EFFECTS OF SMALL MOLECULE BINDING TO mRNA ON GENE EXPRESSION

Julie Marie Sullivan

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Chemistry.

Chapel Hill
2008

Approved by:
H. Holden Thorp
Linda L. Spremulli
Kevin M. Weeks
Thomas J. Meyer
Marcy L. Waters
ABSTRACT

Julie Marie Sullivan: Effects of Small Molecule Binding to mRNA on Gene Expression
(Under the direction of H. Holden Thorp)

RNA plays significant roles in many cellular processes and exhibits attractive features for use as a small molecule targeted therapeutic. This study focuses on using RNA as a drug target, specifically investigating the process by which small molecules can regulate gene expression through their interaction with secondary structural elements found in the 5’ untranslated region (UTR) of two cellular messenger RNA (mRNA) targets, human ferritin and human preproinsulin.

Previous work has identified compounds that specifically and selectively bind in the hairpin loop of the iron responsive element (IRE) in human ferritin mRNA using a chemical footprinting method. Ferritin is regulated at the translational level by IRP binding to the IRE, preventing assembly of the translational initiation complex. One small molecule, promazine, was identified that binds to an internal bulge in the IRE sequence where an iron regulatory protein (IRP) binds. The binding site identified during the initial screening was confirmed through an enzymatic footprinting assay. Using an electrophoretic mobility shift assay, we have shown that this compound can disrupt the binding of IRP proteins to the IRE sequence. In an in vitro translation assay, binding of promazine leads to up-regulation of ferritin protein synthesis by greater than 60% while showing no effect on control mRNA.
The 5’UTR of the human preproinsulin mRNA exists in two forms, the native form, and a splice variant which is up-regulated in response to chronic glucose stimulation. Small molecules were screened for their ability to bind to, or change the secondary structure of, the 5’UTR of the preproinsulin mRNAs using RNase cleavage, and the transition metal cleavage agents Ru(tpy)(bpy)O2+ (RuO2+, bpy = 2,2’-bipyridine, tpy = 2,2’,2″-terpyridine), and Ru(bpy)3 2+. Two small molecules, neomycin B and kanamycin B, have been identified as binding to the mRNAs using these screening methods. Binding of neomycin B, and kanamycin B decreased the amount of preproinsulin protein synthesized for both the native and splice variant forms in an *in vitro* translation assay. A decrease in the translation of a control RNA was also seen, indicating that the binding of these aminoglycosides are not specific to secondary structures found in preproinsulin mRNA.
ACKNOWLEDGEMENTS

Five years ago I moved from Buffalo to Chapel Hill. At the time I was unsure if I would adjust, if grad school was really the place I wanted to be. It was hard to move from the only place I had lived, away from all my family and friends. Five years later, I know that it was one of the best decisions I’ve ever made. I’ve learned more about myself and about chemistry that I’d ever known before. Learning chemistry could happen anywhere, but I’m happy that it happened here.

I came to the chemistry department during the summer to work in Holden’s lab, and have been there ever since. There I met Patty, Jen, Dom and Jared, who helped me with my experiments, my orals, and just finding my way in the department and in Chapel Hill. Patty was a surrogate mom, worried when you were sick and giving advice where needed, an amazing teacher, and without her I would have never been able to write this dissertation, or have the confidence to tackle new areas of science. Jen taught me techniques that I still use and was a great help when figuring out what to do next, and Dom and Jared helped make lab a fun place to come every morning, and are people I can still ask questions. After my first year here, Dana, Jenn and Chrissie joined the lab and have been there for all the transitions. Dana is my benchmate and fellow RNA chemist. She bore the brunt of my procrastination while writing, and was always there for the random questions I had or to tell me that my figures looked okay. Jenn and Chrissie always had words of encouragement when I was
panicking, and constantly supported and took interest in what I was doing. Without these three people, I wouldn’t have made it through the last 4 years.

Working for and with Holden was an experience that I know I couldn’t have gotten anywhere else. He let me gain independence, always listened to my interpretation of data, supported me through anything I wanted to do or try, and understood that the amount of time spent in lab wasn’t as important as getting results and overall being happy with what you were doing. He encouraged me to do things outside of the lab, to travel to Europe, to become involved with the entrepreneurship class he was teaching, to try new things. Who else would call you in the middle of the day to go be part of a “focus group” for the planetarium to make sure their Pink Floyd laser show was appropriate, or invite you to see him perform in a musical, or his band play at Weaver Street Market? It helped make him less of a “boss” and more someone who really wanted to see you succeed, which of course was the ultimate goal.

It takes much more than getting experimental results to make it through graduate school, and it is something that I could not have done by myself. At UNC I’ve made great friends, and without them these past five years would have been incredibly boring. I’ve been lucky to have great housemates who have made day to day issues easier. The past two years, Abby and Stephanie, have been great companions, fellow TV watchers, and provided someone to talk to after long hours in the lab. They’ve been there to provide advice, their knowledge of inorganic chemistry, and were always up for doing something. Abby has always been there to help me when I’ve needed it, and I appreciate her friendship more than she knows. I’ve also been surrounded by friends that I met my first few years, Abby, Anne, Dee, Dana, Nick, Jason, Matt, Becca and Heather, who have been a constant source of support, and overall a great group of people to take trips to the beach, go to dinner, or hang
out with. Even though Heather moved to Boston, we kept in touch through constant email conversations, and trips, and she is my number one cheerleader. I also have been fortunate to have met a great group of girls who have made girls nights that much more fun.

Through random chance I met Pete while working in the Schoenfisch lab on SPR experiments. While the actual experiments didn’t yield any good results, I met someone who has become a more than a great friend, who has provided endless support, and is the person I look forward to spending my future with.

Along with my friends at UNC, I’ve been lucky to have support from friends from Buffalo, specifically Heather and Jenn. Heather took the first 12 hour drive down here, through the scary mountains of West Virginia, to help me move. She has always been someone that I could count on, that I can call up in any situation and she’ll be there. Jenn is my travel companion, and she encourages me to do things I otherwise wouldn’t think to do. Additionally, I have made friends along the way, who have been there to add a little more happiness into my life. Rob I met at UB and we have kept in touch while finishing our respective programs. He has been helpful with all aspects of laboratory life, and has helped most recently with my search for what I want to do after graduate school. He was also always up for getting outdoors and was a much needed visitor during my early time here. I met Sid at U of R as a freshman. We became good friends, and I enjoy the funny anecdotes and stories that he provides, and am thankful for all the times he has been there for me.

I’ve been very blessed to have the family that I have. They have been there to support me through whatever decisions I make, are happy for me when I succeed, and pick me back up when I fail. Although they may not understand exactly what I do, they have always encouraged me, and been there when I needed them. Vicki is my sounding board,
and never fails to remind me of who I am. My younger brothers and sisters provide entertainment when I visit, and I am amazed by what and who they are becoming each time I return. Mary always reminds me how far I have come, and my Dad is always there to keep me on track. My Uncles, Gary and Rob, are the ones who I remember introducing me to science, and have always shown me that science is exciting, even at its most basic level. There are no words to describe my mother, how much she means to me, how I think of her everyday, and how I hope to one day be as wonderful a person as she was.

I am sure that there are people that I have missed but they are no less important than any of the others listed here. Each person I know has impacted my time at UNC and I am extremely grateful for all that they have done.
TABLE OF CONTENTS

LIST OF TABLES ................................................................................................................................. xiii
LIST OF FIGURES ............................................................................................................................... xiv
LIST OF ABBREVIATIONS AND SYMBOLS ......................................................................................... xviii

Chapter 1. Introduction: Targeting RNA with Small Molecules .......................................................... 1

1.1 RNA as a Drug Target .................................................................................................................... 1
1.2 RNA Structure ............................................................................................................................... 2
1.3 Chemical and Enzymatic Mapping Techniques for RNA ................................................................. 5

1.3.1 Enzymatic Mapping Techniques ............................................................................................... 5

1.3.2 Chemical Mapping Techniques ............................................................................................... 6

1.3.2.1 Nucleic Acid Reactivity to Oxidation ................................................................................... 6

1.3.2.2 Metallonucleases as RNA Mapping Agents .......................................................................... 8

1.3.2.3 Ru(tpy)(bpy)O^{2+} as an RNA Chemical Mapping and Screen for Small Molecule Binding ......... 9

1.3.2.4 Ru(bpy)_{3}^{2+} as a Chemical Mapping Agent: The Flash-Quench technique ............................. 11

1.4 Molecular RNA Targets ................................................................................................................ 12

1.4.1 Antiviral Targets: Transactivation Response RNA (TAR RNA) ............................................ 12

1.4.2 mRNA Targets: Inhibiting Ribosome Screening ....................................................................... 14

1.5 Principles of RNA Binding ............................................................................................................ 15

1.6 Dissertation Focus .......................................................................................................................... 17
Chapter 2. In Vitro Up-regulation of Ferritin Protein Synthesis by Small Molecules that Target the Iron Responsive Element in Ferritin mRNA

2.1 Abstract

2.2 Introduction

2.2.1 Cellular Iron Homeostasis

2.2.1.1 Iron Absorption and Transport

2.2.1.2 Cellular Iron Uptake: Transferrin-mediated Endocytosis

2.2.1.3 Cellular Iron Storage: Ferritin

2.2.2 Regulation of Cellular Iron Homeostasis: Regulation of Transferrin and Ferritin by Iron Regulatory Proteins (IRPs) and Iron Responsive Elements (IREs)

2.2.3 Identification of Small Molecules that Disrupt the IRP:IRE Interaction

2.3 Experimental

2.3.1 Materials

2.3.2 Methods

2.3.2.1 Transcription of IRE RNA (50-mer)

2.3.2.2 5’ RNA End-labeling

2.3.2.3 3’ RNA End-labeling

2.3.2.4 Internal Labeling of RNA with $[^{32}P]UTP$

2.3.2.5 Expression, Lysis, and Purification of IRP1 and IRP2

2.3.2.6 Electrophoretic Mobility Shift Assay

2.3.2.7 RNase Cleavage Assay

2.3.2.8 Transcription of Full-Length mRNA: Human
Chapter 2. Identification of Small Molecules that Target RNA and Mediate Human Preproinsulin Protein Synthesis

2.3.2.9 In Vitro Translation of mRNA: Rabbit Reticulocyte Lysate Translation Assay

2.4 Results and Discussion

2.4.1 Small Molecule:Ferritin mRNA Interaction: RNase Cleavage Assay

2.4.2 Effect of Promazine on the IRP:IRE Interaction: Electrophoretic Mobility Shift Assay

2.4.3 Regulation of Protein Synthesis by Small Molecules: In Vitro Translation Assay

2.4.4 Binding Constant Determination

2.5 Conclusions

2.6 Tables

2.7 Figures

2.8 References

Chapter 3. Identification of Small Molecules that Target RNA and Mediate Human Preproinsulin Protein Synthesis

3.1 Abstract

3.2 Introduction

3.2.1 Regulation of Glucose Homeostasis

3.2.2 Insulin

3.2.2.1 Human Insulin Biosynthesis

3.2.2.2 Insulin Secretion from Pancreatic β-cells
3.2.2.3 Mechanism of Insulin Action in the Body ........................................90

3.2.3 Control of Insulin Gene Expression by Glucose ..............................91

3.2.4 Identification of Small Molecules that Bind to Human Preproinsulin mRNA .................................................................93

3.3 Experimental .........................................................................................94

3.3.1 Materials .............................................................................................95

3.3.2 Methods ................................................................................................95

3.3.2.1 Cloning of the Full-length Human Preproinsulin Splice Variant (SPV) .................................................................97

3.3.2.2 Cloning of the 5’UTR of Human Preproinsulin Splice Variant (SPV Hairpin) ..........................................................99

3.3.2.3 Creation of the Full-length and 5’UTR Native Human Preproinsulin Constructs .........................................................100

3.3.2.4 Transcription of NAT and SPV Hairpin RNA ............................100

3.3.2.5 RNA End-labeling ......................................................................101

3.3.2.6 RNase Cleavage Assay ...............................................................102

3.3.2.7 RuO²⁺ Small Molecule Binding Assay .........................................103

3.3.2.8 Flash-Quench Oxidation ..............................................................104

3.3.2.9 Transcription of Full-length mRNA: Human NAT Preproinsulin and Human SPV Preproinsulin ..........................104

3.3.2.10 In Vitro Translation of mRNA ....................................................104

3.4 Results and Discussion ........................................................................105

3.4.1 Secondary Structure Analysis of $^{32}$P-End-labeled Human Preproinsulin Hairpin RNA ........................................................105

3.4.1.1 Native Preproinsulin mRNA Structure ....................................105

3.4.1.2 Preproinsulin mRNA Splice Variant Structure ........................107
3.4.2 Screening Compounds to Find a Human Preproinsulin Hairpin RNA Binder .................................................................108

3.4.3 Comparison of Methods Used to Identify a Human Preproinsulin Hairpin RNA Binder ..........................................................110

3.4.4 Regulation of Protein Synthesis by Small Molecules: In Vitro Translation Assay ..............................................................111

3.4.4.1 Human Preproinsulin Native and Splice Variant mRNA Constructs ..................................................................................111

3.4.4.2 Wild-type Luciferase mRNA Construct .........................................................111

3.5 Conclusions .................................................................................................112

3.6 Future Directions ........................................................................................114

3.7 Tables ..........................................................................................................115

3.8 Figures ..........................................................................................................118

3.9 References ....................................................................................................155
LIST OF TABLES

Table 2.1  Names and structures of compounds found to interact with the 
IRE through the RuO$_2^{2+}$ assay.................................................................64

Table 3.1  Oligonucleotides used in human preproinsulin cloning .......................115

Table 3.2  PCR conditions used in human preproinsulin cloning............................116

Table 3.3  Names and basic structures of compounds tested in the 
preproinsulin binding assay ........................................................................117
LIST OF FIGURES

Figure 1.1  Schematic representation of four general classes of RNA secondary structure.........................................................................................................................18

Figure 1.2  Structures of the two metal complexes used as metallonucleases in this work ..................................................................................................................19

Figure 1.3  Proposed mechanism of guanine oxidation in DNA........................................20

Figure 1.4  Common aromatic ligands used to change the binding properties of nucleases..................................................................................................................21

Figure 1.5  Schematic representations of the secondary structure of RNAs mapped with metallonucleases .........................................................................................22

Figure 1.6  Proposed secondary structure of tRNA^{Phe}..................................................23

Figure 1.7  Flash-quench reaction scheme and structures of Ru(bpy)$_3^{2+}$ and Co(NH$_3$)$_5$Cl$_2^{2+}$ ........................................................................................................24

Figure 2.1  Cellular iron transport: the transferrin cycle..................................................65

Figure 2.2  Sequence and structures of the human ferritin iron responsive element (IRE)..................................................................................................................66

Figure 2.3  Schematic representation of ferritin translation regulation by IRP binding .............................................................................................................................67

Figure 2.4  Schematic representation of transferrin receptor translation regulation by IRP binding.........................................................................................68

Figure 2.5  Effect of promazine on RNA cleavage by RNase I in human IRE RNA........69

Figure 2.6  Effect of yohimbine on RNA cleavage by RNase I in human IRE RNA...........70

Figure 2.7  Effect of promazine (A) and yohimbine (B) on RNase I cleavage of the stem loop region of human IRE RNA ..................................................71

Figure 2.8  Effect of promazine on the IRP1:IRE interaction........................................72

Figure 2.9  Effect of promazine on the IRP2:IRE interaction........................................73
Figure 2.10 Effect of yohimbine on the IRP1:IRE interaction.................................74
Figure 2.11 Effect of yohimbine on the IRP2:IRE interaction.................................75
Figure 2.12 *In vitro* translation assay to determine the effect of promazine on translation................................................................................................76
Figure 2.13 Effect of promazine on IRE-dependent mRNA translation in cell-free extracts as measured by incorporation of [35S]methionine.............77
Figure 2.14 Effect of promazine on the translation of a mixture of ferritin and luciferase mRNA in cell-free extracts as measured by incorporation of [35S]methionine.................................................................78

Figure 3.1 Schematic of the human insulin sequence and secondary structure........................................................................................................118
Figure 3.2 Human insulin biosynthesis........................................................................119
Figure 3.3 Proposed secondary structures of the 5'UTRs of the human preproinsulin mRNA..................................................................................120
Figure 3.4 Schematic representation of preproinsulin full-length and hairpin RNA constructs................................................................................121
Figure 3.5 Preproinsulin full-length mRNA and hairpin RNA constructs...............122
Figure 3.6 Comparison of guanine cleavage sites in the native and splice variant preproinsulin RNA hairpins..............................................................123
Figure 3.7 Analysis of the native preproinsulin mRNA hairpin structure using RNase cleavage................................................................................124
Figure 3.8 Analysis of the native preproinsulin mRNA hairpin structure using Ru(tpy)(bpy)O2+ mediated guanine oxidation........................................125
Figure 3.9 Analysis of the native preproinsulin mRNA hairpin structure using the flash-quench technique.................................................................126
Figure 3.10 Analysis of the preproinsulin splice variant hairpin RNA structure by RNase cleavage........................................................................127
Figure 3.11 Analysis of the preproinsulin splice variant hairpin RNA structure by Ru(tpy)(bpy)O2+ mediated guanine oxidation..........................128
Figure 3.12 Analysis of the preproinsulin splice variant hairpin RNA structure by guanine oxidation using the flash-quench technique ..................129

Figure 3.13 Effect of promazine and yohimbine on RNase cleavage of the native human preproinsulin hairpin RNA .................................................130

Figure 3.14 Effect of promazine and yohimbine on the RNase cleavage of the human preproinsulin splice variant hairpin RNA ......................131

Figure 3.15 Effect of neomycin B and Hoechst 33258 on the RNase cleavage of the native human preproinsulin hairpin RNA ............................................132

Figure 3.16 Effect of neomycin B on RNase cleavage of the native preproinsulin hairpin RNA ............................................................................133

Figure 3.17 Effect of neomycin B on the RNase cleavage of the human preproinsulin splice variant hairpin RNA ..............................................134

Figure 3.18 Effect of neomycin B on RNase cleavage of the human preproinsulin splice variant hairpin RNA .....................................................135

Figure 3.19 Effect of neomycin B on guanine oxidation in the native preproinsulin hairpin RNA ............................................................................136

Figure 3.20 Effect of neomycin B on guanine oxidation using the RuO$^{2+}$ assay for native human preproinsulin hairpin RNA ..............................137

Figure 3.21 Effect of neomycin B on guanine oxidation in the preproinsulin hairpin splice variant RNA ........................................................................138

Figure 3.22 Effect of neomycin B on guanine oxidation using the RuO$^{2+}$ assay for human preproinsulin splice variant hairpin RNA ..................139

Figure 3.23 Effect of kanamycin B on the RNase cleavage of the native human preproinsulin hairpin RNA ..............................................................140

Figure 3.24 Effect of kanamycin B on RNase cleavage of the native human preproinsulin hairpin RNA ..............................................................141

Figure 3.25 Effect of kanamycin B on the RNase cleavage of the human preproinsulin splice variant hairpin RNA ..............................................142

Figure 3.26 Effect of kanamycin B on RNase cleavage of the human preproinsulin splice variant hairpin RNA .....................................................143
Figure 3.27  Effect of kanamycin B on guanine oxidation in the native preproinsulin hairpin RNA .................................................................144

Figure 3.28  Effect of kanamycin B on guanine oxidation using the RuO\textsuperscript{2+} assay for native human preproinsulin hairpin RNA ........................................145

Figure 3.29  Effect of kanamycin B on guanine oxidation in the preproinsulin hairpin splice variant RNA .................................................................146

Figure 3.30  Effect of kanamycin B on guanine oxidation using the RuO\textsuperscript{2+} assay for human preproinsulin splice variant hairpin RNA ........................................147

Figure 3.31  Effect of Hoechst 33258 on the RNase cleavage of the human preproinsulin splice variant hairpin RNA .................................................................148

Figure 3.32  Effect of hygromycin B on the RNase cleavage of the native human preproinsulin hairpin RNA .................................................................149

Figure 3.33  Effect of hygromycin B on the RNase cleavage of the human preproinsulin splice variant hairpin RNA .................................................................150

Figure 3.34  Effect of neomycin B on guanine oxidation using the flash-quench technique in the preproinsulin hairpin splice variant RNA .................................................................151

Figure 3.35  \textit{In vitro} translation assay to determine the effect of neomycin B and kanamycin B on protein translation .................................................................152

Figure 3.36  Effect of neomycin B and kanamycin B on human preproinsulin mRNA translation in cell-free extracts as measured by incorporation of [\textsuperscript{35}S]methionine .................................................................153

Figure 3.37  Effect of promazine and yohimbine on human preproinsulin mRNA translation in cell-free extracts as measured by incorporation of [\textsuperscript{35}S]methionine .................................................................154

