A LIGHT-HARVESTING CATALYTIC DNA FOR THYMINE DIMER REPAIR

by

Daniel J.F. Chinnapen

B.Sc., Simon Fraser University, 1998

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

DOCTOR OF PHILOSOPHY

In the Department of

Molecular Biology and Biochemistry

SIMON FRASER UNIVERSITY

Spring 2005

APPROVAL

Name: Daniel J.F. Chinnapen

Degree: Doctor of Philosophy (Science)

Title of Thesis: A Light-Harvesting Catalytic DNA for Thymine Dimer Repair

Examining Committee:

Chair: Dr. Jamie Scott
Professor of Molecular Biology and Biochemistry, SFU

Dr. Dipankar Sen
Senior Supervisor
Professor of Molecular Biology and Biochemistry, SFU

Dr. Erika Plettner
Committee Member
Assistant Professor of Chemistry, SFU

Dr. Peter Unrau
Committee Member
Assistant Professor of Molecular Biology and Biochemistry, SFU

Dr. Vance Williams
Committee Member
Assistant Professor of Chemistry, SFU

Dr. Melanie O'Neill
Internal Examiner
Assistant Professor of Chemistry
Department of Chemistry

Dr. Steven Rokita
External Examiner
Professor of Chemistry & Biochemistry
University of Maryland, MD

Date Defended/Approved: January 7, 2005

SIMON FRASER UNIVERSITY

PARTIAL COPYRIGHT LICENCE

ABSTRACT

Since the discovery of ribozymes in the 1980's, there have been many examples of catalytic DNA or RNA that can catalyze a variety of reactions, including phosphodiester bond cleavage, porphyrin metallation and nucleotide synthesis. A novel catalytic DNA (DNAzyme) was generated by *in vitro* selection that could perform a photochemical reaction, the photoreversal of thymine dimers upon irradiation of light. Photolyase enzymes found in nature catalyse the photorepair of thymine dimers with the aid of sensitizers, such as FADH and tryptophan. Curiously, one DNAzyme, UV1C, could repair thymine dimers without the aid of photosensitizers. UV1C was found to be active when using light that was significantly red-shifted relative to that of normal DNA absorbance. Using spectroscopic techniques, it was determined that the red-shifted action spectrum and rate enhancement profile were caused by the presence of a guanine quadruplex in the catalytic core.

Further characterization of UV1C revealed interesting information on the folding, substrate specificity and enzyme mechanism. Using sequence mutant constructs, it was determined that the 3'-binding arm of the enzyme could be removed with very little effect on repair rates. Cross-linking reactions determined that the thymine dimer was situated in close proximity, or stacked with the putative guanine quartet. Interestingly, substrates containing uracil dimers with deoxyribose sugar rings were efficiently repaired by UV1C, whereas thymine dimer substrates containing ribosyl sugar rings were not. Oxidative

damage to guanines was not detected, which implied that a fast rate of back electron transfer may have occured after the thymine dimer photosplitting process.

DNA "aptamers" (binders) were generated by *in vitro* selection that could simultaneously bind to the electron transfer protein cytochrome *c*, and to the small metalloporphyrin, hemin. Aptamers selected contained guanine-rich sequences located centrally with respect to the overall sequence. Chemical probing analysis determined that a central guanine-quadruplex formed that was responsible for binding to hemin. Through footprinting analysis, the binding site for cytochrome *c* on the DNA was determined. The ternary complex represents a potential model for the study of electron transfer between artificial and naturally-occurring heme electron transfer systems.

DEDICATION

To my family for their inspiration and motivation through the years and Himani for her encouragement and support.

ACKNOWLEDGEMENTS

I would like to express my appreciation and gratitude for Dr. Dipankar Sen for his support and guidance during my graduate career, as well as giving me the chance to do research at a high level. I also thank members of my supervisory committee, Dr. P. Unrau, Dr. M. O'Neill, Dr. E. Plettner, and Dr. V. Williams for important suggestions and advice for my research projects.

I would also like to thank present members of the Sen laboratory for their discussions and for their help during my graduate research: Yong Liu, Ed Leung, Becky Thorne, Yi-Jeng Huang, Kelly Chapple and Hyun-Wu Lee.

I am grateful to past members of the Sen Lab for their inspiration and guidance: Dr. Ron Geyer, Dr. Yingfu Li, Dr. Dennis Wang, Dr. Anat Feldman, Dr. Paola Travascio and Carlo Sankar.

A special thanks to Dr. Unrau for important feedback and key suggestions on the thymine dimer project, and to Dr. M.O'Neill for her invaluable help with photochemistry.

I thank Gary Leach for use of the nitrogen laser during for the thymine dimers experiments.

I would like to thank friends and staff in the M.B.B. department that have helped me through the years.

I like to thank Dr. Grant Mauk for discussions on cytochrome c, as well as Dr. Dominic Zichi (Somalogic, Boulder, CO) for his advice on sequence analysis algorithms.

I thank Dr. David Baillie for his encouragement in the early years of my research career.

I would also like to thank the M.B.B. department and the university for their financial support through the years, as well as NSERC Canada and CIHR Canada funding.

I thank Himani for her help editing my thesis.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ATP	adenosine 5'-triphosphate
CH6, CH6A, CH6B	sequences derived from the cytochrome-hemin selection
CPD	cyclobutane pyrimidine dimer
DEPC	diethyl pyrocarbonate
DMS	dimethyl sulfate
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dTdU	TDP with a thymidine-deoxyuridine dimer
dTrT	TDP with the 5' sugar of the dimer being ribose
dUdU	TDP with deoxyuridine dimers instead of thymidine dimers
dUdT	TDP with a deoxyuridine-thymidine dimer
GTP	guanosine 5'-triphosphate
hemin	iron(III) protoporphyrin IX
FMN	flavin mononucleotide
H_2O_2	hydrogen peroxide
kJ/mol	kilojoules per mole – unit of thermodynamic energy
NHE	Normal Hydrogen Electrode – reference electrode
NTP	nucleotide 5'-triphosphate
PCR	polymerase chain reaction
pCR2.1	Invitrogen plasmid used for cloning of DNA library
PS2M	a small guanine-rich oligonucleotide (5'-GT GGT AGG GCG GGT TGG-3')
RNA	ribonucleic acid
R-U	RNA sequence 5' to a uracil dimer, with the remainder being DNA
rTrT	TDP with ribothymidine dimers instead of thymidine dimers
rTdT	TDP with the 3' sugar of the dimer being ribose
RDP (R-R)	an all-RNA version of TDP
SCE	Standard Calomel Electrode – reference electrode
SELEX	Systematic Evolution of Ligands by Exponential Enrichment

TDP	thymine dimer-containing primer
TE	Tris-Chloride-EDTA
Tris	tris(hydroxymethyl) aminomethane
U-R	DNA sequence 5' to a uracil dimer, with the remainder being RNA
UDP (U-U)	TDP substrate with a ribo-uridine dimer instead of thymidine dimer
UV1A,UV1B,UV1C	sequences derived from the serotonin-independent selection and chosen for further characterisation
V	volts

CHAPTER 1

INTRODUCTION TO NUCLEIC ACID CATALYSIS, IN VITRO SELECTION AND THYMINE DIMER REPAIR

1.1 Introduction

The work described in this thesis covers the *in vitro* selection and characterization of: 1) a catalytic DNA that photorepairs thymine dimers, and, 2) a DNA aptamer that simultaneously binds to the electron transfer protein cytochrome c and to the small molecule hemin.

This introductory chapter will provide the reader with some background information and help put into context the work presented herein. First, I will give a brief introduction to catalytic nucleic acids and their significance to the RNA world hypothesis. Guanine quadruplex structures as well as the technique of *in vitro* selection will be discussed; both are common threads in my projects. Finally, an overview on the mechanism of thymine dimer repair by photolyase enzymes will be presented.

1.2 Catalytic RNA

The discovery of ribozymes by Thomas Cech and Sydney Altman over twenty years ago (Guerrier-Takada *et al.*, 1983; Kruger *et al.*, 1982) changed the conventional thinking that biological catalysis was the sole responsibility of proteins. It had classically been thought that the role of nucleic acids in the cell was primarily to serve as a means of

storage for genetic information, whereas all cellular enzymatic processes were the exclusive responsibility of protein enzymes. As a catalyst, one would assume that a nucleic acid polymer would be significantly disadvantaged when compared to a protein. Proteins have a relatively large diversity of chemical functionalities with which to catalyse diverse reactions. Many of the twenty amino acid side chains have desirable properties for catalysis, such as the near-neutral pKa values of histidine, positive and negatively charged groups (e.g. lysine and aspartic acid, respectively), reducibility (cysteine) and varying degrees of hydrophobicity with which to form a hydrophobic active site. Nucleic acid bases, on the other hand, consist of four relatively inert components and are relatively poor acids and bases at neutral pH. Also, a protein can form many secondary and tertiary structural motifs such as binding pockets and catalytic domains, wherein the position of the side chains interact with the substrates in a manner suitable for catalysis. Nucleic acids, on the other hand, primarily form hydrogen bonds between complementary bases on opposite strands to produce mainly double helical elements that orient the four possible functional groups towards the interior of the structure. But in spite of these shortcomings, many researchers have effectively demonstrated that nucleic acids can indeed be catalytic (Breaker, 1997; Joyce, 2004; Wilson and Szostak, 1999) and can fold to form very complex three-dimensional structures (reviewed in (Moore, 1999)).

Naturally occurring ribozymes have been found to catalyse reactions involving either phosphodiester cleavage or phosphodiester transfer reactions. They are categorized as either "large" or "small" ribozymes based on their length being greater than, or less than 200 nucleotides respectively (McKay and Wedekind, 1999; Sigurdsson

et al., 1998). The "small" ribozymes catalyse the cleavage of an RNA phosphodiester backbone, and are comprised of the hammerhead (Tuschl *et al.*, 1995), the hairpin (Walter and Burke, 1998), hepatitis delta virus (HDV) (Shih and Been, 2002), and Neurospora VS ribozymes (Saville and Collins, 1990). The cleavage reaction results in the formation of a 2', 3'-cyclic phosphate on the 5' RNA strand, and a 5'-hydroxyl on the 3' RNA strand (Figure 1-1). The "large" ribozyme family is comprised of the group I and group II introns found in Tetrahymena, and RNAse P [Reviewed in (Cech *et al.*, 1992; Cobaleda and Sanchez-Garcia, 2001; Fedorova *et al.*, 2002; Takagi *et al.*, 2001)]. Both the group I and group II intron ribozymes catalyse

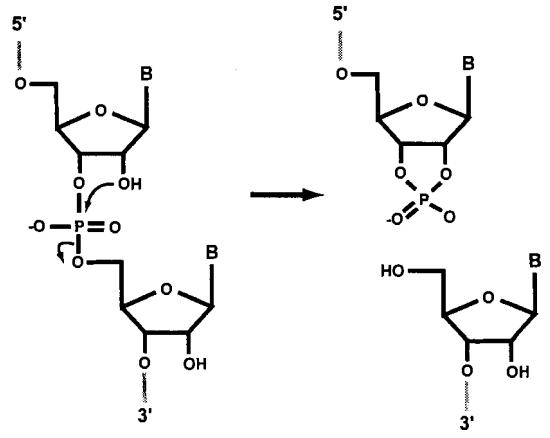


Figure 1-1 Mechanism of RNA cleavage by naturally occurring "small" ribozymes.

self-splicing via a transesterification reaction. In the group I intron, an external guanosine facilitates the cleavage at a 5'-intron-exon junction to yield a new 3'terminus, shown in Figure 1-2. A similar mechanism is seen with of the group II intron, wherein an internal 2'-hydroxyl group acts as a nucleophile instead. The RNA component of RNAse P

catalyses the maturation of transfer-RNA's by trimming off bases from the 5'end using water as a nucleophile instead of a 2'-hydroxyl (Cobaleda and Sanchez-Garcia, 2001).

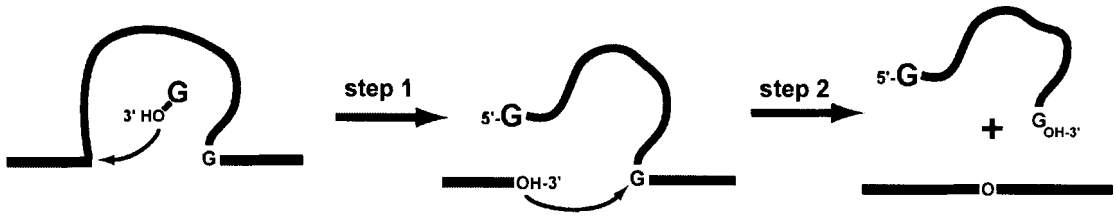


Figure 1-2 Self-splicing by the Tetrahymena group I intron.

1.3 Ribozymes and the RNA World

Early attempts to formulate a scientific theory for how life began on Earth revolved around proteins (Miller, 1974). It is believed the first forms of life may have taken the form of very simple molecules capable of catalysis and self-replication within the harsh reducing environment of early Earth. Stanley Miller and Harold Urey's famous experiments in the 1950's (Miller, 1953; Miller and Urey, 1959) demonstrated that simple amino acids could spontaneously form when a mixture containing ammonia, methane, hydrogen and water were subjected to extreme conditions simulating that of the early Earth atmosphere. It was extrapolated from these experiments that amino acids could possibly assemble and evolve into protein enzymes within the primordial soup. Although it is conceivable that there was a spontaneous formation of amino acids that led to the first biocatalysts in the early Earth, they would also need to replicate and pass on genetic information, a process not attributed to proteins. Nucleic acids are well suited for this task, but until their catalytic potential was realised in the early 1980's, there remained a "chicken and the egg" paradox. If a favourable assemblage of amino acids were the first biocatalysts, eventually evolving into the multicellular and diverse organisms of today,

how could their genetic information be encoded and passed to successive generations in the absence of nucleic acids? Conversely, if nucleic acids were the first forms of life on Earth, how could they perform enzymatic functions for their own replication and survival in the absence of protein enzymes?

The discovery of ribozymes gave full sail to the idea that the first biocatalysts on Earth may have been RNA molecules (Benner and Ellington, 1991; Benner *et al.*, 1989; Gilbert, 1986). The "RNA world hypothesis" envisions that the first forms of life on Earth came in the form of RNA or RNA-like molecules due to: 1) its ability to encode genetic information, and 2) to catalyse reactions. Within this "primordial soup", RNA or RNA-like molecules were proposed to perform catalysis that could ensure the survival of these RNA. Over the course of billions of years, as the need for more complex forms of catalysis arose, RNA would eventually evolve polypeptide enzymes, and their genetic information could be stored by a more stable nucleic acid analogue, DNA (Bartel and Unrau, 1999). In effect, it could be thought that RNA would "delegate" the tasks of catalysis and genetic storage to more qualified biopolymer molecules, that of proteins and DNA respectively. By the late 1990's, many examples of ribozymes had been found in nature and effectively demonstrated to catalyse phosphodiester bond cleavage and RNA splicing reactions. Although, these examples would provide compelling evidence for an RNA world, up to that point there were no ribozymes found that could catalyse reactions that could be essential for RNA world survival, such as nucleotide synthesis and RNA replication. This begged the question: could artificial ribozymes be created that could perform catalysis essential for self-survival in the RNA world?

Recent work by Bartel and co-workers demonstrated that ribozymes could be successfully evolved with the remarkable ability of to perform tasks related to the survival and replication of RNA in the primordial Earth (Johnston *et al.*, 2001; Unrau and Bartel, 1998). The class I ligase ribozymes, derived from random RNA sequences, catalysed the template-directed extension of an oligonucleotide with the addition of GTP located at the first residue of the ribozyme (Ekland and Bartel, 1996). It was also found that the free nucleotide triphosphates (NTP's) could be incorporated specifically at the 3'-end of the ribozyme in a template-dependent manner. But, because this ribozyme could only extend the primer two to three nucleotides at a time, *in vitro* evolution was employed to further evolve it from an RNA ligase into a RNA polymerase (Johnston *et al.*, 2001). Eighteen rounds of selection were carried out to generate a functional ribozyme (named 18.12.23) containing 189 nucleotides. The 18.12.23 ribozyme could extend a detached primer when supplied with template by up to fourteen nucleotides with high fidelity. In a separate study, the same research group had used *in vitro* selection (explained later in this chapter) to select for a ribozyme capable of pyrimidine nucleotide synthesis at the 3' terminus using a tethered 5-phosphoribosyl-1-pyrophosphate (pRpp) and 4-thiouracil base (Unrau and Bartel, 1998). Interestingly, mechanistic studies on the glycosidic bond formation of the ribozyme indicated a mechanism more closely resembling that of hydrolytic enzymes than that of protein-based nucleotide synthases (Unrau and Bartel, 2003). In essence, the above two ribozymes supply important pieces of the puzzle in the realization of a putative RNA World. Added to this is the large number of new ribozymes that have been derived from *in vitro* selection experiments, summarized in Table 1-1 (adapted from (Bartel and Unrau, 1999)).

Table 1-1 New ribozymes derived from *in vitro* selection experiments.

Ribozyme Activity	Ref	Ribozyme Activity	Ref
Phosphodiester cleavage	(Williams *et al.*, 1995)	Cyclic phosphate hydrolysis	(Pan and Uhlenbeck, 1992)
RNA ligation	(Bartel and Szostak, 1993)	Polymerisation	(Ekland and Bartel, 1996; Johnston *et al.*, 2001)
Porphyrin Metallation	(Conn *et al.*, 1996)		
RNA phosphorylation	(Lorsch and Szostak, 1994)	Tetraphosphate cap formation	(Chapman and Szostak, 1995)
RNA branch formation	(Tuschl *et al.*, 1998)	RNA aminoacylation	(Illangasekare *et al.*, 1995)
Acyl transfer	(Lohse and Szostak, 1996)	Amide Bond formation	(Lohse and Szostak, 1996)
Peptide bond formation	(Zhang and Cech, 1997)	RNA alkylation	(Wilson and Szostak, 1995)
Sulfur alkylation	(Wecker *et al.*, 1996)	Diels Alder addition	(Jaschke and Seelig, 2000)
Glycosidic bond formation	(Unrau and Bartel, 1998)	Bridged biphenyl isomerization	(Prudent *et al.*, 1994)

The recent crystal structure of the ribosome and subsequent studies have revealed that the large subunit of the ribosome is indeed a ribozyme (Ban *et al.*, 2000; Nissen *et al.*, 2000). The ribosome is a large ribonucleoprotein complex responsible for the synthesis of proteins in the cell. Comprised of both RNA and protein components, it was previously thought that the protein component was responsible for the peptidyl transferase reaction, and that the RNA component played simply a structural role. The high-resolution structure showed that within 18 Å of the catalytic centre, there was the absence of any protein moiety to catalyse the transferase reaction, leading to the conclusion that RNA is performing the catalysis. In essence, the ribosome is a ribozyme, and it is most likely the protein component that is providing the structural role for the

RNA catalyst and not vice versa (Nissen *et al.*, 2001). More importantly, this fuels the idea that RNA could have evolved and synthesized the first protein enzymes.

1.4 Catalytic DNA

Not long after the discovery of ribozymes there was speculation whether DNA could also be catalytic as found with their RNA counterparts (Cech and Bass, 1986). In the cell, DNA exists primarily within a double stranded context and serves primarily as means of genetic information storage. Due to the limited flexibility imposed by a double helical structure, DNA within the cell has a restricted ability to fold into higher-ordered structures. RNA, on the other hand, had been known to be able to fold into more complex structures and form complex motifs, as exemplified with high-resolution structures of transfer-RNAs and early *in vitro* selection experiments for RNA that bound to small molecule targets such as ATP (Dieckmann *et al.*, 1997; Dieckmann *et al.*, 1996; Jiang *et al.*, 1996; Sassanfar and Szostak, 1993). Also, catalysis performed by the naturally-occurring RNA enzymes involve the 2'-hydroxyl in the phosphodiester bond cleavage step [reviewed in (Tuschl *et al.*, 1995)]. It is interesting to note that no deoxyribozymes to this date have been found in nature. But even if an artificial deoxyribozyme could be synthesized, as was done with their RNA counterparts using *in vitro* selection, would it be nearly as fast as any of the known ribozymes?

Just ten years ago, Breaker and Joyce presented the first artificial deoxyribozyme (DNAzyme), which had the property to cleave RNA in the presence of Pb^{2+} (Breaker and Joyce, 1994). *In vitro* selection (Ellington and Szostak, 1990; Tuerk and Gold, 1990), discussed in later in this chapter, was used to generate a single-stranded DNA that accelerated the cleavage at a ribonucleotide position within an otherwise all DNA

substrate. The rate enhancement of the background phosphodiester cleavage was increased by an impressive 100,000 fold with a k_{cat} of 1 min^{-1} and K_m of 2 μM. It was found that DNA helped the positioning of Pb^{2+} to allow for easy deprotonation of the 2'-hydroxyl next to the site of cleavage, which eventually leads to the hydrolysis of the internucleotide phosphate backbone (Joyce, 2004).

The 10-23 deoxyribozyme selected by Santoro and Joyce (Santoro and Joyce, 1997) best exemplifies the catalytic potential and versatility of catalytic DNA. Their goal was to generate DNAzymes capable of cleaving any sequence of RNA to be potentially used as therapeutic agents for gene silencing. Using *in vitro* selection, they successfully generated deoxyribozymes that could catalyse the cleavage of any given RNA sequence, including biologically important messenger RNA, under physiological salt conditions. The DNAzyme could be "tailor made" to cleave any RNA sequence by simply designing complementary "binding arms" to that of the target sequence. The versatility of the 10-23 DNAzyme is demonstrated by the fact that it has indeed been successfully applied as a therapeutic agent *in vivo* (reviewed in (Cairns *et al.*, 2002; Khachigian, 2004), as a diagnostic agent for "quantitative" PCR (Impey *et al.*, 2000), as well as a logic gate operator for use in molecular computing (Stojanovic *et al.*, 2002). Another important feature of the 10-23 DNAzyme is that it has a high the catalytic efficiency (k_{cat}=0.1 min^{-1}, K_m=1 nM corresponding to a k_{cat}/K_m of 10^8 M^{-1} min^{-1}), which compares favorably to those of the known RNA-cleaving RNA enzymes. Interestingly, a DNAzyme with similar properties, but with different sequence, has recently been found by Feldman and Sen (Feldman and Sen, 2001).

Apart from phosphodiester bond cleavage and ligation reactions, deoxyribozymes have been selected that can perform a variety of different types of reactions. These include DNA deglycosylation (Sheppard *et al.*, 2000), peroxidation (Travascio *et al.*, 1998), DNA phosphorylation (Wang *et al.*, 2002), DNA capping (Li *et al.*, 2000) and porphyrin metallation (Li and Sen, 1996). These are summarized in Table 1-2. The last example shown, a DNAzyme capable of the photorepair of thymine dimers in a light-driven reaction (Chinnapen and Sen, 2004), is subject of this thesis.

Table 1-2 Examples of reactions catalysed by catalytic DNA.

DNAzyme Activity	Ref	DNAzyme Activity	Ref
RNA phosphodiester cleavage	(Breaker and Joyce, 1994; Breaker and Joyce, 1995; Geyer and Sen, 1997)	DNA phosphodiester cleavage	(Carmi *et al.*, 1998)
DNA ligation	(Cuenoud and Szostak, 1995)	DNA deglycosylation	(Sheppard *et al.*, 2000)
DNA phosphorylation	(Li and Breaker, 1999)	Adenosylation	(Li *et al.*, 2000)
Porphyrin metallation	(Li and Sen, 1996)	Peroxidation	(Travascio *et al.*, 1998)
Thymine dimer repair	(Chinnapen and Sen, 2004)		

1.5 *In Vitro* Selection

In vitro selection, or SELEX (**S**ystematic **E**volution of **L**igands with **Ex**ponential Enrichment), is a very powerful that has been used to generate many artificial nucleic acid catalysts, or aptamers (nucleic acid binders) respectively. Independently developed by the labs of Gold (Tuerk and Gold, 1990), Szostak (Ellington and Szostak, 1990) and Joyce (Joyce, 1989), impact of this technique has broadened our understanding on the capabilities of nucleic acids, as well as provided researchers with a tool to generate new

therapeutic reagents (Davis *et al.*, 1996; Gold *et al.*, 1997; Tian *et al.*, 1995). Although the specific details of the individual application used can vary, the general principle is the same. (Figure 1-3) A large combinatorial pool of unique DNA or RNA sequences is synthesized, typically with a sequence diversity in the order of 10^{14}. These sequences are then folded, under desired salt, pH and temperature conditions, and put through a screening process by which the successful sequences for catalysis, or binding, can be separated from those unsuccessful sequences.

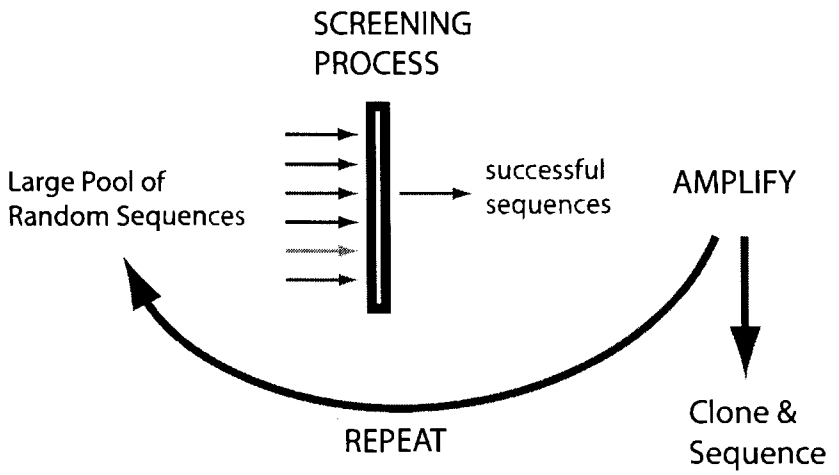


Figure 1-3 A General Selection Scheme.

These successful "catalysts" are then amplified by PCR to generate double-stranded DNA, which is later made into single-stranded DNA or RNA to re-generate a starting pool of oligonucleotides. This is subject to iterative cycles (rounds) until there is an increase in the signal to noise ratio. Typically, the stringency of the reaction can be increased during the course of the selection. For example, in the case of nucleic acid catalysts, an incubation time of 30 minutes from the starting rounds can be gradually

decreased down to 1 minute in the final rounds. By doing this, the selection pressure "weeds out" the slower catalysts from the selection pool. After a satisfactory amount of enrichment is achieved, the resultant pool of sequences are then finally cloned and individually sequenced. Each sequence, or a representative number of sequences, can then be individually tested for activity.

1.5.1 Selection Design

An important consideration when designing the initial pool of random sequence macromolecules is the amount of genetic diversity desired. The total possible number of unique sequences that can be made from nucleic acids of length n is represented by 4^n. A random sequence of all the possible 20-mers corresponds to 10^{12} unique sequences. One drawback in using a random sequence pool of this size is that twenty nucleotides may not provide the complexity needed to fold and catalyse, or to bind a specific target. Conversely, if a very large sequence length is used, e.g. 80 nucleotides (4^{80}), only a very small fraction of the total of number of possible sequences can be represented since it is physically impossible from obtain enough DNA from a synthesis to cover all possible 80-mers. Also, the synthesis length limit of modern automated DNA synthesis machines to generate quality oligonucleotides is, at maximum, approximately 120 nucleotides. For this reason, selections have typically used DNA constructs containing randomised regions in the range of 40 nucleotides, with flanking constant regions of 15-20 nucleotides each for PCR. An alternative method that has been used to incorporate higher numbers of randomised positions is to join together multiple random regions on a single strand. This approach has been successfully used to generate libraries for many ribozyme selections [reviewed in (Joyce, 1989)].

There are a number of different methods that have been used to generate DNA/RNA aptamers and catalysts. Two general approaches can be used for the selection of catalytic nucleic acids; one involves the binding to a transition-state analogue, and the other involves a more direct approach. In the selection for either catalysts or binders, the separation of the successful from the unsuccessful sequences can be accomplished using either a column or a gel-based method.

1.5.1.1 Indirect versus Direct Selections for Nucleic Acid Catalysts

For the selection of nucleic acid catalysts, two general strategies exist: indirect selection, and direct selection. With the "indirect" strategy for selection, the nucleic acid is selected first to bind to a transition-state analogue of the intended reaction. Transition-state theory states that the substrate for a reaction goes through an unstable intermediate, or transition-state, that then is transformed into the product (Jencks, 1987). Theoretically, enzymes increase the rate of reaction by preferentially stabilising the transition-state, thus lowering the free energy required for catalysis. Stable versions of transition states, called transition-state analogues, exist for a variety of reactions. These have been employed as ligands for the development of "catalytic antibodies", wherein an antibody generated to bind to a given transition-state analogue is subsequently able to catalyse the represented reaction (Cochran *et al.*, 1988). By analogy, if a nucleic acid could be selected that binds to the transition-state analogue of a given chemical reaction, there is the possibility for it to also catalyse the reaction. An advantage of this method is that the catalysts derived from selection can immediately be used to perform multiple turnover catalysis, *i.e.* a single catalyst can use multiple substrates. This strategy has been successfully applied to generate a DNAzyme for porphyrin metallation, wherein DNA aptamers were generated

to bind to N-methylmesoporphyrin (NMM). In the presence of an empty porphyrin molecule, the aptamer was able to insert metal ions into the center to produce metalloporphyrins (Li *et al.*, 1996; Li and Sen, 1996; Li and Sen, 1998). Although a conceptually attractive strategy, this approach has not been very successful for the selection of nucleic acid catalysts, with the above example being one of the few examples derived by this method to date (Conn *et al.*, 1996; Li and Sen, 1996; Prudent *et al.*, 1994).

The "direct" strategy for *in vitro* selection is the most used and by far the most successful method for generating ribozymes or deoxyribozymes. In this method, the experiment is designed such that the catalytic event self-modifies the nucleic acid directly so that it can be separated from those sequences that cannot perform catalysis. An example of a construct that can be used for the selection of RNA-cleaving DNA is shown in Figure 1-4. The DNA construct contains a central random sequence (typically 20-50 nucleotides) and constant regions on either side to act as PCR priming sites for

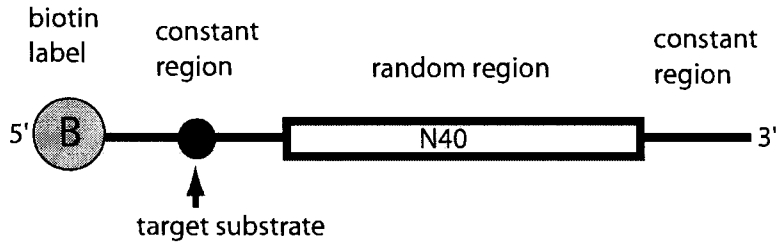


Figure 1-4 A typical construct used for selection.

amplification. The 5' constant region contains the target substrate, in this case, a ribonucleotide in an otherwise all-DNA sequence. During the selection process, any catalytic sequence that cleaves at this ribonucleotide position results in either its own

liberation from an immobilised column, and/or a size difference that can be separated by gel or other method. A typical selection scheme is shown in Figure 1-5. Additionally, a "negative selection" step can be performed wherein any background levels of activity (noise) can be removed during every round of selection. This step is crucial for the generation of high-specificity catalysts.

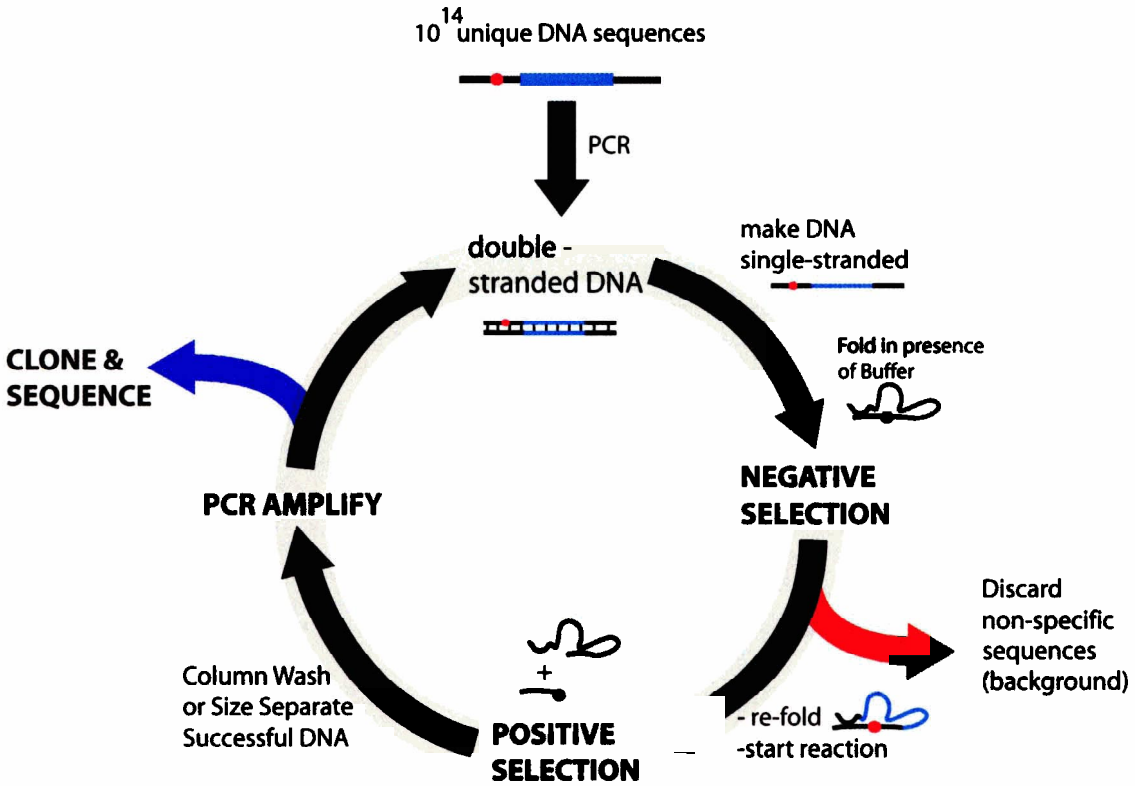


Figure 1-5 A typical selection scheme for self-cleaving DNA.

1.5.1.2 Column-based Methods

For the selection of aptamers (binders) for a specific target ligand, a column-based method can be used. Typically the strategy would involve the covalent attachment of the target ligand (e.g. organic molecule or protein), to a solid support or bead, such as agarose beads (shown in Figure 1-6). The folded random DNA/RNA pool is then added to the column, and is allowed to bind to the column. When the column is washed through, the successful binders are retained and the non-binders are washed through. The bound sequences can then be eluted off in the presence of a denaturant, such as urea. These eluted sequences can then be amplified by PCR to generate a new starting pool, and the process is repeated for many rounds. The first demonstrations of SELEX used similar methods wherein RNA aptamers were selected to bind to organic dyes (Ellington and Szostak, 1990) and to ATP (Sassanfar and Szostak, 1993).

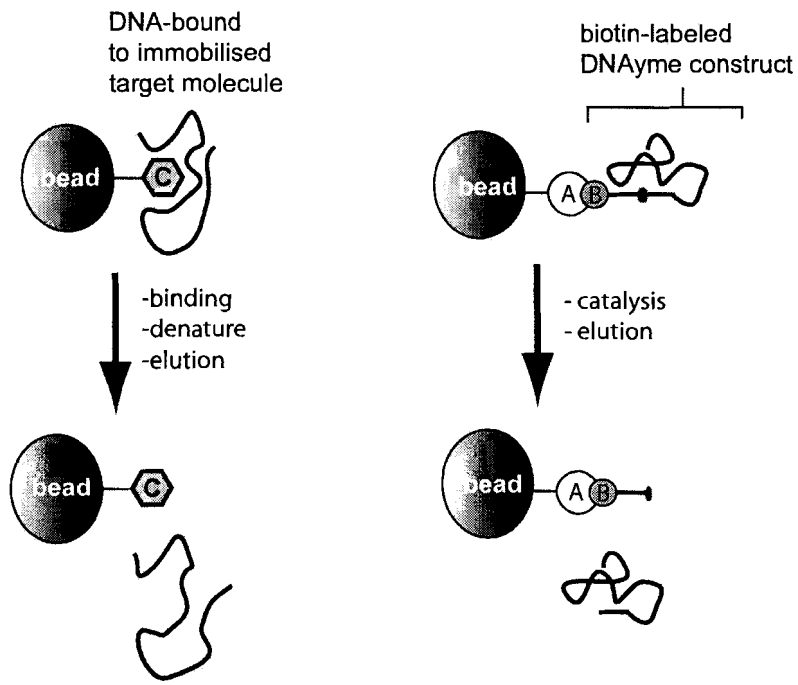


Figure 1-6 Column methods used in SELEX for aptamers (left) and catalysts (right).

The column-based method can also be applied to the selection for ribozymes or DNAzymes. Although there are many examples of more elegant column-based strategies that have been used recently, a simple method similar to the selection for the lead-dependent deoxyribozyme will be explained. Typically, the construct used contains a biotin label at the 5'-end of the DNA construct (Figure 1-4), which allows the DNA to be immobilised onto an avidin-conjugated agarose bead due to the very high affinity between these two (dissociation constants in the femtomolar range). Depending on the design of the experiment, a catalytic event can result in the sequence to be liberated from the column, which can then be eluted off and amplified (Figure 1-6).

1.5.2 Post-Selection Characterization

After a selection has been carried out and the successful sequences have been isolated, each sequence is normally tested for activity. By design, most selection strategies generate catalytic nucleic acids that act on themselves, e.g. are self-splicing, self-cleaving, or self-modifying. Although these nucleic acids have characteristics of being catalytic, it can be argued that they are not true catalysts by strict definition. To be a "true" enzyme, one has to meet the following criteria: 1) accelerate the rate of reaction, 2) does not cause a thermodynamically unfavorable reaction to become favorable, and 3) must not be consumed or modified in the reaction. This third criterion is not met with nucleic acids that are immediately derived from a selection as the ribozyme/DNAzyme is usually modified by itself in the reaction. Essentially, a *cis*-acting ribozyme/DNAzyme is considered a single turnover catalyst, whereas one that can act *in trans,* i.e. use an external substrate, is considered a multiple turnover catalyst. In order to engineer a multiple turnover enzyme, or a "true enzyme", first the catalytic regions of the enzyme

are identified through experimentation, as well as binding and substrate recognition regions. The enzymatic portion can then be separated from the substrate region, usually by cleavage within a looped region, illustrated in Figure 1-7. Post-selection modification as such have been used to generate many examples of multiple-turnover DNAzymes, including one presented in this thesis (Breaker and Joyce, 1994; Feldman and Sen, 2001; Santoro and Joyce, 1997).

