INTERACTIONS BETWEEN PHOTOSWITCHES AND NUCLEIC ACIDS AND THE SELECTION OF NUCLEOTIDE SYNTHASE RIBOZYMES

by

Kelly Elizabeth Chapple Cadieux
B.Sc., Simon Fraser University 2001

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

In the
Department of Molecular Biology and Biochemistry

© Kelly Elizabeth Chapple Cadieux 2005

SIMON FRASER UNIVERSITY

Fall 2005

All rights reserved. This work may not be
reproduced in whole or in part, by photocopy
or other means, without permission of the author.

APPROVAL

Name:	**Kelly Elizabeth Chapple Cadieux**
Degree:	**Master of Science**
Title of Thesis:	**Interactions Between Photoswitches and Nucleic Acids and the Selection of Nucleotide Synthase Ribozymes**

Examining Committee:

Chair: **Dr. Felix Breden**
Associate Professor

Dr. Dipankar Sen
Senior Supervisor, Professor

Dr. Neil Branda
Supervisory Committee Member, Professor

Dr. Jenifer Thewalt
Supervisor Committee Member, Associate Professor

Dr. Peter Unrau
Supervisory Committee Member, Assistant Professor

Dr. Nancy Forde
Internal Examiner
Assistant Professor

Date Defended: November 29th 2005

SIMON FRASER UNIVERSITY library

DECLARATION OF PARTIAL COPYRIGHT LICENCE

The author, whose copyright is declared on the title page of this work, has granted to Simon Fraser University the right to lend this thesis, project or extended essay to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users.

The author has further granted permission to Simon Fraser University to keep or make a digital copy for use in its circulating collection, and, without changing the content, to translate the thesis/project or extended essays, if technically possible, to any medium or format for the purpose of preservation of the digital work.

The author has further agreed that permission for multiple copying of this work for scholarly purposes may be granted by either the author or the Dean of Graduate Studies.

It is understood that copying or publication of this work for financial gain shall not be allowed without the author's written permission.

Permission for public performance, or limited permission for private scholarly use, of any multimedia materials forming part of this work, may have been granted by the author. This information may be found on the separately catalogued multimedia material and in the signed Partial Copyright Licence.

The original Partial Copyright Licence attesting to these terms, and signed by this author, may be found in the original bound copy of this work, retained in the Simon Fraser University Archive.

Simon Fraser University Library
Burnaby, BC, Canada

ABSTRACT

The RNA World hypothesis suggests prebiotic enzymes were RNA molecules (ribozymes) rather than proteins. Previously, *in vitro* selection identified a ribozyme that synthesized the glycosidic bond in a pyrimidine nucleotide. A selection to minimize size and increase catalytic efficiency was concluded and kinetic behaviour of one isolate was examined. In preparation for two new selections for purine and pyridine nucleotide synthase ribozymes, constructs representing theoretical end products were synthesized and used to test the feasibility of the selective procedure.

Dithienylethenes are molecules that interconvert between two isomers (each with unique properties) in response to different light stimuli. It was our interest to control DNA binding and RNA cleavage abilities of novel switches by light exposure. One switch successfully bound DNA, with each isomer displaying the same association constant. Another project featured a histidine-functionalized switch to mimic RNase A. However, no significant difference in cleavage was observed between the two isomers.

For my husband, Jay and my parents, Harv and Chris

ACKNOWLEDGEMENTS

My senior supervisor, Dr. Dipankar Sen, is thanked for his guidance during my graduate studies. Dr. Peter Unrau is thanked for his supervision during Part I of this thesis, and his helpful collaboration with Part II. Dr. Neil Branda is thanked for guidance and supervision during Part II. The other member of my supervisory committee, Dr. Jenifer Thewalt, is thanked for her helpful comments and encouragement during my studies.

Dr. Jonathan Chaires (University of Louisville) is thanked for providing me with advice and help with the DNA binding studies. Thanks also to Fan Sozzi-Guo for assistance and providing reagents and equipment during the bacterial staining experiments. Technical assistance provided by Mrs. M. Tracey (NMR) and Mr. M.K. Yang (CHN microanalysis) is gratefully acknowledged.

My colleagues in the Unrau, Sen and Branda labs are thanked for helpful discussions and support. Sue Tsai and Elaine Chiu are thanked for work on Part I. Matthew Lau is thanked for his continuation of the 6-GMP nucleotide synthase selection. Brian Gorodetsky, David Sud, Bettina Wuestenberg and Corbin Giesbrecht are thanked for providing compounds and their characterization.

Finally, Simon Fraser University is thanked for financial support.

TABLE OF CONTENTS

Approval ii
Abstract iii
Dedication iv
Acknowledgements v
Table of Contents vi
List of Figures viii
List of Tables xi
List of Abbreviations xii

Part I

1 Introduction to Part I 1
- 1.1 Naturally Occurring Catalytic RNA 1
- 1.2 RNA World Hypothesis 2
- 1.3 *In Vitro* Selection 3

2 Nucleotide Synthase Ribozyme Selections 6
- 2.1 4-Thiouridine Nucleotide Synthase Ribozyme Combinatorial Minimization Selection 6
  - 2.1.1 Introduction 6
  - 2.1.2 Materials and Methods 14
  - 2.1.3 Results and Discussion 18
  - 2.1.4 Conclusion 24
- 2.2 Novel Nucleotide Synthase Ribozyme Selections 24
  - 2.2.1 Introduction 24
  - 2.2.2 Materials and Methods 31
  - 2.2.3 Results and Discussion 37
  - 2.2.4 Conclusion 48

Part II

3 Introduction to Part II 50
- 3.1 Molecular Switches and Photochromism 50
- 3.2 Dithienylethenes 53
- 3.3 Biological Motivations 54

4 Biological Applications of Pyridinium Substituted Dithienylethene Switches 58

4.1 Introduction 58
4.2 DNA Binding Studies 59
  4.2.1 Introduction 59
  4.2.2 Materials and Methods 64
  4.2.3 Results and Discussion 72
4.3 Bacterial Staining Studies 88
  4.3.1 Introduction 88
  4.3.2 Materials and Methods 91
  4.3.3 Results and Discussion 93
4.4 Concluding Remarks 100

5 Toward Photo-Controllable RNA Cleavage 105

5.1 Introduction 105
  5.1.1 Ribonuclease A 106
  5.1.2 RNase A Mimics 106
  5.1.3 Motivation 109
  5.1.4 Design of Novel Switches 111
  5.1.5 RNA Substrate 114
5.2 Materials and Methods 115
  5.2.1 General 115
  5.2.2 Synthesis 116
  5.2.3 RNA Substrate and Standards Preparation and Purification 121
  5.2.4 Cleavage Assays 124
5.3 Results and Discussion 127
  5.3.1 *Bis*-Benzimidazole Dithienylethene 127
  5.3.2 *Bis*-Histidine Dithienylethenes 139
5.4 Concluding Remarks 157

6 Future Directions 161

Appendix 164

Reference List 166

LIST OF FIGURES

Figure 1: The goal (a) and strategy used (b) for the original selection for ribozymes capable of synthesizesing a pyrimidine nucleotide. 7

Figure 2: Structure of the thiophilic compound APM. 8

Figure 3: RNA pool design of pyrimidine nucleotide synthase ribozyme selections. 10

Figure 4: Isolate a.6.10 from the degenerate selection. 12

Figure 5: Selection scheme for the deletion selection. 19

Figure 6: Proposed secondary structure of isolate 7.02. 20

Figure 7: Progression of length and deletion through the selection process 22

Figure 8: Biosynthesis of GMP. 27

Figure 9: 6-Thioguanosine monophosphate 27

Figure 10: Overview of 6-thioguanosine monophosphate marker synthesis. 29

Figure 11: a) NAD+ and b) thionicotinamide mononucleotide 30

Figure 12: Overview of thionicotinamide mononucleotide marker synthesis. 31

Figure 13: Mercury gel of thioNAD$^+$ and App^{4s}U ligation with 7.02 39

Figure 14: Gel shift capabilities of ^{4s}U and thioNMN ligated RNA in a) regular and b) mercury gels 41

Figure 15: APM gel of App^{4s}U and App^{6s}G ligation with 7.02 RNA. 46

Figure 16: General scheme of photochromism. 50

Figure 17: UV-Vis absorption spectra of a typical photoisomerization reaction between differently coloured isomers **A** and **B**. 52

Figure 18: Structures of the *bis*-pyridinium dithienylethene switches studied in this chapter. 59

Figure 19: Methyl viologen 60

Figure 20: Fluorinated *bis*-pyridine dithienylethene derivatives reported by Lehn and co-workers (Gilat *et al.*, 1995) 61

Figure 21: Non-fluorinated *bis*-pyridine dithienylethene derivatives. 62

Figure 22: Structures of the known DNA binders netropsin (a) and ethidium bromide (b). 63

Figure 23: Structure of **13-NO$_3$**. 66

Figure 24: Structure of **13-PF$_6$** 67

Figure 25: Synthesis of fluorinated *bis*-pyridinium switch **13-PF$_6$**. 73

Figure 26: UV-Vis spectral changes of **14o** in BPES buffer (2.5 × 10^{-5} M) upon irradiation with 365 nm light 75

Figure 27: Spectral changes of the non-fluorinated *bis*-pyridinium switch **14** isomers (black traces) upon addition of 1 mM CT-DNA (grey traces). 80

Figure 28: UV-Vis absorbance data from the titration of open (a) and closed (b) non-fluorinated *bis*-pyridinium switch **14** with CT-DNA 83

Figure 29: Curve fits for titrations of DNA with each a) open and b) closed non-fluorinated *bis*-pyridinium switch **14**. 86

Figure 30: Crystal violet (Gram's stain). 89

Figure 31: Bacterial staining experiment using fluorinated *bis*-pyridinium switch **13**. 96

Figure 32: Single bacterial cells stained by closed *bis*-pyridinium switch **13c**. 97

Figure 33: Colour change as seen through a light microscope of fluorinated *bis*-pyridinium switch **13** bound to bacteria. 99

Figure 34: General scheme of a RNase A mimic dithienylethene 110

Figure 35: Structure of the *bis*-benzimidazole dithienylethene switch (**26**). 112

Figure 36: Structure of the *bis*-histidine dithienylethene switch (**27**). 113

Figure 37: Abilities of *bis*-benzimidazole and *bis*-histidine switches to adopt RNase A distance between imidazole rings. 114

Figure 38: Structure of **28** 116

Figure 39: Structure of **26** 118

Figure 40: Structure of **27** 119

Figure 41: Structure of **29** 121

Figure 42: Synthesis of *bis*-benzimidazole switch **26**. 128

Figure 43: Absorbance spectra changes of *bis*-benzimidazole switch **26o** (2 × 10^{-5} M, in methanol) upon irradiation with 313 nm light. 129

Figure 44: Absorbance spectra changes of a 2 × 10^{-5} M solution of *bis*-benzimidazole switch **26o** in DMSO upon irradiation with 313 nm light 131

Figure 45: Titration of a 1.2 × 10^{-5} M solution of *bis*-benzimidazole **26o** (1% DMSO, in BPES buffer) with 0.01-0.05% Triton X-100 detergent. 133

Figure 46: RNA cleavage assay with *bis*-benzimidazole switch **26**. 138

Figure 47: Synthesis of *bis*-histidine and mono-histidine switches **27** and **29**, respectively (as performed by David Sud) 141

Figure 48: UV-Visible Absorbance spectra changes of ring-open histidine-functionalized switches (1.75 × 10^{-5} M in H_2O) after irradiation with 254 nm UV light.... 143
Figure 49: Gel results of the second RNA cleavage assay featuring histidine switches **27** and **29** using internally labelled substrate. 147
Figure 50: Triplicate end-labelled cleavage assay featuring histidine switches **27** and **29**. 156
Figure 51: Aligned sequences of clones from the deletion selection. 165

LIST OF TABLES

Table 1: Data obtained from the non-fluorinated *bis*-pyridinium **14** titrations 87

Table 2: Comparison of nucleotide synthase ribozyme selection and *bis*-benzimidazole switch **26** cleavage assay concentrations 136

Table 3: Quantitation of RNA cleavage assay featuring *bis*-benzimdazole switch **26**. 139

Table 4: Quantitated results of two internal cleavage assays using histidine switch 148

Table 5: Quantitated results of end-labelled cleavage assay #1, performed in triplicate. 151

Table 6: Quantitated results of the triplicate cleavage assay #2. 157

LIST OF ABBREVIATIONS

4^sU	4-thiouridine
4^sUra	4-thiouracil
6^sG	6-thioguanine
6^sGMP	6-thioguanosine monophosphate
6^sGua	6-thioguanosine
7.02	second isolate from round seven of deletion selection
A	adenine
Å	Angstrom (0.1 nm)
AMP	adenosine monophosphate
APM	*N*-acryloylaminophenylmercuric acetate
App	adenylated
ATP	adenosine triphosphate
bp	basepair
BPE	sodium phosphate buffer (16 mM NaCl)
BPES	sodium phosphate buffer (201 mM NaCl)
br	broad
BSA	bovine serum albumin
c	closed
C	cytosine

Calcd.	calculated
CT-DNA	calf thymus deoxyribonucleic acid
δ	chemical shift
d	doublet
dd	doublet of doublets
DDQ	2,3-dichloro-5,6-dicyano-1,4-benzoquinone
dec.	(melts with) decomposition
DMF	*N, N*-dimethylformamide
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
DTE	dithienylethene
DTT	dithiothreitol
ε	molar extinction coefficient
EDTA	ethylenediamine tetraacetic acid
G	guanine
HEPES	*N*-(2-hydroxyethyl)piperazine-*N′*-(2-ethanesulfonic acid)
His	histidine
HPLC	high pressure/performance liquid chromatography
Hz	Hertz (cycles per second)
ImpA	adenosine 5′-phosphorimidazolide
J	coupling constant
k_{cat}	catalytic ability (turnover)

K_m	Michaelis constant
Lys	lysine
m	multiplet
MALDI-TOF	matrix-assisted laser ionization time of flight spectroscopy
mRNA	messenger RNA
NAD+	nicotinamide adenine dinucleotide
nm	nanometer
NMN	nicotinamide mononucleotide
NMP	nucleotide monophosphate
NMR	nuclear magnetic resonance
nt	nucleotide
NTP	nucleotide triphosphate
o	open
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PEG	poly(ethylene glycol)
Phe	phenylalanine
ppm	parts per million
pRpp	phosphoribosyl 1-pyrophosphate
Rf	ratio of fronts
Rpm	rotations per minute
RNA	deoxyribonucleic acid
rRNA	ribosomal RNA

RT-PCR	reverse transcription-polymerase chain reaction
s	singlet
TEAA	triethyl ammonium acetate
TE-buffer	tris EDTA buffer
TEAB	triethyl ammonium bicarbonate
thioNAD$^+$	thionicotinamide adenine dinucleotide
TLC	thin layer chromatography
Tris	tris(hydroxymethyl)methylamine
tRNA	transfer deoxyribonucleic acid
U	uracil
UPRTase	uracil phosphoribosyltransferase
UTP	uracil triphosphate
UV	ultraviolet
Vis	visible

1 INTRODUCTION TO PART I

1.1 Naturally Occurring Catalytic RNA

In the early 1980s, several naturally occurring RNA catalysts, called ribozymes, were discovered. The first catalytic RNA isolated was from the organism Tetrahymena (Kruger *et al.*, 1982). This group I intron self-splices ribosomal RNA (rRNA), producing a mature transcript. Shortly after this finding, the RNA portion of the ribonucleoprotein enzyme RNase P (which hydrolyzes precursor RNA substrates such as transfer RNA) was found to possess catalytic ability (Guerrier-Takada *et al*, 1983). Following this was the discovery of four small RNA catalysts: the hammerhead, hepatitis delta virus, hairpin, and *Neurospora* Varkud satellite (VS) ribozymes (Prody *et al.*, 1986; Buzayan *et al.*, 1986; Sharmeen *et al.*, 1988; Saville and Collins, 1990). These ribozymes process the products of rolling-circle replication. These ribozymes cleave at specific sites, due to base-pairing and other interactions (reviewed in Doudna and Cech, 2002).

The self-cleaving and self-splicing ribozymes catalyze two types of phosphodiester bond cleavage reactions (reviewed in Fedor and Williamson, 2005). These produce two different types of products. The small self-cleaving RNAs catalyze a reversible cleavage reaction which produces a 2′-3′ -cyclic phosphate and a 5′-hydroxyl terminus. RNase P and the self-splicing introns catalyze phosphodiester cleavage and ligation reactions which produce a 5′-

phosphate and 3′-hydroxyl termini. These cleavage reactions both proceed with a S_N2-type mechanism.

Two other ribonucleoproteins were recently found to contain catalytically active RNA components. The ribosome is the organelle responsible for protein synthesis from a messenger RNA (mRNA) template. In 2000, the crystal structure of the ribosome placed the location of the rRNA components in the ribosome active site (Nissen *et al.*, 2000). The ribosome was, in fact, a ribozyme. It also was found that the RNA portions of the eukaryotic spliceosome were capable of catalysis alone (Valadkhan and Manley, 2001). These RNA components, called the U2 and U6 small nuclear RNAs, do not catalyze the natural reaction. The rate of this reaction is also much slower than that of the complete spliceosome. However, it expands the range of known natural catalytic RNA molecules. The fact that RNA molecules could catalyze such complicated processes were major findings.

1.2 RNA World Hypothesis

In the mid 1980s, after the discovery of RNase P and the *Tetrahymena* group I intron, Gilbert speculated on the role of RNA in the prebiotic world (Gilbert, 1986). It was wondered if RNA was capable of performing chemistry beyond the cleavage and ligation of RNA phosphodiester bonds. Before protein enzymes developed fully, RNA could have been not only a source of genetic information but also the enzymes of this early environment. This hypothesis is termed the “RNA World” (reviewed in Joyce and Orgel, 1999). Over the course of evolution, proteins would have replaced ribozymes as catalysts. The discovery that the ribosome's catalytic residues are RNA added more support to this hypothesis. This showed that RNA was capable of catalyzing a major step in metabolism. It seemed that structures such as the ribosome and the spliceosome were remnants of RNA catalysis. Much research is ongoing to try to support, or refute, the RNA World theory.

1.3 *In Vitro* Selection

The discovery of evidence to support the RNA World theory has been aided by the development of *in vitro* selection. This method screens a large, diverse pool of RNA or DNA sequences for a particular function, such as catalytic activity (ribozymes and DNAzymes) or ligand binding (aptamers). Depending on the length, completely random pools generally contain over 10^{15} sequences. The diversity of the pool is dependent on the volume used during the amplification step. The limiting reagent is the DNA template, which is optimally amplified at 50 nM concentrations. The assumption is made in this technique that only a very small subset of sequences will have the correct tertiary structure to accomplish the given task. Using a method such as affinity chromatography or polyacrylamide gel electrophoresis (PAGE), nucleic acids that perform this goal are separated from the unreacted molecules. As the successful molecules emerging from the selection are such a small proportion of the population, these sequences are amplified using PCR (polymerase chain reaction). Then the molecules enter the cycle again, where each repetition is considered a round.

This process is reiterated until the successful molecules are a larger proportion or activity is obtained.

In vitro selection is subjected to a number of constraints. First, the researchers need to estimate how long a DNA/RNA molecule should be to perform the desired catalysis or binding. This could bias the selection to molecules that are not the optimal length. Secondly, the separation technique may pose problems such as lack of efficiency and loss of sequences. Thirdly, some selections are contrived and may not reflect how the chemistry may be performed, such as catalysts that are tethered to their products. Fourthly, experimental error may lose pool sequences or prevent the selection from being performed optimally. This can lead to false negative results, when the selection might have actually been capable of finding catalysts if performed properly. The failure to obtain catalytic molecules for a given task may reflect poor design or execution of the project rather than an inability of nucleic acids to perform the desired chemistry.

In the past two decades there have been many catalytic DNA/RNA molecules discovered using this technique. The identification of selected ribozymes has drastically expanded the catalytic abilities of RNA beyond the natural backbone cleavage reaction. As well, *in vitro* selection has demonstrated that DNA molecules can catalyze a variety of reactions (reviewed in Peracchi, 2005). Many of the *in vitro* selections were undertaken to identify ribozymes that would support the RNA World hypothesis. Ribozymes have been shown to catalyze reactions that would have been important in an “RNA World”, such as amide bond synthesis (Lohse and Szostak, 1996), the formation of a nucleotide from a base and a sugar (Unrau and Bartel, 1998), synthesis of the cofactor acyl-coenzyme A (Jadhav and Yarus, 2002), and RNA-dependent RNA polymerization (Johnston *et al.*, 2001). While many metabolic steps crucial in the RNA World have not yet been shown to be catalyzed by RNA, the identification of ribozymes through *in vitro* selection has helped advance this theory.

2 NUCLEOTIDE SYNTHASE RIBOZYME SELECTIONS

2.1 4-Thiouridine Nucleotide Synthase Ribozyme Combinatorial Minimization Selection

2.1.1 Introduction

In 1998, the RNA World hypothesis gained further support from the finding that pyrimidine synthesis could be catalyzed by ribozymes (Unrau and Bartel, 1998). The goal was to isolate RNA molecules capable of catalyzing the synthesis of a pyrimidine nucleotide, in a manner similar to the enzyme uracil phosphoribosyltransferase (UPRTase, Figure 1a). The pool for this selection (containing approximately 1.5×10^{15} different sequences) was 228 nucleotides long, and comprised of three 76 nucleotide random-sequence regions (Figure 3a). The strategy used to accomplish isolation of ribozymes is shown in Figure 1b. To mimic the actions of UPRTase, these RNA molecules were modified to contain phosphoribosyl 1-pyrophosphate (pRpp) at the 3′ terminus. This activated group releases pyrophosphate when the base attacks the anomeric position. The selection was based on a separation technique that relied on the affinity of sulfur for mercury. By modifying uracil to replace a carbonyl with a thione, it retained the structure of the nucleotide while possessing a “tag” for capture. This activated pool RNA was incubated with 4-thiouracil (^{4s}Ura). Ribozymes that synthesized a tethered 4-thiouridine at the 3′ end were purified from unreacted RNA using the thiophillic mercury compound 4-[(*N*-

acryloyl)amino]phenylmercuric chloride (APM) in polyacrylamide gels (Igloi, 1988).

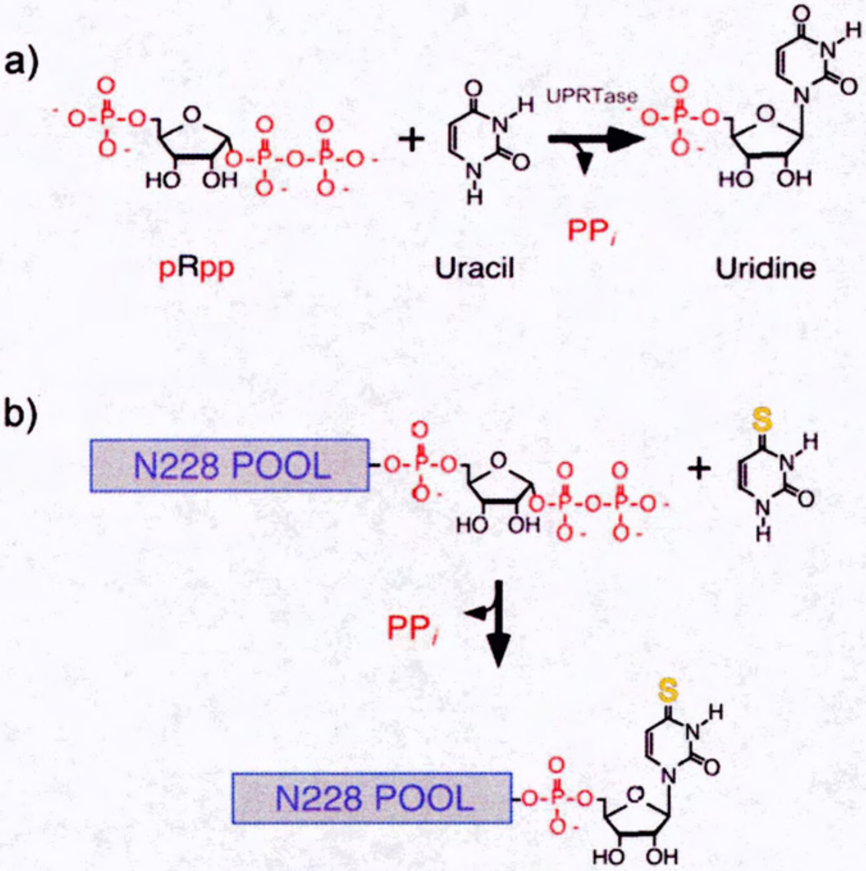
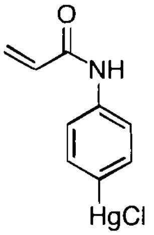


Figure 1: The goal (a) and strategy used (b) for the original selection for ribozymes capable of synthesizesing a pyrimidine nucleotide.
a) The goal was to find ribozymes capable of acting like the enzyme UPRTase, which synthesizes uridine from an activated ribose (pRpp) and uracil. b) To achieve this, a pool of RNA containing 228 random nucleotides (grey box, N denotes random sequence) was tethered at the 3′ end with pRpp. Incubation of this pool with 4-thiouracil allowed catalytic RNA to synthesize a 4-thiouridine nucleotide (including the sulfur tag, in yellow) at the 3′ terminus. Ribozymes were isolated by "capturing" the sulfur tag.

Sulfur is considered a soft Lewis acid, while mercury is a soft Lewis base. Soft acids and soft bases are known to have an affinity for one another (Rodgers, 1994). Igloi developed a way to incorporate the mercury compound in a polyacrylamide gel (Igloi, 1988). The structure of APM is shown in Figure 2. During the polymerization process, APM can be incorporated into the polyacrylamide gel because it contains the same acryloyl group as acrylamide. As a result, APM is crosslinked within the structure of the polyacrylamide gel (Igloi, 1988). The sulfhydryl group of 4-thiouracil has enough of an affinity for the mercury in APM that the ^{4s}U-tethered RNA sequences move slower in the gel than unmodified sequences.



APM

Figure 2: Structure of the thiophilic compound APM.

After eleven rounds of selection, 35 clones were isolated and sequenced. Analysis of these sequences showed the presence of three different "families" of related sequences, termed family A, B, and C. Isoates from each of these three families were capable of nucleotide synthesis with apparent second-order rate constants (k_{cat}/K_m, which describes the efficiency of the enzyme) over 10^7 fold greater than the uncatalyzed reaction. The specificity of these ribozymes was tested by incubation with other thio-substituted pyrimidine bases, however no product was detected. While reaction with uracil was possible, it was very slow. This indicates that this ribozyme was capable of discerning between a sulfur and oxygen in the 4-position.

The random pool used for this selection is a relatively long sequence. It can be advantageous, and sometimes necessary, to use such a long pool when trying to isolate ribozymes capable of catalyzing challenging reactions. This is because as the size of a pool increases, the ways the core elements of a potential ribozyme motif can be arranged increase "combinatorially" (Sabeti *et al.*, 1997). However, regions that are not essential for catalysis can be interspersed with important segments. To identify these nonessential segments, sequences of catalytically active isolates can be compared (Ellington and Szostak, 1990; Ekland and Bartel, 1995). Based on these principles, a new selection was designed to define the core motif of this ribozyme.

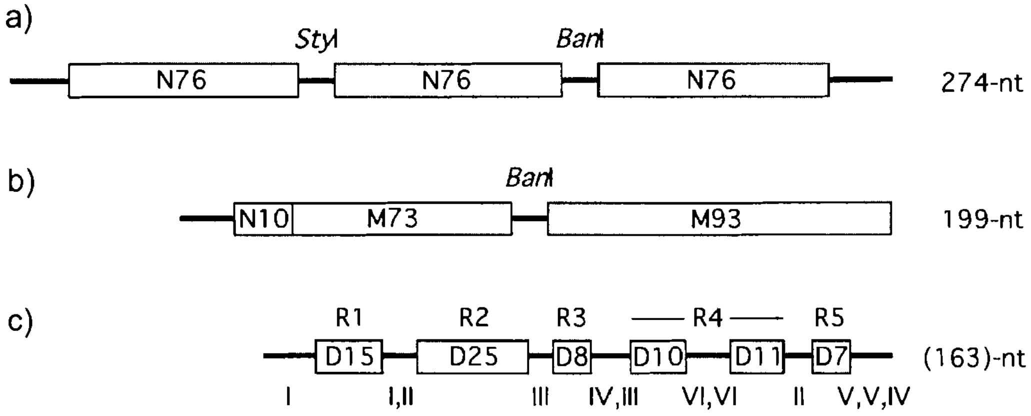


Figure 3: RNA pool design of pyrimidine nucleotide synthase ribozyme selections. Lines indicate fixed RNA sequence regions, while boxes represent degenerate or random RNA sequences. a) Original random-sequence pool for the initial selection (Unrau and Bartel, 1998). Three random sequence regions (N76) were joined by *Sty*I and *Ban*I restriction sites. b) degenerate-sequence pool based on isolate a15 from the initial selection (a). 73 and 93 nt degenerate regions (M73 and M93) contained 20% mutation levels from a15. N10 denotes a 10 nt random region. c) Variable sequence and length pool for the combinatorial minimization selection, based on the isolate a.6.10 (163 nt) from the degenerate selection (b). Two pools were synthesized, each containing five regions (R1-R5) containing both deletions and mutations. The mutation frequency was fixed at 20% and the deletion frequencies were either 17% or 25% depending on the pool. The sequences comprising the ribozyme helical regions stems (I through VI) were kept constant. The arms of these helices are denoted by Roman numerals. Equimolar amounts of the two pools were combined to created a selection pool with approximately 5 × 10^{13} unique sequences. Reproduced in part with permission from RNA, volume 9, number 10, page 1209. Copyright 2003 The RNA Society.

The strategy was to first remove useless sequence, and then create a pool with mutations all the way to the 3′ end. The sequence that served as the template for pool construction was isolate a15 (from family A) from the original selection. As previous truncation experiments demonstrated 98 nucleotides at the 5′ end could be removed with no effect on activity, the selection pool had this portion of sequence omitted. The pool of RNA molecules for this selection contained portions of fixed, random and mutagenized (at a frequency of 20%) sequence (Figure 3b). The 199 nucleotide construct had two separate degenerate regions (joined by a *BanI* restriction site), such that 80% of the residues would feature the same sequence as a15, and 20% would be composed of the other 3 possibilities. The RNA pool had variation all the way to the 3′ end, allowing the sequence to be tested just adjacent to the site of attachment. The selection was completed after six rounds, generating much faster ribozymes than the parent version. There was an approximately 25-fold gain in efficiency in the fastest versions of this selection over a15 (Chapple *et al.*, 2003). The sequences of 30 clones isolated from this selection were aligned, and a secondary structure model for these ribozymes was determined. This featured at least five helical regions, three loop regions, and seven interhelical joining segments. Despite the diversity of sequences, each were able to able to adopt the same fold. The inferred secondary structure of one specific isolate is shown in Figure 4.

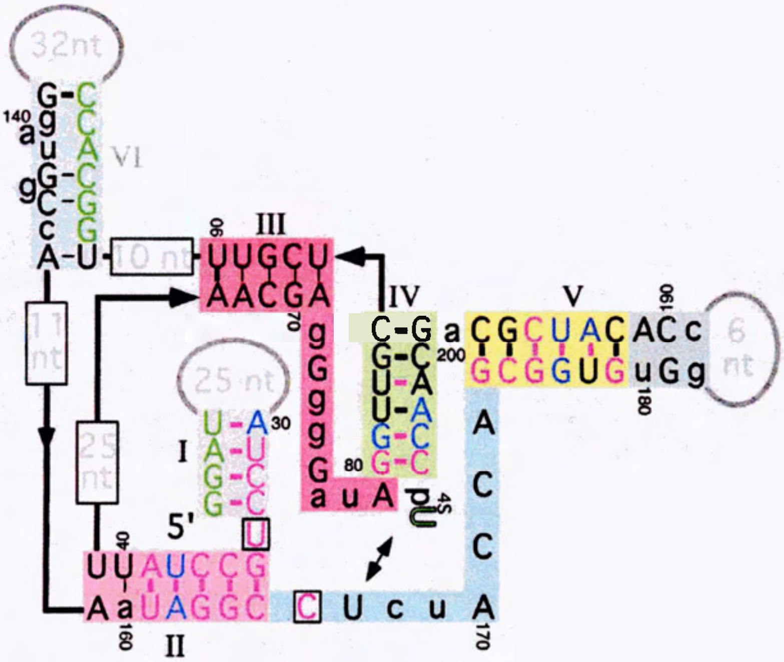


Figure 4: Isolate a.6.10 from the degenerate selection.
The ribozyme is shown containing the product of the reaction, ^{4s}U. Double-pointed arrow indicates the observed crosslink between this product and residue U167. Helical stems are indicated by roman numerals I-VI. Nucleotide colours: green, sequence fixed according to pool construction; magenta, no sequence variation observed (probability due to chance, $P = 0.0012$); blue, always covaries with partner (includes G-U wobble pairs). Uppercase black, ≤ 4/30 variations from consensus; lowercase black, > 4/30 variants from consensus. Conserved potential for pairing is indicated as follows: thick purple bar, always base-paired ($P = 3.6 \times 10^{-5}$ if both residues can vary); thick black bar, 29/30 bp ($P = 4.8 \times 10^{-4}$); thin black bar, 26/30 or more bp ($P<0.04$). Reproduced in part with permission from RNA, volume 9, number 10, page 1211. Copyright 2003 The RNA Society.

With this information about the core motif of the ribozyme, a new selection was designed to create a more efficient, compact version. The strategy for this selection was to simultaneously remove regions of sequence nonessential for catalysis, and introduce mutations to improve function. Therefore, design of the pool needed to contain both targeted deletions and mutations in a combinatorial approach. Nonessential regions such as sequence at the extreme 5′ and 3′ ends of the ribozyme and in loops can generally be successfully deleted. However the deletion of interhelical joining regions can be more challenging, as a minimum length is required to retain the overall secondary structure of the ribozyme.

The sequence used as a template for the selection is a variation of the a.6.10 ribozyme from the degenerate sequence selection. Several changes were made to a.6.10, dropping its size from 204 to 163 nucleotides. The 32 nucleotide loop at the end of helix V was replaced with a tetraloop. Three base pairs at the end of helix VI and a further six nucleotides in the loop were also replaced with a tetraloop. The loop that closes stem I was not reengineered since this sequence had been shown to be important for ribozyme function. Instead, this region was subjected to mutation and deletion variability. Two further changes based on previous consensus information were made: repairing the G-U wobble in stem V, and reversing the stem IV base pair from 203C:81G to 203G:81C. As a result of these changes, the pool construct (Figure 3c) was comprised of seven fixed sequence regions and six regions subject to mutation and deletion. Two separate pools were constructed for this selection, both featuring mutation frequencies throughout these six segments of approximately 20% but featuring either 17% or 25% deletion frequencies. These two pools were then combined for the selection.

2.1.2 Materials and Methods

2.1.2.1 General

α-UTP (800 Ci/mmol) and γ–ATP (6000 Ci/mmol) were purchased from New England Nuclear/Perkin Elmer. T4 RNA ligase was from Amersham. T7 RNA polymerase was from USB. T4 polynucleotide kinase (and corresponding 5× Exchange and 5× Forward buffers), basepair ladders (both 10 and 100), and Superscript II reverse transcriptase were purchased from Invitrogen. tRNA was from provided as a gift from the Brandhorst lab. Taq polymerase was a laboratory preparation. APM, AppRpp and App^{4s}U were prepared by Peter Unrau as published (Unrau and Bartel, 1998).

DNA was synthesized on an Applied Biosystems (ABI) 392 DNA/RNA synthesizer using 0.2 μmol 2000 Å control pore glass (CPG) columns from Glen Research. Individual dA,dC,dG, and dT phosphoramidites were purchased from Applied Biosystems. PCR experiments and temperature-controlled incubations were performed using a PTC Thermal Controller from MJ Research Inc. UV-visible absorbance spectra were obtained using a Shimadzu UV-1601 spectrometer. Radioactive gels were visualized using the Storm Phosphorimager system (Amersham) and quantitated with ImageQuant and ImageTools software (Amersham). Centrifugations were undertaken using a Eppendorf 5415 D centrifuge. Ethanol precipitations were performed by adjusting the volume of the nucleic acid sample to 300 mM sodium chloride, adding 2.5× of the total volume of ethanol, vortexing, and then cooling at -20° for at least 20 minutes. Pellets were obtained by centrifuging tubes at 13,200 rotations per minute for at least 20 minutes at room temperature.

2.1.2.2 Selection Protocol

DNA synthesis, sequences, and pool construction were as described in Chapple *et al.*, 2003. DNA was transcribed into RNA in the presence of [^{32}P]-α-UTP, 40 mM Tris pH 8.6, 26 mM $MgCl_2$, 10 mM dithiothreitol (DTT), 2.5 mM spermidine, 0.01% Triton X-100, 8mM GTP, 4 mM each of ATP and CTP, 2 mM UTP, 5 U/μL T7 RNA polymerase at 37° for 1 hour. Transcription was terminated by the addition of an equal volume of denaturing gel loading buffer (50 mM EDTA, 90% formamide) and the transcript was purified by 6% PAGE. The transcript was excised and eluted from the gel in 300 mM NaCl, and then ethanol precipitated. The RNA was then ligated to pRpp using adenylated pRpp (AppRpp) and T4 RNA Ligase. Ligation conditions were 50 mM HEPES pH 8.3, 10 mM $MgCl_2$, 3.3 mM DTT, 8.3% v/v glycerol, 0.01 mg/mL BSA, 65 μM AppRpp, 1 U/μL ligase for 4 hours at 23°. Ligations were stopped by adjusting the reaction to 4 mM EDTA, phenol chloroform extracted, and ethanol precipitated. pRpp-ligated RNA was prepared for incubation with ^{4s}Ura by heating to 80° for 1 minute, cooling to 4° and leaving at 23°. Incubation conditions were ~1 uM RNA, 50 mM Tris-HCl pH 7.6, 150 mM KCl, 25 mM $MgCl_2$. Incubation times and ^{4s}Ura concentrations varied per round and were as follows: R1, 120 min, 1 mM; R2, 12 min, 0.2 mM; R3, 0.5 min, 0.2 mM; R4, 0.1 min, 0.1 mM; R5, 0.1 min, 0.1 mM; R6, 2 min, 1 mM; R7, 0.25 min, 1 mM; R8, 1 min, 1 mM. Incubation solutions were terminated by addition of a 7-fold excess of denaturing gel loading buffer (50 mM EDTA, 90% formamide, 10 μM tRNA) containing tRNA (to prevent streaking in the gel) and analyzed on a 6% denaturing polyacrylamide gel containing ~40 μM *N*-acryloylaminophenylmercuric acetate (APM). In early rounds a radiolabelled marker molecule was ligated to ^{4s}U to denote the shortest possible sequences. After each round, the DNA derived from the RT-PCR was transcribed and the rates were compared between longer and shorter RNA molecules. The shortest half was carried on for selection. In rounds 6, 7 and 8, the shortest half of reacted ribozymes were excised from the mercury gel and carried forward. Reacted RNA was excised from the gel, eluted, and ethanol precipitated. RNA was reverse transcribed in the presence of a 16-nt RT primer (5′-gcttgctgcggtgcga) and Superscript II reverse transcriptase at 48° for 1 hr. The solution was heated with 100 mM KOH, neutralized with Tris-HCl, and the resultant DNA PCR-amplified as described (Chapple *et al.*, 2003).