Figure 3.38  Effect of neomycin B and kanamycin B on luciferase mRNA translation in cell-free extracts as measured by incorporation of [\textsuperscript{35}S]methionine .................................................................155

xvii
# LIST OF ABBREVIATIONS AND SYMBOLS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>[5'-32P]pCp</td>
<td>Cytidine 3’, 5’-bis(phosphate)</td>
</tr>
<tr>
<td>8-oxo-guanine</td>
<td>7,8-dihydro-8-oxoguanine</td>
</tr>
<tr>
<td>Å</td>
<td>angstrom</td>
</tr>
<tr>
<td>A</td>
<td>absorbance; adenine; binding equation variable</td>
</tr>
<tr>
<td>aa</td>
<td>amino acid</td>
</tr>
<tr>
<td>ADP</td>
<td>adenosine diphosphate</td>
</tr>
<tr>
<td>ATP</td>
<td>adenosine triphosphate</td>
</tr>
<tr>
<td>γ³²P-ATP</td>
<td>adenosine 5’-gamma³²P-triphosphate</td>
</tr>
<tr>
<td>β</td>
<td>beta</td>
</tr>
<tr>
<td>bp</td>
<td>basepair</td>
</tr>
<tr>
<td>bpy</td>
<td>2,2'-bipyridine</td>
</tr>
<tr>
<td>BSA</td>
<td>bovine serum albumin</td>
</tr>
<tr>
<td>°C</td>
<td>degrees Celsius</td>
</tr>
<tr>
<td>C</td>
<td>cytosine</td>
</tr>
<tr>
<td>Ca²⁺</td>
<td>calcium ion</td>
</tr>
<tr>
<td>cDNA</td>
<td>complimentary deoxyribonucleic acid</td>
</tr>
<tr>
<td>cm</td>
<td>centimeter</td>
</tr>
<tr>
<td>CTP</td>
<td>cytidine triphosphate</td>
</tr>
<tr>
<td>Cys</td>
<td>cysteine</td>
</tr>
<tr>
<td>DMT-1</td>
<td>divalent metal transporter 1 protein</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>ds</td>
<td>double strand</td>
</tr>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
</tr>
<tr>
<td>ε</td>
<td>extinction coefficient</td>
</tr>
<tr>
<td>E. coli</td>
<td>Escherichia coli</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>eIF-2</td>
<td>eukaryotic initiation factor 2</td>
</tr>
<tr>
<td>EMSA</td>
<td>electrophoretic mobility shift assay</td>
</tr>
<tr>
<td>E.T.</td>
<td>electron transfer</td>
</tr>
<tr>
<td>FL-NAT</td>
<td>full-length native human preproinsulin mRNA</td>
</tr>
<tr>
<td>FL-SPV</td>
<td>full-length human preproinsulin splice variant mRNA</td>
</tr>
<tr>
<td>FPLC</td>
<td>fast protein liquid chromatography</td>
</tr>
<tr>
<td>G</td>
<td>guanine</td>
</tr>
<tr>
<td>H-chain</td>
<td>heavy chain of human ferritin</td>
</tr>
<tr>
<td>HCl</td>
<td>hydrochloride</td>
</tr>
<tr>
<td>HEPES</td>
<td>N-[2-Hydroxyethyl]piperazine-N’-[2-ethanesulfonic acid]</td>
</tr>
<tr>
<td>HIV-1</td>
<td>human immunodeficiency virus type 1</td>
</tr>
<tr>
<td>hr</td>
<td>hour</td>
</tr>
<tr>
<td>IC&lt;sub&gt;50&lt;/sub&gt;</td>
<td>inhibitory concentration 50%</td>
</tr>
<tr>
<td>IPTG</td>
<td>isopropyl-beta-D-thiogalactopyranoside</td>
</tr>
<tr>
<td>IRE</td>
<td>iron responsive element, iron responsive element hairpin RNA 50-mer</td>
</tr>
<tr>
<td>IRE-luc</td>
<td>luciferase gene with the IRE sequence upstream</td>
</tr>
<tr>
<td>IRP1</td>
<td>iron regulatory protein 1</td>
</tr>
<tr>
<td>IRP2</td>
<td>iron regulatory protein 2</td>
</tr>
<tr>
<td>Symbol</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>------------------------------------</td>
</tr>
<tr>
<td>kan</td>
<td>kanamycin</td>
</tr>
<tr>
<td>$K_{apparent}$</td>
<td>apparent binding constant</td>
</tr>
<tr>
<td>kb</td>
<td>kilobase</td>
</tr>
<tr>
<td>KCl</td>
<td>potassium chloride</td>
</tr>
<tr>
<td>$K_d$</td>
<td>dissociation constant</td>
</tr>
<tr>
<td>kDa</td>
<td>kilodalton</td>
</tr>
<tr>
<td>L</td>
<td>liter</td>
</tr>
<tr>
<td>L-chain</td>
<td>light chain of human ferritin</td>
</tr>
<tr>
<td>LB</td>
<td>luria broth</td>
</tr>
<tr>
<td>M</td>
<td>molar</td>
</tr>
<tr>
<td>Met</td>
<td>methionine</td>
</tr>
<tr>
<td>MgCl$_2$</td>
<td>magnesium chloride</td>
</tr>
<tr>
<td>µ</td>
<td>micro</td>
</tr>
<tr>
<td>µg</td>
<td>microgram</td>
</tr>
<tr>
<td>µl</td>
<td>microliter</td>
</tr>
<tr>
<td>µM</td>
<td>micromolar</td>
</tr>
<tr>
<td>µs</td>
<td>microsecond</td>
</tr>
<tr>
<td>mg</td>
<td>milligram</td>
</tr>
<tr>
<td>ml</td>
<td>milliliter</td>
</tr>
<tr>
<td>min</td>
<td>minute</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger ribonucleic acid</td>
</tr>
<tr>
<td>N</td>
<td>any of the nucleotides A, C, G, U</td>
</tr>
<tr>
<td>NaCl</td>
<td>sodium chloride</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>NAT</td>
<td>native human preproinsulin mRNA</td>
</tr>
<tr>
<td>NAT-HP</td>
<td>native human preproinsulin hairpin RNA</td>
</tr>
<tr>
<td>ng</td>
<td>nanogram</td>
</tr>
<tr>
<td>NH₄OAc</td>
<td>ammonium acetate</td>
</tr>
<tr>
<td>nm</td>
<td>nanometer</td>
</tr>
<tr>
<td>NMR</td>
<td>nuclear magnetic resonance</td>
</tr>
<tr>
<td>ns</td>
<td>nanosecond</td>
</tr>
<tr>
<td>nt</td>
<td>nucleotide</td>
</tr>
<tr>
<td>OD</td>
<td>optical density/absorbance</td>
</tr>
<tr>
<td>OH</td>
<td>hydroxyl</td>
</tr>
<tr>
<td>P. pastoris</td>
<td>Pichia pastoris</td>
</tr>
<tr>
<td>PAGE</td>
<td>polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PC</td>
<td>prohormone convertase</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>Phe</td>
<td>phenylalanine</td>
</tr>
<tr>
<td>pKₐ</td>
<td>-\log [Kₐ]</td>
</tr>
<tr>
<td>pM</td>
<td>picomolar</td>
</tr>
<tr>
<td>PNK</td>
<td>polynucleotide kinase</td>
</tr>
<tr>
<td>ppIGE</td>
<td>preproinsulin mRNA glucose-responsive translation element</td>
</tr>
<tr>
<td>Q</td>
<td>quencher</td>
</tr>
<tr>
<td>RER</td>
<td>rough endoplasmic reticulum</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>RNase</td>
<td>ribonuclease</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>RRL</td>
<td>rabbit reticulocyte lysate</td>
</tr>
<tr>
<td>rRNA</td>
<td>ribosomal RNA</td>
</tr>
<tr>
<td>RuO$^{2+}$</td>
<td>Ru(tpy)(bpy)O$^{2+}$</td>
</tr>
<tr>
<td>Ru-OH$_2$</td>
<td>Ru(tpy)(bpy)OH$_2$</td>
</tr>
<tr>
<td>Ser</td>
<td>serine</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulfate</td>
</tr>
<tr>
<td>SELEX</td>
<td>systematic evolution of ligands by exponential enrichment</td>
</tr>
<tr>
<td>SPR</td>
<td>surface plasmon resonance</td>
</tr>
<tr>
<td>SPV</td>
<td>human preproinsulin splice variant mRNA</td>
</tr>
<tr>
<td>SPV-HP</td>
<td>human preproinsulin splice variant hairpin RNA</td>
</tr>
<tr>
<td>ss</td>
<td>single strand</td>
</tr>
<tr>
<td>T</td>
<td>thymidine</td>
</tr>
<tr>
<td>TAR</td>
<td>transactivating response region</td>
</tr>
<tr>
<td>TBE</td>
<td>tris, boric acid, EDTA buffer</td>
</tr>
<tr>
<td>Tf</td>
<td>transferrin protein</td>
</tr>
<tr>
<td>TfR</td>
<td>transferrin receptor</td>
</tr>
<tr>
<td>tpy</td>
<td>2,2',2''-terpyridine</td>
</tr>
<tr>
<td>Tricine</td>
<td>N-tris(hydroxymethyl)methylglycine</td>
</tr>
<tr>
<td>Tris</td>
<td>tris(hydroxymethyl)aminomethane</td>
</tr>
<tr>
<td>tRNA</td>
<td>transfer RNA</td>
</tr>
<tr>
<td>tRNA$^{phe}$</td>
<td>transfer RNA encoding the phenylalanine amino acid</td>
</tr>
<tr>
<td>U</td>
<td>uracil</td>
</tr>
<tr>
<td>$\alpha$$^{32}$P-UTP</td>
<td>uridine 5'-alpha-$^{32}$P-triphosphate</td>
</tr>
</tbody>
</table>
UTR untranslated region
UTP uridine triphosphate
W watt
Chapter 1

Introduction: Targeting RNA with Small Molecules

1.1 RNA as a Drug Target

Expression of the numerous genes in a cell must be tightly regulated to provide appropriate levels of RNA and protein production at all times. The cell must interpret many different chemical and physical signals in order to respond to environmental and cellular needs. Organisms make extensive use of protein-based control systems to regulate gene expression; modulating mechanisms of transcription, translation, and mRNA processing or degradation. RNA is essential for transcriptional and translational regulation, protein function, and catalysis. These recently discovered functions of RNA dramatically expand its role in the cell and highlight the potential of targeting RNA for the treatment of a number of disease states.

Currently, most small molecule drug discovery efforts are focused on proteins involved in biological pathways associated with disease, although more emphasis is being placed on targeting small molecules to RNA. The unique structures formed by RNA are thought to provide suitable binding pockets for small molecules,\(^1\) and most small molecules that target RNA have been shown to selectively bind to noncanonically paired regions.\(^2\) Over the next two chapters, my work will focus on investigating the use of small molecules to modulate gene expression through their interaction with structured regulatory elements found in messenger RNA (mRNA).
RNA is an attractive therapeutic target because of its many functions in the cell and its lack of a cellular repair mechanism. RNA also provides a target for drugs where a protein target is unavailable or inaccessible, as is seen in many viral and bacterial targets. Drugs which bind to mRNA structures involved in post-transcriptional regulation can either up- or down-regulate expression of a gene. Up-regulation of gene expression is not often accomplished in protein targeting because drugs are normally used to inhibit protein function. The ability of small molecules such as the macrolide and aminoglycoside antibiotics to interact with bacterial ribosomal RNA has been known for some time, and have set a precedent for targeting structured domains of RNA. More recently, riboswitches, mRNA structures that regulate gene expression in bacteria through the use of small molecules, have been suggested as a new cellular target for antibacterial drug discovery. Pursuing RNA as a therapeutic target increases not only the number of available targets in a cell, but also increases the potential number of small molecules that can be considered RNA ligands. However, our knowledge of RNA recognition by small molecules is still very basic and has not progressed as quickly as that of DNA because RNA in solution forms less predictable structures.

1.2 RNA Structure

RNA exhibits many features similar to those possessed by DNA and/or proteins, which makes them attractive from a targeting standpoint. Its hydrogen bonding capabilities allow RNA to fold into complex structures that often provide scaffolding for specific protein recognition. RNA folds so that the short sequences that are complementary form double helices to form the maximum thermodynamically stable structure possible. The same
single stranded RNA sequence can fold to produce many distinct structures, including regions of canonical and noncanonical base pairing, bulges, internal loops and hairpin loops, and various tertiary interactions (Figure 1.1). RNA molecules usually contain short double-helical regions connected by single stranded stretches. The dominant secondary structural element formed when an RNA chain folds back on itself is the hairpin loop or stem-loop (Figure 1.1D). This structure places complementary sequences spatially close to each other to create base pairs that stabilize the overall RNA structure through hydrogen bonding and base stacking interactions (Figure 1.1A). These base paired regions form A-type helices, with 11-12 base pairs per turn in a right-handed, anti-parallel double-helix. The A-type helix has a very deep major groove and a shallow minor groove, which unlike the major and minor grooves of DNA are less conducive to small molecule binding. The major groove of RNA contains the discriminatory edges of the nucleobases, while the defining feature of the minor groove is the presence of the 2’-hydroxyl groups.

Although areas of complementarity can be found in the RNA sequence, there are also areas that cannot form canonical pairs. In addition to the formation of noncanonical base pairs, helical stems are often interrupted by bulged bases or bulged loops (Figure 1.1C), where bases on one strand have no partner to pair with on the other strand, and by internal loops (Figure 1.1B), in which sequences on both strands cannot be paired. These structures produce defined pockets suitable for binding to other RNAs, proteins, and small metabolites. Bulge loops can be either extrahelical or intrahelical, with larger bulges causing increased destabilization of the RNA structure. One-nucleotide bulges are the most common, with single purine bulges tending to stack intrahelically causing the helix to bend, while single pyrimidine bulges tend to be extrahelical. Internal loops have been
identified as being important in tertiary interactions, and in protein recognition. Even if a stem-loop contains no interrupted regions in its stem, it must contain a terminal loop, sometimes termed a hairpin loop, which links the 5’- to the 3’-strand of its stem. Terminal loops can vary in size, from short sequences that play the same role in RNA as β-turns in proteins, to loops that are large enough to contain stem-loops of their own. RNAs also contain tertiary structure, with areas of secondary structure interacting with nucleotides distant in the primary sequence. While advances in the prediction of RNA secondary structure have been made, it is difficult to predict the three-dimensional structure of RNA due to its complex and dynamic behavior.

Experiments have shown the feasibility of targeting RNA using ligands that recognize the above mentioned structural regions in order to modulate gene expression, but there are few examples in the literature. Years of research have resulted in the development of a paradigm for small molecule-duplex DNA binding based on the primary sequence of the DNA. However, research is very far from designing an RNA ligand based only on the knowledge of an RNA sequence because, unlike helical DNA, RNA is recognized more by its structural features rather than its sequence. Understanding how to enhance or prevent RNA-ligand or RNA-protein interactions by blocking their interaction, or by disrupting secondary and tertiary structure, will open up new opportunities for drug design. In order to target small molecules to RNA, a basic understanding of the overall RNA structure is necessary. Determination of the structure of large RNA molecules through crystallographic and nuclear magnetic resonance (NMR) techniques is notoriously difficult. Fortunately, these techniques can be used to characterize the structure of smaller fragments of the RNA that contain important structural elements. Many chemical and enzymatic
techniques have also been used to characterize the structural features of RNA. An important
advantage of these techniques is that the quantity of RNA required for analysis (nanograms)
is much smaller than that required (milligrams) for structural studies by NMR and
crystallography.\textsuperscript{42} Turning these techniques into a screen for ligands that bind in a specific
structural region will enhance the ability to identify small molecules that could modulate
gene expression.

1.3 Chemical and Enzymatic Mapping Techniques for RNA

In RNA mapping studies, the RNA of interest is exposed to a reagent and the extent
of the reaction is quantified at affected nucleotides in the RNA. The reactivity of the
nucleotide depends on its structural constraints in solution. There are two broad classes of
reagents commonly used for RNA mapping, RNases used in selective enzymatic degradation,
and reactive small molecules. Mapping of the entire RNA structure often requires multiple
reagents be used in parallel. The choice between using chemical and enzymatic methods for
probing structural features of RNA depends on what and how information is obtained. There
are advantages and disadvantages to both techniques, but their combined use can give
complimentary information important for determining the overall structure.\textsuperscript{43}

1.3.1 Enzymatic Mapping Techniques

There are a number of RNases that vary in their specificity for single or double
stranded regions of RNA, or the sequences recognized.\textsuperscript{44} In general, enzymatic reactions
allow more flexibility in the choice of reaction conditions. The conditions used can be
similar to physiologically significant conditions unlike those used in chemical mapping.\textsuperscript{43} In
practice, sequence and structural information is determined using a number of enzymes, each
of which is used to identify a specific nucleotide or structural region. Reaction conditions
used maintain the secondary structure of the RNA and allow only single hit kinetics to occur.
The most commonly used RNases are RNase T1, RNase A, RNase I, and RNase V1.
Cleavage occurs with RNase T1 at unpaired guanine residues while RNase A cleaves at
unpaired cytidine and uridine residues. RNase I cleaves after all four residues equally in
unpaired regions and RNase V1 shows a preference for paired or stacked nucleotides.
Dabrowiak and coworkers,45-47 as well as others,48-52 have used RNases to map the binding
sites of small molecules on RNA.

1.3.2 Chemical Mapping Techniques

Chemical nucleases can be used as a good compliment to the enzymatic nucleases
commonly used to probe RNA-small molecule interactions. These nucleases have a specific
geometry, relatively rigid shapes, and are particularly sensitive to the tertiary structure or
shape of the RNA binding site.44 They are also smaller in size than protein nucleases,
allowing them to intimately react with the DNA or RNA substrate. Chemical modification
reagents such as dimethyl sulfate and diethyl pyrocarbonate selectively modify base
positions,53 whereas small molecule reagents such as hydroxyl radicals,45 lead,54 and metal
complexes,44,55-57 often target the phosphoribose backbone of RNA.56 The following section
will focus on two metallonucleases, Ru(tpy)(bpy)O2+ and Ru(bpy)32+, (Figure 1.2), that were
used for chemical mapping in this work. Both compounds interact with RNA by oxidizing
the guanine nucleobase.

1.3.2.1 Nucleic Acid Reactivity to Oxidation

Nucleic acids are susceptible to damage in cells by a variety of mechanisms including
oxidation. Damage via oxidation can occur either at the sugars through hydrogen
abstraction, or at the nucleobases via radical reactions or electron transfer.\textsuperscript{58,59} Sugar oxidation occurs nearly five times slower in RNA than in DNA due to the stabilizing effect of the 2’-OH present in the ribose sugar,\textsuperscript{60,61} and RNA is much more susceptible to hydrolytic cleavage of the phosphodiester bond.\textsuperscript{23,62,63} Base oxidation pathways are similar in both DNA and RNA.

The reactivity of the nucleobases toward oxidation is of the order G > A > C \approx T / U, corresponding to the oxidation potentials of the bases.\textsuperscript{64} Since guanine has the lowest oxidation potential, this is where base damage usually occurs \textit{in vivo}, occurring either directly at the guanine position or through migration to a region containing guanine. Nucleobase cleavage requires a second chemical step, often alkali treatment, to cause complete strand scission.\textsuperscript{59} In DNA, this alkali treatment causes hydrolysis of the glycosidic bond, leaving an abasic site, which can undergo $\beta$-elimination of the 3’ phosphate.\textsuperscript{59} In RNA, all phosphodiester linkages are subject to alkaline hydrolysis due to the presence of the 2’-hydroxyl. Hydrolysis of the phosphodiester linkages occurs faster in RNA than the creation of an abasic site and $\beta$-elimination.\textsuperscript{59} In practice, aniline is used to facilitate complete strand scission in RNA, while piperidine is used for DNA.

One electron oxidation of RNA bases leads to the production of a guanine radical cation\textsuperscript{59} that is a relatively strong acid (pKa = 3.9).\textsuperscript{65} In the mechanism of oxidation determined in DNA (\textbf{Figure 1.3}), the guanine radical undergoes rapid deprotonation to generate the neutral radical. This is then followed by $\text{O}_2$ addition and the eventual production of the products imidazolone and oxazolone. Alternatively, the guanine radical cation can become hydrated followed by one electron oxidation to form 8-oxo-guanine.\textsuperscript{59} 8-oxo-guanine is the only product detected \textit{in vivo} and has become an important marker for
DNA damage. The preferential degradation pathway of the guanine radical cation (hydration vs. deprotonation) depends on the context of the guanine. Nucleosides and single-strands of nucleic acid are expected to favor the imidazolone and oxazolone pathway, whereas base paired guanines are expected to favor the 8-oxo-guanine pathway due to stabilization of the radical cation.

1.3.2.2 Metallonucleases as RNA Mapping Reagents

The ability for metallonucleases to interact with DNA and cause chemical modification has been known for some time, and their mechanism of action and resulting cleavage products have been studied in great detail. Many of these same compounds have also been used to map the structure of RNA, although the mechanism has not been studied as in depth, and has in many cases been assumed to occur in the same manner as that of DNA. Some metallonucleases are able to interact directly with the nucleobase due to their high oxidation states that are generated either chemically or electrochemically, while other cleavage agents are activated photochemically. An appealing characteristic of photochemical reagents is that all the components of a system to be studied can be mixed together without initiating the chemical reaction until the sample is irradiated. Provided that the chromophore of the photocleavage agent is sensitive to light greater than 300 nm in wavelength, select excitation of the cleavage agent will occur. The ability to “tune” the nuclease by changing the ligands allows complexes to be synthesized to match the desired nucleic acid target. Various metal ligands can be added to the metallonuclease to change their photophysical, photochemical, and redox properties (Figure 1.4). These can be intercalating ligands such as dipyridophenazine, electrostatic or solvent bound ligands such as bipyridine, or groove binders such as phenanthroline. Thorp and coworkers,
among others, have explored many DNA and RNA structures using transition metal oxidation chemistry. The following discussion focuses on two metallonucleases, Ru(tpy)(bpy)O$_2^+$ (RuO$_2^+$, bpy = 2,2'-bipyridine, tpy = 2,2',2''-terpyridine) and Ru(bpy)$_3^{2+}$, and their use as probes of RNA structure.

1.3.2.3 Ru(tpy)(bpy)O$_2^+$ as an RNA Chemical Mapping Agent and Screen for Small Molecule Binding

Meyer and coworkers first reported the synthesis of Ru(tpy)(bpy)O$_2^+$ in 1981 and later reported that the complex could oxidize organic hydrocarbons and alcohols. The complex is made as the aqua complex and is converted via bulk electrolysis to the oxidized oxo complex through a coupled two proton, two electron process. Since the discovery that RuO$_2^+$ could oxidize DNA under anaerobic conditions, RuO$_2^+$-nucleic acid oxidation chemistry has been studied extensively. Due to its size (6 Å) and charge (2+), the RuO$_2^+$ compound shows preferential binding to solvent accessible guanines that are prone to cation binding. It is postulated that oxidation occurs through an inner-sphere Ru(III)-O-G intermediate. Hairpin structures such as the transactivation response (TAR) RNA sequence, tRNA$^\text{Phe}$, and the iron responsive element (IRE) sequence found in human and bullfrog ferritin have been oxidized with RuO$_2^+$, and show the greatest extent of damage at loop guanines (Figures 1.5 and 1.6).

tRNA$^\text{Phe}$ (Figure 1.6) was studied using RuO$_2^+$ chemistry due to the available X-ray crystal structure and the ability to control the RNA conformation via salt concentration. In the folded state, tRNA$^\text{Phe}$ was cleaved mainly at guanines and adenines in the D- and anticodon loops. Guanines in the TAR RNA sequence from HIV-1 (Figure 1.5C) were
also studied using RuO$_2^{2+}$ chemistry, and it was observed that guanine nucleobases found in the loop (G$_{17}$, G$_{18}$, and G$_{19}$) were oxidized more efficiently than guanines found in the stem.

The iron responsive element has also been studied using RuO$_2^{2+}$ oxidation chemistry. In the first study, the IRE in full-length bullfrog ferritin mRNA (Figure 1.5A) was probed with RuO$_2^{2+}$. Cleavage was seen only at position G$_{16}$ (the IRE has been renumbered since the original publication, and was originally termed G$_{14}$), and not at any other unpaired guanine residues indicating that this residue was highly solvent exposed. It was concluded that guanine residues other than G$_{16}$ were spatially inaccessible to the reagent.

Lack of cleavage at other unpaired guanine residues was later explained by NMR structural data that suggested that the guanines predicted to be unpaired in the secondary structure were involved in tertiary interactions with other parts of the RNA molecule. Further studies confirmed that the G$_{16}$ residue is flipped out of the RNA structure and is solvent exposed, allowing it to make important contacts with the iron regulatory protein (IRP). Ciftan et al. studied oligonucleotides corresponding to the 30 base IRE hairpin structure and 55 base IRE hairpin with flanking region. RuO$_2^{2+}$ cleavage revealed more cleavage sites than in the previous study although G$_{16}$ was still oxidized to the greatest extent. Other unpaired guanine residues were also oxidized, which was consistent with the expectations from the predicted secondary structure.

Tibodeau used RuO$_2^{2+}$ oxidation chemistry to map the human ferritin IRE RNA structure (Figure 1.5B), and further explored the use of RuO$_2^{2+}$ to screen for small molecules that interact with IRE RNA. The development of RuO$_2^{2+}$ chemistry into a footprinting technique would be a useful screen for identifying functional small molecule-RNA interactions. The ability of thirteen small molecules to interact with a specific binding pocket
in the IRE RNA was probed using RuO$_2^+$
. Three molecules that bound to this pocket were identified, and the effect of one small molecule, yohimbine, on ferritin gene expression was examined. This small molecule was able to up-regulate ferritin protein synthesis in vitro, validating the use of RuO$_2^+$ as a small molecule screen.

1.3.2.4 Ru(bpy)$_3^{2+}$ as a Chemical Mapping Agent: The Flash-Quench Technique

Photolysis of a ruthenium chemical mediator to cause oxidative damage in DNA has been studied extensively by Barton and coworkers.$^{91,92}$ This oxidative damage occurs primarily through base-labile guanine oxidation. The metal complexes have been used to study events such as mismatch recognition$^{93,94}$ and electron transfer from peptides to DNA.$^{95,96}$ Transition metal complexes such as ruthenium and rhodium can act as direct photooxidizers in the flash-quench experiment,$^{91}$ the reaction scheme for which is shown in Figure 1.7. The flash-quench cycle is first initiated by visible light which excites the Ru$_2^{2+}$ complex. The excited ruthenium(II) complex *Ru$_2^{2+}$ is then quenched by a nonintercalating electron acceptor (Q), to form Ru$_3^{3+}$. The Ru(III) species can be reduced back to the Ru(II) form either through recombination with the reduced quencher (Q$^-$), or through electron transfer with a nearby guanine base resulting in an oxidized guanine radical. The oxidized guanine radical can return to the resting state by reaction with the reduced quencher or undergo further reaction to form the guanine oxidation products shown in Figure 1.3. The flash-quench technique was first used to characterize electron transport in proteins,$^{97}$ and was later modified for DNA oxidation.$^{91}$ Recently, Dana Holcomb in our lab has adapted this technique for its use in the oxidation of RNA. (unpublished results)

This flash-quench reaction for oxidation of guanine nucleobases in RNA uses the Ru(bpy)$_3^{2+}$ complex as the photooxidizer and Co(NH$_3$)$_5$Cl$_2^{2+}$ as the quencher. Ru(bpy)$_3^{2+}$ is a
common example of an external RNA binder, and its binding is governed by both solvent binding and electrostatic contributions.\textsuperscript{74,98} The electron transfer between guanine and Ru(bpy)\textsubscript{3}\textsuperscript{3+} occurs in less than 200 ns, and once formed, nearly all the guanine radicals decay within 100 µs.\textsuperscript{91,92,95} The irreversible reaction of the guanine radical with either oxygen or water yields oxidative lesions that can be analyzed by gel electrophoresis.\textsuperscript{95} The flash quench-method of RNA mapping showed oxidation of the guanine residues in the human ferritin IRE in a similar pattern to those identified using the RuO\textsubscript{2+} compound. Preferential oxidation was seen at those residues which were readily solvent accessible and not residues whose flexibility allowed them to become solvent exposed over time. The use of Ru(bpy)\textsubscript{3}\textsuperscript{2+} as a footprinting technique continues to be examined in this work.