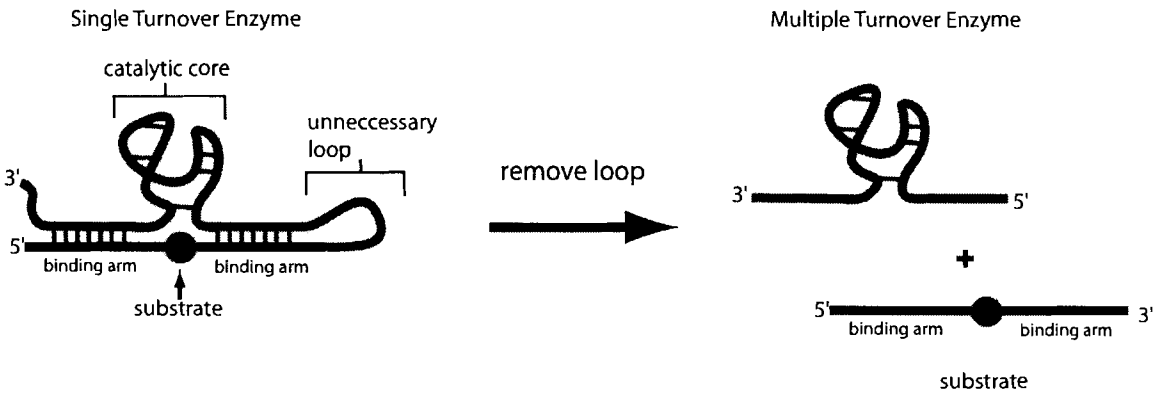


Figure 1-7 Creating of a multiple turnover catalyst from a cis-acting ribozyme/DNAzyme.

1.6 RNA and DNA "Aptamers"

In addition to the selection for novel catalysts, *in vitro* selection allows for the generation of single-stranded DNA or RNA "aptamers", or binders, for a specific target ligand [reviewed in (Wilson and Szostak, 1999)]. Aptamers could be regarded as nucleic acid versions of antibodies, and have been shown to bind, with high affinity and specificity, to diverse ligands ranging from simple ions (Ciesiolka and Yarus, 1996) to entire cells (Morris *et al.*, 1998). For example, an RNA aptamer that was generated to target theophylline with high affinity was found not to bind to caffeine, which differs

from theophylline by only one methyl group. Also, the 2'-hydroxyl has been shown to play an important role in helix stability and interaction with the target ligand (Wilson and Szostak, 1999).

The tertiary structures of simple aptamers tend to have purine-rich loops. These purines can participate in non-standard base pairing with the target ligand and/or within the aptamer itself. An example of this is seen with the ATP aptamer generated from random sequence RNA (Sassanfar and Szostak, 1993). From the NMR structure it was found that the purine-rich loop is highly ordered and consists of three consecutive turns (Dieckmann *et al.*, 1997; Dieckmann *et al.*, 1996; Jiang *et al.*, 1996). This, as well as other non-conventional structural elements such as a mismatched G-G base pair, forms a binding pocket that is stabilised upon the binding of ATP. A DNA aptamer generated to bind to ATP possessed no similarity in sequence or secondary structure with its RNA counterpart, however, high resolution NMR structures revealed some commonalities in the local binding site between the two structures (Huizenga and Szostak, 1995; Lin and Patel, 1997). Common to both aptamers, the ATP ligand hydrogen bonds with the minor groove face of a G residue. Also, the ATP ligand base stacks on top of a G-G base pair in both structures. It can be generalised that, in most cases, DNA or RNA must create an enclosed binding cavity in order to bind to small ligands. The case is different with protein ligands.

As target ligand molecules, proteins contain many surfaces and depressions with numerous hydrogen bond donors and acceptors. This allows for a greater potential to successfully select nucleic aptamers for proteins. To date there are many aptamers that have been selected against proteins. The first example of an aptamer for a protein was

for a single-stranded DNA for thrombin (Bock *et al.*, 1992), a glycoprotein that is instrumental in the blood-clotting cascade. The aptamer consisted of a very compact and guanine-rich 15-nucleotide sequence (dGGTTGG TGTGG TTGG), and tightly bound to thrombin with a dissociation constant in the range of 25-200 nM. From the NMR (Schultze *et al.*, 1994) and X-ray crystal structures (Mathews *et al.*, 1994), it was found to form a two-layered guanine quadruplex. That such a small motif can specifically recognise a protein is particularly impressive, although the specific interactions of the aptamer with the protein are not yet fully understood.

1.7 Guanine Quadruplex Structures

Guanine-rich DNA sequences, under certain buffer conditions, can fold to form a cyclical arrangement for four guanine bases, called a guanine-quadruplexes or guanine-quartets (Henderson *et al.*, 1987; Sen and Gilbert, 1988; Sen and Gilbert, 1990; Williamson, 1994; Williamson *et al.*, 1989). These G-quartets participate in Hoogsteen base-pairing and are stabilised by Na^+ and K^+ ions (Figure 1-8). Multiple "layers" of guanine quartets can stack upon each other for form longer guanine-quartet structures, but as with standard double-helical DNA, the G-quartet layers are slightly rotated with respect to one another in a right-handed manner (Keniry, 2000). Within the biological context they are suspected to be present at telomere ends and play a stabilising role for chromosomes (Williamson, 1994).

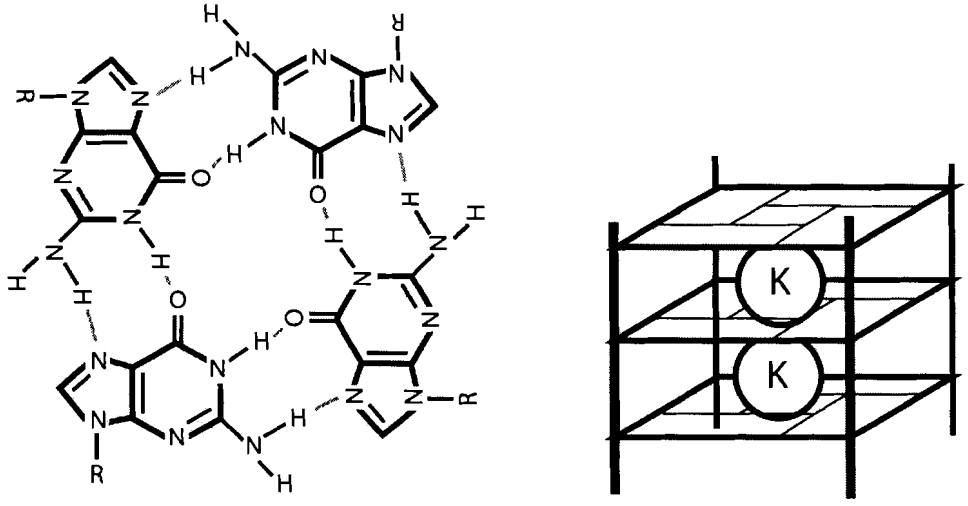


Figure 1-8 Guanine-quadruplex structures.
(Left) A cyclical arrangement of guanine residues. (Right) Monovalent cation stabilised parallel 4-stranded structure.

Depending on the location and sequence context of the guanine residues, guanine-quartets can assemble in a variety of different ways. Guanine residues on four separate strands can combine to form an intermolecular complex, shown in Figure 1-9a. Alternatively, multiple stretches of guanine residues on a given strand can combine to form guanine-quartets consisting of dimeric-like complexes Figure 1-9b. Finally, if a monomer has four separate stretches of guanine residues, these can form an intramolecular complex Figure 1-9c. Also, the overall folding topologies of intramolecular structures have been shown to vary with respect to loop positioning dictated by the orientation of the phosphate backbone, illustrated in Figure 1-10.

Guanine quadruplexes are able to discriminate between different types of metal ions based on formation. The central cavity that is formed can accommodate only metal cations with the correct size and geometry. Potassium and sodium are preferred over the other alkali earth metals because of this. The order of preference is as follows: $K^+ > Na^+$

> Rb^+ > Cs^+, and for a larger group of cations: Sr^{2+} > Ca^{2+} > Na^+ > Mg^{2+} > Li^+ (Hardin *et al.*, 1992) Interestingly, a recent study has found that Pb^{2+} ions can stabilise G-quartet formation as well (Smirnov and Shafer, 2000).

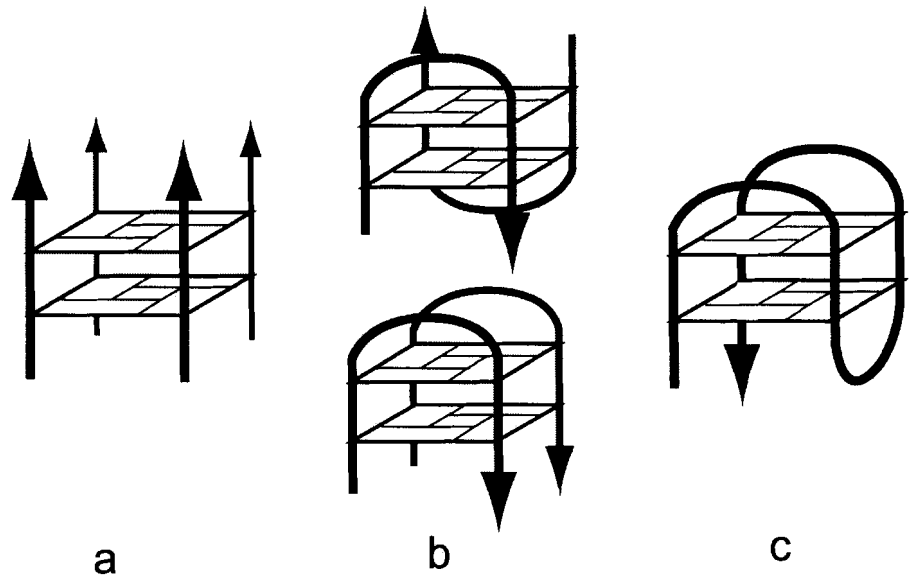


Figure 1-9 Structural polymorphisms of guanine-quartet structures.
a) Parallel structure formed from four separate strands b) Dimeric G-quartet structures formed from two strands. c) Single-stranded intramolecular G-quartet.

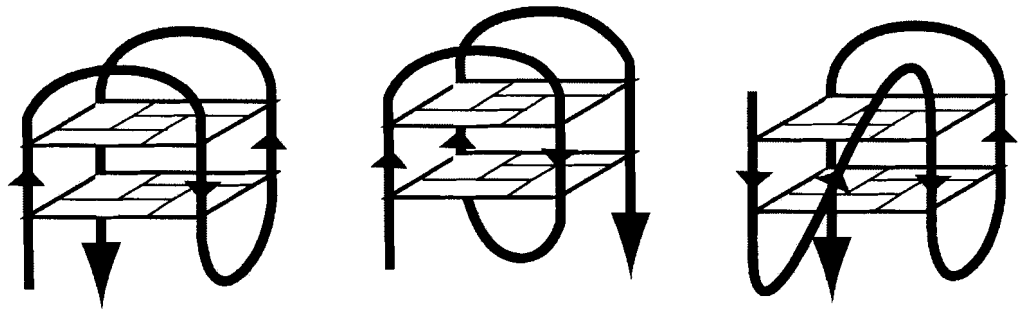


Figure 1-10 Folding topologies of structures containing intramolecular of guanine-quadruplexes. (Wellinger and Sen, 1997)

1.7.1 Guanine-Quartets in Aptamers and DNAzymes

Guanine quartet structures have been found in aptamer sequences as well as in deoxyribozymes (Li and Sen, 1996; Macaya *et al.*, 1993). The high-resolution structures of the thrombin aptamer have shown that an intramolecular structure forms consisting of two layers of guanine-quartets (Figure 1-11) (Macaya *et al.*, 1993; Schultze *et al.*, 1994; Wang *et al.*, 1993a; Wang *et al.*, 1993b). Although there is some discrepancy between the NMR and crystal structures regarding the orientation of the loops, there is agreement in the presence of a two-layered quartet in the central core. PS5.M, another G-quartet containing small-motif aptamer mentioned earlier, was found by Li and Sen to bind to N-methylmesoporphyrin (NMM, the transition-state analogue for porphyrin metalation), as well as other porphyrins (Li *et al.*, 1996; Li and Sen, 1996; Li and Sen, 1998). Proposed to contain a three-layered quartet (Figure 1-11), it is believed that the porphyrin molecule binds this 24-nucleotide aptamer by either end-stacking with the guanine quartet, or by intercalation between quartets. Although impressive at binding its target ligand ($K_D \sim 0.5$ μM), the most interesting feature of this aptamer was its catalytic ability.

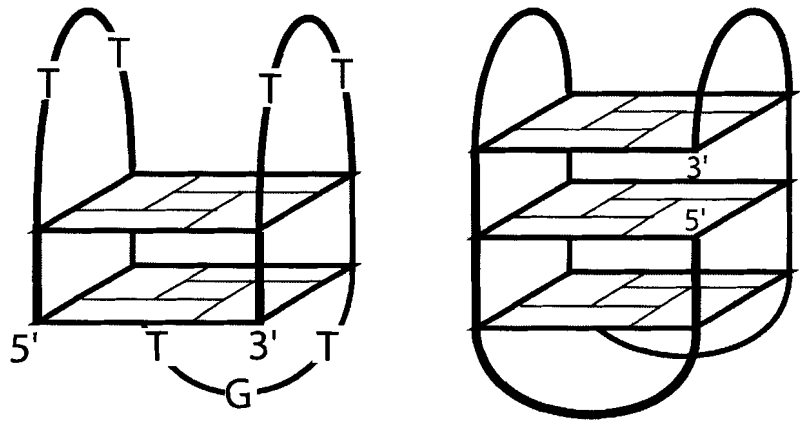


Figure 1-11 Guanine quartet- containing functional nucleic acids.
The crystal/NMR structure of the thrombin aptamer (left) and the proposed folding topology of PS5M DNAzyme (right).

PS5.M was the first example of a deoxyribozyme successfully derived using the "indirect" approach to *in vitro* selection. In the presence of K^+ and Na^+ ion, it was able to catalyse the insertion of copper and zinc ions into the center of an empty porphyrin (mesoporphyrin IX, MPIX). Thermodynamic analysis of PS5.M revealed that guanine-quadruplex provided a hydrophobic environment to bind to the substrate porphyrin, and subsequently catalyse the reaction by distortion of the porphyrin molecule in the transition state of the reaction (Geyer and Sen, 2000). Another porphyrin aptamer derived from the same selection, PS2.M was found to catalyse a peroxidation reaction in the presence of a bound heme (iron protoporphyrin IX) (Travascio *et al.*, 1998). It was proposed that axial coordination of the iron center of the heme molecule, when bound to the guanine-quartet, helped perform catalysis similar to that of the naturally occurring peroxidase enzymes (Travascio *et al.*, 1999; Travascio *et al.*, 2001).

1.8 **Repair of Thymine Dimers**

Thymine (or pyrimidine) dimers are the major lesions formed in DNA as a result of exposure to ultraviolet light. Photoexcitation of DNA results in a [2 + 2] cycloaddition reaction between two adjacent thymine (or pyrimidine) residues (Figure 1-12). This results in covalent linkages between the 5-5 and 6-6 positions of thymines to form a very stable cyclobutane ring. Another photoproduct that forms as a result of ultraviolet light exposure is the (6-4) photoproduct, shown in Figure 1-12 (Matsunaga *et al.*, 1991). The presence of these types of cross-links in DNA can be troublesome for the cell if not dealt with, as they are known to cause genetic mutations and can lead to cancers (Vink and Roza, 2001).

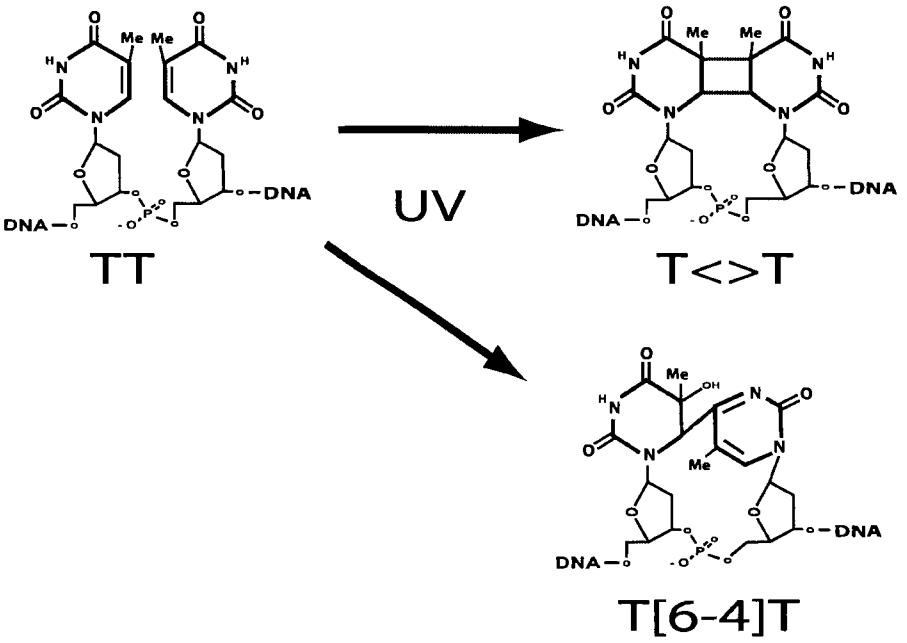


Figure 1-12 Thymine photolesions formed from UV radiation.

Luckily, nature has evolved photochemical and non-photochemical strategies to repair these photolesions. Mammalian cells use "excision repair" machinery, either in the form of "nucleotide excision repair" [reviewed in (Costa *et al.*, 2003)] or "base-excision repair" [reviewed in (Fortini *et al.*, 2003)]. Both of these processes are summarized in Figure 1-13. In the nucleotide excision repair process an "exinuclease repair complex" first binds to the DNA and removes nucleotides on both sides of the damaged DNA, including the thymine dimer lesion. This results in a gap of eight nucleotides to the 5' side of the damage and four to five nucleotides to the 3' side. Following this, DNA polymerase machinery then fills in the missing nucleotides, and the end of the DNA is ligated (sealed) to complete the process. The net result is the removal and replacement of 12-13 nucleotides.

In the "base excision repair" process only two bases, the thymine dimers, are removed and replaced. In the first step the glycosidic bond of the 5' offending thymine is removed by N-glycosylase. Second, the ribose sugar moiety and the 5' phosphate of the newly abasic site is removed by AP-endonuclease. Third, 5'endonuclease removes thymine

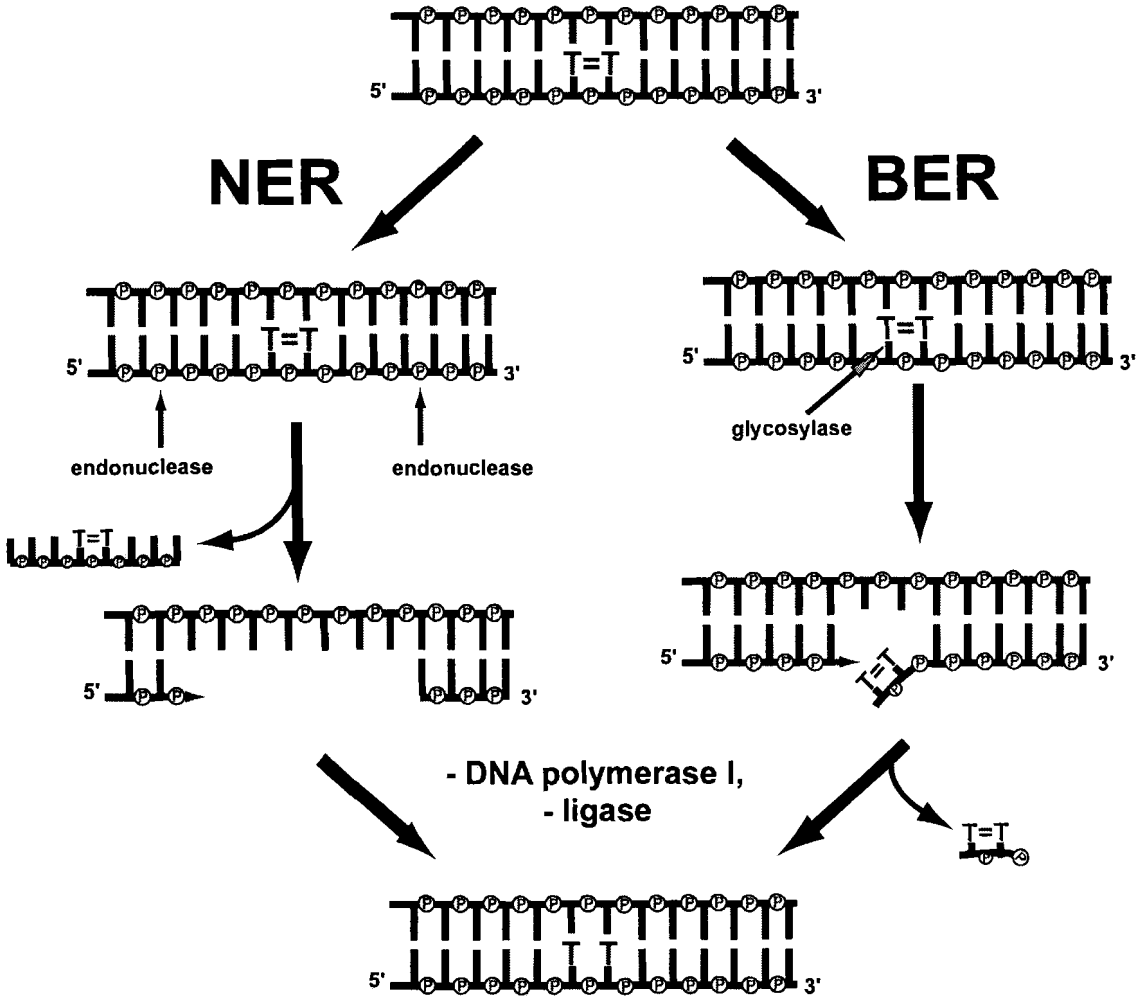


Figure 1-13 Thymine dimer repair by nucleotide excision repair (NER) and base excision repair (BER) pathways.

dimer-containing dinucleotide on the new 5' end. The last step, as seen with the NER mechanism, involves the resynthesis of the bases to complete the process. In both of the

aforementioned repair processes, the cyclobutane thymine dimer is completely removed from the strand, and replaced with new bases. However in the photochemical method of repair performed by photolyase enzymes, no bases are removed or added in a process called "photoreactivation".

1.8.1 Photoreversal of Thymine Dimers by CPD Photolyase Enzymes

The photolyase enzymes harness light of lower energy (longer wavelength) than the natural absorption of thymine dimers (>250 nm wavelength) to reactivate the dimers back to monomers. Photolyase enzymes for the repair of both cyclobutane pyrimidine dimers (CPD) and (6-4) photoproducts have been studied extensively (Carell *et al.*, 2001; Matsunaga *et al.*, 1991). The mechanisms for both types of photolyases are similar in terms of excitation and electron donation, therefore the remainder of the chapter will concern the CPD photolyases.

The cyclobutane pyrimidine dimer photolyases harness a broad spectrum of light by means of two primary, noncovalently bound chromophore cofactors, shown in Figure 1-14. Common to all photolyase enzymes is flavin adenine dinucleotide (FAD), which is primary cofactor responsible for the photoreactivation. As for the "second chromophore", photolyase enzymes fall into one of two classes based on the type of chromophore bound. The folate class of photolyases contain 5, 6-methenyl-tetrahydrofolate (MTHF), whereas those of the deazaflavin class contain 8-hydroxyl-7, 8-didemethyl-5-deazariboflavin (8-HDF) (Sancar, 2003). The main role of the second chromophore is to absorb light energy and transfer it to the flavin molecule, although the flavin itself can also absorb a certain amount of light energy as well.

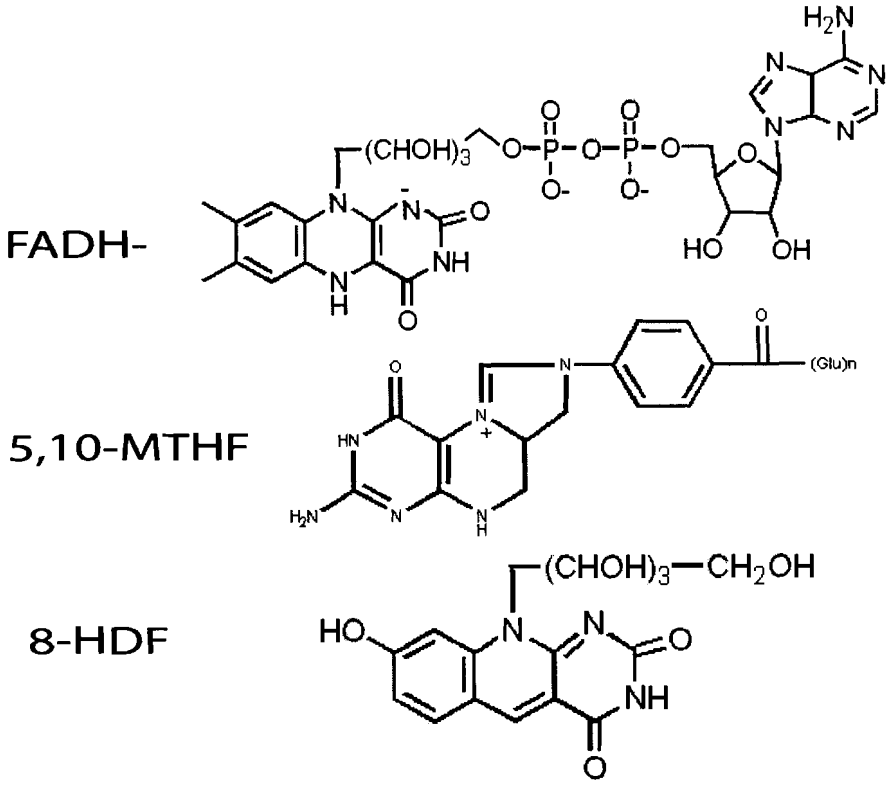


Figure 1-14 Chromophores found within Photolyase enzymes.

The photoreactivation mechanism, shown in Figure 1-15, is as follows: In the first step, a photon in the 300-500 nm wavelength region photoexcites the MTHF chromophore, which then transfers its excitation energy to the reduced form of FADH. Next, the flavin, in its fully reduced excited singlet state, donates an electron directly to the thymine dimer substrate (Payne *et al.*, 1987). Next, the 5-5 and 6-6 bonds of the dimer are destabilised, leading to the reversion to base monomers. The last step is believed to involve "back electron transfer" from one of the photoreversed thymine bases to $FADH^{\circ}$, which is restored back to $FADH^{-}$. This results in a cyclic transfer of electrons. Since there is no net loss or gain of electrons, the photoreversal mechanism is

not considered a redox reaction. The entire process is very efficient, with a quantum yield near unity.

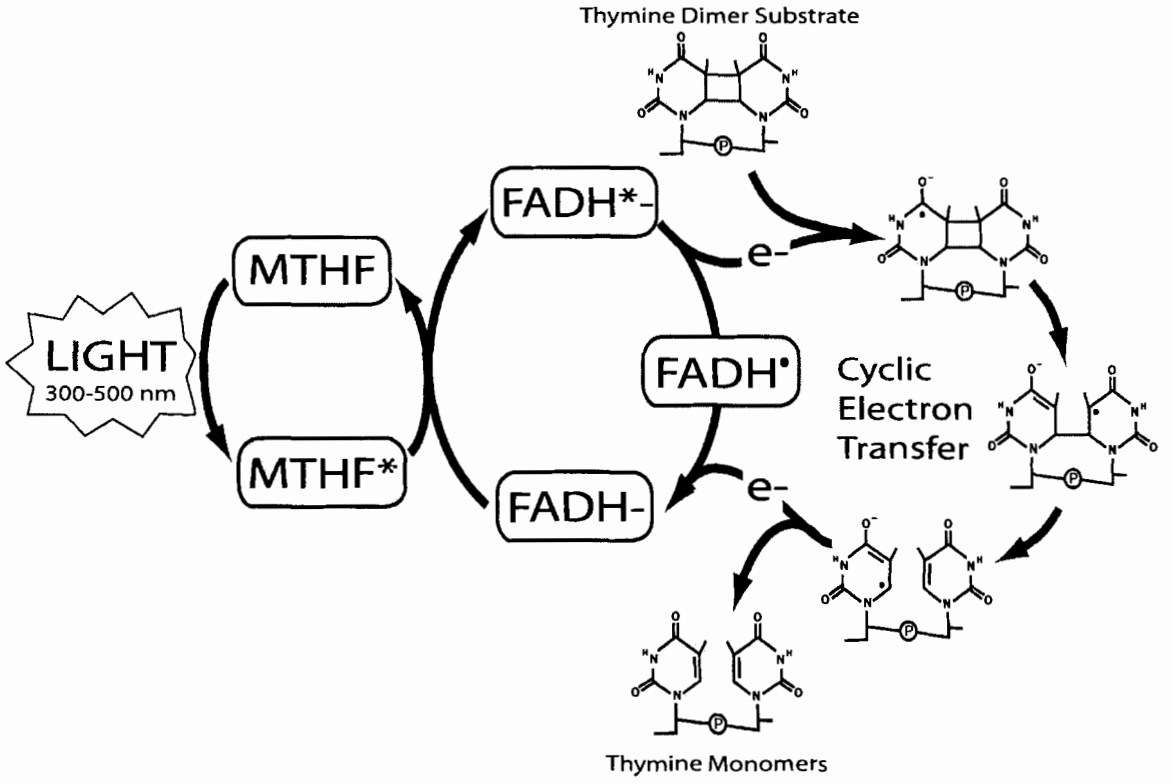


Figure 1-15 Mechanism of photoreactivation by *E.coli* photolyase.

Interestingly, Sancar and coworkers had identified a "third chromophore" within the active site of the enzyme (Kim *et al.*, 1992). Direct photoexcitation of a key tryptophan residue (Trp-277) near the active site of the enzyme was found to be sufficient for photorepair of thymine dimers, even in the absence of the FADH and MTHF cofactors. Due to the fact that the absorption spectra of tryptophan is beyond that of visible light, it was unclear whether or not this finding was relevant to its natural function.

1.8.2 Repair by Charge-Transfer Through DNA

In model studies to understand charge-transfer through double-stranded DNA (Behrens and Carell, 2003; Dandliker *et al.*, 1997; Dandliker *et al.*, 1998), it was found that thymine dimers could be photosplit by either oxidative charge transfer, or by excess electron transfer through DNA. Oxidative charge conduction through double-helical DNA, pioneered by Barton and co-workers in the 1990's, is believed to occur either by a distance-dependent "superexchange" mechanism (Marcus and Sutin, 1985), or by a distance-independent "hole hopping" mechanism (Giese, 2002b; Jortner *et al.*, 1998), wherein guanine residues act as "stepping stones" for the positive charge through the DNA [reviewed in (Giese, 2002a; Giese, 2002b)]. This is possible because guanine is the most easily oxidizable base in DNA, often resulting in damage (Steenken and Jovanovic, 1997). In these experiments, an electron hole is typically generated by the photoexcitation of a tethered oxidant, such as rhodium or ruthenium bipyridal compound. Barton and co-workers demonstrated that when thymine dimers are incorporated at various positions within double-stranded DNA, the induction of oxidative charge transfer results in repair (Dandliker *et al.*, 1997; Dandliker *et al.*, 1998). In this system, repair was accomplished by an oxidative process. Interestingly, Schuster and coworkers had found that thymine dimers could not be repaired by oxidative processes initiated by photoexcited anthraquinone (Dotse *et al.*, 2000).

Recently, Carell and co-workers have demonstrated the repair of thymine dimers by the ejection of an excess electron into a DNA double helix (Behrens and Carell, 2003). Excess electron transfer through DNA has only recently been studied by the laboratories of Carell, Rokita and Wagenknecht [a short review in (Wagenknecht, 2003)], and is

believed to occur using pyrimidines as "stepping stones" (Giese, 2002b; Giese *et al.*, 1999; Ito and Rokita, 2003; Ito and Rokita, 2004) similar to the "hole hopping" mechanism proposed for oxidative charge transfer. In the study by Carell (Behrens and Carell, 2003), a flavin molecule capped the end of a DNA hairpin containing a thymine dimer residue lacking the phosphodiester linkage between the two thymines. Photoexcitation of the flavin resulted in the ejection of an excess electron to the DNA and the repair the thymine dimer. This is presumably more physiologically relevant as repair is accomplished by a reductive mechanism in the naturally occurring photolyase enzymes.

1.8.3 Repair by Catalytic Antibodies

Schultz and co-workers generated catalytic antibodies that could photosplit thymine dimers upon irradiation with ultraviolet light (Cochran *et al.*, 1988; Jacobsen *et al.*, 1995). Using a transition-state analogue approach, antibodies were generated against thymine dimers. These "catalytic antibodies" repaired with a quantum yield of 0.08 at 300 nm wavelength light. Mechanistic studies indicated that one of the tryptophan residues near the binding site donates an electron to the thymine dimer, thus resulting in photoreversion.

1.9 Thesis Overview

My thesis covers two projects from my graduate research, both demonstrate the potential of functional DNA and are significant to the RNA World hypothesis. The early Earth atmosphere was proposed to be largely reducing, and thus lacked a stratospheric ozone layer (Miller, 1953). This would result in a high degree of ultraviolet radiation that

could be damaging to a predominantly RNA population. Is it possible there existed nucleic acids that were capable of self-repairing photolesions? Since many types of reactions were demonstrated to be catalysed by nucleic acids, we wanted to determine if DNA could catalyse a new type of reaction, that of photochemistry as seen with photolyase enzymes. Three chapters are devoted to the development and characterization of a deoxyribozyme that can photorepair thymine dimers.

The first project of my graduate studies, the *in vitro* selection for DNA aptamers for cytochrome c that could be controlled with the addition of a second ligand, hemin. Many biological electron transfer processes utilise cytochrome proteins, which contain one or many heme cofactors to transduce electrons. One example is found in the association of cytochrome c to cytochrome b_5, both containing heme cofactors. Since previous studies showed that DNA aptamers could bind to porphyrin molecules, we wished to determine if a DNA-heme complex could be an electron transfer partner with cytochrome c. This would raise the possibility of primordial nucleic acid electron transfer "cytochromes" that may have interfaced with extant cytochromes. This project is discussed in chapter 5.

Chapter 2 describes use of *in vitro* selection that yielded deoxyribozymes that were able to use light to repair thymine dimers, either in the presence or absence of a light-absorbing sensitizer. A representative the sensitizer-independent selection pool was subsequently made into *trans*-acting multiple turnover catalysts by the separation of the catalytic from the substrate domains. Subsequent characterisation of the sensitizer-independent deoxyribozyme determined that thymine dimer repair correlates to the folding of an intramolecular guanine quadruplex.

Chapter 3 continues the structural characterisation and the determination of the folding topology of the putative guanine quadruplex found within the deoxyribozyme. Deletion constructs of the enzyme, which truncated sequences from either end, showed that the 3' binding arm of the enzyme could be removed while maintaining catalytic activity. From enzyme mutant constructs, it was found that completely complementary binding arms could be made to the substrate. Cross-linking studies indicated that the thymine dimer was situated close to the guanine residues of the guanine quadruplex. Using a combination of enzymatic and chemical probing, a new model for the overall folding is presented.

Chapter 4 addresses the substrate specificity of the deoxyribozyme as well as clues to its photochemical mechanism. It was found that the DNAzyme could not effectively repair RNA-containing substrates containing uracil dimers. The efficiency of repair was greatly affected by type of the sugar moiety present on the pyrimidine dimer, showing highest preference to deoxyribose sugars. Evidence is presented that indicates that electron donation from a photoexcited excited guanine is thermodynamically feasible and that a fast rate of back electron transfer to the guanine quadruplex may occur following thymine dimer repair.

Chapter 5 describes the *in vitro* selection and characterisation of a DNA aptamer that simultaneously binds to the electron transfer protein cytochrome c, and the iron-containing porphyrin, hemin. Chemical probing studies indicate the presence of guanine quadruplex in the structure with complementary stems. Spectroscopic experiments indicate that the DNA aptamer affects the local environment of the heme in cytochrome c

in a manner similar of that to its natural protein-binding partner. Hydroxyl radical footprinting studies revealed the binding site of the cytochrome c onto the DNA.

CHAPTER 2

A PHOTOLYASE DEOXYRIBOZYME THAT HARNESSES LIGHT TO REPAIR DNA DAMAGE

2.1 Introduction

The RNA World Hypothesis (Gilbert, 1986) postulates that RNA or RNA-like polymers, capable of genetic as well as catalytic function, may have constituted primitive "life" in the course of evolution. Currently, *in vitro* selection (Ellington and Szostak, 1990; Tuerk and Gold, 1990) experiments from random sequence DNA and RNA libraries permit the identification of novel catalytic activities for nucleic acids, in support of the RNA World Hypothesis. To date, such selections have indicated a substantially broader range of catalytic ability for RNA and DNA than found in naturally occurring ribozymes, as discussed in the previous chapter (Jaschke, 2001). It would be of interest to determine whether reactions that utilize light energy could be catalyzed by nucleic acid enzymes.

The enzyme photolyase repairs thymine dimers by using light of longer wavelengths (300 to 500 nm) than the natural absorption of thymine dimers (<250 nm) to reactivate the dimers back to monomers. The cyclobutane (CPD) photolyases harness a broad spectrum of light, using a number of chromophores such as methenyl-tetrahydrofolate (MTHF), flavin nucleotides, and tryptophan side-chains (Sancar, 2003). Photoexcitation culminates in electron donation from the excited-state flavin directly to

the thymine dimer, leading to destabilization of the 5-5 and 6-6 bonds of the dimer, and thus, reversion to base monomers. Interestingly, studies have shown that photoexcitation of a tryptophan residue in the *Escherichia coli* enzyme's active site is able to provide a significant photoreactivation, even in the absence of the FADH and MTHF cofactors (Kim *et al.*, 1992).