2.1.2.3 Kinetics

The incubation conditions for the time courses were the same as those used in selections and featured 1 mM ^{4s}Ura. At least five timepoints were used for each assay. Timepoints (containing 10 uM tRNA) were analyzed on 6% analytical PAGE containing ~40 uM APM. Gels were visualized using the Storm phosphorimaging system (Amersham) and quantitated using the ImageQuant program (Amersham). Rates were determined using Kaleidagraph software by fitting k and β to the equation $F = \beta(1-e^{-kt})$ where F is the fraction of RNA molecules reacted, *k* is the apparent first-order rate, and β is the fraction of RNA molecules able to react.

One ribozyme clone (7.02) was chosen for kinetic analysis. In this experiment the concentration of ribozyme was held constant and the rates of nucleotide synthesis recorded for 6 different ^{4s}Ura concentrations. The ^{4s}Ura concentrations chosen were 0.1, 2, 4, 6, 10 and 12 mM. Each time course consisted of 6 timepoints and were analyzed as in the previous section, with all rates possessing R^2 values of over 0.99. The graph of rates vs ^{4s}Ura concentration did not fit a standard Michaelis-Menten curve, but instead featured a linear increase and then linear decrease. This experiment was repeated 6 times all with similar results.

2.1.2.4 Cloning

Pool DNA from rounds 4-7 were cloned into *E. coli* using the TOPO TA Cloning kit version L (012601) and plated onto LB/ampicillin plates. Individual colonies were picked and the sequences PCR amplified using the primers M13F (5′-GTT TTC CCA GTC ACG AC) and M13R (5′-CAG GAA ACA GCT ATG AC). Primers were removed from the PCR products using QIAquick spin columns (QIAGEN). Samples were sent to the University of Calgary DNA Sequencing Laboratory for sequencing. Sequences were aligned using MegAlign.

2.1.3 Results and Discussion

A total of eight rounds of selection were completed according to the scheme in Figure 5. I completed the final three rounds of this selection. For every following round, the band of reacted sequences in the APM gel were divided into upper and lower halves, representing shorter and longer sequences, respectively. These two halves were separately excised from the mercury gel, reverse transcribed, and amplified. In order to successfully select both smaller and faster ribozymes, the rates of both shorter and longer molecules were compared. As long as the shorter ribozymes were as fast or faster, they were carried through to the next round. Activity of the pool was first observed in round four. The eighth round was the final one because shorter ribozymes were slower than the longer ones. At this point it became evident that the decrease in size was detrimental to ribozyme activity. The seventh round was considered the most successful compromise between size and activity. Several short isolates from this round were individually examined to determine their catalytic ability. These isolates displayed first-order rate constants in the range of 0.04-0.13 min^{-1} (using ^{4s}Ura concentrations of approximately 1 mM).

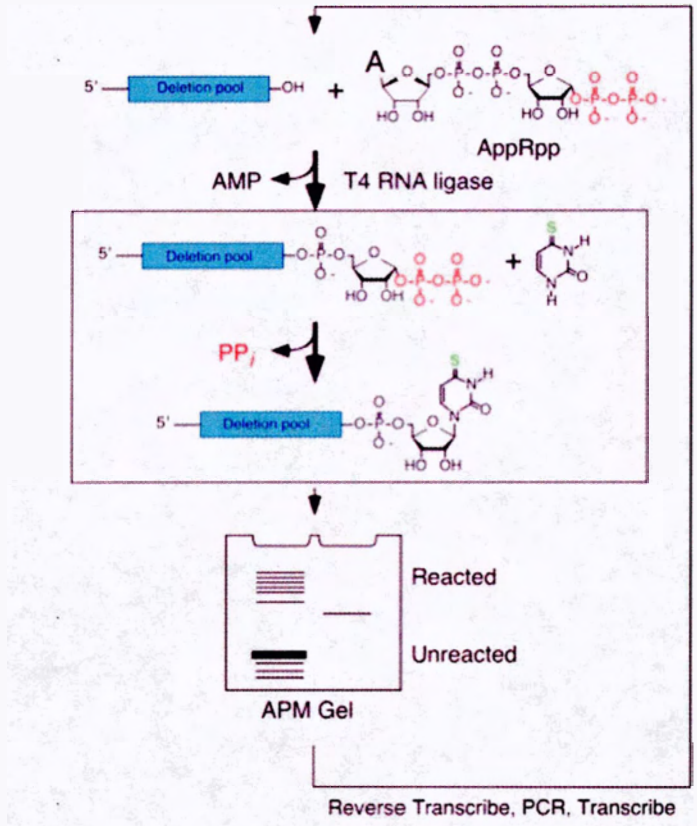


Figure 5: Selection scheme for the deletion selection.
The first step ligates pool RNA to pRpp using T4 RNA ligase, which recognizes AMP. This activated RNA is incubated with 4-thiouracil containing a sulfur as a tag (shown in green). Catalytic RNA molecules synthesize the glycosidic bond of the nucleotide at their 3′-terminus, releasing PP*i* in the process (shown in red). Reacted sequences are separated from unreacted in an APM (mercury containing) polyacrylamide gel and visualized using autoradiography. In early rounds when the signal was too low for detection, a marker indicating the size of reacted ribozymes was used (right lane). Catalytic sequences are converted into DNA, amplified using PCR, and are transcribed into RNA to begin the next round.

After each of round 4, 5, 6, and 7, pool DNA was cloned and approximately 20 isolates were sent for sequencing. In addition to the optimal compromise between size and activity of round seven, another signal the selection was over was the fact one sequence appeared twice. The isolation of one unique sequence more than once can signal the selection is over. To obtain in-depth information about the selection progress at the sequence level, clones from rounds 4,5, 6, and 7 were aligned in relation to the parent sequence 163.1. The results of this alignment are shown in Appendix I. The seven constant regions (I-VI) serve as the consensus sequence in the alignment. These seven constant regions are colour coded according to the secondary structure of the ribozyme, as demonstrate by the isolate da.7.02 (Figure 6). 7.02 was a particularly short ribozyme, only 124 nucleotides. Interestingly, 7.02 experienced a deletion of stem VI yet was able to retain catalytic activity.

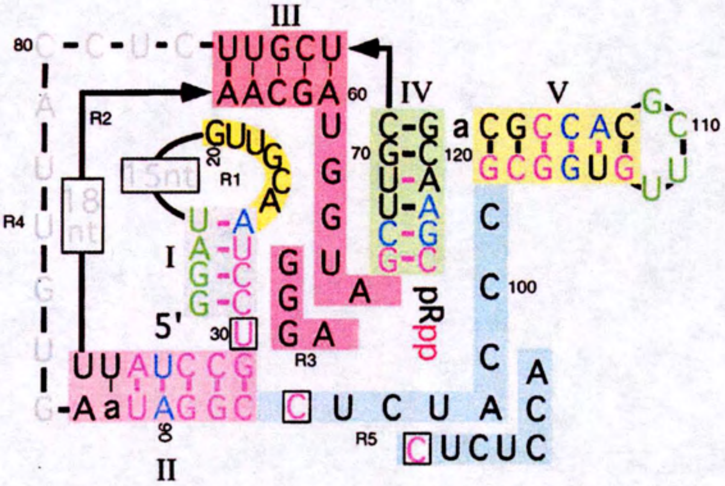


Figure 6: Proposed secondary structure of isolate 7.02.
Sequence conservation in loop I is highlighted in yellow. Nucleotide colours: green, sequence fixed according to pool construction; magenta, no sequence variation observed (probability due to chance, $P = 0.0012$); blue, always covaries with partner (includes G-U wobble pairs). Uppercase black, ≤ 4/30 variations from consensus; lowercase black, > 4/30 variants from consensus. Residues showing poor sequence conservation are coloured grey. Conserved potential for pairing is indicated as follows: thick purple bar, always base-paired ($P = 3.6 \times 10^{-5}$ if both residues can vary); thick black bar, 29/30 bp ($P = 4.8 \times 10^{-4}$); thin black bar, 26/30 or more bp ($P<0.04$). Reproduced in part with permission from RNA, volume 9, number 10, page 1211. Copyright 2003 The RNA Society.

These clones provided valuable information about the size and sequence identity of selected ribozymes. The size steadily decreased in rounds 4 through 7 as shown in Figure 7a. These intervening variable regions (R1-R5) were examined for deletions and mutations. The deletion frequency (averaged for the 20 clones) and deletion frequency variance was plotted versus selection round to observe how the regions accommodated deletions over the course of the selection (Figure 7b, c). Except for R5, these regions displayed a linear increase in deletion frequency over these four selection rounds. The low deletion frequency of R5 is understandable, considering its proximity to the ribozyme active site. It is apparent that the length of R5 was essentially optimized before these four rounds of selection. R5 displays the consensus sequence ***C****UCUCCAGC*, which differs from the a.6.10 sequence ***C****UCUACCAGC* (bold indicating the absolutely conserved C166; italics represent the variable region of the pool) by the elimination of one adenosine.

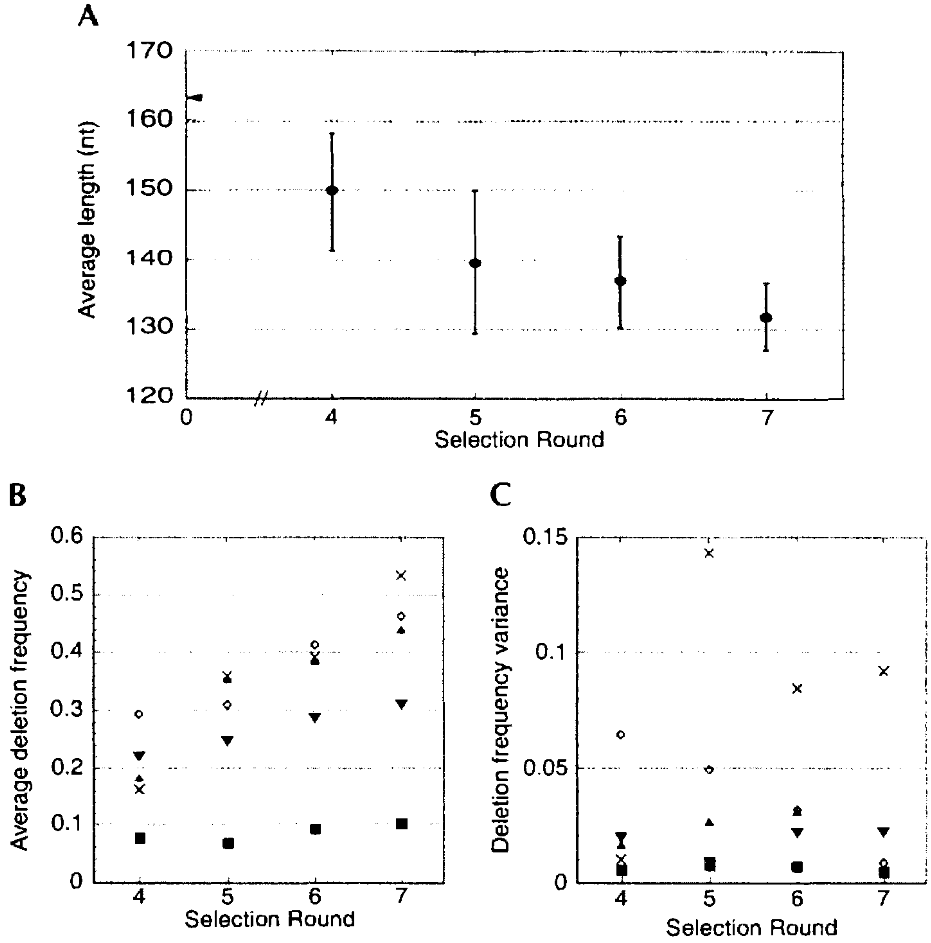


Figure 7: Progression of length and deletion through the selection process. a) Length of cloned ribozymes over several rounds of selection. The numbers of cloned ribozymes were 17, 19, 20 and 20 for rounds 4,5,6 and 7, respectively. The full length template sequence (163 nt) is indicated on the y-axis with an arrow. b) Deletion frequencies of the five variable regions of clones from Appendix I plotted against selection round. c) Deletion frequency variance for variable regions over the course of the selection. Symbols for (b) and (c) are as follows: (▼), R1; (▲), R2; (◇), R3; (×), R4; (■), R5. Reproduced in part with permission from RNA, volume 9, number 10, page 1214. Copyright 2003 The RNA Society.

To get a sense of how these ribozymes compared to the original 4-thiouridine selection, we attempted to obtain the Michaelis constant (K_m) of this ribozyme. The K_m is defined as the concentration of substrate at one-half the maximum reaction velocity. To obtain this value, a series of time courses are undertaken in which catalyst concentration is kept constant and substrate concentration is increased. By plotting reaction rate of each time course versus substrate concentration, the K_m can be extrapolated. As round seven was the best compromise between size and efficiency, we reasoned an isolate from this round would possess the most favourable K_m. The isolate chosen was da.7.02, which was slightly faster compared to other isolates. Another interesting aspect of this sequence was the fact it was identified twice during sequencing analysis of round seven. In order to calculate the K_m of the da.7.02 ribozyme, a series of time courses were undertaken keeping ribozyme concentration constant and varying the ^{4s}Ura concentration from 0.1 mM to 12 mM.

Instead of displaying typical Michaelis-Menten behaviour, the rate increased almost linearly from 0.1 mM to 4 mM ^{4s}Ura, where it reached a maximum apparent first-order rate constant of 0.21 min^{-1}, before dropping (also with linear character) to 0.08 min^{-1} at 12 mM ^{4s}Ura. The time courses over these concentrations produced rates with R^2 values consistently over 0.99, demonstrating each substrate concentration was capable of producing reliable rates. Also, this experiment was completed many times giving reproducible results. The failure of the ribozyme to maintain a constant rate over high ^{4s}Ura concentrations was rationalized to be due to substrate inhibition, as seen in earlier selections. Therefore measurement of the K_m of 7.02 was not possible. However, comparison of ribozyme efficiencies between selections was possible. Isolate da.7.02 displayed an efficiency of 150 $M^{-1}min^{-1}$ at 0.1 mM ^{4s}Ura, approximately 35-fold faster than the a15 isolate (4.3 $M^{-1}min^{-1}$) from the original

4-thiouridine synthase ribozyme selection. The da.7.02 isolate was slightly faster than isolate a.6.10 (120 $M^{-1}min^{-1}$) from the secondary structure selection.

2.1.4 Conclusion

This selection allowed the minimization of pyrimidine nucleotide synthase ribozymes, while increasing activity. The *in vitro* selection technique allowed us to simultaneously isolate the shortest and most active sequences. This approach avoided the laborious design of individual variants (containing both deletions and mutations), and testing the activity of each. This selection produced unexpected isolates which lacked one of the core helices. As expected, minimization eventually took place at the expense of catalysis. Seven rounds of selection produced an isolate approximately 35 times faster than one from the original selection. However, much of the gains in efficiency had come from the preceeding degenerate sequence selection. The modest increase in efficiency from the combinatorial minimization selection is especially impressive considering the decrease in ribozyme length from 272 to 124 nucleotides.

2.2 Novel Nucleotide Synthase Ribozyme Selections

2.2.1 Introduction

After the conclusion of the combinatorial minimization selection of the 4-thiouridine nucleotide synthase ribozyme, our attention turned to initiating new selections. As RNA had been shown to catalyze the formation a pyrimidine nucleotide, was it possible for it to form purine ones? Another pressing question was whether RNA could catalyze the formation of other necessary glycosidic linkages besides those in pyrimidine and purine nucleotides. Identifying novel nucleotide synthase ribozymes would have large implications on the RNA World hypothesis, as the RNA enzymes of the prebiotic world would need to make metabolically relevant molecules such as nucleotides.

The intent was to use the same selection protocol for the 4-thiouridine nucleotide synthase selection. It would be the logical method since the conditions had already been optimized. The critical step in each selective round is the use of a mercury gel to separate reacted from unreacted RNA molecules. Therefore, the design of any novel selection requires a base analogue containing a sulfur tag. Before beginning any selection, an RNA construct tethered to the hypothetical nucleotide product must be generated. This sequence must be run into an APM gel to see if its migration is slowed (relative to unmodified sequences) sufficiently for isolation. Thus, we could test the feasibility of this selection scheme before even beginning the selection. If not possible by this scheme, another separation method such as affinity chromatography would be necessary.

We decided to use two very different random sequence pools for any selection for a novel nucleotide synthase ribozyme. There were two types of pools designed for these selections. The first is a completely random sequence of a defined length. This is a common approach when selecting for novel ribozymes and DNAzymes. A second approach we considered was a pool based on the sequence of the 4-thiouridine selection ribozyme. In this latter design, the assumption is made that the secondary structure adopted by the pyrimidine nucleotide synthase ribozymes is conducive to the formation of other types of nucleotides. The sequence chosen for this is that of isolate da.7.02, from the deletion selection. The five helical regions were kept semi-conserved, with mutations occurring at a rate of 10 and 20%. We designed two different mutated pools, each with completely random segments in the same spots, but the remaining mutated sequences being either 80% or 90% wildtype (based on the sequence of 7.02). The sequences in between these helical regions allowed for complete randomization. By using two types of pools, it would allow us to find novel sequences capable of catalysis while also testing the hypothesis that the structure of the minimized 4-thiouridine nucleotide synthase ribozyme is conducive to the catalysis of other glycosidic linkages.

In modern metabolism, the synthesis of GMP utilizes the activated ribose sugar pRPP analogous to the UMP synthesis (Figure 8). The enzyme responsible for this is hypoxanthine guanine phosphoribosyltransferase (HGPRTase, EC 2.4.2.7). Purines, in addition to pyrimidine bases, have been shown to be made in prebiotic processes (Joyce, 1989; Orgel, 1998). Therefore, RNA enzymes capable of making purine nucleotides could have made use of these as substrates.

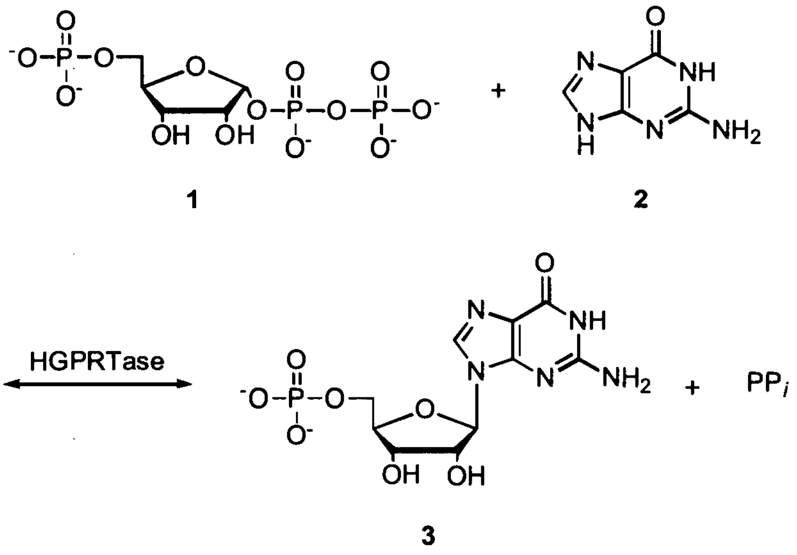
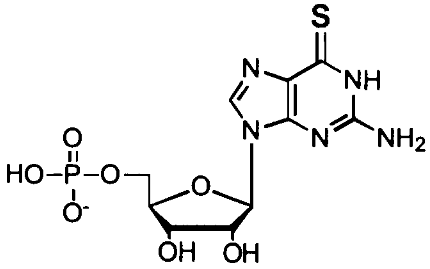


Figure 8: Biosynthesis of GMP.
Hypoxanthine guanine phosphoribosyltransferase (HPRTase, EC 2.4.2.7) catalyzes the synthesis of GMP from pRPP and guanine.

The goal was to find ribozymes capable of creating a purine nucleotide from pRPP and guanine, thus mimicking HGPRTase. The target nucleotide chosen was 6-thioguanine monophosphate (Figure 9), which contains a thiocarbonyl like 4-thiouridine. Fortunately, 6-thioguanine was commercially available.



4

Figure 9: 6-Thioguanosine monophosphate

The synthesis of the marker for the purine nucleotide synthase ribozyme relied on the use of chemical and enzymatic methods (Figure 10). The route chosen was to prepare an AMP-linked 6-thioguanosine monophosphate molecule, to act as the substrate for the final T4 RNA ligase step. The first step would be the synthesis of the nucleotide 6-thioguanosine monophosphate, which is not commercially available. This molecule could be synthesized by chemically phosphorylating 6-thioguanosine. The second step would be generation of the adenylated construct, App^{6s}G. This could be done by reacting ^{6s}GMP with imidazole activated AMP. Once prepared, the ^{6s}G portion of App^{6s}G can be ligated onto the end of test RNA using T4 RNA ligase.

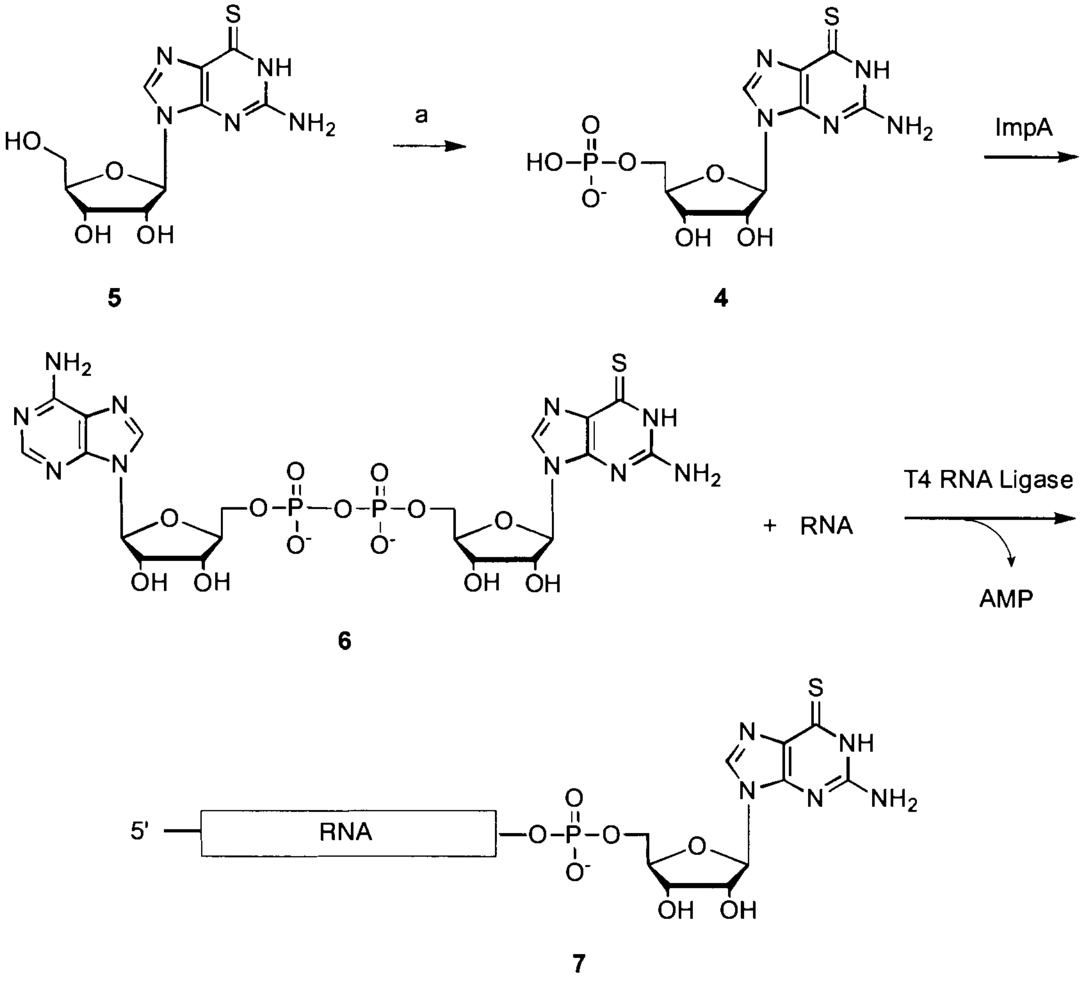
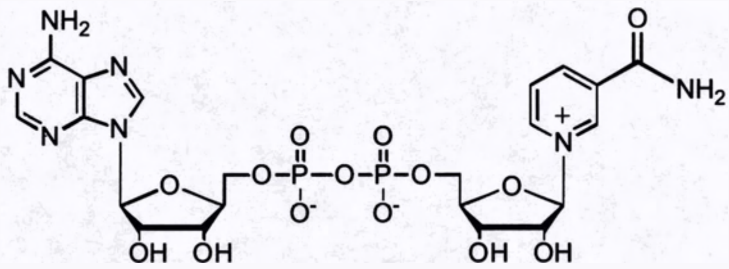


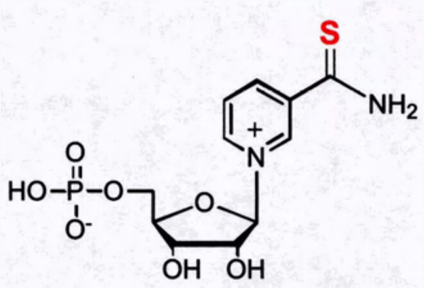
Figure 10: Overview of 6-thioguanosine monophosphate marker synthesis.
Commercially available 6-thioguanine (**5**) could be phosphorylated in the 5′ position using a chemical reagent such as phosphorous oxychloride. 6-thioguanosine monophosphate (**4**) can be reacted with ImpA, generating adenylated ^{6s}GMP (**6**). Ligation of **6** with a suitable RNA sequence would create RNA tethered at the 3′ end to ^{6s}GMP (**7**). The RNA construct **7** mimics the theoretical end product of potential ribozymes.

The second hypothetical selection features the pyridine nucleotide portion of the metabolism cofactor nicotinamide adenine dinucleotide (NAD$^+$) (Figure 11a). The selection would attempt to find ribozymes that synthesize thionicotinamide mononucleotide (Figure 11b). NAD$^+$ is composed of nicotinamide mononucleotide (NMN) joined to AMP through their respective phosphate groups (as a pyrophosphate linkage). Thionicotinamide, a sulfur analogue of nicotinamide, is widely available and would act as the base substrate. As it contains a thiocarbonyl group, it could conceivably shift in a mercury gel; however, the thiocarbonyl is not in the ring as in 6-thioguanine and 4-thiouridine. This may result in different tautomerization behaviour and unsuccessful shifting in a mercury gel. As thioNMN is a sulfur analogue of part of an important enzyme cofactor, a selection of this nature would further expand the range of metabolically important substrates whose formation is catalyzed by RNA.



a

8



b

9

Figure 11: a) NAD+ and b) thionicotinamide mononucleotide

The marker molecule would be easily attainable for the thionicotinamide synthase selection scheme. Thionicotinamide adenine dinucleotide contains the AMP portion recognized by the enzyme T4 RNA ligase. By using thioNAD$^+$ as a substrate, T4 RNA ligase would transfer thionicotinamide mononucleotide to the

3′ end of marker RNA (Figure 12). Therefore, the RNA-nicotinamide mononucleotide construct would be attainable in just one step. The RNA chosen would need a 3' C nucleotide as it is a requirement of T4 RNA ligase. Also, the size of RNA should be approximately the same size as the pool to more accurately represent the hypothetical result of the selection.

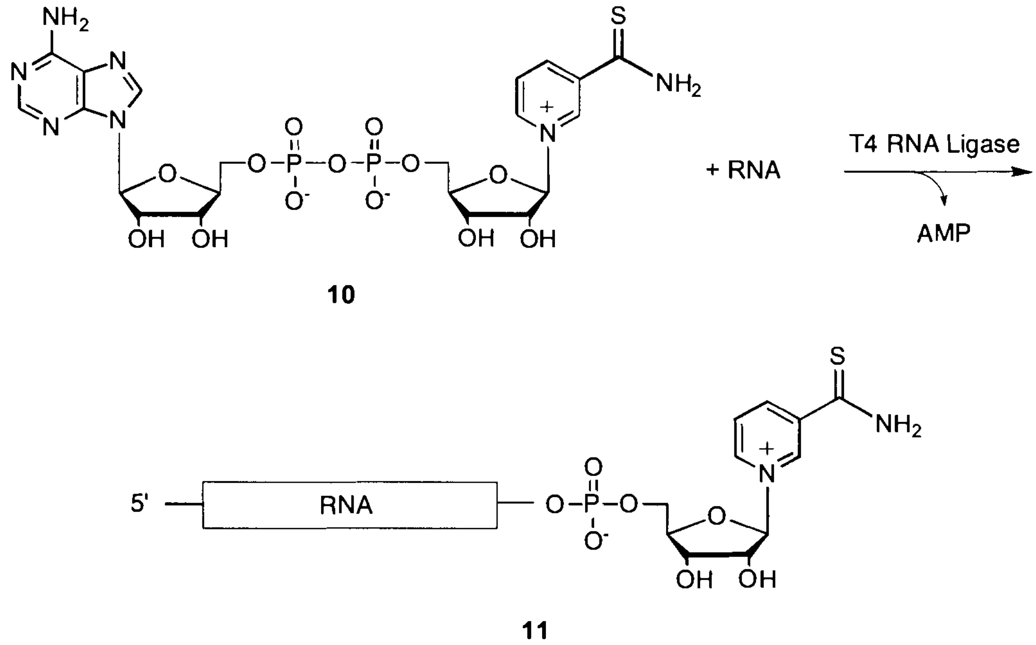


Figure 12: Overview of thionicotinamide mononucleotide marker synthesis. Thionicotinamide adenine dinucleotide (thioNAD+, **10**) serves as a substrate for T4 RNA ligase, which transfers thionicotinamide mononucleotide to the 3' end of RNA. This construct (**11**) serves as the hypothetical end product of the selection and can be used to test shift ability in the APM gel.

2.2.2 Materials and Methods

2.2.2.1 General

6-thioguanosine, phosphorous oxychloride, and thionicotinamide adenine dinucleotide were purchased from Sigma. α-UTP (800 Ci/mmol) and γ–ATP

(6000 Ci/mmol) were purchased from New England Nuclear/Perkin Elmer. T4 polynucleotide kinase (and corresponding 5× Exchange and 5× Forward buffers) and basepair ladders (both 10 and 100) were purchased from Invitrogen. tRNA was from provided as a gift from the Brandhorst lab. A-25 DEAE Sephadex was purchased from Sigma. APM, ImpA, AppA, and App^{4s}U were prepared by Peter Unrau (Unrau and Bartel, 1998). GpC and ApApC oligos were from Sigma. T7 RNA polymerase was a laboratory preparation. T4 RNA ligase was prepared as published (Wang and Unrau, 2002).

PCR experiments and temperature-controlled incubations were performed using a PTC Thermal Controller from MJ Research Inc. HPLC analysis and purification was performed using a Vydac C18 reverse phase column and a Waters HPLC using Millenium software. UV-visible absorbance spectra were obtained using a Shimadzu UV-1601 spectrometer. Radioactive gels were visualized using the Storm Phosphorimager system (Amersham) and quantitated with ImageQuant and ImageTools software (Amersham). Centrifugations were undertaken using either a Eppendorf 5415 D centrifuge (for 1.5 mL volumes or lower) or with a Sorval HB-4 rotor. Ethanol precipitations were performed by adjusting the volume of the nucleic acid sample to 300 mM sodium chloride, adding 2.5× of the total volume of ethanol, vortexing, and cooling at -20° for at least 20 minutes. Pellets were obtained by centrifuging tubes at 13 200 rotations per minute for at least 20 minutes at room temperature. Cellulose TLC plates containing fluor were purchased from J.T. Baker. Mass spectra were obtained using a MALDI-

TOF MS spectrometer (PerSeptive Voyager DE STR from PE Applied Biosystems). Nuclear magnetic resonance spectra (proton and phosphorous) were obtained on a Bruker AMX 400 FT spectrometer (operating frequencies: 1H, 400.13 MHz; 31P, 161.98 MHz) at ambient temperature. Chemical shifts (δ) for all compounds are listed in parts per million downfield from tetramethylsilane using the residual non-deuterated solvent peak as an internal reference.

2.2.2.2 Purification of thioNAD$^+$ (10)

Thionicotinamide adenine dinucleotide (5 mg) was dissolved in 50 mM HEPES (pH 7.4) to make a 50 mM solution and was titrated to neutrality with sodium hydroxide. A 2.5×10^{-4} mmol aliquot of this impure solution was purified by HPLC using a C18 column and monitoring absorbance at 260 nm. Elution conditions were 50 mM triethyl ammonium acetate (TEAA) pH 6.5 in 50% acetonitrile, 50% water followed by a linear gradient to 50 mM TEAA pH 6.5 in 25% acetonitrile, 75% water. The eluent was lyophilized to give an orange solid that was resuspended at 808 μM in 50 mM HEPES pH 7.4. Due to the presence of buffer salts, the yield of product obtained (81%, 2.02×10^{-4} mmol) was determined spectrophotometrically using $\varepsilon(259) = 16\ 700\ M^{-1}cm^{-1}$ (Joppich-Kuhn and Luisi, 1978).

2.2.2.3 6-Thioguanine 5′-monophosphate (4) synthesis and purification

6-Thioguanosine was phosphorylated in the 5′ position with adaptations as described in the literature (Breter and Mertes, 1990). 6-Thioguanosine (30 mg, 100 μmol) was stirred in triethyl phosphate under a nitrogen atmosphere at 4 °C. 1 equivalent of phosphorous oxychloride (9 uL, 100 μmol) was added and the solution stirred an additional 2 hours. The phosphorous oxychloride was evaporated in vacuo and 50 mL chloroform was added to the remaining solution to produce a flocculate. This was centrifuged (5 000 rpm, Sorval HB-4 rotor, 25 minutes) and the supernatant removed. The pellet was resuspended in water and the acidic solution neutralized with NaOH. This crude reaction mixture was analyzed by TLC (cellulose TLC plates presoaked in saturated NH_4SO_4, developed with 80% ethanol). The product was purified on a A-25 DEAE Sephadex column using a 50 mM to 400 mM triethylammonium bicarbonate (TEAB) pH 7.5 gradient at 4°. 6-Thioguanosine 5'-monophosphate eluted at 250-300 mM TEAB. Absorbance of fractions at 342 nm were monitored and those containing the product were combined and lyophilized to a yellow solid. One distinct elution before the product with a much smaller absorbance at 342 nm was attributed to unreacted starting material and not analyzed. To facilitate the removal of volatile buffer salts, the sample was repeatedly resuspended in water and re-lyophilized. The yield (56%) was determined spectrophotometrically using the extinction coefficient $\varepsilon(342) = 24\ 800\ M^{-1}cm^{-1}$ (Fox, 1958). HPLC analysis revealed the product to be 99.1% pure (conditions in 2.2.2.2). Verification of the product was assessed through MALDI-TOF spectrometry (m/z 382) and ^1H and

^{31}P NMR. **^1H NMR** (400 MHz, D_2O) δ 3.98-4.02 (m, 2H), 4.28-4.32 (m, 1H), 4.47 (dd, J = 5.1, 3.3 Hz, 1H), 5.91 (d, J = 6.1 Hz, 1H), 8.30 (s, 1H); **^{31}P NMR** (162 MHz, D_2O) δ 3.66 (br s). Some of this was dissolved in water to prepare a 502 mM stock (which was divided into aliquots) for App^{6s}G preparation. ^{6s}G and the ^{6s}G stock in water were kept at -80° until further use.

2.2.2.4 App^{6S}G (6) preparation and purification

Adenylated 6-thioguanosine-5′-monophosphate (App^{6S}G) was synthesized by reacting 100 mM (1.5×10^{-3} mmol) ^{6S}GMP (2.2.2.3) with 50 mM (7.5×10^{-4} mmol) adenosine 5′-phosphorimidazolide in the presence of 100 mM (1.5×10^{-3} mmol) $MgCl_2$ and 200 mM (3.0×10^{-3} mmol) HEPES (pH 7.4) at 50 °C for 3 hours. The reaction was stopped by the addition of 200 mM EDTA (3.0×10^{-3} mmol). 2.13×10^{-4} mmol of the crude App^{6s}G reaction mixture was purified by HPLC (conditions in section 2.2.2.2) by monitoring absorbances at both 260 and 342 nm (elution time 24.6 minutes). The product was recovered in 9% yield (2.0×10^{-5} mmol) by using the extinction coefficient for the ^{6s}GMP portion of the product. The fact the product absorbed at both 260 and 342 (λ_{max} for adenine and 6-thioguanine, respectively) with relatively equal intensities, as well as the correct mass for the dimer shown in MALDI-TOF spectrometry (m/z 709), indicated the product App^{6s}G was obtained. A 1.91 mM stock in water was prepared for RNA ligations, divided into aliquots, and kept at -80° until further use.