\textbf{1.4 Molecular RNA Targets}

Promising results have been obtained using a number of small molecules to target RNA, offering a complimentary approach to the targeting of proteins. Small molecule ligands have been identified with three major classes of RNA targets in mind: antibacterial, antiviral, and mRNA. Within each of these classes, various avenues have been pursued to achieve gene expression regulation, including inhibiting RNA-protein interactions and preventing protein production by binding to a functionally relevant RNA. To date, all clinically approved RNA-targeting drugs exert their effect by binding to ribosomal RNA.\textsuperscript{26}

\textit{1.4.1 Antiviral Targets: Transactivation Response RNA (TAR RNA)}

Once it was determined that small molecules could regulate gene expression both \textit{in vitro} and \textit{in vivo}, work progressed on endogenous RNA targets. RNAs with highly structured regions were identified, with a focus on those structured regions involved in an
RNA-protein complex. One example of this interaction is between the TAR RNA and the Tat protein involved in the regulation of HIV-1.

The HIV-1 genome is made up of a ∼9 kb genomic RNA encoding for 15 proteins. After integration into the host cell genome, the initial phase of the viral life cycle begins with transcription from the 5’ end of the RNA. Although transcription is critical to viral proliferation, RNA polymerase II transcribes poorly from the viral promoters. The HIV-1 protein Tat helps facilitate efficient transcription of the viral genome. In the presence of the Tat protein, the rate of transcription increases nearly 100-fold to produce full length genomic RNA. Inhibition of the Tat-TAR interaction by binding to TAR, in order to slow or halt production of the viral transcript, has been sought as a potential anti-HIV strategy. Tat has been shown to bind to a specific bulged region in the TAR RNA hairpin loop at the beginning of the viral transcripts.

A variety of techniques have been used to identify compounds able to inhibit the Tat-TAR interaction, including gel mobility shift assays, absorption and fluorescence spectroscopy, mass spectrometry, footprinting experiments, scintillation proximity assays, NMR, and computational screening. Initially, a collection of aminoglycosides was screened for their ability to disrupt the Tat-TAR interactions. Neomycin was found to be the most potent aminoglycoside with an IC50 of 0.92 µM. In subsequent experiments it was demonstrated that the binding site for neomycin was located below the bulged region in TAR RNA. It has been suggested that binding of neomycin induces a conformational change in the TAR RNA, increasing the off-rate of the Tat-TAR complex. Later, Mei and coworkers developed a high-throughput in vitro screening method and were able to screen approximately 150,000 compounds for their ability to interrupt the Tat-TAR interaction.
From this screen, two promising compounds were identified, quinoxaline and tetraaminoquinozaline. The James group has used computer-aided drug design to identify compounds that are candidates for blocking the Tat-TAR interaction.\textsuperscript{101,110} Acetopromazine was identified through the use of a docking program and was found to completely inhibit formation of the Tat-TAR complex at a concentration of 100 nM.\textsuperscript{101} An NMR structure was obtained, showing that acetopromazine bound to TAR RNA in an area containing a trinucleotide bulge.\textsuperscript{110} It was later shown that acetopromazine was able to bind internal bulges and terminal loops, but not double stranded RNA or tetraloops.\textsuperscript{111}

1.4.2 mRNA Targets: Inhibiting Ribosome Scanning

Translation in eukaryotes proceeds first with the binding of the small ribosomal subunit to the eIF-2-GTP-Met-tRNA\textsubscript{i} complex, followed by binding to the 5’ cap of mRNA.\textsuperscript{112} The 5’ end of the mRNA molecule is recognized by the presence of its 5’ cap and two bound initiation factors, eIF4e and eIF4G. The small ribosomal subunit complex then begins scanning the transcript, searching the 5’UTR for the first AUG codon. After recognition of the codon by Met-tRNA\textsubscript{i}, the complex pauses and waits for the large ribosomal subunit to associate before translation begins. Secondary structure in the 5’UTR is known to pause or inhibit the scanning process, resulting in translation inhibition. By binding of a small molecule within the 5’UTR of a specific mRNA, inhibition of translation and modulation of gene expression can occur.

The \textit{in vitro} selection process known as SELEX (systematic evolution of ligands by exponential enrichment) has been used to identify ligands that are able to regulate gene expression \textit{in vivo}.\textsuperscript{4,34,113} Werstruck and Green\textsuperscript{34} have identified an aptamer for Hoechst dye 33258, which when inserted in to the 5’UTR of a mammalian β-glactosidase mRNA,
decreases the \textit{in vivo} expression of β-galactosidase activity by greater than 90%. Internal controls determined that the effects of Hoescht 33258 were not due to general translation inhibition. In the same manner, Grate and Wilson\textsuperscript{113} have placed an aptamer to malachite green in the 5′UTR of the \textit{CLB2} gene, which encodes for a cyclin that directs the cell cycle transition from G2 to mitosis in budding \textit{S. cerevisiae}. They have shown that ligand-induced folding of the 5′UTR limits binding by the 40S ribosomal subunit and results in a reduction in the rate of translational initiation.

\textbf{1.5 Principles of RNA Binding}

Recent work has focused on determining the biochemical and biophysical rules governing small molecule-RNA binding.\textsuperscript{16,21,107,114} Compounds that interact with RNA show marked differences from those typically identified as being “drug like” by Lipinski’s rules and can make targeting a specific RNA difficult.\textsuperscript{1,14-16,21,115} The small molecules often exceed the established molecular weight limits and are more hydrophobic or polar in nature than common drugs, due to the need to interact with the negatively charged RNA. Both of these features could cause a reduction in the bioavailability of the compound. Most small molecules also have only modest selectivity and affinity for their RNA target when compared to their protein counterparts.\textsuperscript{26} The binding constants for many RNA-small molecule interactions are relatively weak, falling in the low micromolar range.\textsuperscript{14} These drugs can however be sufficiently potent if binding causes or inhibits a conformational change in the RNA.

The general affinity of the aminoglycosides for many different RNAs leads to their often being the first compounds investigated for binding to a new RNA target. Due to their
promiscuity, the compounds are often useful for in vitro studies, but less useful for cell culture or in vivo work. Aminoglycosides can also be very toxic at high concentrations. Still, a large amount of data has been collected on the aminoglycoside-RNA interaction, and certain basic principles have emerged that will aid in designing future RNA ligands. Aminoglycosides have been found to bind pockets created by bases in and around internal loops and bulged regions or those created by noncanonical interactions. They have shown a strong dependence for electrostatic interactions, as well as nonionic interactions. They also have shown that shape complementarity (how well the ligand and RNA fit together) as well as conformational adaptation (binding induced changes in the ligand and RNA) are a major contributor of the specificity of small molecules for RNA.

Using the knowledge of binding principles provided by the aminoglycosides, the Hergenrother group have shown that subsets of deoxystreptamine dimers are able to selectively bind RNA tetraloops and octaloops through the use of a combinatorial library of 105 deoxystreptamine dimers. Ligand binding was assessed using fluorescein labeled RNA to obtain an estimate of the dissociation constant and the ligand binding site was determined using RNase footprinting. Two compounds with nanomolar affinity for hairpin loops were identified using this screen, one showing preferential binding to a tetraloop, and the other to an octaloop. These were the first compounds reported to show discrimination between RNA hairpin loops of various sizes. More work needs to be done to identify the RNA motifs that small molecules prefer so that compounds can be efficiently designed. In this same area, understanding the characteristics of the small molecules that increase their effectiveness on gene expression modulation is important.
The rational discovery of small molecule ligands for RNA is still in its infancy compared to the use of these small molecules to bind proteins and DNA. The complex structure of RNA increases the difficulty in the identification of potential ligands. To date, the small molecules identified have little in common as far as their structures and modes of interaction, making it hard to generalize what qualifies as a good RNA binder. Detailed characterization of small-molecule RNA interactions, along with methods for high throughput screening are necessary to progress toward the goal of selectively targeting RNA with small molecules.

1.6 Dissertation Focus

This dissertation focuses on the use of secondary structures in mRNA as a small molecule drug target for regulation of gene expression. Two model systems are investigated, the iron responsive element in the human ferritin mRNA, which is studied in Chapter 2, and the preproinsulin glucose regulatory element found in the human preproinsulin mRNA studied in Chapter 3. Chapter 2 focuses on the ability of a compound identified using oxidation chemistry to regulate ferritin gene expression. In Chapter 3 the secondary structure of the preproinsulin glucose regulatory element is first mapped with transition metal complexes and RNases, then the mapping agents are used to screen for small molecules that bind to the preproinsulin mRNA. The effect of these compounds on gene regulation is further studied using an in vitro cell-free expression system.
Figure 1.1. Schematic representation of four general classes of RNA secondary structure. A) Base paired RNA forming a helix or duplex structure. B) In an internal loop secondary structure, sequences on both strands cannot be paired. C) A bulged region in RNA occurs when bases on one strand have no partner on the other strand to base pair with. D) Hairpin loop, or stem loop region consisting of a base paired “stem” region with an unpaired terminal loop. The bold line represents a canonical base pair, whereas the circle represents a noncanonical G-U wobble pair.
Figure 1.2. Structures of the two metal complexes used as metallonucleases in this work

Ru(tpy)(bpy)O2⁺  
Ru(bpy)₃²⁺
Figure 1.3. Proposed mechanism of guanine oxidation in DNA (adapted from Burrows\textsuperscript{59}).
Figure 1.4. Common aromatic ligands used to change the binding properties of metallonucleases. Bipyridine (bpy) ligands (A) are an example of an electrostatic ligand and also known as external binders. 3,4,7,8-tetramethylphenanthroline (B) and phenanthroline (C) are known groove binders in DNA. 9,10-phenanthrenequinone diimine (D), dipyridophenazine (E), and 4,7-diphenyl-1,10-phenanthroline (F) are known intercalators.
Figure 1.5. Schematic representations of the secondary structure of RNAs mapped with metallonucleases. A) Proposed bullfrog ferritin IRE secondary structure. Bold residues represent bases that make direct contacts with the IRP protein. B) Proposed structure of the human ferritin IRE. There is a base pair across the hairpin loop between C\textsubscript{14} and G\textsubscript{18}. C) Proposed secondary structure of the TAR element found in HIV-1. Bold residues represent bases critical for binding of the Tat protein.
Figure 1.6. Proposed secondary structure of tRNA$^{\text{Phe}}$. Most guanine oxidation occurs in the D- and anticodon loops.
Figure 1.7. Flash-quench reaction scheme and structures of Ru(bpy)$_3^{2+}$ and Co(NH$_3$)$_5$Cl$_2^{2+}$. In this scheme, G = Guanine, Q = Quencher. Visible light excites the Ru(bpy)$_3^{2+}$ species to form *Ru(bpy)$_3^{2+}$. This excited complex is then quenched by the nonintercalating Co(NH$_3$)$_5$Cl$_2^{2+}$ electron acceptor to form Ru(bpy)$_3^{3+}$. The Ru(bpy)$_3^{3+}$ can then be reduced back to Ru(bpy)$_3^{2+}$ either through recombination with the reduced quencher or through electron transfer with a nearby guanine base resulting in a guanine radical. The guanine radical can return to its resting state by reacting with the reduced quencher, or can undergo further reaction to form 8-oxo-guanine or one of its reaction products.
1. 8 References


(114) Hermann, T. Biopolymers 2003, 70, 4-18.


Chapter 2

*In Vitro* Up-regulation of Ferritin Protein Synthesis by Small Molecules that Target the Iron Responsive Element in Ferritin mRNA

2.1 Abstract

RNA is an attractive therapeutic target because of its many functions in the cell and its lack of a cellular repair mechanism. Previous work with a Ru(tpy)(bpy)O$_2^+$ (RuO$_2^+$, bpy = 2,2'-bipyridine, tpy = 2,2',2''-terpyridine) screen has identified four small molecules that are able to specifically and selectively bind to the iron responsive element (IRE) structure in human ferritin mRNA. One small molecule, promazine, has been identified that binds to an internal bulge region in the IRE sequence that makes important contacts with the iron regulatory protein (IRP). An RNase I footprinting assay has confirmed that promazine binds in the bulge region of the IRE. Using an Electrophoretic Mobility Shift Assay (EMSA), we have shown that this compound can disrupt the binding of IRP1 and IRP2 to the IRE sequence. Finally, binding of promazine significantly increases the production of ferritin in an *in vitro* translation system, and is also able to up-regulate synthesis of the luciferase protein from an engineered IRE-luciferase gene.
2.2 Introduction

2.2.1 Cellular Iron Homeostasis

Iron (Fe) is both an essential nutrient and a potential toxin that plays a role in the metabolism of all cells.\(^1\) It is critical to many diverse functions of the cell, such as DNA replication, electron transfer, and oxidative stress control. Paradoxically, the presence of high levels of iron can be harmful, causing inflammation, tissue damage, and toxicity. At physiological pH, ferric iron (Fe\(^{3+}\)) is extremely insoluble and its ability to react with oxygen to produce dangerous free radicals leads to extremely low free iron concentrations in the cell, with much of the Fe\(^{3+}\) complexed with proteins.\(^2\) It is the balance between the beneficial aspects of cellular iron and its toxicity that is a challenge. This leads to tight regulation of the acquisition, storage and utilization of cellular iron.

An intricate system has been developed to handle the iron needs of mammalian cells. This system involves iron transport and storage proteins (transferrin and ferritin, respectively), RNA-binding proteins (iron regulatory proteins 1 and 2; IRPs) that regulate the transport and storage proteins, and specific RNA sequences/structures (iron responsive elements; IREs) that are the recognition sites for the RNA-binding proteins. Much work has been done by Theil,\(^3-6\) Munro,\(^7-9\) Rouault,\(^10-12\) Leibold,\(^13,14\) and others to elucidate the details and mechanisms of this system. While there are still some unanswered questions, this system is one of the most well understood regulatory systems to date.

2.2.1.1 Iron Absorption and Transport

The human body contains \(~3-5\) g iron, with only 0.5-2 mg entering and leaving the body daily.\(^15\) The vast majority of iron in the body is used in erythroid cells for heme and hemoglobin synthesis. In mammals, non-heme iron (Fe\(^{3+}\)) is absorbed from food by
enterocytes in the duodenum. The ferric iron is first reduced by a ferric reductase to Fe$^{2+}$ and is transported into the enterocyte by the transmembrane iron transporter DMT1 (divalent metal transporter 1). Once inside, excess iron is either stored in ferritin, or is exported outside the cell through the basolateral membrane to reach plasma by an iron exporter, ferroportin1. The Fe$^{2+}$ is then oxidized to Fe$^{3+}$ by a ferroxidase, likely either the plasma ferroxidase ceruloplasmin, or its membrane associated analogue hephaestin.$^{15-18}$ Ferric iron is then loaded into apo-transferrin to give holo-transferrin, although the mechanism by which this occurs is not yet understood. In this form the insoluble Fe$^{3+}$ can circulate throughout the body in the plasma and is no longer toxic.

2.2.1.2 Cellular Iron Uptake: Transferrin-mediated Endocytosis

There are four cell types that have special functions in iron handling; the duodenal enterocytes that function to absorb iron from food sources, erythroid precursors which incorporate all the iron they uptake into hemoglobin, hepatocytes which store the iron not needed elsewhere, and the reticuloendothelial macrophages which carry out iron recycling and also store iron.$^{15}$ Each of these cell types handles iron in a different way. A number of mechanisms have evolved for the uptake of iron, but the major method is the transferrin-mediated endocytosis pathway (Figure 2.1).$^{18}$

The transferrin protein (Tf) has a molecular mass of approximately 80 kDa and binds tightly to the Fe$^{3+}$ ions with an approximate $K_{\text{apparent}}$ of $10^{-20}$ M$^{-1}$, making the iron difficult to extract forcing the entire complex to be brought into the cell before iron is released.$^{19}$ Holo-transferrin binds the transferrin receptors (TfR1 and TfR2; generally TfR) at the cell surface. The TfR/Tf complexes localize to clathrin-coated pits, which are endocytosed and form a specialized endosome. After the loss of the clathrin coat, the endosomes are acidified to pH
5.5 through the use of a proton pump. This leads to conformational changes of the transferrin protein and release of the iron. The apo-transferrin and TfR are recycled to the cell surface where they dissociate at neutral pH and can be used in further cycles of iron binding and uptake. Subsequently, the released Fe$^{3+}$ is reduced to Fe$^{2+}$ inside the endosome and is transported by DMT1 into the cytoplasm where it is either utilized by the cell or stored.

2.2.1.3 Cellular Iron Storage: Ferritin

The portion of iron taken into the cell that is not needed for immediate use is stored as a reserve in case body iron levels become low. Because free iron would aggregate and form toxic precipitates, the iron is stored in the protein ferritin. Ferritin (480 kDa) is a ubiquitous, highly conserved binding protein that can store up to 4500 iron atoms in its hydrated ferric oxide core. In mammals, ferritin is a heteropolymer made up of 24 subunits of two types, heavy (H-, 21 kDa) and light (L-, 19 kDa), which are encoded by their corresponding genes. The ratio of H to L subunits is dependent on the tissue in which the ferritin is expressed. The H subunit contains inherent ferroxidase activity, converting Fe$^{2+}$ found in the cytosol to Fe$^{3+}$ as the iron is internalized into the mineral core. The L subunit has been shown to play a role in the formation of the iron core. Degradation of ferritin resulting in iron release helps to mobilize iron for utilization by the cell. Although little is known about the degradation mechanism, lysosomal and proteosomal pathways have been implicated. The importance of ferritin is demonstrated by the fact that ferritin H gene deletions are lethal in mice and mutations in the gene encoding the L subunit causes defects in the central nervous system in mice.
2.2.2 Regulation of Cellular Iron Homeostasis: Regulation of Transferrin and Ferritin by Iron Regulatory Proteins (IRPs) and Iron Responsive Elements (IREs)

Together ferritin and transferrin regulate the concentration of cellular iron. They are under tight genetic control and their expression can be up- or down-regulated depending on iron availability. When presented with iron overload, transferrin receptor synthesis is down-regulated to stop iron uptake, and ferritin synthesis is increased to sequester excess iron. In times of iron deficiency, a higher number of TfRs are presented at the cell surface to increase the amount of internalized iron, and synthesis of the ferritin storage protein is stopped thereby increasing iron availability. The balance between ferritin and transferrin synthesis is achieved by regulatory mechanisms that include modulating transcription, mRNA stability, translation, and post-translational modification. Post-transcriptional regulation is currently the best characterized. This system includes trans-acting iron regulatory proteins (IRPs) interacting with iron responsive elements (IREs), conserved secondary structures found in untranslated regions of the mRNA.

There are two known IRPs, IRP1 and IRP2. Both proteins are universally expressed and have evolved to sense cellular iron concentration and modify gene expression accordingly. IRP1 is involved in basal mammalian iron homeostasis, while IRP2 is the predominant post-transcriptional regulator of iron metabolism in response to the level of iron in the cell. Human IRP1 is reported to share 57-61% sequence identity and 75-79% similarity with human IRP2. Each protein is made up of four domains with domains 1-3 connected to the fourth domain by a flexible hinge region. This arrangement creates an active site cleft between the three compact domains and the fourth. Mutagenesis and
footprinting of the IRE-IRP1 interaction have shown that residues close to the entrance of the active site are important in IRE binding.

IRP1 (previously called translation regulator protein P-90, iron-responsive element-binding protein 1, iron regulatory factor, or ferritin repressor protein) is a 98 kDa protein coded in human chromosome 9 and has been more thoroughly characterized than IRP2. It is a bifunctional protein, which can act as a functional cytosolic aconitase, converting citrate to isocitrate, or an RNA-binding protein. A recent crystal structure of IRP1 reveals that the structure is very similar to that of the mitochondrial aconitase. The modulation between the aconitase form and the RNA-binding form of IRP1 depends on the presence of an iron-sulfur [4Fe-4S] cluster that blocks RNA binding. The iron-sulfur cluster is coordinated in the active site by three cystines, leaving the fourth residue to bind solvent and substrate. When the iron-sulfur cluster is bound in the active site, IRP1 loose the ability to bind to RNA. In iron deficient cells or those under oxidative stress, the cluster disassembles and apo-IRP1 gains RNA binding activity. This RNA binding activity can act to either stabilize TfR mRNA, or repress ferritin synthesis in cells. IRP1 also responds to other stimuli. Exposure of cells to hydrogen peroxide (H2O2) or nitric oxide (NO) promotes removal of the [4Fe-4S] cluster and induces IRE-binding activity.

The active site for RNA binding has been shown to overlap that of the aconitase active site. The crystal structure of IRP1 holo-protein shows a closed conformation (domain 4 closing onto domains 1-3) that does not allow access of substrates to the active site. The IRE has a shape similar to the portion of domain 4 that interacts with domains 1-3 leading to the hypothesis that the IRE binding is able to substitute for domain 4 binding, resulting in a widened cleft region. A more detailed characterization of the IRE binding
site in IRP1 has become available with the crystallization of IRP1:ferritin H-chain IRE RNA. This work has also lead to a greater understanding of the RNA-protein contacts that are made between IRE RNA and IRP1.

IRP2 is a 104 kDa protein that is highly homologous to IRP1 but lacks aconitase activity most likely due to its inability to assemble an [4Fe-4S] cluster. IRP2 is distinguished from IRP1 by a 73 amino acid segment that is responsible for functioning as an iron sensor and mediating proteolytic cleavage of the protein in iron replete cells, though the details of this mechanism are not yet known. In the presence of iron, proteosomes bind to the IRP2 protein and degrade it, preventing its interaction with IREs. Mice that lack IRP2 develop anemia due to insufficient erythroid expression of TfR and overexpression of ferritin that depletes cells of iron needed for heme synthesis. Mice exhibit a 200-fold increase in synthesis of the porphyrin heme precursor protoporphyrin IX than wild-type mice; in adulthood these mice develop progressive neurodegenerative disease. IRP2 has been shown to dominate regulation of iron homeostasis and can compensate for the loss of IRP1.

IRP1 and IRP2, in their RNA binding forms, have been extensively studied. The interaction between the protein and RNA is very strong with the dissociation constant between IRP1 and the wild-type ferritin IRE being between 10 and 90 pM. Both IRPs bind with equally high affinity and specificity to a consensus IRE sequence, although each IRP binds a different subset of IRE-like sequences.

A number of genes that code for proteins involved in iron homeostasis contain a non-coding sequence of 28-30 nucleotides called the iron responsive element (IRE). IREs are highly conserved stem loop RNA structures that act as IRP binding sites. There are
several aspects of IRE secondary structure that are important for high-affinity binding to the IRPs. The consensus IRE sequence contains a conserved 6-nt loop with the sequence 5’-C_{14}A_{15}G_{16}U_{17}G_{18}N-3’ where N can be A, C, or U (Figure 2.2). The hexalooop contains a C_{14}-G_{18} base pair across the loop which helps stabilize the loop structure, whereas G_{16} is flipped out and solvent exposed. This structure is supported by both NMR and chemical cleavage studies. In vitro evolution experiments have shown that the highest affinity loop structure for the IRP proteins is one where there is a hydrogen bond between the 14 and 18 positions, allowing for nucleotides 15-17 to be accessible for protein binding.

Additionally, mutational analysis has shown that nucleotide substitutions at either the C_{14}, G_{18} position, or at the C_{8} bulge of the stem, decreases IRP1 binding 36-99% depending on the specific mutation.

This loop region is flanked by an upper and lower stem region separated by a small bulged region containing an unpaired cytosine on the 5’ side. The upper stem region is a 5-bp long helix with variable sequence, and the lower base paired region does not contain sequence or length constraints. Point mutations that disrupt the upper stem helix were found to be non-functional whereas mutations that did not disrupt the double stranded nature of the stem were functional. Fully functional IREs are likely to have longer lower stems either to allow more extensive interactions of the RNA with the IRPs or to stabilize the conformation of the stem-loop. It has also been indicated that the stem region is predicted to be a potential target for small molecule drugs, allowing discrimination and selectivity between sequences.