To investigate whether a photolyase deoxyribozyme could be selected for, *in vitro* selection was carried out from a random-sequence DNA library. Our library (seven copies of 10^{14} different sequences, each containing a 40-nucleotide random sequence flanked by 35- and 20-nucleotide primer-binding sequences) incorporated the following design feature (Figure 2-1): the 35-nucleotide 5' element ("TDP") consisted of 20- and 15-mer oligonucleotides covalently linked through a thymine dimer, but lacking a connecting phosphodiester linkage. This TDP substrate was incorporated by PCR into the random-sequence DNA library, providing a means for selecting out catalytically active molecules from the pool, since DNA molecules capable of efficient self-repair decreased in length from 95 to 80 nucleotides, making size-based purification possible. Since tryptophan had been shown to be an adequate sensitizer for photoreactivation, the indole-containing compound, serotonin (whose absorbance spectrum is red-shifted relative to that of tryptophan), was chosen as a cofactor for our selection (Figure 2-2). Early model studies of photoreactivation (Helene and Charlier, 1977) had demonstrated the ability of photoexcited serotonin to repair thymine dimers in double-stranded DNA. Initially, my goal was to use *in vitro* selection to generate a DNAzyme that could photorepair thymine dimers with the aid of serotonin. It was hoped that a DNA sequence could be selected that could fold in such a way as to position a serotonin sensitizer sufficiently close to the

thymine dimer. Thus photoexcitation could result in electron donation to the dimer and result in photorepair.

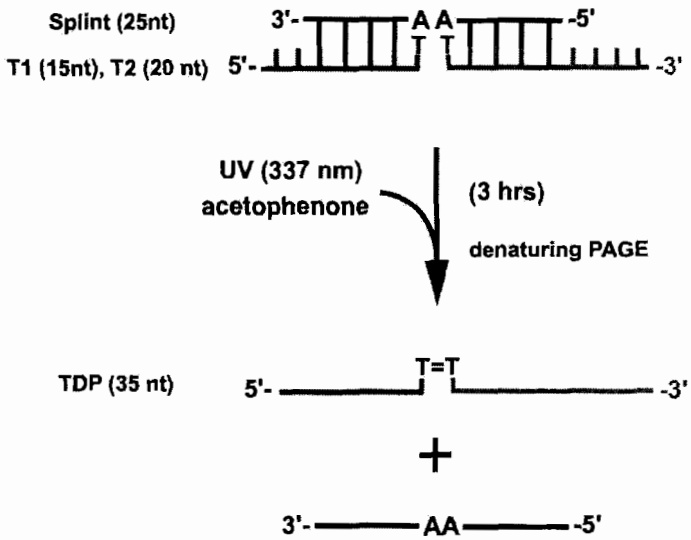


Figure 2-1 Thymine dimer substrate design and synthesis.
Synthesis of a thymine dimer-containing DNA oligonucleotide (TDP) lacking a key internal phosphodiester. A 25-nt complementary "splint" was used to align a 5'-^{32}P-labelled 15-nt oligonucleotide (T1) and a 20-nt oligonucleotide (T2), lacking 3' and 5' terminal phosphate groups, respectively.

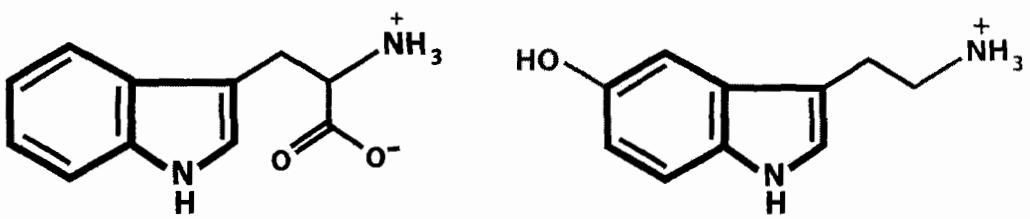


Figure 2-2 Chemical structures of indole-containing compounds tryptophan (left) and serotonin (right).

2.2 Materials and Methods

2.2.1 DNA Oligonucleotides

DNA oligomers were obtained from University of British Columbia NAPS Unit, and size-purified on denaturing polyacrylamide gels. The starting library of random DNA sequences was obtained from University of Calgary Core DNA Services, and had the sequence 5'-AGGAT CTACA TGTAT TGTGT GCGTA CGAGT ATATG G N_{40} GTCTC AATCG GTCTG TATC-3'.

2.2.2 Synthesis of Thymine Dimer-containing Oligonucleotide

For TDP formation, unlabeled T2 and 5'-^{32}P-labeled T1 oligonucleotides were annealed at 400 μM concentrations to a complementary splint oligonucleotide, in buffered 40 mM $MgCl_2$. Solutions were degassed by freezing and thawing under vacuum three times. 5 mM of the triplet sensitizer acetophenone (shown specifically to favour the formation of cyclobutane thymine dimers (Wang, 1976)) and 5% acetone. The DNA was irradiated for 3 hours in a quartz cuvette with a Photon Research Associates LN-1000 nitrogen laser with spectral output at 337 nm and a pulse rate of 7 Hz. The TDP construct was now size-purified on 12% denaturing polyacrylamide gels and tested for the presence of the cyclobutane (rather than the (6-4)) thymine dimer by direct photoreversal experiments at 254 nm. At this wavelength almost quantitative yield of the regenerated T1 and T2 strands was obtained.

2.2.3 DNA Library Preparation and Selection Procedure

The TDP primer was incorporated into the 5'-end of the random sequence library using PCR. Owing to the inability of Taq polymerase to extend past the thymine dimer

on the complementary strand during extension (Wellinger and Thoma, 1996), the two resulting strands had different sizes and could be cleanly separated using denaturing polyacrylamide gel electrophoresis. For the "negative" selection step, 100 μL of 1 μM of purified DNA was folded in 20 mM sodium phosphate, pH 7.0, and 40 mM NaCl, and irradiated on a Fotodyne transilluminator (with a 300 nm output at 3.4×10^{-9} einsteins min^{-1} illumination, where one einstein unit is defined as one mole of photons) with an intervening polystyrene filter, such that the light transmitted had wavelengths of >300 nm. The distance of the irradadiated solution from the light source filter was approximately 1 mm based on the thickness of the bottom well of the ELISA plate. The irradiated pool was then gel purified to retrieve unmodified (95 nt) single-strands, which were then subjected to "positive" selection by irradiation in the presence of 10 μM serotonin (in the first five rounds of selection, a PCR amplification step was carried out between the negative and positive selection steps). Stringency was enhanced in the course of the selection by decreasing the time of irradiation, from 2 hours (round 1) down to 10 seconds by round 25.

2.2.4 Cloning and Sequencing

Following round 25, the pools were cloned, using standard protocols. The amplified DNA was gel purified in an 8% non-denaturing gel and ligated into the pCR2.1 plasmid using a TA cloning kit (Invitrogen). The plasmid was then transformed into INVαF' (Invitrogen) competent *E. coli* cells. Thirty-eight recombinant clones were picked for analysis and sequenced using the Thermo-sequenase dideoxy kit (USB Amersham).

2.2.5 Kinetic Analysis

The kinetic properties of UV1C (5'- GGAGA ACGCG AGGCA AGGCT GGGAG AAATG TGGAT CACGA TT) were measured in 20 mM sodium phosphate, pH 7, 240 mM NaCl, at room temperature. 100 μL solutions containing 5' ^{32}P-labeled DNA were irradiated in a quartz cuvette and cell holder using an Ushio Xenon short arc lamp, with wavelengths isolated with a SLM MC200 model monochromator, and bandwidth set at 4 nm. Time points were taken, and the DNA, following separation on 12% denaturing gels, was quantified using Molecular Dynamics ImageQuant software. Single turnover reactions were carried out with 2.1 μM UV1C and 20 nM TDP substrate, in 20 mM sodium phosphate, pH 7.0, with 240 mM NaCl. Multiple turnover kinetics were irradiated at > 300 nm with a Fotodyne transilluminator at a fluence rate of 3.4×10^{-9} einsteins min^{-1} illumination with 20 nM UV1C and varying concentrations of TDP substrate (0.1 μM to 5.0 μM), in 20 mM sodium phosphate buffer, pH 7.0, with 240 mM NaCl. Background rates of reaction were measured at longer timepoints (1-2 hours) for reproducible results. Experiments were performed in triplicate and error values were calculated from the fit of the curves using GraphPad Prism4 software.

Samples were irradiated in a quartz cuvette with a 1 cm path-length at different wavelengths using a 4 nm bandwidth. The photon flux of the lamp source at given wavelengths was determined by standard ferrioxalate actinometry (Calvert and Pitts, 1966), and used for light intensity corrections and quantum yield calculations. Quantum yields of thymine dimer repair were calculated using the slope obtained from a least squares fit of the initial rates of reaction obtained from quantitation of gel data. Quantum

yield calculations for the formation of product were calculated using the equation: Φ = (# moles of product formed) / (# moles of photons absorbed) (Calvert and Pitts, 1966).

2.2.6 Spectrophotometric Analysis

The difference spectra were recorded in a Cary dual beam UV spectrophotometer, on DNA folded in 50 mM Tris-Cl, pH 8.0, supplemented with either 200 mM NaCl or 200 mM LiCl. The absorbance spectrum of DNA in lithium buffer was subtracted from the spectrum of DNA in sodium buffer, and signal averaged over 10 scans, to obtain the final difference spectrum. For the rate enhancement data, reaction mixtures contained 4 μM UV1C and 40 nM TDP substrate in 20 mM sodium phosphate, pH 7.0, plus 240 mM NaCl. Initial rates of reaction were divided by the background rates (measured on 40 nM TDP in the presence of 4 μM of an unrelated DNA oligomer).

2.3 Results and Discussion

2.3.1 Selection Progress and the UV1 DNAzyme

Figure 2-3 summarizes the *in vitro* selection cycle. In a negative selection step, the TDP-containing DNA pool was irradiated at >300 nm (Step 4), in the absence of serotonin (Step 8); DNA molecules surviving this step intact were subjected to positive selection in the presence of 10 μM serotonin. In the initial rounds of selection, the DNA pool was irradiated at >300 nm for two hours in the presence of serotonin. Figure 2-4 shows the polyacrylamide gel from round one indicating negligible repair by the DNA pool. To increase the stringency of the selection the irradiation time was reduced to 1 hour following round 3. Curiously, after 5 rounds of selection, a pool of DNA molecules

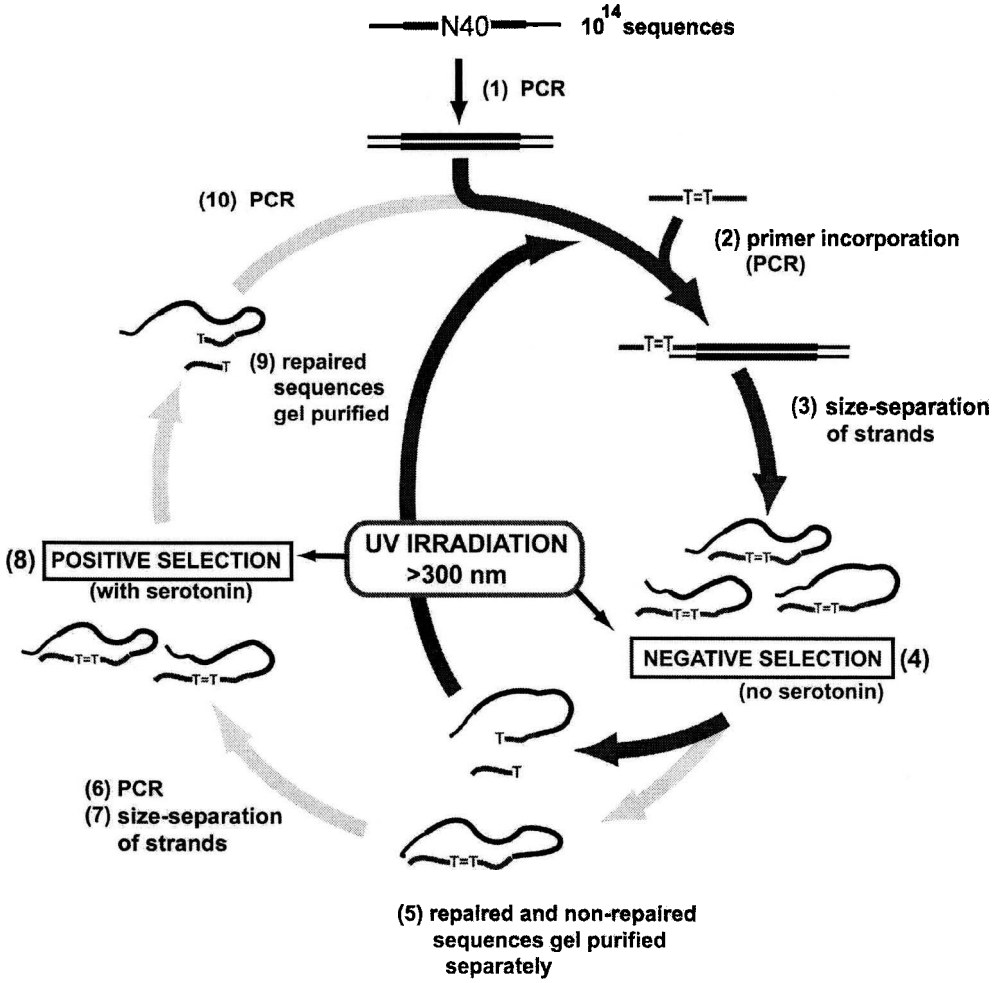


Figure 2-3 The *in vitro* selection cycle.

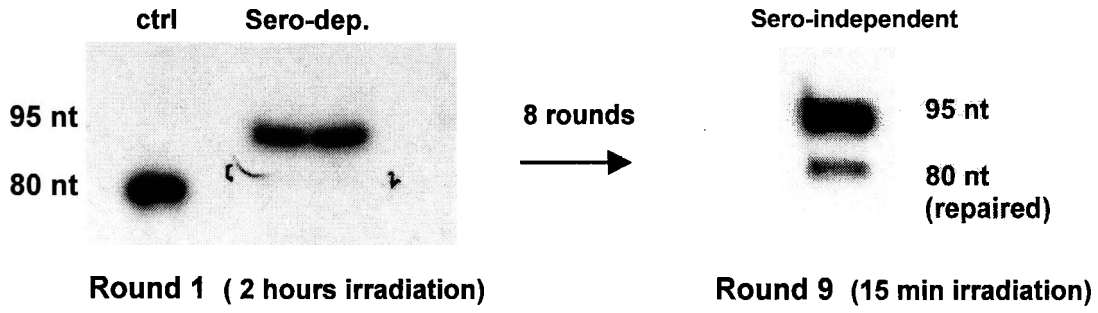


Figure 2-4 Denaturing polycacrylamide gels showing repair efficiency of Rounds 1 and 9. (Left) Gel of round 1 with 80 nucleotide control lane, and two lanes (replicates) of the serotonin-dependent pool after two hours irradiation time. The area between the brackets were cut out from the gel and DNA eluted for the next round. (Right) Round 9 gel of the serotonin-independent pool showing repair after 15 minutes irradiation.

emerged that robustly self-repaired in the negative selection step; this was surprising because neither DNA nor thymine dimers absorb significantly at >300 nm. The serotonin-dependent and -independent pools were now separately taken through 20 additional selection rounds, and each pool was cloned and sequenced. From rounds 6 to 21, the irradiation time was gradually decreased from 30 minutes down to 10 seconds. Figure 2-4 shows the enhanced efficiency of repair in round nine for the serotonin-independent pool after fifteen minutes irradiation. By round 21, both selection pools displayed a great enhancement of photorepair. To further optimise the catalytic efficiency, a mutagenic PCR step was introduced to provide genetic variability into the pool, and the selection continued for a further four rounds. Following this, both pools were cloned and individually sequenced. The serotonin-independent pool, the subject of this chapter, yielded only one dominant sequence, UV1A (whose N_{40} sequence is Figure 2-5). The purine-rich sequence of UV1A was distinct from two major sequences

I25-1	AGAACGCGAGGCAAGGCTGGGAGAAATGTGGATCACGATT	
I25-7	AGAACGCGAGGCAAGGCTGGGAGAAATGTGGATCACGATT	
I25-8	AGAACGCGAGGCAAGGCTGGGAGAAATGTGGATCACGATT	
I25-9	AGAACGCGAGGCAAGGCTGGGAGAAATGTGGATCACGATT	
I25-11	AGAACGCGAGGCAAGGCTGGGAGAAATGTGGATCACGATT	
I25-14	AGAACGCGAGGCAAGGCTGGGAGAAATGTGGATCACGATT	UV1A
I25-18	AGAACGCGAGGCAAGGCTGGGAGAAATGTGGATCACGATT	
I25-19	AGAACGCGAGGCAAGGCTGGGAGAAATGTGGATCACGATT	
I25-20	AGAACGCGAGGCAAGGCTGGGAGAAATGTGGATCACGATT	
I25-2	AGAACGCGAGGCAAGGCTGGGAGAAATGTGGATCACGGTT	
I25-12	AGAACGCAAGGCAAGGCTGGGAGAAATGTGGATCACGATT	
I25-13	AGAACGCAAGGCAAGGCTGGGAGAAATGTGGATCACGATT	
I21-3	ACGTGGTGAAGGCA-GGGTGG AATATGATCCTGGACATACC	
I25-4	ACGTGGTGAAGGCA-GGGTGG AATATGATCCTGGACATACC	UV2A
I25-5	ACGTGGTGAAGGCA-GGGTGG AATATGATCCTGGACATACC	
I25-10	ACGTGGTGAAGGCA-GGGTGG AATATGATCCTGGACATACC	

Figure 2-5 Sequences of the random region from the serotonin-independent selection.

obtained from the serotonin dependent pool, shown in Figure 2-6.

```
S25-1       TAAGTCGGATGGAGGGATCCGTTGGCACATGTAGTCACGT
S25-7       TAAGTCGGATGGAGGGATCCGTTGGCACATGTAGTCACGT    Sero1
S25-12      TAAGTCGGATGGAGGGATCCGTTGGCACATGTAGTCACGT

S25-3       AGCACAGTCGCAAGACGATATGCAGGAACTTGGACAGCCG
S25-9       AGCACAGTCGCAAGACGATATGCAGGAACTTGGACAGCCG    Sero2
S25-14      AGCACAGTCGCAAGACGATATGCAGGAACTTGGACAGCCG
S25-8       AGCACAGTCGCATGACGATATGCAGGGACTTGGACAGCCC

S25-20      TAGGCGACAGCTGCTTGG-CGGGAATGAAGATCAGGATACTC
S25-11       AGGCGACAGCTGCGTGG-CGGGAATGAAGATCAGGATACAC  Sero3
S25-16     AATGGC-TCGGGACGGTGGTCGGGAAGGTAGATATGG-TACA
```

Figure 2-6 Sequences of the random region from the serotonin-dependent selection.

Preliminary characterization of UV1 and UV2 sequences showed that it did indeed self-repair at an accelerated rate, relative to controls, with light of wavelength >300 nm (shown in Figure 2-7). Initial rates were 0.22 ± 0.02 min^{-1}, 0.09 ± 0.01 min^{-1}, and $3.2 \times 10^{-4} \pm 0.3 \times 10^{-4}$ min^{-1} for UV1, UV2 and round 0, respectively. Due to its faster catalytic rate, the UV1 sequence was chosen for further characterization.

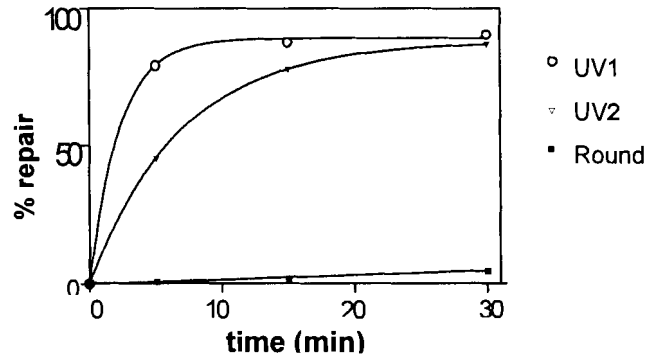


Figure 2-7 Rates of self-repair of UV1 and UV2 DNAzymes.
Initial rates of repair for UV1 = 0.22 ± 0.02 min^{-1}, UV2 = 0.09 ± 0.01 min^{-1} and Round 0 = $3.0 \times 10^{-4} \pm 0.3 \times 10^{-4}$ min^{-1} at a photon flux of 3.4×10^{-9} einsteins min^{-1}.

Chemical probing analysis was carried out with dimethyl sulfate (DMS) and diethylpyrocarbonate (DEPC) on the UV1 DNAzyme. Dimethyl sulfate preferentially methylates the N7 position of guanines, accessible from the major groove within Watson-Crick double helices (Maxam and Gilbert, 1977). Dimethyl sulfate -reactivity provides information about the participation of specific guanines in guanine-quartet formation (guanines in G-quartets are not methylated by dimethyl sulfate) (Sen and Gilbert, 1988). Protection is not conclusive evidence on its own, since there could be other non-quadruplex hydrogen-binding interactions with the N7 position. Normally, additional experiments are needed for confirmation, for example, monovalent cation-dependence and spectroscopy. Diethylpyrocarbonate reacts with adenine N7 positions, but preferentially with adenines that are present in either single-stranded, unstructured, or solvent-exposed regions of DNA secondary/tertiary structure.

The methylation protection data with dimethyl sulfate, shown in Figure 2-8, reveals stretches of guanines that are partially protected from methylation (shown with arrows). Partial protection of guanines could be interpreted as a "loosely" formed guanine-quadruplex, which may allow for intermittent N7 modification by dimethyl sulfate. Since many of the guanines in the UV1 sequence occur in doublets, it is possible that a two-layered quartet forms, which are as stable as three or more guanine-quartet stacks. Cation dependence and spectroscopic experiments will be discussed later in this chapter, as well as their significance.

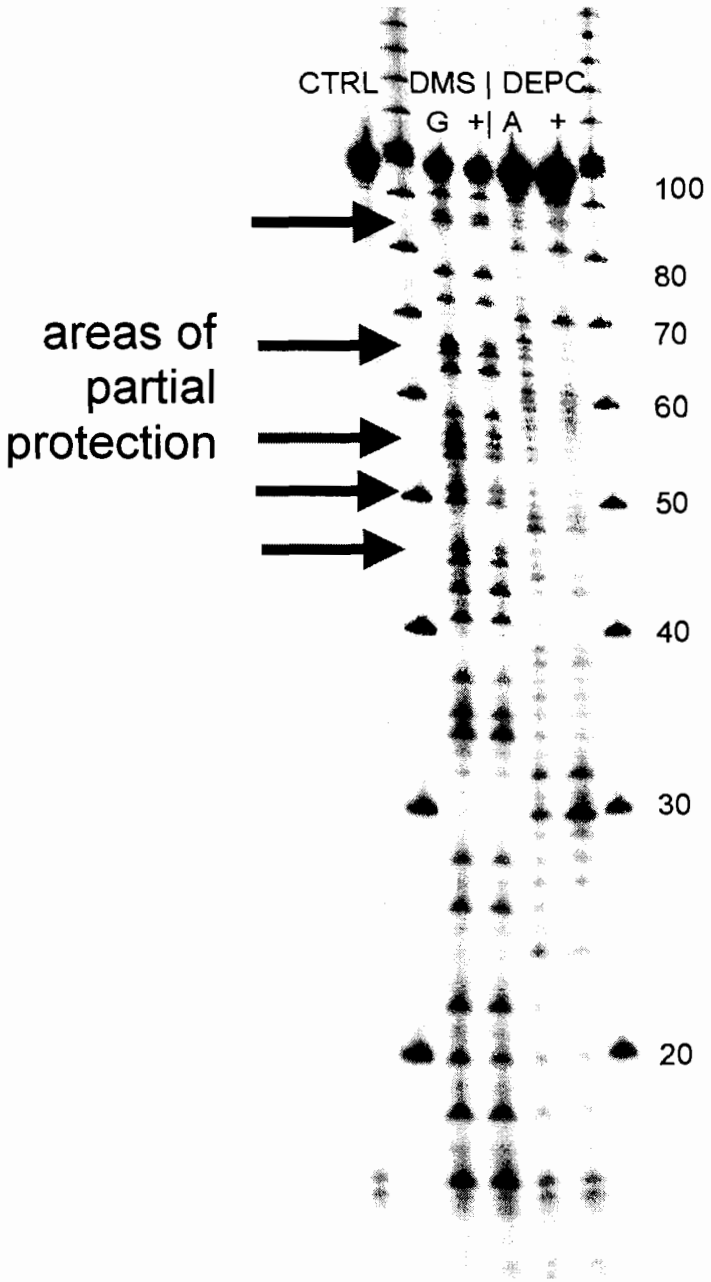


Figure 2-8 DMS and DEPC chemical probing of UV1 sequence.
Arrows indicate areas of partial protection from dimethyl sulfate methylation of the folded UV1 DNA ("+" lane - Left) with respect to the unfolded DNA ("G" lane): DEPC modification of folded DNA ("+" lane - Right) does not indicate a significant changes in protection with respect to unfolded DNA ("A" lane) with exception to hyper-reactivity to adenine 28.

2.3.2 The Creation of a Multiple Turnover DNAzyme

Essentially, the UV1 sequence was a 95-nucleotide self-repairing DNAzyme that used light to perform single-turnover catalysis. I wished to determine whether UV1 could now be engineered into a multiple turnover enzyme to act on an external substrate. To test whether a "substrate" component within the UV1 sequence could be separated from a "catalytic" component, the full-length DNAzyme was separated from TDP (the substrate sequence used in PCR) to create UV1B, and further separated from the former 3' constant region (LT2) to create UV1C (42 nucleotides, incorporating the 40 formerly random nucleotides), shown Figure 2-9. UV1 was thus renamed UV1A. Both sequences were found to be fully catalytic with repair rates of 0.10 ± 0.02 min^{-1} and 0.11 ± 0.01 min^{-1} for UV1B and UV1C, respectively. The photoreversal rates were 366-fold faster than the background rate for the TDP substrate alone, shown in Figure 2-10. Also, as an

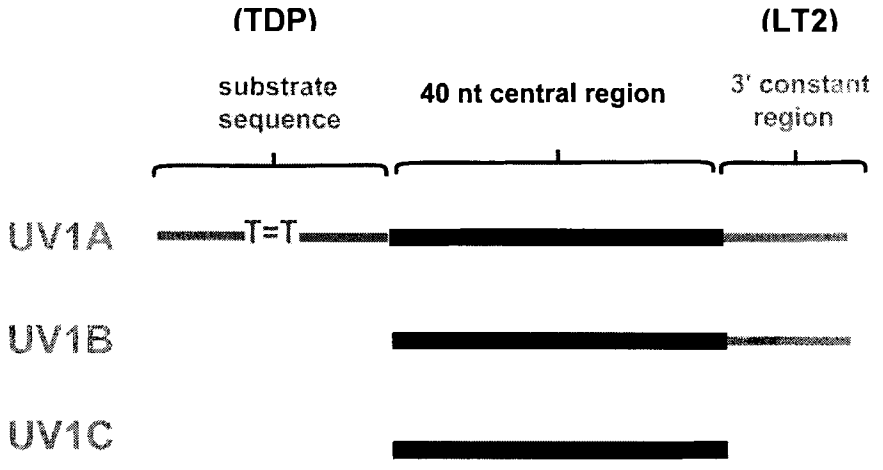


Figure 2-9 Constructs used for the engineering of a *trans*-acting DNAzyme.

added control to test whether the presence of the LT2 sequence had any effect on UV1C, the LT2 sequence was added to UV1C and irradiated in the presence of TDP substrate, shown in Figure 2-10. It was observed that the LT2 sequence did not have a significant effect on the repair rate of UV1C (0.10 ± 0.02 min^{-1}). All subsequent analysis was thus carried out with the 42-nucleotide UV1C deoxyribozyme.

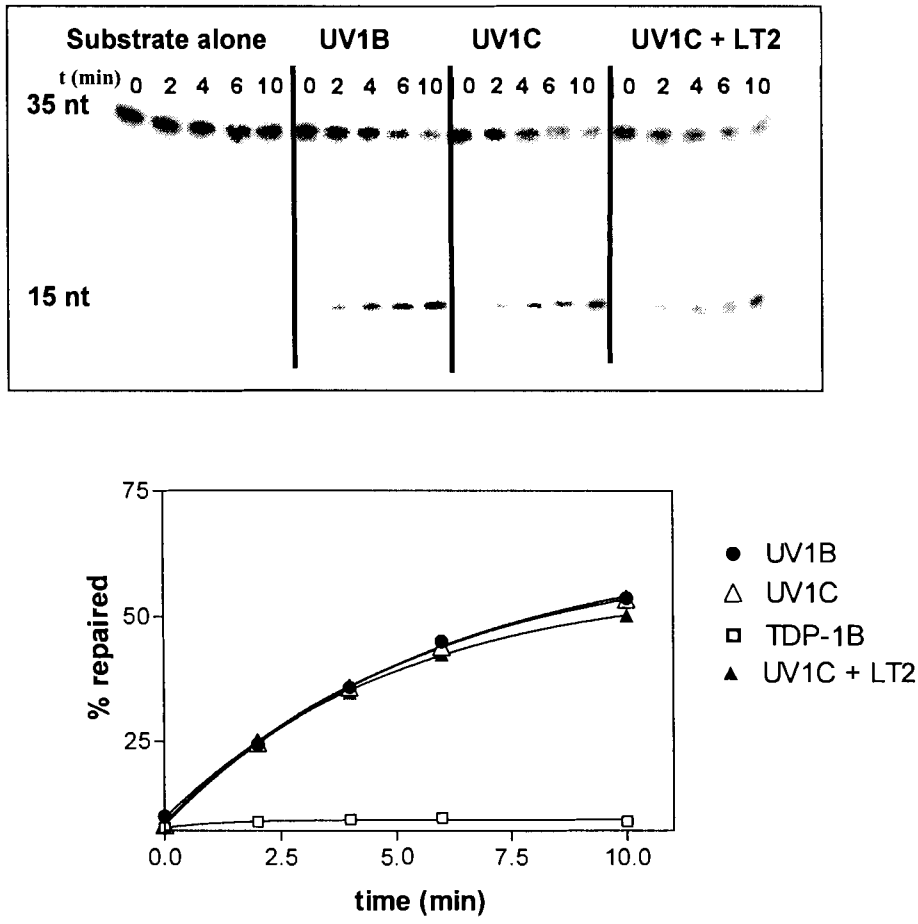


Figure 2-10 Repair of external substrate (TDP) by separation of the catalytic region from substrate region. (Top) 12 % Denaturing polyacrylamide gel showing photorepair by UV1B and UV1C. (Bottom) Plot of percent repair versus time for constructs used. Rates of repair for both UV1B and UV1C were 0.11 ± 0.02 min^{-1}. Repair rate for the UV1C sequence with LT2 oligomer was and 0.10 ± 0.01 min^{-1}. The TDP background reaction for photoreversal was $3.0 \pm 0.6 \times 10^{-4}$ min^{-1}.

Multiple turnover kinetic analysis (Figure 2-11) yielded a k_{cat} value of 4.5 min^{-1} and K_M of 0.58 ± 0.14 μM at 3.4 x 10^{-9} einsteins min^{-1} illumination (the uncatalyzed photoreactivation rate of TDP was 0.00018 min^{-1}). Both of these parameters compare favourably with those of a catalytic antibody for thymine dimer photoreactivation (1.2 min^{-1} and 6.5 μM at 1.26 x 10^{-7} einsteins min^{-1}) (Cochran *et al.*, 1988; Jacobsen *et al.*, 1995); although poorer than *E. coli* photolyase itself (Kim and Sancar, 1991).

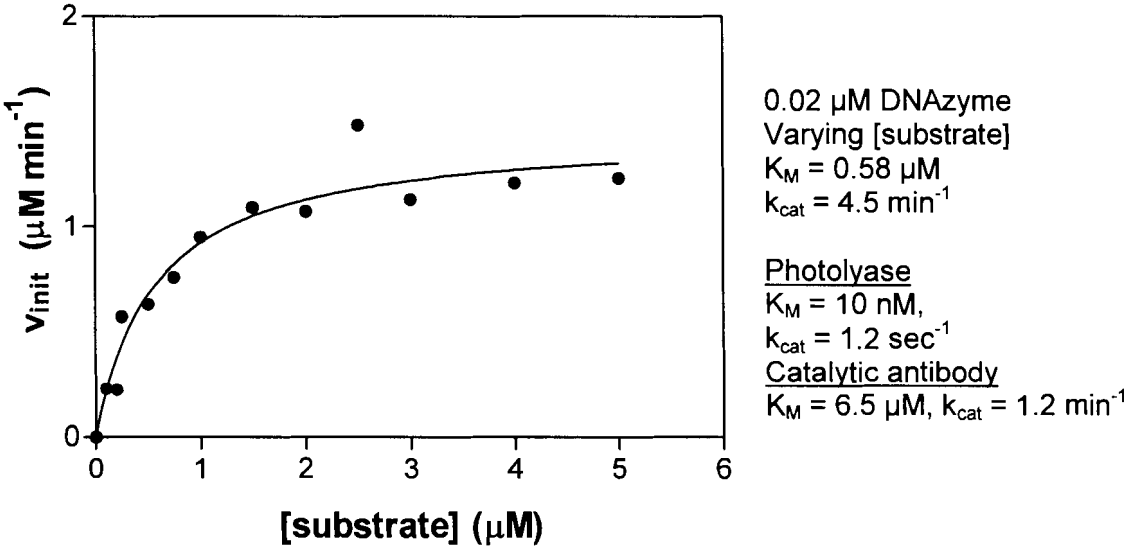


Figure 2-11 Multiple turnover kinetics of UV1C.

2.3.3 Wavelength Dependence of UV1C

Figure 2-12 shows the action spectrum of UV1C, under single turnover conditions, within the 250-320 nm spectral range (normalized for constant light intensity). 0.02 μM TDP was irradiated in the presence of an excess (2.1 μM) of (a) UV1C; (b) the 25-nt splint (Figure 2-12) that renders TDP double-stranded; and, (c) an unrelated, 42-nucleotide, single-stranded DNA. While the splint contributed to

somewhat enhanced TDP repair, especially between 250-270 nm, UV1C caused significantly higher reactivation rates, across the spectrum, with a notable "shoulder" at >275 nm.

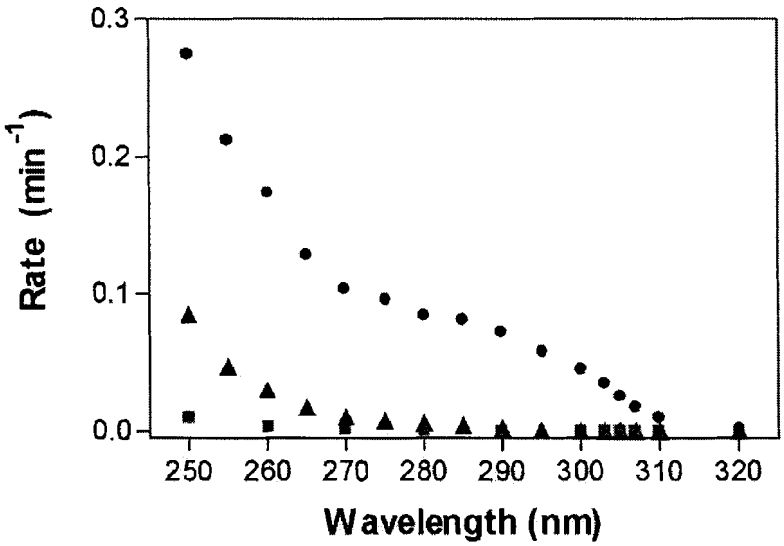


Figure 2-12 Wavelength dependence for the catalytic activity of the UV1C deoxyribozyme. Action spectrum for photoreactivation of 20 nM TDP substrate by excess (2.1 μM) UV1C (circles); 2.1 μM of the 25-nt complementary splint oligomer (triangles); and, 2.0 μM of an unrelated 42 nt single-stranded DNA (squares). All measurements were made under single turnover conditions, and all rates are shown corrected for variations in light intensity at the different wavelengths.

Figure 2-13 shows Φ, or quantum yield, for TDP reactivation in the presence of excess UV1C; and, of excess complement (splint) DNA. The UV1C quantum yield of 0.05, although low compared to those of CPD photolyases (~0.7) (Kim and Sancar, 1991) was of the order of that for (6-4) photolyases (~0.1) (Hitomi *et al.*, 1997). The heightened efficiency of UV1C-dependent reactivation between 280 and 320 nm is particularly evident.

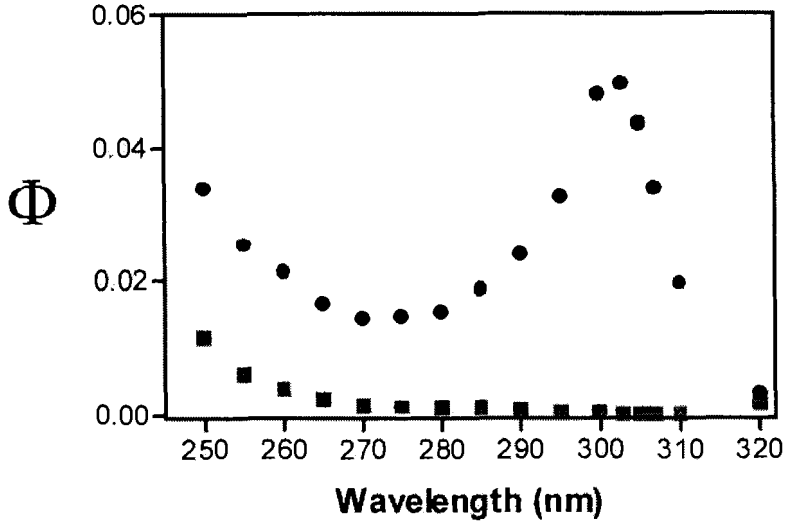


Figure 2-13 Quantum yield for the catalytic activity of the UV1C deoxyribozyme. Quantum yield (Φ) of photoreactivation by UV1C is shown in circles, and by the complementary splint oligomer in squares.