2.2.2.5 7.02 RNA ligation protocol

3.3 μM 7.02 RNA and 62.5 μM thioNAD$^+$ (2.2.2.2), 65 μM App^{6s}G (2.2.2.4) or App^{4s}U were incubated in the presence of ligation buffer (50 mM HEPES, 10 mM $MgCl_2$, 3.3 mM DTT, 10 μg/mL BSA, 8.3% v/v glycerol, 15% PEG 8000, pH 8.0) and incubated for three hours at 23 °C in the presence of 0.5 U/μL T4 RNA ligase (Wang and Unrau, 2002). Ligations in the presence of AppA were the same as above except the reactions contained 32.5 μM of each App^{6s}G and AppA (source, 1 mM stock in water). The reaction was terminated with addition of 2.4 mM EDTA and 300 mM NaCl, and ethanol precipitated. The pellet was resuspended in an equal of volume of water, an aliquot was diluted with a seven-fold excess of tRNA loading buffer (50 mM EDTA, 90% formamide, 10 μM tRNA), and analyzed by 6% PAGE containing 40 μM APM. Samples ligated to ^{6s}G were analyzed in gels containing 40 μM, 20 μM, 10 μM and 8 μM APM.

2.2.2.6 Preparation and Ligation of GpC, ApApC oligos to thioNAD$^+$, App4^{s}U or App^{6s}G

GpC and ApApC (10 μM each) were each radiolabelled in the 5' position in the presence of γ -[^{32}P] ATP, 100 mM KCl, 70 mM Tris-HCl pH 7.6, 10 mM $MgCl_2$, 1 mM 2-mercaptoethanol (supplied as 5× Forward reaction buffer), and 1 unit/μL T4 polynucleotide kinase for 30 minutes at 37 °C. After this time, the enzyme was heat inactivated at 60 °C for 25 minutes. Kinased oligos were analyzed by 20% PAGE. 5' endlabelled oligos (10 μM each) were ligated with thionicotinamide adenine dinucleotide (2.2.2.2), App^{6s}G or App^{4s}U in conditions as described in section 2.2.2.5. Aliquots were removed over this incubation time and stopped with excess loading buffer. These samples were loaded onto a 20% polyacrylamide gel. AAC ligation timepoints, which showed successful ligation to both ^{4s}U and thionicotinamide mononucleotide, were also analyzed by 20% PAGE containing 40 μM APM. Lanes containing kinased samples with no incubation were included as a zero timepoint.

2.2.3 Results and Discussion

2.2.3.1 Thionicotinamide Mononucleotide Synthase Ribozyme Selection Marker

Due to the prominent role of T4 RNA ligase in the previous 4-thiouridine selection, it was very simple to attain the RNA-thionicotinamide mononucleotide construct. Since thioNAD$^+$ contains the T4 RNA ligase recognition element AMP, it was used to transfer the pyridine nucleotide to the 3′ end of substrate RNA (Figure 12). Since we commonly observed ligation efficiencies of approximately 50%, both modified and unmodified sequences would still be present in the reaction. Therefore, analysis of this reaction mixture on an APM gel would display two bands if retardation of the ligated sequences was successful. The RNA sequence chosen was the 7.02 clone from the 4-thiouridine ribozyme synthase selection. The catalytic activity of this sequence played no role in the reaction or the subsequent gel-shift assays - it was chosen because of its length of 124 nucleotides. Since both novel selections would be based on this sequence, it was important to try the gel shift on a sequence of a similar size.

The first ligation using thioNAD$^+$ and 7.02 RNA was performed under standard conditions for ligating 7.02 and App^{4s}U. As a control, a separate ligation of App^{4s}U and 7.02 RNA was performed to confirm ligase activity of the enzyme. Also, the shift pattern of 7.02 RNA-^{4s}U relative to unmodified 7.02 RNA would confirm APM conditions are high enough to show retention (at least of the 7.02 RNA-^{4s}U). These parallel ligations were analyzed on a 6% denaturing polyacrylamide gel containing 40 uM APM. The results are shown in Figure 13. As expected, the App^{4s}U ligation lane showed two bands roughly equal in intensity (corresponding to the roughly 50:50 ratio of ligated and unligated RNA). As well, an even slower migrating band was also present. This band, which was just beneath the wells of the gel, corresponds to 7.02 RNA that is ligated to two ^{4s}U nucleotides. Given that the 7.02-^{4s}U control showed shifted bands, we deduced that the APM gel conditions were appropriate for retarding the mobility of thiolated material. Unfortunately, no shifted RNA was present in the thioNAD$^+$ ligation reaction. The thioNAD$^+$ lane contained one band that migrated to the unligated position as compared to the control. This single band had approximately double the intensity of the unreacted band in the control. A small amount of degradation (smaller than unmodified 7.02 RNA) was seen in each lane. There were two possible reasons for lack of shifted bands in the thioNAD$^+$ lane. The first (and most likely) is that any 7.02-thionicotinamide construct (assuming it was formed during the ligation) was incapable of shifting in an APM gel (of this concentration). The second is that no, or very little, ligation was achieved during the course of the reaction.

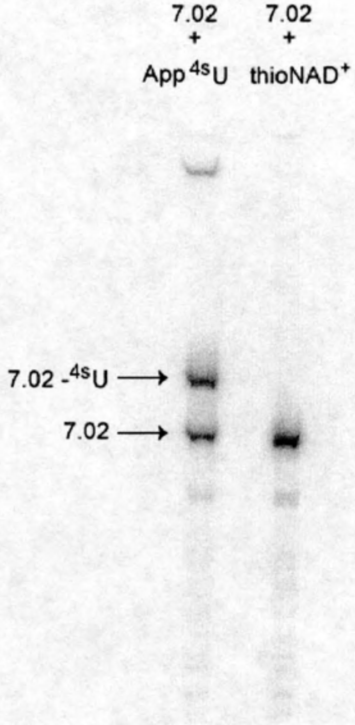
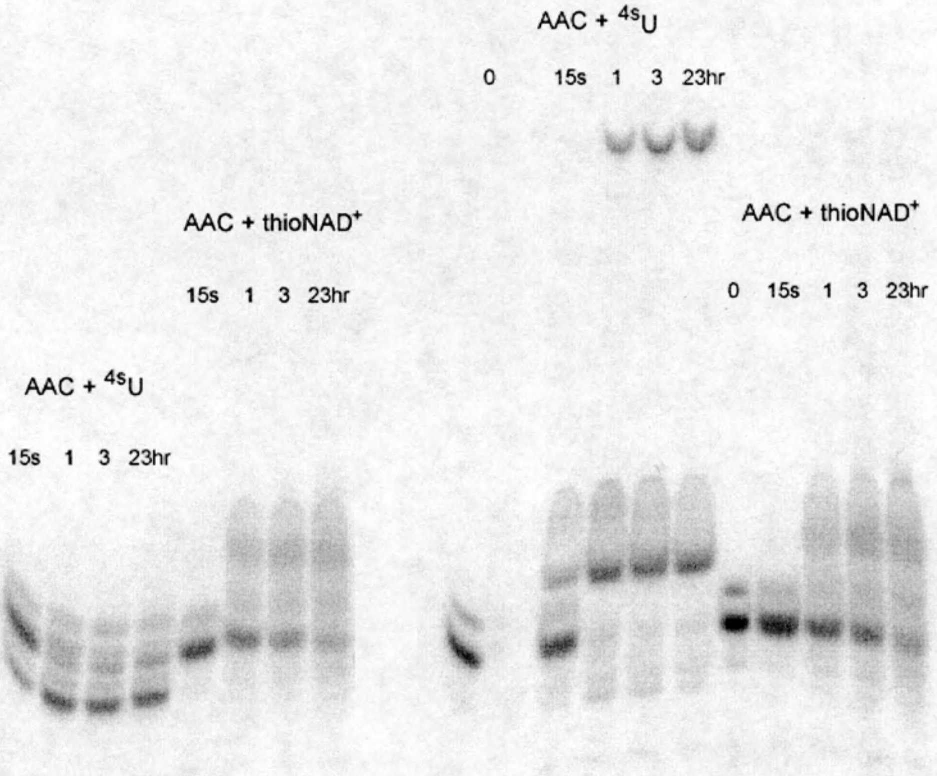


Figure 13: Mercury gel of thioNAD$^+$ and App^{4s}U ligation with 7.02

As the preliminary gel was inconclusive, it was necessary to investigate whether thionicotinamide mononucleotide was capable of being transferred to the 3′ end of the 7.02 sequence under our conditions. As there is a one-nucleotide length difference between unligated and ligated 7.02 (124 and 125 nucleotides, repectively), PAGE cannot easily resolve these sequences. Two small oligos (GpC and ApApC) were employed as substrates for thionicotinamide ligation. These oligos allowed us to visualize any length increase resulting from ligation to thionicotinamide using a high-percentage polyacrylamide gel. The advantage of this experiment is that it allows us to follow the ligation results based both on size difference (regular PAGE) and mercury affinity (APM gel).

Both oligos (from stocks already present in the lab) were examined for purity in a polyacrylamide gel prior to ligation. While GpC was a single band, ApApC was a mixture of the three nucleotide band and two minor contaminants (slightly larger and smaller). As we were only interested in observing qualitative size changes in the reaction, the ApApC sample was not purified. Each oligo was ligated in parallel experiments to either thionicotinamide mononucleotide or 4-thiouridine as in the previous experiment. The reactions were allowed to progress at room temperature over a twenty-three hour period, with several aliquots removed and quenched with EDTA. ApApC was successfully ligated to each of ^{4s}U and thionicotinamide mononucleotide over the course of the reaction (Figure 14a). However, GpC proved to be a poor substrate compared to ApApC, ligating only to ^{4s}U and not thionicotinamide mononucleotide (results not shown). While the AAC-thionicotinamide construct migrated slower than AAC in the gel as expected, the GC-^{4s}U and AAC-^{4s}U sequences actually migrated faster than their unmodified versions. Potentially, the tautomerization behaviour of the thione of 4-thiouracil could impart more negative character to the ^{4s}U ligated RNA, causing faster migration during electrophoresis. As the two AAC ligation reactions were shown to both be successful, the remaining volumes of these samples were run into an APM gel. That way, it could be directly investigated whether the AAC-thionicotinamide product was capable of shifting. In conjunction with the previous experiment, the presence of mercury did not change the mobility of the AAC-thionicotinamide product (Figure 14b). It also confirmed the identity of the fast moving product in the ^{4s}U ligation as RNA-^{4s}U, as these bands disappeared and higher-moving bands were present in the APM gel. Therefore, APM gels (at 40 μM concentration) were considered incapable of retarding the mobility of RNA sequences tethered to thionicotinamide mononucleotide at the 3′ end.



a) - APM

b) + APM

Figure 14: Gel shift capabilities of ^{4s}U and thioNMN ligated RNA in a) regular and b) mercury gels

Without the shift needed to select potential reactive ribozymes away from unreacted sequences, a selection for thionicotinamide nucleotide synthase ribozymes using our scheme was likely not possible. We did not investigate a possible requirement for higher APM concentrations. In the future, any selection of this sort could feature a column capable of binding reacted sequences. After encountering this setback for a thionicotinamide selection, we instead turned our attention to investigating the proposed selection for 6-thioguanosine monophosphate synthase ribozymes.

2.2.3.2 6-Thioguanosine Monophosphate Synthase Ribozyme Selection Marker

The first in the three-step process to create the marker for the selection was the phosphorylation of 6-thioguanosine. 6-Thioguanosine 5′-monophosphate was prepared as described in the literature (Breter and Mertes, 1990) using phosphorous oxychloride. TLC analysis of the crude reaction showed a mixture of two spots. The predominant spot had the lower Rf, the same as that of the reference nucleotide GMP. The faint spot with the higher Rf co-migrated with the standards 6-thioguanine, 6-thioguanosine and guanosine, and was likely starting material. Due to the change in mobility as a result of varying pH, all standards and samples were dissolved in HEPES pH 7.4. These results indicated the reaction had gone close to completion and no attempts were made to optimize the synthesis.

To purify the product from nucleoside, a Sephadex column was employed. We tried three different elution conditions using such a column. As the HCl conditions used by Breter and Mertes seemed harsh, we used a TEAA gradient. However, the buffer salts were not volatile and could not be easily removed. As a result, NMR analysis was impossible because the signals from the salts overwhelmed those of the product. After a second synthesis, the HCl column as in the publication was used. However, the NMR spectrum indicated it was not clean. Analysis by HPLC confirmed it was impure (only 36% was attributed to the ^{6s}GMP). After a third synthesis of ^{6s}GMP, a gradient of TEAB was used to elute the nucleotide from the Sephadex column. The advantage of using this buffer was that the salts were more volatile and were removed after resuspending in water and re-lyophilyzing. The absorbances of all fractions were measured and plotted as a function of fraction number allowing the peaks to be more acurately visualized. One minor distinct peak that eluted just before the nucleotide was attributed to nucleoside starting material. Tubes at the edges of this major peak were not included. The yield would have probably been higher had a fraction collector failure (during the product elution) been discovered earlier. HPLC analysis of this product showed it to be present in 99.1%. Only one other minor peak was visible (0.9%), which eluted after the product. This is in agreement with the report of Breter and Mertes, who also observed a later-eluting minor contaminant after Sephadex purification. This sample was used as is.

While Breter and Mertes reported no further characterization of their product beyond HPLC, we were interested in using NMR spectroscopy for identification. We not only wanted to verify that the phosphorous was present, but also that it was in the correct position on the ribose. Since no NMR spectrum of 6-

thioguanosine monophosphate was reported in the literature, a comparison to guanosine monophosphate was used. The proton NMR spectrum of our product compares very well to GMP. Each spectrum indicates one aromatic proton, corresponding to the lone guanine base proton. Around 6 ppm, both spectra show a signal split into a doublet. This is probably the H-1 protons, split into a doublet from H-2. Also it is the furthest downfield of all the ribose protons because of its proximity to the nitrogen. The remaining ribose protons show up in the 3-5 ppm region. Two peaks overlap around 4 ppm, probably the diastereotopic H-5 protons. Of the remaining signals, the more complex multiplet is attributed to the H-4. This proton is not only subject to splitting from the two H-5s but also H-3. Both spectra lack a sixth ribose signal, which is probably buried under the large HOD peak.

Phosphorous NMR spectrum was performed to confirm the presence of phosphorous on the molecule. The spectrum showed a peak at around 4 ppm relative to the phosphoric acid reference. This spectrum was obtained by allowing coupling to occur, in order to determine if the splitting between C-5 protons and the phosphorous was present. Theoretically, the phosphorous signal should be split (into a triplet), as was observed in the GMP reference. In our sample the peak was broad, indicative of substitution in the C-5 position. Possibly, the cleaner nature of the GMP standard (which did not contain any buffer salts) allowed splitting to be clearly visible. Also, the NMR sample was probably more concentrated due to the abundance of this commercially available compound. MALDI-TOF analysis gave further evidence that ^{6s}GMP was indeed phosphorylated.

Adenosine 5′-phosphorimidazolide imidazole-activated AMP (ImpA) was successfully used for preparation of substrates for T4 RNA ligase in the original selection (Unrau and Bartel, 1998). This procedure was again used to prepare App^{6s}G (**6**). This synthesis was done under the same conditions as the App^{4s}U synthesis for the 4-thiouridine synthase selection. The crude reaction was purified by HPLC on a reverse phase column, as the product was only about 20% of this mixture. To identify the correct peak, absorbances at 260 and 342 nm were followed simultaneously. This ensured the product had both adenine and 6-thioguanine bases. The formation of this molecule was confirmed by MALDI-TOF spectroscopy. As the following ligation step was enzymatic (which involves substrate specificity), we were not concerned about the purity or identification of this compound.

As described in the previous thionicotinamide section, the marker RNA chosen was the 7.02 clone from the 4-thiouridine ribozyme synthase selection. 7.02 RNA internally labelled with ^{32}P was incubated with the substrate App^{6s}G and T4 RNA ligase. A lane containing unligated RNA was also included to demark the migration of unmodified 7.02. Luckily, two retained bands appeared in the App^{6s}G ligation samples in addition to the unligated band (Figure 15). These products were assumed to be the RNA singly and doubly ligated to ^{6s}GMP. The

^{6s}G-tethered RNA mobility is greatly reduced compared with the ^{4s}U-tethered RNA. In App^{4s}U ligations, the singly ligated product shows up a short distance above unmodified, and doubly ligated product almost at the wells. However in the App^{6s}G ligations, the singly ligated product is almost at the location of the RNA doubly ligated to ^{4s}U. The doubly ligated is at the top of the wells and did not even move. The affinity between RNA tethered to one thiolated guanosine monophosphate and APM is much greater than that of RNA tethered to 4-thiouridine monophosphate.

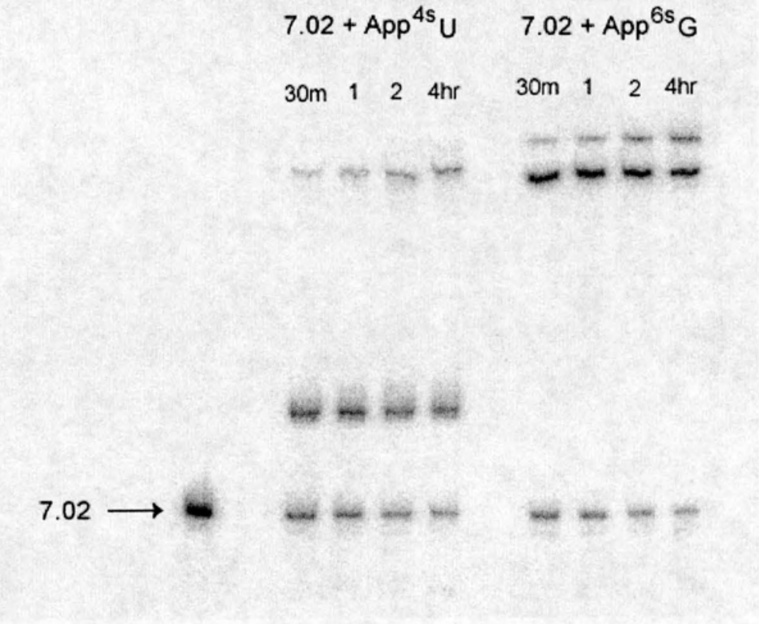


Figure 15: APM gel of App^{4s}U and App^{6s}G ligation with 7.02 RNA.

In another experiment, we proved that a product of ligation was responsible for the shifted bands in the mercury gel. In addition to App^{6s}G, an equimolar amount of the competitive substrate AppA was included to bring the overall total of adenylated substrate to the same final concentration as the previous experiment.

As a result, the ligated RNA should be a combination of RNA-^{6s}G and RNA-AMP products. Only the ^{6s}G tethered RNA sequences should shift in a mercury containing gel. As expected, a decrease in intensity of the shifted bands could be seen, due to the decrease of sulfur-containing RNA (results not shown). Compared to regular ^{6s}G ligation, the intensity of unshifted RNA is approximately double. We concluded that it was indeed the product of ligation that led to these low mobility bands in the gel.

A more direct way to examine the gel shifting of ^{6s}G terminated RNA was to use small oligos as substrates, as in the thionicotinamide experiments. As before, GpC and ApApC oligos were end labelled with ^{32}P and used as substrates for the ligation assays. The substrates were ligated to either App^{6s}G or App^{4s}U. GC again proved to be the poorer substrate of the two, with ligation occurring at a very low amount according 20% PAGE. The GC and AAC ligation products to ^{4s}U showed new fast moving bands appearing in the lanes over the course of the ligation, similar to the thionicotinamide controls. Over the course of the App^{6s}G ligation, two additional slow moving bands appear above the parent bands. These appear to be one and two nucleotides higher than the larger of the two parent bands. To directly confirm these new bands were from the ligation, the same samples were also run into an APM gel. As expected, these two bands are shifted drastically in the gel (results not shown). These are most likely the products singly and doubly ligated to ^{6s}G. These results led us to conclude that ligation to ^{6s}G slows the mobility of RNA in a mercury gel. A very positive aspect in these experiments was the efficient separation of ^{6s}G-tethered RNA from unmodified RNA in APM gels. This allowed for much lower concentrations of the mercury compound (20 μM) to achieve the same amount of separation as the previous selections (40 μM).

The different abilities of the marker RNA constructs to shift in mercury gels was rationalized to be due to the natures of the sulphurs in each of these compounds. The thiocarbonyl (C=S) can tautomerize to its corresponding sulfhydryl (-SH) functional group, which shows a higher affinity for mercury. The thiocarbonyl in thionicotinamide likely does not have sufficient sulfhydryl character to allow shifting in the APM gel. However, the sulfhydryl form is probably more highly stabilized in ^{6s}G than in ^{4s}U, which leads to slower migration of 7.02-^{6s}G than even 7.02-^{4s}U in mercury gels.

After the marker shift experiment confirmed that the selection was capable of using the mercury separation method, the selection for 6-thioguanosine monophosphate nucleotide synthase ribozymes was carried out by Matthew Lau as described in Lau *et al.*, 2004.

2.2.4 Conclusion

The ability of ribozymes to catalyze pyrimidine nucleotide synthesis has previously been established, indicating nucleotide synthesis was possible in an early “RNA World” where enzymes were RNA, not proteins. With this discovery, the next step was to investigate whether RNA was capable of catalyzing the synthesis of other nucleotides. In this hypothetical RNA World, ribozymes would have needed to synthesize purine nucleotides, in addition to pyrimidine nucleotides, to complete the genetic code. Synthesis of other glycosidic bonds would have also been necessary in early metabolism, such as that of the pyridine nucleotide in the cofactor nicotinamide adenine dinucleotide. Identification of such ribozymes would complement the role of previously discovered ribozymes in the RNA World. The selection of a ribozyme capable of synthesizing thionicotinamide mononucleotide was not feasible according to our selection protocol. However, a selection for ribozymes catalyzing the synthesis of the purine nucleotide 6-thioguanosine monophosphate was possible. This selection was eventually completed, isolating ribozymes capable of this catalysis. This finding demonstrates that RNA enzymes could have been capable of creating both classes of nucleotides in the "RNA World", lending more support to this hypothesis.

3 INTRODUCTION TO PART II

3.1 Molecular Switches and Photochromism

Molecular switches are compounds that reversibly interconvert between two isomers **A** and **B** (as shown in Figure 16) in response to an external stimulus (reviewed in Feringa, 2001). Each isomer displays unique properties such as shape, absorption spectra, emission, and redox potentials. Stimuli that trigger the conversion of one isomer to the other include light, heat, electrons, and chemical reagents. Specifically, photochromism describes the reversible conversion between two isomers displaying different absorption spectra in response to two different wavelengths of light. All of the molecular switches featured in this thesis are photochromic.

$$\mathbf{A} \underset{h\nu_2}{\overset{h\nu_1}{\rightleftharpoons}} \mathbf{B}$$

Figure 16: General scheme of photochromism.

The photoisomerization process can be followed by UV-Vis absorption spectroscopy. An example of the absorption spectra of a typical isomerization process between isomers **A** and **B** is shown in Figure 17. The absorption spectrum of each isomer is used to determine the irradiation wavelength necessary to trigger its conversion to the other isomer. As isomer **A** is irradiated at its absorption band in the UV region ($h\nu_1$), this decreases and a new band appears in the visible region as isomer **B** is generated. Reversion back to **A** is accomplished by irradiating **B** at its absorbance in the visible region ($h\nu_2$). Spectra obtained over the course of the isomerization process display isosbestic points as shown in Figure 17. At this wavelength, the sum of the absorbances of all species are equal for each sample. Generally, this indicates the photoisomerization process is clean, meaning no degradation or formation of side products. One characteristic of successful photochromic systems is two isomers with separated absorption spectra, such that each isomer is independently addressable. This means that irradiation with a particular wavelength will only affect one isomer. For example, if both isomers absorb within energy range $h\nu_1$, stimulation of both **A** to **B** and **B** to **A** isomerizations can result. The conversion to isomer **B** is not complete; instead, an equilibrium (called the photostationary state) is reached between these two competing reactions.

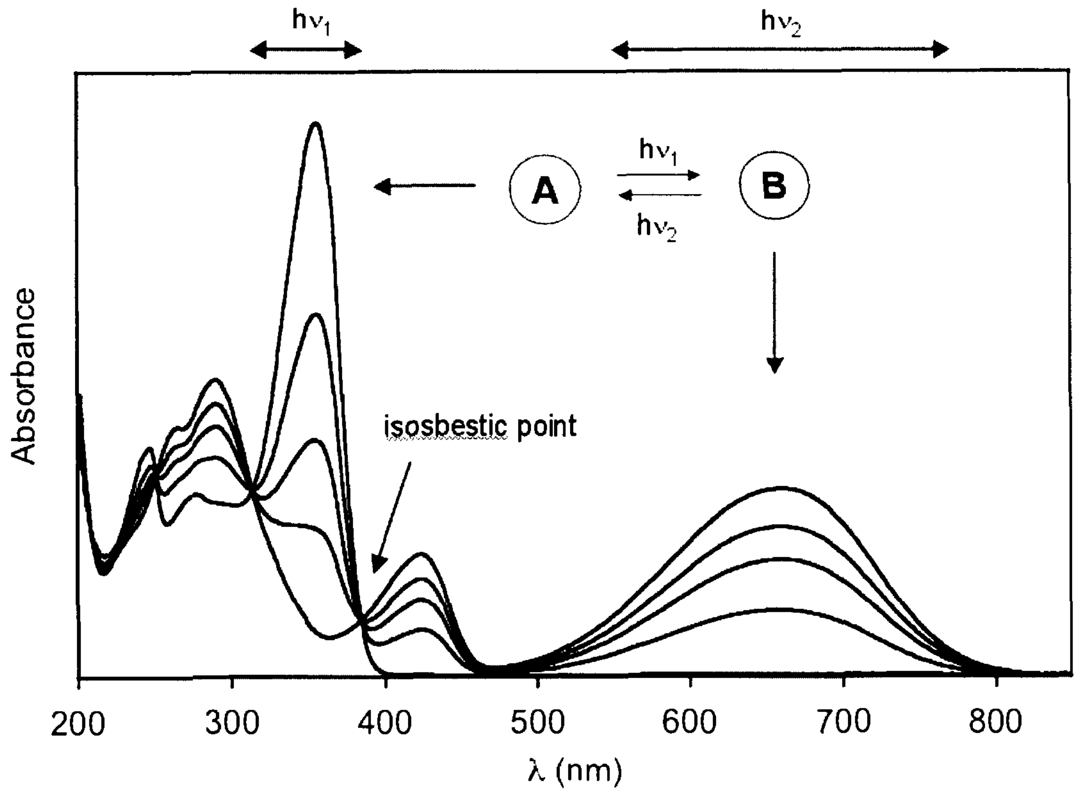


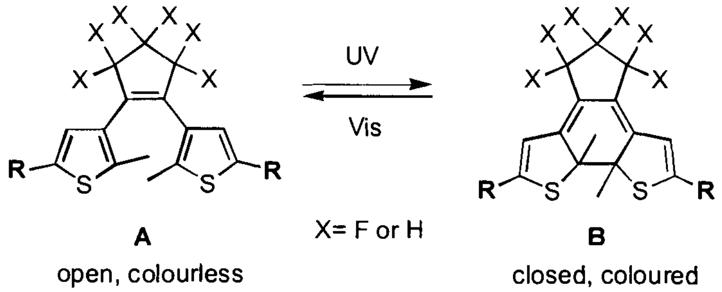
Figure 17: UV-Vis absorption spectra of a typical photoisomerization reaction between differently coloured isomers **A** and **B**.

The ability to control the physical properties of these molecular switches makes them appealing candidates for use in a range of applications such as filters, sensors, information storage and other optoelectronic and photonic devices. In particular, photoresponsive compounds have been incorporated into eyewear capable of reverting between dark and light lenses in response to light conditions. Although the definition of desirable molecular switching characteristics varies according to the application, several properties are often necessary. For example, it is critical that a given system is able to interconvert between its isomers many times without significant degradation (fatigue resistance). Also, it is desirable that the quantum yield (the proportion of photons absorbed by the molecule that trigger isomerization to the other isomer) be high. Another requirement for successful photochromic systems is that the isomers can be independently addressable, as previously described. A common obstacle with photochromic systems is the thermal conversion of one isomer to the other. It is considered desirable for photoresponsive compounds to exibit thermal stability of each form (bistability).

3.2 Dithienylethenes

Dithienylethenes (DTEs) are a class of photochromic compounds that exhibit bistability, high fatigue resistance and synthetic versatility. Due to these properties, DTEs have been the subject of intense academic and industrial research. All photoswitches used in this thesis are dithienylethenes. Dithienylethenes can interconvert between two isomers, "open" (**A**) and "closed" (**B**), as shown in Scheme 1. The open isomer is colourless, and has rotation through two of its bonds making it flexible. Exposing the open form to ultraviolet light induces a ring-closing reaction. This creates a conjugated backbone between the **R** substituents (Scheme 1). This extended pattern of conjugation causes the closed molecule in almost all cases to be highly coloured. As the closed isomer is cyclized, it is more rigid than the open version. The closed isomer possesses a characteristic absorbance in the visible region. Reversal to the open form is achieved by irradiation of the closed form with visible light.

The dithienylethene scaffold contains few positions that can be varied synthetically, the easiest being the **R** substituent on each thiophene. The nature of the substituents on dithienylethenes can affect the colour of the ring-closed isomer.



Scheme 1: Photochemical reaction of dithienylethene switches.
Conversion from the flexible "open" isomer (**A**) to the rigid "closed" isomer (**B**) is achieved by irradiation with UV light. The open isomer is regenerated by exposure of the closed form to visible light.

3.3 Biological Motivations

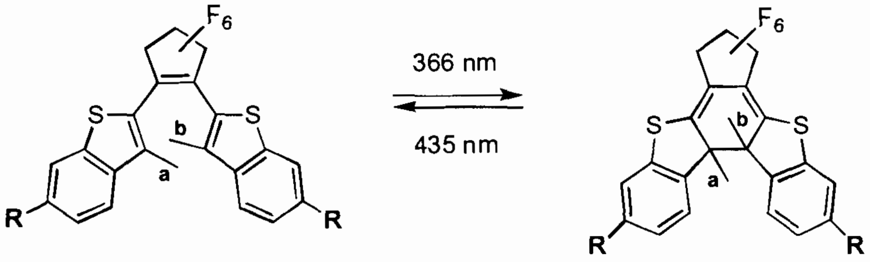
Controlling interactions between synthetic compounds and biological molecules is very desirable in both *in vitro* and *in vivo* systems. Specifically, our goals focused on modulating potential interactions between dithienylethene molecular switches and nucleic acids. Biologically desirable dithienylethene target molecules would contain substituents (in the **R** position) possibly able to confer biological functions to these novel switches. With DNA binding, the switch could be bound (blocking transcription) or unbound (allowing it). This would be applicable to *in vitro* roles such as control of gene transcription. For RNA cleavage, mRNA could be degraded. This would prevent production of a protein product.

Another potential application of dithienylethenes is photodynamic therapy. Photodynamic therapy is a treatment of dermatological, oncological, and ophthalmic diseases (Levy and Obochi, 1996). This is based on the use of photosensitive compounds, known as photosensitizers. These therapeutic compounds are either topically applied, injected or delivered to diseased tissues beneath the skin (subcutaneous). Irradiation with the appropriate wavelength of light activates the photosensitizers, causing them to generate reactive species (such as singlet oxygen), which destroys surrounding tissues. The hydrophobic nature of these drugs makes subcutaneous administration difficult. The development of new delivery systems is an area of ongoing research (reviewed in Konan *et al.*, 2002). The specific location of these drugs in affected tissues, leaving healthy surrounding tissues intact, makes photodynamic therapy an appealing alternative to treatments such as chemotherapy and surgery.

Photochromic compounds also have attractive features compared to other molecular switches. Reversible changes in colour (and potentially, fluorescence) may allow development of such compounds into sensors or visualization aids in biological settings. There has only been one report of a photochromic compound interacting with DNA under biological conditions (Starcevic *et al.*, 2005). One form of this molecule does not bind DNA but the other isomer intercalates into the double helix. The authors speculate on its usefulness in medical research as an anti-cancer drug in photodynamic therapy. However, it is important to note that the compound used undergoes a permanent conversion after UV light exposure. In this sense, it is not a reversible switch.

The dithienylethene scaffold presents several attractive features for use in biological systems. These switches display fatigue resistance and high quantum yields compared to other photochromic molecular switches. Secondly, synthetic variation of the thiophene substituents can vary many of these properties, often without destroying isomerization ability. This also allows researchers to make dithienylethenes soluble in either hydrophobic or hydrophilic conditions tailoring them for the biological environment being studied. Also, the small scaffold size relative to biological macromolecules such as nucleic acids and proteins allows use of these in such systems.

To date there has only been one literature report of interactions between dithienylethenes and nucleic acids. Saito and colleagues reported (during the course of this thesis) that DNA molecules, when non-covalently incorporated into a dithienylethene film, alter the refractive index properties of the dithienylethene units (Saito *et al.*, 2003). They hypothesized that the DNA intercalated based on the perpendicular orientation of the switches into the film. This group speculated that one aromatic substituent of the closed molecule intercalated into the DNA basepairs. The two methyl groups ‘a’ and ‘b’ (shown in Scheme 2) limit intercalation past this point. This report, as well as that by Starcevic *et al.*, are the only two literature references of photochromic compounds interacting with DNA.



12*o* R = CH_2OH **12c**

Scheme 2: Dithienylethene switch bound to DNA-containing film. F_6 denotes fluorine substituents at every position in the cyclopentene ring. Abbreviations 'c' and 'o' in the compound names refer to the open and closed isomers, respectively.

4 BIOLOGICAL APPLICATIONS OF PYRIDINIUM SUBSTITUTED DITHIENYLETHENE SWITCHES

4.1 Introduction

The goal of this work was to investigate potential biological applications of dithienylethene molecular switches. The first project was to investigate possible binding of dithienylethene switches to DNA. Ultimately, the most desirable outcome would be the development of a switch that binds to DNA when in one isomeric form, but not in the other. It was our hope that the structural changes between DTE isomers would allow the identification of a switch capable of binding DNA in only one isomer. This would present an intriguing scenario, in which binding to DNA could be turned "on" or "off" with the appropriate wavelength of light. Another project was the staining of bacteria in the procedure known as the Gram stain. This procedure utilizes a cationic dye that permeates through the bacterial cell wall and is held in place by a fixative. However, unlike the current procedure, this would allow a stain capable of reverting between colourless and brightly coloured states. Two different *bis*-pyridinium dithienylethene derivatives (Figure 18), featuring a number of attractive properties for each application, were utilized. We used the fluorinated switch **13** and the novel non-fluorinated switch **14** for applications in both of these areas. Success in such projects would be the first reported examples of dithienylethene switches studied in biological environments.

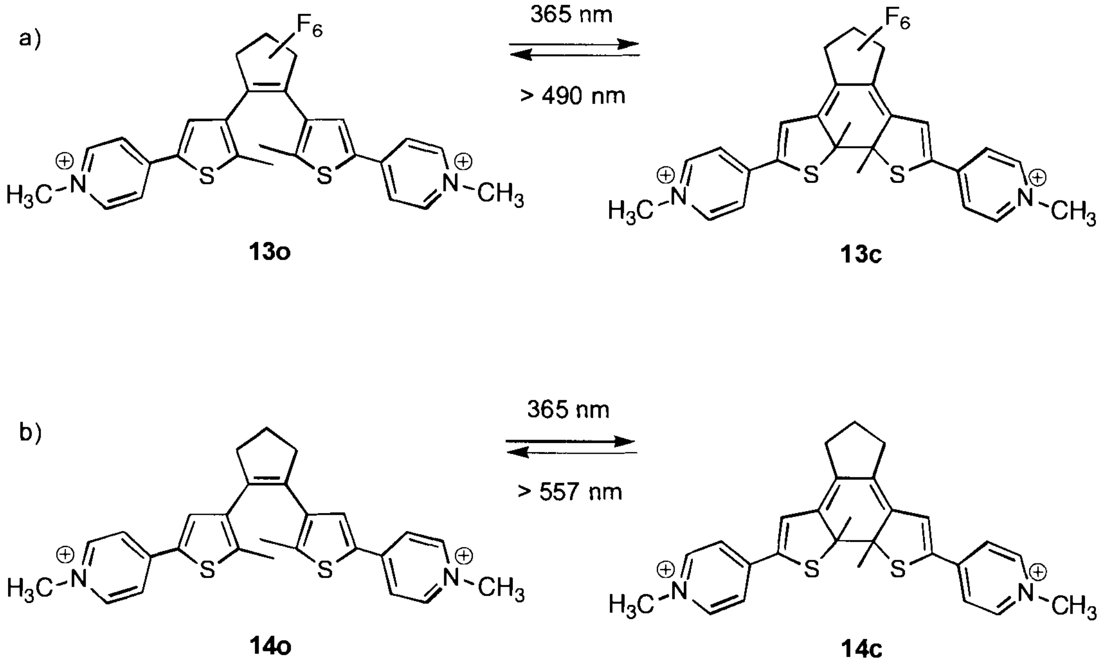
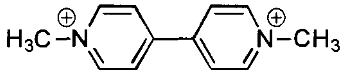


Figure 18: Structures of the *bis*-pyridinium dithienylethene switches studied in this chapter. a) The fluorinated *bis*-pyridinium **13** reported by Gilat *et al.*, 1995. b) The novel non-fluorinated *bis*-pyridinium **14**, synthesized by Brian Gorodetsky (Branda group, SFU).

4.2 DNA Binding Studies

4.2.1 Introduction

The first candidate we considered for DNA binding was the *bis*-pyridinium dithienylethene switch shown in Figure 18a. The synthesis and characterization of this molecule was first reported by Lehn and colleagues (Gilat *et al.*, 1995). This switch is based on the structure of methyl viologen (Figure 19). Viologens are two pyridinium groups joined with a conjugated linker at the 4-position in each ring. Methyl viologen (**15**) is capable of undergoing a reduction at each of the two pyridinium positions. Lehn proposed use of a methyl viologen unit in this switch to control the ability of these groups to conduct electrons.



15

Figure 19: Methyl viologen

Gilet *et al.* created an extended analogue of methyl viologen, in which the pyridinium substituents are separated by a photoswitchable unit. The pyridinium groups are only in conjugation in the closed form. They displayed in this paper that the redox properties of the two forms, open and closed, are very different. In the open form the two distal pyridinium substituents behave independently, being reduced at the same potential (E_{red} = -1.04 V). However in the closed form, backbone conjugation allows each pyridinium to electrochemically "sense" the other. As a result, reduction of the first pyridinium is at a less negative potential (E_{red} = -0.23 V), followed by a reduction at a more negative potential of the second (E_{red} = -1.47). More recently, the *bis*-pyridinium **13** has been studied in the Branda laboratory. The ring-closing reaction has been shown to be induced by electrochemical reduction in addition to light stimulus (Gorodetsky *et al.*, 2004). As well, the corresponding unmethylated *bis*-pyridine switch **16** (Figure 20) has been derivatized with porphyrins (Norsten and Branda, 2001) and polymerized using zinc coordination (Matsuda *et al.*, 2001) to allow photocontrol of supramolecular complexes.