IRE conservation for the same gene between different species is >95% while conservation between different genes in the same species is only 36-55%. The ferritin IRE has an internal loop in the stem region made up of U_{6}, G_{7}, C_{8}, and
The ferritin mRNA contains one copy of the IRE sequence in its 5’ untranslated region (UTR) that is used to regulate ferritin synthesis. The position of the IRE relative to the location of the 5’ UTR cap structure of the mRNA is important, with locations >60 nt from the cap leading to a reduction of translational control. In cases of low cellular iron levels, the IRP binds to the IRE preventing ferritin protein synthesis, and preventing the storage of iron needed elsewhere. Binding of the IRE by the IRP prevents the recruitment or scanning of the 40S ribosomal subunit to the ferritin mRNA interfering with translational initiation. In times of a high cellular iron level, the IRP is unable to bind to the IRE allowing ferritin protein synthesis to occur (Figure 2.3). This type of post-transcriptional regulation is also present for mammalian mitochondrial aconitase, erythroid amino levulinic acid synthase, (eALAS) which catalyzes the first step in heme synthesis, and the iron exporter ferroportin 1. There is also evidence of an IRE-like structure in the 5’ UTR of the Alzheimer’s amyloid precursor protein transcript.

The transferrin receptor is regulated in a slightly different fashion, although still by IREs. Five copies of the IRE are found in the 3’ UTR region of the transferrin receptor transcript and mediate mRNA stability thereby regulating receptor levels in the cell (Figure 2.4). In iron deficient cells, translation of the transferrin receptor is up-regulated to increase the amount of iron brought into the cell. This occurs by the binding of IRPs to the IRE in the 3’UTR, which protects the transcript from endonucleolytic cleavage and degradation by blocking the cleavage sites. In cases of high cellular iron, transferrin receptor synthesis is shut off by the inability of the IRPs to bind to the IRE, either through proteosomal cleavage
of IRP2 or iron-sulfur cluster formation in IRP1. This reduces the number of TfRs at the surface of the cell resulting in a decrease in the amount of iron brought into the cell. Regulation of DMT1, a divalent metal transporter involved in iron homeostasis also occurs via the mediation of mRNA stability through an IRE located in the 3’ UTR of its transcript.\textsuperscript{17}

The binding of IRP1 to the ferritin IRE has been extensively studied.\textsuperscript{37,38,51} Footprinting reactions, using both chemical and enzymatic nucleases, were performed on the consensus IRE (28-mer) as well as a 55-nt sequence containing the IRE and flanking regions.\textsuperscript{27} The IRP1 binding site was found to completely cover the IRE and this interaction resulted in an increase in the helical nature of the stem as shown by increased sensitivity to RNase V1.\textsuperscript{27,28,54} Mutagenesis of the IRE consensus sequence has revealed that disrupting the base pairing of the stem structure reduces protein binding 400-500 fold.\textsuperscript{9} Also, replacing G\textsubscript{18} in the hexaloop by an A, or deletion of C\textsubscript{8} in the internal bulge, completely shut off protein binding.\textsuperscript{3} More recently, a crystal structure of the rabbit Cys\textsuperscript{437}Ser/Cys\textsuperscript{503}Ser double-mutant IRP1 in complex with the bullfrog ferritin H IRE was solved, and has led to a more detailed understanding of the specific contacts made between the IRP1 protein and the IRE RNA.\textsuperscript{37,38}

2.2.3 Identification of Small Molecules that Disrupt the IRP:IRE Interaction

Previous work has shown that chemical and enzymatic footprinting techniques are able to identify the precise binding site of small molecules on an RNA target.\textsuperscript{55,64-68} In the body only about half of the available ferritin mRNA is activated.\textsuperscript{43} A strategy for increasing the pool of available ferritin is to block the binding of the IRP protein to the IRE RNA with a small molecule. The challenge is to selectively target the ferritin IRE isoform as opposed to the IRE isoforms found in other proteins involved in iron homeostasis such as the transferrin receptor. When compared with other IRE sequences the human ferritin IRE contains a
unique internal bulged region created by the unpaired U₆ and C₈ residues. Specific RNA:protein contacts are made in the areas of the hairpin loop and lower stem internal bulge that will allow a compound that selectively binds in this bulged region to disrupt the IRE-IRP complex, increasing the rate of ferritin synthesis in the cell.

The chemical cleavage agent Ru(tpy)(bpy)O²⁺ (RuO²⁺) has been used to find small molecules that bind specifically and selectively to the pocket created by the bulged region of the ferritin IRE. The structure of the IRE has been studied previously using RuO²⁺ oxidation chemistry. RuO²⁺ oxidizes solvent accessible guanine nucleobases that are prone to cation binding and its small size relative to enzymatic nucleases make it an excellent probe of RNA secondary structure. Four molecules, yohimbine, Hoechst 33285, acetopromazine, and promazine, were identified as binding to this bulge region, and experimental evidence showed that yohimbine was able to disrupt the human ferritin IRE-IRP1 complex causing the up-regulation of ferritin protein synthesis in an in vitro translation assay. The binding constant for the IRE hairpin:yohimbine interaction is 4 µM.

The goal of this research was to examine the ability of promazine, a small molecule identified by the RuO²⁺ footprinting technique, to block the IRP:IRE interaction and determine its effect on ferritin protein synthesis. The hypothesis, based on the ability of yohimbine to up-regulate the production of ferritin by binding in the internal bulge region of the IRE, was that binding in the internal bulge region by promazine would also disrupt the IRP:IRE interaction, allowing ferritin synthesis to be regulated.
2.3 Experimental

2.3.1 Materials

Proteinase K (20 mg / ml), phenol/chloroform, acid phenol/chloroform, linear acrylamide (5 µg/µl), DEPC-treated water, Superase-In RNase inhibitor (20 U / µl), T7 RNA Polymerase Plus™ (200 U / µl), RNase A (1 µg / ml), RNase T1 (1 U / µl), RNase I (100 U / µl), alkaline hydrolysis buffer, Gel Loading Buffer II, TURBO DNase I, MEGAscript™ T7 Kit, MEGAscript™ SP6 Kit, MEGAscript™ T7 Kit and MEGAclean™ transcription reaction purification kit were purchased from Applied Biosystems. Promazine HCl, chlorpromazine HCl, promethazine HCl, yohimbine HCl, heparin, desferoxamine, dithiothreitol (DTT) and CellLytic™ Y Plus kit (lyticase) were purchased from Sigma-Aldrich. T4 DNA ligase, mini Quick Spin RNA Columns, and Complete Protease Inhibitor Cocktail Tablets were purchased from Roche. All restriction enzymes, calf intestinal phosphatase (CIP), T4 polynucleotide kinase, T4 RNA ligase, and corresponding buffers were purchased from New England Biolabs. The pUC19K vector was a gift from Robert Nicholas of the Department of Pharmacology at UNC-CH. Novex® Tris-Glycine Gels, and NuPAGE® Novex® Bis-Tris Gels were purchased from Invitrogen. [5’-32P]pCp (0.01 mCi / µl), [α-32P]UTP (0.01 mCi / µl), [γ-32P]ATP (0.01 mCi / µl), and Easy Tag EXPRESS-[35S] Protein Labeling Mix (0.011 mCi / µl) were purchased from Perkin Elmer. Ultrafree-MC centrifugal filter devices were purchased from Millipore. Plasmid Maxi and Mini prep kits were purchased from Qiagen. Nonidet P-40® was purchased from USB Corporation. XL10-Gold® ultracompetent cells were purchased from Stratagene. The pGEM®-luc vector, and Rabbit Reticulocyte Lysate were purchased from Promega. Y-PER® Yeast Protein Extraction Reagent was purchased from Pierce Scientific. HiTrap Heparin HP affinity
column and HiTrap Q FF ion exchange column were purchased from GE Healthcare. Dye used to determine protein concentration was purchased from Bio Rad. The pCRII-TOPO vector containing the full-length ferritin mRNA sequence (human H-chain; 5’UTR is 268 bases and contains the IRE) was a gift from Dr. Elizabeth Theil of the Childrens Hospital Oakland Research Institute. IRP1 and IRP2 P. pastoris clones were a gift from Dr. Tracey A. Rouault from the National Institute of Child Health and Human Development at the National Institutes of Health in Bethesda, MD.

2.3.2 Methods

2.3.2.1 Transcription of IRE RNA (50-mer)

Creation of vectors used in this project was previously described. A stock of cells carrying the plasmid containing the IRE (50-mer) sequence (referred to here as hIRE) was streaked on LB/Kan agar plate. A single colony was picked and grown overnight in LB/Kan media (37°C, with shaking at 250 rpm). Plasmids were isolated and purified using a Qiagen Plasmid Maxi Kit according to the protocol provided by Qiagen. Final concentration and purity of plasmid DNA were determined by the absorbance at 260 nm and 280 nm. The plasmid was digested with Dra I (100 µl; in 1X NEB Buffer # 4, 20 µg hIRE plasmid, 200 U Dra I) for 2 h at 37°C, treated with 0.5% SDS and proteinase K (10 µg), for 2 h at 50°C, then purified by phenol/chloroform extraction and ethanol precipitation. To generate the IRE RNA a MEGAscript™ T7 Kit was used according to the protocol provided by Applied Biosystems, with the addition of 200 U T7 RNA Polymerase Plus™ to increase yields, and 20 U Superase·In to inhibit RNase degradation. DNA in the resulting sample was digested using 30 U TURBO DNase I at 37°C for 15 min. Production of the RNA transcript was verified on a 15% denaturing (8 M Urea) gel. The RNA was purified using a mini Quick
Spin RNA Column (following the protocol provided by Roche), followed by acid phenol/chloroform extraction, and ethanol precipitation. Absorbance at 260 nm was used to determine the final RNA concentration.

2.3.2.2 5’ RNA End-labeling

RNA (~14 µg) was first dephosphorylated at the 5’ end using 30 U calf intestinal phosphatase at 50 °C for 1 h. RNA was then run through a mini Quick Spin Column (following the protocol provided by Roche), acid phenol/chloroform extracted, and ethanol precipitated. RNA was 5’ end labeled with 50 – 80 µCi [γ-32P]ATP (0.01 mCi / µl) using 80 U T4 polynucleotide kinase at 37 °C for 1 h. The kinase was inactivated by heating the reaction at 65 °C for 20 minutes. Labeled RNA was then run through a mini Quick Spin Column, and ethanol precipitated with linear acrylamide (2.5 µg). RNA was purified on a small 10% (7 M Urea) polyacrylamide gel. RNA bands were visualized by autoradiography and excised from the gel. Gel slices were crushed and soaked in elution buffer (0.5 M NH₄OAc, 1 mM EDTA, 0.1% SDS) with Superase·In RNase inhibitor (20 U) in Utrafree-MC filters at 37 °C overnight. Samples were spun down at 4,000 x g and ethanol precipitated overnight with linear acrylamide (2.5 µg). The concentration of RNA was determined using the absorbance at 260 nm.

2.3.2.3 3’ RNA End-labeling

RNA (30 µg) was 3’ end labeled with 50 – 80 µCi [5’-32P]pCp using 60 U T4 RNA ligase overnight at 4 °C. RNA was run through a mini Quick Spin RNA column and further purified on a small 10% (7 M Urea) polyacrylamide gel. RNA bands were visualized by autoradiography and excised from the gel. Gel slices were crushed and soaked in elution buffer (0.5 M NH₄OAc, 1 mM EDTA, 0.1% SDS) with 20 U Superase·In RNase inhibitor in
Utrafree-MC filters at 37°C overnight. Samples were spun down at 4,000 x g and ethanol precipitated overnight with linear acrylamide (2.5 µg). The concentration of RNA recovered was determined using the absorbance at 260 nm.

2.3.2.4 Internal Labeling of RNA with [α-32P]UTP

RNA was transcribed according to the procedure described in Section 2.3.2.1 with the addition of 40 – 60 µCi [α-32P]UTP. Excess radiolabel was removed from the transcribed RNA using a mini Quick Spin RNA column, and ethanol precipitated overnight. RNA was further purified on a small 10% (7 M Urea) polyacrylamide gel. RNA bands were visualized by autoradiography and excised from the gel. Gel slices were crushed and soaked in elution buffer (0.5 M NH₄OAc, 1 mM EDTA, 0.1% SDS) with 20 U RNase inhibitor in Utrafree-MC filters at 37°C overnight. Samples were spun down at 4,000 x g and ethanol precipitated overnight with linear acrylamide (2.5 µg). The concentration of recovered RNA was determined using the absorbance at 260 nm.

2.3.2.5 Expression, Lysis, and Purification of IRP1 and IRP2

Expression of IRP1 and IRP2 proceeded via the method reported by Allerson et al., using methanol to activate the alcohol oxidase promoter to induce protein synthesis. After harvesting the yeast, the pellets were stored at -80°C until needed for purification. Yeast pelleted after IRP expression were thawed on ice and resuspended in 2 volumes (volume/weight) of breaking buffer (25 mM Tris-HCl, pH 8.0, 20 mM KCl, 1 mM EDTA, 0.1 M desferoxamine, 0.75 mM DTT, 1 tablet of Complete Protease Inhibitor). Lysis was performed to create spheroplasts using lyticase (1250 U / g cells) from the CellLytic™ Y Plus Kit by allowing the reaction to continue until the OD₈₀₀ nm reached 40% to 60 % of the pre-lysed sample reading. The spheroplasts were disrupted by adding Y-PER® Yeast Protein
Extraction Reagent (1875 µl/ 500 mg cells) and agitating at room temperature for 20 min. Cells were pelleted at 14,000 x g for 10 min, the lysate was collected and immediately subjected to chromatographic purification.

The purification proceeded in the general manner reported by Allerson et al. with the following modification; only two columns, heparin and Q-sepharose, were needed to obtain protein purity levels necessary for the subsequently performed experiments.74 Chromatographic buffers were: Buffer A-no salt (25 mM Tris-HCl, pH 8.0, 0.5 mM EDTA, 0.1 M desferoxamine, 1 mM DTT) and Buffer B-high salt (25 mM Tris-HCl, pH 8.0, 3 M NaCl, 0.5 mM EDTA, 0.1 M desferoxamine, 1 mM DTT). To decrease degradation by proteases, one protease inhibitor tablet was added per 25 ml buffer before use. Desferoxamine was used to chelate free iron, and all columns were run at 4°C. Elution of protein from each column was monitored by absorbance at 280 nm. Fractions were eluted on an ÄKTAPrime FPLC system (GE Healthcare) with typical flow rates of 4 ml/min. Crude lysates were first loaded onto a 2 x 5 ml HiTrap Heparin HP column. The column was washed with Buffer A, then eluted with a gradient to Buffer A with 50% Buffer B. IRP1 eluted at ~390 mM NaCl and IRP2 eluted at ~525 mM NaCl. The appropriate fractions (as determined by SDS-PAGE) were pooled, diluted to 40 mM NaCl with Buffer A, and loaded onto a 1 x 5 ml HiTrap Q Sepharose FF column. The column was washed with Buffer A with 1% Buffer B, and then eluted with a gradient to Buffer A with; 30% Buffer B for IRP1 and 50% Buffer B 1PR2. IRP1 eluted at ~ 450 mM NaCl and IRP2 at ~780 mM NaCl. The appropriate fractions were pooled, dialyzed two times against no salt buffer (25 mM Tris-HCl, pH 8.0, 0.5 mM EDTA), and concentrated using an Amicon Centricon® Plus 20 centrifugal filter according to the protocol provided by Millipore. Protein concentrations
were determined with a Bio-Rad protein assay, using BSA as a standard following the procedure provided by Bio-Rad. Samples of protein were stored as 50% glycerol stocks at -20°C. The concentration of the IRP1 stock was 1.1 mM, while the concentration of the IRP2 stock was 19 µM. Ability of IRP1 and IRP2 to bind to the IRE were evaluated using an Electrophoretic Mobility Shift Assay (EMSA).

2.3.2.6 Electrophoretic Mobility Shift Assay

IRE hairpin RNA internally labeled with [α-32P]UTP was transcribed according the protocol described in Section 2.3.2.1. A stock solution of labeled RNA (100 µl; 200 nM RNA, in 10 mM HEPES-KOH pH 7.2, 40 mM KCl, 20 U Superase·In) was folded by heating to 90°C for 5 min followed by slow cooling to room temperature. An aliquot of the stock solution (5 µl) containing labeled RNA (~50 nM), binding buffer (10 mM HEPES-KOH pH 7.5, 40 mM KCl, 0.3% nonidet P-40®, 5% glycerol, 3 mM MgCl2), and varying concentrations of either yohimbine or promazine (only for drug interaction studies) were incubated at room temperature for 20 min. Concurrently, IRP protein (IRP1 1 µM, IRP2 2 µM) was incubated with DTT (1 mM) for 20 min. The two solutions were then mixed, heparin (5 µg) was added to prevent any non-specific binding, and incubated for 20 min at room temperature. The reaction was terminated by adding 80% glycerol/bromophenol blue dye. The RNA-protein complexes were separated from unbound RNA on 6% native polyacrylamide gels (29:1 acrylamide/bisacrylamide) at 4°C in 1X TBE. Gels were dried and phosphorimaged overnight. Gels were imaged using a Storm® 860 imaging system. Band intensities were quantified using the ImageQuant™ software package (Molecular Dynamics/GE Healthcare).
2.3.2.7 RNase Cleavage Assay

Human ferritin IRE RNA (50 base hairpin) was transcribed using a MEGAscript T7 Kit\textsuperscript{TM} according to the procedure described in Section 2.3.2.1. RNA was 3'-end labeled with [5'-\textsuperscript{32}P]pCp using T4 RNA ligase overnight at 4\degree C as described in Section 2.3.2.3. Footprinting solutions were made by adding a trace amount of radiolabeled IRE RNA (~ 0.5 \mu M) to a stock solution (100 \mu l; containing 11 \mu M unlabeled RNA, in 80 mM Tris-Cl / 20 mM KCl, pH 7.4). The RNA solution was then folded by heating to 90\degree C for 5 min followed by slow cooling to room temperature. Aliquots of the stock solution (5 \mu l) were used in the RNase cleavage reactions. The footprinting solutions were then incubated with varying concentrations of either promazine HCl or yohimbine HCl for 20 min at room temperature (9 \mu l total volume). RNase I (0.001 U, 1 \mu l) was added and the reaction was allowed to proceed for 30 sec before being quenched by adding 5 \mu l Gel Loading Buffer II. Samples were separated on 20\% (7 M urea) polyacrylamide gels, exposed to a phosphorimaging screen overnight and analyzed using a Storm\textsuperscript{®} 860 phosphoimager. Band intensity was quantified using ImageQuant\textsuperscript{TM} software. RNA ladders were constructed using the 3’ labeled RNA via enzymatic cleavage by RNase TI, RNase A, and alkaline hydrolysis according to the protocol provided by Applied Biosystems.

2.3.2.8 Transcription of Full-Length mRNA: Human H-chain Ferritin, IRE-luciferase, and pGEM\textsuperscript{®}-luciferase

The human H-chain ferritin (FL-HF) plasmid encodes for the ferritin gene (1049 bases) with the IRE sequence in the 5’UTR (286 bases) in the pCRII-TOPO vector under the SP6 promoter. The wild-type plasmid was digested with Xho I (100 \mu l; containing 40 \mu g FL-HF plasmid, in 1X NEB Buffer #2, 0.1 \mu g BSA, 200 U Xho I) at 37\degree C for 2 h, followed
by treatment with 0.5 % SDS, 10 µg proteinase K for 2 h at 50°C, and purification using phenol/chloroform extraction and ethanol precipitation. Concentration of linearized vector was determined using the absorbance at 260 nm. Full length human H-chain ferritin mRNA was transcribed from the linearized vector using the SP6 promoter. Transcription was performed using a MEGAscript SP6 Kit™, for 3 h following the protocol provided by Applied Biosystems. Transcribed RNA was purified using a MEGAclean Kit™ (according to the protocol provided by Applied Biosystems) followed by acid phenol/chloroform extraction and ethanol precipitation. Production of the correct size mRNA was confirmed by electrophoresis on a Lonza precast 1.25% agarose RNA gel. RNA concentrations were determined using the absorbance at 260 nm.

Creation of the IRE-luciferase (IRE-luc) plasmid, containing a 1739 bp fragment coding for the luciferase protein with the human ferritin H-chain IRE sequence 18 bases upstream of the luciferase start codon in the pUK19K vector, was described in a previous work.55,74 IRE-luciferase mRNA was transcribed from the Xho I linearized vector (digested following the protocol described in the above paragraph) using the T7 promoter. Transcription was performed using a MEGAscript T7 Kit™ following the protocol provided by Applied Biosystems. Transcribed RNA was purified using a MEGAclean Kit™ (according to the protocol provided by Applied Biosystems) followed by acid phenol/chloroform extraction and ethanol precipitation. Production of the correct size mRNA was confirmed by electrophoresis on a Lonza precast 1.25% agarose RNA gel. RNA concentrations were determined using the absorbance at 260 nm.

The pGEM®-lac vector (wild type luciferase gene under the SP6 promoter) was digested with Xho I following the procedure outlined at the beginning of this section.
MEGAscript SP6 Kit™, for 3 h following the protocol provided by Applied Biosystems.

Transcribed RNA was purified using a MEGAclear Kit™ (according to the protocol provided by Applied Biosystems) followed by acid phenol/chloroform extraction and ethanol precipitation. Production of the correct size mRNA was confirmed by electrophoresis on a Lonza precast 1.25% agarose RNA gel. RNA concentrations were determined using the absorbance at 260 nm.

2.3.2.9 In Vitro Translation of mRNA: Rabbit Reticulocyte Lysate Translation Assay

Full length wild-type ferritin, wild-type luciferase, or IRE-luciferase mRNA was transcribed as indicated above. A stock solution of full length mRNA (1 µg/µl in 10 mM Tris-HCl, pH 8.0, 1 mM EDTA) was folded by heating the solution to 85 °C for 5 min, followed by slow cooling to room temperature. Translation of mRNA (1 µg) occurred after a 20 min pre-incubation (10 mM Tris-HCl, pH 8.0, 1 mM EDTA, 30 °C) with or without the small molecule for 20 min at 30°C using Rabbit Reticulocyte Lysate and [35S]methionine according to the protocol provided by Promega (final reaction volume 12.5 µl). Synthesized protein was analyzed by electrophoresis on 16% Tris-Glycine gels in Tris-Glycine SDS buffer (pH 8.3). The gels were fixed in a 50% methanol/ 10% acetic acid solution and exposed to a phosphorimaging screen overnight. Protein band intensities were quantitated based on [35S]methionine incorporation using ImageQuant™ software.

A competition assay with two mRNAs, wild-type ferritin and wild-type luciferase, was used to determine the specificity of promazine for the IRE in a more complex mixture. The procedure was identical to the above, except that 1µg of each RNA was preincubated in solution with promazine before the addition of lysate.
2.4 Results and Discussion

2.4.1 Small Molecule:Ferritin mRNA Interaction: RNase Cleavage Assay

Previous work has identified commercially available small molecules that bind near the internal bulge of the ferritin IRE using the chemical cleavage agent Ru(tpy)(bpy)O$_2^{2+}$.\textsuperscript{55} Ru(tpy)(bpy)O$_2^{2+}$ cleaves solvent accessible guanines in nucleic acids, and small molecule binding causes inhibition of this cleavage.\textsuperscript{55,68,75,76} Compounds identified were those that inhibited guanine oxidation near the internal bulge of the IRE structure, while guanine oxidation of the hairpin loop remained constant. Thirteen compounds were initially screened, and four of these, yohimbine, Hoechst 33258, promazine, and acetopromazine, showed favorable binding characteristics; selective inhibition of the cleavage at G23 and G27 guanines near the internal bulge of the IRE structure, with little effect on the cleavage at G16 found in the hairpin loop.\textsuperscript{55} The structures of these compounds are shown in Table 2.1.