Figure 2-14 plots the ratio of the quantum yields reported in Figure 2-13 (equivalent, also, to $k_{UV1C\text{-}catalyzed}/k_{uncatalyzed}$, under single-turnover conditions). Notably, the peak of UV1C catalytic activity, at 305 nm, is a significantly longer wavelength than the absorbance maximum of folded DNA (shown in red) and also far from the absorption (<250 nm) of thymine dimers themselves. Such a property, of utilization of near ultraviolet light by UV1C, is reminiscent of protein photolyase enzymes. That UVIC was in fact catalysing photoreactivation (and not a nuclease or glycosidase activity) was verified by the successful reutilization of the UV1C- cleaved (using 303-307 nm light)

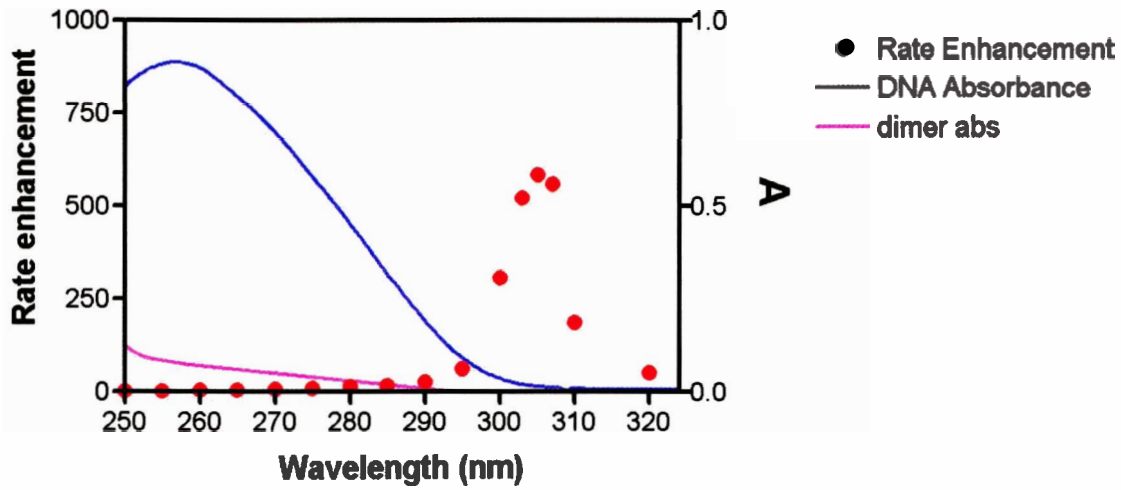


Figure 2-14 Catalytic rate enhancement for photoreactivation by the UV1C deoxyribozyme. The photoreactivation rate enhancement (red circles) by UV1C, over that by an unrelated, control DNA oligomer. The absorption spectrum of the DNAzyme-substrate complex (shown as a blue line) helps to highlight the red shift in the spectral region for optimal rate enhancement relative to the absorption maximum of DNA. Thymine dimers alone have low absorbance in the UV range (purple line).

T1 product to regenerate the active substrate TDP shown in Figure 2-15 (using the protocol shown in Figure 2-1).

To explore whether the catalytic role of UV1C was primarily to orient the TDP thymine dimer to a conformation more amenable to photoreactivation, a number of model oligonucleotide constructs were synthesized and their reactivation rates were measured in the absence of UV1C. TDP was base-paired to (a) a perfectly complementary splint; (b) a complementary splint as in (a), but lacking the two adenines expected to base pair

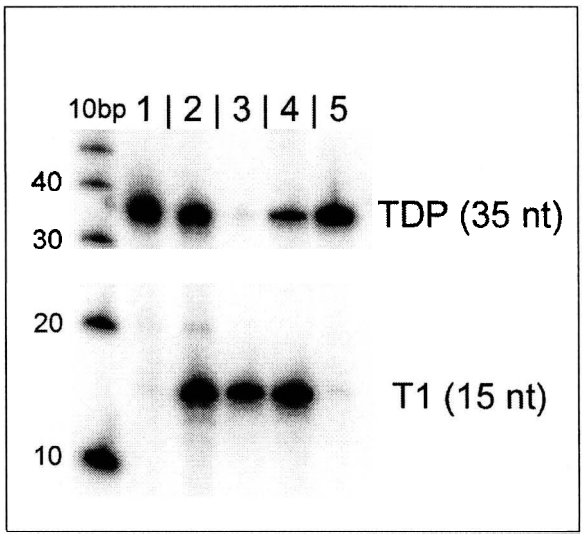


Figure 2-15 Re-formation of TDP substrate from catalytic products.
Lane 1, 5' ^{32}P- labelled TDP substrate (35 nt). Lane 2, formation of T1 product (15 nt) after enzymatic photoreactivation at wavelengths >300 nm in the presence of 10 μ M UV1C. Lane 3: Size-purified T1 product from lane 2. Lane 4, re-formation of TDP substrate from T1 product by irradiation in the presence of T2 and splint DNA with 5 mM acetophenone. Lane 5, size-purified TDP DNA from lane 4. All lanes were loaded with equal radioactive counts.

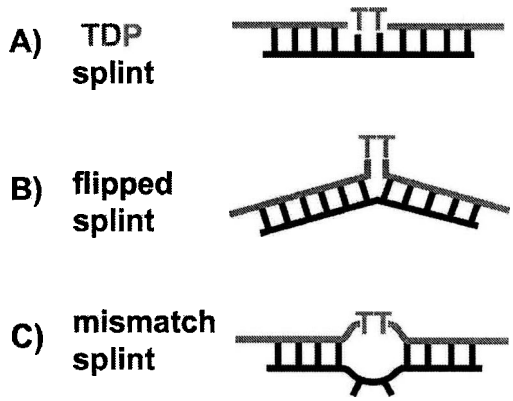


Figure 2-16 Oligonucleotide constructs used to test whether UV1C induced specific conformations in the TDP substrate.TDP hybridized to: a, A perfectly complementary, 25-nt splint. b, A "flipped" splint in which the two adenines complementary to the TDP thymine dimer have been deleted. c, A "mismatch" splint wherein the abovementioned adenines have been replaced by either two thymines or two cytosines, creating mismatches with the thymine dimer. Rates of repair of these constructs were identical to the background reaction (3.0×10^{-4} min^{-1}), all irradiated at > 300 nm at a photon flux of 3.4×10^{-9} einsteins min^{-1}.

to the thymine dimer, shown in Figure 2-16. It was anticipated that the thymine dimer might bulge out of the double helix in a manner preferred by various DNA repair enzymes (Bhattacharyya and Lilley, 1989; Vande Berg and Sancar, 1998); and (c) a complementary splint, containing two cytosines (or thymines) to replace the abovementioned adenines. Interestingly, all three constructs, in the absence of UV1C, underwent photoreactivation at or close to the background, uncatalyzed rate. Therefore, it was concluded that the catalytic role of UV1C was not solely to orient the thymine dimer in any one of the above conformations.

A clue to UV1C's mode of catalysis emerged from analysis of its metal ion requirements. I tested the activity of UV1C when folded in buffers containing only lithium, sodium, potassium or magnesium. The results are shown in Figure 2-17. UV1C required sodium or potassium ions for activity but were inactive when folded in buffers containing lithium or magnesium ions. Furthermore, methylation protection analysis

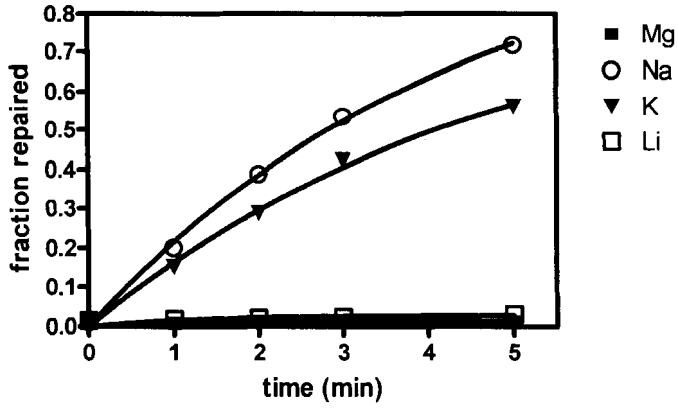


Figure 2-17 Ionic requirement for catalysis. UV1C was folded in buffers containing 50 mM Tris-Cl pH 8, and one of: 260 mM NaCl, 260 mM KCl, 260 mM LiCl, or 40 mM MgCl2. Activity was seen in reactions containing sodium or potassium buffers ($k = 0.11\ min^{-1}$ and $0.10\ min^{-1}$, respectively), and not lithium or magnesium (rates equal to background $k = 0.0004\ min^{-1}$).

of folded UV1A revealed the protection of blocks of guanines in sodium but not lithium-containing buffers (Figure 2-8). These observations, taken together with the high guanine-content of UV1C and its parent, UV1A, suggested that they fold to form guanine-quadruplexes (which are stabilized specifically by Na and K, but not by Li or Mg ions (Sen and Gilbert, 1990)). Recently, Mergny *et al.* (Mergny *et al.*, 1998) have shown, using difference spectral analysis, that G-quadruplex absorption incorporates a modestly enhanced "tail" in the 290-305 nm region.

Figure 2-18 shows a difference spectrum, ΔA, recorded for UV1C (spectra recorded in buffered 200 mM LiCl subtracted from those recorded in NaCl). The shape of this difference spectrum is fully consistent with UV1C folding to a G-quadruplex in sodium but not in lithium solution. Figure 2-19, which plots ΔA/A (the difference

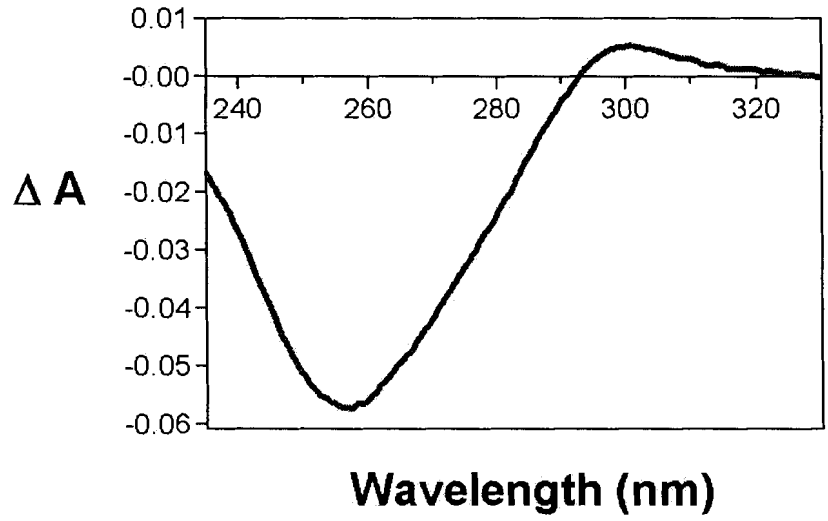


Figure 2-18 Absorbance difference spectra of the deoxyribozyme-substrate complex recorded in sodium-containing and lithium-containing buffers. The difference spectrum was obtained by subtracting the lithium spectrum from the sodium spectrum. 2 μM UV1C was folded in the presence of 2 μM of a pseudo-substrate (a 35-mer oligonucleotide of the sequence of TDP, but containing two undimerized thymines in place of the thymine dimer) in buffered 200 mM NaCl or 200 mM LiCl. Absorbance measurements were taken using a dual beam UV spectrophotometer, in a quartz cuvette with a 1 cm path length.

spectrum standardized to absorbance), shows a dominant peak at ~305 nm, a feature that strikingly resembles the peaks at 305 nm in both the quantum yield (Figure 2-13) and rate enhancement (Figure 2-14) plots. The markedly efficient reactivation by UV1C at 305 nm therefore appears to have a basis in the deoxyribozyme's spectral absorption. To test whether contaminating acetophenone from the synthesis of the thymine dimer in TDP played a part in the observed catalysis, acetophenone was deliberately added to UV1C and substrate. However, no change in rate was observed (data not shown).

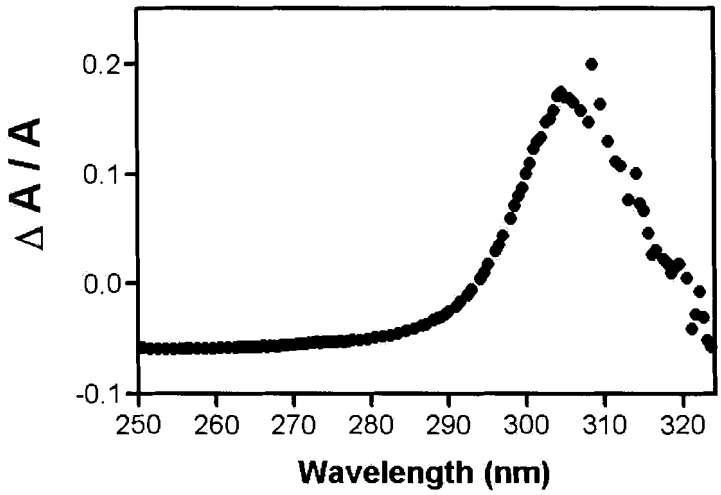


Figure 2-19 Normalized difference spectrum.
The difference spectrum shown in Figure 2-18, normalized with respect to overall absorbance.

Since the presence of a guanine-quadruplex structure appeared to be responsible for the enhanced rate of repair of thymine dimers, I wanted to test whether the presence of any unrelated guanine-quadruplex-forming DNA oligonucleotide could catalyse the repair of the TDP substrate. When PS5.M DNA, previously known to catalyse porphyrin metallation (Li and Sen, 1996), was irradiated with TDP, no repair was seen (Figure 2-20). It was also determined that UV1C

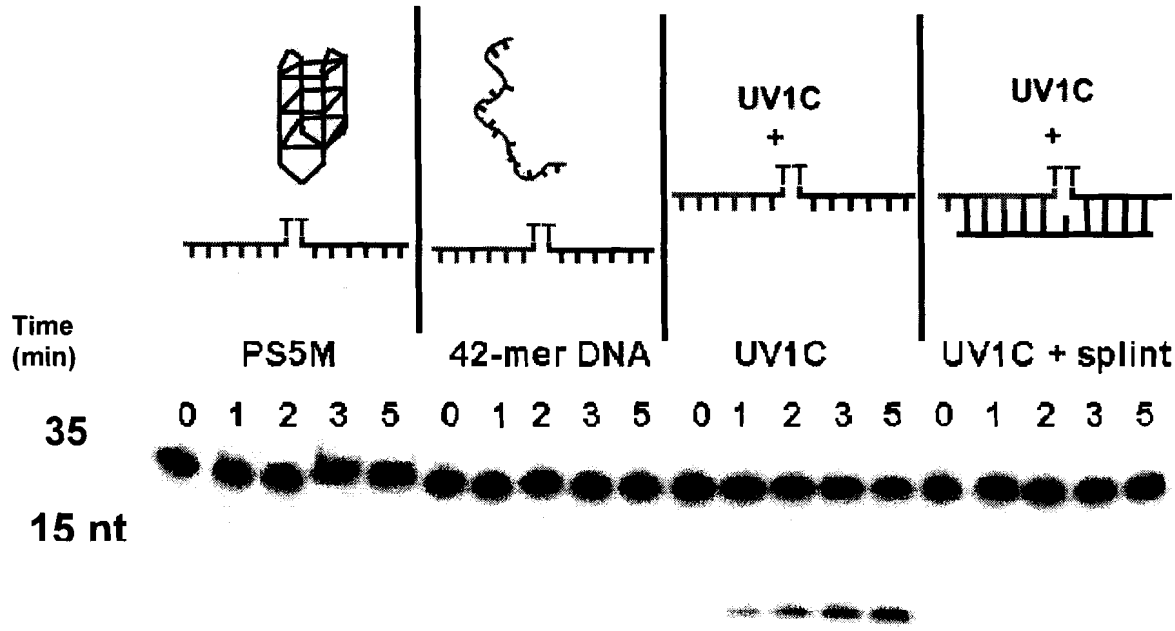


Figure 2-20 Further controls for photorepair by UV1C.
Reaction buffers contained trace concentrations of substrate DNA and 8 uM non-substrate DNA.(Far left) PS5.M, a guanine-quartet containing DNAzyme. (Left center) unrelated 42-nucleotide DNA sequence. (Right center) UV1C DNAzyme. (Far right) UV1C presented with TDP substrate within a double-stranded context.

could not repair the TDP substrate within a double stranded context when the Splint oligonucleotide used to form the substrate was added (far right in Figure 2-20). Figure 2-21 shows a model that was proposed for the folded structure of UV1C, based on methylation protection data obtained from UV1A. The recognition between the TDP

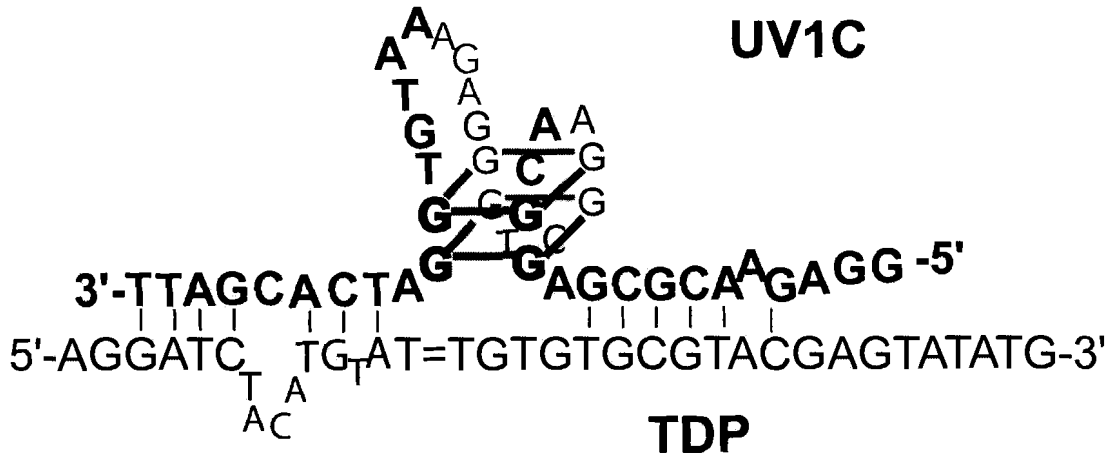


Figure 2-21 A model for the folded structure of the UV1C deoxyribozyme. Possible Watson-Crick base pairs may be formed between it and the substrate, TDP.

substrate and the UV1C deoxyribozyme appeared to involve a significant degree of tertiary interaction, since only modest stretches of Watson-Crick complementarity was found between them. Figure 2-22 illustrates a three-dimensional model for the hypothesized mechanism of UV1C. The guanine quadruplex within the folded structure of UV1C acts as an antenna for "long wavelength" (near ultraviolet) light, whose energy is then utilized with relative efficiency for the repair of the thymine dimer substrate. The precise mechanism of repair remains yet unclear; it is conceivable, by analogy with photolyase protein enzymes, that the photoexcited G-quadruplex donates one or more electrons towards the reactivation of the thymine dimers (of the four natural bases in DNA, guanine is known to be the most easily oxidised (Steenken and Jovanovic, 1997)).

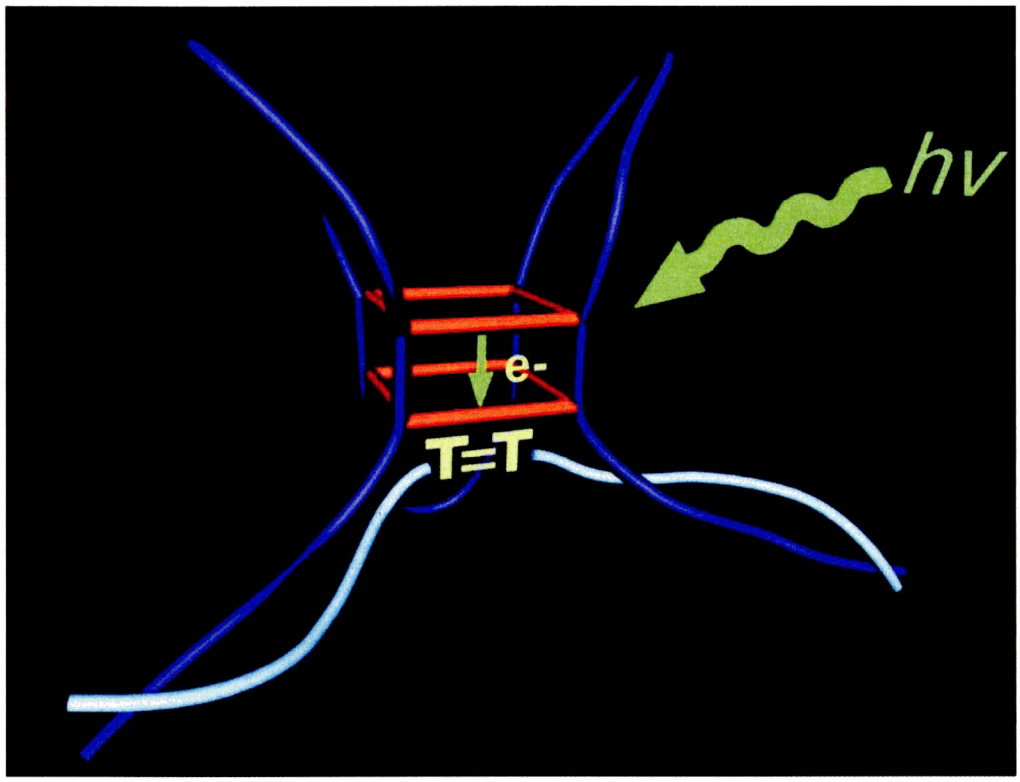


Figure 2-22 A model for the mode of catalysis by the UV1C deoxyribozyme.
A short guanine quadruplex within the folded deoxyribozyme is able to absorb near-ultraviolet light at 305 nm wavelength. The absorbed energy is then transmitted efficiently to the bound substrate, possibly in the form of one or more electrons donated from the guanines of the quadruplex to the thymine dimer, which may be stacked next to the quadruplex.

2.4 Summary

In the search for light-activated deoxyribozymes that could repair thymine dimers, I had found a population of DNA that could accomplish this task without the aid of any extraneous sensitizer. From this sensitizer-independent DNA pool, a self-repairing DNAzyme, called UV1A, was chosen for further characterization. Designing sequence constructs, I was able to separate the "substrate" sequence from the "enzymatic"

sequence to generate a "multiple turnover" catalyst (UV1C); one that could repair an externally supplied substrate. Observing the wavelength dependence on the rate of repair, the spectrum was somewhat "red shifted" with respect to that of the natural absorbance of DNA. A plot of the rate enhancement and quantum yield for repair exhibited a maximal peak of activity at 305 nm. This was unusual because DNA, as well as thymine dimers, have very low absorbances at this wavelength. Difference spectroscopy of UV1C folded with sodium versus lithium ions indicated the presence of guanine quadruplex. Furthermore, this spectrum contained a maximal peak at ~ 305 nm that matched the rate enhancement, as well as the quantum yield spectra. It was determined that the guanine quadruplex was responsible for the observed red-shifted spectrum in repair, and proposed that the guanine quadruplex may be acting as a photoantenna to absorb longer wavelength photons to donate an electron directly to the thymine dimer for repair. In the next chapter, I will examine the sequences that are necessary for catalysis and attempt to identify the folding topology of the guanine quartet core.

CHAPTER 3

INVESTIGATING THE STRUCTURE AND FOLDING TOPOLOGY OF UV1C

3.1 Introduction

In the previous chapter, I demonstrated the intriguing property of a deoxyribozyme that repaired thymine dimers upon irradiation of >300 nm light without the aid of any added sensitizers. It was determined that the presence of a guanine quadruplex directly correlated to the enhanced rate of repair at wavelengths above 300 nm. It was postulated that the G-quadruplex in the UV1C DNAzyme may be closely positioned to the thymine dimer to allow for efficient electron donation to occur aided by secondary and tertiary interactions. The proposed model suggested that the DNAzyme may bind to the substrate using a small degree of sequence complementarity to help strategically position the T<>T substrate close to the G-quartet. In this chapter, I will examine the minimal sequence requirement needed for photorepair using a series of deletion constructs for both the enzyme and the substrate. Chemical and enzymatic probing was used to provide structural information of the enzyme. Also, the overall folding topology and strand orientations of the guanine-quartet core are investigated using 5-iodouracil as a photo-crosslinking agent. From these experiments, it was possible to approximate the location of the thymine dimers with respect to the guanine

quadruplex. A revised model for the folding topology of the guanine-quartet is proposed in this chapter.

3.2 Materials and Methods

3.2.1 DNA Oligonucleotide and Reagents

DNA oligonucleotide constructs were ordered from University Core DNA Synthesis Lab (UCDNA) at the University of Calgary. The TDP substrate had the sequence 5'-AGGAT CTACA TGTAT=TGTGT GCGTA CGAGT ATATG-3'. The LDP "pseudosubstrate" sequence is identical the normal TDP substrate with the exception that it is a continuous DNA strand containing normal thymines, i.e. non-crosslinked thymines. 5-iodouracil-containing DNA was synthesized by UCDNA using 5'-dimethoxytrityl-5-iodo-2'-deoxyuridine, 3'-[(2-cyanoethyl)- (N, N-diisopropyl)]-phosphoramidite. Single base substitutions were made on both the DNAzyme and substrate.

3.2.2 Synthesis of Thymine Dimer-containing Oligonucleotide

Synthesis of the TDP substrate was performed as described in section 2.2.2. Unlabeled T2 and 5'-^{32}P-labeled T1 oligonucleotides were annealed at 400 μM concentrations to a complementary splint oligonucleotide, in buffered 40 mM $MgCl_2$. Solutions were degassed by freezing and thawing under vacuum three times. Photocrosslinking was performed in 5 mM of the triplet sensitizer acetophenone and 5% acetone. The DNA was irradiated for 30 minute in a 96-well ELISA plate with a transilluminator with a polystyrene filter to cut out wavelengths shorter than 300 nm. The TDP construct was ethanol precipitated and size-purified on 12% denaturing

polyacrylamide gels. (The TDP construct was size-purified on 12% denaturing polyacrylamide gels.)

3.2.3 Kinetic Analysis

Reactions contained 8 μM UV1C DNA enzyme in the presence of trace amounts of ^{32}P end-labelled substrate. DNA was heated to 90° C for one minute and allowed to cool slowly to room temperature in the presence of reaction buffer: 50 mM Tris-Cl, pH 8.0, 200 mM NaCl. 50 μL reaction mixtures were placed into a flat-bottom 96-well polystyrene ELISA plate (the surface of which was pre-blocked with glycogen) and irradiated on a Fotodyne transilluminator with samples exposed to 3.4×10^{-9} einsteins min^{-1} illumination. 5 μL aliquots were removed at set time points and separated on 12% denaturing polyacrylamide gels. Fragments were quantitated using Molecular Dynamics ImageQuant software. Rates were calculated using Prism4 GraphPad software using a first order rate equation and initial rates of reaction. Percent relative rates of activity are stated with respect to UV1C catalysed photoreversal (0.11 ± 0.01 min^{-1}) with a standard error of ± 3%.

3.2.4 Chemical and Enzymatic Probing

For methylation reactions, a 1.2% dimethyl sulfate (DMS) solution in water was freshly prepared and 2 μL added to 10 μL of ^{32}P 5'-end-labelled DNA folded in reaction buffer (50 mM Tris-Cl, pH.8, 200 mM NaCl). For diethyl pyrocarbonate (DEPC) reactions, 5 μL of 3% DEPC was added to 10 μL of a DNA solution in the same buffer. Both reactions were allowed to proceed for 30 minutes at room temperature, and terminated by addition of 7% β-mercaptoethanol followed by ethanol precipitation. The

recovered DNA was dissolved in 10% piperidine and heated for 30 minutes at 90° C, followed by lyophilization three times. The samples were then dissolved in denaturing gel loading buffer (containing 90% formamide v/v and 5 mM EDTA, pH 8.0) and analyzed in 12% denaturing polyacrylamide gels.

Enzymatic cleavage reactions were performed using 1 unit of Mung Bean Nuclease added to UV1C and LDP pseudosubstrate folded in buffer containing 50 mM NaOAc, 60 mM NaCl, 1 mM $ZnSO_4$ pH 5.0. DNAse I digestions were performed in 10 mM Tris-Cl, pH 7.8, 2.5 mM $MgCl_2$, 0.5 mM $CaCl_2$, 60 mM NaCl. Enzymatic digestions were incubated for 30 minutes at 25° C followed by phenol-chloroform extraction. The DNA was loaded onto 12% denaturing gels and analysed using Molecular Dynamics ImageQuant phosphorimager software.

3.2.5 Iodouracil Cross-linking

For cross-linking reactions involving iodouracil enzyme constructs, ^{32}P-end labelled LDP substrate construct (4 μM) was folded in reaction buffer in the presence of 4 μM 5-iodouracil substrate. Solutions were irradiated for 30 minutes using a Fotodyne transilluminator with maximum output at 300 nm and filtered to block out wavelengths shorter than 300 nm. Solutions were ethanol precipitated, loaded onto 12% denaturing polyacrylamide gels and exposed to both X-ray film and phosphor screens. Crosslinked species were cut out and the gel pieces eluted in a buffer containing 10 mM Tris-Cl, pH 8, 0.1 mM EDTA pH 8, 300 mM NaOAc pH 8. The eluted DNA was ethanol precipitated and resuspended in 100 μL of 10% piperidine v/v and heated to 90° C for 30 minutes. The DNA solutions were lyophilised and fragments separated and analyzed on 12% denaturing gels.

3.3 Results and Discussion

3.3.1 Truncation of UV1C From the 3' and 5' ends

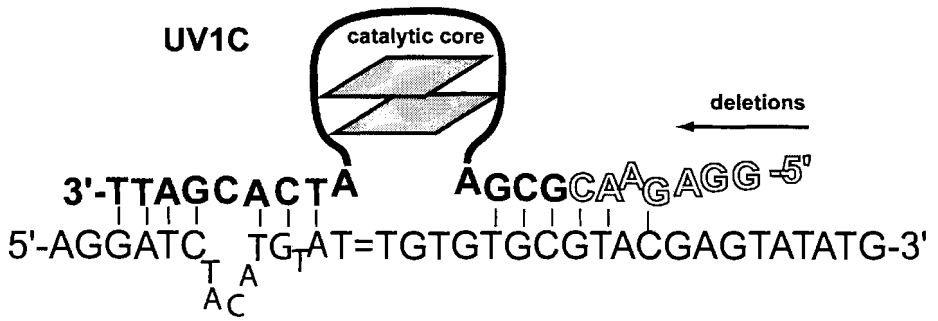
To gain a general understanding of sequences that were important for catalytic activity, a series of deletion mutants were designed that eliminated large stretches of sequences (7-16 nucleotides) from their 5' and 3'-ends. These are shown in Figure 3-1 with their relative percent activity compared to the full-length enzyme (UV1C). Out of the ten constructs tested, only one (UV1F, shown as a boxed sequence) was found to be active (83 % activity). UV1F corresponded to a sequence in which seven nucleotides from the 3'-end were removed from the "wild-type" DNAzyme. Interestingly, enzyme activity was lost after removal of seven or more bases from the 5'-end (represented by UV1D). Loss of activity observed after deletion of nucleotides normally implies that

	rel. activity	sequence
UV1C	100%	GGAGAAC GCGA**GG**C AA**GG**CT **GGG**AGAAA TGT**GG**AT CACGATT
UV1D	0%	GCGA**GG**C AA**GG**CT **GGG**AGAAA TGT**GG**AT CACGATT
UV1E	0%	AA**GG**CT **GGG**AGAAA TGT**GG**AT CACGATT
UV1F	83%	GGAGAAC GCGA**GG**C AA**GG**CT **GGG**AGAAA TGT**GG**AT
UV1G	0%	GAGAAC GCGA**GG**C AA**GG**CT **GGG**AGAAA
UV1H	6%	GCGA**GG**C AA**GG**CT **GGG**AGAAA TGT**GG**AT
UV1I	0%	AA**GG**CT **GGG**AGAAA
UV1J	5%	GCGA**GG**C AA**GG**CT **GGG**AGAAA
UV1K	0%	AA**GG**CT **GGG**AGAAA TGT**GG**AT
UV1L	0%	GAGAAC GCGA**GG**C AA**GG**CT
UV1M	0%	**GGG**AGAAA TGT**GG**AT CACGATT

Figure 3-1 Deletion construct sequences and their relative catalytic activity.

these sequences provide crucial interactions with the substrate. As expected, the deletion of sequences containing guanines believed to form G-quartets, represented by UV1E, UV1G and UV1I to UV1M, resulted in complete loss of activity.

To further investigate the loss of activity observed when UV1C was truncated by seven nucleotides (represented by UV1D), additional constructs were designed and tested. These corresponded to single nucleotide deletions from the 5'-end of UV1C, shown in Figure 3-2. When three bases (5'-GGA-3') were removed from the 5'-end of the sequence (represented by the C.2 construct), no significant decrease in rate was observed. This is in agreement with the proposed model shown in Figure 3-2, where these three nucleotides were proposed not to participate in hydrogen bonding with the

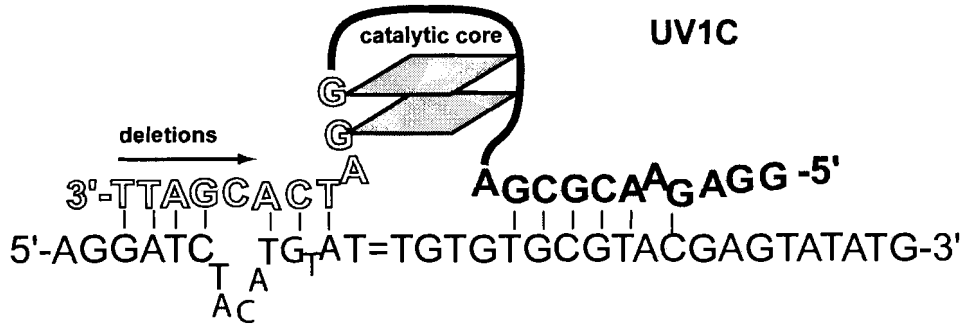


	rel.activity	SEQUENCE
UV1C	100%	GGAGAAC GCGAGGC AAGGCT GGGAGAAA TGTGGAT CACGATT
C.6	92%	GAGAAC GCGAGGC AAGGCT GGGAGAAA TGTGGAT CACGATT
C.7	110%	AGAAC GCGAGGC AAGGCT GGGAGAAA TGTGGAT CACGATT
C.2	90%	GAAC GCGAGGC AAGGCT GGGAGAAA TGTGGAT CACGATT
C.3	60%	AAC GCGAGGC AAGGCT GGGAGAAA TGTGGAT CACGATT
C.4	40%	AC GCGAGGC AAGGCT GGGAGAAA TGTGGAT CACGATT
C.5	34%	C GCGAGGC AAGGCT GGGAGAAA TGTGGAT CACGATT
UV1D	0%	GCGAGGC AAGGCT GGGAGAAA TGTGGAT CACGATT

Figure 3-2 Truncation construct sequences that omitted bases from the 5'end of UV1C. (Top) Diagram illustrating proposed secondary structure interactions with TDP substrate. Outlined bases indicate nucleotides that were deleted from the 5'-end of the enzyme.

substrate. When three additional bases were removed from this end (5'-GAA-3'), a 70 % decrease in rate was observed. After the seventh nucleotide was deleted (represented by UV1D) the rate dropped to background levels. These results agreed with the proposed model, where each base deleted would resulted in a weaker association with the substrate. Three base pairing interactions were removed out of the six that were proposed to form. This would result in two G-C pairings and one unstable G-T "wobble" pair, which would contribute negligibly to overall stability of base- pair stems. For helix formation, at least three base pairs are required for nucleation (Saenger, 1984). It is unlikely that a two base pair association would occur under normal conditions due to a positive free energy (Saenger, 1984). It is also possible that tertiary interactions may be present, but this cannot be predicted without high resolution NMR or crystallographic techniques.

Using the same approach, deletion constructs for the 3'-end were tested, which removed single nucleotides up to and including two guanines contained in the putative guanine-quartet. These are shown in Figure 3-3. Removal of three bases from the 3'-end (5'-ATT-3', corresponding to the C.10 construct) had no effect on activity. Deletion of four additional bases from the 3'-end (5'-CACG-3'), as well as the two guanine bases from the 5'-end (to form UV1N) resulted in a 30 % loss in activity. From the working model, shown in Figure 3-3, these bases were proposed to participate in base pairing with the 5'-end of the substrate. One would expect these deletions to have a greater effect on activity, since it would leave only a single A-T base pair and thus are not likely to form under these conditions. As expected, no activity was seen with the further deletion of nucleotides, which included the two guanines proposed to be in the quadruplex (represented by UV1N.3).



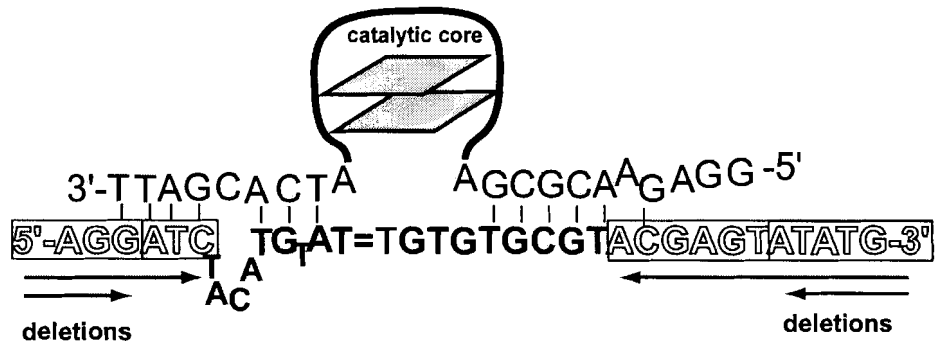
rel.activity		SEQUENCE
UV1C	100%	GGAGAAC GCGAGGC AAGGCT GGGAGAAA TGTGGAT CACGATT
C.8	89%	GGAGAAC GCGAGGC AAGGCT GGGAGAAA TGTGGAT CACGAT
C.9	102%	GGAGAAC GCGAGGC AAGGCT GGGAGAAA TGTGGAT CACGA
C.10	99%	GGAGAAC GCGAGGC AAGGCT GGGAGAAA TGTGGAT CACG
F	70%	GGAGAAC GCGAGGC AAGGCT GGGAGAAA TGTGGAT
N	71%	AGAAC GCGAGGC AAGGCT GGGAGAAA TGTGGAT
N.2	35%	AGAAC GCGAGGC AAGGCT GGGAGAAA TGTGGA
N.3	6%	AGAAC GCGAGGC AAGGCT GGGAGAAA TGTGG
N.4	4%	AGAAC GCGAGGC AAGGCT GGGAGAAA TGTG
N.5	2%	AGAAC GCGAGGC AAGGCT GGGAGAAA TGT

Figure 3-3 Truncation constructs removing bases from the 3'-end of UV1C. (Top) Diagram illustrating proposed secondary structure interactions with TDP substrate. Outlined bases indicate nucleotides that were deleted from the 5'-end of the enzyme.