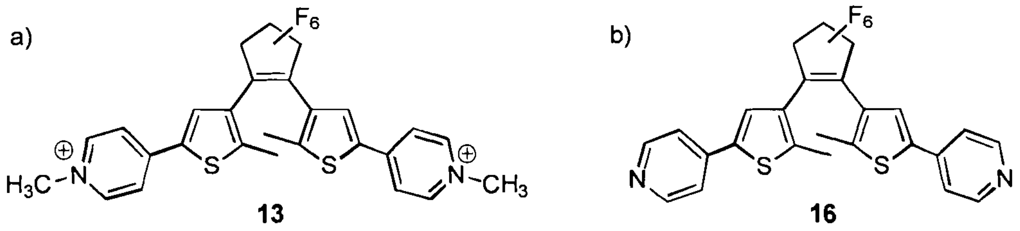
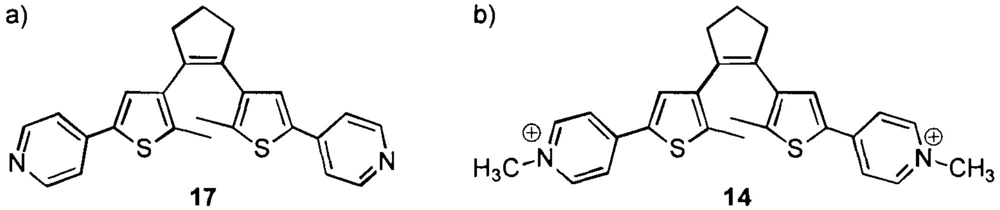


Figure 20: Fluorinated *bis*-pyridine dithienylethene derivatives reported by Lehn and co-workers (Gilat *et al.*, 1995).
a) Fluorinated *bis*-pyridinium switch **13** used in this work. b) Fluorinated *bis*-pyridine switch **16**.

This *bis*-pyridinium dithienylethene switch was attractive for use in DNA binding studies for a number of reasons. The synthesis, photochromism and redox properties are well-studied. Methyl viologen derivatives have been shown to interact with nucleic acids (Pang and Abruna, 2000). As well, the photocontrollable redox properties of *bis*-pyridinium **13** could potentially be used to disrupt the electron transfer abilities of DNA if bound to the double helix.

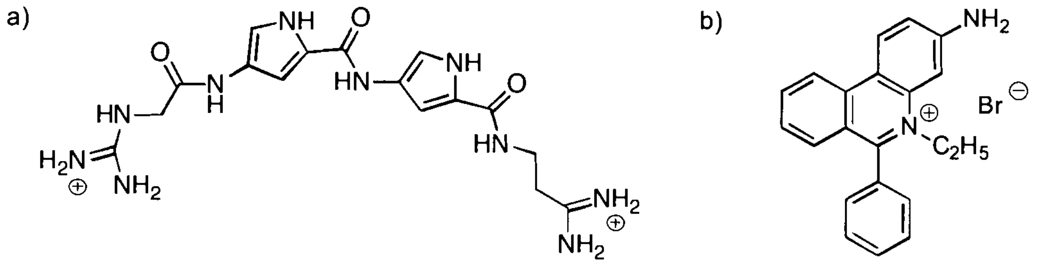
In 2003, Tian and co-workers reported the preparation of a non-fluorinated *bis*-pyridine (**17**), shown in Figure 21a (Qin *et al.*, 2003). Unlike the fluorinated *bis*-pyridine **16**, this compound is fluorescent. The photochromism and fluorescence of this compound are enhanced by coordination to zinc. During the course of this thesis, the novel *bis*-pyridinium analogue of this compound (**14**) was synthesized by another student in the Branda lab (Brian Gorodetsky, Ph.D. thesis). This compound is shown in Figure 21b. The open form is colourless and the closed form is blue-green. The extinction coefficients of the open and closed forms at their respective λ_{max} wavelengths are very similar. This makes the non-fluorinated switch **14** a good molecule for DNA studies since several assays rely on UV absorbance. The similar absorbances of the two isomers allow direct comparison and similar error. The open molecule displays fluorescence when irradiated with UV light, but the closed does not. Therefore, not only colour, but fluorescence can be controlled by light stimuli. However one drawback is that the excitation wavelength for fluorescence (approximately 410 nm) stimulates the ring-closing reaction. The redox properties have not been studied but are presumed to be analogous to the fluorinated compound.



17 **14**

Figure 21: Non-fluorinated *bis*-pyridine dithienylethene derivatives.
a) *Bis*-pyridine switch **17** reported by Tian. b) Novel *bis*-pyridinium **14** studied in this thesis.

Small molecules can bind DNA by binding the minor groove or by intercalation in between base pairs. Molecules have also been synthesized that contain both intercalation and groove binding moieties, allowing a combination of binding modes. Minor groove binding requires a crescent-like shape of the ligand to snake along the curve of the groove. One such example is netropsin (**18**), shown in Figure 22a. The crescent shape of the *bis*-pyridinium switches, combined with its distal charges, might allow it to groove bind similarly to netropsin. Although both isomers are crescent shaped, the open form is more flexible compared to the closed switch. Due to this flexibility, it may be able to adopt a configuration for binding the minor groove. The closed isomer may be too rigid to bind a flexible DNA helix.



18 **19**

Figure 22: Structures of the known DNA binders netropsin (a) and ethidium bromide (b). Netropsin binds the minor groove of DNA due to its crescent shape and hydrogen bonding ability. The planar, charged structure of ethidium bromide allows it to intercalate into the base-pairs of DNA.

On the other hand, it may be able to intercalate between the base-pairs, such as ethidium bromide (**19**, Figure 22b). As pyridinium is an aromatic group, it may be able to intercalate in between the base-pairs of DNA. Any intercalation ability could be aided by the presence of the positive charge at the edge of each pyridinium ring, allowing favourable interactions with the negatively charged backbone. Since the pyridinium groups in the closed molecule are fixed at distal ends of the molecule, only one would be able to intercalate into one DNA helix at a given time. However in the open form, the rotation may place one pyridinium at an orientation that prevents the other from binding due to steric bulk.

The ideal outcome of this project would be that the dramatic contrast in shapes of the open and closed isomers would allow binding in one isomer, but negate it in the other. Another possibility is that the closed form might bind in one method, but open in another. For example, the open form may prefer groove binding over intercalation; however switching into the closed form could prevent this and the molecule would revert to intercalation. Another possibility is that the binding of one isomer might prevent the isomerization reaction. The intercalation of one pyridinium (which features free rotation with respect to the dithienylethene scaffold) would not be affected by the ring closing or opening reaction. However a groove binding isomer may not switch to the closed isomer as that could remove the binding ability.

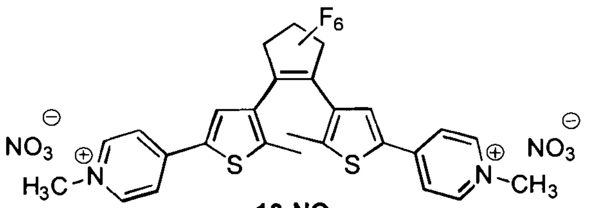
4.2.2 Materials and Methods

4.2.2.1 General

Calf thymus DNA (type XV, activated), RNase A, and Proteinase K were purchased from Sigma. Dialysis membrane was purchased from SpectraPor. Samples were sonicated in a Ultrasonik (NEY) water-bath sonicator. Samples were centrifuged in a Fisher Scientific 228 centrifuge (max 3400 rpm). Solvents were dried by passing them through steel columns containing activated alumina under nitrogen (using an MBRAUN solvent purification system). Octafluorocyclopentene was supplied by the Nippon Zeon Corporation. NMR solvents were obtained from Cambridge Isotope Laboratories. 1,2-*bis*((2′-methyl-5′-(*N*-methylpyrid-4″-yl)-3’-thien-yl) cyclopentene, *bis*-ammonium hexafluorophosphate salt (**14-PF$_6$**) was prepared by Brian Gorodetsky (unpublished results). All other chemicals were purchased from Sigma. Silica gel column chromatography was performed using 230-400 mesh silica gel from Silicycle. Visualization of TLC plates was performed using Spectroline (470 mW/cm^2) hand-held UV lamps. Melting points were measured on Fisher-Johns melting point apparatus. Nuclear magnetic resonance spectra were obtained on either a Bruker AMX 400 FT (operating frequency: 1H, 400.13 MHz) or Varian AS 500 (operating frequency: 1H, 499.77 MHz) spectrometer at ambient temperature. Chemical shifts (δ) for all compounds are listed in parts per million downfield from tetramethylsilane using the residual non-deuterated solvent peak as an internal reference. Mass spectra were obtained using a MALDI-TOF MS spectrometer (PerSeptive Voyager DE STR from PE Applied Biosystems). Microanalysis data was obtained from a Carlo Erba Model 1106 CHN analyzer.

The Spectroline (470 mW/cm^2) hand-held UV lamps were also used to perform the ring-closing reaction (at 254, 313 or 365 nm wavelengths) of dithienylethene-containing samples. During the ring-closing reaction, these samples were either in (Hellma, 3 mL or 1 mL, all with a 1 cm path length) cuvettes (with the lamp placed next to it) or in Eppendorf tubes with the top open (and the lamp placed over top). Ring-opening reactions of these samples were performed by passing the light from a 150-W tungsten source through 490 nm or 550 nm cutoff filters (for compounds **13** and **14**, respectively). UV-visible spectra were obtained on a Cary 300 Bio instrument from Varian.

4.2.2.2 **Synthesis**

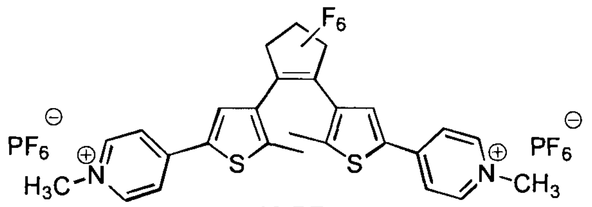


13-NO$_3$

Figure 23: Structure of **13-NO$_3$**.

Synthesis of 1,2-*bis*((2′-methyl-5′-(*N*-methylpyrid-4″-yl)-3′-thien-yl)perfluorocyclopentene, *bis*-nitrate salt (13-NO$_3$).

The *bis*-pyridine **16** starting material was synthesized in four steps as described in the literature (Gilat *et al.*, 1995). *Bis*-pyridine **16** (150 mg, 0.287 mmol) and dry acetonitrile (30 mL) were stirred under nitrogen at ambient temperature in the dark. Methyl iodide (1 mL, 6.57 mmol) was added and the solution was brought to reflux. After 24 hours the reaction contents were evaporated *in vacuo* to remove any remaining methyl iodide. The *bis*-pyridinium product was purified from the trace mono-pyridinium side product by flash chromatography on a silica column. A gradient (10:1:0.01 acetonitrile / water / saturated potassium nitrate increased to 10:1:0.1) was used to elute the product **13-NO$_3$**. The yield was obscured by excess potassium nitrate in the product. Although the synthesis and purification was performed in the dark, the product obtained was a pale blue solid due to the presence of a small amount of the ring-closed isomer **13c-NO$_3$**. Before NMR analysis, the NMR tube containing this product was subjected to the ring-opening reaction. ^{1}**H NMR** (500 MHz, CD$_3$OD) δ 2.18 (s, 6H), 4.32 (s, 6H), 8.16 (s, 2H), 8.24 (d, *J*= 6.7 Hz, 4H), 8.77 (d, *J* = 6.6 Hz, 4H).



13-PF$_6$

Figure 24: Structure of **13-PF$_6$**.

Synthesis of 1,2-*bis*((2′-methyl-5′-(*N*-methylpyrid-4″-yl)-3′-thien-yl)perfluorocyclopentene, *bis*-ammonium hexafluorophosphate salt (13-PF$_6$): 140 mg **13-NO$_3$** was dissolved in a minimum amount of water and stirred. Excess ammonium hexafluorophosphate (10 equivalents, 141 mg) was dissolved in a minimum amount of water and added to the solution. This suspension was stirred for 2 hours in the dark and resultant light blue precipitate filtered. After rinsing the precipitate with water and ether, it was dried in *vacuo* (122 mg, yield unobtainable by excess of potassium nitrate in starting material). **M.p.** 228 °C (dec.); ^{1}**H NMR** (400 MHz, CD_3OD) δ 2.17 (s, 6H), 4.32 (s, 6H), 8.15 (s, 2H), 8.24 (d, J= 7.0 Hz, 4H), 8.76 (d, J = 7.0 Hz, 4H); **Anal.** Cacld. for $C_{27}H_{22}N_2P_2S_2F_{18}$: C, 38.49; H, 2.63; N, 3.32; Found: C, 38.43; H, 2.75; N, 3.08.

4.2.2.3 DNA Binding Studies

The DNA preparation was modified from Chaires *et al.*, 1982 (personal communication). 108 mg calf thymus DNA was resuspended in BPE buffer (6 mM Na_2HPO_4, 2 mM NaH_2PO_4, 1 mM sodium EDTA, pH 7.0) at 2 mg/ml concentration. This unsonicated DNA had an average length of ~2000 bp when visualized with ethidium bromide in a 1% agarose gel. This solution was used for preliminary DNA binding assays. 25 mL of the DNA solution was gently sonicated in an ice-water bath sonicator for a total of 1.5 hours. The sonication was performed by sonicating 5 minutes, bubbling nitrogen through the solution 5 minutes, and repeating until the total sonication time was reached. After each period of sonication and bubbling, a small aliquot of the solution was removed. The aliquots were analyzed on 2% agarose gels. The DNA length decreased to a steady size distribution of approximately 250 bp. This sonicated DNA was adjusted to 200 mM NaCl and treated with 0.5 mg/mL RNase A at 37 °C for 30 min. The solution was then treated with 0.5 mg/mL Proteinase K at 37 °C for 1.5 hours. The DNA was then washed in a separatory funnel with 25 mL phenol and centrifuged (in a desktop centrifuge) to separate aqueous and organic layers. The aqueous layer was removed, washed with 3 × 25 mL chloroform, and the aqueous layer removed after centrifugation as above. The DNA solution was then placed into a 6-8,000 molecular weight cutoff dialysis bag and dialyzed for two days against BPE or BPES buffer (6 mM Na_2HPO_4, 2 mM NaH_2PO_4, 1 mM EDTA, 185 mM NaCl, pH 7.0). The concentration (in basepairs) was determined by absorbance in a UV/Vis Spectrometer at 260 nm using the assumption 0.05 μg/μL = 1 absorbance unit/ concentration in base-pairs, 660 g/(mol base-pair).

Preliminary binding experiments were performed by incubating solutions of either the open or closed isomer of **13** and **14** (each at 1×10^{-5} M) with and without 1 mM CT-DNA (unsonicated DNA dissolved in BPE). The UV-visible absorbances of each sample were read, and referenced against a cuvette containing the appropriate background (including DNA).

These titrations were performed in the dark, to prevent the isomerization of one isomer to the other. Non-fluorinated *bis*-pyridinium switch **14-PF$_6$** was dissolved in BPE buffer (125.6 μM) and was switched into either the open or closed form. The titration procedure was the same for both the open and closed isomers. The open and closed solutions of **14** were each aliquoted out into a series of 1.5 mL tubes in the dark, with each tube containing the final concentration of 15.7 uM **14**. BPE buffer was added to each tube to bring the volume to one-half the final amount. The range of DNA examined was 1 nM to 1 mM bp. A series of stock tubes containing DNA concentrations from 2 nM to 2 mM (double the final concentration desired in the samples) were prepared by serially diluting the prepared calf thymus DNA with BPE buffer. DNA from the stock tubes was added to each aliquot of switch and the tubes vortexed. The reactions were allowed to incubate for at least 30 minutes. This minimum time needed for incubation was determined by spectrophotometrically analyzing a mixture of 1 mM DNA and switch. After 20 minutes there was no further change in the spectrum, indicating samples had reached equilibrium. Samples were analyzed on a Cary UV/Vis Spectrometer. Each sample was referenced against a cuvette containing the appropriate DNA concentration (diluted with BPE buffer).

The process used for analyzing data is described in the publication "Analysis of Drug-DNA Binding Data" (Qu and Chaires, 2000). The absorbances at λ_{max} for open and closed switch with 1 nm DNA were used for all plots. The absorbance for each trace at this wavelength was used, even though samples containing higher DNA concentrations experienced a red-shift (shift to longer wavelength) of the λ_{max}. The binding process can be described with the equilibrium expression:

$$D+S \rightleftharpoons C \qquad (1)$$

where D is ligand (switch), S is the DNA binding site (assumed to be a base pair), and C is the complex between switch and DNA. The absorbance observed in the titration is assumed to be the sum of the absorbance of both free and bound switch, each weighted by their concentrations:

$$A = A_0(C_t - C_b) + A_bC_b \qquad (2)$$

where A is the apparent absorbance observed at each DNA concentration, A_0 is the absorbance of free ligand, and A_b is the absorbance of bound ligand, C_t is the concentration of total ligand, and C_b is the concentration of bound ligand. The interaction of the ligand D with its DNA binding site S can be expressed as the quadratic equation:

$$KC_b^2 - C_b(KS_0 + KD_0 + 1) + KS_0D_0 = 0 \qquad (3)$$

where K is the association constant, S_0 is the total site concentration, and D_0 is the total ligand concentration. Titration data for each isomer was plotted in the form absorbance vs. log ([DNA], M (bp)). This data was fit by Dr. Jonathan Chaires using nonlinear least squares methods to obtain K, A_0 and A_b (using a fitting function in FitAll software).

Percent hypochromicity (decrease in absorbance) was calculated separately for each isomer using the equation:

$$\%H = [(A_f - A_b)/A_f] \times 100 \qquad (4)$$

Where A_f is absorbance of the free ligand and A_b is the absorbance of the ligand in the bound complex.

The change in λ_{max} (separately determined for each isomer) between the DNA-bound switch and free switch is given by the equation:

$$\Delta\lambda_{max} = \lambda_{max}(\mathbf{14} + \text{DNA}) - \lambda_{max}(\mathbf{14}) \qquad (5)$$

The values of λ_{max} used were 381 nm (open) and 669 nm (closed).

4.2.3 Results and Discussion

4.2.3.1 Synthesis

The synthesis for the fluorinated *bis*-pyridinium switch **13-PF$_6$** was achieved in six steps. *Bis*-pyridine **16** was synthesized from 2-methylthiophene (**20**) in four steps as described in the literature (Gilat *et al.*, 1995). Methylation of **16** with methyl iodide generated the *bis*-pyridinium switch. In order to separate the trace amounts of mono-methylated *bis*-pyridinium switch, a potassium nitrate gradient was successfully used as the eluent during silica gel chromatography. The product contained an excess of nitrate salt, which prevented weights to be accurately measured. To remove the excess salt, *bis*-pyridinium switch **13** (as the nitrate salt) was dissolved in water, and an excess of ammonium hexafluorophosphate was dissolved in water and added to the solution. As the nitrate counterions of the switch were exchanged for hexafluorophosphate, the compound became insoluble in water and precipitated out of solution. Any remaining switch with nitrate counterions remained in solution. This product was filtered and washed remove the remaining salt. The NMR spectrum of this product showed symmetry of the molecule, indicating both pyridiniums were associated with a PF_6 counterion. Although the PF_6 associated switch was less soluble in water than the nitrate salt, the concentrations needed for the DNA binding assays were so low (in the micromolar range) that it was not a problem.

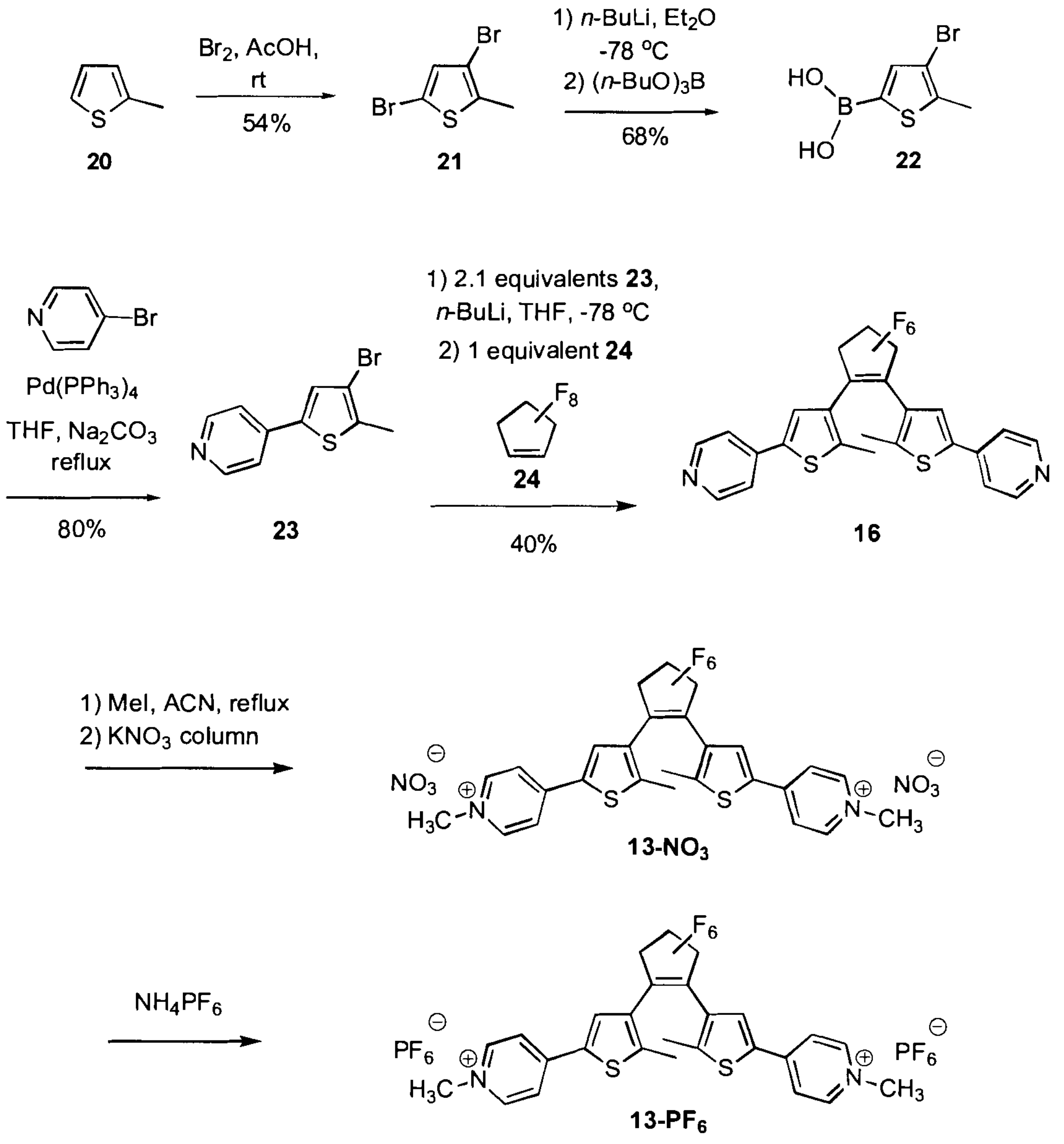

Br2, AcOH, rt
54%
1) n-BuLi, Et2O -78 °C
2) (n-BuO)3B
68%
20
21
22
Pd(PPh3)4
THF, Na2CO3
reflux
80%
23
1) 2.1 equivalents 23, n-BuLi, THF, -78 °C
2) 1 equivalent 24
24
40%
16
1) MeI, ACN, reflux
2) KNO3 column
13-NO3
NH4PF6
13-PF6

Figure 25: Synthesis of fluorinated *bis*-pyridinium switch **13-PF$_6$**. *Bis*-pyridine **16** was synthesized from 2-methylthiophene (**20**) in four steps as described in Gilat *et al.*, 1995. Methylation of **16** created the *bis*-pyridinium switch, and purification included an excess of nitrate counterions. Exchanging NO$_3$ counterions for PF$_6$ allowed isolation of **13-PF$_6$**.

4.2.3.2 UV-Vis Absorbance Spectroscopy

Since all binding experiments needed to be performed under biological conditions (to retain DNA secondary structure), the properties of these dithienylethene switches in these buffers needed to be assessed. The absorbance spectra of each isomer can vary depending on the solvent. As a result, the wavelengths needed to generate each isomer might vary. Also, the switching speeds of the two forms can vary depending on solvent. Since assays for DNA binding rely on decrease in switch absorbance upon complexation, lack of solubility would lead to a flawed experiment.

The photochemical isomerization between the two isomers of the non-fluorinated *bis*-pyridinium **14-PF$_6$** was followed by UV-Vis absorption spectroscopy (shown in Figure 26). A 2.5×10^{-5} M solution of **14** was prepared in BPES buffer. Due to the light green colour of the solid, the cuvette sample was converted to the ring-open form by irradiation with greater than 557 nm light. 30 minutes saw a complete disappearance of a minor peak at 669 nm. Irradiation of the solution of **14o** with 365 nm light resulted in the appearance of absorption bands at 669 and 436 nm, due to the production of the ring closed isomer **14c**. To the eye, the initially colourless solution turned blue as the ring-closing reaction proceeded. Three distinct isosbestic points were visible in the spectra, indicating the conversion from open to closed isomers was clean. To regenerate the ring-open isomer, the solution was irradiated with light greater than 557 nm for 30 minutes.

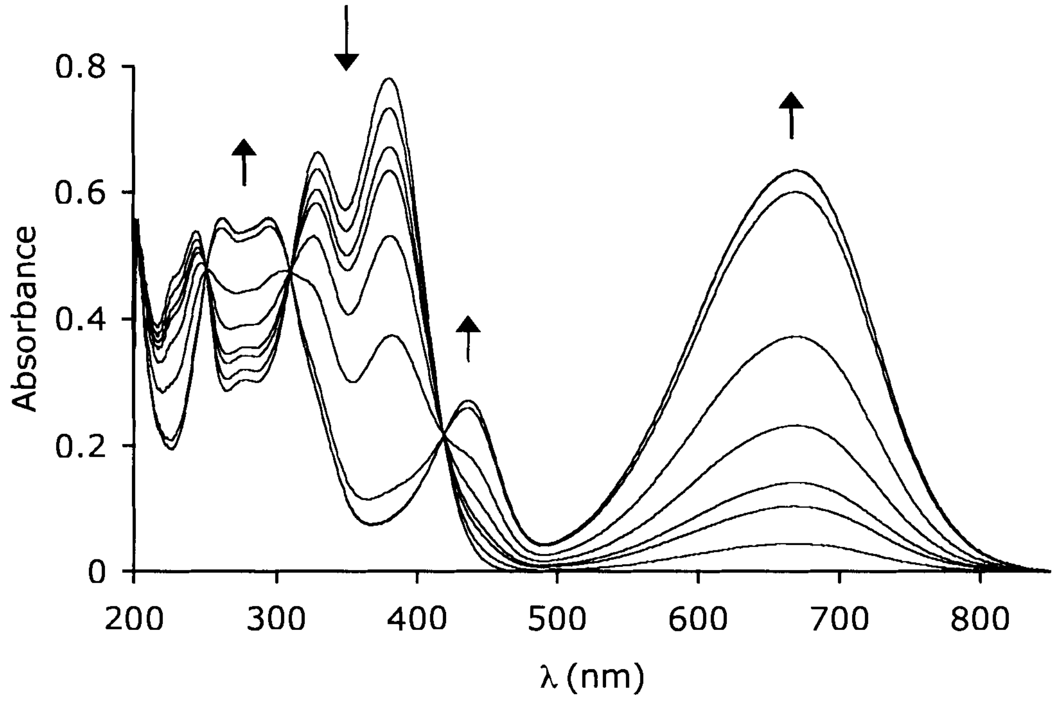


Figure 26: UV-Vis spectral changes of **14o** in BPES buffer (2.5 × 10^{-5} M) upon irradiation with 365 nm light.
Total irradiation periods are 10, 20, and 30 seconds, and 1, 2, 6, 12 and 15 minutes.

4.2.3.3 DNA Binding Studies

The first step when investigating possible binding of a small molecule to DNA is to see if DNA changes the absorbance spectra of the molecule. Interaction of the ligand chromophore with DNA can induce a change in the spectrum of the former's chromophore. These changes can include hypochromicity (absorbance decrease), hyperchromicity (absorbance increase), and blue or red shifts (shifts to lower or higher wavelengths, respectively). However, not all interactions cause a UV change. Therefore, absorbance changes are indicative of an interaction, but the absence of one does not remove the possibility of binding.

The absorbance spectrum of a given compound is observed both with and without DNA. In each sample, the concentration of the small molecule is the same. The concentration of ligand is the lowest concentration possible that still gives a reliable absorbance signal – generally in the range of 1-10 μM for DNA binding drugs (Qu and Chaires, 2000). The amount of DNA used is large enough to be considered a saturating concentration (usually 1 mM base pairs) in relation to the DNA binding ligand. In order to perform this experiment, it is critical that the ligand molecule has an absorbance that does not overlap with the DNA peak at 260 nm. Since DNA will be present at a much higher concentration than the ligand, it will saturate the signal. Fortunately, the open spectrum of each of the fluorinated (**13**) and non-fluorinated *bis*-pyridinium (**14**) switches contains a large peak far above 260 nm. The closed spectrum of each contains a λ_{max} in the visible region, far from the absorbance of DNA. Therefore, DNA binding assays were possible with both switches **13** and **14**.

The DNA chosen for these binding experiments was calf thymus in origin, since it is commonly used in the literature. This DNA is quite long (over 1000 bp), giving a very viscous solution when dissolved in buffer. It is necessary to sonicate the DNA, to break it into smaller fragments (approximately 150 bp). This reduces the viscosity of the solution, allowing more accurate pipetting. The next step is to treat the solution with the enzyme RNase A to remove any trace of RNA that could interfere with assessment of DNA binding. The preparation of DNA is performed in a sodium phosphate buffer containing 201 mM Na^+ (called BPES).

However, the sonication step requires a lower sodium concentration buffer (16 mM; called BPE) to decrease stabilization of the helix and to allow breakage. The DNA preparation procedure was adapted from the method of Chaires (Chaires *et al.*, 1982) and modified according to personal communication.

Initial absorbance studies were undertaken with the fluorinated *bis*-pyridinium switch **13**. The UV-Vis absorbance spectra of each of the closed and open isomers were examined both in the presence and absence of DNA. Both isomers showed hypochromicity upon addition of DNA, indicating binding was occurring (results not shown). While the open switch showed more hypochromicity upon DNA addition than the closed, the closed showed red shifting in the presence of DNA. In order to determine how fast these binding complexes formed, the hypochromicity was followed over time. This would indicate the minimum incubation time required for the equilibrium between free and DNA-bound switch to form. A 1 mM DNA concentration was chosen, as this is considered the saturating concentration of DNA in binding experiments. The sample was prepared and immediately placed into a cuvette. The absorbance of this sample was read repeatedly until there was no further change in the spectra. Over the course of these scans, the absorbance decreased until a steady spectrum was reached. This final spectrum was obtained after 20 minutes of incubation. Therefore, all further samples were incubated at least 20 minutes before any UV-Vis absorbance spectra were examined, to ensure equilibrium was reached.

While these experiments show an interaction between both switch isomers and DNA, it is not evident from this work whether the open and closed isomers have the same binding constants. Even though both isomers show a spectroscopic response at high DNA concentrations, one may have a much stronger affinity for DNA than the other. While it is tempting to make conclusions about the difference between open and closed binding based on the preliminary UV-Vis absorbance results, more data is needed. In order to get this information, a titration is performed in which DNA covering a large range of concentrations is added over a constant concentration of ligand. During the preparation of the DNA for titrating the fluorinated switch **13**, the DNA was dissolved in BPE (the first step). At this time, the non-fluorinated derivative **14-PF$_6$** was provided by another member of the lab (Brian Gorodetsky). Although the DNA preparation had not yet been completed, it was used for a preliminary spectral study with the non-fluorinated switch.

As with the fluorinated switch, these experiments showed hypochromicity of both the open and closed isomers in the presence of DNA. There was also red-shifting in the spectra, especially for the closed. The hypochromicity of both isomers upon addition of DNA was greater than for the fluorinated *bis*-pyridinium **13**. However, it wasn't clear whether this was attributable to the non-fluorinated **14**, or the effect of the lower sodium chloride concentrations in the BPE buffer (compared to BPES). It has been observed that higher salt concentrations can inhibit binding of small molecules with DNA (for an example see Chaires *et al.*, 1982). To determine this, additional samples were prepared that were supplemented with NaCl to equal the final sodium chloride concentration of BPES buffer. The increase in Na^+ concentration from 16 to 201 mM had a measurable effect, reducing both the hypochromicity and red shifting of both isomers (results not shown). Due to the greater spectral response of these isomers to DNA in BPE buffer, we decided to proceed with the titration in this buffer. Although the use of BPE was not ideal for comparison with other literature findings, our main goal was to determine if there was a difference in binding between the two isomers. The DNA preparation was completed, and in preparation for the titration, the DNA was dialyzed against BPE. The overall spectral effects of the addition of this prepared DNA to the switch (also in BPE) is shown in Figure 27.

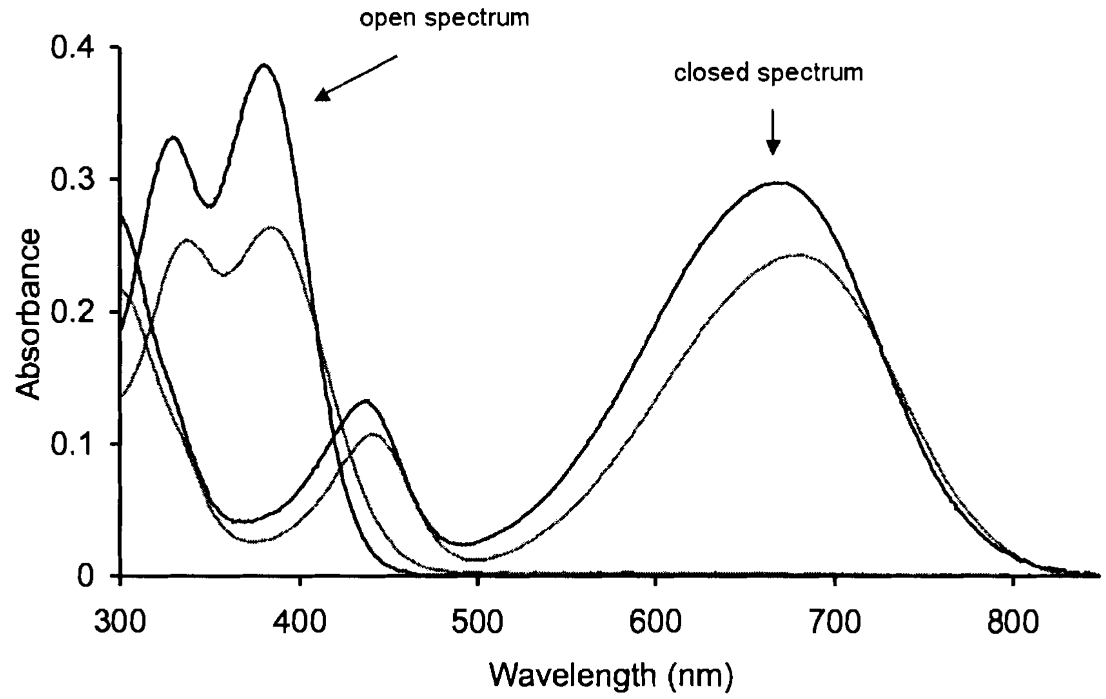


Figure 27: Spectral changes of the non-fluorinated *bis*-pyridinium switch **14** isomers (black traces) upon addition of 1 mM CT-DNA (grey traces).

The non-fluorinated *bis*-pyridinium **14** appeared to be the better candidate of the two switches for DNA titrations. Controlling for the effects of sodium chloride concentrations, it showed a slightly more pronounced UV change than the fluorinated switch **13**. Also, the similar extinction coefficients of both isomers of the non-fluorinated switch **14** allow the titrations to be done at the same ligand concentration, allowing a direct comparison between titration curves. The DNA concentrations for the titration ranged from 1nM to 1mM, with about 25 logarithmic concentrations being measured. DNA prepared and dialyzed in BPE was serially diluted using BPE buffer. Due to the incredibly wide range of DNA concentrations needed, it was not possible to titrate a cuvette directly with DNA. Instead, the concentrated prepared DNA was serially diluted and stocks (double the final desired concentration) were prepared. Separate reaction tubes were prepared for each DNA concentration.

The concentration of compound used is typically the lowest necessary to generate a reliable signal, taking into consideration any hypochromism that may result. A concentrated solution of switch **14** was prepared in BPE, and aliquots switched separately to the open and closed photostationary states. Despite the fact the extinction coefficients had been measured in BPES, these values were used for calculation of concentration of the BPE stock. Although the extinction coefficients in BPES may vary from those in BPE, the main focus of the titration was to ensure the two isomers were at the same concentration. The limited amount of **14-PF$_6$** provided prevented the repeat of the UV studies in BPE buffer.

Since the ligand we were interested in titrating is a molecular switch, it complicates the titration process. Each isomer must be titrated separately. In order to address the binding of each isomer properly, the titrations need to be performed in the dark to prevent any conversion to the other form during the course of the experiment. Prior to each titration, a solution of the desired isomer (containing more than enough for all samples) was generated by exposure of the cuvette containing it to the correct wavelength. Dilutions of these solutions were measured by UV-Vis spectroscopy. Achievement of the closed photostationary state was verified by no change between the spectra. The open solution was confirmed by the absence of an absorbance band in the visible region. To keep the isomers in the correct form, the entire titration and reading of samples on the UV-Vis spectrometer was done in the dark. The solution of starting isomer was found to be in the same form at the conclusion of the titration; this excluded the possibility that reversion to the other form caused the observed hypochromicity. As the closed switch had a slightly lower extinction coefficient (leading to higher error), additional data points were taken.