Yohimbine and Hoechst 33258 were further screened for their ability to block the IRP:IRE interaction. Hoechst 33258 was shown to indiscriminately bind to all areas of the IRE RNA and was used only as a non-specific binder in future studies.\textsuperscript{55,74} The natural product yohimbine displayed ideal binding characteristics, binding only in the stem/bulge region of the IRE. Enzymatic footprinting with RNase I was used to confirm the results of Ru(tpy)(bpy)O$_2^{2+}$ footprinting (Figures 2.5 and 2.6). RNase I cleaves equally after all four nucleotides, exhibits reproducible cleavage patterns, and has been used previously to footprint RNA binding sites.\textsuperscript{65,66} As expected, both promazine and yohimbine showed an inhibition of cleavage in the area of nucleotides 21-25, with no change in cleavage intensity seen in the hairpin loop region (Figure 2.7). This finding confirms the viability of the Ru(tpy)(bpy)O$_2^{2+}$ assay as an RNA-ligand screen. Due to the variability in the sequence and
base-pairing surrounding the C8 found mid-helix, targeting drugs to this region provided an increased selectivity for the ferritin IRE compared to IREs found in other mRNAs.9,44,46

2.4.2 Effect of Promazine on the IRP:IRE Interaction: Electrophoretic Mobility Shift Assay

Protein-nucleic acid interactions are commonly studied using the observation that protein:nucleic acid complexes migrate more slowly in a native polyacrylamide gel than nucleic acids alone. By observing the retardation of a radiolabeled nucleic acid strand when it is bound to the much larger protein compared to the free nucleic acid one can determine if a protein or mixture of proteins is capable of binding to a given nucleic acid sequence. This experiment is known as an electrophoretic mobility shift assay (EMSA) and can also be referred to as a gel shift assay, gel mobility shift assay, band shift assay, or gel retardation assay. We used this type of assay to determine the effect of promazine on the IRP:IRE interaction using radiolabeled IRE RNA and purified IRP1 and IRP2 proteins. The IRP:IRE interaction has been previously studied using EMSA to elucidate the structure and sequence constraints for high affinity binding of IRP1 and IRP2 with the IRE.9,44,54,77,78 These studies have revealed that the determinants for binding are similar for the two proteins, but not exactly the same. Both proteins are able to bind the wild type IRE sequence with similar affinity5,44, but sequences were identified that specifically bind to either IRP1 or IRP2.

Here, gel shift analysis was used to detect the activity of the purified IRP1 and IRP2 proteins and the disruption of the IRP:IRE interaction by promazine and yohimbine. At constant IRP and IRE concentrations, increasing concentrations of small molecule were added, and the resulting IRP:IRE complexes were observed on a native polyacrylamide gel (Figures 2.8A, 2.9A, 2.10A and 2.11A). Quantitations of the bands revealed an approximate increase in free RNA of about 25% for promazine at 50 µM with IRP1 (Figure 2.8B), and
7% for IRP2 (Figure 2.9B). Yohimbine showed an increase in free RNA of 12% for IRP1 (Figure 2.10B) and 4% with IRP2 (Figure 2.11B) at a concentration of 50 µM. Variations in binding and quantitation result from impurities in the purified protein. Two conclusions can be drawn from these results. First, both promazine and yohimbine are able to displace a limited amount of IRP1 and IRP2 from the IRE. Second, there may be an additional population of a ternary complex formed involving IRE, IRP, and small molecule. This complex was first suggested by Tibodeau^74 to explain the positive translation assay results with saturating concentrations of yohimbine. These same positive translation assay results are seen in the presence of promazine (see section 2.4.3). This new evidence suggests that using a small molecule to disrupt the IRP:IRE interaction is more complex than first thought.

A crystal structure of the interaction between bullfrog ferritin IRE RNA and rabbit IRP1 shows that contacts occur between the protein and RNA at two separate sites. Direct protein:RNA contacts are made in the areas of the hexaloop and the RNA lower stem with protein binding centered around the C8 nucleotide. The conservation of the IRE structure and sequence across species allows us to use this crystal structure to explain our experimental results. The binding of the IRE through these two separate binding sites contributes to the specificity of the IRP:IRE interaction. It is also shown that these two areas contain the bases that have been shown to interact specifically with yohimbine and promazine in the human ferritin IRE RNA. Disrupting the binding at either site with a small molecule would cause a weakened IRP:IRE interaction that would allow the ribosome to scan through this area and increase the amount of ferritin biosynthesis. If the structural change due to this binding was not large enough to displace all binding, it would be possible for ferritin synthesis to occur, albeit at a decreased rate.
2.4.3 Regulation of Protein Synthesis by Small Molecules: In Vitro Translation Assay

2.4.3.1 Ferritin Translation

IRP proteins were first identified as repressors of ferritin synthesis when it was noted that ferritin mRNAs translate very poorly in rabbit reticulocyte lysates (RRL) relative to other RNAs, and that this deficiency was not seen in wheat germ lysates. A repressor protein, now known as IRP1, that bound the IRE was purified from the reticulocyte lysate. Since then, reticulocyte lysates have been used to assess translation of ferritin and other iron-related transcripts in the presence of IRPs. We wanted to use this system to determine whether promazine could retain its IRE-binding ability in a more complex in vitro translation mixture, by competing for binding of the IRP.

A time-course translation assay was performed to ensure that ferritin synthesis rates were linearly dependent on incubation length at the chosen time (20 min). In addition to the wild-type ferritin mRNA, the effect of promazine on the regulation of two other sequences were studied: luciferase mRNA which contains no regulatory sequence, and IRE-luc, containing the human H-chain ferritin IRE regulatory element 18 bases upstream of the luciferase gene. Representative SDS-PAGE gels for the three translations are shown in Figure 2.12. Promazine, like yohimbine, was able to increase the rate of ferritin synthesis in the in vitro translation system. The wild-type ferritin mRNA translation was increased by 66% in the presence of promazine compared to 40% in the presence of yohimbine. The positive up-regulation of ferritin synthesis by promazine occurs at lower concentrations than yohimbine, indicating tighter binding of promazine to the IRE RNA, which is consistent with the gel mobility shift assay and RNase footprinting experiments. At high concentrations of
promazine, a decrease in overall translation occurs for all transcripts tested. This is most likely due to promazine interacting with the translational machinery in the RRL.

2.4.3.2 IRE-luciferase and Wild-type Luciferase Translation

To determine if the results observed with the increase of ferritin synthesis by promazine were dependent on the identity of the transcript, mRNA encoding luciferase with the IRE upstream of the start codon was used. This construct is still able to bind to IRPs indicating that its synthesis should be modulated by promazine binding. After a time-course assay to determine the linear region for IRE-luciferase translation, the assay was repeated in the presence of promazine. Translation in the RRL showed that the promazine interaction with the IRE-luciferase RNA caused an increase in synthesis of the luciferase protein by about 70% over baseline levels. (Figure 2.13) In contrast, there was no significant change in the amount of wild-type luciferase translated with the addition of promazine, even at high concentrations. These results were very similar to the ferritin results where a 66% increase was seen. The reported value for the yohimbine interaction is an increase of 40%. This result, with the earlier ferritin findings, indicates that the observed up-regulation in the presence of promazine does not depend on the identity of the mRNA, but instead is a direct result of the IRE-small molecule interaction. It also supports the hypothesis that promazine binds to the IRE tighter than yohimbine.

2.4.3.3 Ferritin and Wild-type Luciferase Competition Assay

Once it was established that promazine was able to increase the synthesis of protein in the presence of an IRE, we wanted to determine its ability to specifically increase production of one protein in a mixture of mRNAs. To test this we used the same RRL assay, with a modified procedure: the promazine was pre-incubated with both ferritin mRNA containing
Determination of another measurement of selectivity, the binding constant \( (K_d) \) for the promazine-RNA interaction was attempted by various methods. Our initial attempts included: luminescence spectroscopy, surface plasmon resonance (SPR), fluorescence polarization anisotropy, and absorbance spectroscopy. Binding constant data can also be determined using electrophoretic mobility shift assays. However, due to the impurities in the purified IRP proteins, a reasonable binding constant could not be determined from this data. A binding constant range was determined by fitting a 1:1 binding isotherm (Equation 1) to the RNase I cleavage data.

\[
Y = A \frac{1}{([RNA]/Kd) + 1}
\]

From the data a two binding event model was predicted, with the first binding event occurring around 30 nM, and the second around 3.5 \( \mu \)M. Because these values are derived from RNase cleavage data, it can not be determined if the first binding constant represents an
actual binding event or a structural rearrangement of the RNA due to an interaction of the
positively charged molecule with the negatively charged phosphate backbone. It is also
possible that there is more than one molecule of promazine binding to the IRE at higher
concentrations. However, evidence for this was not seen in any of the completed footprinting
experiments. The binding constant for yohimbine with the ferritin IRE 50-mer was
determined to be 4 μM. The results of the in vitro translation assays where ferritin
synthesis is up-regulated at lower concentrations of promazine compared to yohimbine
supports the tighter binding of promazine and the determined binding constant range. Many
small molecules that bind to RNA have been shown to interact in this low micromolar
range. These include RNAs selected via SELEX (systematic evolution of ligand by
exponential enrichment) for a specified interaction, and the TAR RNA element found in
HIV-1. Promazine has been shown to interact with TAR RNA in a bulged site that is
important for Tat protein binding with a binding constant of > 5 mM.

2.5 Conclusions

Two main conclusions can be drawn from these experiments. First, the use of the
Ru(tpy)(bpy)O²⁺ footprinting method as an assay for small molecule-RNA binding was
validated through the use of an RNase cleavage method. The area identified as interacting
with the small molecules by Ru(tpy)(bpy)O²⁺ oxidation also showed reduced cleavage in the
RNase method. Second, binding of small molecules to specific three-dimensional structures
in mRNA can alter protein synthesis. The modulation of gene expression through the
targeting of mRNA with small molecules has previously been shown, but all prior
molecules have resulted in a down-regulation of protein synthesis. To our knowledge the up-
regulation of ferritin synthesis by the addition of promazine and yohimbine is the first reported case of enhancement of gene expression by small molecules. However, this up-regulation is likely to be seen where a protein:RNA interaction, such as the IRP:IRE interaction, blocks formation of the translation initiation complex. Disruption of the complex would allow ribosome scanning and protein synthesis.

These studies indicate that promazine binds to the ferritin IRE selectively, and this binding event up-regulates the synthesis of ferritin through the disruption the IRP:IRE interaction. There are three main possibilities for the origin of this result. First, the free RNA generated by promazine blocking the IRP:IRE interaction is responsible for the increase in translation. Second, a ternary complex formed between promazine, IRP, and RNA, can be translated, unlike the IRP:IRE complex. Third, the increased synthesis seen could be a combination of these two events. It would be difficult to ascertain the absolute level of synthesis occurring in the lysate that would be needed to sort out which of these three possibilities is occurring. Additionally, in a more complex mixture of two RNAs, one containing a regulatory element and one without an element, addition of promazine does not detectably increase the rate of ferritin synthesis at any concentration, suggesting that considerably greater selectivity would be required to achieve up-regulation of gene expression \textit{in vivo}. Nonetheless, these results point to the feasibility of specific RNA-small molecule interactions as modalities for up-regulation.

There are two potential means by which a small molecule binding to RNA could disrupt the protein-RNA interactions. Binding could occur in the RNA minor groove and block the ability of the proteins to make specific contacts, or binding could occur in an intercalative manner, causing a disruption in base pairing or stacking, -- resulting in a change
in secondary or tertiary structure. Promazine binding to HIV-1 TAR RNA, which contains a structural element similar to the IRE hairpin, has been studied by the James group. They identified the binding of promazine to possess some intercalative qualities, although the non-planar structure of phenothiazine prevents binding solely by intercalation. Binding through a semi-intercalative manner would cause distortion in the helical stem area of the IRE and disrupt its secondary structure. This could prevent contacts from being formed correctly between the IRP protein and IRE RNA suggested by the crystal structure. It is likely a combination of both the disruption of essential protein contacts and the distortion of the IRE structure by promazine and yohimbine that causes the increase protein synthesis levels shown in this work.

More work is needed in the drug discovery aspect of this project. Although two promising small molecules were identified from the Ru(tpy)(bpy)O$_2^+$ footprinting screen, these compounds were not selective in a more complex system. With the advances in computer aided drug discovery, this method could be used to aid in the search for a small molecule with favorable binding characteristics. With a crystal structure of the IRP:IRE interaction now available, the specific contacts that need to be inhibited have been identified. Using the available NMR structure of the IRE RNA a database of available chemicals could be screened for binding to the three-dimensional RNA structure. Once identified, the small molecules could then undergo a footprinting method to validate their ability to bind in the desired region.

A second future direction of this project would be to test a phenothiazine library as well as derivatives of yohimbine to determine modifications to the small molecule that would
enhance the selectivity and specificity of binding. Small regulatory RNAs (small interfering RNAs and microRNAs) could also be evaluated for their ability to regulate translation.
### 2.6 Tables

Table 2.1. Names and structures of compounds found to interact with the IRE through the RuO$_2^{2+}$ assay.

<table>
<thead>
<tr>
<th>Compound Name</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yohimbine</td>
<td><img src="image1" alt="Yohimbine Structure" /></td>
</tr>
<tr>
<td>Hoechst 33258</td>
<td><img src="image2" alt="Hoechst 33258 Structure" /></td>
</tr>
<tr>
<td>Promazine</td>
<td><img src="image3" alt="Promazine Structure" /></td>
</tr>
<tr>
<td>Acetopromazine</td>
<td><img src="image4" alt="Acetopromazine Structure" /></td>
</tr>
</tbody>
</table>
Figure 2.1. Cellular iron transport: the transferrin cycle. Schematic showing the receptor-mediated endocytosis pathway used for cellular iron transport. Holo-Tf binds to the transferrin receptor. Receptors are brought into the cell via clathrin-coated pits. Once inside the cell specialized endosomes form and become acidified through proton pumps. The pH change causes a conformational change in the transferrin protein and iron is released. Iron is exported out of the endosome by DMT-1. Iron that is not immediately used by the cell is stored in ferritin. Subsequently, apo-transferrin and the transferrin receptor are returned to the cell surface where they dissociate at neutral pH. The recycled apo-transferrin and transferrin receptor can be involved in numerous cycles of iron transport. Adapted from Andrews.15
Figure 2.2. Sequence and structures of the human ferritin iron responsive element (IRE).  

A) Secondary structure of the human ferritin IRE predicted by mfold.92,93 The base pair between C14 and G18 is indicated.  
B) NMR structure of the human ferritin IRE.33 The important residues in the structure are indicated. Blue residues are those found in the hairpin loop, while residues in green are those found in the internal bulge. (PDB accession number 1AQO)  
C) Sequence of the human ferritin IRE (50-mer) used in this study. Bases in bold are found in the hexalooop.
Figure 2.3. Schematic representation of ferritin translation regulation by IRP binding. In iron-depleted cells, IRP binding to the IRE in the 5’ UTR interferes with translation initiation. In iron-replete cells the IRP does not bind the IRE and the ribosome can translate the mRNA. This type of regulation is also seen in mammalian mitochondrial aconitase, and erythroid amino levulinic acid synthase. Adapted from Rouault. 33
Figure 2.4. Schematic representation of transferrin receptor translation regulation by IRP binding. In iron-depleted cells, IRP binding to the IREs in the 3’ UTR protects the mRNA from endonucleolytic cleavage, increasing mRNA stability. Protein can then be translated. In iron-replete cells the IRP does not bind the IRE leaving the mRNA susceptible to nuclease cleavage. This type of regulation is also seen in DMT-1. Adapted from Rouault.33
Figure 2.5. Effect of promazine on RNA cleavage by RNase I in human IRE RNA. Lane 1, alkaline hydrolysis ladder; lane 2, a C/U ladder generated using RNase A; lane 3, a G ladder generated using RNase T1; lane 4, RNA alone. Lanes 5-21, [RNA] = 5.6 μM, [RNase I] = 0.001 U. Lanes 5-21 contain promazine concentrations (μM) as follows: lane 5, 0; lane 6, 0.0025; lane 7, 0.005; lane 8, 0.01; lane 9, 0.025; lane 10, 0.05; lane 11, 0.1; lane 12, 0.25; lane 13, 0.5, lane 14, 1.0; lane 15, 5.0; lane 16, 10.0; lane 17, 25.0; lane 18, 50.0; lane 19, 75.0; lane 20, 100.0; lane 21, 2000.0.
Figure 2.6. Effect of yohimbine on RNA cleavage by RNase I in human IRE RNA. Lane 1, alkaline hydrolysis ladder; lane 2, a C/U ladder generated using RNase A; lane 3, RNA alone. Lanes 5-20, [RNA] = 5.6 µM, [RNase I] = 0.001 U. Lanes 5-21 contain yohimbine concentrations (µM) as follows: lane 5, 0; lane 6, 0.0025; lane 7, 0.005; lane 8, 0.01; lane 9, 0.025; lane 10, 0.05; lane 11, 0.1; lane 12, 0.25; lane 13, 0.5; lane 14, 1.0; lane 15, 5.0; lane 16, 10.0; lane 17, 25.0; lane 18, 50.0; lane 19, 75.0; lane 20, 100.0; lane 21, 2000.0.
Figure 2.7. Effect of promazine (A) and yohimbine (B) on RNase I cleavage of the stem region of human IRE RNA. Decrease in cleavage at higher small molecule concentrations indicates binding in the stem region. Quantitations are from one experimental trial using ImageQuant™ software and are normalized to the C<sub>14</sub> residue. Legend values on the right indicate final concentration of small molecule in the assay. Data represents one cleavage experiment.
Figure 2.8. Effect of promazine on the IRP1:IRE interaction. A) Electrophoretic mobility shift assay. [RNA] = 50 nM. [IRP1] = 1 µM. B) Quantitation of $[^{32}\text{P}]$ RNA bands. Legend values indicate the final small molecule concentration in the assay. An overall increase in unbound RNA is seen with increasing concentrations of promazine. Data represents one trial.
Figure 2.9. Effect of promazine on the IRP2:IRE interaction. A) Electrophoretic mobility shift assay. [RNA] = 50 nM. [IRP2] = 2 µM. B) Quantitation of [³²P] RNA bands. Legend values indicate the final small molecule concentration in the assay. An overall increase in unbound RNA is seen with increasing concentrations of promazine. Data represents one trial.
Figure 2.10. Effect of yohimbine on the IRP1:IRE interaction. A) Electrophoretic mobility shift assay. [RNA] = 50 nM. [IRP1] = 1 µM. B) Quantitation of [32P] RNA bands. Legend values indicate the final small molecule concentration in the assay. An overall increase in free RNA is seen with increasing concentrations of yohimbine. Data represents one trial.
Figure 2.11. Effect of yohimbine on the IRP2:IRE interaction. A) Electrophoretic mobility shift assay. [RNA]= 50 nM, [IRP1]= 1 µM. B) Quantitation of [32P] RNA bands. Legend values indicate the final small molecule concentration in the assay. An overall increase in free RNA is seen with increasing concentrations of yohimbine. Data represents one trial.
Figure 2.12. *In vitro* translation assay to determine the effect of promazine on protein translation. [*35*S]methionine SDS-PAGE gel depicting translation of (A) wild-type ferritin mRNA, (B) IRE-luciferase mRNA, and (C) wild-type luciferase mRNA. [mRNA] = 1 μg. [promazine]: Lane 1, 0 μM; Lane 2, 0.5 μM; Lane 3, 1 μM; Lane 4, 10 μM; Lane 5, 20 μM; Lane 6, 40 μM; Lane 7, 80 μM.
Figure 2.13. Effect of promazine on IRE-dependent mRNA translation in cell-free extracts as measured by incorporation of $[^{35}S]$methionine. Data are representative of at least three different trials on at least two separate preparations of RNA. These results are expressed as a percentage of the control (no promazine added, normalized to 100%), and the error bars represent standard deviation.
Figure 2.14. Effect of promazine on the translation of a mixture of ferritin and luciferase mRNA in cell-free extracts as measured by incorporation of $[^{35}\text{S}]$methionine. The reactions contained 1 µg each of ferritin and luciferase mRNA. Data are representative of at least three different trials on at least two separate preparations of RNA. These results are expressed as a percentage of the control (no promazine added, normalized to 100%), and the error bars represent standard deviation.
2.8 References


(43) Theil, E. C.; Eisenstein, R. S. J. Biol. Chem. 2000, 275, (52), 40659-40662.


(49) Laing, L. G.; Hall, K. B. Biochemistry 1996, 35, (42), 13586-13596.


Chapter 3

Identification of Small Molecules that Target RNA and Mediate Human Preproinsulin Protein Synthesis

3.1 Abstract

Insulin resistance and deficiency in insulin secretion are the main causes of diabetes mellitus, making the ability to regulate insulin synthesis an area of great interest. Two forms of the 5’ untranslated region (UTR) of the human preproinsulin mRNA exist, a native form (NAT), and a splice variant form (SPV) that is up-regulated in response to chronic glucose stimulation. To investigate the ability of small molecules to modulate preproinsulin synthesis, small molecules that bind to 5’UTR of the human preproinsulin mRNAs were identified, and their effect on protein synthesis was determined. The secondary structure of the NAT and SPV 5’UTR was examined using RNase cleavage, and the transition metal complexes Ru(tpy)(bpy)O²⁺ (RuO²⁺, bpy = 2,2’-bipyridine, tpy = 2,2’,2”-terpyridine), and Ru(bpy)₃²⁺. RuO²⁺ and Ru(bpy)₃²⁺ are chemical nucleases that oxidize guanine bases in RNA sequences. The ability of small molecules to bind to, or change the secondary structure of, the 5’UTR of the preproinsulin mRNAs was investigated using the same techniques. Two small molecules, neomycin B and kanamycin B, have been identified as binding to the mRNAs using these screening methods. An in vitro translation system showed that these small molecules decreased the amount of preproinsulin protein synthesized for both the native and splice variant forms of mRNA.
3.2 Introduction

3.2.1 Regulation of Glucose Homeostasis

Glucose is required as a metabolic fuel by the brain under physiologic conditions while other organs use fatty acids as well as glucose to generate energy. Plasma glucose concentrations are maintained within a narrow range; the production, uptake and utilization is regulated by an intricate network of hormones, neural pathways and metabolic signals. Insulin, a peptide hormone, is produced in the \( \beta \)-cells of the pancreatic islets in response to glucose and is the key regulator of glucose homeostasis.

The body’s major source of glucose is the diet. In the fasting state, plasma glucose levels are maintained primarily by the breakdown of glycogen in the liver, and the production of glucose in the liver and kidneys.\(^1\) Low insulin levels cause decreased glycogen synthesis, reduce glucose uptake into insulin-sensitive tissues, and promote mobilization of stored precursors. Post-prandially, glucose elicits a rise in insulin secretion from \( \beta \)-cells, reversing the processes identified above. Insulin enhances cellular glucose uptake by increasing the number of glucose transporters in the plasma membrane of all cells.\(^1\) Once inside the cell, the transported glucose is either used as metabolic fuel or promotes energy storage by conversion of cellular substrates into glycogen, triglycerides, and proteins. Abnormalities in insulin secretion lead to misregulation of plasma glucose levels resulting in the disease diabetes mellitus.\(^2,3\)

There are two broad categories of diabetes mellitus, type 1 and type 2.\(^4\) Type 1 diabetes results in the lack of, or severe reduction in insulin secretion due to autoimmune destruction of pancreatic \( \beta \)-cells.\(^4\) Current treatment by administration of insulin does not prevent long-term complications including failure of major organs.\(^3\) Type 2 diabetes is the
more prevalent form in society. It is a multifactorial polygenic disease, though the mechanism is currently not understood.\textsuperscript{2,4,5} The pathogenesis involves progressive development of insulin resistance, a relative deficiency in insulin secretion, and increased glucose production.\textsuperscript{4} Much work is being put into understanding insulin regulation and its effects on the body.