3.3.1.1 Substrate Deletion Constructs

Substrate deletion constructs were synthesized analogous to the enzyme deletion constructs, in which nucleotides from both the 5' and 3' ends were removed. These are shown in Figure 3-4.

Deletion of five bases from the 3'-end of the substrate (construct T1T4), and a further six bases (construct T1T6), resulted in small changes in activity (78 % activity compared to the full-length substrate). This is in agreement with the proposed model, where these bases are not likely to participate in a significant degree of hydrogen bonding to the enzyme, shown in Figure 3-4.



	rel. activity	SEQUENCE
TDP	100%	5'- AGG ATC TAC ATG TA **T=T** GT GTG CGT ACG AGT ATA TG –3'
T1T4	85%	5'- AGG ATC TAC ATG TA **T=T** GT GTG CGT ACG AGT –3'
T1T6	78%	5'- AGG ATC TAC ATG TA **T=T** GT GTG CGT –3'
T3T2	56%	5'- ATC TAC ATG TA T=T GT GTG CGT ACG AGT ATA TG –3'
T5T2	0%	5'- TAC ATG TA T=T GT GTG CGT ACG AGT ATA TG –3'
T3T4	34%	5'- ATC TAC ATG TA T=T GT GTG CGT ACG AGT –3'
T5T6	0%	5'- TAC ATG TA T=T GT GTG CGT –3'

Figure 3-4 Substrate Deletion Mutants
(Top) Diagram illustrating proposed secondary structure interactions with TDP substrate.

Interestingly, truncation from the 5' end of the substrate showed a more dramatic result, where the removal of three to six nucleotides resulted in drastic reduction in activity. This would not be expected since, according to the model, only one of the bases deleted was proposed to participate in base pairing.

3.3.1.2 Model Revision of the 3' end Binding

Since the substrate 5'-end deletion results did not agree with the enzyme 3'-end deletion results, it was clear that the proposed model for base pairing required further refinement. Deletion data from the 5'-end the enzyme was consistent with the results of the deletions from the 3'-end of the substrate. On the other hand, deletion data from the 5'-end of the substrate was not consistent with base-pairing with the 3'-end of UV1C. A logical conclusion would be that no significant base pairing occurred with the 3'-end of

the enzyme to the 5'end of the substrate. If this were the case, then where would the 5'-end of the substrate bind? Also, the purpose of the 3'-end was unclear.

A possible clue on how this might occur would come from revisiting the initial engineering of a "multiple turnover" enzyme (UV1C) from the parent DNAzyme, UV1A, in the previous chapter (Section 2.3.2). The first construct designed, UV1B, consisted of the formerly 40 nucleotide random region plus the 3' constant primer region, called LT2 (Figure 2-9). The second construct, UV1C, was made by the removal of the LT2 sequence from the 3'end. Upon revisiting the sequences of the construct when the LT2 sequence was added back to the 3'end of UV1C (to form UV1B), the LT2 sequence would potentially form a high degree of base pairing to the 3'-end. This is illustrated in Figure 3-5, with the LT2 sequence represented by open letters. This also would allow only a small degree of base pairing to occur (three base pairs) with the 5'end of the substrate. Thus, in the separation of the primer sequence from the UV1B to form UV1C,

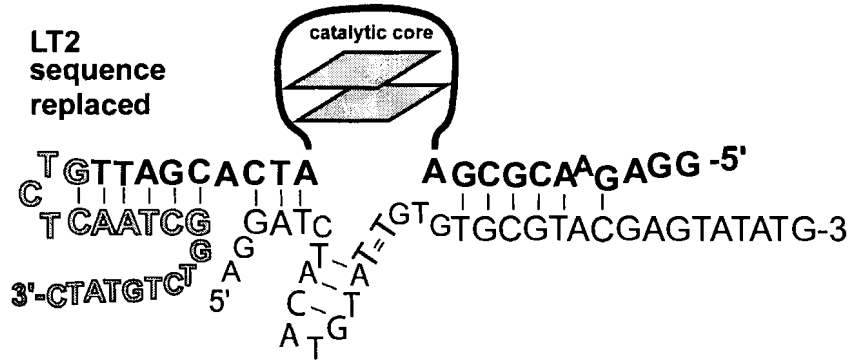


Figure 3-5 Completed sequence by re-addition of 3'-constant region forms base pair complementarity with 3' overhang.

sequences that may have formed the stem loop structure were removed. Although this would assume that an important component of the DNAzyme structure was removed, the

catalytic rates were identical for UV1C, UV1B, as well as UV1C with the LT2 sequence added, previously shown in Figure 2-10.

The substrate could associate with the UV1C enzyme as shown in Figure 3-6A, leaving a six nucleotide overhang. This may also agree with the substrate deletion data from the 5'-end, in which removal of three to six bases could drastically affect binding. It is important to note that there is no experimental evidence to suggest the presence of a three base-pair stem loop, nor the presence of the stem loop structure.

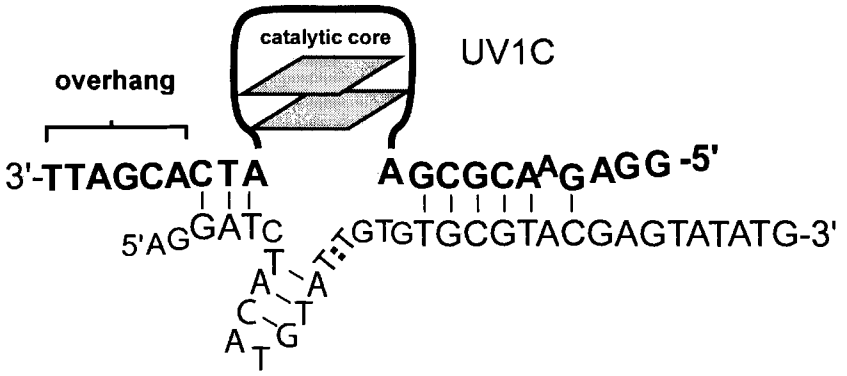


Figure 3-6 Model of substrate interaction with 3' end of UV1C.

3.3.1.3 Creation of Complementary Binding Arms

Based on the data found from the truncation constructs, I wanted to determine if it were possible to make a fully complementary binding arm to the 5'-end, or to both the 3' and the 5'-ends. Three mutant enzymes were designed, shown in Figure 3-7, in which sequences on the enzyme were made fully complementary on the 5'-side (UV1C.5'ARM) and on both sides (UV1C.2ARMS). When tested, both constructs demonstrated activity comparable to the original DNAzyme. Although it was unclear how the 3'-arm of the

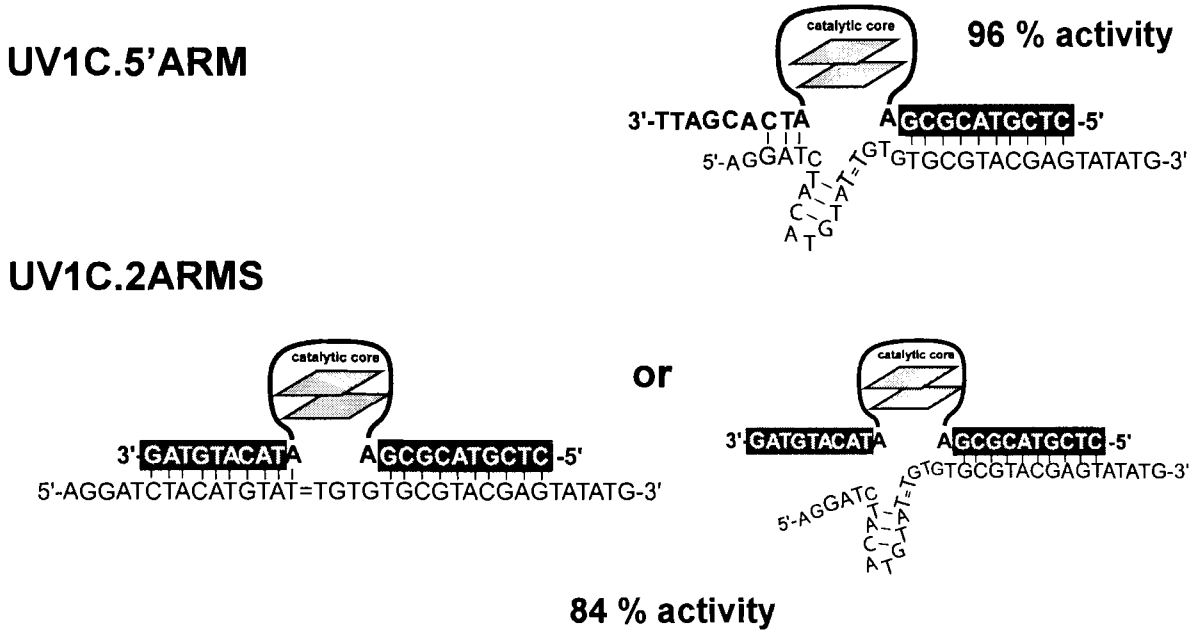


Figure 3-7 Creation of complementary binding arms to substrate.
(Top) UV1C.5'ARM makes the entire 5' end of the DNAzyme completely complementary with the substrate. (Bottom) UV1C.2ARMS construct may either make the 5' and 3' ends of the enzyme complementary with the substrate, or only the 5' end.

2ARMS construct base paired with the substrate, a slight reduction in activity (16 %) may be an indication that the complementary sequence may slightly interfere with the folding of the substrate. To confirm whether or not the 3'-end of the enzyme is associating with the substrate, further experiments could be designed where the sequences on this "3'-overhang" could be changed to different sequences. If no base pair interactions occur with the substrate, the catalytic rate would not be effected.

3.3.2 Examining the Central Core

Next, I wanted to determine whether the proximity of the guanine quartet to the thymine dimer was important for the catalysis. To test this, I designed UV1C.2A shown

in Figure 3-8, in which two adenines were added to hypothetically "distance" the quartet away from the thymine dimer. When tested for activity, I found almost a complete loss in photorepair (92 % reduction). Assuming that the quartet is positioned as depicted in

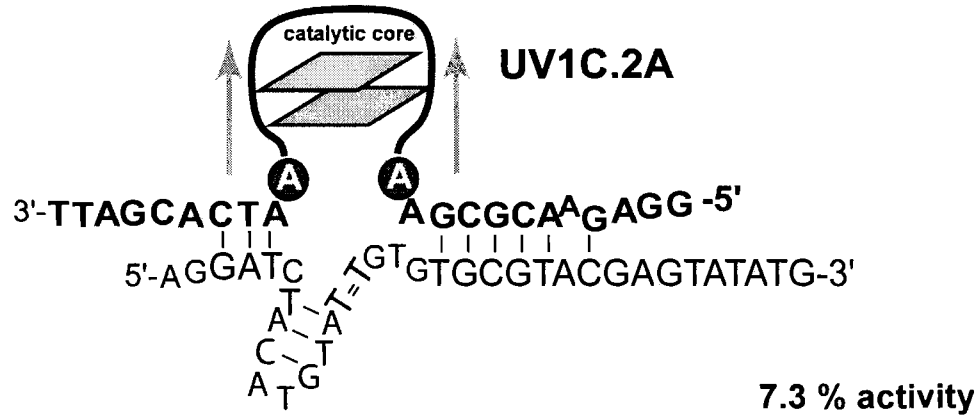
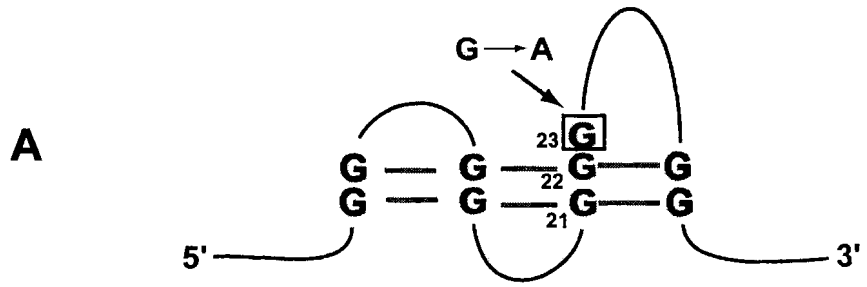


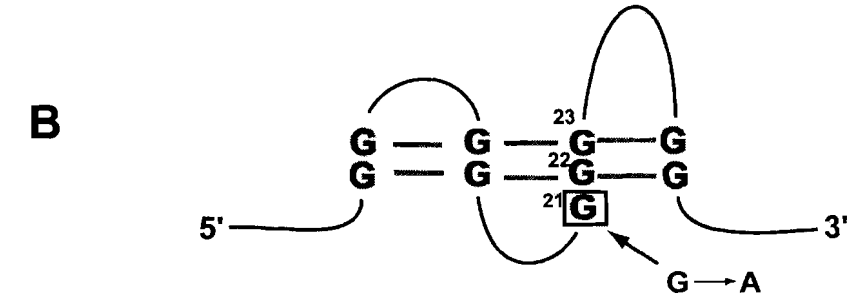
Figure 3-8 The addition of two adenines to theoretically "push away" the catalytic core from the thymine dimers. A 92 % decrease in activity was observed.

Figure 3-8, this large effect on activity would support the idea that the thymine dimers are closely coupled to the guanine quartet core.

Since a two-layered guanine quadruplex was proposed to form, I wished to determine which of the three guanines contained in the guanine triplet (G21, G22, or G23) were participating in quadruplex formation. Of these, only two can participate in quartet formation. This is illustrated in Figure 3-9. To test this, I designed two mutants with single base changes: One construct, G21A, substituted adenine for guanine 21. The other construct, G23A, substituted adenine for guanine 23. A loss of activity of a substitution at one position, and full activity at the other position, would imply that the latter position was not participating in guanine-quartet formation, and vice versa. When tested for activity, it was found that the G21A mutant was relatively non-functional ($3.5 \times 10^{-3} \pm 0.5 \times 10^{-4}$ min^{-1}) when compared to the G23A construct displayed full activity



G21A 5'- GGAGAACGCGAGGCAAGGCTAGGAGAAATGTGGATCACGATT –3'



G23A 5'- GGAGAACGCGAGGCAAGGCTGGAAGAAATGTGGATCACGATT –3'

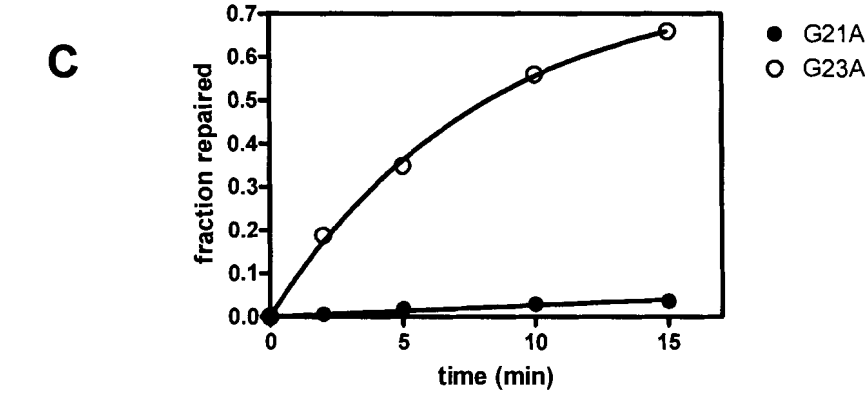


Figure 3-9 The participation of guanine 21 versus guanine 23 in the central guanine-quartet. An open view of the guanine quadruplex core showing two possible arrangements of the guanine triplet with the sequence mutants shown underneath. A) Guanine 21 is included in the quartet resulting in G23 being excluded. B) Guanine 23 is included in the quartet resulting in G21 exclusion. C) Plot of repair by G21A and G23A mutant constructs. G23A repairs at a rate of 0.12 ± 0.01 min^{-1} which is much faster than G21A at 3.5 x10^{-3} ± 0.5 x10^{-4} min^{-1}.

(0.12 ± 0.01 min^{-1}). This strongly suggested that G21 participated in the guanine quartet, whereas G23 is did not.

3.3.3 5-IodoUracil Cross-linking

The use of halogenated analogs of thymine has previously been used to crosslink nucleic acids to proteins (Ogata and Gilbert, 1977; Willis *et al.*, 1993). The most common is that of 5'bromodeoxyridine (BrdU) to induce double-stranded breaks, alkali labile to bonds and crosslinks proteins (Barbier *et al.*, 1984; Wick and Matthews, 1991). The advantage of using these analogs to replace thymine lies in the similarity of the Van der Waals radii of the halogen to that of the methyl group. Bromine has a van der Waals radius of 1.94 Å, and iodine is slightly larger at 2.15 Å, which are both close to that of the methyl group (2.0 Å) (Willis *et al.*, 1993). The structures of 2'-deoxythymidine and 5-iodo-2'-deoxyuridine are shown in Figure 3-10. Due to their high degree of similarity, substitution of these halogenated uracil analogs into DNA or RNA minimally affects nucleic acid structure, as well as hydrogen bonding interactions. In addition, they are considered "zero length" crosslinkers requiring very close contact of the photoreactive

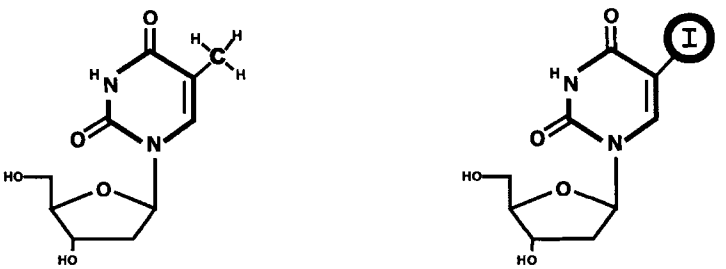


Figure 3-10 Chemical structures of 2'-deoxythymidine (left) and 2'-deoxy 5-Iodouridine (right).

base in order for crosslinking to occur. Photoirradiation of light >300 nm results in radical formation upon leaving of the halogen atom. This radical then reacts with nearby molecules to form crosslinked species. Iodouracil substitutions within the diagonal loops of guanine-quartet containing DNA have been demonstrated to form 2'-deoxyribonolactone residues with the release of a nucleic acid base after the irradiation by Sugiyama and coworkers (Xu and Sugiyama, 2004). From this study, it was proposed that iodouracil could be used to determine the presence of antiparallel guanine quadruplex formation using an HPLC assay. The photoreactivity of halogenated uracil has also been useful in understanding RNA and DNA interactions with proteins, as well as applied to *in vitro* selection techniques to select for high affinity nucleic acids to proteins (Golden *et al.*, 2000). Interestingly, protein-RNA crosslinking studies indicated that 5-iodouracil was a more efficient crosslinker than 5-bromouracil (Willis *et al.*, 1993). I wished to determine if iodouracil could be used to probe for the interactions of the substrate with the enzyme.

To investigate the bases on the UV1C enzyme that may interact with the thymine dimer substrate, I designed 5-iodo-2'-deoxyuracil constructs that substituted the thymine dimer positions with iodouracil (Figure 3-11). Also, substitutions were made for thymine residues at various positions on the enzyme. The crosslinking of these bases to positions on the enzyme (which need to be in close contact) would result in a slower migrating species on a polyacrylamide gel. Treatment of the crosslinked DNA with piperidine could then result in strand cleavage, which could then be identified on a second polyacrylamide

gel. Using this strategy structural information, such as strand orientations and the overall folding topology of the guanine quartet can be determined.

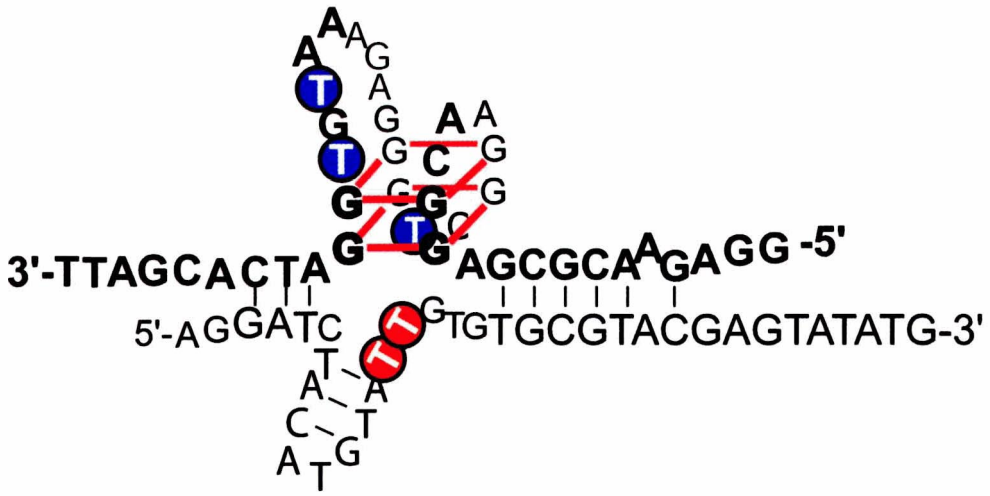


Figure 3-11 Positions of iodouracil substitutions.
Each construct created had a single substitution for a thymine residue on either the substrate or the enzyme.

^{32}P - end labelled enzyme was folded in the presence of 5-iodouracil – containing substrates (T15-IdU and T16-IdU) and irradiated at 300 nm for thirty minutes to form cross-linked species. The DNA strands were size separated and cut out from a denaturing gel, shown in Figure 3-12. A slower migrating crosslink is observed which corresponds to the substrate cross-linking to the enzyme (Lanes 3 and 6). These are not observed in the corresponding "dark" control lanes (Lanes 2 and 5) where the sample was not irradiated. This crosslinked species was then isolated from the gel and subjected to piperidine treatment. The DNA was then run on denaturing gels and piperidine-labile fragments were mapped to positions on the DNA strand, shown in Figure 3-13.

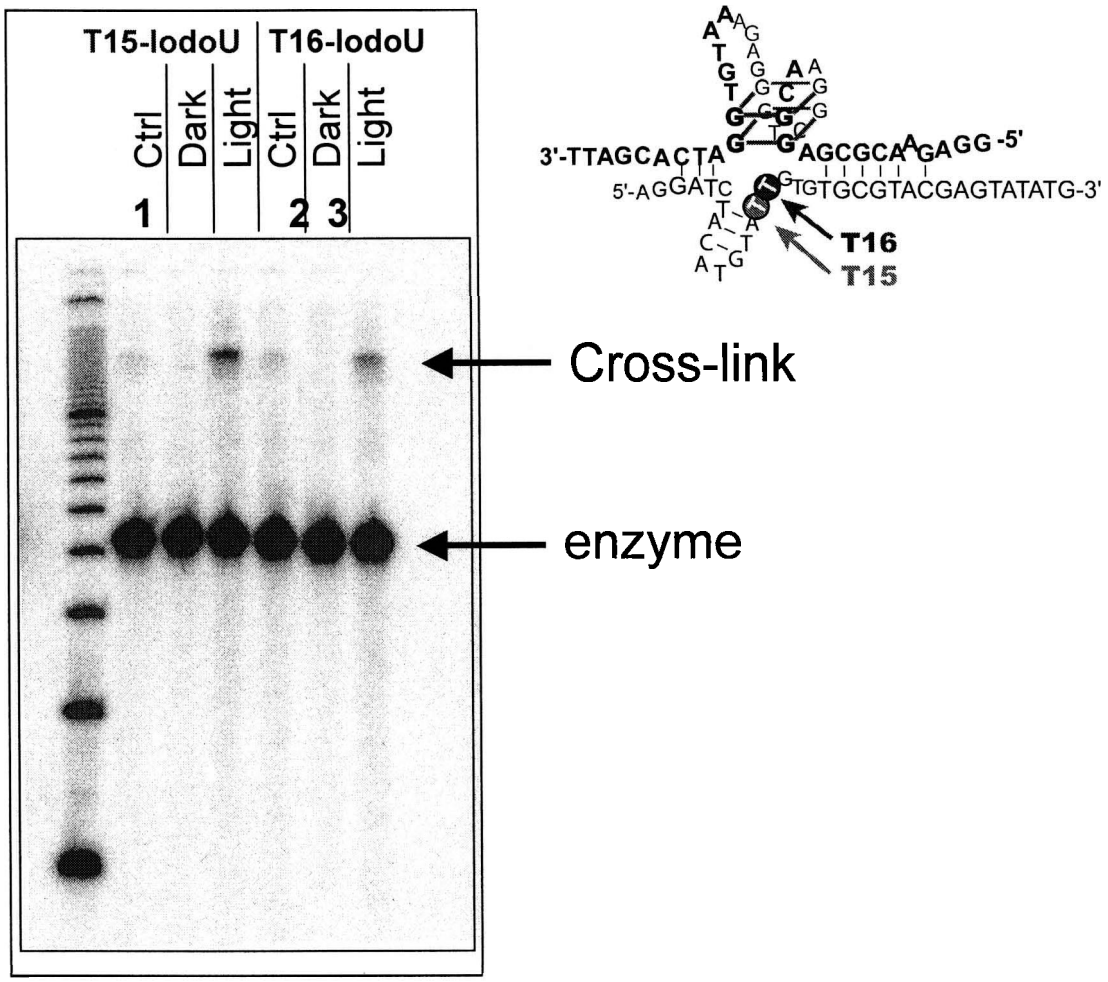
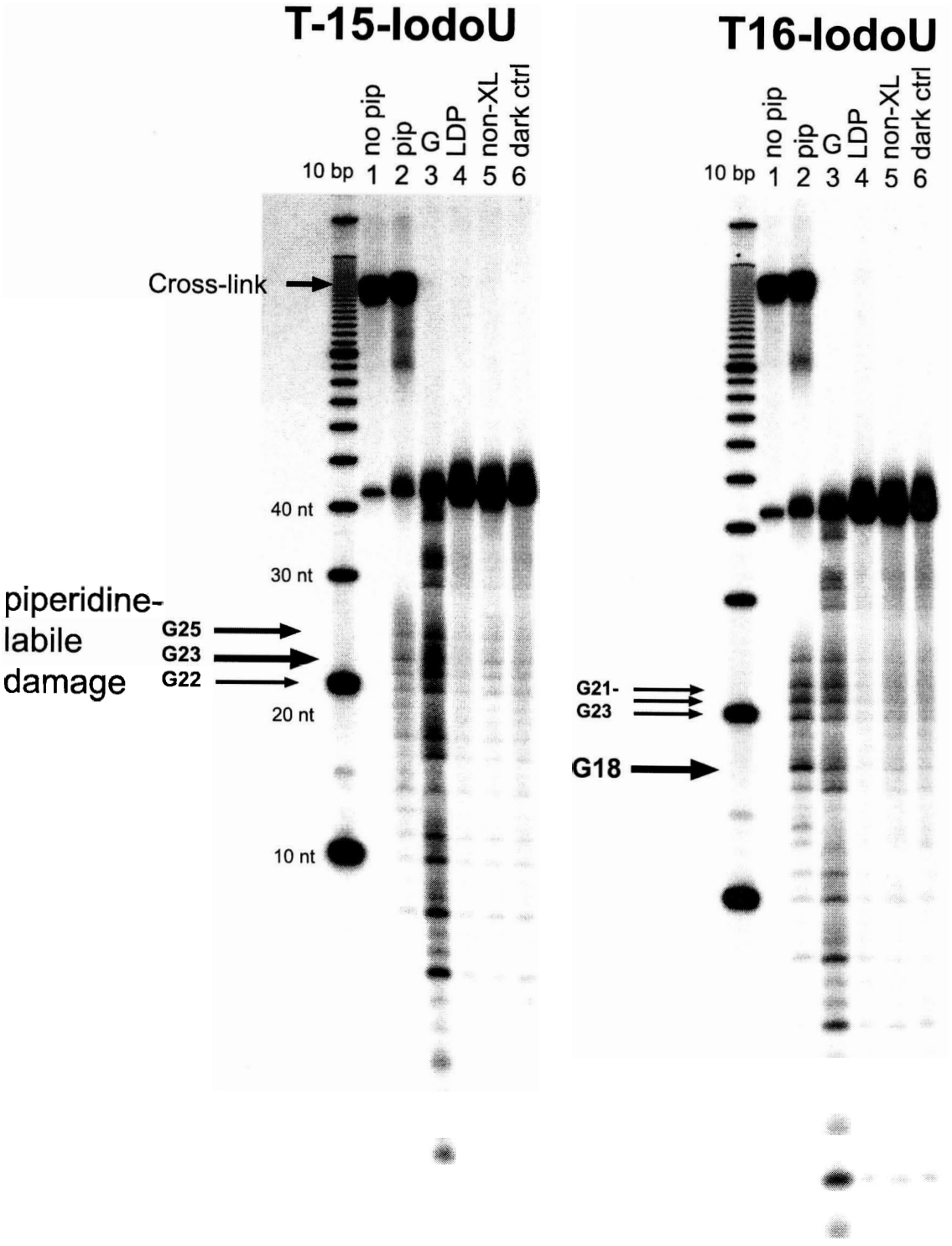


Figure 3-12 Photochemical Cross-linking of UV1C to 5-Iodouracil-containing substrates. (Inset): Positions of Iodouracil substitutions on the TDP substrate strand. Lanes 1 & 4: Control lanes of unsubstituted,UV1C irradiated with LDP (pseudosubstrate). Lanes 2 & 5: Non-irradiated T15IdU and T16IdU substrate construct in the presence of UV1C. Lanes 3 & 6: Irradiated T15IdU and T16IdU constructs in the presence of UV1C. Crosslinked species are seen as slower migrating bands.

The T15-IdU and T16-IdU constructs, which correspond to the positions of the thymine dimer, reveal the bases located within the catalytic core. Damage is observed with the T15 construct to guanine bases 22, 23 and 25. Of these, guanine 22 is believed to form the guanine quadruplex. Interestingly, the T16-IdU construct shows a different pattern of damage, where strong damage is observed at guanine 18, and weaker damage

Figure 3-13 Mapping of Cross-linked damage following piperidine treatment.
Denaturing polyacrylamide (12 %) gels showing piperidine labile crosslinked positions of T15-IdU (Left) and T16-IdU (Right) substrate constructs relative to UV1C. In both gels, Lane 1 shows the crosslinked DNA without piperidine treatment. Lane 2 shows DNA fragments formed after piperidine treatment of the crosslinked DNA relative to the G ladder in Lane 3. Damage is seen primarily to guanines 23 and 25 in the T16 construct, whereas damage mainly occurs to guanine 18 and to a lesser extent, guanines 21 - 23. (Lanes 4-6 are control lanes) Lane 4 contains piperidine treated LDP pseudosubstrate as a control. Lane 5 contains UV1C DNA that was non-crosslinked in the gel containing the crosslinked species. Lane 6 contains UV1C that was folded with the iodouracil construct, but not irradiated.


T-15-IodoU
10 bp
1 no pip
2 pip
3 G
4 LDP
5 non-XL
6 dark ctrl
Cross-link
40 nt
30 nt
20 nt
10 nt
piperidine-
labile
damage
G25
G23
G22
T16-IodoU
10 bp
1 no pip
2 pip
3 G
4 LDP
5 non-XL
6 dark ctrl
G21-
G23
G18

to guanines 21-23. Of these, guanine 18 is proposed to participate in the guanine quartet.

Based on the information from these crosslinking studies, a tentative folding topology of the guanine quartet can be proposed. A potential means by which these damage patterns could occur is illustrated in Figure 3-14. The middle two strand orientations are parallel to each other, resulting in a diagonal loop consisting of C and T bases. The outer strands are also parallel to each other, but antiparallel with respect to the inner strands. This strand orientation allows for T16 to be in close contact with G18, as well as G23 and the other two guanine bases within the quadruplex. Additional information from the T16 gel indicate that the CT bases in the diagonal loops were not damaged. This is also in agreement with this proposed strand orientation.

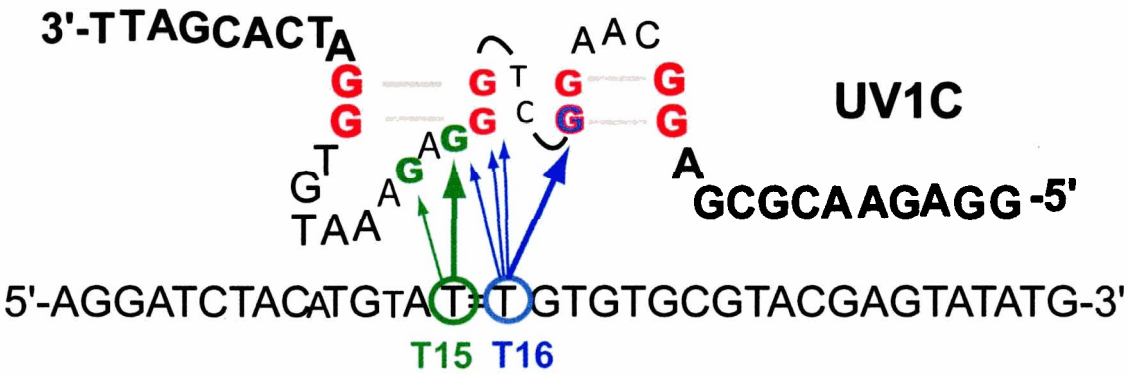


Figure 3-14 Potential topological folding based on substrate 5-iodouracil crosslinking data. Thicker arrows indicate greater degree of crosslinking. The strand orientation suggests a diagonal loop consisting of cytosine and thymine between the second and third quartets.

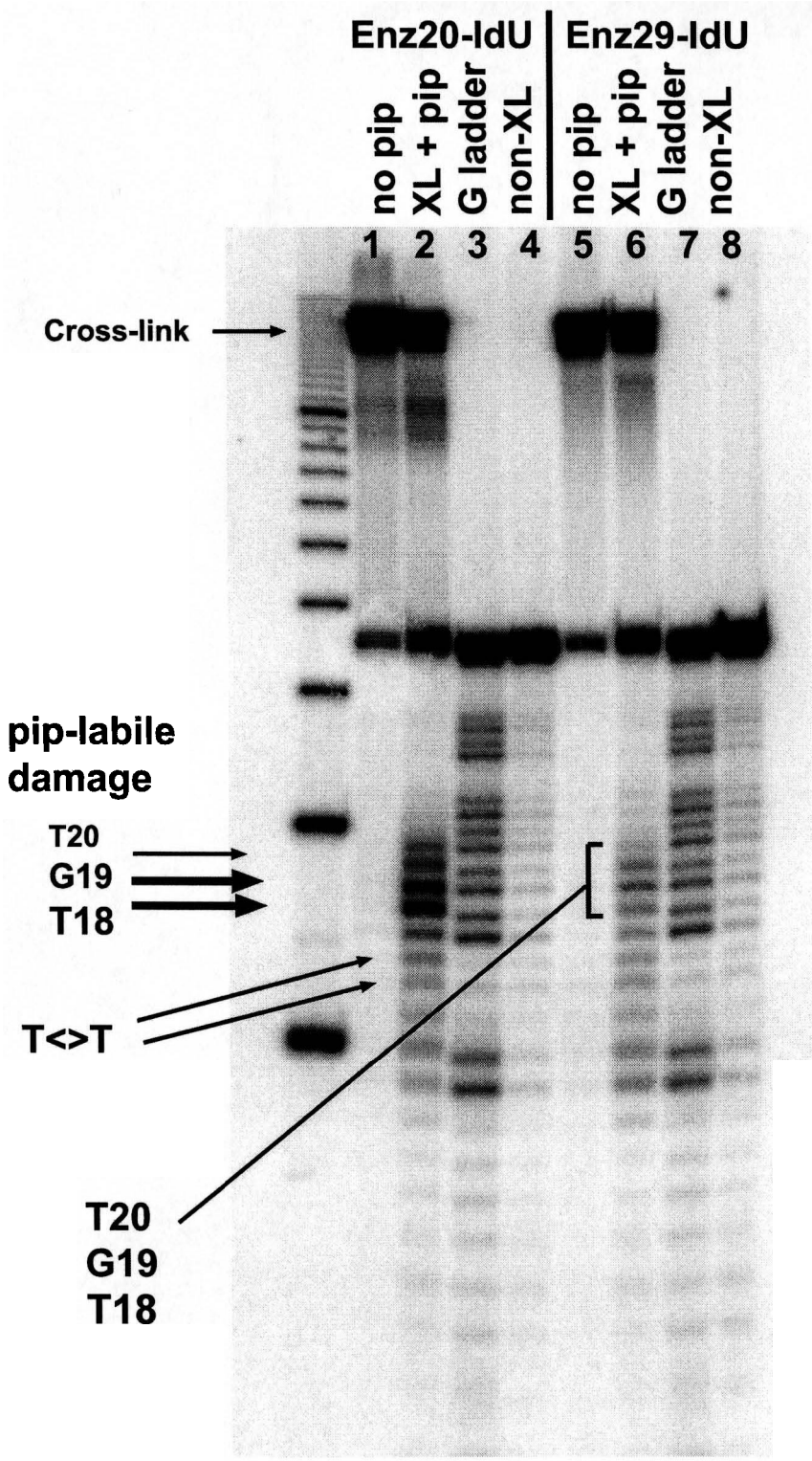
Following the same methodology, thymines located on the enzyme, shown in Figure 3-11, were also replaced with 5-iodouracil. These correspond to positions T20 (located in the now proposed diagonal CT loop), T29 and T31 of the large loop. Photoirradiation of end-labelled substrates with the iodouracil enzyme constructs resulted in crosslinks that were then subjected to piperidine treatment and run on 12% denaturing gels, shown in Figure 3-15. The Enz20-IdU mutant resulted in strong crosslinking to thymine 18 and guanine 19 on the substrate strand, which is one position away from the thymine dimer. Weaker crosslinks occurred to bases T20 and G21. The Enz29-IdU construct shows crosslinking to guanine 19 and thymine 20. The crosslinking data from the T31-IdU demonstrated similar results to that of the T29 mutant (data not shown). This is expected considering the substituted positions were one nucleotide position apart.