The primary absorbance data from the two non-fluorinated pyridinium titrations are shown in Figure 28. An isobestic point is identifiable in each isomer titration, but neither is very concise. A clear isosbestic point in titrations indicates one distinct mode of binding is taking place between the small molecule and DNA. Therefore the ligand exists in an equilibrium between two forms; bound and unbound. One possible reason for the imperfect isosbestic point in these titrations is that there may be multiple binding modes for each switch. If several modes of binding occur, the isosbestic point can be unclear. A more likely reason is that all samples were prepared in separate tubes, resulting in more pipetting error. In all UV-Vis absorbance spectra of the open titration, the visible region showed no absorbance. This indicated the switch in all samples had remained in the open form during the experiment, hence the reduction in absorbance was not due to isomerization.

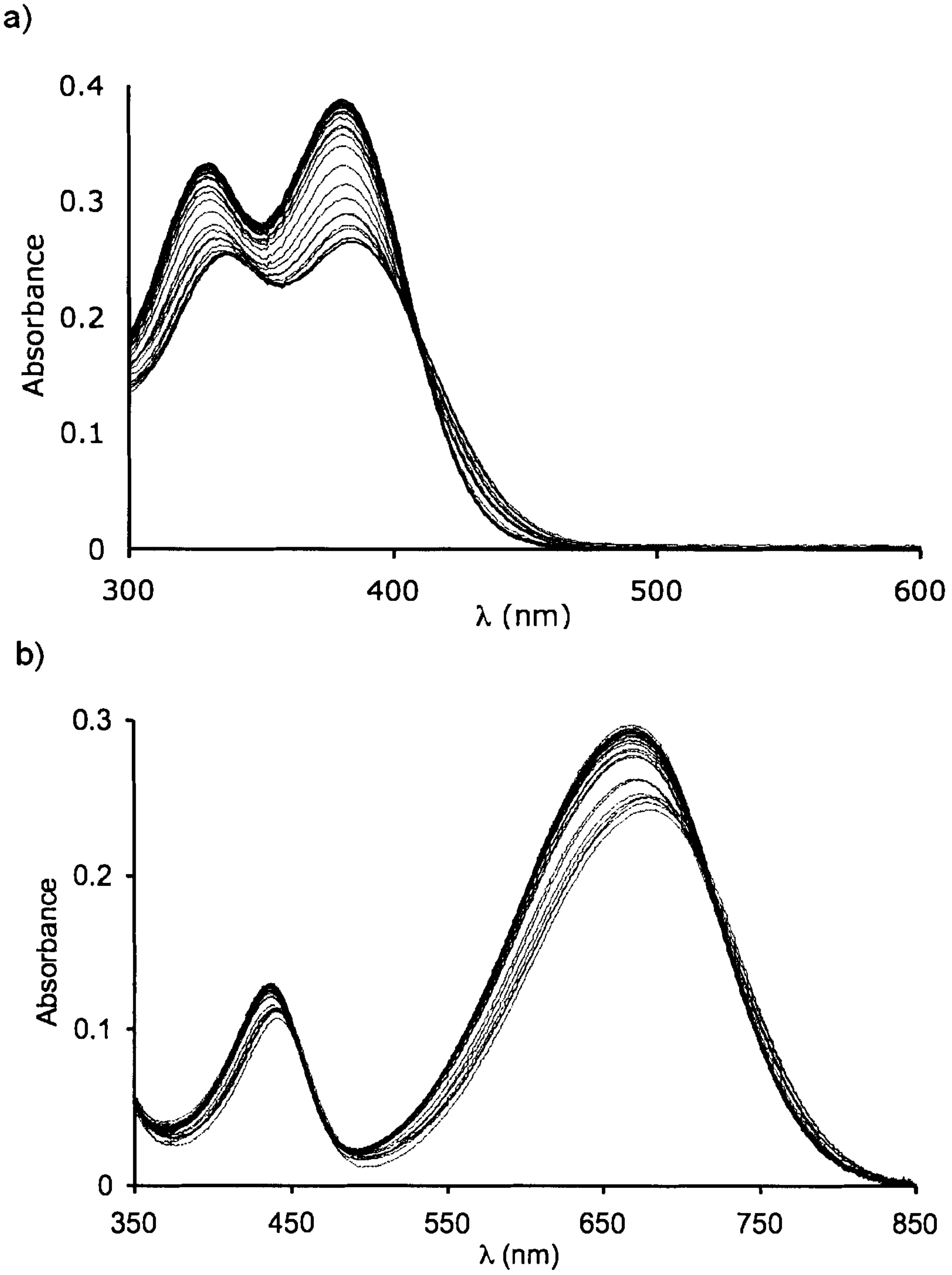


Figure 28: UV-Vis absorbance data from the titration of open (a) and closed (b) non-fluorinated *bis*-pyridinium switch **14** with CT-DNA.

To obtain a binding constant from this data, it is necessary to plot the primary data and fit it to equation. Our approach was to analyze the data using the procedure discussed in the publication "Analysis of Drug-DNA Binding Data" (Qu and Chaires, 2000). This approach is the most direct and rigorous way to analyze data obtained by this titration method. It allows the data to be fit directly without any graphical transformation (which can make the data noisier by distorting the error). This method is advantageous because it determines not only the binding constant K, but also the values of A_0 and A_b (the limiting absorbance values of free and bound forms of ligand, respectively). This method assumes the binding between ligand and DNA follows a bimolecular interaction. The equilibrium of such a reaction is described by equation 1 in section 4.2.2.3. The data was plotted in the form of absorbance response vs. log DNA concentration. Equation 2 is represents the absorbance response of the ligand. The optical properties A_0 and A_b can be calculated using equation 2. The quadratic equation 3 (section 4.2.2.3) can be solved to obtain the binding constant.

The absorbances at the λ_{max} of each isomer (381 nm for open, 669 nm for closed) were identified from the 1 nM DNA samples, and recorded from each sample spectrum. Figure 29 shows the plots generated for open and closed forms of the non-fluorinated pyridinium switch. Each curve shows a similar shape. This data was processed by Jonathan Chaires (University of Louisville). The information obtained using this method is given in Table 1. Information concerning hypochromicity and red shifting behaviour is also summarized in the table. The binding constant (expressed as $K/10^4$) obtained for open is 6.5 (± 0.6) M^{-1} and for closed it is 5 (± 0.8) M^{-1}. Unfortunately, the values do not differ significantly. There is a small deviation from the curve around 5.5 M log DNA. Chaires speculated that it could either be a concentration deviation, or an indication of stronger binding. Since the binding constants are relatively low, this probably indicates that the affinity of each isomer for DNA is not strong enough to prevent the ring closing or opening reaction. While this is attractive because the switch is still addressable in this environment, it is probably due to the fact that the binding is not very strong.

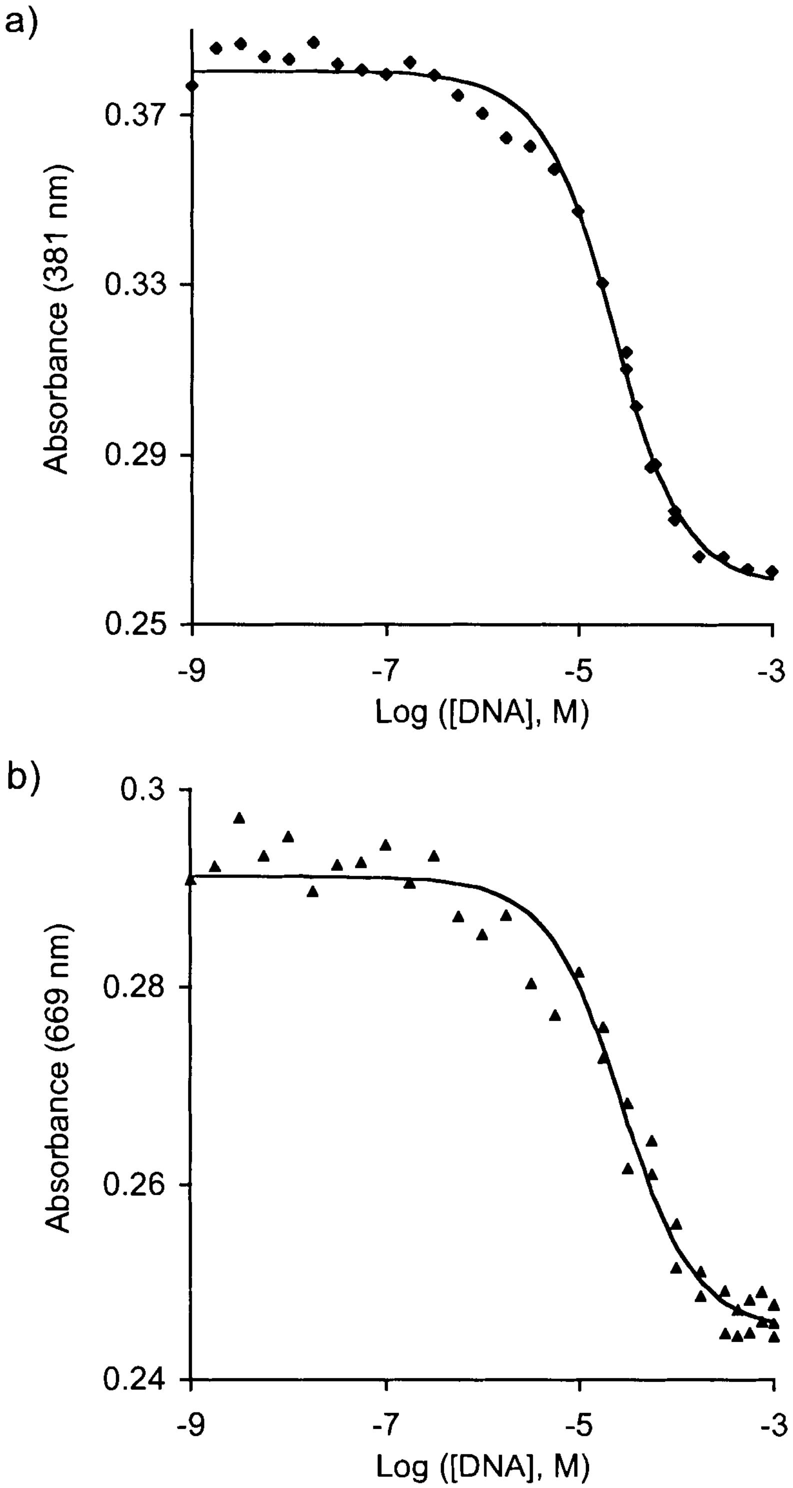


Figure 29: Curve fits for titrations of DNA with each a) open and b) closed non-fluorinated *bis*-pyridinium switch **14**.

Table 1: Data obtained from the non-fluorinated *bis*-pyridinium **14** titrations.

Isomer	**%H**[a]	**$\Delta\lambda_{max}$**[b]	**$\Delta\lambda_{max}$**[c]	**$K/10^4$, M^{-1}**	**A_0**	**A_b**
Open	32	+4 nm	_	6.5 (±0.6)	0.381	0.259
Closed	16	_	+10 nm	5.0 (±0.8)	0.292	0.245

a %H refers to hypochromic effect (equation 4 in section 4.2.2.3). $\Delta\lambda_{max}$ = λ_{max}(**14** + DNA) - λ_{max}(**14**); *b* λ_{max} 381 nm. *c* λ_{max} 669 nm.

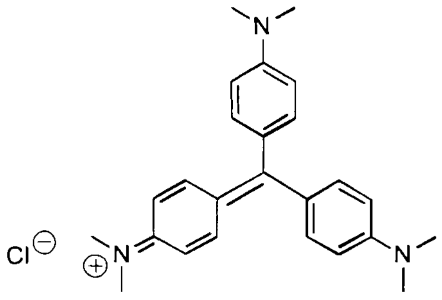
This experiment confirms that the non-fluorinated *bis*-pyridinium switch is the first reversible photochromic compound that shows an interaction with DNA under biological conditions, with both open and closed isomers binding with similar strength. While this is interesting, our hope of photoswitchable binding is not attainable with this molecule. The fluorinated analogue of this switch shows similar effects on its spectra at saturating DNA concentrations. While it was not examined by the same titration procedure, one would expect similar binding constants of the two isomers. Due to the similar structures of the two switches, they would probably bind in the same manner. At this point it is not known how the DNA binding properties of the non-fluorinated *bis*-pyridinium switch can be compared to other dithienylethene derivatives.

4.3 Bacterial Staining Studies

4.3.1 Introduction

Although the binding of the non-fluorinated *bis*-pyridinium DTE switch **14** to calf thymus DNA was not controlled by light stimulation, its optical properties made it a candidate for another biological application. As this switch is able to toggle between colourless, blue, and fluorescent states depending on wavelength, it is ideal for a visualization role. One established staining technique in cell biology is the Gram stain. This method relies on the migration of a cationic dye into the cell wall of bacteria. The differential ability to retain the dye through an organic wash distinguishes between two classes of bacteria (deemed Gram negative or positive). The successful adaptation of the DTE switch **14** as a bacterial stain would mean a new multistate stain, one capable of rendering the bacteria it is bound to colourless, blue or fluorescent. Also attractive is the possibility of reversibly switching the state of the stain bound on the bacteria. Controllable fluorescence would be particularly desirable in cell imaging, in which visualization of one individual cell/organelle/macromolecule tagged with a fluorescent molecule is not possible due to surrounding fluorescence signals. Potentially, staining with a dithienylethene fluorescent molecule would allow control over the extent of fluorescence. The composition of the switch (proportion of closed) bound to the bacteria could be adjusted by UV or visible light exposure, giving either a stronger or weaker fluorescence signal. Also, the amount of colour could also be adjusted by this method, using both fluorinated and non-fluorinated (**13**) *bis*-pyridinium analogues. Developing a dithienylethene stain into a technique such as the Gram stain procedure could allow a level of sensitivity and control not currently available to dyes.

The Gram staining technique is a very old, yet reliable method for differentially staining two major classes of bacteria, Gram positive and Gram negative (reviewed in Beveridge, 2001). Discovered by Hans Christian Gram in 1884, it relies on the permeability of the cell wall to take up, and retain, a dye known as crystal violet (Gram's stain). Gram positive cells have a thick outer cell wall, containing peptidoglycan. The Gram negative organisms have a thinner cell wall (which includes a layer of peptidoglycan), an outer wall comprised of lipids, and a periplasmic space which separates the two. Crystal violet (Figure 30) has a very large extinction coefficient (112 000; Zollinger, 1991) and is dark purple in colour. These optical properties make it a very good stain. The cationic structure of crystal violet is responsible for its binding to the cell wall, as the teichoic acids in the wall make it negatively charged.



ε = 112 000

25

Figure 30: Crystal violet (Gram's stain).

To begin the Gram stain procedure, fresh culture of the bacteria is smeared onto a microscope slide, and bound to the slide by passing it through a flame. Both gram negative and positive cells take up the crystal violet. To “fix” the dye, iodine (Gram’s iodine) is applied to the slide. A large insoluble complex forms between crystal violet and the iodine counterion, preventing the crystal violet from permeating back out into the surroundings. The next step is the differential step. “Decolourization” is achieved when a solvent, usually aqueous alcohol or acetone, is added to the smear. This dissolves away the thin lipopolysaccharaide cell wall of Gram negative bacteria. However in Gram positive cells, the solvent dehydrates the gram positive cell wall, concentrating it. As a result, this leaves the Gram positive cells purple and the Gram negative colourless. In the final step, the red-pink dye safranin is applied to visualize the gram negative cells. Safranin binds to the gram positive as well, however the dark purple appearance of these cells does not change. This technique is generally accurate, however several factors can lead to false findings. Overgrowth of colonies, over or under staining, and excessive/insufficient decolourization can lead to false positive or false negative results. Therefore, since these studies followed the experimental technique for Gram staining, care was taken to ensure that it was being followed properly. Colonies were grown between 20 and 24 hours, allowing large enough cells but preventing anomalous results that occur from overgrowth.

Before starting these experiments, there was no evidence suggesting that this compound would bind the bacteria. We relied on the structural similarity between the *bis*-pyridinium dithienylethenes (both fluorinated and non-fluorinated) and crystal violet. The cationic structures of both *bis*-pyridinium switches should allow them to be attracted to the cell wall the same way as crystal violet. Also, the comparable size should not only allow permeability, but also formation of an iodine complex large enough to prevent its dissociation from the wall. It was also not known if iodine would complex the same way as Gram's crystal violet. Lastly, it is possible that the green-blue colour, relatively low extinction coefficient and potentially weaker binding affinity would result in a weak stain. Even though fluorescence would ultimately be a more interesting approach for Gram staining, we examined slides with a visible light microscope only. Irradiating to generate fluorescence closes the molecule, losing the fluorescence signal. Also, we did not have access to a fluorescence microscope with the appropriate excitation/emission filters. Therefore it was easier to investigate any bacterial staining based on visible colour change of the cells.

4.3.2 Materials and Methods

4.3.2.1 General

Tryptic soy agar (Difco), bacteria (ATCC), and solutions of Crystal violet, Gram's iodine (BDH) and safranin were kindly made available from the SFU Biological Sciences teaching laboratory. Crystal violet (Anachemia) was from prepared as an 8 mM stock in water, and safranin (Fisher) was prepared as a 28.5 mM stock in 1:1 water/ absolute ethanol. Slides containing bacteria were visualized on a Zeiss light microscope and photographed on a digital camera.

Two Gram positive strains (*Bacillus subtillis*, *Staphylococcus epidermis*) and two Gram negative strains (*Citrobacter freundii*, *Enterobacter aerogenes*) were inocculated onto individual tryptic soy agar plates using septic technique. The plates were incubated at 37° for no longer than 24 hours. Individual colonies were picked and mixed with a drop of water to create Gram smears in on microscope slides. The smears were left until the water evaporated, and then heat fixed by passing through a flame. Solutions of dithienylethene switches **13** and **14** were prepared by weighing each switch (each as PF_6 salts) and dissolving the appropriate amounts of water to make 6 mM solutions. This resulted in a moderately soluble solution. As the nitrate salts of both pyridinium switches are more water soluble, excess mole equivalents of potassium nitrate were added to the vials and stirred vigorously. Small amounts of solid were left at the bottom of the vials, however this was not seen as a concern because the crystal violet solution is also filtered. The vials containing **13** and **14** solutions were converted to either open or closed forms using the appropriate wavelengths of light (over 550 nm for **14** and over 490 nm for **13**). The smear slides were stained using the either crystal violet as a control or the dithienylethene switches. Staining was achieved by adding several drops of the solutions from a glass pipet on to the smear for 3 minutes (up to a maximum of 10 minutes) for the switches and 1 minute for crystal violet. The smears were then rinsed with water until colour stopped leaching from the slide (approximately 10 seconds). Excess water was removed by shaking the slide and the area surrounding the smear was wiped dry with a tissue. For the fixing step, a solution of Gram's Iodine was flooded over the smear for 20 seconds, followed by water rinse and removal of excess water as above. The decolourization step was performed by dropping 80% ethanol over the smear for 20 seconds, followed by water rinse as above. For control experiments, the counterstain safranin was applied to the smear for 10 seconds, rinsed with water and dried by shaking and with a tissue. To prepare the slides for viewing, the slide (containing a small amount of water on the smear after each rinsing step) had a coverslip placed over the smear and lens oil was placed on top. The samples were viewed by oil immersion under the 100x power lens of a light microscope. Coverslips were used to prevent the bacteria from being dislodged over repeated viewings. Slides were observed after every step (however the same slide could not be used as the preparation for microscope viewing is not reversible).

4.3.3 Results and Discussion

The first series of experiments served as a control to confirm that correct Gram stain results could be reproduced. Even though the differential staining purpose and counterstaining was not an objective of this work, it was vital to run these controls to ensure the process was correct. The bacterial strains and all Gram stain reagents were obtained in a microbiology teaching lab, where this is a common procedure in an undergrad course. Two different strains of each Gram negative and positive classes were chosen. *Bacillus subtillis* (+) and *Citrobacter* *freundii* (-) were the best strains of each to use. *Staphylococcous epidermis* (+) is a slow growing strain resulting in small cells at the time the smear is created. This made the cell colour more difficult to see compared to the faster growing *Bacillus*. *Enterobacter aerogenes* had some potential contamination as two different sizes of colonies were obtained on the media plates. We did not attempt to find an alternate Gram negative strain. Successful results were obtained in the control Gram staining, as the Gram positive strains stained purple and Gram negative stained red-pink (results not shown).

The first attempts to substitute each *bis*-pyridinium switch (**13** and **14**) as the stains (in place of crystal violet) in the full Gram stain procedure gave inconclusive results. Both Gram negative and positive cells stained a unique orange-red colour, distinctive from the colour of the safranin control (results not shown). We concluded the counterstain was so strong that it essentially masked the colour of any bound switch. In order to prove this, I proceeded with every step in Gram staining (without the initial crystal violet staining) and I observed this same orange-red colour in both Gram negative and positive bacteria. So at that point it appeared that neither *bis*-pyridinium bound, or the safranin counterstain, which binds both types of cells, overwhelms any bound switch due to the latter's lower extinction coefficient. It was necessary to determine which of the two it was.

In order to see if **13** and **14** bound bacteria at all, cells were observed after the staining step (with the closed isomer of each switch as they are coloured). In order to investigate this, slides containing *Bacillus subtillis* (the easiest to determine colour under the microscope due to its large size) had several drops of the closed **13** or **14** switch solution applied to it. This volume used was comparable to that of crystal violet used in the control experiments. Rather than the one minute staining time for crystal violet, the switch stains were allowed to sit for ten minutes. After this time each slide was rinsed the same amount of time as crystal violet, to remove unbound switch. As with crystal violet, at the end of rinsing, no more colour was seen leaching from the slides. The smears of bacteria on both slides were visible to the eye as green spots; however the fluorinated switch generated a darker stain. Since fluorescence wasn't being investigated, and binding mechanisms of the two switches was expected to be similar, it was advantageous to use the darker version. In addition to staining with closed, a control was performed in which *Bacillus subtillis* was stained with open **13o** switch. Open staining resulted in a off-white clump of bacteria on the slide. The observed results from staining with both open and closed fluorinated *bis*-pyridinium switch **13** are shown in Figure 31.

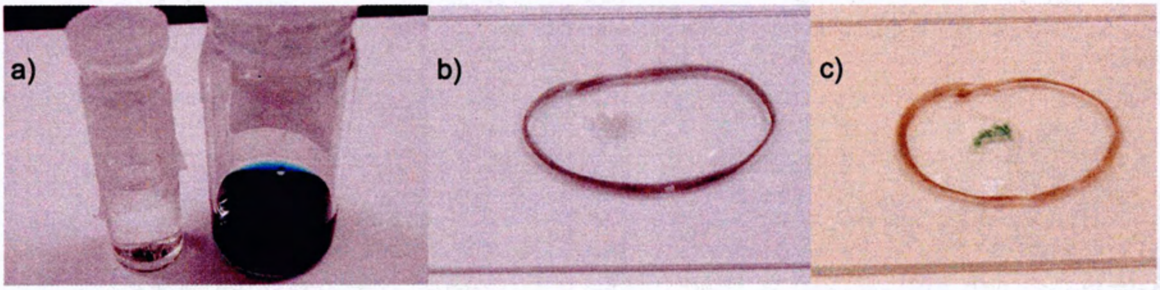


Figure 31: Bacterial staining experiment using fluorinated *bis*-pyridinium switch 13. a) Open (left) and closed (right) switch solutions used for staining. Colour differences visible to the eye of bacteria on microscope stained with either open (b) or closed (c) switch. Bacterial smears were stained 10 min, rinsed with water, and photographed.

The stained bacteria were examined under the microscope. In the slide stained with closed switch, the bacteria had a green colour (Figure 33a). The colour was most obvious in the large masses of bacteria in regions of the smear that were dense. While the intensity of colour was not very dark at all, the cells at least appeared to be stained. The open staining yielded opaque cells though the microscope. As an unstained control, smears of *Bacillus subtillis* were rinsed with water (to mimic the end result of the staining and rinsing process) and viewed. Even though the cells were unstained, they had a shadowy effect under the light of the microscope. Unfortunately this is not as attractive a control as one that was completely colourless.

It was possible that the large masses of bacteria in the smears were only trapping the stain solution, leading to the green colour. To investigate this, single cells were located on the slide (as shown in Figure 32). The cells are definitely green in colour, and the areas between closely located bacteria appear to be free of coloured solution. As well, cells that are perpendicular to the slide are darker green in colour, which supports the hypothesis the bacteria are stained. All bound switch throughout the full length of the bacteria would be viewed end-on, concentrating the colour. Therefore we concluded that the bacteria were indeed stained, and not just trapping green solution.

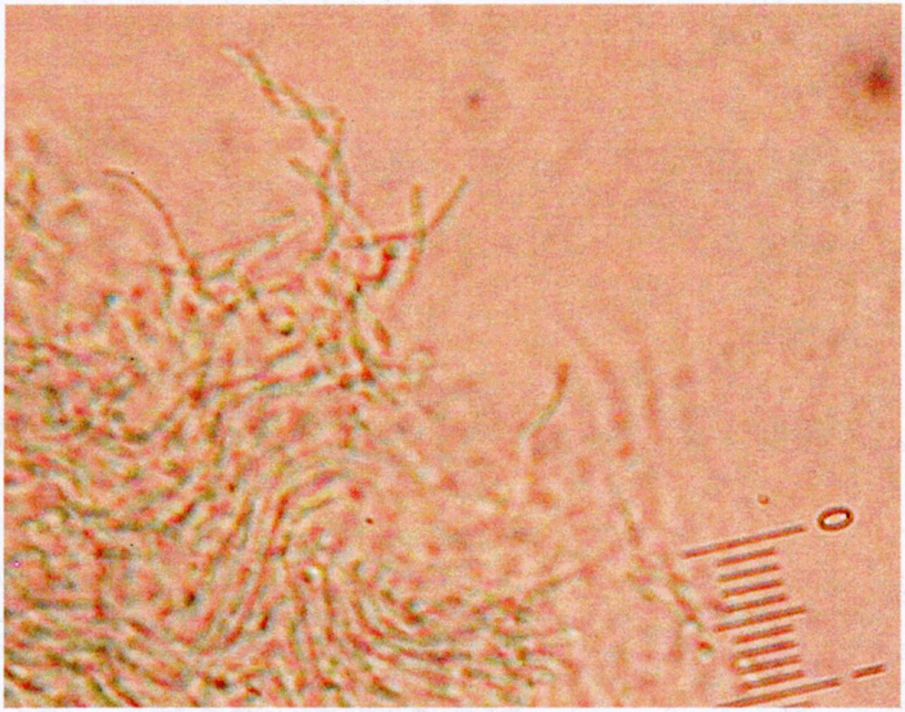


Figure 32: Single bacterial cells stained by closed *bis*-pyridinium switch **13c**.

We then wanted to investigate if the switch was capable of changing isomers while bound to bacteria. Since we knew the closed switch was bound due to the green colour of the cells, this slide was subjected to ring-opening conditions. The stained slide was placed (perpendicularly) in front of the tungsten light source with the correct filter to provide 490 nm light. As the slide is made of glass, it would not absorb in this region. After 30 minutes, the bacteria on the slide was more beige in colour. As shown in Figure 33b, the cells were much fainter green.

The phototochromism of the switch allowed interesting possibilites, but also led to some difficulties while performing these experiments. After staining with closed switch in the previous experiment, many photographs though the microscope were taken. Since the microscope shines light on the slide (and this contains light in the visible region), the switch appeared to get less green over time. As well, the open stain appeared to turn light green the longer it was viewed. This was expected to be a problem with the fluorescence experiments, although on a much faster scale.

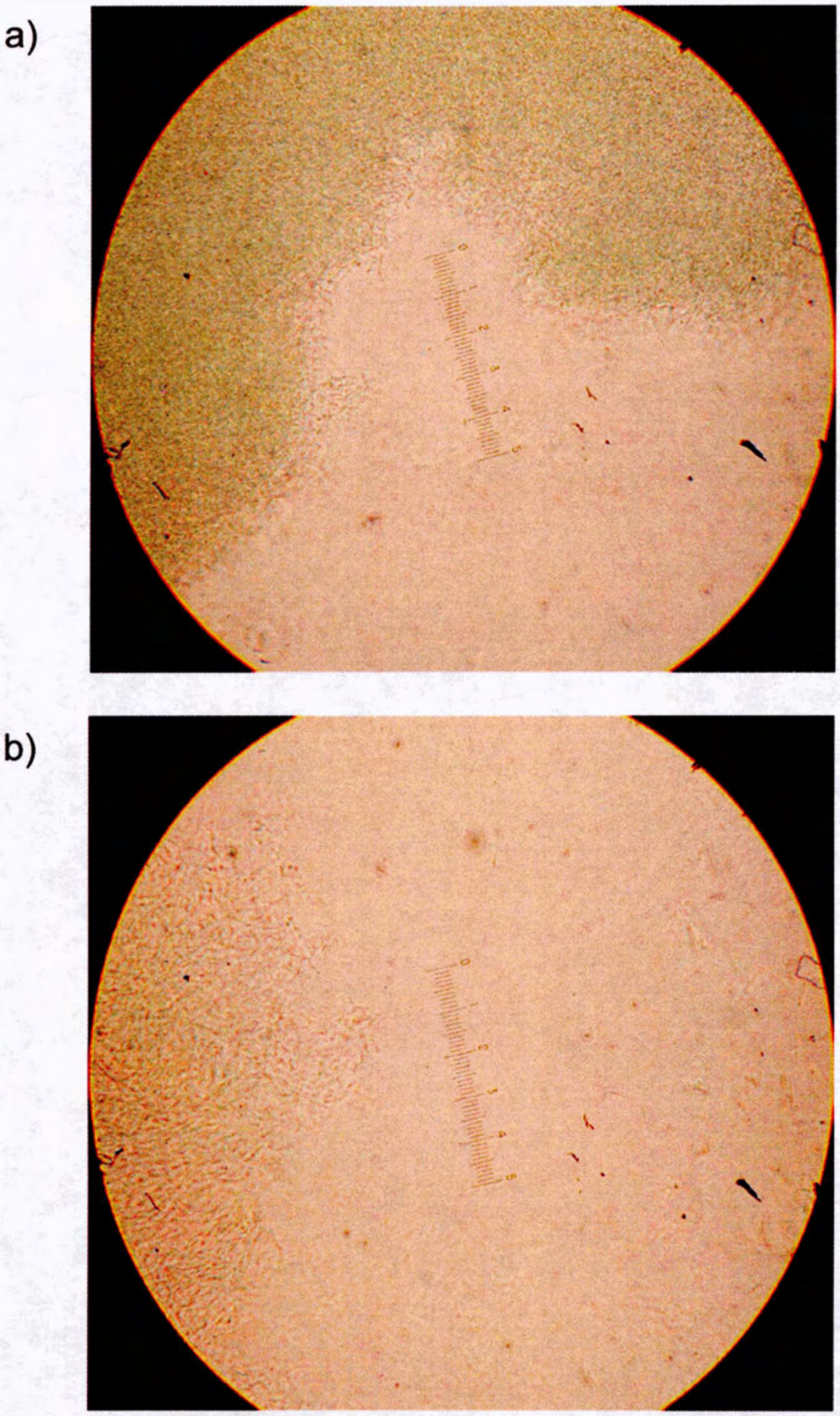


Figure 33: Colour change as seen through a light microscope of fluorinated *bis*-pyridinium switch **13** bound to bacteria.
a) *Bacillus subtillis* (fixed to a slide) stained with closed *bis*-pyridinium **13c**. b) Same slide exposed to >490 nm light for 30 minutes.

Staining of the Gram negative bacteria *Citrobacter freundii* was also attempted, however the cells were so small that determination of colour was difficult. Since the stain provided by switch **13c** was so weak, further investigation into the Gram staining procedure was difficult. *Bacillus subtillis* stained with closed switch was observed after the iodine fixing and decolourization steps. The cells still appeared to be stained, however as more of a yellow-green colour. One possible reason could be that the colour of the pyridinium changes after iodine addition, as we saw. There are two additional water rinses in these steps, which should further remove the pyridinium if not specifically bound. While this indicated the switch was bound the same way as crystal violet, it is not conclusive due to the light colour of the stain.

4.4 Concluding Remarks

This work has identified a dithienylethene switch that binds to calf thymus DNA. Although the ideal outcome of this project would have been the binding of only one isomer of the pyridinium switch (and the ability to turn binding on or off), titrations of each isomer with DNA revealed that both bind with the same affinity. This is interesting given the shape difference between the two switch isomers. It was initially thought that the different flexibilities would lead to variability in the modes or strengths of binding. Although it is impossible to determine the binding nature of a small molecule to DNA using UV-Vis spectral data only, the hypochromicity and red-shifting (although both mild compared to other intercalators) indicate there is some intercalation occurring. The most likely hypothesis is that both isomers bind via intercalation of one pyridinium ring into the DNA basepairs. This is a reasonable assumption, due to the planar charged nature of the pyridinium group. If this was the binding method, then the flexibility difference between open and closed does not impact the availability of the pyridinium to bind.

The aromatic substituents are the intercalative sites implicated by Saito in studies with the dithienylethene switch **12** and a DNA-containing film (Saito *et al.*, 2003). In the case of the pyridinium switches **13** and **14**, each pyridinium substituent is always able to rotate with respect to the neighbouring thiophene. The distance between the two pyridinium subsituents is too wide to fit both into the DNA helix.

While exploiting the structural changes of the *bis*-pyridinium switch **14** is not possible in this system, there is another possible application of this compound in the context of DNA binding. It may be advantageous that both forms bind DNA with similar strength. While bound to DNA, two other reversible properties could be manipulated. First, the differential fluorescence may be an attractive option for visualization of DNA. The reversible nature of this fluorescence could be useful, as the switch could be turned “off” (converted to closed) to visualize other stains present. The switch could then be activated for fluorescence (converted to open) and the fluorescence of the switch observed. The blue colour of the molecule could be used as well, however fluorescence is a more sensitive technique.

To briefly investigate a fluorescence application by this switch, the staining of agarose gel DNA samples with ethidium bromide was used as a model. DNA bands in agarose gels are usually visualized by using the fluorescent DNA-binding molecule ethidium bromide. This is often achieved by adding a solution of ethidium bromide (in buffer) to a container, and allowing the electrophoresed agarose gel to float in this solution with gentle shaking. The ethidium bromide molecules permeate through the agarose gel matrix during this staining procedure, and intercalate into the DNA (which is double-stranded) in the gel. The gel is next destained by soaking the gel in a solution of buffer alone, which helps diffuse unbound ethidium molecules out of the gel (preventing background fluorescence). Ethidium bromide experiences an increase in its fluorescence when intercalated in between the base pairs. The concentration of the fluorescent ethidium in the DNA bands, as well as the increased fluorescence over unbound molecules, allow visualization of these bands.

The staining of an agarose gel with the open isomer of *bis*-pyridinium **14** was attempted. Despite the addressability problem of the fluorescence, and the relatively low binding contants of both switch isomers with DNA, it was attempted to see if there was any success at all. To attempt this, an agarose gel was run containing DNA base-pair ladder. A solution was prepared of non-fluorinated switch **14** in 1× TBE buffer (at the same concentration as the ethidium bromide stain, 0.5 μg/mL) and exposed to over 550 nm light to generate the open form. Although maximum excitation wavelength for fluorescence is ~410, a light box of this wavelength was not available in the department. The light boxes available were 302 nm and 365 nm. An aliquot of the open solution was placed into an eppendorf tube, and irradiated with 365 nm light (~300 nm light was not attempted) to test if it was sufficient for fluorescence. Fluorescence was visible to the eye, however during irradiation the solution turned blue. This is similar to initial experiments with the fluorimeter, which showed that irradiating the cuvette of open switch (even with the maximum wavelength, ~410 nm) triggered the ring-closing reaction, enough to see a visible colour change to a dark blue solution (results not shown). The gel was stained in the solution of open switch for 10 minutes, and destained 5 minutes in 1× TBE buffer. To prevent isomerization of switch, everything was kept in the dark using tinfoil. The gel was visualized using the 360 nm light box. No fluorescent bands were observed, and there did not appear to be much fluorescence in the gel (although it is difficult to tell, as UV light is shone from below). At this point, I was concerned that the switch had not permeated the gel at all. However, after about 30 seconds of irradiation of the gel with the UV light, the whole gel turned blue. This proved that the solution had indeed permeated the gel in a uniform manner, at a concentration allowing colour detection by the eye. There were many possibilities why this qualitative experiment did not work. The low association constants may prevent meaningful binding. It may be that longer destaining was necessary, or the switch has no appreciable affinity for the DNA bands above that of the agarose. An additional concern is the salt conditions of TBE may negate binding, as a salt dependence has already been observed. Another possibility is that switch was diluted over the many bands present in the DNA ladder. Perhaps sample composed of a single band would have allowed a larger concentration of the switch and a higher fluorescence signal. Given the addressability problem of the fluorescence, this experiment was not pursued further.

While the bacterial staining experiments with the pyridinium switches was interesting and featured some success, it does not realistically provide an advantage over current systems. The faint colour, due to either weak binding or low extinction coefficients, is useless compared to conventional stains. However, higher concentrations of switch may be possible if prepared from an alcohol stock, as safranin is prepared. If the colour of these stains was more intense, the ability to switch them on the bacteria would be more useful. The fluorescence of the open non-fluorinated switch **14** posseses addressability problems, making it impractical for use beyond extrememly short visualization periods. Future studies of a dithienylethene stain for biomolecules could also attempt to covalently link the switch to a tag known to bind bacteria or a specific organelle.

5 TOWARD PHOTO-CONTROLLABLE RNA CLEAVAGE

5.1 Introduction

Ribonuclease A is an enzyme that catalyses the breakdown of the RNA backbone. This enzyme contains two histidine residues in the active site, which are responsible for catalytic activity. Many small synthetic constructs have been created which mimic the function of RNase A. The identification and improvement of enzyme mimics capable of degrading RNA is important for therapeutic applications, as RNA is an intermediate in gene expression and the genetic material of many pathogenomic viruses. The majority of these mimics have employed the use of functionalized cyclodextrins, complementary nucleic acid sequences, or small organic binding compounds to achieve substrate recognition. Some RNase A mimics have no methods of binding and have relied on hydrogen bonding or electrostatic interactions. However, none have possessed a level of control to cleavage ability to be turned on or off following introduction of the mimic into the reaction. In order to attempt this goal, we have created two potential mimics, each based on the dithienylethene molecular switch scaffold. If successful, it would be possible to change the configuration and resulting catalytic activity of an RNase A mimic by exposure to the correct light stimulus. The ability to switch this RNA cleavage activity on or off would be a significant development in the design of RNase A mimics.