3.2.2 Insulin

3.2.2.1 Human Insulin Biosynthesis

Expression of the insulin gene by pancreatic $\beta$-cells is mediated by a number of circulating nutrients and hormones, of which glucose is the predominant physiological stimulus.\textsuperscript{6} Mature human insulin is a 6 kDa peptide-hormone produced in the $\beta$-cells of the pancreatic Islets of Langerhans.\textsuperscript{7-10} It is made up of two polypeptide chains, the A- (21 amino acids) and the B-chain (30 amino acids) connected by two disulfide bonds (residues A7 to B7, and A20 to B19)\textsuperscript{8,9} (Figure 3.1). The A-chain also contains an internal disulfide bridge (residues A6 to A11)\textsuperscript{8,9}. The positions of these disulfide bonds do not vary across mammalian species.\textsuperscript{9}

Insulin is initially synthesized as a single-chain, 86-amino-acid precursor polypeptide, known as preproinsulin, which undergoes subsequent proteolytic processing to create the hormone insulin (Figure 3.2).\textsuperscript{4,11} Preproinsulin is first converted to proinsulin, and later to mature insulin by proteases. Preproinsulin is a 12 kDa protein that contains a 24 residue N-terminal signal peptide, characteristic of proteins that enter the secretory pathway.\textsuperscript{12-14} The signal sequence facilitates transit into the lumen of the rough endoplasmic reticulum (RER).\textsuperscript{15} This signal peptide is cleaved cotranslationally in the RER, generating proinsulin.\textsuperscript{16} Proinsulin is a 9 kDa protein that contains the A- and B- chain peptides of insulin, joined by
a connecting C-peptide (31 amino acids). The C-peptide is thought to aid in the folding of the molecule, allowing for the correct formation of disulfide bridges. It is also packaged in the vesicles and co-secreted with mature insulin. There is considerable variation in the sequence of the C-peptide between species, and its exact function is unknown. After being produced in the RER, proinsulin is transferred into the cis-Golgi, where it is sorted along with its processing enzymes and other components through the trans-Golgi into clathrin coated secretion granules. Maturation of this granule involves the removal of the proinsulin C-peptide by prohormone convertase (PC) 1/3 and PC2 as well as carboxypeptidase E/H. PC1/3 and PC2 are trypsin-like endoproteases that cleave at dibasic sites. PC2 shows preference for cleavage at the A-/C-chain junction while PC1/3 cleaves at both sites with a preference for the B-/C-chain junction. Carboxypeptidase E/H is an exopeptidase that removes C-terminal basic residues after cleavage by PC2 and PC1/3. The mature secretory granules are stored in β-cells until their fusion with the plasma membrane is stimulated by secretory signals.

3.2.2.2 Insulin Secretion from Pancreatic β-cells

Under normal conditions, only a small portion of stored insulin is released after each meal, even under maximum stimulation. This allows for the replenishment of the stored pool of insulin and maintenance of constant insulin content in β-cells. Insulin secretion is regulated primarily by changes in blood glucose levels. Uptake and metabolism of glucose in β-cells generates intracellular signals for insulin secretion and activates the electrical activity of β-cells. Glucose regulates the β-cell electrical activity through changes in the activity of ATP-sensitive K⁺ (K⁺ATP) channels. Glucose is brought into the β-cell through the glucose transporter GLUT2, which allows entry via facilitated diffusion, and a rate of
influx proportional to its extracellular concentration. Once inside the cell, glucose is phosphorylated to glucose-6-phosphate by glucokinase. Further metabolism of glucose leads to an increase in the ATP/ADP ratio causing closure of $K^+_{\text{ATP}}$ channels.\textsuperscript{24} Closure of these channels leads to a depolarization of the plasma membrane, and initiation of $\beta$-cell electrical activity leading to the opening of voltage gated $Ca^{2+}$ channels.\textsuperscript{10,25} The influx of $Ca^{2+}$ ions results in an increased intracellular concentration that triggers fusion of the vesicles containing the stored insulin pools to the plasma membrane. Although other hormones are able to increase the intracellular $Ca^{2+}$ concentration, they do not induce insulin secretion in the absence of glucose.\textsuperscript{3} Insulin secretory profiles reveal a pulsed pattern of hormone release, with small bursts occurring about every 10 min, with larger amplitude oscillations occurring every 80 to 150 min.\textsuperscript{4} Meals or other major stimuli of insulin secretion induce large (four- to five-fold increases over baseline) bursts that usually last 2 to 3 h before returning to baseline.

3.2.2.3 Mechanism of Insulin Action in the Body

Once insulin is released into the venous system ~50\% is degraded by the liver.\textsuperscript{4} The remaining insulin binds to insulin receptors in target sites. The insulin receptor is a heterotetrameric protein consisting of two extracellular $\alpha$ subunits and two transmembrane $\beta$ subunits.\textsuperscript{26-28} The $\beta$ subunits contain an intrinsic tyrosine kinase activity that is stimulated by binding of insulin to the $\alpha$ subunit. Upon insulin binding, the receptors autophosphorylate and activate their kinase activity, allowing binding of downstream signaling molecules.\textsuperscript{29-32} These molecules are involved in a variety of signaling cascades that are involved in gene expression, glycogen and protein synthesis, and lipogenesis.\textsuperscript{4} An example of one of the pathways activated is the phosphatidylinositol-3'-kinase (PI-3-kinase) pathway that
stimulates the translocation of glucose transporters such as GLUT4 to the cell surface. These transporters move circulating glucose from the plasma into muscle, hepatic and adipose tissues. Attenuated GLUT4 translocation and glucose uptake by muscle and fat cells following insulin stimulation represent a prime defect observed in insulin resistance. The mechanism underlying insulin-mediated regulation of glucose concentration is dependent on glucose itself regulating insulin biosynthesis and secretion.

### 3.2.3 Control of Insulin Gene Expression by Glucose

In humans, a single insulin gene is located on chromosome 11 and was first sequenced by Bell and coworkers in 1980. The gene consists of three exons and two introns, with the coding region starting in exon 2. In this manner, exon 1 and part of exon 2 comprise the 5’UTR of human insulin. Expression of the insulin gene in adult humans is restricted to the islet beta-cell, and is regulated at both the transcriptional and translational levels. In the short term (< 2 h), glucose-induced proinsulin synthesis is regulated at the translational level. During longer periods of glucose stimulation (> 6 h), there is an additional element of transcriptional control and an increase in mRNA stability.

Positive and negative regulation of insulin gene transcription relies on the interaction of sequence motifs in the promoter with islet cell specific transcription factors. There are a number of cis-acting DNA elements and trans-acting factors involved in regulating insulin gene transcription including glucose. Inducible transcription most likely depends on sequences in the promoter and/or enhancer region that control the cell-type specificity of the insulin gene. It has been suggested that factors binding to sites in this region are modulated by glucose, although the mechanism is not well understood. Increasing glucose
concentration 10-fold has been shown to increase the level of preproinsulin mRNA in cells 2.3-fold.\textsuperscript{42}

Glucose-mediated insulin protein synthesis is predominately regulated through control of preproinsulin translation at the post-transcriptional level, however the mechanism is not well understood. It is known that insulin production parallels insulin secretion to maintain intracellular stores.\textsuperscript{48} Glucose stimulates general protein synthesis in \( \beta \)-cells approximately 2-fold\textsuperscript{49}, while glucose-stimulated proinsulin synthesis can be increased up to 10-fold.\textsuperscript{39,45,50} Glucose exerts its effects largely through mechanisms requiring specific sequences in the UTRs of preproinsulin mRNA, especially in the 5’UTR.\textsuperscript{13,14,20} Examples of these processes are the up-regulation of general translation initiation factor activity, the rate of translational elongation, and signal recognition particle effects.\textsuperscript{50-52} The preproinsulin 5’UTR is predicted to form a stem loop structure similar to the 5’UTR of the ferritin mRNA.\textsuperscript{53}

The 5’UTR of the human and mouse insulin genes contain a secondary structural element which can exist in two forms: a native preproinsulin mRNA form (NAT), and a preproinsulin splice variant (SPV) (Figure 3.3).\textsuperscript{54,55} The splice variant mRNA has shown increased translational efficiency over the native constructs in both human and mouse islet cells.\textsuperscript{54,55} In humans it was shown that the alternatively spliced mRNA was induced 10-fold in times of chronic glucose stimulation. Recently, an 8 nt region of the native preproinsulin 5’UTR, called the preproinsulin mRNA glucose-responsive translation element (ppIGE, glucose regulatory element), was found to specifically bind an unidentified factor present in the islets of Langerhans (referred to here as ppIGE regulatory factor).\textsuperscript{37} This interaction has been shown to be a key component of the glucose-mediated regulation of insulin production.
While the exact binding effect of this factor is unknown, it is believed that binding has an inhibitory effect upon translation, similar to what is seen in the human ferritin system discussed in Chapter 2. When this 8 nt region is disrupted either by mutagenesis or as in the SPV mRNA, glucose stimulated proinsulin biosynthesis was abolished. Production of this SPV mRNA is up-regulated only in times of chronic glucose stimulation, such as hyperglycemia. This may indicate that the normal glucose-regulated proinsulin synthesis could be circumvented in times of high insulin need. Identifying a small molecule that could bind to this region in the NAT mRNA would help determine whether binding of the ppIGE binding factor stimulates or inhibits insulin biosynthesis.

3.2.4 Identification of Small Molecules that Bind to Human Preproinsulin mRNA

Chemical and enzymatic footprinting methods can be used to identify precise binding sites of small molecules on an RNA target. Previously a RuO\(^{2+}\) footprinting method was used to identify small molecules that bound specifically and selectively to the iron responsive element (IRE) in human ferritin mRNA. Another nucleic acid oxidation method, the flash-quench reaction, has also been developed and its use as a chemical footprinting method is currently being investigated. To date, there have been no known structural studies of either the NAT or SPV mRNA secondary elements, only computational structure predictions using mfold.

The goal of this research was to use chemical and enzymatic footprinting methods to study the secondary structure of both the native and splice variant forms of human preproinsulin mRNA. With this knowledge, these techniques could then be used to identify functional small molecule-RNA interactions. The hypothesis, based upon the previous work on the ferritin IRE system, was that a small molecule that has an effect on the secondary
structure of the human proinsulin mRNA would be able to modulate proinsulin translation in vitro.

3.3 Experimental

3.3.1 Materials

Proteinase K (20 mg/ml), phenol/chloroform, acid phenol/chloroform, linear acrylamide, DEPC-treated water, Superase-In RNase inhibitor (20 U/µl), T7 RNA Polymerase Plus™ (200 U/µl), RNase A (1 µg/ml), RNase T1 (1 U/µl), RNase I (100 U/µl), alkaline hydrolysis buffer, TURBO DNase I, Gel Loading Buffer II, MEGAscript™ T7 Kit, MEGAscript™ T7 Kit and MEGAClear™ transcription reaction purification kit were purchased from Applied Biosystems. Promazine HCl, carbenicillin disodium salt, yohimbine HCl, dithiothreitol (DTT), neomycin B trisulfate hydrate, kanamycin B sulfate, hygromycin B, and Hoechst 33258 were purchased from Sigma-Aldrich. Calf Intestinal Phosphatase (CIP), T4 DNA ligase and mini Quick Spin RNA Columns were purchased from Roche. All restriction enzymes, T4 polynucleotide kinase, Klenow, T4 RNA ligase, and corresponding buffers were purchased from New England Biolabs. Wheat germ extract and Rabbit Reticulocyte Lysate were purchased from Promega. ChargeSwitch® PCR Clean-up Kit, Novex® Tricine-SDS running buffer, DryEase® Mini-Gel Drying Kit and Novex® Tricine Gels were purchased from Invitrogen. [5',-32P]pCp (0.01 mCi/µl) and Easy Tag EXPRESS-[^35S] Protein Labeling Mix (0.011 mCi/µl) were purchased from Perkin Elmer. Ultrafree-MC centrifugal filter devices were purchased from Millipore. Plasmid purification Maxi- and Miniprep kits were purchased from Qiagen. XL10-Gold® ultracompetent cells, QuickChange™ II XL Site-Directed Mutagenesis Kit, pBluescript II SK(-), Herculase® II
Fusion DNA Polymerase, and PfuUltra® High-Fidelity DNA polymerase were purchased from Stratagene. Reliant® FastLane® Pre-cast agarose gels, SeaPlaque® low-melting agarose and NuSieve® GTG® agarose were purchased from Lonza. Polyspring glass inserts used in the flash-quench procedure were purchased from National Scientific. PCR primers were purchased from the Nucleic Acid Core Facility at the UNC Lineburger Comprehensive Cancer Center. The pBR328 vector containing the full-length human insulin cDNA clone (510 bases) was a gift from Dr. Graeme I. Bell of the University of Chicago.

3.3.2 Methods

3.3.2.1 Cloning of the Full-length Human Preproinsulin Splice Variant (FL-SPV)

Creation of full-length human preproinsulin cDNA using PCR: The primers used in this study are listed in Table 3.1 (Primers 1-6). A schematic representation of the constructs created is shown in Figure 3.4. The plasmid received from G. I. Bell (referred to here as hINS plasmid) contained the full length human insulin cDNA sequence and was used as the DNA template during round 1 of PCR. In order to produce a DNA template for transcription of the human preproinsulin splice variant sequence (SPV) PCR was used to add 74 bases of the 5’UTR sequence on the 5’ end of the human insulin cDNA sequence to form full-length SPV DNA. After creation of the SPV DNA template, BamH I and EcoR I restriction sites were added to the 5’ and 3’ ends of the sequence during PCR amplification to facilitate cloning into the pBluescript II SK(-) vector (Primers 5c and 6). The PCR conditions used for each round of amplification are found in Table 3.2, and generally follow the procedures provided by Stratagene for the respective polymerases. During the first round of amplification 7.2 ng of the hINS plasmid was used as a DNA template. After each round of amplification, the resulting products were visualized on a 3% NuSieve® GTG® agarose gel,
and gel plugs containing the desired PCR product were used as the template for future rounds. Products from the final round of amplification were purified using a ChargeSwitch® PCR Clean-up Kit then used as a template for cloning.

**Cloning procedure:** Primers (600 pmol) used for round 5c of amplification (Primers 5c and 6) were phosphorylated with T4 PNK (15 U; in 1X PNK Buffer, pH 9.0, 50 mM DTT, 0.6 mM rATP) at 37°C for 45 min, followed by heat inactivation at 70°C for 10 min. Phosphorylated primers were used in a PCR reaction (50 µl; containing 1X Herculase® II reaction buffer, 200 µM each dNTP, 250 pmol each forward and reverse phosphorylated primer, 18-36 ng purified DNA template, 1 µl DMSO (100 %), 0.5 µL Herculase® II fusion DNA polymerase; denaturation at 98°C, 20 sec; annealing 65°C, 20 sec; and elongation 72°C, 90 sec; 25 cycles). After first trying to clone the full-length SPV sequence into pBluescript II SK (-) with the EcoR I and BamH I sites was unsuccessful, Sma I digestion of the vector was used to facilitate cloning through blunt-end ligation. pBluescript II SK (-) vector (1.5 µg) was digested with 20 U Sma I (30 µl; 1X NEB buffer #4) at 25°C for 2 h. Complete digestion of the plasmid was verified using a 1% Reliant® FastLane® agarose gel (1X TBE for 30 min at 175 V). The digested vector was then purified using phenol/chloroform extraction followed by ethanol precipitation, and brought up in 7 µl sterile water. Digested vector (7 µl) was then dephosphorylated using calf intestinal phosphatase (CIP) (10 µl; containing 1X CIP buffer, 2 U CIP) at 37°C for 45 min. The resulting product was purified on a 1% SeaPlaque® low-melting agarose gel (in 1X TBE) at 100 V. The band containing the dephosphorylated, digested vector was excised from the gel after visualization using UV light. To facilitate the ligation of the digested vector with the phosphorylated PCR product, the agarose containing vector DNA was heated in a microcentrifuge tube to 65°C.
and added to the ligation reaction before addition of T4 DNA ligase (11 µl; consisting of 4.0 µl agarose containing vector DNA, 5.0 µl phosphorylated insert PCR product, in 1X T4 DNA ligase buffer, 400 U T4 DNA ligase). After incubation at 4°C overnight, the ligation reaction was melted at 68°C for 10 min, then diluted with 40 µl 1X TE, and the ligation incubated at 37°C until needed for the transformation. The construct was transformed into XL-10 Gold® ultracompetent cells and plated on LB/carbenicillin plates overnight. The orientation and sequence of all constructs were verified by sequencing. This plasmid is referred to in this work as FL-SPV plasmid.

3.3.2.2 Cloning of the 5’UTR of Human Preproinsulin Splice Variant (SPV-HP)

Creation of the human preproinsulin intermediate splice variant hairpin template (INT-SPV-HP) using PCR: Primers used in this procedure are listed in Table 3.1 (Primers 5c and 9). A schematic representation of the construct is seen in Figure 3.4. SPV plasmid containing the full-length human preproinsulin splice variant sequence (FL-SPV) was used as a template. PCR was used to amplify out the 5’UTR sequence and add an Xba I restriction site on the 3’ end to facilitate cloning and in vitro transcription. Primers were phosphorylated using T4 polynucleotide kinase (as described in Section 3.3.2.1), and used in the amplification reaction with Herculase® DNA polymerase (50 µl; containing 1X Herculase® reaction buffer, 200 µM each dNTP, 400 pmol each forward and reverse primer, 2.6 ng FL-SPV plasmid as a template, 1 µl DMSO (100%), 0.5 µl Herculase® II fusion DNA polymerase; 5 cycles of denaturation at 98°C, 20 sec; annealing 60°C, 20 sec; extension 72°C, 30 sec, followed by 25 cycles of denaturation at 98°C, 20 sec; annealing 63°C, 20 sec; extension 72°C, 30 sec) as suggested by Stratagene. The size of the resulting PCR products (referred to here as intermediate SPV hairpin PCR products) was checked on a Reliant®
FastLane® precast 2.5 % agarose gel. To remove the remaining SPV parent plasmid in the PCR product, digestion with Dpn I was done using the protocol from the Stratagene QuickChange™ II XL Site Directed Mutagenesis Kit. The resulting solution was cleaned using a ChargeSwitch® PCR Clean-up Kit. pBluescript II SK(-) vector was digested with EcoR I/Xba I or Sma I (as described in Section 3.3.2.1), and then dephosphorylated using calf intestinal phosphatase (as previously described). Digested pBluescript II SK(-) vector and intermediate SPV hairpin PCR product were purified on a 1% SeaPlaque® low-melting agarose gel, and ligated together using T4 RNA ligase at 4°C overnight (as described previously). The construct (INT-SPV-HP) was transformed into XL-10 Gold® ultracompetent cells according to the protocol provided by Stratagene, and plated on LB/carbenicillin plates overnight. The orientation and sequence of all constructs were verified by sequencing.

**Construction of the plasmid containing the human preproinsulin splice variant hairpin sequence (SPV-HP):** To remove extra base pairs between the T7 promoter and the start of the human preproinsulin 5’UTR sequence, the previously cloned plasmid (INT-SPV-HP) was digested with Pst I (100 µl; containing 31.5 µg INT-SPV-HP plasmid, 1X NEB Buffer #3, 0.1 µg BSA, 200 U Pst I) for 2 h at 37°C, the purified by phenol/chloroform extraction and ethanol precipitation. The resulting plasmid was brought up in 10 µl sterile water. The Pst I digested plasmid was then digested with Kpn I (100 µl; containing 10 µl Pst I digested plasmid, 1X NEB Buffer #1, 0.1 µg BSA, 100 U Kpn I) for 2 h at 37°C. The digested plasmid was then purified by phenol/chloroform extraction, and ethanol precipitation. The digested plasmid was brought up in 10 µl sterile water. After digestion, the remaining 3’ overhang sequence was removed with Klenow (50 µl; containing 1X NEB
buffer #2, 40 µM each dNTP, 20 U Klenow) by incubation for 15 min at 25°C. The reaction was stopped by adding 1 µl 0.5 M EDTA and heating to 70°C for 20 min. The plasmid was re-ligated, transformed into XL-10 Gold® ultracompetent cells, and the sequence verified. The resulting plasmid contained the preproinsulin 5’UTR with 19 bases between the T7 promoter and the start of the UTR. The folding of this construct was checked with mfold\textsuperscript{63,64} and RNAstructure\textsuperscript{65} and was shown to have the same secondary structure as the original SPV hairpin RNA. This structure is referred to in this work as SPV-HP.

3.3.2.3 Creation of the Full-length and 5’UTR Native Human Preproinsulin Constructs (FL-NAT and NAT-HP)

Loop-out mutagenesis using a Stratagene QuickChange\textsuperscript{TM} II XL Site Directed Mutagenesis Kit was performed on the FL-SPV and SPV-HP plasmids to remove the 26 bases present in the SPV hairpin RNA that are not present in the native preproinsulin (NAT, and NAT Hairpin) RNAs. Primers were constructed so that the deletion fell in the middle of the sequence and are listed in Table 3.1 (Primers 7 and 8). A schematic representation of the construct is seen in Figure 3.4. PCR amplification was performed using PfUUltra\textsuperscript{®} DNA polymerase, and the methylated parental DNA template was degraded using Dpn I. The mutants were transformed into XL-10 Gold\textsuperscript{®} ultracompetent cells. DNA was sequenced, cells were grown up, and plasmids purified using a Maxi-prep kit. The resulting plasmids contained the sequences for either the full-length native human preproinsulin mRNA, or the native hairpin RNA sequence. These plasmids are referred to in this work as FL-NAT and NAT-HP, respectively.
3.3.2.4 Transcription of NAT and SPV Hairpin RNA

Plasmid containing the desired sequence was digested with Xba I (100 µl; containing ~ 100 µg plasmid (NAT-HP or SPV-HP), 1X NEB Buffer #2, 0.1 µg BSA, 200 U Xba I) for 2 h at 37°C, treated with 0.5% SDS and 10 µg proteinase K for 2 h at 50°C, and then purified by phenol/chloroform extraction and ethanol precipitation. The concentration of the digested plasmid was determined by taking the absorbance at 260 nm, and purity was assessed by using the A260/A280 ratio. To generate hairpin RNA a MEGAscript<sup>TM</sup> T7 Kit was used according the protocol provided by Applied Biosystems for 4 h, with the addition of high concentration T7 RNA polymerase (200 U T7 Polymerase Plus<sup>TM</sup>) to increase yields, and 20 U Superase·In to inhibit RNase degradation. DNA in the resulting sample was digested using 30 U TURBO DNase provided with the MEGAscript<sup>TM</sup> Kit at 37°C for 15 min. Production of the RNA transcript was verified on a 10% denaturing (8 M Urea) polyacrylamide gel. The RNA was purified using a mini Quick Spin RNA Column (following the protocol provided by Roche), followed by acid phenol/chloroform extraction, and ethanol precipitation. Absorbance at 260 nm was used to determine the final RNA concentration.

3.3.2.5 RNA End-labeling

RNA was 3’ end labeled with 50 µCi [5'-32P]pCp using T4 RNA ligase (30 µl; containing 1X T4 RNA ligase buffer, ~150 µM RNA template, 30 U Superase·In, 3 µl DMSO (100%), 60 U T4 RNA ligase) overnight at 4°C. RNA was run through a mini Quick Spin RNA column (following the protocol provided by Roche) and further purified on a small 8 % (7 M Urea) polyacrylamide gel for 30 min at 200 V. RNA bands were visualized by autoradiography and excised from the gel. Gel slices were crushed and soaked in elution
buffer (0.5 M NH₄OAc, 1 mM EDTA, 0.1% SDS) with 20 U Superase·In RNase inhibitor in Ultrade-MC filters at 37°C overnight. Samples were spun down at 4,000 x g for 15 min then ethanol precipitated overnight with 5 µg linear acrylamide. RNA concentration was determined using the absorbance at 260 nm.

3.3.2.6 RNase Cleavage Assay

Footprinting solutions were made by adding a trace amount (~ 0.53 µM) of radiolabeled NAT-HP or SPV-HP RNA to a stock solution (130 µl; containing 3.9 µM unlabeled RNA, in 80 mM Tris / 20 mM KCl, pH 7.4). The RNA solution was then folded by heating to 95°C for 5 min followed by slow cooling to room temperature. Aliquots of the stock solution (5 µl) were used in the RNase cleavage reactions. To determine the correct concentration of RNase for single hit kinetics, the footprinting solutions (10 µl; containing 1 µl RNase, 4 µl ddH₂O) were then incubated for 20 min at room temperature. Varying concentrations of RNase A (0.00001 - 0.001 pg) and RNase T1 (.001 - 1 U) were used, and the reaction stopped by adding 10 µl Gel Loading Buffer II. Samples were separated on 10% (7 M urea) denaturing polyacrylamide gels, exposed to a phosphorimaging screen overnight and analyzed using a Storm® 860 phosphoimager.

Once the correct concentrations of RNase were determined, footprinting reactions were performed to determine the small molecule binding site. RNA was folded as described above, and footprinting solutions were incubated with varying concentrations of the small molecule for 20 min at room temperature. RNase A (0.0001 U, 10 min, or 0.001 U, 5 min) and RNase T1 (0.2 U, 10 min) were added and incubated for the indicated time. The cleavage reaction was stopped by adding 10 µl Gel Loading Buffer II. Samples were separated on 10% (7 M urea) denaturing polyacrylamide gels, exposed to a phosphorimaging
screen overnight and analyzed using a Storm® 860 phosphoimager. Band intensity was quantified using ImageQuant® software. RNA ladders were constructed using the 3’ labeled RNA via enzymatic cleavage by RNase T1, and alkaline hydrolysis using protocols from Applied Biosystems.

