14 and name C1') (resid 14 and name C2')
      1 6 15 2
! C14 nu1
assign (resid 14 and name O4') (resid 14 and name C1')

```

```

      (resid 14 and name C2') (resid 14 and name C3')
      1 -25 15 2
! C14 nu2
assign (resid 14 and name C1') (resid 14 and name C2')
      (resid 14 and name C3') (resid 14 and name C4')
      1 37 15 2
! C14 nu3
assign (resid 14 and name C2') (resid 14 and name C3')
      (resid 14 and name C4') (resid 14 and name O4')
      1 -33 15 2
! C14 nu4
assign (resid 14 and name C3') (resid 14 and name C4')
      (resid 14 and name O4') (resid 14 and name C1')
      1 17 15 2
! C15 alpha
assign (resid 14 and name O3') (resid 15 and name P )
      (resid 15 and name O5') (resid 15 and name C5')
      1 -65 15 2
! C15 beta
assign (resid 15 and name P ) (resid 15 and name O5')
      (resid 15 and name C5') (resid 15 and name C4')
      1 178 15 2
! C15 gamma
assign (resid 15 and name O5') (resid 15 and name C5')
      (resid 15 and name C4') (resid 15 and name C3')
      1 54 15 2
! C15 epsilon
assign (resid 15 and name C4') (resid 15 and name C3')
      (resid 15 and name O3') (resid 16 and name P )
      1 200 40 2
! C15 zeta
assign (resid 15 and name C3') (resid 15 and name O3')
      (resid 16 and name P ) (resid 16 and name O5')
      1 -71 15 2
! C15 chi
assign (resid 15 and name O4') (resid 15 and name C1')
      (resid 15 and name N1 ) (resid 15 and name C2 )
      1 -158 20 2
! C15 nu0
assign (resid 15 and name C4') (resid 15 and name O4')
      (resid 15 and name C1') (resid 15 and name C2')
      1 6 15 2
! C15 nu1
assign (resid 15 and name O4') (resid 15 and name C1')
      (resid 15 and name C2') (resid 15 and name C3')
      1 -25 15 2
! C15 nu2
assign (resid 15 and name C1') (resid 15 and name C2')
      (resid 15 and name C3') (resid 15 and name C4')
      1 37 15 2
! C15 nu3

```

```

assign (resid 15 and name C2') (resid 15 and name C3')
      (resid 15 and name C4') (resid 15 and name O4')
      1 -33 15 2
! C15 nu4
assign (resid 15 and name C3') (resid 15 and name C4')
      (resid 15 and name O4') (resid 15 and name C1')
      1 17 15 2
! C16 chi
assign (resid 16 and name O4') (resid 16 and name C1')
      (resid 16 and name N1 ) (resid 16 and name C2 )
      1 -158 30 2
! C16 nu0
assign (resid 16 and name C4') (resid 16 and name O4')
      (resid 16 and name C1') (resid 16 and name C2')
      1 6 15 2
! C16 nu1
assign (resid 16 and name O4') (resid 16 and name C1')
      (resid 16 and name C2') (resid 16 and name C3')
      1 -25 15 2
! C16 nu2
assign (resid 16 and name C1') (resid 16 and name C2')
      (resid 16 and name C3') (resid 16 and name C4')
      1 37 15 2
! C16 nu3
assign (resid 16 and name C2') (resid 16 and name C3')
      (resid 16 and name C4') (resid 16 and name O4')
      1 -33 15 2
! C16 nu4
assign (resid 16 and name C3') (resid 16 and name C4')
      (resid 16 and name O4') (resid 16 and name C1')
      1 17 15 2

```

UUCG 16mer Planarity Restraints

```

! G1-C16 WC
!-----
group
selection= ((resid 1 and name N1) or (resid 1 and name N3) or
            (resid 1 and name C5) or (resid 16 and name N1) or
            (resid 16 and name N3) or (resid 16 and name C5))
weight = $pscale end

! G2-C15 WC
!-----
group
selection= ((resid 2 and name N1) or (resid 2 and name N3) or
            (resid 2 and name C5) or (resid 15 and name N1) or
            (resid 15 and name N3) or (resid 15 and name C5))
weight = $pscale end

```

```
! G3-C14 WC
!-----
group
selection= ((resid 3 and name N1) or (resid 3 and name N3) or
            (resid 3 and name C5) or (resid 14 and name N1) or
            (resid 14 and name N3) or (resid 14 and name C5))
weight = $pscale end
```

```
! C4-G13 WC
!-----
group
selection= ((resid 4 and name N1) or (resid 4 and name N3) or
            (resid 4 and name C5) or (resid 13 and name N1) or
            (resid 13 and name N3) or (resid 13 and name C5))
weight = $pscale end
```

```
! G5-C12 WC
!-----
group
selection= ((resid 5 and name N1) or (resid 5 and name N3) or
            (resid 5 and name C5) or (resid 12 and name N1) or
            (resid 12 and name N3) or (resid 12 and name C5))
weight = $pscale end
```

```
! C6-G11 WC
!-----
group
selection= ((resid 6 and name N1) or (resid 6 and name N3) or
            (resid 6 and name C5) or (resid 11 and name N1) or
            (resid 11 and name N3) or (resid 11 and name C5))
weight = $pscale end
```