5.1.1 Ribonuclease A

This 124 amino acid (13 686 Da) molecule is a relatively small enzyme. Much structural information has been gathered through NMR studies and crystallization (reviewed in Raines, 1998). RNase A contains two histidine residues that act as a general acid and base to cleave the RNA backbone. The catalysis by RNase A follows a bell shaped profile, attributed to the two titrable histidine residues. The reaction mechanism is by Raines in 1998. The non-protonated imidazole on His12 abstracts the proton from the substrate's ribose 2′-oxygen, aiding the in-line attack on the phosphorus atom. His119 facilitates the departure of the leaving group by protonating the 5′ oxygen. The reaction is thought to occur through a pentavalent phosphorus atom, which is stabilized by Lys41 and Phe120. The resolution of this intermediate generates two RNA fragments, a 2′, 3′-cyclic phosphodiester and free 5′ OH. The hydrolysis of the cyclic phosphodiester is a relatively slow process. The deprotonated His119 abstracts a proton from the water solvent, which nucleophilically attacks the cyclic phosphate. Through the same intermediate, His12 protonates the leaving 2′-oxygen. Lys41 has been shown to be an important residue in activity. Its role has been attributed to stabilizing the negative character that forms on the nonbridging oxygens of the intermediate.

5.1.2 RNase A Mimics

In the last few decades, several research groups have worked towards designing synthetic molecules that mimic the activity of Ribonuclease A. Generally, they have featured two imidazole or histidine residues to catalyze the breakdown of the RNA substrate. There have been a variety of different strategies used to approach this goal. In some designs, these functional groups have been incorporated onto scaffolds such as cyclodextrin, a complementary sequence of RNA or DNA, or binding (intercalating or groove binding) compounds for recognition of an RNA or model substrate. Others have improved existing DNA/ribozymes, including chemical groups for catalysis. Some have been small compounds that do not have any substrate recognition properties beyond several hydrogen bonds. Other designs not bearing imidazoles employ metal complexes, photochemical nucleases or complementary sequences tethered to a truncated enzyme domain. Nucleases of this type will not be discussed in this thesis.

The first RNase A mimic, created in 1978 (Breslow *et al.*, 1978), contained a cyclodextrin core. β-cyclodextrin is composed of seven glucose residues, circularly linked. This creates a concave shaped bowl. Due to the outward orientation of the hydroxyl groups on the saccharide units, the inner portion of this ring is hydrophobic. Cyclodextrin has been a popular scaffold for enzyme mimics. The pocket is used to accommodate hydrophobic side chains of the model substrates, and the rim is functionalized with groups to perform chemistry on the docked substrate. The mimics created by Breslow featured two imidazole groups, on various positions of the rim. However one drawback is the substrate is a simple cyclic phosphate, which does not resemble nucleic acids at all. This technique served more to probe the mechanism of cleavage by this mimic, not to actually mimic RNase A on natural substrates.

More recently, new approaches have been used to design RNase A mimics. Mimics which contain a motif for substrate binding in addition to a "catalytic site" have been quite successful. These methods have achieved success with actual RNA molecules (such as tRNA) as substrates. The first mimics utilizing this design featured tripeptides (to achieve catalysis), joined with a lysine linker to the known intercalator acridine (Tung *et al.*, 1992). Vlassov and coworkers synthesized two *bis*-histidine molecules, joined with variable length linkers (Podyminogin *et al.*, 1993). Although these molecules demonstrated low levels of cleavage of an RNA substrate, they found much higher rates of cleavage when these molecules were each tethered to a phenazine derivative. The phenazine portion of the construct intercalates into the double stranded portions of the tRNA molecule. This binding allows the *bis*-imidazole molecule to cleave the adjacent single stranded regions at much higher rates. While cleavage has been observed with mono-histidine analogues, the *bis*-histidine compounds are capable of catalysis at higher efficiency. Lown's group created RNase A mimics than work due to the specificity of the binding of lexitropsin to DNA. By modifying the lexitropsin backbone to contain modified imidazoles, they were able to observe RNA cleavage (Kissinger *et al.*, 1987). Although lexitropsin is a DNA binding molecule, the RNA substrate provided enough similarity to bind the mimic, but optimization is necessary (Helm *et al.*, 2001). Other mimics have oligos chemically linked to cleavers that are complementary to a single-stranded RNA sequence. As the oligo binds through base-pairing, the chemical cleavers cut the RNA backbone. Vlassov devoloped a spermine-imidazole construct capable of RNA cleavage (Vlassov *et al.*, 1995). Unlike the other methods developed, this molecule does not feature any sequence specific or pocket based binding preorganization. Since the molecule only features one imidazole, the other imidazole must be supplemented in the buffer.

5.1.3 Motivation

These attempts to effect and control cleavage of RNA are based on the scaffold of the dithienylethene molecular switches. Three imidazole-based dithienylethene constructs were synthesized to investigate any RNA-cleaving properties of these molecules. Each molecule had advantages and drawbacks. The first molecule synthesized was a *bis*-benzimidazole dithienylethene. The second molecule, prepared by David Sud, was *bis*-histidine dithienylethene. A mono-histidine dithienylethene was also synthesized to act as a control for the *bis*-histidine studies.

The goal of this project was to design a dithienylethene derivative containing two groups capable of cleavage on opposite ends of the molecule, as substituents on the thiophene rings (Figure 34). Ideally, catalysis would be performed by one isomer of the switch only; the other form would be rendered inactive. This would allow the process of RNA cleavage to be turned on of off by stimulation with the appropriate wavelength (either UV or visible) of light. The conformational change induced by light would change the orientation of the two groups. The two moieties may be positioned in approximately the right position for cleavage of the RNA backbone. Although there has been limited success in creating RNase A mimics that do not contain binding regions, we proceeded with this because it can dissociate/be turned off in solution, as opposed to ones which remain bound.

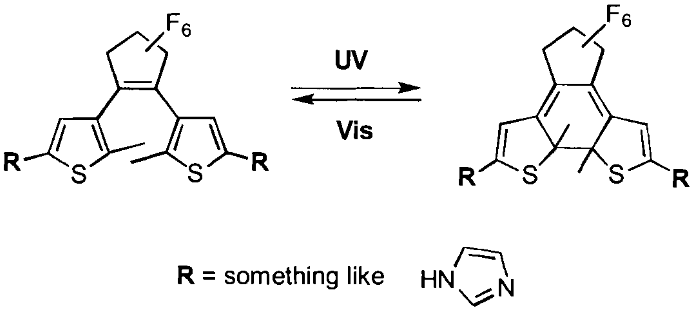


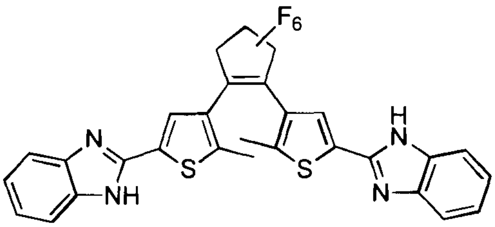
Figure 34: General scheme of a RNase A mimic dithienylethene

There were two possibilities in which catalysis would be allowed by one isomer only. One uses a strictly structural argument. As the open isomer is flexible, with rotation allowed throughout several bonds, this may allow the molecule to attain the proper geometry of the catalytic groups. Otherwise, the closed molecule may be too constrained in its rigid structure to allow for proper positioning. The opposite also may be true: perhaps the closed molecule is fixed in a more optimal configuration, while the open form is too flexible. A second argument relies on the potential change of pKa of the system upon ring closing. When open, there is no communication between the distal histidine residues. However in the closed form there is conjugation between the backbone, and the two histidines are in communication. Potentially, the pKa vales of the two histidines may be different in the closed form

5.1.4 Design of Novel Switches

Because of the dithienylethene backbone, there are limitations in the number and location of functional groups. It was necessary to keep the structure of the basic dithienylethene molecule to retain switchability, but also add groups such as in RNase A. The first target molecule would have contained two imidazole groups, in the **R** position. An advantage of this molecule would be the biological pKa of these histidines. A second benefit would be the structural differences between the open and closed forms of this switch. It is feasible that since the histidines in the closed isomer would be in conjugation, the pKa of the residues could be altered from that of the open form. Perhaps in this scenario, the pKa of the cleavage reaction could be turned on or off, thus controlling the cleavage reaction. However, this would be synthetically more challenging and would require more steps, as the nitrogens would need to be protected.

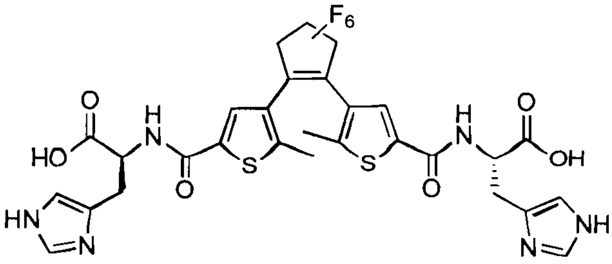
Due to the limitations mentioned above, the first target molecule synthesized for this project was a novel dithienylethene containing one modified imidazole group on each end of the molecule (Figure 35). Synthetically, this was a much easier molecule to obtain. The benzimidazole groups are positioned directly next to the thiophene rings, placing them in conjugation. Also, the structural differences are large between the open and closed forms. While there are some benefits to this molecule, there are also some drawbacks. First, while benzimidazoles are well reported in the literature to be present in many DNA binding agents, there are no mentions of cleavage activity. There may be several reasons: one may be the lowered pKa compared to histidine. Another is that the benzimidazole may be chemically very different from imidazole. A third potential problem is the connection of the benzimidazole to the switch between the two nitrogens of the substituted imidazole ring may alter the pKa even more. Another problem is that the nitrogens may be more hindered, not being on the ends of the molecule. Also problematic is the lack of polar subsituents on the molecule.



26

Figure 35: Structure of the *bis*-benzimidazole dithienylethene switch (**26**).

The second RNase A mimic molecule designed was a *bis*-histidine variant (Figure 36). As histidines perform the chemistry in RNase A, they are a logical choice. The pKa of histidine is neutral, which is ideal for biological studies. Also, there is literature precedent for histidine functional groups on RNA cleaving molecules. The *bis*-histidine switch features more polar substituents than the benzimidazole, which improves solubility in aqueous environments. However, this switch does not have some of the desirable properties of the benzimidazole version. As the histidines are connected by five bonds to the thiophenes (not one as in benzimidazole), there is still flexibility in the closed form. Also, this extra distance in connectivity between the distal imidazole rings makes the possibility of a pKa change between open and closed unlikely.



27

Figure 36: Structure of the *bis*-histidine dithienylethene switch (**27**).

The crystal structure of RNase A complexed to a substrate was solved in 1994 (Zegers *et al.*, 1994). The distance between the nitrogens of each imidazole (acting as the general acid and general base) was shown to be 7.4 Å. Each isomer of both the *bis*-benzimidazole and *bis*-histidine switches were viewed in the program Chem 3D. This program allows the user to rotate individual bonds throughout the molecule. The open isomer of the *bis*-benzimidazole switch was shown to be able to bring the imidazoles 7.4 Å apart. The closed *bis*-benzimidazole switch can bring the two imidazole nitrogens only 9.6 Å apart. As Figure 37 shows, the difference in flexibility of the thiophene rings and their substituents between the open (a) and closed (b) isomers is critical. On the other hand, both *bis*-histidine isomers are capable of placing the imidazole rings the correct distance apart (Figure 37c and d).

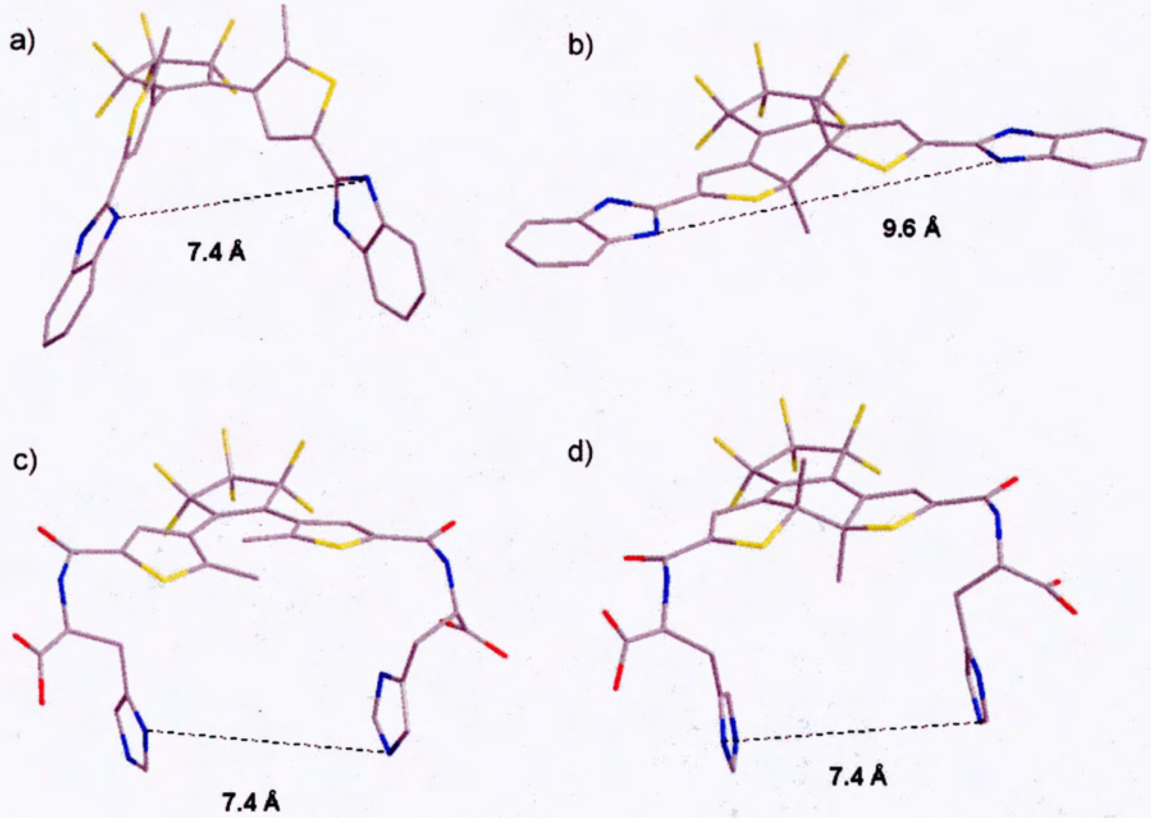


Figure 37: Abilities of *bis*-benzimidazole and *bis*-histidine switches to adopt RNase A distance between imidazole rings.
bis-Benzimidazole and *bis*-histidine switches were visualized using the program Chem 3D and subjected to rotation throughout bonds. *bis*-Benzimidazole switch isomers open and closed are shown in (a) and (b), respectively. The closest approach achievable by the closed isomer is 9.6 Å. *bis*-Histidine switch isomers open and closed are shown in (c) and (d) respectively. Atom colours are as follows: carbons are grey, nitrogens are blue, oxygens are red, and sulfurs and fluorines are yellow. Hydrogen atoms are not shown.

5.1.5 RNA Substrate

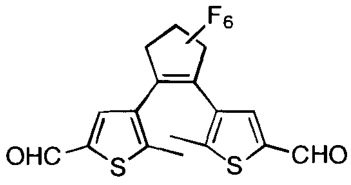
There were a number or factors that determined the nature of RNA substrate used in these assays. First, it was imperative to use a single sequence rather than a mixture. This way, exact cleavage sites can be mapped by using an end labelled RNA sequence. If using a mixture of sequences, cleavage at many sites could result in a smear of products rather than concise bands, and mapping studies would not work. Secondly, we wanted to use a substrate with some known secondary structure information. This would allow us to identify possible binding locations of the switches by analyzing any cleavage site patterns. Lastly, we wanted to choose a substrate within the department, allowing easier coordination of reagents. We chose the clone 7.02 isolated from the 4-thiouridine nucleotide synthase deletion selection (Section 2.1.2.3). This clone, one of the fastest isolated, had already been the subject of kinetic and structural studies (Chapple *et al.*, 2003). This collaboration was advantageous as I was familiar with the experimental protocols with this RNA. The RNA cleavage assays undertaken with this ribozyme did not involve its catalytic properties. However, we did use buffer conditions conducive to folding to allow it to form its secondary structure.

5.2 Materials and Methods

5.2.1 General

Histidine (Aldrich) was used as obtained. T4 polynucleotide kinase (and corresponding 5× Exchange and 5× Forward buffers) and basepair ladders (both 10 and 100) were purchased from Invitrogen. Calf intestinal phosphatase and 10× NEBuffer 3 were purchased from New England Biolabs. T7 RNA polymerase was previously made in the laboratory with activity comparable to commercial enzyme. α-UTP (800 Ci/mmol) and γ–ATP (6000 Ci/mmol) were purchased from New England Nuclear/Perkin Elmer. T1 RNase (Fermentas), tRNA, transcription reagents and 7.02 DNA were generous gifts from the Unrau lab. Ethanol precipitations were performed by adjusting the volume of the nucleic acid sample to 300 mM sodium acetate, adding 2.5× of the total volume of ethanol, vortexing and cooling at -20° for at least 30 minutes. Pellets were obtained by centrifuging tubes at 13,200 rotations per minute in an Eppendorf 5415 D centrifuge for at least 30 minutes in a 4° cold room. Ribozyme samples were heated and cooled in a Perkin Elmer thermal cycler. Radioactive polyacrylamide gels were visualized using a Typhoon Phosphorimager (Amersham). Quantitation analysis was performed using ImageQuant software (Amersham) and Microsoft Excel. Molecular modeling was performed using the CS-MOPAC Pro software package as implemented in Chem3D Ultra 8.0 (CambridgeSoft, Inc., Cambridge, MA, USA). Models were subjected to a semi-empirical energy minimization at the AM1 (closed-shell restricted) level of theory (Dewar *et al.*, 1985)

5.2.2 Synthesis



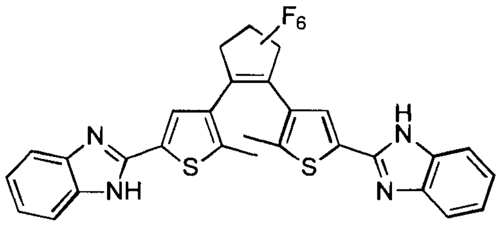
28

Figure 38: Structure of **28**.

Synthesis of 1,2-*bis*(5'-formyl-2'-methyl-3'-thien-yl)perfluoropentene (28)

This synthesis of **28** was performed under reaction conditions listed in Lucas *et al.*, 1999. The known intermediate 1,2-*bis*(5′-chloro-2′-methylthienyl-3′-

yl)perfluorocyclopentene **30** (Figure 42; Lucas *et al.*, 1999) was prepared as described (Peters, 2003) and used as starting material. This dichloro switch **30** (156.6 mg, 0.36 mmol) was dissolved in dry tetrahydrofuran (25 mL) at –78 °C under nitrogen in the dark. *t*-Butyllithium (1.7 M in pentane, 1.5 mL, 2.7 mmol) was added dropwise until starting material disappeared by TLC to give a dark brown solution. Dry dimethylformamide (500 μL, 6.4 mmol) was added quickly and the reaction stirred 15 minutes. The reaction was then allowed to warm to room temperature and then stirred with and equal volume of saturated ammonium chloride for 10 minutes. Layers were separated and the aqueous layer was extracted with ether (3 x 50 mL) and the combined organic layers washed with water (2 x 50 mL), and brine (2 x 50 mL). The layers were dried (Na_2SO_4), filtered and evaporated under reduced pressure to give the crude product as a brown oil. The pure product was purified using column chromatography through silica gel (dichloromethane:hexane = 2:1) to yield dialdehyde **28** as a yellow solid (54.8 mg, 36% yield). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 2.03 (s, 6H), 7.74 (s, 2H), 9.85 (s, 2H).



26

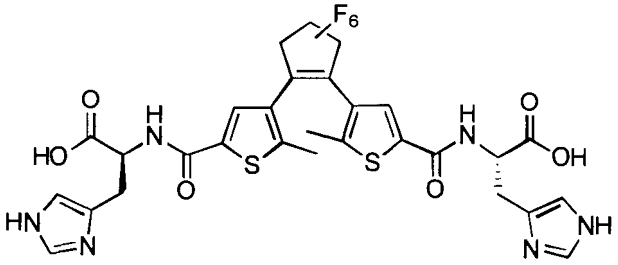
Figure 39: Structure of **26**.

Synthesis of 1,2-*bis*((benzimidazoyl-2''-yl)2'-methyl-thien-3'-yl)perfluorocyclopentene (26)

Adapted from Vanden Eynde *et al.*, 1995.

54.8 mg dialdehyde **28** (0.126 mmol) was dissolved in a minimal amount of *N,N*-dimethylformamide (DMF, ~10 mL) and transferred to a flask containing 2.2 equivalents *o*-phenylene diamine ((**31**; Figure 42), 29.9 mg, 0.277 mmol) and 20 mL acetonitrile. After 10 minutes of stirring in the dark at room temperature 2.2 equivalents of 2,3-dichloro-5,6-dicyano-1,4-benzoquinone (DDQ, 62.7 mg, 0.277 mmol) were added. The reaction was left to stir for 24 hours covered with a septum to prevent solvent loss. Then 2.2 more equivalents of each diamine and DDQ were added and the reaction continued another 24 hours. After this time 40 mL of water was added and the contents stirred overnight. The layers were separated and the aqueous layer extracted with 3 x 50 mL ethyl acetate and the combined organic layers washed with 3 x 50 mL each of water and brine. The organic layers were dried with Na_2SO_4, filtered and dried in vacuo. The crude mixture was dry loaded onto a silica gel column and the product isolated using the eluent methanol / dichloromethane (2.5% methanol increasing to 5%). The pure product was recrystallized from 95% ethanol as beige powdery crystals

(21.1 mg, 27% yield). **M.p.** > 300 °C; **^{1}H NMR** (500 MHz, CD_3OD) δ 2.08 (s, 6H), 7.24-7.29 (m, 4H), 7.56 (br s, 4H), 7.86 (s, 2H). **Anal.** Calcd. for $C_{29}H_{18}N_4S_2F_6$: C, 57.99; H, 3.02; N, 9.32; Found: C, 57.99; H, 3.66; N, 8.65. MALDI: 601 $C_{29}H_{18}N_4S_2F_6$ (600.61).



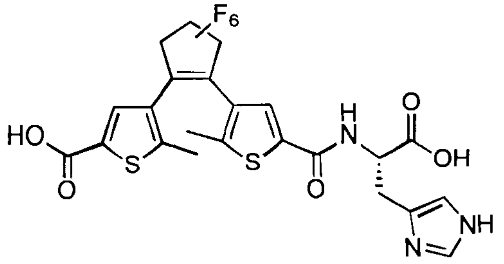
27

Figure 40: Structure of **27**.

Synthesis of 1,2-*bis*((2"(*S*)-(5'''-imidazoylmethyl)carboxylethyl)carboxamido-5'-yl-(2'-methylthien-3'-yl)))perfluorocyclopentene (27)

Adapted from Yuan *et al.*, 2002; synthesized by David Sud.

A CH_2Cl_2 solution of oxalyl chloride (1.53 mL, 2M, 3.07 mmol) was added dropwise over 15 minutes to a CH_2Cl_2 solution (10 mL) of diacid intermediate **32** (Figure 47; 100mg, 0.22 mmol) cooled at 0°C in an ice-water bath. A catalytic amount of DMF (20 µL) was added to the reaction mixture. The cooling bath was left to warm to room temperature and the reaction mixture was stirred overnight. The solvent and unreacted oxalyl chloride were removed under vacuum. The crude acid-chloride product **33** (Figure 47) was placed under high vacuum overnight to remove traces of DMF. To a suspension of L-histidine (67.96 mg, 0.438 mmol) in H_2O (0.5 mL) cooled with an ice-salt bath, solid potassium hydroxide (KOH) (28.91 mg in 85% content, 0.438 mmol). After all solids were dissolved, THF (2.5 mL) was added. Using an injection pump, a THF (2 mL) solution of the crude acid chloride (0.22 mmol) and a H_2O (0.5 mL) solution of KOH (28.91 mg in 85% content, 0.438 mmol) were added simultaneously over 6 hours. Cooling of the reaction mixture with an ice-salt bath was maintained during the addidtion and for an additional 3 hours. The reaction was then left to warm to room temperature and stirred overnight. THF was removed by evaporation under vacuum. Formic acid was added dropwise until the pH of the solution reached 4. The grey brown solid formed was washed with CH_2Cl_2 (7 mL), CH_3OH (7 mL) and acetone (7 mL) to yield a light grey powder. The product was then dissolved in a mixture of CH_3CN (15 mL), H_2O (2 mL) and 28% NH_4OH (1 mL) and dry-loaded onto silica. The crude was further purified by flash chromatography (silica, gradient from 9:0.5:0.5 $CH_3CN/H_2O/$ 28 %NH_4OH to 9:1:1 $CH_3CN/H_2O/$ 28 %NH_4OH) to yield *bis*-histidine **27** as a white powder (40 mg, 25%). **M.p.** 242 °C (dec.); **^{1}H NMR** (500 MHz, $(CD_3)_2SO$), δ 1.88 (s, 6H), 2.96 (dd, *J*= 17, 9 Hz, 2H), 3.05 (dd, *J*= 15, 5 Hz, 2H), 4.48-4.60 (m, 2H), 6.81 (s, 2H), 7.57 (s, 2H), 7.89 (s, 2H), 8.93 (d, J= 7 Hz, 2H).



29

Figure 41: Structure of **29**.

Synthesis of 1-((2"(*S*)-(5'''-imidazoylmethyl)carboxyethyl)carboxamido-5'-yl-(2'-methyl-thien-3'-yl)))-2-(5'-carboxy-2'-methylthien-3'-yl)perfluorocyclopentene (29)

The mono switch **29** was obtained during the synthesis of **27** as a white powder (13 mg, 10%). **M.p.** 195 °C (dec.). **^{1}H NMR** (500 MHz, $(CD_3)_2SO$) δ 1.89 (s, 3H), 1.93 (s, 3H), 2.94-3.10 (m, 2H), 4.52-4.59 (m, 1H), 6.85 (s, 1H), 7.58 (s, 1H), 7.64-7.68 (m, 1H), 7.89 (s, 1H), 8.95 (d, *J* = 7.9 Hz, 1H).

5.2.3 RNA Substrate and Standards Preparation and Purification

5.2.3.1 RNA Substrate

The 124 basepair DNA sequence of the 7.02 isolate from the 4-thiouridine synthase ribozyme selection was amplified through polymerase chain reaction (sequences of primers and conditions in section 2.1.2.2). This DNA was made available by the Unrau laboratory. Analysis of this product was performed by gel electrophoresis of the sample in a 2% agarose gel and visualized with ethidium bromide. The product showed up as one clean band of the appropriate size in comparison with a 100 basepair ladder. The DNA was precipitated and the pellet was resuspended at 10× PCR concentration in transcription buffer (contents in section 2.1.2.2). The sequence was transcribed to yield internally labelled RNA as described in section 2.1.2.2. The resulting RNA was ethanol precipitated to remove much of the unincorporated isotope. The RNA pellet was resuspended at a 1x concentration in water and an equal volume of loading dye (bromophenol blue, xylene cyanol, formamide, 10 mM EDTA) was added. The RNA was purified by loading the sample in an 8% denaturing polyacrylamide gel and excision from the gel after visualization with autoradiography. The sample was eluted from the gel slice by immersion in TE buffer (10 mM Tris-HCl, 0.1 mM EDTA, pH 8.0) and ethanol precipitated. The pellet was recovered as above and resuspended in TE at 1× transcription concentration. The RNA concentration was determined by absorbance in a Cary UV/Vis Spectrometer at 260 nm using the assumption 40 μg/ml ssRNA = 1 absorbance unit. RNA without a 5′-^{32}P label was obtained by transcribing the PCR product as above except without isotope. This RNA was purified as above and visualized by UV shadowing. This band was eluted from the gel, precipitated, resuspended at 6.37 uM in water and kept in aliquots in the freezer. 100 pmol of this stock was treated with 1 unit/μL calf intestinal phosphatase in the presence of 100 mM NaCl, 50 mM Tris-HCl pH 7.9, 10 mM $MgCl_2$ and 1mM dithiothreitol (supplied as 10× NEBuffer 3) at 37° for 1 hour. After this time another aliquot of enzyme was added (to bring the overall enzyme concentration to 1.8 units/μL) and the reaction left for one further hour at 37°. The phosphatase was removed by two consecutive phenol chloroform extractions (in the presence of 300 mM sodium acetate) and the RNA was ethanol precipitated. 20 pmol phosphatase-treated RNA was end-labelled by incubation with γ-ATP ^{32}P, 100 mM KCl, 70 mM Tris-HCl pH 7.6, 10 mM $MgCl_2$, 1 mM 2-mercaptoethanol (supplied from 5× Forward reaction buffer) and 0.4 units/μL T4 polynucleotide kinase at 37° for 30 minutes. This reaction was ethanol precipitated two consecutive times to remove unincorporated isotope, resuspended in TE and formamide loading buffer and purified on an analytical 8% denaturing polyacrylamide gel. Kinased RNA was eluted from the gel, ethanol precipitated, resuspended in TE (10 mM Tris-HCl, 0.1 mM EDTA). All RNA samples were kept at -20° until further use.

5.2.3.2 Ladder Preparation

0.8 μg 10 basepair DNA ladder was end-labelled using γ-ATP ^{32}P, 50 mM imidazole-HCl pH 6.4, 12 mM $MgCl_2$, 1 mM 2-mercaptoethanol, 70 μM ADP (supplied from 5× exchange reaction buffer) and 0.5 units/μL T4 polynucleotide kinase at 37° for 30 minutes. Sample was precipitated as above and resuspended in TE buffer (10 mM Tris-HCl, 0.1 mM EDTA, pH 8.0) and loading buffer. Appropriate dilutions as necessary for different strengths of radioactivity were made using loading buffer. A hydrolysis ladder of 5′-labelled 7.02 RNA was generated by incubating a trace amount of end labelled RNA with 50 mM $NaHCO_3$ and 19 uM tRNA for 30 minutes at 90°. The sample was neutralized with 200 mM Tris-HCl and diluted in loading buffer. A G-ladder of 7.02 was prepared by incubating a trace amount of end labelled RNA with 50 mM Tris-HCl and 0.2 units/μL T1 RNase for 15 seconds. The incubation was phenol-chloroform extracted two times to remove the enzyme, ethanol precipitated, and resuspended in TE and loading buffer. All ladders were kept at -20° until further use.

5.2.4 Cleavage Assays

5.2.4.1 *bis*-Benzimidazole Cleavage Assays

Erroneous secondary structures were removed from RNA samples by heating at 80° for 1 minute, cooling to 4° transiently and then holding at room temperature until reaction initiation. Aliquots of concentrated stocks of the *bis*-benzimidazole dithienylethene switch **26** (2.4 mM, in DMSO) were converted to either the ring open or closed photostationary state under the appropriate conditions (as determined by UV studies). These aliquots were kept enclosed in foil. Cleavage incubations were 0.35 μM internally labelled RNA, 10-24 μM open or closed switch (except control), 1% DMSO, 50 mM HEPES (pH 6), 50 mM NaCl, 10 mM $MgCl_2$, and 0.05% Triton X-100. Samples were prepared by adding HEPES, NaCl, $MgCl_2$, and Triton (all added as a stock solution) to the tubes containing RNA. All tubes were then moved to a dark room and reactions were initiated by adding either closed switch, open switch, or DMSO (control containing no switch). The cleavage assays were allowed to progress at either room temperature or 37° (in a PCR machine), covered in aluminum foil. Timepoints were removed and quenched with addition of EDTA (to a final concentration of 10 mM) and an equal volume of loading buffer. All timepoints were immediately stored in the dark at -20°. Timepoints were loaded into a 12% analytical denaturing polyacrylamide gel along with a 10 basepair standard (above) and electrophoresed until the full length RNA was approximately 1/3 through.

Radioactive gels were exposed to phosphor screens for two different periods of time. A short exposure allowed the main uncleaved band in each lane to be quantitated. An intermediate signal (an unsaturated band in the 10 bp ladder) was also quantitated, and the amount of background corrected signals of the two related by the fraction (intermediate band / main band). The cleavage products were quantitated by exposing the gels for a longer period of time (generally several days). The cleavage bands to be quantitated (those between the 40 and 100 base markers in the ladder) were arbitrarily chosen. The same ladder band (also unsaturated in this exposure) as in the short exposure was also quantitated. These two background corrected signals were related by the equation (cleavage products / intermediate band). Finally, the amounts of cleavage and uncleaved product in each lane were related by multiplying the two equations to obtain the ratio of cleavage products / uncleaved substrate. This ratio was multiplied by 100 for each sample. For simplicity, this value is referred to as the cleavage ratio.

5.2.4.2 *bis*-Histidine Cleavage Assays

Reaction RNA concentrations were 0.35 μM, composed of either internally labelled or cold (supplemented with a trace amount of end labelled) 7.02. RNA samples (8 μL, all in TE buffer) were prepared by heating at 80° for 1 minute in a PCR machine, cooling to 4° transiently and then holding at room temperature until reaction initiation. All reactions (except the TE control) contained 50 mM

HEPES pH 7, 50 mM NaCl, 10 mM $MgCl_2$, 0.2 mM EDTA, 2.5% DMSO and 0.05% Triton X-100. Dithienylethene samples contained 150 μM of either bis- or mono-switch, in either open or closed photostationary states. Two histidine control reactions were performed, with equal histidine concentrations to the bis (300 μM, deemed 2× histidine) and mono (150 μM, deemed 1× histidine) dithienylethene switches. Fresh stocks of mono and bis and histidine were prepared immediately prior to each assay. Mono and bis histidine switches were each dissolved in 1:1 mixtures of DMSO and ddH_2O. A 3 M stock of histidine was prepared in water immediately prior to the experiment, and the concentration verified by UV spec (ε(211nm) = 5 700 $M^{-1}cm^{-1}$; Fasman, 1976). One control reaction contained just the RNA (prepared as above) supplemented with TE. Everything was in the dark. 1.5× volume stocks of all reaction components (except the RNA) were prepared for each sample. Those samples containing dithienylethenes were not irradiated (to keep the switch open) or switched to the closed form by shining 254 nm light from a hand-held TLC lamp above into the open 600 μL tubes for 6 minutes. The irradiated tubes were intermittently vortexed to ensure the whole solution was being exposed. Reactions were initiated by adding 92 μL of the appropriate 1.5× stock to each tube containing RNA. 25 μL timepoints were stopped by adding to a 2-fold excess of loading buffer (10 mM EDTA, 10 mM Tris base, xylene cyanol, bromophenol blue in formamide) and EDTA to adjust the entire sample to 12 mM. All timepoints were kept at –20° in the dark until run on a gel. Quantitation was performed in the same manner as the *bis*-benzimidazole assay.

5.3 Results and Discussion

5.3.1 *Bis*-Benzimidazole Dithienylethene

5.3.1.1 Synthesis

The synthesis of the *bis*-benzimidazole dithienylethene switch, a new compound, is shown in Figure 42. It consisted of two steps from a dichlorinated dithienylethene (**30**). This compound is a common precursor for the synthesis of modified dithienylethenes, as the chlorines are easily displaced for synthetic variation. This allowed the synthesis to be completed in fewer steps than starting with the parent 2-methylthiophene. The first step used *t*-butyllithium and dimethylformamide to generate the known dialdehyde dithienylethene **28**. The second step was successfully adapted from a procedure to synthesize benzimidazoles from aldehydes and benzene-1,2-diamines (Vanden Eynde *et al.*, 1995). This is a one-pot reaction in which 2,3-dichloro-5,6-dicyano-1,4-benzoquinone (DDQ) is used in both catalytic and stoichiometric capacities. DDQ first catalyzes the formation of an imine that can tautomerize to a cyclized benzimidazole precursor. In the second stage DDQ oxidizes this precursor, forming the double bond of the imidazole portion of benzimidazole. Although this reaction was expected to take two equivalents of DDQ, four equivalents were required for the reaction to go to completion.

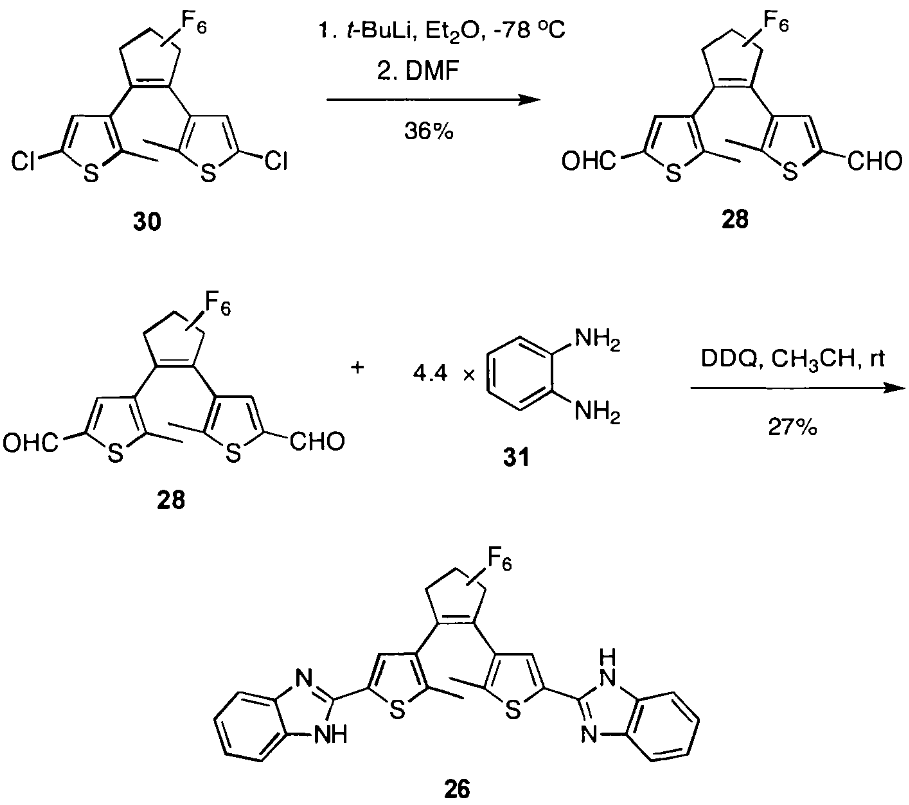

F6
Cl
S
S
Cl
30
1. t-BuLi, Et2O, -78 °C
2. DMF
36%
OHC
CHO
28
+
4.4 ×
NH2
NH2
31
DDQ, CH3CH, rt
27%
N
NH
H
N
N
26

Figure 42: Synthesis of *bis*-benzimidazole switch **26**.