3.3.2.7 RuO$^2+$ Small Molecule Binding Assay

The RuO$^2+$ oxidant was generated from Ru-OH$_2$ by bulk electrolysis, holding the aqueous solution at 0.85 V with rapid stirring for 10 min.$^{66,67}$ Generation of RuO$^2+$ was visually evident through a color change from brown to yellow, and by leveling off of the current. Concentration of RuO$^2+$ was determined by the absorbance of Ru-OH$_2$ at 476 nm ($\varepsilon = 9600 \text{ M}^{-1}\text{cm}^{-1}$). The Ru-OH$_2$ complex was generated by reduction of RuO$^2+$ with an excess of L-ascorbic acid. A stock solution of $^{32}$P 3’ end labeled RNA ($\sim 1.5 \mu M$) was folded in 40 mM Tris/ 10 mM KCl (pH = 7.4) by heating to 95°C for 5 min followed by slow cooling to room temperature. Aliquots of folded, labeled RNA solution (5 µl, $\sim 0.75 \mu M$ labeled RNA) were used in each reaction. Labeled RNA solution (5 µl) and drug were incubated at room temperature for 20 min (reaction volume 7 µl). Concurrently, RuO$^2+$ was generated as described above. RuO$^2+$ (100 µM) was then added and reacted with the folded RNA solution for 5 min at room temperature (10 µl total reaction volume). The reaction was quenched with 15 µl ethanol, and speed vacuumed to dryness. Lyophilized RNA was then aniline treated with 20 µl 1 M aniline-acetate (pH 4.5), in the dark for 20 min at 60°C. The solution was frozen to -80°C, lyophilized to remove the aniline, then washed with water, frozen, and lyophilized two more times. Samples were resuspended in 5 µl Gel Loading Buffer II, separated on 10 % (7 M urea) denaturing polyacrylamide gels, exposed to a phosphorimaging screen overnight and analyzed using a Storm® 860 phosphoimager. Band intensity was
quantified using ImageQuant® software. RNA ladders were constructed using 3’ labeled RNA via enzymatic cleavage by RNase T1, and alkaline hydrolysis using protocols from Applied Biosystems.

3.3.2.8 Flash-Quench Oxidation

A stock solution of $^{32}$P-pCp 3’ end labeled RNA (~4.5 µM) was folded in 40 mM Tris-Cl (pH = 7.6) by heating to 95°C for 5 min followed by slow cooling to room temperature. Aliquots of 5 µl labeled RNA solution (~0.5 µM labeled RNA) and drug were incubated in glass polyspring inserts at room temperature for 20 min. After incubation, 100 µM Ru(bpy)$_3^{2+}$ and 1 mM Co(NH$_3$)$_5$Cl$_2^{2+}$ were added to the RNA (total reaction volume 20 µl) and the solutions were kept in the dark to prevent photolysis. The flash-quench method$^{68,69}$ was then used to generate oxidative lesions on the RNA using Ru(bpy)$_3^{3+}$ as an electron donor, and Co(NH$_3$)$_5$Cl$_2^{2+}$ as a quencher. Photolysis was performed with a 350 W Hg lamp (Oriel) with a 265 nm cutoff filter, and water filter, for 10 min. The reaction was quenched with 60 µl ethanol, and then ethanol precipitated with 2.5 µg linear acrylamide and 0.5 µl glycoblue, followed by treatment with aniline (see Section 3.3.2.7). The samples were resuspended in 5 µl Gel Loading Buffer II, and separated using a 10% (7 M urea) denaturing polyacrylamide gel, exposed to a phosphorimaging screen overnight and analyzed using a Storm® 860 phosphoimager. Band intensity was quantified using ImageQuant® software. RNA ladders were constructed using the 3’ labeled RNA via enzymatic cleavage by RNase T1, and alkaline hydrolysis using protocols from Applied Biosystems.
3.3.2.9 Transcription of Full-length mRNA: Human NAT Preproinsulin and Human SPV Preproinsulin

The human preproinsulin plasmids FL-NAT and FL-SPV encode for the preproinsulin mRNA with the native (NAT, 59 bases) or splice variant (SPV, 85 bases) sequence in the 5’UTR, in the pBluescript II SK(–) vector under the T7 promoter, respectively. The plasmids were digested with BamH I (100 µl; containing ~100 µg plasmid in 1X BamH I buffer, 0.1 µg BSA, 200 U BamH I) at 37°C for 2 h. All digestions were followed by treatment with 0.5 % SDS, 10 µg proteinase K for 2 h at 50°C, and purification using phenol/chloroform extraction and ethanol precipitation. Concentration of linearized vector was determined using the absorbance at 260 nm. Full length human preproinsulin mRNA was transcribed from the linearized vector using the T7 promoter. All transcriptions were performed using a MEGAscript T7 Kit™ for 3 h at 37°C (following the protocol provided by Applied Biosystems) with addition of 20 U Superase·In to inhibit RNA degradation. Excess plasmid DNA was removed by incubation with TURBO DNase (40 U, 15 min at 37°C) followed by purification using a MEGAclear Kit™ (according to the protocol provided by Applied Biosystems), acid phenol/chloroform extraction, and ethanol precipitation. Production of the correct size mRNA was confirmed by electrophoresis on a Reliant® FastLane® 1.25% agarose RNA gel. RNA concentrations were determined using the absorbance at 260 nm.

3.3.2.10 In Vitro Translation of mRNA

Translation of mRNA (NAT / SPV, 1.5 µg; luciferase 1 µg; total reaction volume 25 µl) occurred after pre-incubation with or without small molecule for 20 min at 30°C using wheat germ extract and 11 µCi [35S]methionine (following the procedure provided by
Promega) for 60 min at 25°C. Synthesized protein was analyzed by electrophoresis on either 10-20% or 16% Tricine gels in Novex® Tricine-SDS running buffer (pH 8.3) according to the protocol provided by Invitrogen. The gels were fixed in a 50% methanol / 10% acetic acid solution for 10 min, rinsed twice with ddH_2O, dried using a DryEase® Mini-Gel Drying Kit and exposed to a phosphorimaging screen overnight. Protein band intensities were quantitated based on [35S]methionine incorporation using ImageQuant® software.

3.4 Results and Discussion

3.4.1 Secondary Structure Analysis of 32P-End-Labeled Human Preproinsulin Hairpin RNA

The 5’UTR of the human preproinsulin gene contains exon 1 and part of exon 2, which are separated by intron 1 in the unspliced RNA. Any changes in intron 1 will alter the 5’UTR without disrupting the gene product. Chemical and enzymatic footprinting were used to investigate the secondary structure of the native and splice variant forms of the human preproinsulin hairpin RNA. Experiments were carried out under conditions that should produce one cleavage or modification per RNA molecules. RNAs used for the nuclease mapping contained 59 bases of the 5’UTR for the native insulin construct, and 85 bases for the splice variant construct, as well as flanking sequence. The results of the footprinting agreed with the proposed secondary structures (Figure 3.3) and are summarized in Figure 3.5.

3.4.1.1 Native Preproinsulin mRNA Structure

For the native construct, most nucleotides predicted to be in single-stranded regions (loops or bulges) were cleaved with RNase A, T1 or I (Figure 3.6). Some variability in cleavage is seen in positions surrounding loop/bulge regions indicating possible nucleotide
flexibility. An example of this variability is seen in the cleavage with RNase T1 and RNase I around the G49 position. This nucleotide is computationally predicted to be double stranded, but is cleaved at all relevant nuclease concentrations. Cleavage by RNase V1, which is used in the identification of double stranded regions, also occurs around this position, suggesting that there may be flexibility in the RNA structure at this position in solution. Chemical footprinting with RuO$_2^{2+}$, which causes oxidation of solvent-accessible guanines prone to cation binding, produces a cleavage pattern similar to the RNase T1 pattern (Figure 3.7). Prominent cleavage occurred at positions G25, G28, G30 and G31 all found in the terminal loop region. Due to the solvent accessibility of the unpaired bases in this region, cleavage with RuO$_2^{2+}$ also appeared to occur at nucleotides surrounding the guanines, though this could be an artifact of separation by electrophoresis. Use of the flash-quench technique with Ru(bpy)$_3^{2+}$ also causes oxidation of guanine nucleobases. Prominent cleavage occurred at the same positions as found using RuO$_2^{2+}$, along with secondary cleavage seen at G15, G16, G39, G42 and G49 (Figure 3.8). Due to the short lifetime of the excited Ru(bpy)$_3^{3+}$ oxidant, cleavage occurs in a static manner, at regions that are spatially near the oxidant, not those that are able to “breathe” and become solvent accessible over a longer timescale. Cleavage at G39, which was not seen with either RNase T1 or RuO$_2^{2+}$, suggests that this position may not be flexible enough to be in direct contact with either RuO$_2^{2+}$ or the RNases, but oxidation of the guanine may occur through long range electron transfer with Ru(bpy)$_3^{3+}$.

Double-stranded stem regions were probed using RNase V1. Cleavage patterns were similar to those predicted computationally. The region around G49 in the lower stem region produced an area of cleavage overlap with RNase T1, indicating that this lower stem region may be structurally flexible.
A preproinsulin binding factor (ppIGE regulatory factor) has been identified that binds in the area 40-48 nucleotides from the beginning of the 5’UTR and is involved in glucose-regulated translational control. From the predicted secondary structure, this sequence spans a region which includes a four nucleotide bulge, 5’-A₄₃U₄₄C₄₅A₄₆-3’, flanked by base paired nucleotides. This bulged region could be important for making contacts with the ppIGE regulatory factor, similar to what is seen in the IRP:IRE system. Mutation of this sequence, or disruption as seen in the splice variant, abolishes this glucose-regulated translational control.

### 3.4.1.2 Preproinsulin mRNA Splice Variant Structure

The preproinsulin splice variant contains an extra 26 bases than the native hairpin RNA that change the overall secondary RNA structure. Because the inserted sequence is in the middle of the hairpin, the ends of the two constructs share a similar structure. This can be seen in both the predicted secondary structure and the resulting cleavage patterns of the native and splice variant constructs (Figures 3.3 and 3.5).

Overall, the predicted secondary structure of the splice variant differed slightly from the cleavage pattern seen using enzymatic and chemical nucleases (Figures 3.5 and 3.9). Little cleavage was seen around nucleotides 44-45 and 54 which were predicted to be single stranded. There was no evidence of base pairing occurring in this region, as evidenced by the lack of substantial cleavage by RNase V1. Also, the area around G75 which is found in the lower stem region is cleaved by RNase T1, suggesting that this is a more flexible region that exhibits base paired characteristics. The other RNase cleavage patterns agree with the predicted structure.
The RuO$^{2+}$ and flash-quench cleavage patterns are similar to that of RNase T1 (Figures 3.10 and 3.11). Prominent cleavage is seen at positions: G10, G11, G19, G25, G28, G39, G42, G43, G47, G54, G55, G62, and G64 in both methods. Due to the size of the preproinsulin SPV hairpin RNA, it was difficult to get single nucleotide resolution using gel electrophoresis. The structure of a population of folded molecules can vary in solution. It is possible that there is a greater tertiary structure present in the preproinsulin hairpin RNA that accounts for some of the reduced cleavage in single stranded regions. These interactions would not be predicted using computational secondary structure methods. This could also explain some of the regions such as those around G75 which show both single and double stranded characteristics. Once the secondary structures of the preproinsulin mRNAs were probed, this information was used in the search for small molecules that bind preproinsulin hairpin RNA.

3.4.2 Screening Compounds to Find a Human Preproinsulin Hairpin RNA Binder

To determine if our hypothesis was valid, a number of drug-like molecules were screened for binding to both preproinsulin hairpin RNA constructs. The structures of these molecules are given in Table 3.3. There are three classes of compounds that are well-known for their interactions with nucleic acids. These include intercalating agents (i.e. ethidium bromide), minor groove binders (i.e. Hoechst dyes), and major groove binders (i.e. methyl green). Due to the absence of previous work on these RNAs, and little information about their secondary and tertiary structures, we began our investigation by screening compounds that were known RNA binders. Initially, yohimbine and promazine, the two compounds that were able to up-regulate ferritin synthesis (discussed in Chapter 2), were studied for their ability to bind and regulate translation of preproinsulin mRNA. After this initial screen, four
more compounds were tested for their binding ability. Overall the compounds fell into four
categories: 1) phenothiazine derivatives,\textsuperscript{71,72} 2) Hoechst dyes,\textsuperscript{73} 3) indole derivatives,\textsuperscript{71} and 4) aminoglycoside antibiotics.\textsuperscript{59,60}

Compounds were screened against the native or splice variant sequence using RNase
footprinting, the flash-quench technique or the RuO$^{2+}$ footprinting assay. Promising
candidates were determined to be those that inhibited RNA cleavage or guanine oxidation, or
caused a change in the secondary structure of the RNA. Promazine and yohimbine showed
no evidence of binding to either RNA (\textbf{Figures 3.12 and 3.13}). A secondary screen
identified neomycin B (\textbf{Figures 3.14-3.21}) and kanamycin B (\textbf{Figures 3.22-3.29}) as
promising compounds for both constructs, while Hoechst 33258 (\textbf{Figures 3.14 and 3.30}) and
hygromycin B (\textbf{Figures 3.31 and 3.32}) lacked any specific binding characteristics for either
RNA, and were not further studied. Increasing concentrations of neomycin B caused a
decrease in cleavage around nucleotides 7-8, 13, 25-31, 42-47 and 49-50 of the NAT
construct, and kanamycin B changed the amount of cleavage around nucleotides 7-8, 25-31,
36-39 and 44-49 in the NAT construct. For the SPV construct, changes in cleavage were
seen around nucleotides 20-23, 30-31, 40, 42-43, 47, 50-51, 70-71, and 75 with neomycin B.
Kanamycin B produced changes in cleavage at all nucleotides in the SPV construct. Of the
compounds tested, promazine, yohimbine, neomycin B and kanamycin B were studied more
thoroughly. There is some concern that these compounds exhibit non-specific binding due to
their effect on cleavage in multiple regions of the SPV hairpin RNA, although this could be
due to a change in the overall secondary structure. Aminoglycosides such as neomycin B
have been shown to interact with multiple RNAs; they are often called “promiscuous”,
because they bind a range of different targets, and more than one aminoglycoside molecule
can bind per RNA.\textsuperscript{74-81} The majority of RNA binding compounds, including the aminoglycosides, are cationic compounds carrying several positively charged groups. Electrostatic interactions between the negatively charged RNA backbone and cationic groups enhance the binding affinity in RNA-drug complexes, contributing to both specific and nonspecific binding interactions.\textsuperscript{77} Despite all this, aminoglycosides have been shown to display high specificity in eliciting biological function by interaction with a particular RNA structure, such as the bacterial decoding site RNA.\textsuperscript{77}

### 3.4.3 Comparison of Methods Used to Identify a Human Preproinsulin Hairpin RNA Binder

Of the three methods used to identify small molecules that are able to bind to the hairpin RNA constructs, enzymatic cleavage using a combination of RNases was the most beneficial followed by the use of the RuO\textsubscript{2}\textsuperscript{2+} assay. RNase A is able to cleave after single-stranded C and U nucleotides, while RNase T1 is able to cleave after single-stranded guanine nucleotides. Using a combination of RNase A and RNase T1 allowed for more of the hairpin RNA structure to be probed versus the solvent-accessible guanine specific RuO\textsubscript{2}\textsuperscript{2+} cleavage. The flash-quench reaction showed an interesting result; with increased small molecule concentration, the amount of RNA cleavage increased to such a degree that no full-length RNA remained (Figure 3.33). The concentrations of small molecule used in this assay were the same as those used in the RuO\textsubscript{2}\textsuperscript{2+} assay and the same ratio as those used in the RNase experiments. The mechanism behind the enhanced cleavage is unknown, but it is possible that the aminoglycosides are enhancing the ability of the Ru(bpy)\textsubscript{3}\textsuperscript{2+} molecule to be excited, either by direct electron transfer of the small molecule, or through interaction with the Co(NH\textsubscript{3})\textsubscript{5}Cl\textsuperscript{2+} quencher. This would cause an increased number of Ru(bpy)\textsubscript{3}\textsuperscript{3+} molecules in the excited state, resulting in a larger amount of cleaved RNA.
3.4.4 Regulation of Protein Synthesis by Small Molecules: In Vitro Translation Assay

3.4.4.1 Human Preproinsulin Native and Splice Variant mRNA Constructs

After determining that two small molecules, neomycin B and kanamycin B, bound or had an effect on the secondary structure of the preproinsulin native and splice variant hairpins, we sought to determine whether these small molecules could retain their binding characteristics in a more complex in vitro translation mixture. Use of both the rabbit reticulocyte lysate and wheat germ cell-free expression systems was explored, and the wheat germ assay was superior to the reticulocyte lysate for production of preproinsulin. To perform this assay, full length native or splice variant preproinsulin mRNA was pre-equilibrated with or without small molecule before the extract, amino acid mixture and radiolabeled methionine were added. A time course assay was performed to make sure preproinsulin synthesis rates were linearly dependent on incubation time at the time chosen (30 min).

Neomycin B and Kanamycin B showed an effective decrease in preproinsulin synthesis for both the native and splice variant forms at all concentrations tested (Figures 3.34 and 3.35). This result validated our hypothesis that a small molecule that has an effect on the secondary structure of the human preproinsulin mRNA should be able to modulate proinsulin translation. The effect of promazine and yohimbine was also examined, but neither of these small molecules caused a change in the amount of translation (Figure 3.36).

3.4.4.2 Wild-type Luciferase mRNA Construct

To determine if the results observed with the down-regulation of preproinsulin synthesis by neomycin B and kanamycin B were dependent on the identity of the transcript, mRNA encoding wild-type luciferase was used as a control. A decrease in luciferase protein
translation was seen for both small molecules at increasing concentrations (Figure 3.37). Kanamycin B showed a greater reduction in protein synthesis than neomycin B, but both compounds had a profound affect on protein synthesis at concentrations greater than 10 µM. These results indicate that at high concentrations the aminoglycosides used here most likely have an effect on the translational machinery of the wheat germ system. Aminoglycosides have shown some toxicity at high concentrations. At concentrations lower than 10 µM, the effect of these small molecules on preproinsulin synthesis may be greater than the effect on the luciferase construct, but more data is needed to at lower concentrations to support this result.

Recently, the existence of a binding factor similar to the iron regulatory protein found in the ferritin system, was identified. It has been suggested that binding of this factor to the preproinsulin binding element prevents assembly of the translational initiation machinery and blocks preproinsulin protein synthesis. Glucose is thought to interfere with the binding of the regulatory factor, stimulating preproinsulin synthesis. Because the presence of this factor is limited to the extracts of islet cells, these experiments should be repeated in human islet extract or in vivo in human islets to determine the ability of the aminoglycosides to interact with the RNA when the factor is present.

3.5 Conclusions

Two main conclusions can be drawn from these experiments. First, footprinting by enzymatic cleavage with RNases, or chemical cleavage with RuO$_2^{2+}$ are better methods for determining the specific binding sites for small molecules on mRNA, than the flash-quench technique. Together, the RuO$_2^{2+}$ and RNase cleavage methods provide a valuable screen for
binding site identification. However, with these techniques it would be difficult to elucidate binding sites with nucleotide resolution on RNA targets larger than the preproinsulin splice variant used here. Second, differential binding of small molecules to specific three-dimensional structures in mRNA can alter mRNA translation. As seen with other aminoglycoside:RNA interactions, this binding causes the down-regulation of protein synthesis. A future goal of this work is to identify a small molecule would be able to enhance preproinsulin biosynthesis, allowing more protein to be stored in the secretory granules for release upon glucose stimulus.

These studies indicate that neomycin B and kanamycin B bind secondary structures in both the native and splice variant preproinsulin mRNA. Binding of these small molecules causes a conformational change in the secondary structure of the 5’UTR that result in the down-regulation of protein synthesis. Identification of a small molecule that prevents interaction of the binding factor with the glucose regulation element in the native preproinsulin mRNA should allow for the up-regulation of preproinsulin synthesis.

The alternative splicing of the human preproinsulin mRNA results in a higher level of translation than is observed for the normally spliced RNA. The splice variant form of preproinsulin mRNA is stimulated after long term (>24 h) exposure to glucose, indicating its expression may be important in those suffering from hyperglycemia due to insulin resistance or reduced insulin secretion. Because the splice variant disrupts the glucose factor binding site, finding a small molecule that selectively binds to this splice variant form should allow an increased pool of preproinsulin to be synthesized. This would help maintain insulin stores in times of increased insulin secretion.
3.6 Future Directions

Insulin synthesis and secretion from pancreatic β-cells is a complex biological process regulated at many levels. More work on the drug discovery aspect of this project is needed to increase the selectivity of the small molecules for the different preproinsulin forms. With the identification of a glucose regulatory element in the native preproinsulin mRNA, identification of small molecules that bind specifically and selectively in this region should prevent ppIGE regulatory factor binding and allow for the up-regulation of preproinsulin synthesis.

More work could also be done to determine the effect of the ppIGE regulatory factor:ppIGE interaction. The glucose binding element is thought to act solely through its primary sequence. It is hypothesized that binding of this factor acts in a manner similar to that of the IRP:IRE interaction, by preventing translation initiation machinery assembly, resulting in decreased protein synthesis. Another possible effect would be that binding of the glucose binding factor would create a “landing pad” for ribosomes, stimulating the production of preproinsulin. Use of siRNA, peptides, or other molecules that are able to target specific sequences should also be investigated. Insulin production is also regulated at the transcriptional level, so small molecules that enhance the transcription of preproinsulin mRNA could have an effect on insulin secretion.
### 3.7 Tables

**Table 3.1 Oligonucleotides used in human preproinsulin cloning.** Lowercase letters represent nucleotides involved in the restriction site sequence.

<table>
<thead>
<tr>
<th>Sequence (5’ to 3’)</th>
<th>Description</th>
<th>Restriction Site</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. TGC GTC AGA TCA CTG TCC TTC TGC CATGGC CCT GTG GAT</td>
<td>5’ primer for round 1 of amplification</td>
<td></td>
</tr>
<tr>
<td>2. TGT TCC AAG GGC CTT TGC GTC AGA TCA CTG TCC TTC TGC CAT</td>
<td>5’ primer for round 2 of amplification</td>
<td></td>
</tr>
<tr>
<td>3. AAG CAG GTC TGT TCC AAG GGC CTT TGC GTC AGA TC</td>
<td>5’ primer for round 3 of amplification</td>
<td></td>
</tr>
<tr>
<td>4a. GAG GCC ATC AAG CAG GTC TGT TCC AAG GGC</td>
<td>5’ primer for round 4a of amplification</td>
<td></td>
</tr>
<tr>
<td>4b. GCA TCA GAA GGC ATC AAG CAG GTC TGT</td>
<td>5’ primer for round 4b of amplification</td>
<td></td>
</tr>
<tr>
<td>5a. GGA CAG GCT GCA TCA GAA GAG GCC ATC AAG C</td>
<td>5’ primer for round 5a of amplification</td>
<td></td>
</tr>
<tr>
<td>5b. AGC CCT CCA GGA CAG GCT GCA TCA GAA GAG G</td>
<td>5’ primer for round 5b of amplification</td>
<td></td>
</tr>
<tr>
<td>5c. cgaattc AGC CCT CCA GGA CAG GCT GCA TC</td>
<td>5’ primer for round 5c of amplification and SPV</td>
<td>EcoR I</td>
</tr>
<tr>
<td></td>
<td>hairpin construction</td>
<td></td>
</tr>
<tr>
<td>6. gcggatcc TCT AGT TGC AGT AGT TCT CCA GCT GGT AGA GGG AGC</td>
<td>3’ primer for full length preproinsulin cloning</td>
<td>BamH I</td>
</tr>
<tr>
<td>7. GGC CAT CAA GCA GAT CAC TGT CCT TCT GCC</td>
<td>5’ primer for NAT construct creation</td>
<td></td>
</tr>
<tr>
<td>8. GGC AGA AGG ACA GTG ATC TGC TTG ATG GCC</td>
<td>3’ primer for NAT construct creation</td>
<td></td>
</tr>
<tr>
<td>9. getctaga GCC ATG GCA GAA GGA CAG TGA TCT GAC</td>
<td>3’ primer for making SPV hairpin construct</td>
<td>Xba I</td>
</tr>
<tr>
<td>Round of Amplification</td>
<td>Reaction Conditions</td>
<td>Segment</td>
</tr>
<tr>
<td>------------------------</td>
<td>---------------------</td>
<td>---------</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>- 50 µl reaction volume</td>
<td>1-4</td>
</tr>
<tr>
<td></td>
<td>- 1X PFU Ultra® Buffer</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- 400 pM each forward and reverse primers</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>- 200 µM each dNTP</td>
<td>1-4</td>
</tr>
<tr>
<td></td>
<td>- 0.5 µl PFU Ultra® Polymerase</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>4a</td>
<td></td>
<td>1-4</td>
</tr>
<tr>
<td>4b</td>
<td></td>
<td>1-4</td>
</tr>
<tr>
<td>4c</td>
<td></td>
<td>1-4</td>
</tr>
<tr>
<td>5a</td>
<td>- 50 µl reaction volume</td>
<td>1-4</td>
</tr>
<tr>
<td></td>
<td>- 1X Herculase® Buffer</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- 250 pM each forward and reverse primers</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>5b</td>
<td>- 200 µM each dNTP</td>
<td>1-4</td>
</tr>
<tr>
<td></td>
<td>- 1 µl DMSO (100 %)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- 0.5 µl Herculase® II fusion DNA polymerase</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5c</td>
<td></td>
<td>1-4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
</tr>
</tbody>
</table>
Table 3.3. Names and basic structures of compounds tested in the preproinsulin binding assay.