The mapping of these piperidine-labile crosslinks with respect to the tentative folding topology is shown in Figure 3-16. The large loop may be in contact with bases immediately 3' to the thymine dimer. Thymine 20 appears to have direct contact with the same positions. Although it is difficult to conceptualize how this might occur, there may be a high degree of tertiary interactions present that cannot be accounted for.

3.3.3.1 Chemical Probing of UV1C

Chemical probing with permanganate was used to gain further structural information of UV1C. $KMnO_4$ is known to cis-hydroxylate thymine residues at the 5 and 6 positions to form 5,6-dihydroxy-5, 6-dihydrothymine. Subsequent piperidine treatment

Figure 3-15 Damage patterns of 5-Iodouracil enzyme constructs following cross-linking. Denaturing polyacrylamide (12 %) gels showing piperidine labile crosslinked positions of Enz20-IdU (Lanes 1 - 4) and Enz29-IdU (Lanes 5 - 8) enzyme constructs relative to LDP pseudosubstrate. Lanes 1 & 5 show the crosslinked DNA without piperidine treatment. Lanes 2 & 6 show DNA fragments formed after piperidine treatment of the crosslinked DNA relative to the G ladder in Lanes 3 & 7. Damage is seen primarily to guanine 18 and thymine 19 in the Enz20 construct, whereas damage occurs to a lesser extent to thymines 18 & 20 and to guanine 19. Lanes 4 & 8 contains piperidine treated LDP that was folded with the iodouracil construct, but not irradiated.


Enz20-ldU
Enz29-ldU
no pip
XL + pip
G ladder
non-XL
no pip
XL + pip
G ladder
non-XL
1 2 3 4 5 6 7 8
Cross-link
pip-labile damage
T20
G19
T18
T<>T
T20
G19
T18

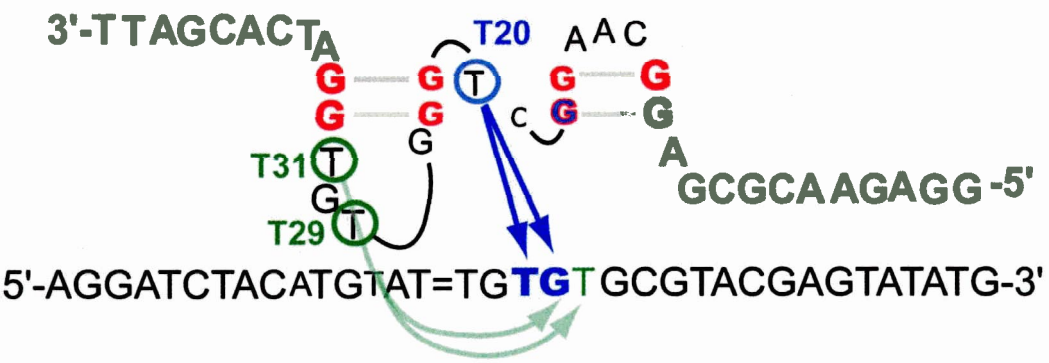


Figure 3-16 Cross-linking due to enzyme 5-iodouracil substitutions.
Thymine 20 on UV1C shows reactivity to thymine 18 and guanine 19 on the subtrate strand. Both thymines 29 and 31 show identical crosslinking patterns with respect to each other and contact guanine 19 and thymine 20 on the substrate strand.

results in degradation, leading to strand scission. Thymines that are more exposed to solvent result in hyper-reactivity, as opposed to thymines within a double stranded context. When UV1C was folded with TDP and subject to permanganate probing, thymine 20 was found to be hyper-reactive. This is shown in Figure 3-17. Thymine 20 was proposed to form the two base diagonal loop between the quartets, which would result in a flipped out thymine. As a result, this thymine would be more susceptible to permanganate treatment and agrees with the proposed diagonal strand configuration. Interestingly, thymine 29, also located within a loop, is shown to be slightly hyper-reactive.

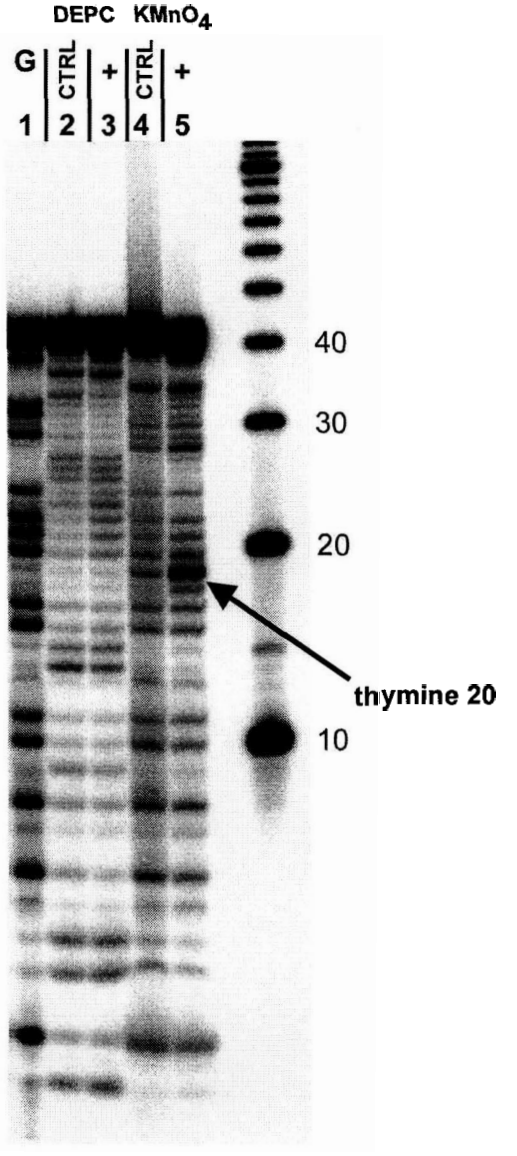


Figure 3-17 Chemical and enzymatic probing of UV1C bound to TDP substrate.
GEL A: Chemical probing - Lane 1: Guanine ladder made with dimethyl sulfate, Lane 2: DEPC control lane with unfolded DNA. Lane 3: DEPC-treated UV1C folded in buffer and bound to TDP substrate. Lane 5: Control lane for KMnO4 with unfolded DNA. Lane 6: KMnO4-treated UV1C folded in buffer and bound to TDP substrate. Hyperreactivity of thymine 20 is clearly visible.

3.3.4 Importance of Loop Sequences

From the crosslinking data, it was determined that certain bases within in the loop of the substrate sequence were in close contact with thymine dimer. To further investigate whether these sequences were important for enzymatic repair, mutant constructs were designed that: 1) replaced all sequences of the loops with adenines, and 2) shortened the large loop to three adenine bases. These are shown in Figure 3-18. Replacement of all loop sequences with adenines (ALOOPS) resulted a loss in activity (2.6 %). In another construct, 3A-LOOP, bases from the large loop were removed to

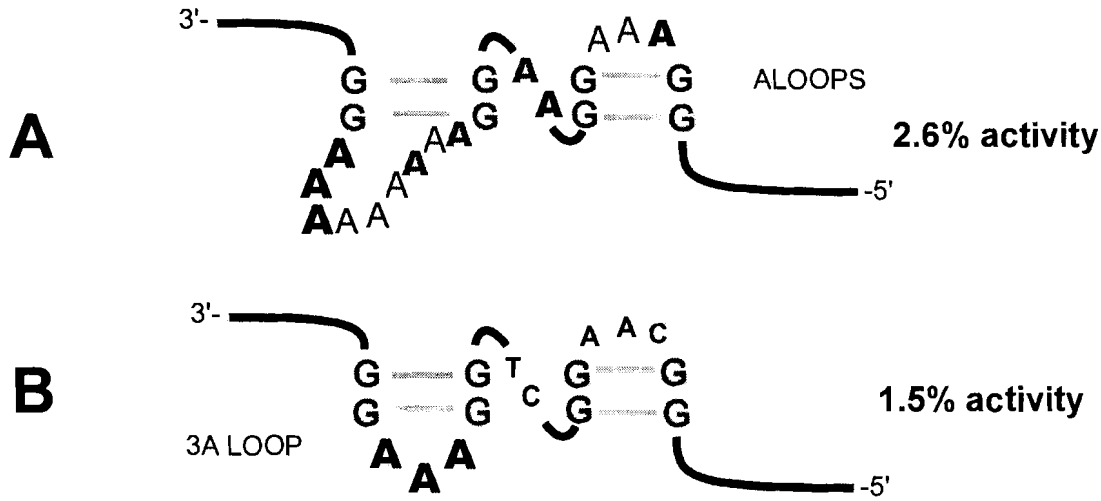


Figure 3-18 Changing the sequence of the loops.
A) ALOOPS construct results from changing the loop sequences to all adenines. This reduces catalytic activity by 97 %. B) 3A LOOP construct shortens the large loop by the removal of all but three nucleotides. Activity is reduced by 98.5 % to background levels.

leave a three adenine loop. This also resulted in a loss in activity (1.5 %), indicating that these bases within the loops were crucial to the catalytic activity of the DNAzyme.

Since it was predicted that the central structure contained a two-layered quartet, the effect of having an additional layer of guanine quartets was also investigated. Construct UV1C.G4 was designed, shown in Figure 3-19, to add one additional layer of quartets. This construct exhibited negligible activity at 1.6 %. It is possible that the DNAzyme may have not folded properly and/or had different folding topology than that of the "wild-type" sequence.

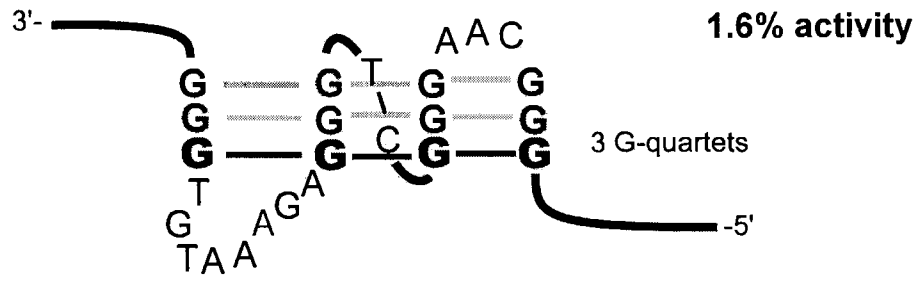


Figure 3-19 One extra layer of guanines added to theoretically form a three-layered quartet. Activity is reduced down to 1.6 % relative to UV1C.

3.3.5 Summary and Towards a Revised Model

The series of experiments explained in this chapter have provided a better understanding of the sequence requirements and folding characteristics of the UV1C DNAzyme. The deletion mutant constructs indicated that many nucleotides from the 3'end of the DNAzyme could be removed without greatly sacrificing activity. The 5'-end of the enzyme appeared to involve base-pairing interactions with the substrate strand and fully complementary sequences could be engineered into the enzyme. From crosslinking data, a number of bases located at or near the quadruplex core were found to

interact with the thymine residues of the substrate. These crosslinks also identify the bases located in the "active site" of the enzyme. The sequences of the "large loop" appeared to play an important role in catalysis as well.

A new model can be constructed based on this data, shown in Figure 3-20. In this model, the thymine dimer is positioned in close proximity to the quartet, making close contacts with those sequences identified from the crosslinking information, shown in green. Also, due to the loop orientations of the quadruplex, the 5' and 3' flanking sequences are situated on opposites side of the quartet, which differs from the original model. The guanine quartet is depicted in a horizontal fashion, which may provide for continuous stacking with the standard base paired sequences on either side of the quartet. It is unclear at this moment how the 5'end of the substrate (shown in red) would be capable of base pairing with the 3'end of the enzyme if it located on the opposite side of the quartet. Such an interaction may require high torsional angles, and the possibility of base-flipping of the thymine dimer out from the strand onto the quartet. The

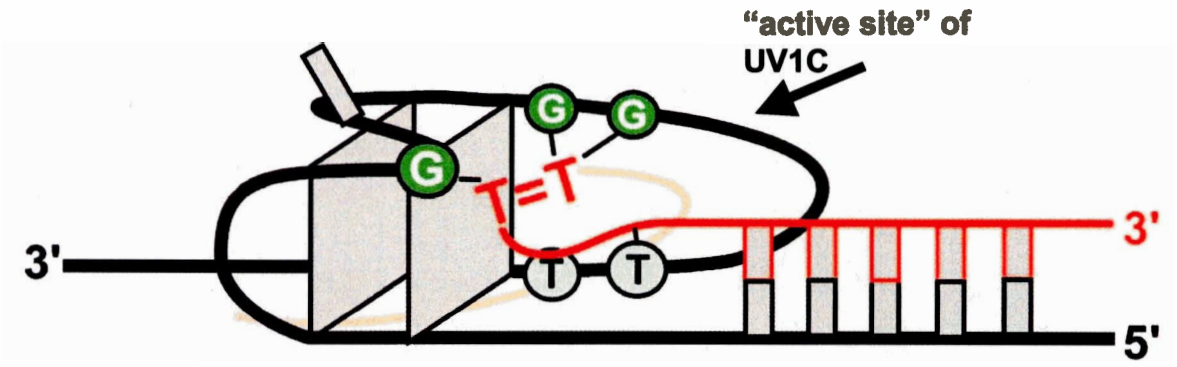


Figure 3-20 Model for interaction of UV1C DNAzyme with TDP substrate.

crosslinking of thymine 20, located in the diagonal loop, cannot be accounted for by this model, although chemical probing with permanganate indicated a high degree of solvent

accessibility. A high degree of tertiary interactions may be present that may fold UV1C into a very compact structure. A recent crystal structure of a naturally occurring photolyase bound to DNA has revealed that many contacts are formed between the DNA double-helical substrate and amino acid side chains of the enzyme (Komori *et al.*, 2001; Mees *et al.*, 2004; Tamada *et al.*, 1997). Hydrogen bonds and salt bridges are formed between photolyase and phosphate backbone of nucleotides neighbouring the thymine dimer in the DNA. These interactions help to flip out the thymine dimer from the duplex into the active site of enzyme, resulting in a DNA bend in the range of 30° (Husain *et al.*, 1988) to 50° (Mees *et al.*, 2004). Also, a number of amino acid side chains in the active site hydrogen bond with the C4-carbonyl and N3-imide groups of the thymines to stabilise the flipped out conformation.

In the DNAzyme complex, the TDP substrate is single stranded, and thus could accommodate a higher degree of bending. It would be tempting to suggest the possibility that the thymine dimer may be flipped out in this system as well, although there is no evidence to suggest this. Future work on the structure of UV1C will concentrate on obtaining high-resolution structures to determine precisely the interactions of UV1C with its substrate. We are currently in the process of crystallizing the UV1C DNAzyme when folded with the TDP substrate. These efforts will hopefully provide us with a better understanding as to the structure and function of UV1C, and perhaps reveal specific contacts in the active site to the thymine dimer bases. In the next chapter, I will address the specificity of UV1C for its substrate, and discuss preliminary data concerning the photochemical mechanism of UV1C.

CHAPTER 4

SUBSTRATE SPECIFICITY AND PHOTOCHEMICAL MECHANISM

4.1 Introduction

In the previous chapter, it was determined that the thymine dimer was in close proximity to the guanine quadruplex and made close contacts to sequences contained in the loops. In this chapter, I will investigate the substrate specificity and the photochemical mechanism of UV1C. In addition to thymine dimers, *E.coli* photolyase has been shown to catalyse the repair of other types pyrimidine dimers (Kim and Sancar, 1991). These include the dimers of uracil, chimeric uracil-thymine, and to a lesser extent, cytosine. It was also found to repair uracil dimers within poly(rU) sequences. I wanted to determine whether UV1C could also repair different types of pyrimidine dimers. I will look at the importance of sugar type (deoxyribose versus ribose), pyrimidine dimer base, and internucleotide phosphate on enzyme activity. The importance of having RNA or DNA as flanking sequences to the thymine dimer will also be examined. This data will be used to formulate a hierarchy for substrate preference. Also, some experiments alluding to the photochemical mechanism will be discussed.

4.2 Materials and Methods

4.2.1 DNA Oligonucleotide Constructs

DNA oligonucleotides and ribothymidine constructs were ordered from University Core DNA Synthesis Lab. RNA oligonucleotides were purchased from Dharmacon. (UCDNA) at the University of Calgary. The TDP substrate had the sequence 5'-AGGAT CTACA TGTAT TGTGT GCGTA CGAGT ATATG-3'. Ribothymidine constructs were purchased from UCDNA. Deoxyuridine constructs were synthesized using 5'-dimethoxytrityl -2'-deoxyuridine, 3'-[(2-cyanoethyl)-(N, N-diisopropyl)]-phosphoramidite by UCDNA. DNA was end-labelled using T4 polynucleotide kinase from Invitrogen.

4.2.2 Synthesis of Thymine Dimer-containing Oligonucleotide

Synthesis of the TDP substrate was performed as described in section 2.2.2. Unlabeled T2 and 5'-^{32}P-labeled T1 oligonucleotides were annealed at 400 μM concentrations to a complementary splint oligonucleotide in buffered 40 mM $MgCl_2$. Solutions were degassed by freezing and thawing under vacuum three times. Photocrosslinking was performed in 5 mM of the triplet sensitizer acetophenone and 5% acetone. In the case of uracil-containing substrates, 5 mM acetone was used as a triplet sensitizer due to the higher triplet energy required for formation. The DNA was irradiated for 30 minute in a 96-well ELISA plate with a transilluminator with a polystyrene filter to cut out wavelengths shorter than 300 nm. The TDP construct was ethanol precipitated and size-purified on 12% denaturing polyacrylamide gels.

4.2.3 Kinetic Analysis

Reactions contained 8 μM UV1C DNA enzyme in the presence of trace amounts of ^{32}P-end-labelled substrate. DNA was heated to 90° C for 1 minute and allowed to cool slowly to room temperature in the presence of 50 mM Tris-Cl, pH 8.0, 200 mM NaCl. 50 μL reaction mixture was placed into a flat-bottom 96-well polystyrene ELISA plate (the surface pre-blocked with glycogen) and irradiated on a Fotodyne transilluminator with > 300 nm light at a photon flux of 3.4×10^{-9} einsteins min^{-1}. 5 μL aliquots were removed at set time points and separated on 12% denaturing polyacrylamide gels. Fragments were quantitated using Molecular Dynamics ImageQuant software.

4.2.4 Chemical Probing

For methylation reactions, a 1.2% dimethyl sulfate (DMS) solution in water was freshly prepared and 2 μL added to 10 μL of ^{32}P 5'-end labelled DNA folded in reaction buffer (50 mM Tris-Cl, pH.8, 200 mM NaCl). For diethyl pyrocarbonate (DEPC) reactions, 5 μL of 3% DEPC was added to 10 μL of a DNA solution in the same buffer. Both reactions were allowed to proceed for 30 minutes at room temperature, and terminated by addition of 7% β-mercaptoethanol, followed by ethanol precipitation. The recovered DNA was dissolved in 10% piperidine and heated for 30 minutes at 90° C, followed by lyophilization. The samples were then dissolved in denaturing gel loading buffer (containing 90% formamide v/v and 10 mM EDTA, pH 8.0) and analyzed in 12% denaturing polyacrylamide gels.

4.2.5 Iridium (IV) Chloride Oxidation Experiments

Reactions involving iridium (IV) chloride treatment contained ^{32}P-labelled DNAzyme (5 uM) with 5 uM TDP substrate were folded in a buffer containing 50 mM Tris-Cl, pH 8, 200 mM NaCl DNA. Enzyme-substrate reaction mixtures were irradiated for one hour in a 96-well ELISA plate using a transilluminator with a maximum wavelength output at 300 nm. The reactions were ethanol precipitated and resuspended in 100 uL of 10% piperidine. The piperidine solutions were heated to 90° C for 30 minutes and lyophilised. The DNA pellets were resuspended in denaturing dye and loaded with equal radioactive counts onto a 12% denaturing polyacrylamide gel. The gel was exposed to phosphorimager screen and scanned with a Molecular Dynamics Typhoon Phosphorimager.

4.3 Results

4.3.1 SECTION 1: Substrate Specificity

In this section, I used a series of substrate constructs designed to dissect the substrate specificity of UV1C. These substrates contain: uracil, thymine, or cytosine dimers, and a mixture of T=U or U=T dimers. Constructs containing deoxyribose, or ribose sugars will also be examined.

4.3.1.1 Uracil Dimer Substrate

An all-RNA version of TDP was synthesised to determine whether UV1C could repair an RNA, shown in Figure 4-1. It was found that RNA was a very poor substrate for UV1C, shown in Figure 4-2. The rate of repair of RDP ($2.3 \times 10^{-4} \pm 1.0 \times 10^{-5}$ min^{-1}) was

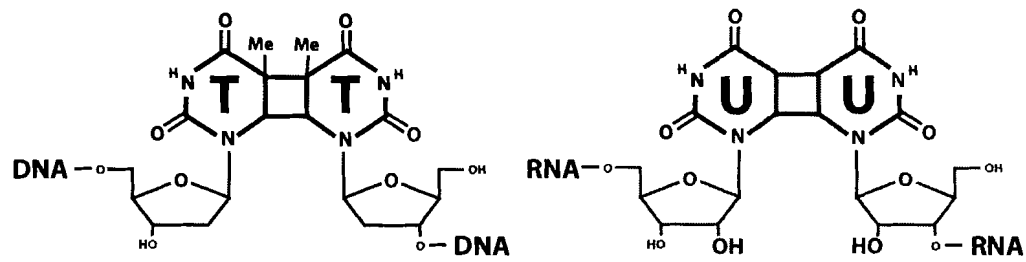


Figure 4-1 Thymine and uracil dimer substrates

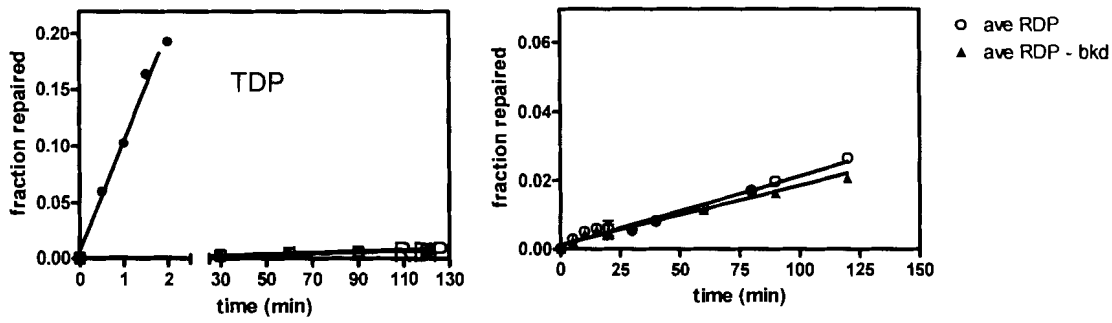


Figure 4-2 Poor repair of an all-RNA (RDP) substrate.
(Left) Rate of repair for RDP was $2.3 \times 10^{-4} \pm 1.0 \times 10^{-5}$ min^{-1} similar to the background rate of $2.1 \times 10^{-4} \pm 1.1 \times 10^{-5}$ min^{-1}. Rate of repair by of TDP substrate was 0.11 ± 0.01 min^{-1}. (Right) Rates of photoreversal for the background and UV1C-present reactions. All kinetics were performed at wavelengths > 300 nm at 3.4×10^{-9} einsteins min^{-1} illumination.

negligible when compared to that of the normal thymine dimer substrate (0.11 ± 0.01 min^{-1}). Furthermore, UV1C was not able to accelerate the background rate for RDP repair ($2.1 \times 10^{-4} \pm 1.1 \times 10^{-5}$ min^{-1}), as is shown in Figure 4-2. Many factors could contribute to inefficient repair in this case, such as the RNA flanking sequences, the ribose sugar of the uracil dimer substrate, even the uracil dimer itself. The RDP substrate was dissected to determine which of these factors contribute of lack of repair.

4.3.1.2 Effect of Flanking Sequence on U=U Dimer Repair

To investigate the importance of the RNA sequences located on either side of the uracil substrate, three constructs were designed to replace the RNA sequence immediately 3' and/or immediately 5' to the uracil dimer substrate. These are shown in Figure 4-3. As a naming convention for the chimeric substrates, "R" signifies an all RNA sequence, including the uracil involved in dimer formation. "U" denotes a DNA sequence that has a ribouridine that forms the dimer. Combinations of these oligonucleotides result in substrates that all contain ribouridine dimers, but with different flanking nucleotides. For example, the R-U construct contained DNA sequence

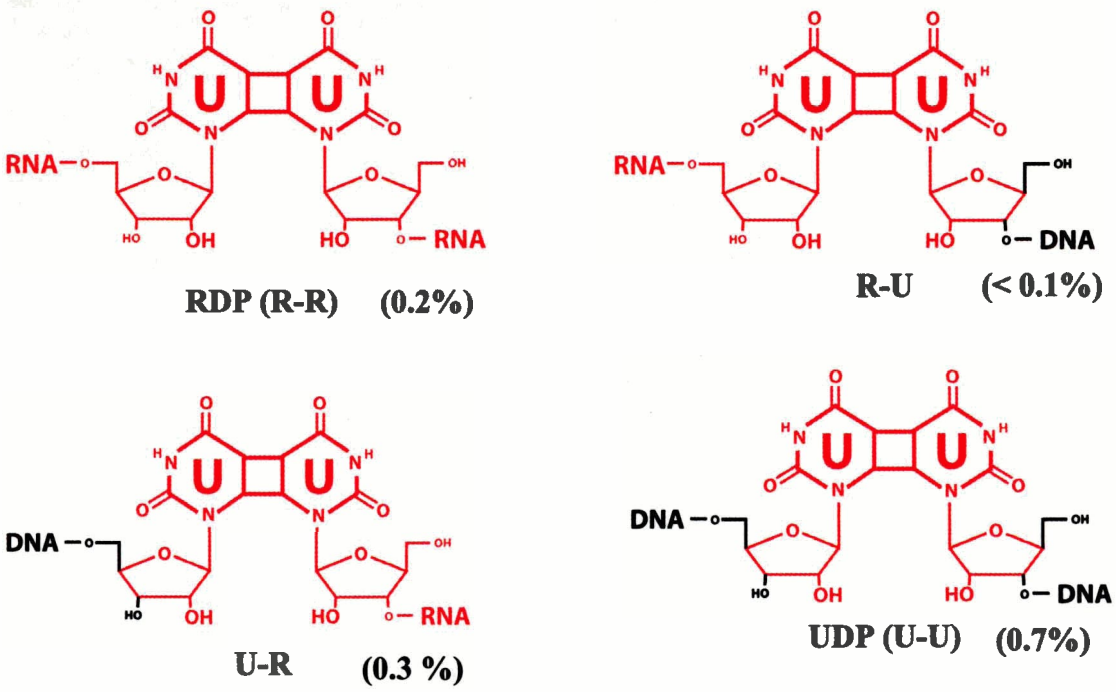


Figure 4-3 Uracil dimer-containing substrates with DNA and/or RNA flanking sequence. The percent relative rates of repair compared to the TDP substrate are shown in brackets.

immediately 3' to the uracil dimer in an otherwise all-RNA substrate. U-R substrate contained DNA sequence immediately 5' to the uracil dimer. The U-U contains a uracil

dimer with ribose sugars with DNA sequence on both sides. When tested for activity, it was found that UV1C could not efficiently repair ribouridine dimers, whether or not the 3' or 5' sequence contained DNA or RNA (shown as percent relative rates in Figure 4-4). Further analysis of these results indicated that the UDP construct displayed the highest rate enhancement over the background, shown in Figure 4-4.

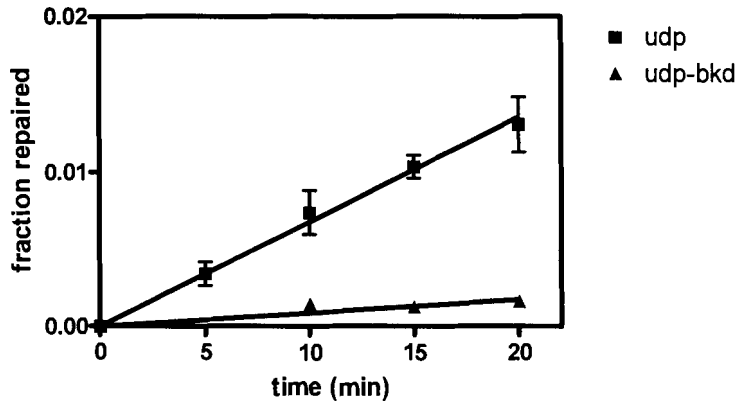


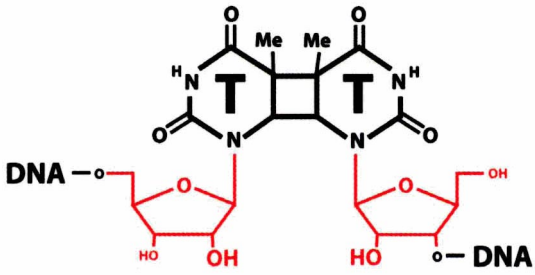
Figure 4-4 Rates of repair of ribo-uridine dimers with DNA flanking sequences (UDP). Data points are a result of triplicate kinetic runs. Rate of repair of UDP substrate was $6.8 \times 10^{-4} \pm 3.2 \times 10^{-5}$ min^{-1}. The background rate of repair was $8.8 \times 10^{-5} \pm 1.1 \times 10^{-5}$ min^{-1}.

It was evident that UV1C could repair ribouridine dimers, although not very efficiently.

4.3.1.3 The Importance of Sugar Ring Type on Catalytic Activity

The UDP substrate contained a uracil dimer with ribose sugar groups in an otherwise all-DNA substrate. The great difference in activity of the normal TDP substrate and the UDP substrate (consisting of DNA-UU-DNA and DNA-TT-DNA respectively), could be due to two factors. One is the presence of the ribose sugar versus 2'-deoxyribose sugar. The other is the type of dimer that is being repaired, i.e. thymine

versus uracil dimers. It is possible that either one or both factors contribute to the lowered activity of the repair. To further investigate this, a construct was synthesised which contained ribothymidine dimers. In such a construct, the substrate is identical the regular TDP substrate, including the thymine dimer, with the exception that the sugars are now ribose sugars as opposed to deoxyribose. This is shown in Figure 4-5.



rTrT (2.7 %)

Figure 4-5 Ribo-thymidine construct (rTrT) and relative rate of repair by UV1C.

Although the rate of repair for rTrT was four-fold faster than the UDP construct, it was still relatively slow (3.0×10^{-3} min^{-1}) when compared to the regular substrate. The presence of the two ribose sugars impaired the catalytic activity even though the substrate contained thymine dimers. To further investigate this, chimeric constructs were made where instead of both positions containing ribose sugars, either the 5' or the 3' sugar of the thymine dimer contained the ribose sugar (Figure 4-6). These constructs are changed from the normal TDP substrate essentially by only one ribose sugar.

When tested for activity, UV1C repaired the dTrT construct where the 5' thymine contained a deoxyribose ring with a greater than three-fold increase in activity (9.7×10^{-3}

min^{-1}) with respect to rTrT, shown in Figure 4-6. When the rTdT substrate was tested, an even greater increase in activity over the rTrT substrate was observed (1.6×10^{-2} min^{-1} corresponding to a five-fold increase). These results reveal that UV1C is very sensitive to the sugar moiety attached to the dimer. Also, that the presence of a 2' hydroxyl on the sugar of the 3' thymine has a larger effect on rate than on the 5' sugar. The reduced rate observed with the rTrT substrate was contributed by both ribose sugars on the thymine dimers.

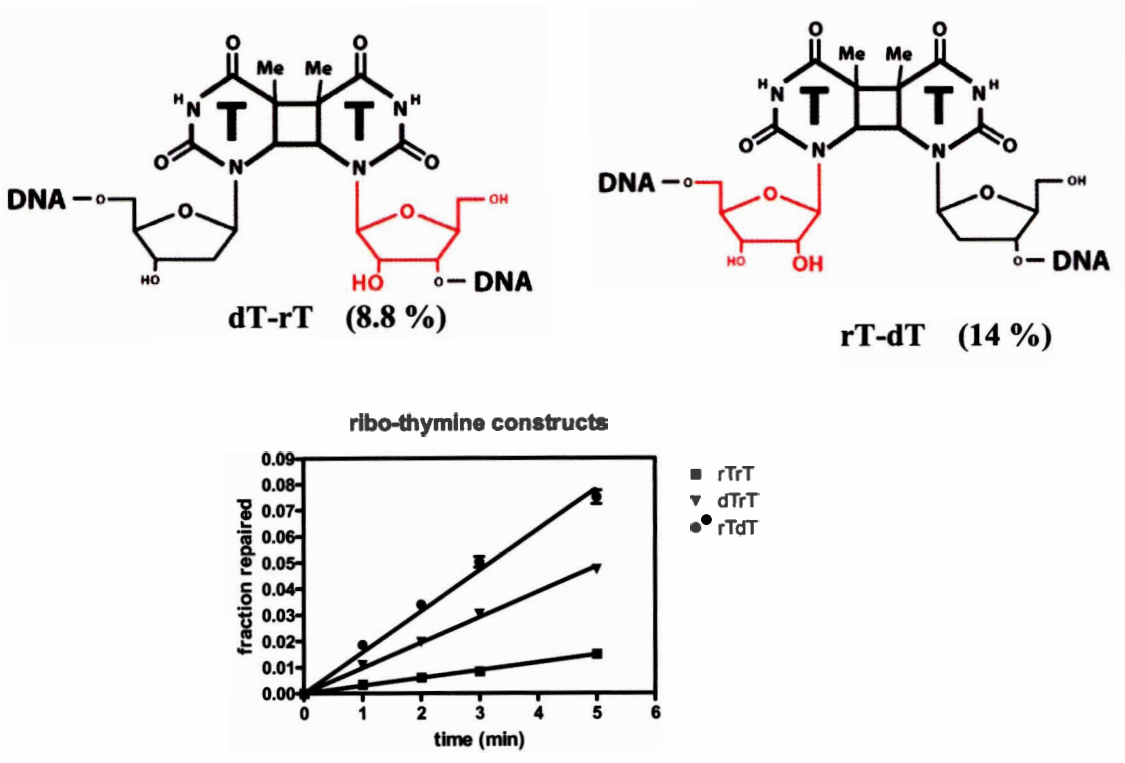


Figure 4-6 Chimeric ribo-deoxyribothymidine containing substrates.
(Bottom) Rates of repair as a result of sugar pentose moiety. Repair rates of were 3.0×10^{-3} min^{-1} for rTrT, 9.7×10^{-3} min^{-1} for dTrT, and 1.6×10^{-2} min^{-1} for rTdT.

4.3.1.4 The Importance of Uracil Dimers on Catalytic Activity

Next, I examined the importance of the type of pyrimidine dimer repaired when attached to deoxyribose sugars. A construct containing a deoxyuracil dimer was synthesized, named dUdU. In this construct, only the substrate dimer bases were different than the regular substrate. When tested for activity, UV1C was found to repair dUdU with appreciable efficiency (20%), which is comparable to that of the regular TDP substrate. This is shown in Figure 4-7.

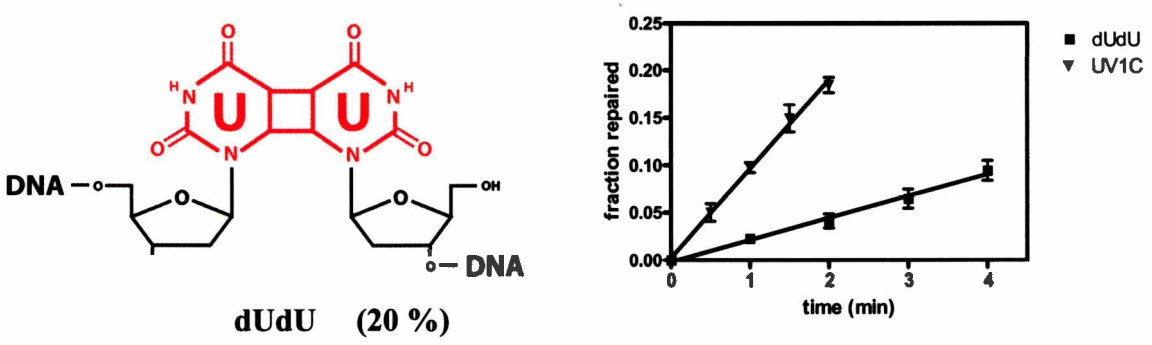


Figure 4-7 Deoxyuridine dimer substrate construct and repair rate.
Subtrate construct dUdU (left) contains all deoxyribose sugars and uracil dimers. The rate of repair is 0.023 min^{-1}, 20 % of that of the TDP (all DNA) substrate.

Next, I wanted to determine how well chimeric deoxy- T<>U and U<>T substrate dimers could be repaired by UV1C. Chimeric substrate constructs were tested containing mixed deoxyuridine and deoxythymidine dimers. The constructs dUdT and dTdU are shown in Figure 4-8. It was found that both of the chimeric substrates were repaired much faster than the dUdU substrate (0.056 min^{-1}). Interestingly, both constructs were

repaired at roughly the same rates with respect to each other. These rates are only 50 % of that of the normal TDP substrate.