5.3.1.2 UV-Vis Absorption Spectroscopy

The photochromism of *bis*-benzimidazole switch **26** was first performed in organic solvent before proceeding to aqueous buffer for the biochemistry assays. A 2 x 10^{-5} M final concentration was prepared in methanol (Figure 43). The ring-open isomer displayed a λ_{max} of 327 nm. Irradiation of this solution with 313 nm light resulted in a band at 611 nm, corresponding to the production of the closed isomer. The solution visibly changed from colourless to blue. The photostationary state was reached within 12 seconds. A sharp isosbestic point was present in all spectra at 345 nm. Regeneration of the ring open isomer was completed by exposing the solution to light of wavelengths greater than 557 nm. The switch also showed little or no degradation upon two cycles. The conjugated position of the benzimidazoles gives the closed product a dark blue colour.

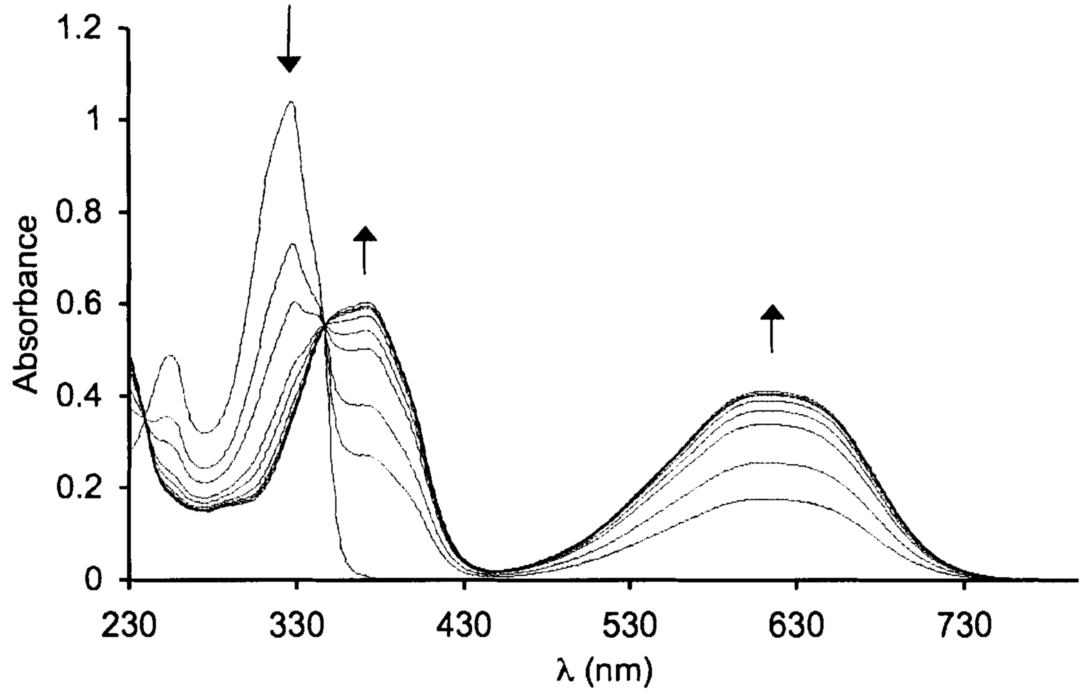


Figure 43: Absorbance spectra changes of *bis*-benzimidazole switch **26o** (2 × 10^{-5} M, in methanol) upon irradiation with 313 nm light.
Irradiation times are every second until 7 seconds, followed by a final time of 12 seconds.

While the spectral properties of the *bis*-benzimidazole switch were very promising in organic solvent, attempts to examine it in aqueous solution failed. The compound was very insoluble in sodium phosphate buffer, giving a jagged spectrum. Often DMSO is used to prepare a concentrated stock of organic compounds that are otherwise insoluble in aqueous solutions. Then small aliquots of this compound are added to the aqueous solution, as long as the final DMSO concentration is between one and five percent. As DMSO is miscible with water, this promotes the solubility of the compound in water. These low concentrations of DMSO are not thought to alter the behaviour of the biological molecules. The *bis*-benzimidazole switch was readily soluble in DMSO, therefore allowing dilutions of this concentrated stock into aqueous solutions. As some of the earlier studies were designed to switch the bis-benzimidazole into either closed or open forms before adding to biological reactions, a UV study was performed in this solvent (as shown in Figure 44). The *bis*-benzimidazole **26** displays different spectral behaviour in DMSO than methanol. The variability in absorption profiles of a compound depending on the nature of the solvent is termed solvatochromism. Irradiation of the ring-open isomer (λ_{max} 330 nm) with 313 nm light resulted in the production of a band at 666 nm. The isosbestic point was red-shifted five nanometers to 350 nm. In addition to the solvatochromism of *bis*-benzimidazole **26**, the kinetics of the ring-closing and ring-opening reactions were slower. The time required to attain the photostationary state was longer (88 seconds) than in methanol (12 seconds).

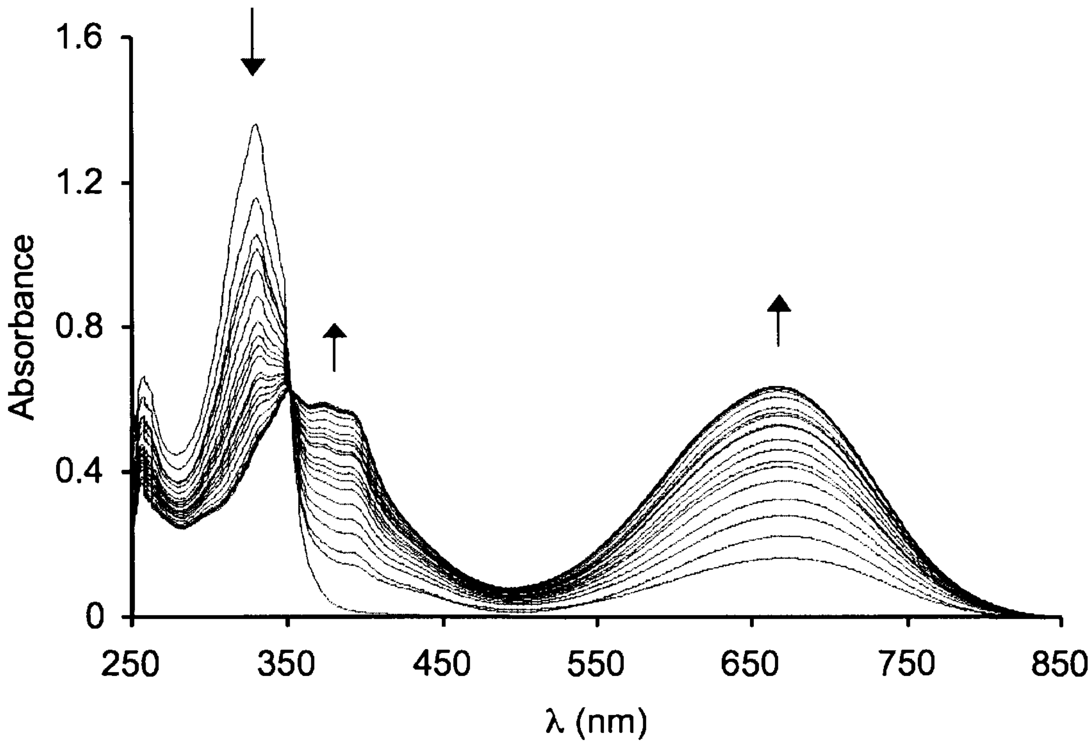


Figure 44: Absorbance spectra changes of a 2 × 10^{-5} M solution of *bis*-benzimidazole switch **26o** in DMSO upon irradiation with 313 nm light.
Irradiation times are every two seconds until 28 seconds, and then every 10 seconds until 88 seconds.

As the *bis*-benzimidazole switch was accessible in DMSO, addition from this stock into aqueous buffer was attempted. Since the buffer for the RNA cleavage assays had not been determined, BPES buffer was chosen. Addition of this DMSO stock into BPES buffer (to give a final concentration of 1% DMSO) did not give the expected absorbance. Additional DMSO was supplemented in individual samples from a total of 2% to 5% with no effect (results not shown). In order to get *bis*-benzimidazole **26** into aqueous solution, the DMSO stock solution of this compound was aliquotted into five tubes (to give final concentrations of 1% DMSO and 1.2 × 10^{-5} M), containing increasing concentrations of the detergent Triton X-100 (final concentrations 0.01-0.05 %). Triton X-100 is commonly used to promote solubility of an organic compound in aqueous solutions, in addition to use of a stock solution in DMSO. As shown in Figure 45, the first concentration of Triton attempted (0.01%) dramatically increased the solubility of the *bis*-benzimidazole switch in BPES buffer. Higher concentrations did not affect the UV-visible absorbance spectrum further. No attempts were made to find the minimal amount of Triton necessary for solubility. Although 0.01% Triton was sufficient, 0.05% was chosen for use in the cleavage assays for two reasons. First, the concentrations of benzimidazole switch were higher than 12 μM. Secondly, 0.05% Triton is standard in biochemical studies containing RNA.

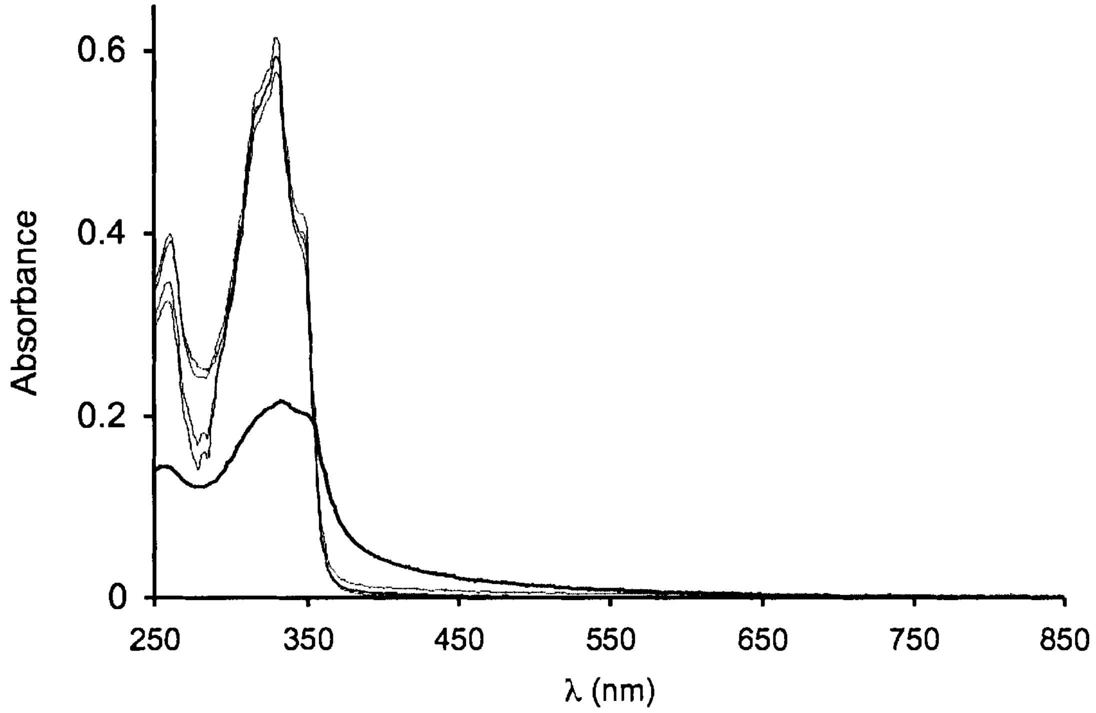


Figure 45: Titration of a 1.2 × 10^{-5} M solution of *bis*-benzimidazole **26o** (1% DMSO, in BPES buffer) with 0.01-0.05% Triton X-100 detergent.
The bold trace is a sample without detergent.

5.3.1.3 RNA Cleavage Assays

To analyze these RNA cleavage assays, we decided to use polyacrylamide gel electrophoresis. This provided an experimental advantage, as previous work with sequence 7.02 featured similar techniques. By using a radiolabelled RNA as the substrate, the cleavage samples could be loaded onto gels and any resulting fragments detected by autoradiography.

There are two different strategies to radiolabel RNA. One way (called internal labelling) is to introduce radiolabelled NTPs during the transcription of RNA from

DNA. The resulting transcript will contain the radiolabelled nucleotide monophosphate as part of the backbone. The molecule can contain more than one rabiolabel, and this abundance depends on two factors. First, it depends on the sequence being generated, and the identity of the hot nucleotide in solution. For example, if radioalabelled UTP is introduced into the transcription reaction, the molecule will contain more hot sites the more uracils there are in the sequence. Secondly, the proportion of radiolabelled versus unlabelled precursor in the reaction will have an impact on the inclusion of radioactive nucleotides. Generating internally labelled 7.02 would take one step from the DNA stock of this sequence from the deletion selection.

The second method is to end label RNA or DNA (at the 5′-end), using the enzyme polynucleotide kinase. This enzyme can either add a radioactive phosphate to the 5′-hydroxyl of the single-stranded sequence (termed the forward reaction), or exchange the unlabelled phosphate on the 5′-end for a labelled one (the exchange reaction). As a result of this, one radiolabel is present at the 5′-end of the molecule.

The end-labelling method provides several benefits. Since all radiolabel is incorporated at the 5′-end, a cut site can be mapped for each cleavage product. Comparing cleavage samples against a hydrolysis ladder (which shows light cuts at every base, due to basic conditions) and a T1 RNase ladder (which cleaves 3′ of G residues) of the same RNA substrate helps identify which residues are cut.

Another advantage of this method relates to the reliability of quantitation. The intensity of each cleavage product bands in a gel is only dependent on its abundance in the sample, since all bands have one label. Therefore, the individual cleavage products can be quantitated. However with internally labelled RNA, cleavage products not only display intensity based on abundance, but also the number of labels in that sequence. Also, the specific activity is much higher for kinasing than for internally labelling, allowing for better signal of fainter cleavage products.

Since 7.02 RNA obtained from transcription contains a 5′-NTP, this is not a suitable kinase substrate for polynucleotide kinase. The RNA must be treated with phosphatase enzyme, which removes the three phosphates from the 5′-end. Lastly, the RNA is subjected to the forward kinase reaction. In order to investigate whether there was any overall effect of incubating 7.02 RNA with open or closed switch, all benzimidazole assays and the first histidine experiments were performed with internally labelled RNA. This labelling technique generates the RNA substrate in just one step as opposed to end-labeling.

The conditions for the benzimidazole cleavage assays were designed to contain similar buffer and salt conditions to the selection for the nucleotide synthase ribozyme selections, as the substrate was an RNA isolate from this selection. Since it was in our interest to obtain cleavage information in context of the proposed secondary structure of the RNA, it was imperative it adapt this fold. However it was necessary to include conditions (for example magnesium concentrations) conducive to cleavage by a small molecule. So the switch had to be soluble (including DMSO and Triton detergent) and also feature a pH that would be at the pKa of the benzimidazole switch. The buffer conditions chosen for these benzimidazole RNA cleavage assays, in addition to those of the selection, are shown in Table 2.

Table 2: Comparison of nucleotide synthase ribozyme selection and *bis*-benzimidazole switch **26** cleavage assay concentrations

Selection conditions	**Cleavage conditions**
50 mM Tris-HCl pH 7.6	50 mM HEPES pH 6.0
150 mM KCl	4 M NaCl
25 mM $MgCl_2$	1 M $MgCl_2$
-	1% DMSO
-	0.05% Triton X-100
~1 μM RNA	0.67 μM RNA

Several preliminary assays were performed in which conditions such as stoichiometry, gel exposure times, and sample loading were explored (results not shown). After this a more detailed assay was performed under improved conditions. Three separate incubations were used to assess cleavage behaviour of the open and closed switches, and no switch (DMSO added in its place). The stoichiometry of switch to RNA was 24 μM : 0.67 μM. The switch stock solution in DMSO was aliquoted into two tubes, each switched to either open or closed. The incubations were done in the dark. The experiment was performed at biological temperature (37°), as this is the temperature that many RNase A mimic assays are done at. Six timepoints were removed over the course of the 25 hour incubation. The results of these assays are shown in Figure 46. All three samples show a steady increase in cleavage bands over the course of the incubation. Qualitatively, there does not appear to be an effect by adding closed or open benzimidazole switch, as compared to the control. The cleavage bands all appear to be similar, with a similar pattern obtained. An artifact approximately 100 bases long is present in all samples from purification. Since the purification of the RNA is done on a lower percentage polyacrylamide gel than the assay gel, surrounding bands are evident in the analytic gel. Ideally, we were hoping for higher rates of cleavage for one isomer compared to the other. Also, it would have been interesting to see a preference of certain cleavage sites over others.

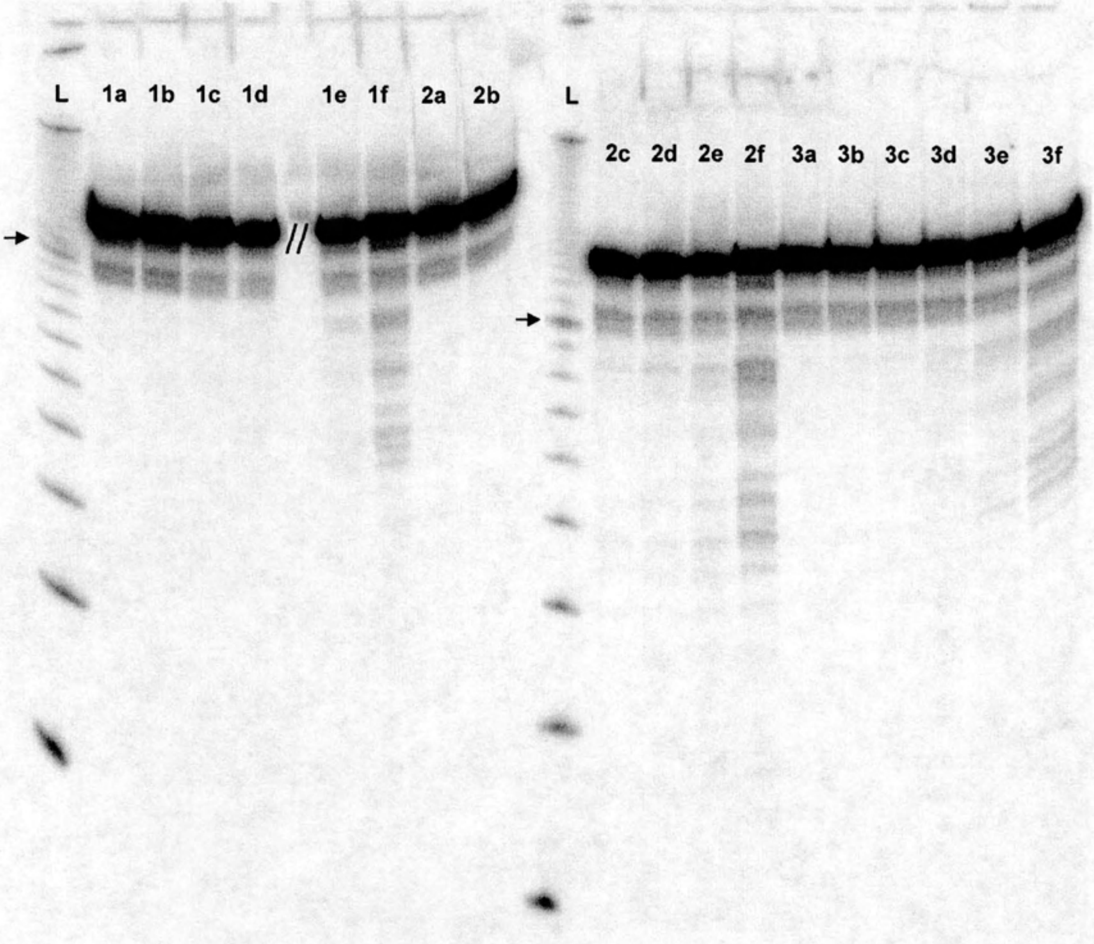


Figure 46: RNA cleavage assay with *bis*-benzimidazole switch **26**.
Samples are closed switch (1), open switch (2), and no switch (3). Lanes a-f refer to timepoints 0, 2, and 30 minutes, and 2, 6, and 25 hours, respectively. L refers to 10 bp ladder with arrows indicating the 100 bp band. The dashed lines (//) indicate discontinuity in the left gel. Quantitated regions are between 40 and 100 bp markers.

Quantitation of these gels was not straightforward, due to the low proportion of signal of cleavage products compared to the main band. Ideally, cleavage fragments are quantitated, and expessed as a percentage of this plus uncleaved substrate. In order to properly quantitate the cleavage, the resulting gels were exposed for a long period (one week) to increase the signal of cleavage products. However in doing so, the uncleaved band signal was saturated.

Therefore, quantitation of the main bands was achieved by obtaining a shorter exposure. Quantitation of an intermediate intensity band, which was not saturated in the long exposure, was used to relate the two different values. These results are shown in Table 3. Unfortunately, there is no significant difference in the ratios of cleavage products/uncleaved substrate between the three samples. Shortly after the completion of this assay we decided to proceed instead with the histidine switch, due to the literature precedent of histidines in RNase A mimics.

Incubation time	Cleavage ratio (Open switch)	Cleavage ratio (Closed switch)	Cleavage ratio (No switch)
0 min	0.062	0.148	0.10
2 min	0.092	0.079	0.11
30 min	0.132	0.17	0.16
2 hours	0.150	0.26	0.22
6 hours	0.254	0.5	0.96
25 hours	0.826	1.34	1.22

Table 3: Quantitation of RNA cleavage assay featuring *bis*-benzimdazole switch **26**. Cleavage values are expressed as product/main band, multiplied by 100.

5.3.2 *Bis*-Histidine Dithienylethenes

5.3.2.1 Synthesis

This procedure was successfully adapted from Yuan *et al.* (2002), which described the synthesis of an amide bond between histidine and an acid chloride substituent. The diacid switch **32** served as the starting material for the preparation of the diacid chloride **33**. By adapting the conditions in Yuan *et al.*, the coupling of the switch to two histidines was achieved (Figure 47), producing **27**. This procedure also produced a mono-substituted histidine switch **29**, which hydrolyzed during workup on the unsubstituted side to the corresponding carboxylic acid. This compound would serve as an important control during the cleavage assays, probing whether the location of a second histidine (as in the *bis* switch) in proximity to one provides an added benefit. Purification using silica gel chromatography was very difficult due to the polar histidine groups on these molecules. A solvent system using ammonium hydroxide was successfully used after reading literature preparations for other histidine-derived RNase A mimics.

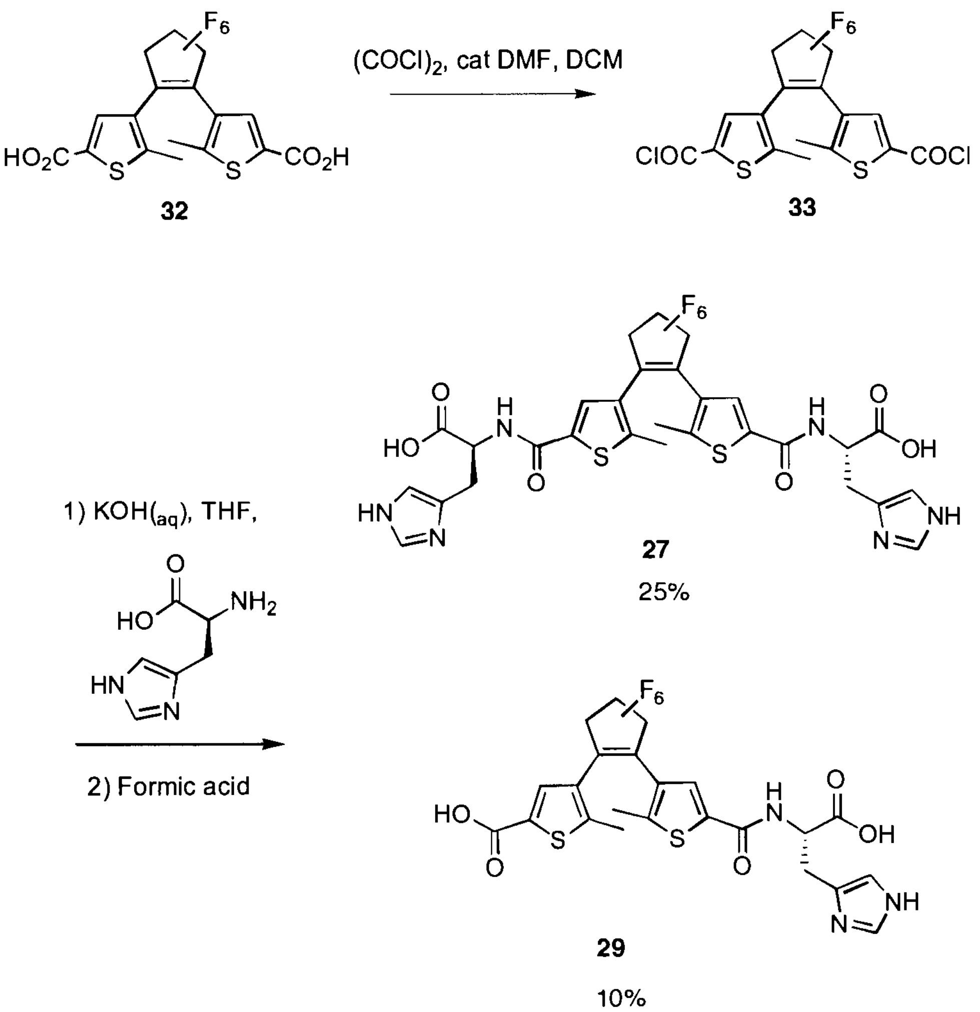

F6
HO2C
CO2H
32
(COCl)2, cat DMF, DCM
ClOC
COCl
33
1) KOH(aq), THF,
HO
NH2
HN
N
2) Formic acid
O
H
N
OH
NH
27
25%
29
10%

Figure 47: Synthesis of *bis*-histidine and mono-histidine switches **27** and **29**, respectively (as performed by David Sud).

5.3.2.2 UV-Vis Absorption Spectroscopy

The UV studies were performed for each switch in water, at final concentrations of 1.75 × 10^{-5} M (17.5 uM). These results are shown in Figure 48. The colourless open form of the mono switch **29** displayed a λ_{max} of 256 nm. Conversion to the ring-closed isomer was achieved by irradiating the solution with 254 nm light. This generated a band at 573 nm and a smaller band at 367 nm, producing a light violet colour. Extinction coefficients were calculated to be ε(256 nm) = 14 500 $M^{-1}cm^{-1}$ for the open isomer, and ε(573 nm) = 3 800 $M^{-1}cm^{-1}$ for the closed molecule. The spectrum of the open *bis*-histidine switch **27** looked quite similar, with a λ_{max} of 260. 254 nm light was also used to induce the ring-closing reaction, which produced bands at 367 and 568 nm and changing to the colourless solution to light violet. The *bis* switch had extinction coefficients of ε(260 nm) = 17 000 $M^{-1}cm^{-1}$ (open) and ε(568 nm) = 4 700 $M^{-1}cm^{-1}$ (closed). Both switches demonstrated isosbestic points, indicating clean conversion from the open to closed isomers.

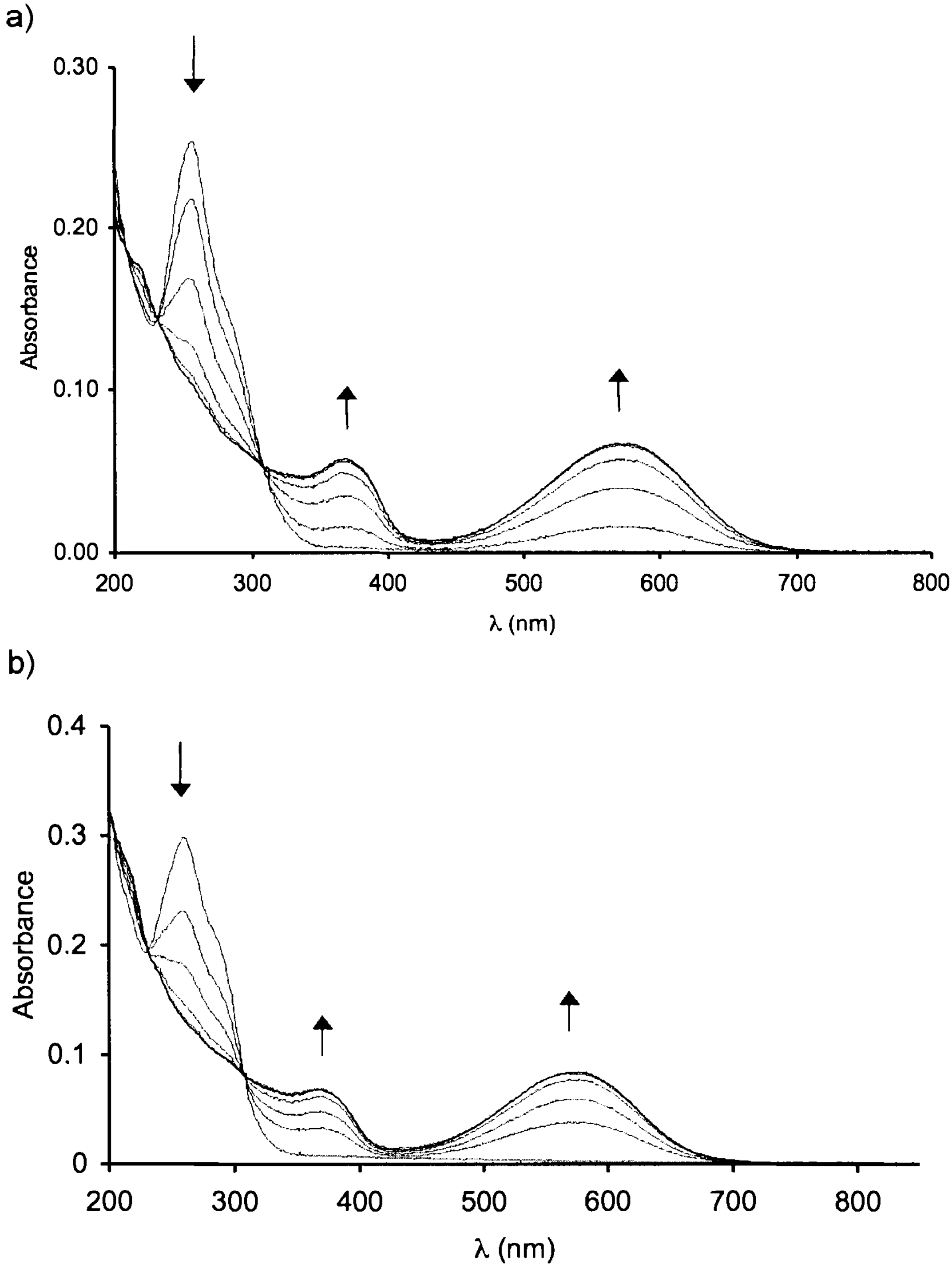


Figure 48: UV-Visible Absorbance spectra changes of ring-open histidine-functionalized switches (1.75×10^{-5} M in H_2O) after irradiation with 254 nm UV light.
a) Mono-histidine **29**; irradiation times are 2, 8, 20, 40, 70, and 115 seconds. b) *Bis*-histidine **27**; irradiation times are 4, 8, 16, 32 seconds, and 1 and 2 minutes.

In order to optimize any cleavage observed, the histidine switch cleavage assays were performed with a higher stoichiometry than the benzimidazole studies (150 μM switch, as opposed to 24 μM). However, the preparation of a stock concentrated enough to add switch to the assays was problematic. The *bis*-histidine switch and mono-switch were found to be sparingly soluble in water and organic solvents during the purification process. Perhaps the combination of both hydrophobic and hydrophilic regions in the molecules causes the switches to self-associate. The best solubility of concentrated solutions (above 100 uM) was in a 1:1 mixture of DMSO and water. While the mono-histidine **29** was soluble at a 3 mM concentration in this mixture, the *bis*-histidine was not. Any other solvent would be undesirable for biochemical assays, as they would be present in the final sample. The concentration of each switch to be used for the cleavage assays was 150 uM. Due to high absorbance, the solubility could not be directly observed at this concentration. Triton detergent at 0.05% concentration was used in the cleavage assays to promote solubility. As this was sufficient for benzimidazole switch solubility (which is very hydrophobic and displayed no solubility in water alone), it was assumed it would be appropriate for the histidine switches.

5.3.2.3 RNA Cleavage Assays

As with *bis*-benzimidazole switch **26**, the first assays used internally labelled substrate to get an overall idea of whether or not there was any effect. However there were several changes. To accommodate the higher pKa of histidine, pH 7

HEPES buffer was used. To reduce the amount of background cleavage, the assays were done at room temperature, and EDTA (at 0.2 mM concentation) was included in the assay buffer. Although the mono switch was fully soluble in the DMSO/water stock, solubility of the *bis*-histidine was poor. Therefore, addition directly from this stock was not possible. Instead, an aliquot was removed from this stock (which was more of a suspension), and added to a stock of all other reaction components, except RNA. The dilution and detergent allowed solubility of the *bis* -switch. The mono switch was treated in the same way.

More controls were performed than the benzimidazole assays. As with the benzimidazole assays, a control was designed to assess cleavage ability of the assay buffer. A second control was performed in which the RNA was incubated in TE buffer alone. TE does not contain magnesium and nucleic acids are stable under these conditions. This shows how labile the substrate RNA is, without the assay buffer. As well, cleavage ability of the amino acid histidine was investigated. This was done to measure the effects of tethering histidine to a dithienylethene scaffold. Two concentrations were used: the first, called 2× histidine, was the same concentration as the histidines in the bis switch (300 μM); the other (150 μM) was equal to the mono switch and called 1× histidine.

Two successive internally labelled assays were performed under identical conditions. The gel results of the second assay are shown in Figure 49. The fragments that result are the same as those in the benzimidazole experiments.

Looking at the gels, it is apparent the amount of cleavage is lower than the final benzimidazole assay. This is most likely a result of the decrease in temperature from 37° to room temperature. As visible in these gels, the fragment bands in all experiments are identical to those in the cleavage control. This indicates that the switches (if they bind) do not cleave in new locations. Also, the relative abundances of the bands look the same, so they do not prefer different sites.

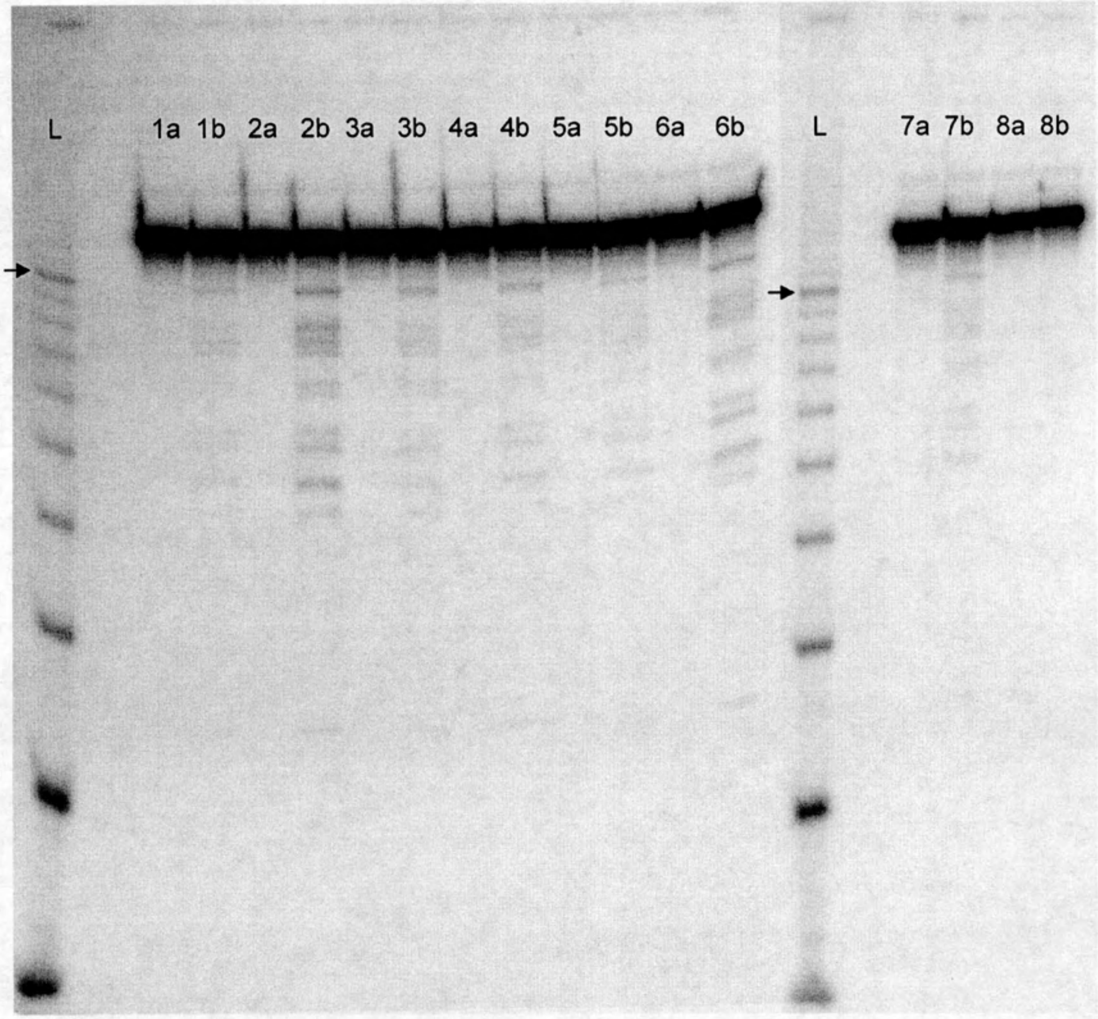


Figure 49: Gel results of the second RNA cleavage assay featuring histidine switches **27** and **29** using internally labelled substrate.
Numbered lanes 1-8 refer to the following: *bis*-histidine switch (open and closed), mono-histidine switch (open and closed), histidine control (2× and 1×), assay buffer control and TE buffer control, respectively. In each case a refers to the zero timepoint and b to the 23.5 hour point. L is radiolabelled 10 bp ladder, and arrows indicate the 100 bp band. Quantitated cleavage products were between 40 and 100 bases.