<table>
<thead>
<tr>
<th>Compound Name</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yohimbine</td>
<td>![Yohimbine Structure]</td>
</tr>
<tr>
<td>Promazine</td>
<td>![Promazine Structure]</td>
</tr>
<tr>
<td>Hoechst 33258</td>
<td>![Hoechst 33258 Structure]</td>
</tr>
<tr>
<td>Hygromycin B</td>
<td>![Hygromycin B Structure]</td>
</tr>
<tr>
<td>Neomycin B</td>
<td>![Neomycin B Structure]</td>
</tr>
<tr>
<td>Kanamycin B</td>
<td>![Kanamycin B Structure]</td>
</tr>
</tbody>
</table>
3.8 Figures

**Figure 3.1.** Schematic of the human insulin sequence and secondary structure. Disulfide bonds (shown here as black lines), are formed between the A- and B-chain at residues A7/B7, and A20/B19. The A-chain contains an internal disulfide bond, between residues 6 and 11.
Figure 3.2. Human insulin biosynthesis. Insulin is initially synthesized as preproinsulin in the rough endoplasmic reticulum (RER). Removal of the signal peptide produces proinsulin. Proinsulin is transferred from the RER, through the golgi, and is deposited into secretory granules. Mature insulin is formed in these granules through removal of the C-peptide by the prohormone convertases (PC) 1/3 and 2, and by carboxypeptidase E/H. Adapted from a figure by Issac Yonemoto.
Figure 3.3. Proposed secondary structures of the 5’UTRs of the human preproinsulin mRNA. The native human preproinsulin 5’UTR (A) and the human preproinsulin splice variant 5’UTR (B) predicted by mfold.\textsuperscript{63,64} The calculated energies (\(\Delta G\)) were -13.3 for the NAT and -24.5 for the SPV 5’UTR. The arrows in (B) mark the 26 base intronic sequence present in the splice variant and the shaded area represents the preproinsulin mRNA glucose-responsive translation element, or ppIGE.
Figure 3.4. Schematic representation of preproinsulin full-length mRNA and hairpin RNA constructs. Schematic representation of (A) the hairpin RNA constructs and (B) the full length preproinsulin constructs. Symbols: (+) indicates residues in the pBluescript SK (-) vector; (-) represents the 5'UTR sequence shown in section C; (+1) refers to the transcription start site and the arrow represents the translation start site. The underlined sequence in (C) is the splice variant sequence found in the FL-SPV and SPV-HP constructs.
Figure 3.5. Comparison of guanine cleavage sites in the native and splice variant preproinsulin RNA hairpins.
Figure 3.6. Analysis of the native preproinsulin mRNA hairpin structure using RNase cleavage. All lanes, [RNA] ~5 µM. Lane 1, folded RNA only. Lane 2, alkaline hydrolysis ladder. Lane 3, guanine Ladder. Lanes 4-6, cytosine, uridine ladder. [RNase], time: Lane 7, RNase A, 0.001 U, 5 min; Lane 8, RNase A, 0.001 U, 10 min; Lane 9, RNase T1, 1 U, 10 min; Lane 10, RNase T1, 1 U, 15 min; Lane 11, RNase T1, 0.2 U, 5 min; Lane 12, RNase T1, 0.2 U, 10 min; Lane 13, RNase I, 0.1 U, 15 sec; Lane 14, RNase I, 0.01 U, 2 min; Lane 15, RNase I, 0.01 U, 2.5 min; Lane 16, RNase I, 0.01 U, 1 min; Lane 17, RNase V1, 0.1 U, 15 min; Lane 18, RNase V1, 0.2 U, 10 min; Lane 19, RNase V1, 0.2 U, 15 min.
Figure 3.7 Analysis of the native preproinsulin mRNA hairpin structure using Ru(tpy)(bpy)O$_2^+$ mediated guanine oxidation. All lanes, [RNA] $\sim 0.75$ µM. Lane 1, RNase T1, 0.2 U, 10 min. Lane 2, RNase A, 0.001 U, 10 min. Lane 3, RNase A, 0.001 U, 5 min. Lane 4, guanine ladder. Lane 5, alkaline hydrolysis ladder. Lanes 6-13, reaction time 5 min. [RuO$_2^+$]: Lane 6, 250 µM; Lane 7, 200 µM; Lane 8, 150 µM; Lane 9, 100 µM; Lane 10, 50 µM; Lane 11, 25 µM. Lane 12, folded RNA with aniline treatment. Lane 13, folded RNA without aniline treatment.
Figure 3.8 Analysis of the native preproinsulin mRNA hairpin structure using the flash-quench technique. Lanes 1-13, [RNA] ~ 1 μM. Lane 1-5, controls: Lane 1, folded RNA with aniline treatment; Lane 2, no photolysis; Lane 3, Ru(bpy)$_3^{2+}$ oxidant only; Lane 4, Co(NH$_3$)$_5$Cl$_2^{2+}$ quencher only; Lane 5 folded RNA with no aniline treatment. [Ru(bpy)$_3^{2+}$, Co(NH$_3$)$_5$Cl$_2^{2+}$], photolysis time: Lane 2, 50 μM, 500 μM, no photolysis; Lane 3, 50 μM, 0 μM, 10 min; Lane 4, 0 μM, 50 μM, 10 min; Lane 5, 50 μM, 500 μM, 10 min; Lane 6, 10 μM, 100 μM, 10 min; Lane 7, 25 μM, 250 μM, 10 min; Lane 8, 50 μM, 500 μM, 10 min; Lane 9, 100 μM, 1 mM, 10 min; Lane 10, 250 μM, 2.5 mM, 10 min; Lane 11, 50 μM, 500 μM, 5 min; Lane 12, 50 μM, 500 μM, 15 min; Lane 13, Guanine ladder.
Figure 3.9. Analysis of the preproinsulin splice variant hairpin RNA structure by RNase cleavage. All lanes, [RNA] ~ 5 µM. Lane 1, folded RNA only. Lane 2, alkaline hydrolysis ladder. Lane 3, guanine Ladder. Lanes 4-6, cytosine, uridine ladder. [RNase], time: Lane 7, RNase A, 0.001 U, 5 min; Lane 8, RNase A, 0.001 U, 10 min; Lane 9, RNase T1, 1 U, 10 min; Lane 10, RNase T1, 1 U, 15 min; Lane 11 RNase T1, 0.2 U, 5 min; Lane 12, RNase T1, 0.2 U, 10 min; Lane 13, RNase I, 0.1 U, 15 sec; Lane 14, RNase I, 0.01 U, 2 min; Lane 15, RNase I, 0.01 U, 2.5 min; Lane 16, RNase I, 0.01 U, 1 min; Lane 17, RNase V1, 0.1 U, 15 min; Lane 18, RNase V1, 0.2 U, 10 min; Lane 19, RNase V1, 0.2 U, 15 min.
Figure 3.10. Analysis of the preproinsulin splice variant hairpin RNA structure by Ru(tpy)(bpy)O$_2^+$ mediated guanine oxidation. All lanes, [RNA] $\sim$ 0.75 µM. Lane 1, alkaline hydrolysis ladder. Lane 2, guanine ladder. Lane 3, RNase A, 0.001 U, 5 min. Lane 4, RNase A, 0.001 U, 10 min. Lane 5, RNase T1, 0.2 U, 10 min. Lane 6, RNase I, 0.01 U, 2 min. Lanes 7-14, reaction time 5 min. [RuO$_2^+$]: Lane 7, folded RNA without aniline treatment; Lane 8, folded RNA with aniline treatment; Lane 9, 25 µM; Lane 10, 50 µM; Lane 11, 100 µM; Lane 12, 150 µM; Lane 13, 200 µM; Lane 14, 250 µM.
Figure 3.11. Analysis of the preproinsulin splice variant hairpin RNA structure by guanine oxidation using the flash-quench technique. Lanes 1-15, [RNA] = 0.7 µM. Lane 1, alkaline hydrolysis ladder. Lane 2, RNase A, 0.001 U, 10 min. Lane 3, RNase T1, 0.2 U, 10 min. Lanes 4-8 controls: Lane 4, folded RNA with aniline treatment; Lane 5, no photolysis; Lane 6, Ru(bpy)$_3^{2+}$ oxidant only; Lane 7, Co(NH$_3$)$_5$Cl$^{2+}$ quencher only; Lane 8 folded RNA with out aniline treatment. [Ru(bpy)$_3^{2+}$, Co(NH$_3$)$_5$Cl$^{2+}$], photolysis time: Lane 5, 50 µM, 500 µM, no photolysis; Lane 6, 50 µM, 0 µM, 10 min; Lane 7, 0 µM, 50 µM, 10 min; Lane 8, 50 µM, 500 µM, 10 min; Lane 9, 10 µM, 100 µM, 10 min; Lane 10, 25 µM, 250 µM, 10 min.; Lane 11, 50 µM, 500 µM, 10 min; Lane 12, 100 µM, 1 mM, 10 min, Lane 13, 250 µM, 2.5 mM, 10 min; Lane 14, 50 µM, 500 µM, 5 min; Lane 15, 50 µM, 500 µM, 15 min.
Figure 3.12. Effect of promazine and yohimbine on RNase cleavage of the native human preproinsulin hairpin RNA. Lane 1, alkaline hydrolysis ladder. Lane 2, guanine ladder. Lane 3, folded RNA only. Lanes 4-8 and 14-18, RNase A, 0.001 U, 10 min. Lanes 9-13 and 19-23, RNase T1, 0.2 U, 10 min. [promazine]: Lane 4, 0 µM; Lane 5, 10 µM; Lane 6, 50 µM; Lane 7, 100 µM; Lane 8, 2 mM; Lane 9, 0 µM; Lane 10, 10 µM; Lane 11, 50 µM; Lane 12, 100 µM; Lane 13, 2 mM. [yohimbine]: Lane 14, 0 µM; Lane 15, 10 µM; Lane 16, 50 µM; Lane 17, 100 µM; Lane 18, 2 mM; Lane 19, 0 µM; Lane 20, 10 µM; Lane 21, 50 µM; Lane 22, 100 µM; Lane 23, 2 mM.
Figure 3.13. Effect of promazine and yohimbine on the RNase cleavage of the human preproinsulin splice variant hairpin RNA. Lane 1, alkaline hydrolysis ladder. Lane 2, guanine ladder. Lane 3, folded RNA only. Lanes 4-8 and 14-18, RNase A, 0.001 U, 10 min. Lanes 9-13 and 19-23, RNase T1, 0.2 U, 10 min. [promazine]: Lane 4, 0 µM; Lane 5, 10 µM; Lane 6, 50 µM; Lane 7, 100 µM; Lane 8, 2 mM; Lane 9, 0 µM; Lane 10, 10 µM; Lane 11, 50 µM; Lane 12, 100 µM; Lane 13, 2 mM. [yohimbine]: Lane 14, 0 µM; Lane 15, 10 µM; Lane 16, 50 µM; Lane 17, 100 µM; Lane 18, 2 mM; Lane 19, 0 µM; Lane 20, 10 µM; Lane 21, 50 µM; Lane 22, 100 µM; Lane 23, 2 mM.
Figure 3.14. Effect of neomycin B and Hoechst 33258 on the RNase cleavage of the native human preproinsulin hairpin RNA. Lane 1, alkaline hydrolysis ladder. Lane 2, guanine ladder. Lane 3, folded RNA only. Lanes 4-8 and 14-18, RNase A, 0.001 U, 10 min. Lanes 9-13 and 19-23, RNase T1, 0.2 U, 10 min. [neomycin B]: Lane 4, 0 µM; Lane 5, 10 µM; Lane 6, 50 µM; Lane 7, 100 µM; Lane 8, 2 mM; Lane 9, 0 µM; Lane 10, 10 µM; Lane 11, 50 µM; Lane 12, 100 µM; Lane 13, 2 mM. [Hoechst 33258]: Lane 14, 0 µM; Lane 15, 10 µM; Lane 16, 50 µM; Lane 17, 100 µM; Lane 18, 2 mM; Lane 19, 0 µM; Lane 20, 10 µM; Lane 21, 50 µM; Lane 22, 100 µM; Lane 23, 2 mM.
Figure 3.15. Effect of neomycin B on RNase cleavage of the native preproinsulin hairpin RNA. A) Quantitation of RNase A cleavage. B) Quantitation of RNaseT1 cleavage. Legend values indicate the final neomycin B concentration in the assay. Data is representative of one trial.
Figure 3.16. Effect of neomycin B on the RNase cleavage of the human preproinsulin splice variant hairpin RNA. Lane 1, alkaline hydrolysis ladder. Lane 2, guanine ladder. Lane 3, folded RNA only. Lanes 4-8, RNase A, 0.001 U, 10 min. Lanes 9-13, RNase T1, 0.2 U, 10 min. [neomycin B]: Lane 4, 0 µM; Lane 5, 10 µM; Lane 6, 25 µM; Lane 7, 50 µM; Lane 8, 100 µM; Lane 9, 0 µM; Lane 10, 10 µM; Lane 11, 25 µM; Lane 12, 500 µM; Lane 13, 100 µM.
Figure 3.17. Effect of neomycin B on RNase cleavage of the human preproinsulin splice variant hairpin RNA. Quantitations were done using ImageQuant™ software. Legend values indicate the final neomycin B concentration in the assay. Data is representative of one trial.
Figure 3.18. Effect of neomycin B on guanine oxidation in the native preproinsulin hairpin RNA. Lane 1, alkaline hydrolysis ladder. Lane 2, guanine ladder. Lane 3, folded RNA only, no aniline treatment. Lane 4, folded RNA only, with aniline treatment. Lanes 5-10, [RuO$_2^+$] = 100 µM, reaction time 5 min. [neomycin B]: Lane 5, 0 µM; Lane 6, 0.01 µM; Lane 7, 1 µM; Lane 8, 5 µM; Lane 9, 10 µM; Lane 10, 50 µM.
Figure 3.19. Effect of neomycin B on guanine oxidation using the RuO$_2^{2+}$ assay for native human preproinsulin hairpin RNA. Legend values indicate the final neomycin B concentration in the assay. Quantitations were done using ImageQuant™ software. Data is representative of one trial.
Figure 3.20. Effect of neomycin B on guanine oxidation in the preproinsulin hairpin splice variant RNA. Lane 1, alkaline hydrolysis ladder. Lane 2, guanine ladder. Lane 3, folded RNA only, no aniline treatment. Lane 4, folded RNA only, with aniline treatment. Lanes 5-10, [RuO$_2^{2+}$] = 100 µM, reaction time 5 min. [neomycin B]: Lane 5, 0 µM; Lane 6, 0.01 µM; Lane 7, 1 µM; Lane 8, 5 µM; Lane 9, 10 µM; Lane 10, 50 µM.
Figure 3.21. Effect of neomycin B on guanine oxidation using the RuO$_2^+$ assay for human preproinsulin splice variant hairpin RNA. Legend values indicate the final neomycin B concentration in the assay. Quantitations were done using ImageQuant™ software. Data is representative of one trial.
Figure 3.22. Effect of kanamycin B on the RNase cleavage of the native human preproinsulin hairpin RNA. Lane 1, alkaline hydrolysis ladder. Lane 2, guanine ladder. Lane 3, folded RNA only. Lanes 4-8, RNase A, 0.001 U, 10 min. Lanes 9-13, RNase T1, 0.2 U, 10 min. [promazine]: Lane 4, 0 µM; Lane 5, 10 µM; Lane 6, 50 µM; Lane 7, 100 µM; Lane 8, 2 mM; Lane 9, 0 µM; Lane 10, 10 µM; Lane 11, 25 µM; Lane 12, 50 µM; Lane 13, 100 µM.
Figure 3.23. Effect of kanamycin B on RNase cleavage of the native human preproinsulin hairpin RNA. A) Quantitation of RNase A cleavage. B) Quantitation of RNaseT1 cleavage. Legend values indicate the final kanamycin B concentration in the assay. Data is representative of one trial.
Figure 3.24. Effect of kanamycin B on the RNase cleavage of the human preproinsulin splice variant hairpin RNA. Lane 1, alkaline hydrolysis ladder. Lane 2, guanine ladder. Lane 3, folded RNA only. Lanes 4-8, RNase A, 0.001 U, 10 min. Lanes 9-13, RNase T1, 0.2 U, 10 min. [kanamycin B]: Lane 4, 0 µM; Lane 5, 10 µM; Lane 6, 25 µM; Lane 7, 50 µM; Lane 8, 100 µM; Lane 9, 0 µM; Lane 10, 10 µM; Lane 11, 25 µM; Lane 12, 50 µM; Lane 13, 100 µM.
Figure 3.25. Effect of kanamycin B on RNase cleavage of the human preproinsulin splice variant hairpin RNA. Quantitations were done using ImageQuant™ software. Legend values indicate the final kanamycin B concentration in the assay. Data is representative of one trial.
Figure 3.26. Effect of kanamycin B on guanine oxidation in the native preproinsulin hairpin RNA. Lanes 5-10, [RuO$_2^{2+}$] = 100 µM, reaction time 5 min. Lane 1, alkaline hydrolysis ladder. Lane 2, guanine ladder. Lane 3, folded RNA only, no aniline treatment. Lane 4, folded RNA only, with aniline treatment. [kanamycin B]: Lane 5, 0 µM; Lane 6, 0.01 µM; Lane 7, 1 µM; Lane 8, 5 µM; Lane 9, 10 µM; Lane 10, 50 µM.
Figure 3.27. Effect of kanamycin B on guanine oxidation using the RuO$_2^{2+}$ assay for native human preproinsulin hairpin RNA. Legend values indicate the final kanamycin B concentration in the assay. Quantitations were done using ImageQuant™ software. Data is representative of one trial.
Figure 3.28. Effect of kanamycin B on guanine oxidation in the preproinsulin hairpin splice variant RNA. Lane 1, alkaline hydrolysis ladder. Lane 2, guanine ladder. Lane 3, folded RNA only, with aniline treatment. Lane 4, folded RNA only, with out aniline treatment. Lanes 5-10, $[\text{RuO}_2^+] = 100 \, \mu\text{M}$, reaction time 5 min. [kanamycin B]: Lane 5, 0 $\mu\text{M}$; Lane 6, 0.01 $\mu\text{M}$; Lane 7, 1 $\mu\text{M}$; Lane 8, 5 $\mu\text{M}$; Lane 9, 10 $\mu\text{M}$; Lane 10, 50 $\mu\text{M}$. 

145
**Figure 3.29.** Effect of kanamycin B on guanine oxidation using the RuO$_2^+$ assay for human preproinsulin splice variant hairpin RNA. Legend values indicate the final kanamycin B concentration in the assay. Quantitations were done using ImageQuant™ software. Data is representative of one trial.
Figure 3.30. Effect of Hoechst 33258 on the RNase cleavage of the human preproinsulin splice variant hairpin RNA. Lane 1, alkaline hydrolysis ladder. Lane 2, guanine ladder. Lane 3, folded RNA only. Lanes 4-8, RNase A, 0.001 U, 10 min. Lanes 9-13, RNase T1, 0.2 U, 10 min. [Hoechst 33258]: Lane 4, 0 µM; Lane 5, 10 µM; Lane 6, 50 µM; Lane 7, 100 µM; Lane 8, 2 mM; Lane 9, 0 µM; Lane 10, 10 µM; Lane 11, 50 µM; Lane 12, 100 µM; Lane 13, 2 mM.
Figure 3.31. Effect of hygromycin B on the RNase cleavage of the native human preproinsulin hairpin RNA. Lane 1, alkaline hydrolysis ladder. Lane 2, guanine ladder. Lane 3, folded RNA only. Lane 4, RNase A, 0.001 U, 10 min. Lanes 5-9, RNase T1, 0.2 U, 10 min. [hygromycin B]: Lane 4, 2 mM; Lane 5, 0 µM; Lane 6, 10 µM; Lane 7, 50 µM; Lane 8, 100 µM; Lane 9, 2 mM.
Figure 3.32. Effect of hygromycin B on the RNase cleavage of the human preproinsulin splice variant hairpin RNA. Lane 1, alkaline hydrolysis ladder. Lane 2, guanine ladder. Lane 3, folded RNA only. Lanes 4-8, RNase T1, 0.2 U, 10 min. [hygromycin B]: Lane 4, 0 µM; Lane 5, 10 µM; Lane 6, 50 µM; Lane 7, 100 µM; Lane 8, 2 mM.
Figure 3.33. Effect of neomycin B on guanine oxidation using the flash-quench technique in the preproinsulin hairpin splice variant RNA. An increase in the concentration of small molecule caused an increase in the amount of guanine oxidation, and resulting RNA cleavage. Lane 1, alkaline hydrolysis ladder. Lane 2, guanine ladder. Lane 3, folded RNA only, with aniline treatment. Lanes 4-9, \([\text{Ru(bpy)}_3^{2+}, \text{Co(NH}_3)_5\text{Cl}^{2+}] = 100 \, \mu\text{M, 1mM; reaction time 5 min. [neomycin B]: Lane 4, 0 \, \mu\text{M; Lane 5, 0.01 \, \mu\text{M; Lane 6, 1 \, \mu\text{M; Lane 7, 5 \, \mu\text{M; Lane 8, 10 \, \mu\text{M; Lane 9, 50 \, \mu\text{M.}}]}
\]
Figure 3.34. *In vitro* translation assay to determine the effect of neomycin B and kanamycin B on protein translation. [*35S*]methionine SDS-PAGE gel depicting translation of (A) native preproinsulin mRNA in the presence of kanamycin B (B) native preproinsulin mRNA in the presence of neomycin B, (C) preproinsulin splice variant mRNA in the presence of kanamycin B, and (D) preproinsulin splice variant mRNA in the presence of neomycin B. [mRNA]= 1ug. [promazine]: Lane 1, 0 µM; Lane 2, 1 µM; Lane 3, 10 µM; Lane 4, 40 µM; Lane 5, 80 µM.
**Figure 3.35.** Effect of neomycin B and kanamycin B on human preproinsulin mRNA translation in cell-free extracts as measured by incorporation of $[^{35}\text{S}]$methionine. Data are representative of at least three different trials on at least two separate preparations of RNA. These results are expressed as a percentage of the control (no small molecule added), and the error bars represent standard deviation.
Figure 3.36. Effect of promazine and yohimbine on human preproinsulin mRNA translation in cell-free extracts as measured by incorporation of [35S]methionine. Data are representative of at least three different trials on at least two separate preparations of RNA. These results are expressed as a percentage of the control (no small molecule added), and the error bars represent standard deviation.
Figure 3.37. Effect of neomycin B and kanamycin B on luciferase mRNA translation in cell-free extracts as measured by incorporation of $[^{35}\text{S}]$methionine. Data are representative of at least three different trials. These results are expressed as a percentage of the control (no small molecule added), and the error bars represent standard deviation.
3.9 References


(53) Knight, S. W.; Docherty, K. *Biochem. Soc. Trans.* 1991, 19, 120S.


(75) Hermann, T. Biopolymers 2003, 70, 4-18.