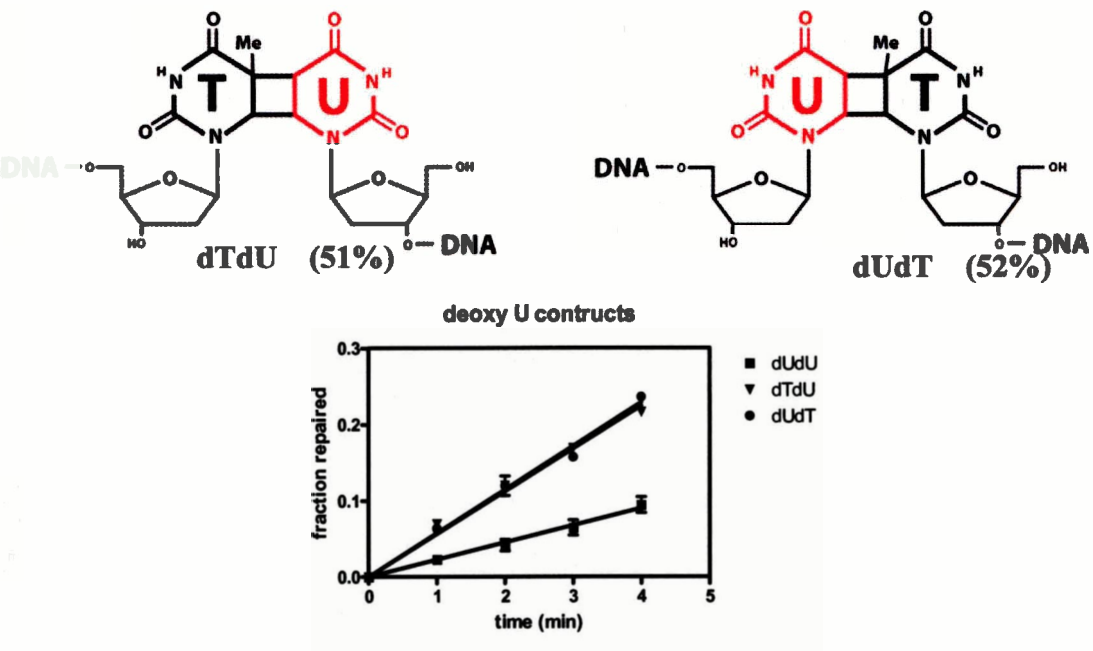
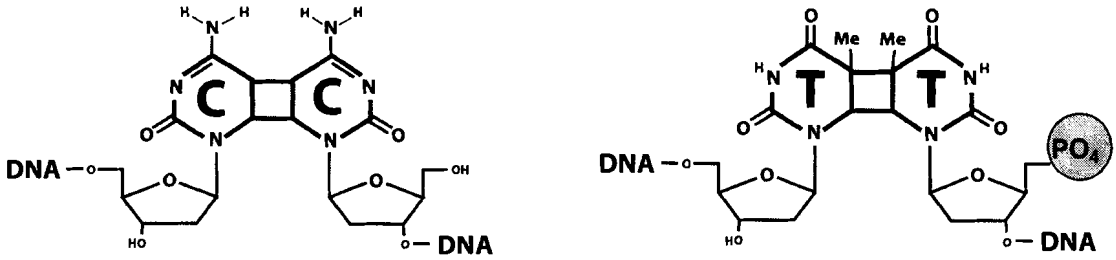


Figure 4-8 Deoxy-uridine substrates dTdU and dUdT.
(Bottom) Rates of repair for the chimeric substrates. dUdU = 0.023 min-1, dTdU = 0.056 min-1, dUdT = 0.057 min-1.

I also tested the ability of UV1C to repair cytosine dimers, and thymine dimers containing an internal phosphate but without the 5'-3' linkage, shown in Figure 4-9. The formation of the cyclobutane-type dimer was confirmed by quantitative photoreversal of the dimer by irradiation at 254 nm. It was found that deoxycytosine dimers were a very poor substrate for UV1C (rate = 0.9 x 10^{-5} min^{-1}). Curiously, the addition of a phosphate group to the 5' end of the 3' thymine dimer nucleotide resulted in a loss in activity (2.1 x 10^{-4} min^{-1}, corresponding to 2% activity).



cytosine dimers (0.1 %) **internal phosphate (2 %)**

Figure 4-9 Cytosine dimer substrate.
Cytosine dimers were repaired at a rate of 0.9 x 10^{-5} min^{-1}. The internal phosphate construct was repaired at 2.1 x 10^{-4} min^{-1}.

From the above studies, it was clear that UV1C was sensitive to both the type of sugar group and the type of pyrimidine dimer found in the substrate strand. The presence of a ribose sugar ring on the pyrimidine dimer appeared to have the greatest effect on the 3' sugar. It is possible that specific contacts were made to the sugar ring that were disrupted by the presence of a 2' hydroxyl group. The addition of a phosphate to the 5' position on the same 3' sugar position may have also caused a degree of hindrance. Also, UV1C appeared to be sensitive to the type of pyrimidine dimer present on the substrate. Based on the above studies, a hierarchy of the preferences of the substrate components can be made. In terms of the nature of the pyrimidine dimers, the efficiency of repair can be ordered as follows: T<>T > (U<>T = T<>U) > U<>U >> C<>C. The preference of the sugar moieties on the pyrimidine dimer are: deoxy—deoxy > ribose—deoxy > deoxy —ribose >> ribose—ribose. Similar studies of the *E. coli* photolyase have shown the a similar preference in substrate pyrimidine dimers (Kim and Sancar, 1991). In their work, they correlated binding data obtained from gel mobility

shift assays with kinetic data from photolysis experiments. They were able to determine that the loss in activity seen with uracil dimer repair was a result of lower binding affinity to the enzyme. But once the uracil dimer was bound, the efficiency of repair was equal to that of thymine dimers. Furthermore, they had found that thymine dimer substrates lacking the internal phosphate backbone were not repaired, nor bound efficiently. In our system, the thymine dimer substrate always lacks the internal phosphodiester bond, and appeared not to repair when it is present. It is difficult to determine binding affinities of the substrate to enzyme in this system directly since the enzyme binds to the substrate via a number of base pairs. This could overshadow any small measurable changes in binding affinity due to the pyrimidine dimer substrate. Also, reactions were performed under single turnover conditions so information, such as K_M, cannot be determined. What can be concluded from this section is that UV1C is highly specific for the thymidine dimer substrate, and is sensitive to the absence or presence of hydroxyl or methyl groups located on the thymine dimer substrate.

4.3.2 SECTION 2: Photochemical Mechanism

In this section, I demonstrate some preliminary experiments that may suggest the photochemical mechanism of repair by UV1C. As mentioned in Chapter 2, it was proposed that a photoexcited quadruplex would donate an electron to the thymine dimer, thus facilitating its repair. The issues regarding the nature of the excited state (singlet or triplet), energy transfer and electron transfer will be discussed.

4.3.2.1 Excited Singlet versus Triplet States

Photoexcitation of ground state species leads to the formation of excited states. This results in the promotion of an electron from the ground state to a higher energy excited singlet state. From here, there are many processes that can occur. The excited electron can move back into the ground state and emit energy in the form of fluorescence. Alternatively, the electron can undergo "intersystem crossing" where the electron changes its spin orientation to become parallel to the electron that it was formerly paired with. This results in a lower energy triplet state. From this triplet state, the electron can then relax back to the ground state and emit energy in the form of phosphorescence.

The presence of oxygen in a photochemical reaction is known to quench the excited triplet state (Wilson *et al.*, 1996). Quenching is the process whereby an excited molecule becomes deactivated by an external component. If the rate of a photochemical reaction is found to increase when oxygen is removed from the reaction and slowed down in oxygen saturated solutions, this is usually indicative of an excited triplet state. If no change is observed in either oxygenated or deoxygenated reactions, this may indicate a mechanism involving the excited singlet state.

Photoreversal reactions were carried out in either deoxygenated solutions, using a freeze-vacuum thaw method followed by nitrogen purging, or under oxygenated conditions, where solutions were purged with oxygen. The rates of repair were compared to standard (semi-oxygenated) reaction conditions. The results, shown in Figure 4-10, indicate that the photoreversal reaction by UV1C was independent of the presence or absence of oxygen in the solutions. Although a single experiment is never definitive proof, this result points to an excited singlet state mechanism.

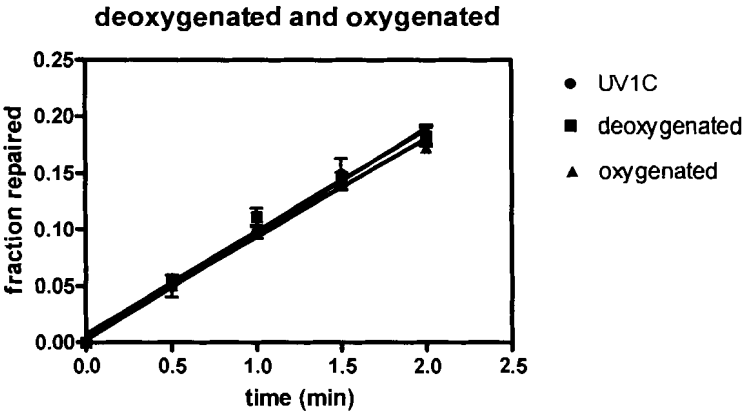


Figure 4-10 The effect of oxygenated, deoxygenated and semi-oxygenated conditions. No significant differences in rate are observed (k = 0.11 min^{-1}) with respect to standard semi-oxygenated conditions.

4.3.2.2 Energy Transfer versus Electron Transfer

Thymine dimers can be split by [2 + 2] cycloreversion with far ultraviolet light (240 nm) (Sancar, 2003). Since the DNAzyme uses ultraviolet light to repair thymine dimers, the possibility of energy transfer from the photoexcited guanine quartet to the thymine dimers must first be addressed. In an energy transfer process, a molecule is excited to the first excited singlet state. This excited donor (D*) can then undergo energy transfer to excite the acceptor molecule (A*) by either electron exchange, or dipole-dipole interactions (Kavarnos, 1993). For this process to occur, there must be enough free energy in the system. The energy required to excite the thymine dimers to promote a [2 + 2] cycloreversion is estimated to be 500 kJ mol^{-1} (Sancar, 2003). The UV1C DNAzyme uses 305 nm light most efficiently, which corresponds to an energy of 393 kJ mol^{-1}. This results in an energy gap of +107 kJ mol^{-1}, and thus does not contain enough energy to excite the thymine dimers. In the photosplitting by UV1C, thymine dimers are

not excited in the reaction. In the photolyase system photosplitting is accomplished by electron donation by a photoexcited flavin. It is feasible that the photoexcited guanine quadruplex of UV1C photorepairs thymine dimers by a similar charge transfer mechanism

4.3.2.3 The Possibility of Electron Donation from the Guanine Quadruplex

In order for charge transfer to occur, the free energy change (ΔG_{ct}) has to be exergonic, or < 21 kJ/mol endergonic (Scannell *et al.*, 1997). The free energy can be determined from the ground-state oxidation potential of the donor (E_{ox}, in V), the reduction potential of the acceptor, (E_{red}, in V) and the excited-state energy of the sensitizers (E_{00} in eV) based on the Weller equation (Kavarnos, 1993). The equation is as follows:

$$\Delta G_{ct} = [E_{ox} - E_{red}] - E_{00} \qquad \text{(eq-1)}$$

Based on available literature values of the reduction/oxidation potentials and excited-state measurements of DNA bases, approximate calculations can be made to determine if charge transfer can occur in the DNAzyme photorepair process. In our system, photoexcited guanine is proposed to be oxidized and donates an electron to the thymine dimer. Unfortunately, the redox potentials for guanines within a guanine-quadruplex have not been directly measured. Measurements of free nucleotides indicate that guanine is the most easily oxidized base with an oxidation potential of 1.34 V versus SCE (Standard Calomel Electrode) (Steenken and Jovanovic, 1997). Thymine, on the other hand, is the most easily reduced base, with a reduction potential of –2.18 vs. NHE (Normal Hydrogen Electrode) (Seidel *et al.*, 1996). The other bases follow with –2.35 for cytosine, -2.52 for adenine and –2.76 for guanine (vs. NHE) (Seidel *et al.*, 1996). The

reduction potential of thymine dimers within the context of DNA has not been directly measured, but is estimated to be –2.2 V versus SCE from direct electrochemical measurements (Scannell *et al.*, 1997). These values have been extensively applied to predict charge transfer processes through double-stranded DNA (Dandliker *et al.*, 1998; Ito and Rokita, 2004; Wagenknecht, 2003). Literature values for the excited state energies of free guanine have been estimated to be 3.8 – 4.01 eV from fluorescence measurements (Preuss *et al.*, 2004; Wang, 1976).

In our system, the guanines within the quadruplex are photoexcited at 305 nm, separate from that of normal base paired DNA bases, and it is presumed that these guanines are more oxidizable than standard base-paired guanines. Assuming that 305 nm wavelength energy roughly corresponds to the E_{00} excitation energy, a value of 4.07 eV can be tentatively attributed to the excited state energy of the quadruplex. Substitution of these values into the free energy equation results in $\Delta G_{ct} = [1.34 - (-2.2)] - 4.07 = -0.53$ eV corresponding to –51.0 kJ mol^{-1}, which is an exothermal process. Thus, the donation of an electron from a photoexcited guanine(s) within the quartet to the thymine dimer is thermodynamically possible. The excited state oxidation potential of the guanine quadruplex can also be estimated from: $E_{ox}^{*} = [E_{ox} - E_{00}]$ (Kavarnos, 1993) which would be –2.73 V vs. SCE, which has enough driving force to reduce thymine dimers (E_{ox} = -2.2 V vs. SCE). Interestingly, this value is comparable to the estimated excited state reduction potential for flavin estimated to be –3.0 to –2.8 V vs. SCE (Scannell *et al.*, 1997; Seidel *et al.*, 1996; Yeh and Falvey, 1992).

4.3.2.4 Inosine Mutants

Inosine, shown in Figure 4-11, is a nucleobase that is structurally very similar to guanine but lacks the amino group at the carbon-2 position. Despite this, they have still been found to form quadruplex structures with other inosine or guanine bases (Smith *et al.*, 1995; Sponer *et al.*, 1996). An important property of inosine is that it is less

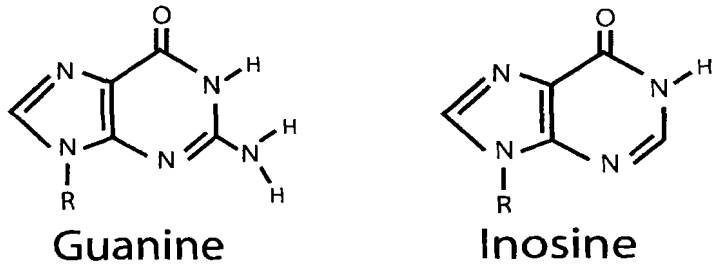


Figure 4-11 Chemical Structures of guanine and inosine.

easily oxidizable than guanine due to its oxidation potential (1.5 V vs. SCE) (Hush and Cheung, 1975; Steenken and Jovanovic, 1997). If inosine were substituted for guanine in the putative guanine quartet of UV1C, a change in catalytic rate may be attributed to the change in the oxidation potential of the quadruplex. Making an assumption that the excited state energy of inosine is similar to that of guanine, an excited state oxidation potential could be approximated at –2.57 V using the Weller equation. Falvey and coworkers determined that an excited state potential of at least –2.4 V is necessary for thymine dimer photoreversal (Scannell *et al.*, 1997), using sensitizers that had excited state oxidation potentials varying from –2.08 V to –3.25 V vs. SCE. Thus, inosine substitution within the quartet of UV1C might impair its ability to repair the thymine dimer.

We used inosine mutants to substitute single guanines within the guanine quartet core, shown in Figure 4-12. When tested for activity, it was found that two inosine mutant constructs displayed the most significant reduction in repair rates. These were I-12 and I-13 constructs, which corresponded to the 5'-most guanines in the guanine-quartet. Curiously, these guanines were not in close contact with the thymine dimer

Figure 4-12 Inosine Mutant Constructs and their relative rates.
The highlighted sequences indicate the greatest loss in activity.

from the iodouracil crosslinking data. As expected, replacement of guanine 23 (I-23), previously determined not to participate in quadruplex formation, did not affect activity (98%). All but one position (I-32) appeared to contribute significantly to photorepair. These positions are illustrated relative to the revised model shown in Figure 4-13. As an added control difference spectroscopy confirmed the presence of guanine quadruplexes in

the DNA (data not shown). These results might suggest that many of the guanines within the quartet may cooperatively help in donating an electron to the thymine dimer.

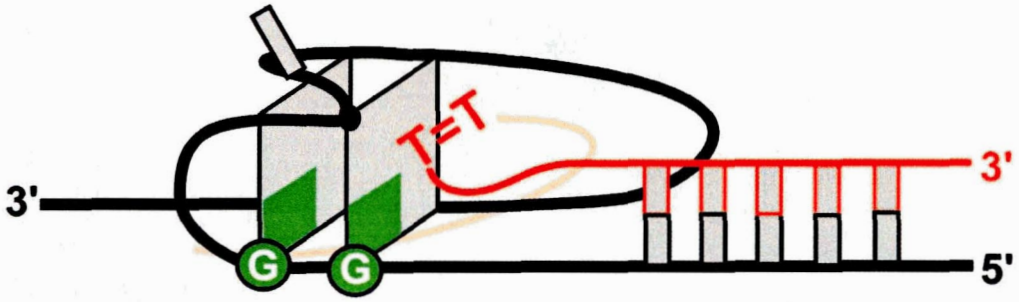


Figure 4-13 Positions of Inosine Substitutions Leading to Decreased Photorepair.

4.3.2.5 Oxidative Damage-KMnO4 - Iridium (IV) Chloride

The effects of oxidative damage to DNA has been well studied [reviewed in (Burrows and Muller, 1998)]. These can lead to oxidised guanines in the form of 8-oxoguanine, shown in Figure 4-14. This can then be further oxidised, resulting in strand breakage that can be detected by piperidine treatment. In our system, thymine dimers may be repaired by photoexcited guanine residues in a quadruplex. Since electron

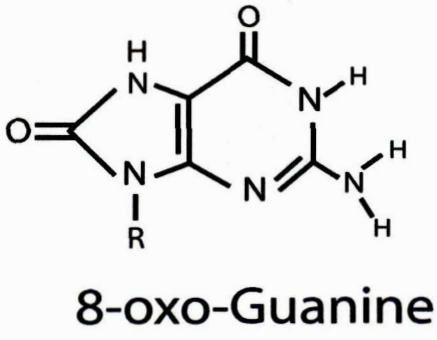
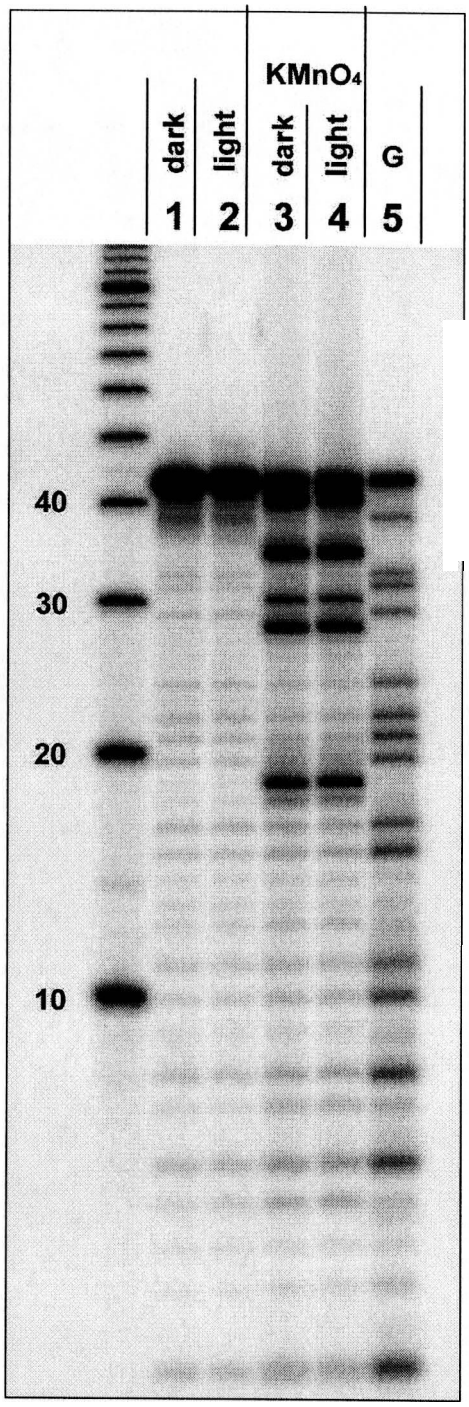


Figure 4-14 The Structure of 8-oxo-Guanine.

donation from guanine residues could result in permanent guanine damage, it would be reasonable to expect 8-oxoguanine, or a similar oxidized guanine species, would form. I wished to investigate if it were possible to detect oxidized guanines that may form from the irradiation of the UV1C DNAzyme with the substrate to stimulate photorepair. Although 8-oxoguanine itself is suspected not to be particularly susceptible to piperidine treatment, it is possible to further oxidise oxoguanine to piperidine-sensitive products with chemical oxidants such as $KMnO_4$ and iridium (IV) chloride compounds (Muller *et al.*, 1998). These can then be detected on a polyacrylamide gel.

After the photoreversal reaction of UV1C in the presence of excess TDP substrate, the DNAzyme was treated with either $KMnO_4$ or iridium (IV) chloride. The DNA was then piperidine treated and size separated on a denaturing polyacrylamide gel. The results are shown in Figure 4-15. No oxidized products are seen with piperidine treatment of photoirradiated UV1C without further oxidation (lane 2). Curiously, oxidation by $KMnO_4$ does not result in detection of oxidized guanine residues (lane 4). As well, when photoirradiated UV1C is treated with iridium chloride followed by piperidine treatment, no detectable damage to guanines were observed, shown in Figure 4-16. Although this could be evidence that guanines are not oxidized in the reaction, it is also possible that there is a fast rate of back electron transfer to the guanine residues, which would prevent oxidation from occurring.

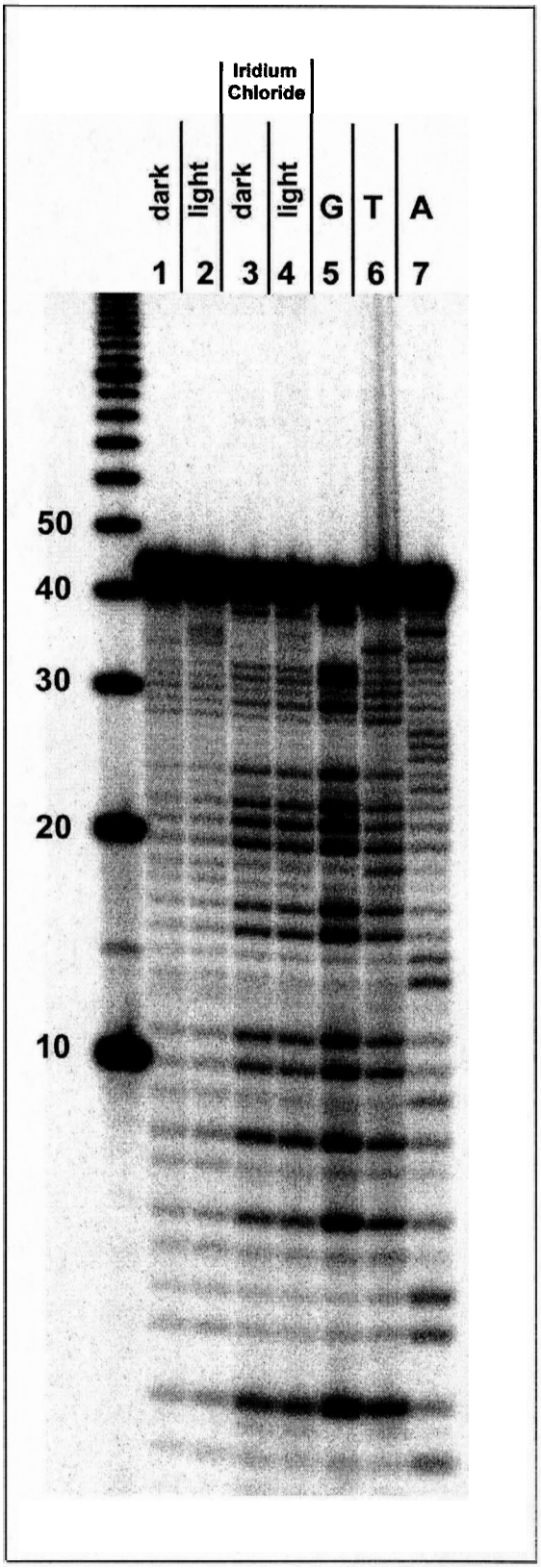
Figure 4-15 Potassium permanganate treatment of photoirradiated end-labelled UV1C DNAzyme. Lane 1: Control dark reaction of UV1C piperidine treated only. Lane 2: Piperidine-treated UV1C following photoirradiation at 300 nm in the presence of TDP substrate. Lane 3: Control reaction of unirradiated UV1C following $KMnO_4$ modification and piperidine treatment. Lane 4: Irradiated UV1C following KMnO4 modification and piperidine treatment. Lane 5: Guanine marker lane.


KMnO4
dark
light
dark
light
G
1
2
3
4
5
40
30
20
10

The excited singlet states of DNA bases are very short-lived, caused by an extremely rapid internal conversion ($k_{ic} = 10^{12}$ s^{-1}) and fluorescent lifetime that is in the picosecond range (Pecourt *et al.*, 2001; Scaiano, 1989). Also, the rates of charge transfer that occurs between bases in DNA has been reported to be very rapid (Wan *et al.*, 2000). Repair of thymine dimers by an excited state guanine could be a rapid photochemical event that results in an ultrafast rate of "back electron transfer", although it is important to note that forward and back electron transfer are not related necessarily.

Observing the femtosecond absorbance of photoexcited aminopurine-containing DNA, Barton and colleagues determined the rate of charge transfer to neighbouring guanine bases to be 10 picoseconds. This ultrafast rate of electron transfer occurring with stacked bases raises the possibility that the thymine dimer may be stacking with the guanine quadruplex in this system. Interestingly, the same study had found that the rate of charge transfer between photoexcited aminopurine and inosine to be 512 picoseconds, which may account for reduced rates observed with the inosine substitution studies. The fact that UV1C can operate in multiple turnovers is evidence that damage does not occur to a significant degree to the enzyme. To better understand the photophysics of this repair process, ultra-fast spectroscopic techniques would have to be employed, such as laser flash photolysis.

Figure 4-16 Iridium (IV) Chloride treatment of UV1C following photoreversal reaction. Lanes 1 & 2: Control lanes of piperidine-treated unirradiated, and irradiated UV1C, respectively. Lane 3: Unirradiated control lane of UV1C following iridium (IV) chloride treatment and piperidine treatment. Lane 4: Iridium (IV) chloride treated DNA following 1 hour photoirradiation and piperidine treatment. Lanes 5, 6 & 7: Marker lanes for guanine, thymine and adenine, respectively.


Iridium Chloride
dark
light
dark
light
G
T
A
1
2
3
4
5
6
7
50
40
30
20
10

4.4 Summary and Future Directions

Like the naturally occurring photolyase enzymes, UV1C exhibited high substrate specificity, and most likely would use an electron transfer mechanism for catalysis. Although UV1C could not efficiently repair ribonucleotide containing pyrimidine dimers, it was able repair uracil dimers if the sugar ring was 2'-deoxyribose. By systematically replacing components of the all-RNA substrate to resemble that of the all-DNA substrate, it was possible to determine an order of substrate preference. In terms of photochemical mechanism, the possibilities of triplet state or energy transfer mechanism are not likely to occur. Based on approximate calculations, it would be thermodynamically favorable for photoexcited of guanines within the quartet to donate an electron to the thymine dimer. Such a mechanism is illustrated in Figure 4-17, where a photoexcited guanine quartet would donate an electron to a ground state thymine dimer. Using the less oxidizable inosine as

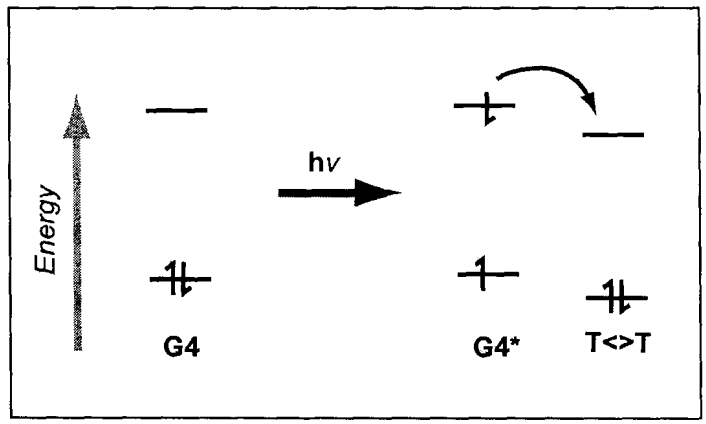


Figure 4-17 Hypothetical orbital energy diagram for electron transfer between guanine-quartets and thymine dimers.

substitutions within the quartet, it appeared that many of the positions within the quartet seemed to contribute to charge transfer. Interestingly, no oxidised damage to guanines were detectable using oxidizing agents such as potassium permanganate or iridium chloride. This may be evidence of a fast rate of "back electron transfer" from the thymine dimer back to the guanine quadruplex.

Future work on UV1C could focus on the photophysics of the repair process spectroscopically. Laser flash photolysis can be used to detect electron transfer in the picosecond timescale (Kavarnos, 1993; Kim *et al.*, 1991). Using a short laser pulse (~30 picoseconds) to photoexcite the guanine residues, the change in the absorbance spectrum can be recorded at intervals in the range of 20 picoseconds to many nanoseconds. Unfortunately, since the fluorescent lifetimes of DNA bases are extremely short (20 ps), the detection of fluorescent decay of the excited state would be difficult. Based on the changes observed in the presence or absence of the thymine dimer, the timescale of the electron transfer reaction can be determined.

The presence of a radical intermediate formed on either the thymine dimer or one on the guanine bases within the enzyme, could be detected by electron paramagnetic resonance (EPR). EPR can be a very effective method for studying the presence of radical species in a reaction when used in combination with spin trapping agents (Rosen, 1999). The trapping agent operates by trapping the "transient" radical in the reaction to form a more stable radical that can then be detected by EPR. Trapping agents, such as 5,5-dimethylpyrroline-N-oxide (DMPO) and phenyl-N-tert-butylnitrone (PBN), have been known to efficiently trap radical species located on the DNA (Hix *et al.*, 2000; Travascio *et al.*, 2001; Witting *et al.*, 2001). Based on the line width and shape of the

EPR signal, the nature of the radical species can be determined. Alternatively, "time-resolved" EPR can be used to measure radical intermediates directly during the catalytic process. In a manner similar to that of laser flash photolysis, after a short (μs) laser flash the EPR signal is analysed at microsecond intervals. This has been used to study electron transfer processes between tryptophan residues and the excited-state flavin in *E. coli* photolyase (Kim *et al.*, 1993).

In the next chapter, I will describe a project involving the selection for a DNA aptamer to binds to both the electron transfer protein cytochrome *c*, and to the small porphyrin hemin.

CHAPTER 5

HEMIN-STIMULATED DOCKING OF CYTOCHROME *C* TO A DNA-CYTOCHOME *C* HEMIN APTAMER COMPLEX

5.1 Introduction

A number of important redox processes in biology involve the transfer of electrons between two docked hemin [iron(III)-protoporphyrin IX]-containing proteins. Cytochrome *c* is one such protein that has been well characterized and is known to bind to many different protein redox partners, such as cytochrome b_5, cytochrome *c* oxidase, and cytochrome *c* peroxidase (Pettigrew and Moore, 1987). Cytochrome c has also been shown to be involved in apoptosis (Cai *et al.*, 1998; Green and Reed, 1998; Mignotte and Vayssiere, 1998). The interaction between cytochrome *c* and another heme-containing protein, cytochrome b_5 (reviewed in (Mauk *et al.*, 1995)), in particular, has been intensively researched and has become a model for the study of interactions between two redox proteins. Recently, novel charge-transfer interactions have also been reported between cytochrome *c* and ruthenium-intercalated double-stranded DNA (Stemp and Barton, 2000).

DNA enzymes have been described, for which porphyrins act as either cofactor or substrate. Li and Sen (Li *et al.*, 1996; Li and Sen, 1996; Li and Sen, 1998) reported a porphyrin-metallating DNA enzyme, PS2.M, 18 nucleotides long, capable of catalytically

inserting zinc and copper(II) ions into mesoporphyrin IX or protoporphyrin IX. Interestingly, this same DNA oligonucleotide, when complexed with hemin, catalyzed peroxidation reactions utilizing hydrogen peroxide (Travascio *et al.*, 1999; Travascio *et al.*, 1998; Travascio *et al.*, 2001; Witting *et al.*, 2001). Given the ability of DNA to accomplish these two types of porphyrin-utilizing reactions normally attributed to proteins, it would be interesting to see if DNA could also participate in electron transfer reactions comparable to those seen with cytochrome proteins.

In this last chapter, I investigate whether a hemin-binding DNA molecule could also bind to cytochrome *c*, specifically and with relatively high affinity. Within cells, the basic protein cytochrome *c* is known to dock with the acidic cytochrome b_5 to initiate electron transfer (Mauk *et al.*, 1995; Stonehuerner *et al.*, 1979). The binding of DNA to cytochrome *c* could be mediated purely by coulombic interactions between the positively charged lysine side chains of cytochrome *c* and the negative phosphate backbone of DNA. However, such an interaction might be non-specific and lead to unpredictable binding stoichiometries. For this reason, it would be important to select for DNA binders of cytochrome c at relatively high salts, to encourage a more specific binding. The goal overall, then, would be to select DNA aptamers capable of binding simultaneously to the small molecule hemin and to the protein cytochrome *c*. From earlier investigations on DNA enzymes utilizing porphyrins as either substrate or cofactor (see above) we had obtained a pool of DNA molecules (the "PS" pool) with a propensity to bind to porphyrins. This unsorted collection of DNA molecules was therefore subjected to a secondary selection process for the ability to bind cytochrome *c*.

5.2 Materials and Methods

5.2.1 Oligonucleotides and Chemicals

Hemin was purchased from Porphyrin Products (Logan, UT) and was used without further purification. Horse heart ferricytochrome c was obtained from Sigma and purified by the method of Brautigan *et al.* (Brautigan *et al.*, 1978) on CM-cellulose. The purified cytochrome *c* was fully oxidized with potassium ferricyanide and the solution passed through a G-25 Sephadex column to remove ferricyanide. Hemin-agarose beads, agarose beads and epoxy-activated Sepharose 6B were purchased from Sigma. Cloning was performed with a TA cloning kit and One-shot competent cells from Invitrogen. γ-^{32}P ATP and the Thermo-sequenase ^{33}P dideoxy sequencing kit were from USB/Amersham. DNA oligonucleotides were purchased from Nucleic Acids Protein Services (NAPS) at the University of British Columbia. The control oligonucleotide PS2M used in spectrophotometric titrations had the sequence 5'-GT GGT AGG GCG GGT TGG-3'. The control peptide ADP-1 was purchased from United Biochemical Research, Inc. (Seattle, WA) and had the sequence: NH_2-CFTTK ALGIS YGRKK RRQRR RPPQG SQTHQ VSLSKQ-COOH.

5.2.2 DNA Library and Oligomers

A pool of DNA molecules (the "PS" pool) from a previous *in vitro* selection (Li *et al.*, 1996) for the binding of N-methyl mesoporphyrin (NMM), was used as the starting pool for this selection. These DNA molecules were of the sequence 5'-GGA TCT TTT TGA TCG GTC GGC ACC -N_{76}- CCT TGG GTC ATT AGG CGA-3'. The primers R1P1-biotin (5'-biotin-TCG CCT AAT GAC CCA AGG) and RLS (5'-GGA TCT TTT TGA TCG GTC GGC ACC-3') were used to amplify these molecules between rounds of

selection, and the amplified DNA was internally labelled with α-^{32}P dATP. Amplified double-stranded DNA was size-purified in 8% non-denaturing gels.

Single-stranded DNA was prepared by binding the amplified double-stranded sequences to an avidin-agarose column in avidin-binding buffer (50 mM Tris-Cl pH 8, 200 mM NaCl, 0.1 mM EDTA). The column was washed with 20 volumes of binding buffer. Single-stranded DNA molecules (the non-biotinylated strand of the column-immobilized duplexes) were eluted with 500 µL of 0.2 M NaOH, directly into 100 µL of 50 mM Tris-Cl containing 1M HCl to neutralize the basicity of the sample. The single-stranded DNA obtained was then ethanol precipitated.

5.2.3 Selection Columns

Hemin-agarose beads were washed and pre-equilibrated in selection buffer A (50 mM Tris, pH 8.0, 120 mM NaCl, 20 mM KCl, 1% DMSO, 0.03% Triton X-100). Cytochrome *c* was attached to activated epoxy Sepharose 6B beads containing a 12-atom spacer. Excess cytochrome *c* was linked to the beads in protein coupling buffer (5 mM sodium phosphate, pH 7.0, 10 mM sodium chloride) by shaking for 72 hours in the dark at room temperature. Protein attachment to the beads was periodically monitored by measuring the decrease in solution protein absorbance at 408 nm. The beads were then washed with 20 column volumes of storage buffer (80 mM sodium phosphate, pH 7), and stored in this buffer. Chemically blocked beads for negative selection columns were prepared by reacting the beads with 4% β-mercaptoethanol and shaking overnight. The beads were then washed and stored in storage buffer at 4°C.

5.2.4 Selection Protocols

The selection strategy involved four columns: two positive columns, containing hemin and cytochrome *c*, respectively, and two negative columns containing the respective beads-- in the case of the Sepharose, chemically blocked. A single round of selection typically involved the following: ^{32}P-end-labeled single-stranded DNA molecules were heated to 95° C for 5 minutes in TE buffer (10 mM Tris-Cl pH 8, 0.1 mM EDTA) and allowed to cool very slowly for 25 minutes. For the first round of selection an equal volume of 2X selection buffer A (see above) was added, and the DNA allowed to fold for 30 minutes at room temperature. The folded DNA was then passed through a negative column containing blocked-agarose beads. DNA molecules that did not bind to this column were directly passed into a hemin-agarose column, the flow-through from this second column being recirculated through the column five times. Following thirty minutes of incubation, the hemin-agarose column was washed with selection buffer B (50 mM Tris, pH 8.0, 240 mM NaCl, 20 mM KCl, 1% DMSO, 0.03% Triton X-100) until no more counts were detectable in the eluate. The DNA molecules bound to this column were then eluted with TE buffer (10 mM Tris-Cl pH 8, 10 mM EDTA).