Since the amounts of cleavage were low relative to the uncleaved substrate, these gels were quantitated using two different exposure times (similarly to the benzimidazole switch quantitation). During the RNA substrate purification, care was taken to not include other bands when excising the band out of the gel.

Although the fragment at 100 bases was not included, the region below the full-length sequence is a bit smeared, indicating this purification was not perfect. However the first assay was not as stringently excised and contains some of this fragment. As this fragment is also a cleavage product, it was included in both quantitations. Therefore, the quantitation results between the two assays cannot be directly compared, since the compositions of the substrates slightly differ. There is a general trend in both of the experiments (shown in Table 4). Within each experiment, the *bis* switch (in each form) provides more breakdown than both forms of the mono switch. The TE buffer control experiences the lowest cleavage, with the assay buffer providing higher levels. The results of the histidine assays are variable. The second assay displays a very low 2× histidine value, possibly the result of an experimental error.

Table 4: Quantitated results of two internal cleavage assays using histidine switch

Sample	**Cleavage ratio (Assay 1)**	**Cleavage ratio (Assay 2)**
bis-switch (open)	14.7	10.3
bis-switch (closed)	16.9	14.6
mono-switch (open)	12.0	6.5
mono-switch (closed)	8.9	5.1
2× histidine	20.8	2.9
1× histidine	6.8	6.3
assay buffer control	0.7	3.9

Sample	Cleavage ratio (Assay 1)	Cleavage ratio (Assay 2)
TE buffer control	0.03	1.6

After these initial assays indicated differing amounts of cleavage, end-labelled substrate was used to obtain more information. Rather than use switch in the concentrated stock, new solutions of each switch were made before each assay. However due to the low solubility of the bis stock in DMSO/water, a new approach was used for this. The switch was dissolved in the DMSO/water stock and then all other reagents for the reaction (scaled up) were added to this stock and the concentration checked. The combination of the dilution and detergents solubilized it, allowing the solution to be soluble. Although the extinction coefficient had been measured in water, the final dilution used for concentration measurements in the UV was mainly water so it did not matter.

A control was performed to identify breakdown of end-labelled 7.02 in TE buffer. It showed very little breakdown of the RNA over the 23 hour period of the assay. In the triplicate assay, the samples were background corrected for the initial breakdown present in the zero timepoint. As TE contains EDTA to sequester any metal ions, the RNA is relatively stable. Although this RNA was selected to fold into a secondary structure with metal ions, it appears that the lack of these conditions does not drastically affect the structure. Evidence for this is the similar breakdown pattern between the TE control and the cleavage buffer samples. If the secondary structure varied between the two conditions, then the cleavage sites would differ. These contain conditions very similar to the selection conditions. It is not surprising that the cleavage buffer samples display higher rates of cleavage than the TE samples. This is probably because of the magnesium and relatively low concentration of EDTA.

As in the internal label assay, no new bands show up as a result of switch or histidine catalysis compared to control breakdown. Also, the relative abundances of the various cleavage bands in each sample are all the same (Table 5). It was possible that cleavage catalyzed by the molecules would favour certain cleavage sites, especially in the case of the switches where sterics or binding location preferences may be present. As well no difference was seen between open and closed molecules, meaning that the steric changes between open and closed do not translate into a significant difference in cleavage location.

There were several drawbacks to this experiment. One was the low solubility of *bis*-switch **27** in the DMSO/water stock. Since the maximum concentration of switch was already achieved, the only way to maximize the stoichiometry was to minimize the RNA. This decreased the signal. One drawback to the experiment was the requirement to "quench" the reaction by adding a two-fold excess of loading buffer. This was thought to stop any cleavage by the addition of formamide in the buffer, which would denature the reaction. This diluted the already faint signal even further.

Table 5: Quantitated results of end-labelled cleavage assay #1, performed in triplicate. Values given are ratio of cleavage fragments (between 40 and 100 bases)/amount of uncleaved RNA in the 23 hour timepoint, multiplied by 100. To control for the amount of breakdown in RNA substrate before beginning the reaction, the amount in the zero timepoint was subtracted from the 23 hour timepoint.

Sample	**Cleavage ratio (Replicate 1)**	**Cleavage ratio (Replicate 2)**	**Cleavage ratio (Replicate 3)**
bis-switch (open)	1.30	1.69	1.83
bis-switch (closed)	63.53	2.60	2.94
mono-switch (open)	1.85	0.55	0.48
mono-switch (closed)	11.58	2.96	6.20
2× histidine control	0.86	0.56	9.28
1× histidine control	4.33	1.20	7.03
assay buffer control	4.45	2.14	2.10
TE buffer control	0.26	0.48	0.27

The large amount of variation in this assay made us consider whether the cause was something experimental. In order to examine this, the experiment was repeated, minimizing variation. The additional histidine controls (1× and 2×) were eliminated until error was reduced. To minimize pipetting error, a given switch was dissolved in the reaction buffer components as a master stock. Therefore the switch and all other components (except RNA) for all six reactions were combined (and were scaled up to contain more than needed). This avoided the individual pipetting of components to each individual tube. Aliquoting from this stock would ensure concentrations were all the same. One possibility for the variation between the closed samples might be have been due to differing successes of the ring closing reaction, as they had been conducted in separate tubes. To prevent this, half of the stock was removed and exposed to UV light. This was vortexed thoroughly and quickly aliquoted out (in the dark). The unexposed stock (which has been shown to be in the open form by UV) was also vortexed and aliquoted out, ensuring equivalent sample compositions.

The variation also could have been related to preparation of the RNA substrate. Previously, all RNA samples had been aliquoted out into the reaction tubes, heated and cooled. There could have been differences in temperature between the tubes, as well the vortexing of the RNA before beginning the reaction could have been different. Rather than proceeding this way, one tube containing the RNA pooled together was heated and cooled. This sample contained more RNA than necessary to ensure the RNA withdrawn for the experiments would not be near the bottom of the tube (where there might be damaged fragments). The RNA used was slightly older than that used for the previous assay. Although the RNA had been stored at –20°, minor breakdown was apparent in the zero timepoints. Very faint bands were visible corresponding to this degradation. Another possibility is that the handling of the RNA, and not the age was responsible for the breakdown. The additional step of pipetting the RNA out after the heating/cooling process might have been a factor. To account for potential differences in RNA composition for each sample (at the beginning of the assay), the quantitation procedure involved subtraction of the ratio of breakdown already present in the initial sample (zero timepoint). As the storage conditions do not contain the buffers and salts necessary for adaptation of the secondary structure, the RNA is susceptible to breakdown at any position.

The results of this triplicate assay are shown in Figure 50. There were two new problems that not been encountered in previous assays. In one of the open *bis*-histidine samples, both timepoints had portions of the sample that failed to migrate out of the wells. This can indicate some sort of precipitate has formed. This is odd, given the common stock from which the RNA came. Another problem was the presence of T1 RNase contamination in two of the mono open samples. One of these is included in the Figure 50. In addition to the regular degradation pattern, several other bands are shown to appear only in the one-day timepoint. The contamination must have been introduced after withdrawing the zero timepoint from the reaction, as the zero timepoint is free of T1 RNase contamination. T1 RNase is a very fast acting enzyme; the T1 ladder was generated by only 15 seconds of exposure to this enzyme. The contamination could have been introduced before withdrawl of this aliquot, as long as it was very minor. The contamination also could have occurred after loading the gel, by cross-contamination in the gel running buffer. However, the location of the contaminated sample several lanes away from the ladder sample makes it less likely.

As in the previous end-labelled assay, the samples incubated in TE buffer show the lowest amounts of breakdown over the one-day time period. Since the cleavage products in the zero timepoint were subtracted from the final timepoint, some samples showed no significant change. In fact, two of the three samples shown a negative change, probably from variation in the amounts of background degradation already present in the RNA. The quantitated values from this experiment are summarized in Table 6. The assay buffer control shows higher amounts of cleavage than TE buffer, consistent with all other assay results. This indicates that the buffer and salt conditions present in this buffer alone are capable of promoting the degradation of the RNA. It is interesting to note that the degradation pattern of the RNA in all zero timepoints appears as a ladder, with cleavage at many different positions. In all samples other than the TE control (where no significant cleavage occurs after one day), only specific bands increase over time. This may indicate that the secondary structure of the RNA has formed, protecting certain regions from cleavage. The TE control does not contain the high salt concentrations that were present during the selection (Table 2), which may prevent the structure from being formed. However since the cleavage is so minor, no conclusions can be made about any cleavage patterns. The assay buffer control shows the same specific bands appearing over the course of the incubation as the previous end-labelled assay (not shown). Two ratio values are approximately similar (1.99 and 1.71) while the third value is much higher (17.22). With so much variation in this control for background levels of cleavage, it is hard to make assumptions about the switches beyond a qualitative assessment.

The two *bis*-histidine open samples (that migrated properly) show similar cleavage ratios in the gel. These values (4.09 and 3.80) are similar to the lowest values for assay buffer. An encouraging aspect was the larger amount of cleavage in the closed *bis*-histidine samples than the open. While two of the triplicate samples are very close (30.42 and 24.91), the third displays a much higher ratio (148.11). This preliminarily indicates the open *bis*-switch displays background levels of cleavage, but switching it into the closed isomer produces higher amounts of cleavage. However the variation between assay control samples, as well as closed *bis*-histidine samples, makes it difficult to say this with much certainty. The presence of T1 contamination in two of the three open mono-histidine samples only allowed quantitation of one sample. This value of 2.79 puts it in the range of the open *bis*-histidine switch and two of the three assay buffer samples. However the loss of the two other replicates prevents any information about variation within this sample. Another successful aspect to this experiment was the low ratio of cleavage for the closed mono-histidine samples (0.95, 0.81 and 2.13). The substitution of a carboxylic acid for a histidine (as in the mono-switch) removes the general acid/base capabilities of one end of the molecule. Therefore, changing the geometry of the molecule should not affect the activity of the molecule. These results seem to support this hypothesis.

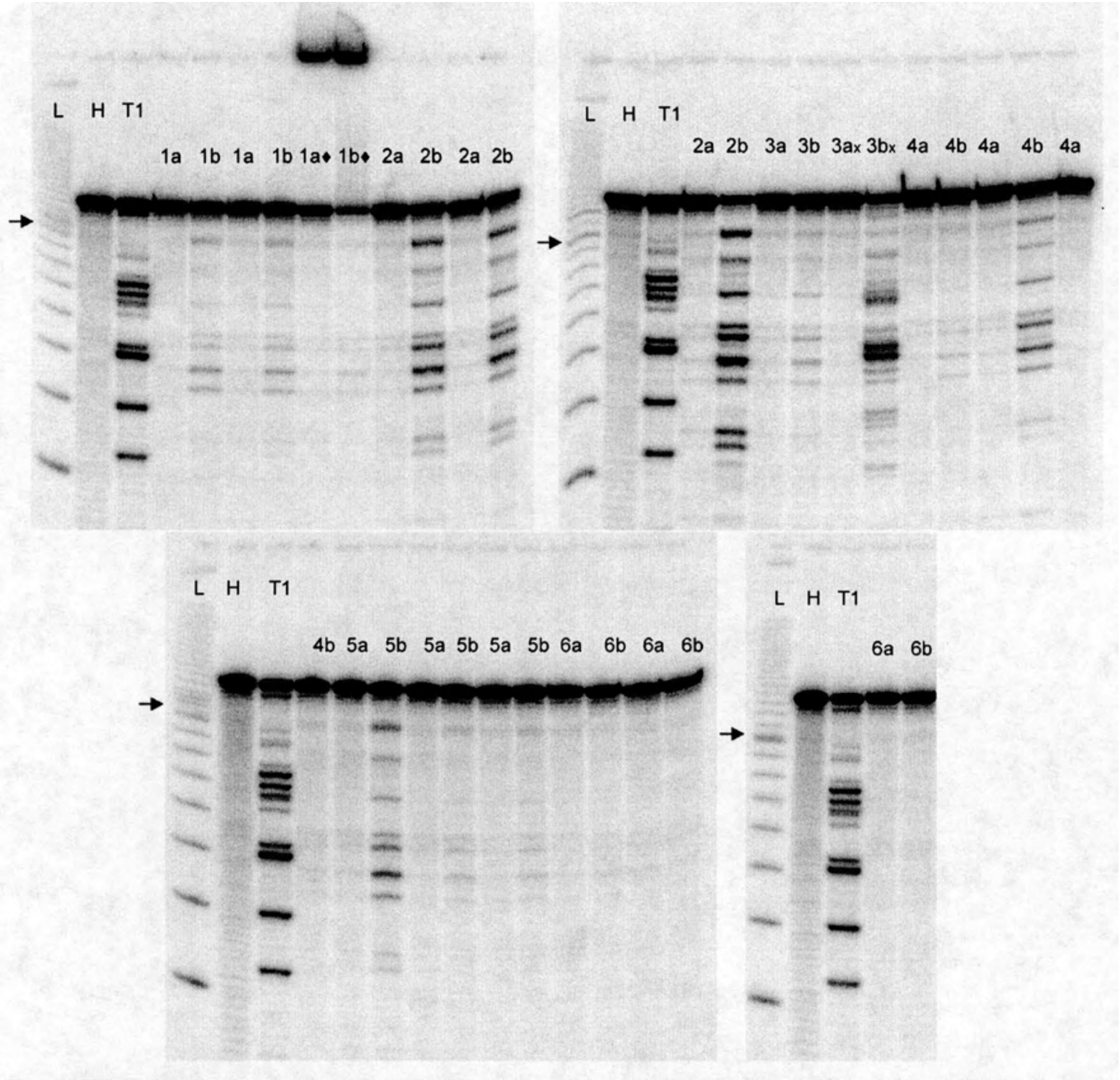

L H T1
1a 1b 1a 1b 1a♦ 1b♦ 2a 2b 2a 2b
L H T1
2a 2b 3a 3b 3ax 3bx 4a 4b 4a 4b 4a
L H T1
4b 5a 5b 5a 5b 5a 5b 6a 6b 6a 6b
L H T1
6a 6b

Figure 50: Triplicate end-labelled cleavage assay featuring histidine switches **27** and **29**. Numbered lanes 1-6 refer to the following: *bis*-histidine switch (open and closed), mono-histidine switch (open and closed), assay buffer control and TE buffer control, respectively. In each case a refers to the zero timepoint and b to the 23 hour point. L is radiolabelled 10 bp ladder, (100 bp band indicated with arrow), H is a hydrolysis ladder of 7.02, and T1 is a T1 RNase cleavage ladder of 7.02. Improper migration out of the wells is observed in one open *bis*-switch replicate marked with ♦. The cleavage pattern in samples marked with x appears to be due to T1 RNase contamination.

Table 6: Quantitated results of the triplicate cleavage assay #2.
♦ indicates sample did not migrate into wells properly, preventing proper quantitation of results. × indicates samples contained T1 contamination (as determined by polyacrylamide gels) and quantitation was not performed. Quantitation was identical to the method presented in Table 5.

Sample	**Cleavage ratio (Replicate 1)**	**Cleavage ratio (Replicate 2)**	**Cleavage ratio (Replicate 3)**
bis-switch (open)	4.09	3.80	no data♦
bis-switch (closed)	30.42	24.91	148.11
mono-switch (open)	no data×	2.79	no data×
mono-switch (closed)	0.95	0.81	2.13
assay buffer control	17.22	1.99	1.71
TE buffer control	-.007	0.53	-0.10

5.4 Concluding Remarks

We have synthesized and investigated the RNA cleavage abilities of imidazole-functionalized dithienylethene switches. The aim of this project was to identify a switch capable of degrading RNA while in one isomer, but not the other. This would allow control over *in vitro* RNA cleavage by using the external stimuli of light, not possible in current RNase A mimics. Such a finding would be useful in the degradation of mRNAs sequences, preventing the translation of their unwanted protein products. Another role could be degradation of viral RNA. Potentially, development into therapeutic applications such as photodynamic therapy may allow the treatment of particular diseases *in vivo*.

Preliminary experiments with the *bis*-benzimidazole dithienylethene **26** did not yield positive results. Not only was there no difference in the amounts of cleavage between open and closed switches, the values were identical to the level of background cleavage. While bis-benzimidazole compounds are capable of strongly binding DNA, they have not been implicated in general acid-base catalysis such as histidine. Therefore, it was not surprising these experiments did not prove successful.

The cleavage assays with the histidine switches **27** and **29** are difficult to interpret, due to the large amount of variability. This variability was present not only in the switch-containing triplicate samples, but also the control triplicate samples. Although it seems that the closed form of the *bis*-histidine switch **27** promotes cleavage more than the open form, the variability in this experiment means this may not be significant. Theoretically, the two histidines in the bis-histidine switch could act as a general acid and base respectively to cleave the phosphodiester bond of the substrate, analogous to RNase A. The higher ratios of cleavage of the closed isomer of the *bis*-histidine switch relative to the open may indicate that the positions of the two histidines are more favourable than the open isomer. In the case of the mono histidine switch (**29**), the low pKa (approximately 3) of carboxylic acids means that a general acid/base role of this substituent is not possible under these conditions. The neutral pH of the incubation conditions means the carboxylic acid substituent would be predominantly deprotonated. Therefore, the position of the two ends of the molecule should not affect cleavage. The configuration of the mono-histidine does not seem to have an impact on the amounts of cleavage, supporting this hypothesis. These results support the theory that cleavage of the RNA substrate by the *bis*-switch is enhanced by the positioning of two histidines (in a particular geometry) near the RNA backbone.

The variation within this experiment may be due to the RNA substrate. The heating and cooling procedure may not produce homogenous RNA samples. Also, the adoption of this complicated fold, and stabilization of it through the incubation, may vary from sample to sample. Another possibility is that the RNA substrate is not a good one, as its stability may prevent cleavage. Also, the compact structure of the 7.02 RNA may not allow accessibility for these molecules. Single tRNA sequences have been successfully used in RNase A mimic studies (Podyminogen *et al.*, 1993; Helm *et al.*, 2001). In the future it may be useful to use such a substrate.

Unlike RNAse A mimics that contain elements for RNA recognition, the switches featured in this thesis likely could only bind by hydrogen bonding to the RNA substrate. It is very possible there is not enough recognition for for histidine switches **27** and **29** to bind the substrate. Also, the negatively charged carboxylic acids on both switches may hinder cleavage, as it would require interaction with the similarly charged RNA backbone. Any future strategies to photocontrol RNA cleavage with a dithienylethene switch would probably be more successful if it included an element for nucleic acid recognition. One strategy would be to include an intercalating molecule somewhere on the dithienylethene scaffold. Once docked to the RNA substrate by intercalation, the histidine portions of the DTE could cleave the substrate at a nearby location. Another scenario is the covalent linkage of a dithienylethene onto an oligonucleotide sequence, allowing it to recognize the RNA due to basepairing. Sequence specificity of the substrate RNA could be achieved by designing the complementary oligonucleotide.

6 FUTURE DIRECTIONS

The dithienylethene molecular switch scaffold undergoes large changes in structure than can be regulated by light stimuli. This thesis describes the use of dithienylethenes in biological roles such as DNA binding, bacterial staining, and RNA cleavage. The ultimate goal of these studies was to identify switches capable of a given function in one isomer, but not the other. While these projects were not completely successful at regulating these functions, they provide a context for future studies.

We identified a dithienylethene switch capable of binding DNA in each isomer. The DNA binding is not as strong as some known DNA binding ligands, and controllable binding is not possible using these compounds. However, these studies have shown that a dithienylethene switch is capable of binding DNA. The development of tighter DNA binding dithienylethene switches in the future may be more successful if the substituents are known groove binders or intercalators. In addition, these groups could be located in other positions on the dithienylethene scaffold such as the methyl thiophene substituents. In the open isomer, these substituents are in the plane of the thiophene rings, while in the closed isomer, these groups are perpendicular to the fused rings. This may provide a more dramatic change in orientation than the switches described in this thesis. Differential binding between isomers may be possible by synthetically modifying these positions

A dithienylethene switch capable of lightly staining bacteria was also described in this thesis. The low intensity of the coloured isomer made it difficult to assess the reversibility of this staining through conversion to the colourless open isomer. This application requires a switch with high optical densities to compete with conventional dyes. To exploit changes in fluorescence and colour, it would be best that the fluorescence excitation and emission wavelengths be separate from those that trigger the photoreactions. That way, the fluorescence and absorbance properties would be independently addressable. It may be best in the future to use another method to attach the switches to bacteria or organelles, rather than relying on binding abilities of the switch.

Further development of an RNase A mimic would ideally feature a switch with some of the properties offered by the compounds used in this thesis. The geometry of the *bis*-benzimidazole switch provides the desired "on" and "off" feature. The closed molecule prevents the two imidazoles from approaching each other close enough to act as a general acid and base simultaneously in the RNA cleavage mechanism. However, the benzimidazoles are not a well known group for RNA cleavage chemistry. Placement of an unmodified imidazole closer to the dithienylethene scaffold than the attachment of the *bis*-histidine switch would allow the geometry to be turned on or off, and would incorporate a functional group known to cleave RNA. Further similarity to RNase A could be achieved by including a lysine side chain or positively charged group on the scaffold. This would help solubility, stabilization of the pentavalent phosphate intermediate, and association with the negatively charged RNA backbone. An additional strategy would be to incorporate a substrate recognition motif, such as one capable of base pairing or intercalation, into the scaffold. This would achieve sequence specificity, which would be important in medical applications such as the targeting of undesirable RNA.

APPENDIX

The DNA sequences of clones isolated from rounds 4,5, 6, and 7 are shown in the following figure. The sequences were aligned with respect to the parent sequence 163.1, which contained no deletions. Sequences were aligned using MegAlign software. Names assigned to isolates sequenced from each round are in the form da.x.y, where da refers to deletion selection (of family A ribzoymes), x is the round, and y is the clone number. The seven constant regions (as summarized in Figure 3) are comprised of stems (each arm denoted as I-VI), including two tetraloops closing stems V and VI. Stems are colour coded with according to the a.6.10 and da.7.02 isolate secondary structures in Figure 4 and Figure 6, respectively. The five variable regions R1-R5 show amounts of deletions over successive selection rounds (refer to Figure 7). The 3′ end of R1 appears to contain a conserved motif in many round seven isolates. Round seven also reveals interactions between the 3′ end of R3 and the 5′ end of R5. R3 and R5 appear to favour particular sequences, while R2 and R4 show no apparent sequence preference.

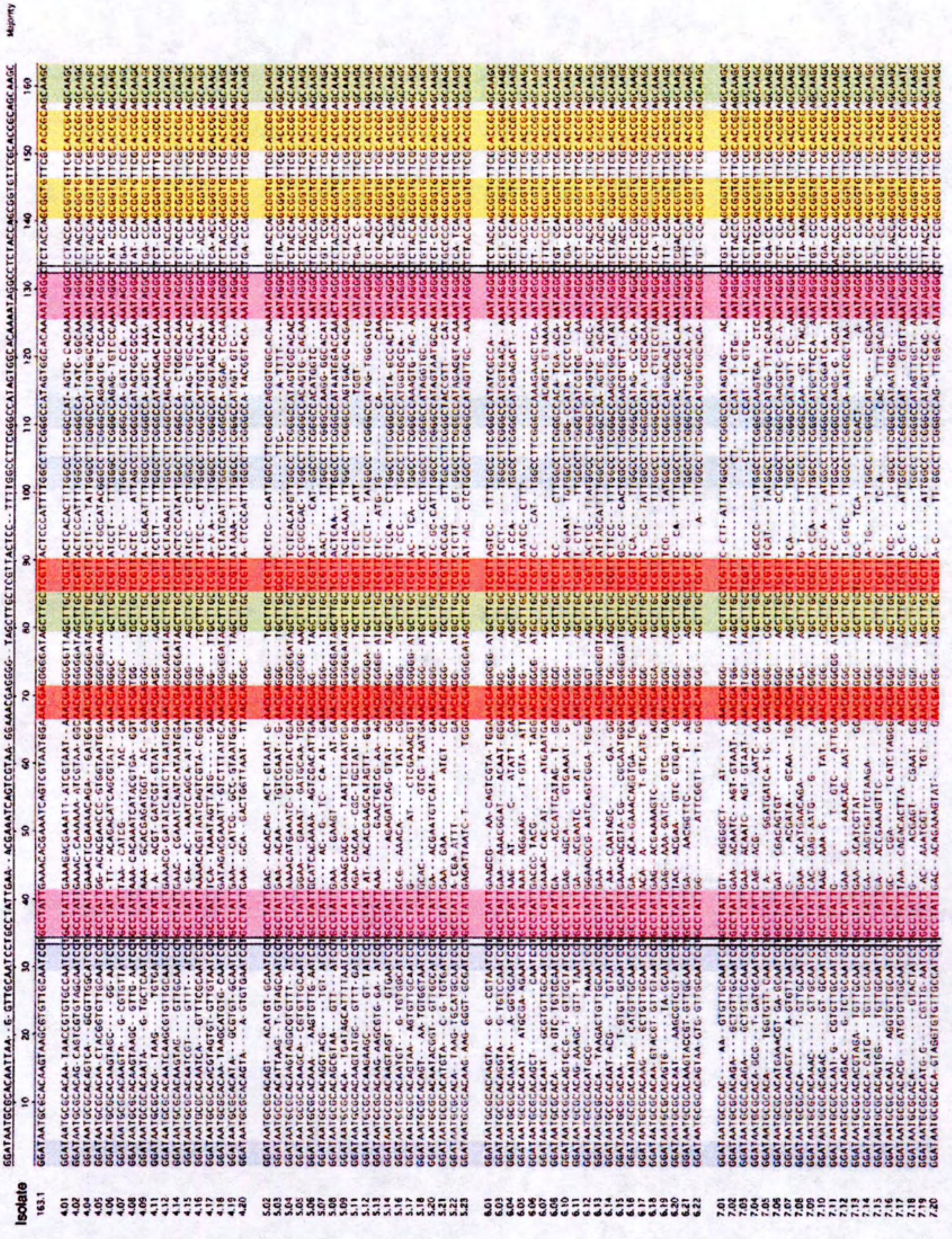


Figure 51: Aligned sequences of clones from the deletion selection.
Reproduced in part with permission from RNA, volume 9, number 10, page 1215.
Copyright 2003 The RNA Society.

REFERENCE LIST

Beveridge, T.J. 2001. Use of the Gram stain in microbiology. *Biotech. & Histochem.* **71:** 111-118.

Breslow, R., Doherty, J.B., Guillot, G., and Lipsey, C. 1978. *β*-cyclodextrinylbisimidazole, a model for ribonuclease. *J. Am. Chem. Soc.* **100:** 3227-3229.

Breter, H.J. and Mertes, H. 1990. The quantitative determination of metabolites of 6-mercaptopurine in biological materials. VII. Chemical synthesis by phosphorylation of 6-thioguanosine 5'-monophosphate, 5'-diphosphate and 5'-triphosphate, and their purification and identification by reversed-phase/ion-pair high-performance liquid chromatography and by various enzymatic assays. *Biochim. Biophys. Acta* **1033:** 124-132.

Buzayan, J.M., Gerlach, W.L., and Bruening, G. 1986. Non-enzymatic cleavage and ligation of RNAs complementary to a plant satellite virus RNA. *Nature* **323:** 349-353.

Chaires, J.B., Dattagupta, N., and Crothers, D.M. 1982. Studies on interaction of anthracycline antibiotics and deoxyribonucleic acid: equilibrium binding studies on interaction of daunomycin with deoxyribonucleic acid. *Biochemistry* **21:** 3933-3940.

Chapple, K.E., Bartel, D.P., and Unrau, P.J. 2003. Combinatorial minimization and secondary structure determination of a nucleotide synthase ribozyme. *RNA* **9:** 1208-1220.

Dewar, M.J.S., Zoebisch, E.G., Healy, E.F., and Stewart, J.J.P. 1985. AM1: A new general purpose quantum mechanical model. *J. Am. Chem. Soc.* **107:**3902-3909

Doudna, J.A. and Cech, T.R. 2002. The chemical repetoire of natural ribozymes. *Nature* **418:** 222-228.

Ekland, E.H. and Bartel, D.P. 1995. The secondary structure and sequence optimization of an RNA ligase ribozyme. *Nucleic Acids Res.* **23:** 3231-3238.

Ellington, A.D. and Szostak, J.W. 1990. *In-vitro* selection of RNA molecules that bind specific ligands. *Nature* **346:** 818-822.

Fasman, G.D. 1976. *Handbook of Biochemistry and Molecular Biology, Proteins, I*, 183-203, CRC Press, 3 ed.

Fedor, M.J. and Williamson, J.R. 2005. The catalytic diversity of RNAs. *Nat. Rev. Mol. Cell. Biol.* **6:** 399-412.

Feringa, B.L. (editor). 2001. *Molecular Switches.* Wiley-VCH, Weinheim.

Fox, J.J., Wempen, I., Hampton, A., and Doerr, I.L. 1958. Thiation of nucleosides. I. Synthesis of 2-amino-6-mercapto-9-β-D-ribofuranosylpurine ("thioguanosine") and related purine nucleosides. *J. Am. Chem. Soc.* **80:** 1669-1675.

Gilat, S.L., Kawai, S.H., and Lehn, J.-M. 1995. Light-triggered molecular devices: photochemical switching of optical and electrochemical properties in molecular wire type diarylethene species. *Chem. Eur. J.* **1:** 275-284.

Gilbert, W. 1986. The RNA World. *Nature* **319:** 618.

Gorodetsky, B. *Ph.D. Thesis* Simon Fraser University (in progress).

Gorodetsky, B., Samachetty, H.D., Donkers, R.L., Workentin, M.S., and Branda, N.R. 2004. Reductive electrochemical cyclization of a photochromic 1,2-dithienylethene dication. *Angew. Chem. Int. Ed.* **43:** 2812-2815.

Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. and Altman, S. 1983. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **35:** 849-857.

Helm, M., Kopka, M.L., Sharma, S.K., Lown, J.W., and Giege, R. 2001. RNase activity of a DNA minor groove binder with a minimalist catalytic motif from RNase A. *Biochem. Bioph. Res. Co.* **281:** 1283-1290.

Igloi, G.L. 1988. Interaction of tRNAs and of phosphorothioate-substituted nucleic acids with an organomercurial. Probing the chemical environment of thiolated residues by affinity electrophoresis. *Biochemistry* **27:** 3842-3849.

Jadhav, V.R. and Yarus, M. 2002. Acyl-coAs from coenzyme ribozymes. *Biochemistry* **41:** 723-729.

Johnston, W.K., Unrau, P.J., Lawrence, M.S., Glasner, M.E., and Bartel, D.P. 2001. RNA-catalyzed RNA polymerization: accurate and general RNA-templated primer extension. *Science* **292:** 1319-325.

Joppich-Kuhn, R. and Luisi, P.L. 1978. Circular dichroic properties and conformation of thionicotinamide dinucleotides. *Eur. J. Biochem.* **83:** 587-592.

Joyce, G.F. 1989. RNA evolution and the origins of life. *Nature.* **338:** 217-224.

Joyce, G.F. and Orgel, L.E. 1999. *The RNA World* (Eds Gesteland, R.F. and Atkins, J.F. (1-25) Cold Spring Harbor Lab., Cold Spring Harbor, NY.

Kissinger, K., Krowicki, K., Dabrowiak, J.C., and Lown, J.W. 1987. Molecular recognition between oligopeptides and nucleic acids. Monocationic imidazole lexitropsins that display enhanced GC sequence dependent DNA binding. *Biochemistry.* **26:** 5590-5595.

Konan, Y.N., Gurny, R., and Allémann, E. 2002. State of the art in the delivery of photosensitizers for photodynamic therapy. *J. Photochem. Photobiol. B* **66:** 89-106.

Kruger, K., Grabowski, P.J., Zaug, A.J., Sands, J., Gottschling, D.E., and Cech, T.R. 1982. Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. *Cell* **31:** 147-157.

Lau, M.W.L., Cadieux, K.E.C., and Unrau, P.J. 2004. Isolation of fast purine nucleotide synthase ribozymes. *J. Am. Chem. Soc.* **126:** 15686-15693.

Levy, J.G. and Obochi, M. 1996. *New applications in photodynamic therapy introduction. Photochem. Photobiol.* **64:** 737-739.

Lohrmann, R. and Orgel, L.E. 1978. Preferential formation of (2′-5′) linked internucleotide bonds in non-enzymatic reactions. *Tetrahedron* **34:** 853-855.

Lohse, P.A. and Szostak, J.W. 1996. Ribozyme-catalysed amino-acid transfer reactions. *Nature* **381:** 442-444.

Lucas, L.N., van Esch, J., Kellogg, R.M., and Feringa, B.L. 1999. A new synthetic route to symmetrical photochromic diarylperfluorocyclopentenes. *Tetrahedron Lett.* **40:** 1775-1778.

Matsuda, K., Takayama, K., and Irie, M. 2001. Single-crystalline photochromism of a linear coordination polymer composed of 1,2-bis[2-methyl-5-(4-pyridyl)-3-thienyl]perfluorocyclopentene and bis(hexafluoroacetylacetonato)zinc(II). *Chem. Commun.* **(4):** 363-364.

Nissen, P., Hansen, J., Ban, N., Moore, P., and Steitz, T. 2000. The structural basis of ribosome activity in peptide bond synthesis. *Science* **289:** 920-930.

Norsten, T.B. and Branda, N.R. 2001. Axially coordinated porphyrinic photochromes for non-destructive information processing. *Adv. Mater.* **13:** 347-349.

Orgel, L.E. 1998. The origin of life-a review of facts and speculations. *Trends Biochem. Sci.* **23:** 491-491.

Pang, D.-W. and Abruna, H.D. 2000. Interactions of benzyl viologen with surface-bound single-and double-stranded DNA. *Anal. Chem.***72:** 4700-4706.

Peracchi, A. 2005. DNA catalysis: potential, limitations, open questions. *Chembiochem* **6:** 1316-1322.

Peters, A.J. 2003. *Ph.D. Thesis* Simon Fraser University.

Podyminogin, M.A., Vlassov, V.V., and Giege, R. 1993. Synthetic RNA-cleaving molecules mimicking ribonuclease A active center. Design and cleavage of tRNA transcripts. *Nucleic Acids Res.* **21:** 5950-5956.

Prody, G.A., Bakos, J.T., Buzayan, J.M., Schneider, I.R., and Bruening, G. 1986. Autolytic processing of dimeric plant virus satellite RNA. *Science* **281:** 1577-1580.

Qin, B., Yao, R., Zhao, X., and Tian, H. 2003. Enhanced photochromism of 1,2-dithienylcyclopentene complexes with metal ion. *Org. Biolmol. Chem.* **1:** 2187-2191.

Qu, X. and Chaires, J.B. 2000. Analysis of drug-DNA binding data. *Methods Enzymol.* **321:** 353-369.

Raines, R.T. 1998. Ribonuclease A. *Chem. Rev.* **98:** 1045-1065.

Rodgers, G.E. 1994. *Introduction to coordination, solid state, and descriptive inorganic chemistry.* McGraw-Hill, Inc.

Sabeti, P.C., Unrau, P.J., and Bartel, D.P. 1997. Accessing rare activities from random RNA sequences: the importance of the length of molecules from the starting pool. *Chem. Biol.* **4:** 767-774.

Saito, M., Yokoyama, Y., and Yokoyama, Y. 2003. Reversible change in optical rotation by photochromism of diarylethenes in the stretched DNA-quaternary ammonium ion complex films. *Chem. Lett.* **32:** 806-807.

Saville, B.J. and Collins, R.A. 1990. A site-specific self-cleavage reaction performed by a novel RNA in Neurospora mitochondria. *Cell* **61:** 685-696.

Sharmeen, L., Kuo, M.Y.-P., Dinter-Gottleib, G., and Taylor, J. 1988. Antigenomic RNA of of human hepatitis delta virus can undergo self-cleavage. *J. Virol.* **62:** 2674-2679.

Starcevic, K., Karminski-Zamola, G., Piantanida, I., Zinic, M., Suman, L., and Kralj, M. 2005. Photoinduced switch of a DNA/RNA inactive molecule into a classical intercalator. *J. Am. Chem. Soc.* **127:** 1074-1075.

Suh, D. and Chaires, J.B. 1995. Criteria for the mode of binding of DNA binding agents. *Bioorg. & Med Chem.* **3:** 723-728.

Tung, C.-H., Wei, Z., Leibowitz, M.J., and Stein, S. 1992. Design of peptide-acridine mimics of ribonuclease activity. *Proc. Natl. Acad. Sci. USA* **89:**7114-7118.

Unrau, P.J. and Bartel, D.P. 1998. RNA-catalysed nucleotide synthesis. *Nature* **395:** 260-263.

Valadkhan, S., and Manley, J.L. 2001. Splicing-related catalysis by protein-free snRNAs. *Nature* **413:** 701-707.

Vanden Eynde, J.J., Delfosse, F., Lor, P., and Van Haverbeke, Y. 1995. 2,3-Dichloro-5,6-dicyano-1,4-benzoquinone, a mild catalyst for the formation of carbon-nitrogen bonds. *Tetrahedron* **51:** 5813-5818.

Vlassov, V.V., Zuber, G., Felden, B., Behr, J.-P., and Giege, R. 1995. Cleavage of tRNA with imidazole and spermine imidazole constructs: a new approach for probing RNA structure. *Nucleic Acids Res.* **23:** 3161-3167.

Wang, Q.S. and Unrau, P.J. 2002. Purification of a histidine-tagged T4 RNA ligase from *E. coli*. *Biotechniques* **33:** 1256-1260.

Yuan, Y., Xiao, R., Gao, G., Su, X.-Y., Yu, H., You., J., and Xie, R.-G. 2002. A direct synthetic approach to tripodal imidazole compounds. *J. Chem. Research (S)* 267-269.

Zegers, I., Maes, D., Dao-Thi, M.-H., Poortmans, F., Palmer, R., and Wyns, L. 1994. The structures of RNase A complexed with 3′-CMP and d(CpA): active site conformation and conserved water molecules. *Protein Sci.* **3:** 2322-2339.

Zollinger, H. 1991. *Colour chemistry: synthesis, properties and applications of organic dyes and pigments*, Weinheim, New York, 2nd rev.