The recovered DNA was ethanol precipitated and refolded in reaction buffer. Hemin was added to a concentration of 8 μM (in at least a 10-fold molar excess over the DNA). These folded DNA molecules in the presence of the excess hemin were then passed through a second negative column containing blocked Sepharose (without cytochrome *c* attached). The unbound DNA was passed directly through a positive column containing cytochrome *c*-linked beads. As above, the flow-through was re-circulated five times and the beads and DNA solution incubated for 30 minutes each

time. The column was then washed with reaction buffer until no more counts were detected, and the bound DNA was eluted with a high salt buffer, 2M NaOAc (pH 6.0) followed by a TE buffer elution, so as not to bias the selection towards electrostatic interaction. In rounds 5 and 6, elution was achieved by incubation of the DNA-bound cytochrome *c* beads with free cytochrome *c* (2.5 mM) in selection buffer B. The eluted DNA-cytochrome solution was then phenol-chloroform extracted, ethanol precipitated, and the purified DNA amplified by PCR using the R1P1-biotin and RLS primers.

5.2.5 Cloning and Sequencing

After round 6 of selection the pool of DNA molecules successfully binding to the cytochrome *c* column was amplified using the cloning primers RLC (5'-CTT GTC TGC AGG GAT CCT TTT GAT CCG GTC GGC) and R1P1C (5'-GAT ATC AAG CTT CTC GAG TCG CCT AAT GAC CAA GG). The amplified DNA was gel purified in an 8% non-denaturing gel and ligated into the pCR2.1 plasmid using a TA cloning kit (Invitrogen). The plasmid was then transformed into INVαF' (Invitrogen) competent *E. coli* cells. Thirty-eight recombinant clones were picked for analysis and sequenced using the Thermo-sequenase dideoxy kit (USB Amersham). Sequence alignment was performed with the aid of the FOLDALIGN algorithm (Gorodkin *et al.*, 1997a; Gorodkin *et al.*, 1997b) to identify secondary structure conservation.

5.2.6 Dissociation Constant Determinations

The binding affinity for hemin of the various DNA aptamers was determined spectroscopically (Li and Sen, 1998; Travascio *et al.*, 1998), by monitoring hyperchromicity of the hemin Soret absorption. All electronic absorption spectra were

recorded in a dual-beam Cary 300-Bio UV-Visible Spectrophotometer, at 22 ° C. The change in absorbance at 404 nm for hemin was monitored at increasing concentrations of DNA (0-20 μM) titrated to a fixed concentration (1.5 μM) of hemin in Selection buffer A. The data obtained were fit to a binding equation by Wang *et al.* (Wang *et al.*, 1997):

$$[DNA]_o = K_d(A - A_o) / (A_\infty - A) + [P_o](A - A_o) / (A_\infty - A_o) \qquad \text{(eq-2)}$$

where $[DNA]_o$ is the starting concentration of DNA; A_o and A_∞ are the absorbances of DNA-hemin at zero concentration of DNA and at saturating concentrations of DNA respectively. $[P_o]$ is the initial concentration of monomeric hemin.

Binding constants for cytochrome *c* were also determined by difference absorption spectroscopy (Erman and Vitello, 1980; Mauk *et al.*, 1982; Michel and Bosshard, 1984) using matched Hellma cuvettes and a dual-beam UV-visible Spectrophotometer (Varian). Measurements in the absence of hemin were carried out with only selection buffer in cuvette A; and, 1.0 μM DNA in the same buffer in cuvette B. 0 - 7.6 μM cytochrome *c* was now added to the DNA in cuvette B and the added absorbance spectrum of both cuvette A and cuvette B recorded as "Spectrum 1". The comparison run consisted of cuvettes C and D, both filled with buffer, with cuvette C containing 1.0 μM DNA. Cytochrome *c* (0 - 7.6 μM) was now added to cuvette D and the added spectrum of cuvettes C and D recorded as "Spectrum 2". Spectrum 2 was subtracted from Spectrum 1 to give the difference spectrum. Due to the sensitivity of the electronic spectra of both hemin and cytochrome *c* to their respective oxidation states (Falk and Smith, 1975), titrations were performed under oxidizing conditions in the presence of 10 μM potassium ferricyanide. All cuvettes were weighed on an analytical

balance to an accuracy of ± 0.0005 g prior to the recording of absorbance values. Spectra measured in the presence of hemin were recorded as above, except that 10 μM hemin was added to cuvette B (with DNA and cytochrome *c* already present) for the sample run; and, was added to cuvette D (containing cytochrome *c* alone) for the blank run. No significant difference spectrum was observed for a titration of a fixed concentration of hemin with cytochrome *c*, relative to buffer, with a maximum ΔA_{408} of 0.0005 (in other words, any putative interaction of hemin with cytochrome *c* did not yield a difference spectrum). Binding constants of cytochrome *c* with DNA or with DNA-hemin were determined using a modified equation by Erman and Vitello (Erman and Vitello, 1980):

$$\Delta A = \Delta A_{\infty} / 2D \left[C + D + 1/K_A - \{(C + D + 1/K_A)^2 - 4CD\}^{1/2}\right] \qquad \text{(eq-3)}$$

where C and D are the concentrations of cytochrome *c* and DNA respectively; ΔA_{∞} is the change in absorbance at saturating concentrations of cytochrome *c*, and K_A is the association constant. Concentrations were corrected for dilution effects resulting from additions to the cuvettes from the cytochrome *c* stock. Three sets of independent measurements were taken.

5.2.7 Chemical Probing with Dimethyl Sulfate and Diethyl Pyrocarbonate

For methylation reactions, a 1.2% dimethyl sulfate (DMS) solution in water was freshly prepared and 1 μL added to 11 μL of 5'-kinased DNA folded in methylation buffer (20 mM potassium cacodylate, pH 8, 120 mM NaCl, 0.5% DMSO and 0.03% Triton). For diethyl pyrocarbonate (DEPC) reactions, 5 μL of 3% DEPC was added to 10 μL of a DNA solution in the same buffer. Both reactions were allowed to proceed for 30 minutes at room temperature, and terminated by ethanol precipitation at dry ice

temperatures. The recovered DNA was dissolved in 10% piperidine and heated for 30 minutes at 90° C, followed by lyophilization. The samples were then dissolved in denaturing gel loading buffer (containing 80 % formamide v/v and 20 mM EDTA, pH 8.0) and analyzed in 10% denaturing polyacrylamide gels.

5.2.8 Fenton Reaction Footprinting

The cleavage of DNA by Fenton reaction-generated hydroxyl radicals was performed using a method modified from Weidner *et al.* (Weidner *et al.*, 1989). 5'-kinased DNA molecules were folded in selection buffer A, containing 20 mM KCl and 120 mM NaCl in a total volume of 10 μL. The solutions were then made up to 50 μM $Fe(NH_4)_2(SO_4)_2$, 100 μM EDTA, 1 mM sodium ascorbate, and the reaction initiated by the addition of 10 mM H_2O_2 (final). After 1 minute of reaction at room temperature, the samples were quenched by the addition of 2 μL of 5 mM thiourea. Solutions containing protein were phenol-chloroform extracted. The DNA solutions were then lyophilized, resuspended in denaturing gel-loading buffer, and analyzed in a 10% denaturing gel. It was ensured that equal radioactive counts were loaded into every well.

5.3 Results

5.3.1 Selection

As described above, prior selections carried out in our laboratory to select for DNA enzymes that utilized porphyrins as either substrate or cofactor (see above) had generated, after thirteen rounds of selection a pool of DNA molecules (the "PS" pool) with a propensity to bind to porphyrins. It was impossible to estimate the number of individual sequences present in this PS pool; however, the sequencing of twenty-five cloned members from this pool had yielded twenty-five different sequences (Li *et al.*, 1996). Individual aptamers from this pool were also shown to be able to bind a variety of porphyrins and metalloporphyrins other than the N-methyl mesoporphyrin (NMM) used for that selection. This present selection was therefore one for a secondary binding activity, to cytochrome *c*, of DNA aptamers already predisposed to binding porphyrins.

Table 5-1 summarizes the binding of the PS pool DNA to the hemin and cytochrome *c* selection columns, as functions of rounds of selection. While the first round of selection was carried out in selection buffer A (50 mM Tris, pH 8.0, 120 mM NaCl, 20 mM KCl, 1% DMSO, 0.03% Triton X-100), in subsequent rounds the sodium chloride concentration was raised from 120 mM to 240 mM (selection buffer B), in an attempt to reduce the possibility of non-specific electrostatic interactions between cytochrome *c* and DNA that could occur at low ionic strengths. Whereas, as expected, the hemin-binding properties of the PS pool were excellent from the start of selection, cytochrome *c* binding increased substantially after the second round, and essentially reached a plateau following the fifth round. Starting from round 5 an additional step of negative selection was incorporated, in which DNA folded in the absence of hemin was applied to the

cytochrome *c* column, and only those DNA molecules not binding to the column in the absence of hemin were allowed to continue through the selection. Following round 6, the DNA pool was cloned. Thirty-seven clones were sequenced and their random region sequences compared in search for sequence homologies that might be important for the binding of both hemin and cytochrome *c* (Figure 5-1).

Table 5-1 Summary of Binding to Selection Columns

	% binding	
round	hemin beads	cyt *c* beads
1	49	< 1
2	74	5[a]
3	73	15
4	ND	37
5	ND	50[b,c]
6	65	53[c]

[a] Total monovalent salt was increased from 140 mM to 260 mM.
[b] Additional step added. DNA folded in the absence of hemin was passed through the cyt *c* column. Less than 5% of the DNA did not bind and these molecules were then amplified for round 6.
[c] 2.5 mM cytochrome *c* in selection buffer B was used to competitively elute the DNA off the column.

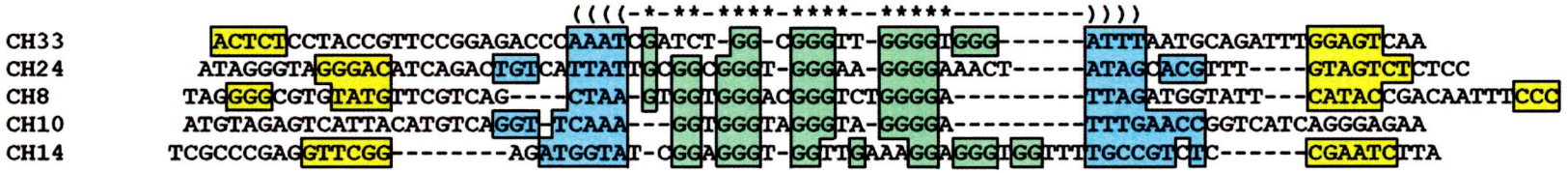
Since the starting library for this selection was the PS pool, described earlier, from a previous selection for porphyrin binding carried out by Li *et al.* (Li *et al.*, 1996), it was not surprising that guanine-rich sequence elements (shown in green boxes in Figure 5-1), consistent with porphyrin binding sites, were found in all clones. However, unlike clones analyzed in the earlier selection (Li *et al.*, 1996), where the G-rich region within

each clone was situated randomly within the N_{76} region, the G-rich regions within these clones appeared to be located centrally within the N_{76} sequence. These new clones also provided a stronger consensus sequence for the G-rich elements and could be classified into three groups based on their guanine consensus sequence. Members from Group A contained $GTG_2A(T)G_3TG_3AG_{2\text{-}3}$. Group B sequences ($GNG_2NG_3NG_3NG_4$) contained one run of four guanines, with two members containing an additional set of guanine triplets. Group C had in common $G_3T_nG_3T_nG_3NG_2$. Previous studies in our lab on DNA aptamer-porphyrin interactions (Li *et al.*, 1996; Travascio *et al.*, 1998) had found that these G-rich elements folded intramolecularly to form guanine-quadruplexes (G-quadruplexes) (Wellinger and Sen, 1997; Williamson, 1994). With the exception of the G-rich motif, however, attempts to find other elements of sequence homology among the new clones using conventional sequence alignment algorithms (Higgins *et al.*, 1996) failed. However, in spite of the absence of overt sequence homology between the clones, when their sequences were subjected to a DNA folding prediction program (Mfold server, (SantaLucia, 1998)), roughly half of the clones were predicted to fold in a similar way, forming one or two predominantly double-helical stems (involving sequences shown as boxed in Figure 5-1), with the guanine-rich, putatively hemin-binding element isolated as a terminal loop. With the aid of the sequence alignment tool FOLDALIGN (Gorodkin *et al.*, 1997a; Gorodkin *et al.*, 1997b), which takes into account secondary structure, we were able to identify conserved regions of base complementarity. Shown in Figure 5-1, in blue and yellow highlight, all sequences contained two sets of base-paired stems. Sequence homology of these stems is highest amongst Group A members, with runs of six to eight consecutive complementary bases. Sequences in Groups B and C had

Figure 5-1 Sequence alignment of selected clones using the FOLDALIGN algorithm. Sequences are shown 5' to 3' and are grouped according to their highly conserved guanines are shown in green. Flanking sequences predicted to form double-stranded regions are highlighted in yellow and in blue to their respective complementary regions, also indicated with parenthesis above the sequence. (*) denotes high sequence conservation.

Group A

Group B



Group C

shorter conserved areas. Although the presence of two duplex stems is ubiquitous in the sequenced pool, the role of these elements in protein binding is unclear. Figure 5-2 shows the predicted secondary structure for a single representative clone, CH6, which was chosen for further characterization.

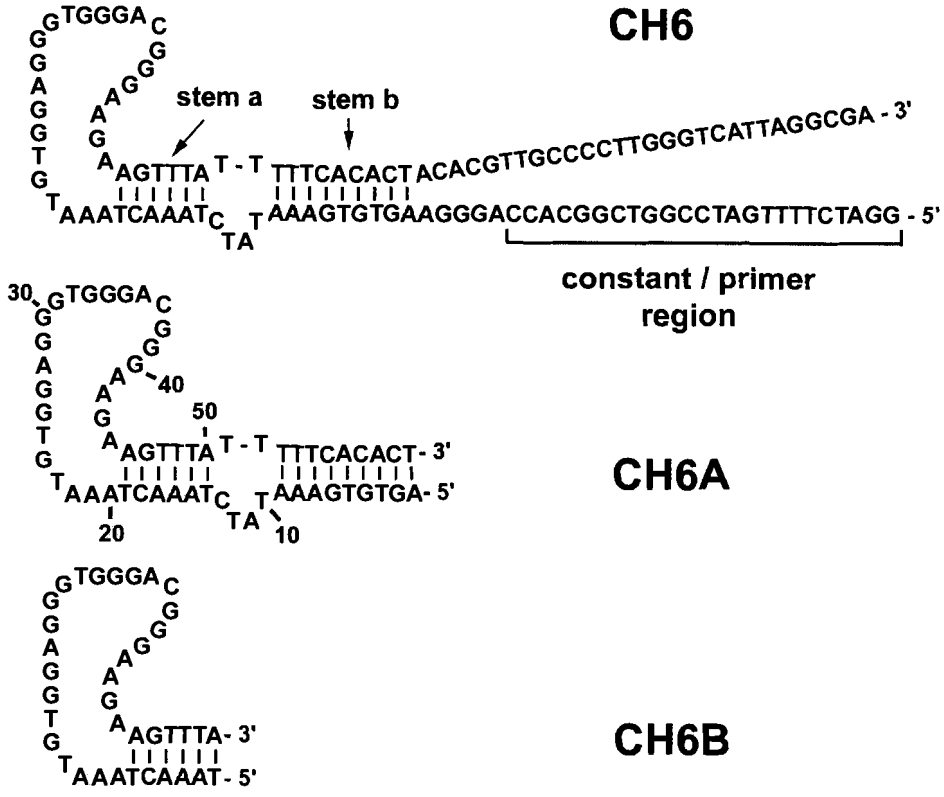


Figure 5-2 The sequence and the predicted folding of the DNA aptamers.
Shown are CH6, and its two deletion mutants, CH6A and CH6B. Two double-helical stems separated by a four base bulge are predicted for CH6, based on sequence complementarity. CH6A lacks the primer-binding sequences of CH6, but maintains the two stems, a and b. CH6B lacks stem b, as well as the primer-binding sequences.

5.3.2 Spectroscopy of Hemin Binding

To investigate the importance for cytochrome- or hemin-binding of the putatively stem regions flanking the G-rich element, two deletion mutants of clone CH6 were synthesized, which contained: (a) the full sequence of clone CH6, except for the 5' and 3' constant regions ("CH6A"), and (b) the full sequence minus the elements required for stem **b** as well as the 5' and 3' constant regions ("CH6B"). These are shown in Figure 5-2.

The binding of CH6 to hemin was measured by UV-visible spectroscopy, following the method of Travascio *et al.,* (Travascio *et al.*, 1998). Figure 5-3 shows that disaggregated hemin binding to CH6 (or, to CH6A or to CH6B-- data not shown) gave rise to a hyperchromicity in the hemin Soret absorbance, along with a small red shift, from 398 nm to 404 nm. In addition, isosbestic points were seen at 420 nm, 493 nm and 545 nm. The change in A_{404} as a function of DNA concentration was used to calculate dissociation constants, using the formalism of Wang *et al*, (Wang *et al.*, 1997). Table 5-2 summarizes the K_d values for hemin binding to the three DNA oligomers. The values obtained for CH6, CH6A, and CH6B were 3.56 ± 0.14 μM, 0.97 ± 0.16 μM, and 0.6 ± 0.22 μM, respectively. The extinction coefficients at 404 nm determined for the DNA-hemin complexes were: $1.94 \times 10^5\ M^{-1}cm^{-1}$, $1.98 \times 10^5\ M^{-1}cm^{-1}$, and $1.56 \times 10^5\ M^{-1}cm^{-1}$ respectively.

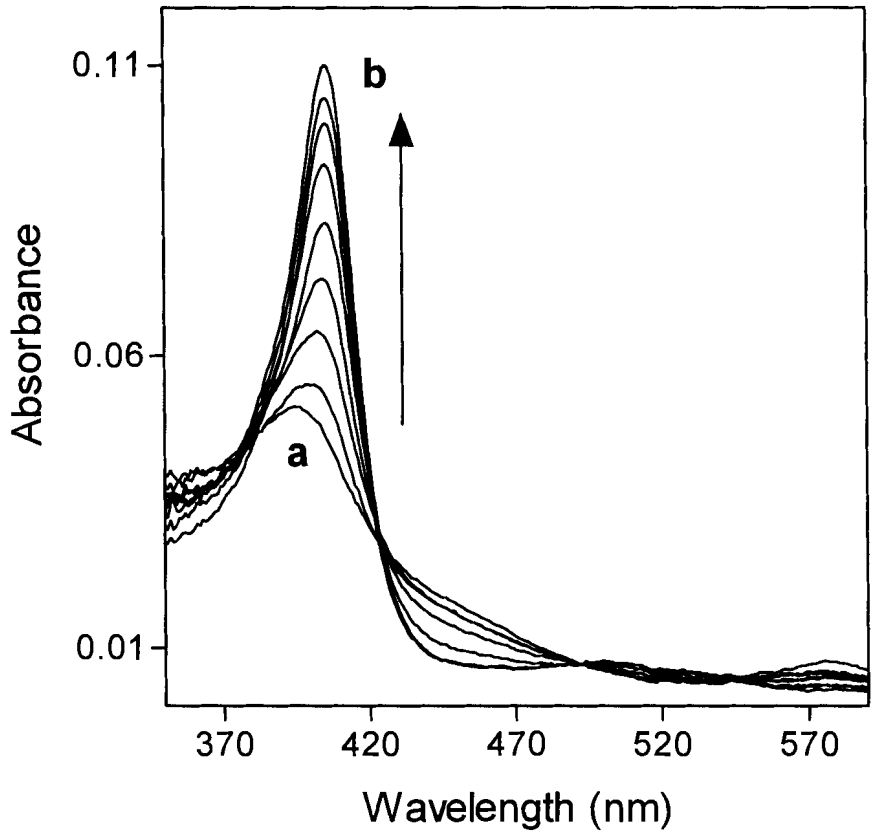


Figure 5-3 Binding of hemin to the aptamer CH6.
The absorbance spectrum of hemin is plotted as a function of increasing DNA concentrations. The hemin concentration was held constant at 0.5 μM, and the DNA concentration varied from 0 to 1, 2, 5, 10, 15, and 20 μM (a = 0 μM; b = 20 μM DNA).

Table 5-2 Summary of Dissociation Constants

Dissociation Constants (K_D)[a]			
	CH6	CH6A	CH6B
DNA to hemin[b]	3.6 ± 0.2	1.0 ± 0.2	0.6 ± 0.2
DNA to cyt c[c]	0.8 ± 0.3	4.6 ± 0.6	ND
DNA to cyt *c* in presence of hemin[d]	0.4 ± 0.2	0.3 ± 0.2	ND

[a] K_D values are stated in μM.
[b] Buffers conditions used: 10 μM hemin, 50 mM Tris-Cl pH 8.0, 120 mM NaCl, 20 mM KCl, 1% DMSO and 0.03% Triton X-100. [c] conditions same as in b without hemin and with 10 μM ferricyanide.
[d] conditions same as in c, with the addition of 10 μM hemin.

Scatchard analysis of hemin binding to CH6, CH6A and CH6B gave ν values of 0.92, 0.86, and 1.12 ± 0.20, respectively (where ν describes the number of moles of hemin bound per mole of DNA). This indicated that for all three DNA oligomers, one hemin molecule bound per molecule of DNA.

5.3.3 Binding of Cytochrome *c* to CH6A monitored by Difference Spectroscopy

The docking of protein partners to cytochrome *c* have traditionally, and most conveniently, been measured using "difference" spectroscopy (Erman and Vitello, 1980; Mauk *et al.*, 1982; Michel and Bosshard, 1984), the measurement of small spectroscopic perturbations in the spectra of the cytochrome's heme concomitant with the docking of another protein to cytochrome *c*. Fundamentally, this involves obtaining the difference

between the spectrum of the cytochrome *c*- docking partner complex and the added individual spectra of cytochrome *c* and of its docking partner.

We used difference spectroscopy in the Soret region of the spectrum to measure the binding affinity of cytochrome *c* to CH6 and to its deletion mutants. The binding of cytochrome *c* to 1 μM CH6A (in both the presence and absence of hemin) resulted in positive difference absorbance spectra. Figure 5-4 shows the spectrum obtained in the presence of hemin (dashed). A major peak was seen at 407 nm (as well as two smaller maxima at approx. 500 nm and 530 nm in the presence of hemin). In the absence of hemin (solid) the difference spectrum was similar, with the exception of a lower absorbance of the Soret peak (409 nm) relative to that seen in the presence of hemin.

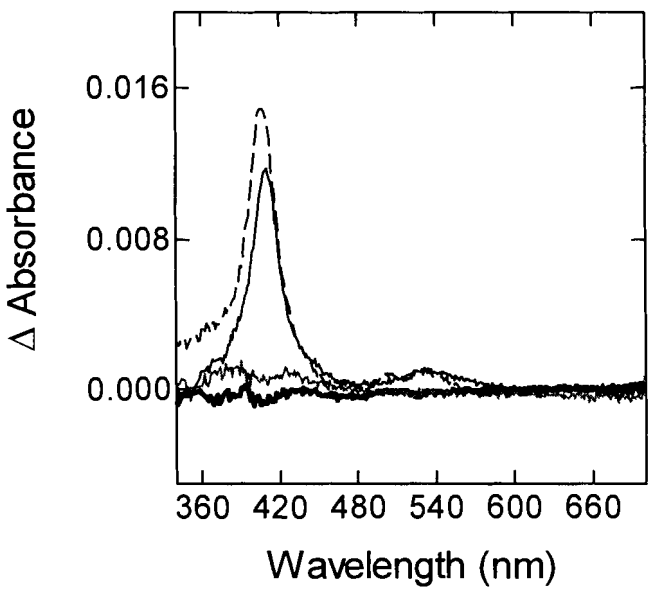


Figure 5-4 Electronic difference spectra of 1 μM CH6A in the presence of cytochrome *c*. A major Soret peak is observed for CH6A folded with 10 μM hemin (dashed line) and for CH6A alone (solid line), but none for CH6B (thick line) or PS2M (grey line), both in the presence of 10 μM hemin.

The maximal change in the spectra amounted to ~2% of the total absorbance (with $\Delta\varepsilon_{408}$ = 1.02 x 10^4 $M^{-1}cm^{-1}$). Figure 5-5 plots ΔA_{408} against the molar ratio of cytochrome *c*:CH6A DNA in the presence (filled circles) and absence (open circles) of hemin. Saturation of the absorbance difference occurred at a protein to DNA molar ratio of ~1, suggesting a binding stoichiometry of one cytochrome *c* molecule to one molecule of CH6A (similarly, analysis of cytochrome binding to the larger CH6 also gave a 1:1 protein: DNA binding stoichiometry-- data not shown). The stoichiometries found agreed with Scatchard analysis and Job plots (Job, 1928) (data not shown). Figure 5-5 shows also that CH6A bound cytochrome *c* with a significantly higher affinity in the presence of hemin (K_d = 0.3 μM) than in the absence of hemin (K_d = 4.6 μM).

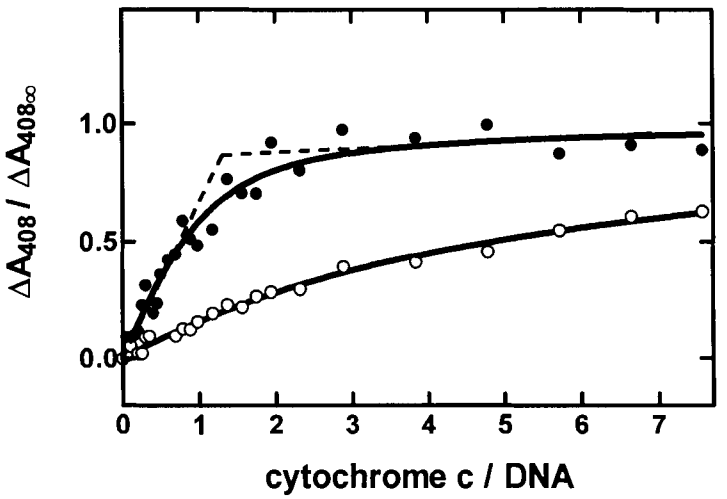


Figure 5-5 Absorbance difference at 408 nm of CH6A-hemin titrated with cytochrome *c*. The absorbance difference saturates at a DNA:cytochrome molar ratio of 1.1. Titrations were carried out on 1 μM CH6A DNA, in the presence (●) and absence (○) of 10 μM hemin, in selection buffer A.

While difference spectra were observed with CH6A and CH6, the oligomer CH6B failed to produce a difference spectrum (Figure 5-4, thick), suggesting a lack of significant

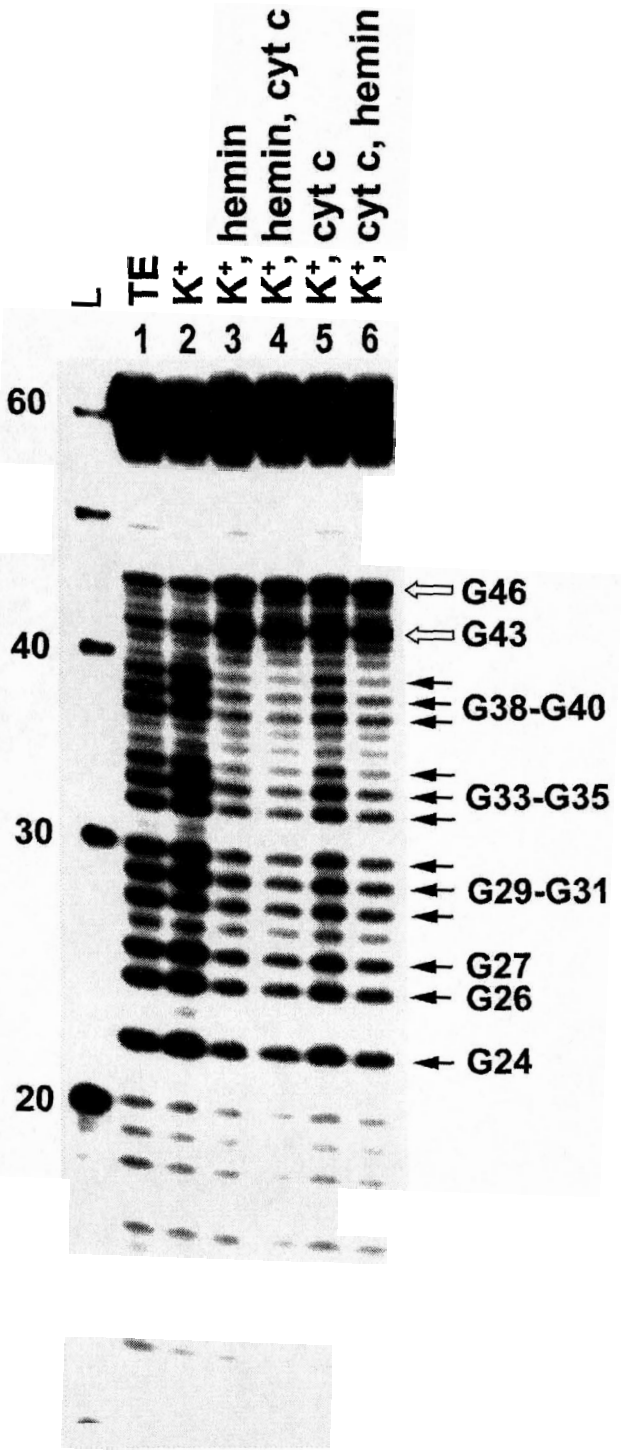
binding of cytochrome *c* to CH6B. This striking difference in the properties between CH6A and CH6B may indicate that either: 1) the stem **b** of the aptamer secondary structure (see Figure 5-2) is directly involved in binding to cytochrome *c*; or, that 2) stem **b** helps stabilize the overall folded structure of the molecule, possibly stabilizing stem **a** such as to facilitate binding to the protein. As a control, we also attempted to determine if PS2M (Travascio *et al.*, 1999; Travascio *et al.*, 1998; Travascio *et al.*, 2001; Witting *et al.*, 2001), a hemin-binding 18-mer that folds to form a quadruplex with no double-stranded elements, was capable of producing a difference spectrum. Both in the presence of hemin (Figure 5-4, gray) and the absence of hemin (data not shown) PS2M did not produce a difference spectrum in the presence of cytochrome *c*. Table 5-2 summarizes the dissociation constants for complexes formed by the aptamer oligomers to hemin alone, to cytochrome *c* alone, and to hemin and cytochrome *c* in a ternary complex. Because of the 1:1 hemin-CH6A binding stoichiometry, as well as the interesting feature of enhanced binding affinity in the presence of hemin, the aptamer oligomer CH6A was chosen for a more in-depth analysis.

5.3.4 DNA Structure Probing

Chemical probing analysis was carried out with dimethyl sulfate (DMS) and diethylpyrocarbonate (DEPC) on CH6A and on complexes of CH6A with hemin and/or cytochrome *c*. Dimethyl sulfate preferentially methylates the N7 position of guanines, accessible from the major groove within Watson-Crick double helices. Therefore, the interaction of proteins with the major groove of DNA duplexes can be footprinted using dimethyl sulfate. However, dimethyl sulfate -reactivity also provides information about the participation of specific guanines in guanine-quartet formation

(guanines in G-quartets are not methylated by dimethyl sulfate) (Sen and Gilbert, 1988). Diethylpyrocarbonate reacts with adenine N7 positions, but preferentially with adenines that are present in either single-stranded, unstructured, or solvent-exposed regions of DNA secondary/tertiary structure. Figure 5-6 shows the dimethyl sulfate modification pattern of CH6A under different conditions. In the low-salt TE buffer (lane 1) all guanines in CH6A showed dimethyl sulfate reactivity. Folding of CH6A in a K^+-containing buffer (lane 2) still left all guanines reactive, even though it is known that the potassium ion supports the formation as well as stabilization of guanine-quartets. However, when hemin was added (lane 3) to CH6A in the potassium buffer a significant protection was seen for the central region of guanines, from G24-G40, consistent with the anticipated hemin-binding, guanine-rich loop (Figure 5-2) folding to form a G-quadruplex. The addition of cytochrome *c* to the CH6A-hemin complex (lane 4) did not significantly alter the guanine modification pattern. However, the addition of cytochrome *c* in the absence of hemin (lane 5) still showed significant guanine reactivity with dimethyl sulfate, suggesting that cytochrome *c* binding did not in itself support G-quadruplex formation. In correlation with this finding were the binding data to CH6A (above), which indicated that the prior, and hemin-assisted, formation of a G-quadruplex resulted in a stronger binding of cytochrome to CH6A. Lane 6 shows that adding cytochrome *c* first, followed by hemin, to CH6A gave a protection pattern essentially indistinguishable from that seen in lane 4 (where hemin was added first, followed by cytochrome). Therefore,

Figure 5-6 Chemical probing analysis of DMS-modified CH6A.
Samples in lanes 1 and 7 were probed in TE buffer. Lanes 2 and 8: DNA folded in methylation buffer (containing 20 mM K^+). Lanes 3 and 9: 5 μM hemin added to DNA folded in methylation buffer. Lanes 4 and 10: 10 μM cytochrome *c* added to DNA-hemin in methylation buffer. Lanes 5 and 11: 10 μM cytochrome *c* added to DNA folded in methylation buffer in the absence of hemin. Lanes 6 and 12: 5 μM hemin added to DNA folded in buffer with 10 μM cytochrome *c*. Lane C is a DEPC control where non-modified DNA is subject to piperidine treatment. Solid arrows indicate areas of protection in (a) and areas of hyper-reactivity in (b). Open arrows indicate bases showing little or no change. Bracketed regions show regions of following the addition of methylation buffer alone. Little change is seen when buffer alone is added to the DNA (lanes 1, 7 versus 2, 8). Protection of central guanines is seen when hemin is added (lane 3) and correlates to the appearance of hyper-reactive adenines 28 and 36 (lane 9).


L
TE
K+
K+; hemin
K+; hemin, cyt c
K+; cyt c
K+; cyt c, hemin
1 2 3 4 5 6
60
40
30
20
G46
G43
G38-G40
G33-G35
G29-G31
G27
G26
G24

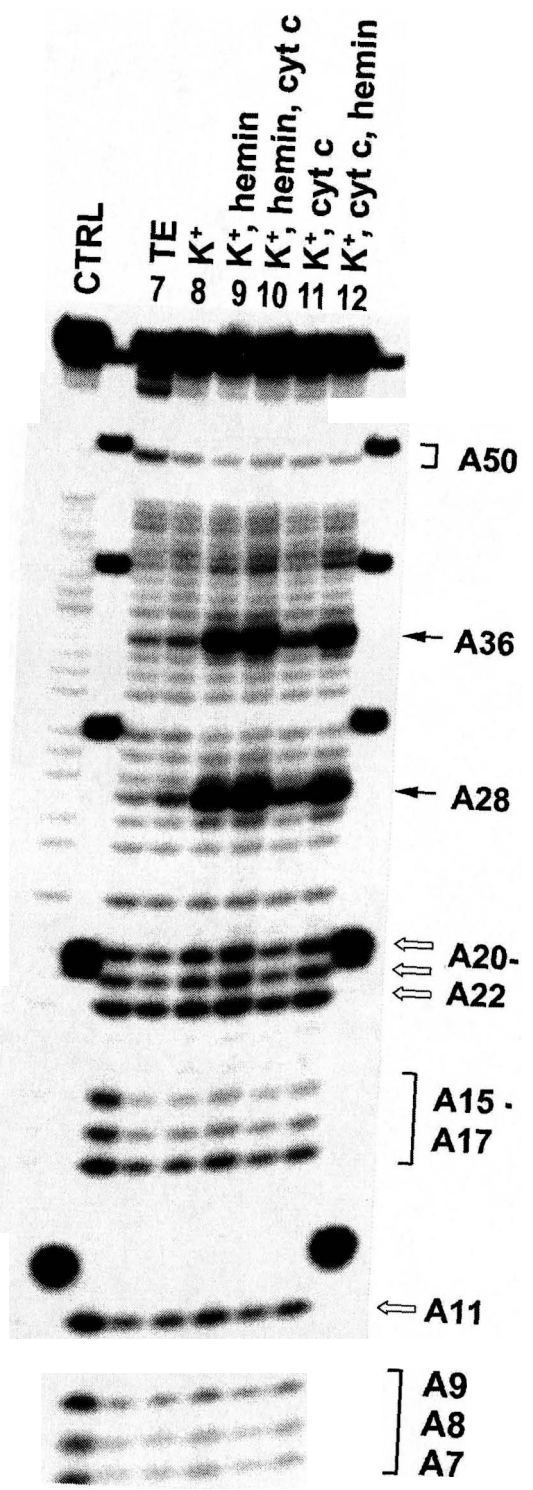
the order of addition of the two ligands, hemin and cytochrome *c*, to CH6A, was probably not important for the proper formation of the ternary complex.

The diethylpyrocarbonate experiments (Figure 5-7) show that A15-A17 became protected when folded in K^+ buffer (lane 8 versus lane 7), consistent with their lying within a duplex region (see Figure 5-2). When hemin was added to this solution (lane 9, Figure 5-7) A28 and A36 within the G-rich region became hyper-reactive, possibly owing to being thrust out into the solvent concomitant with their surrounding guanines base-pairing with each other to form the quadruplex. The addition of cytochrome *c* to this DNA-hemin complex (lane 10) again gave a reactivity/ protection pattern essentially unchanged from that in lane 9. Cytochrome, added in the absence of hemin (lane 11) did not result in the hyperactive adenines seen in lanes 9 and 10, once again suggesting that cytochrome *c* in itself was incapable of supporting a formation of the G-quadruplex. As seen with the dimethyl sulfate experiments, reversing the order of hemin and cytochrome *c* addition (lane 12) did not result in a different protection pattern from the one seen in lane 10.

5.3.5 Fenton Reaction Footprinting

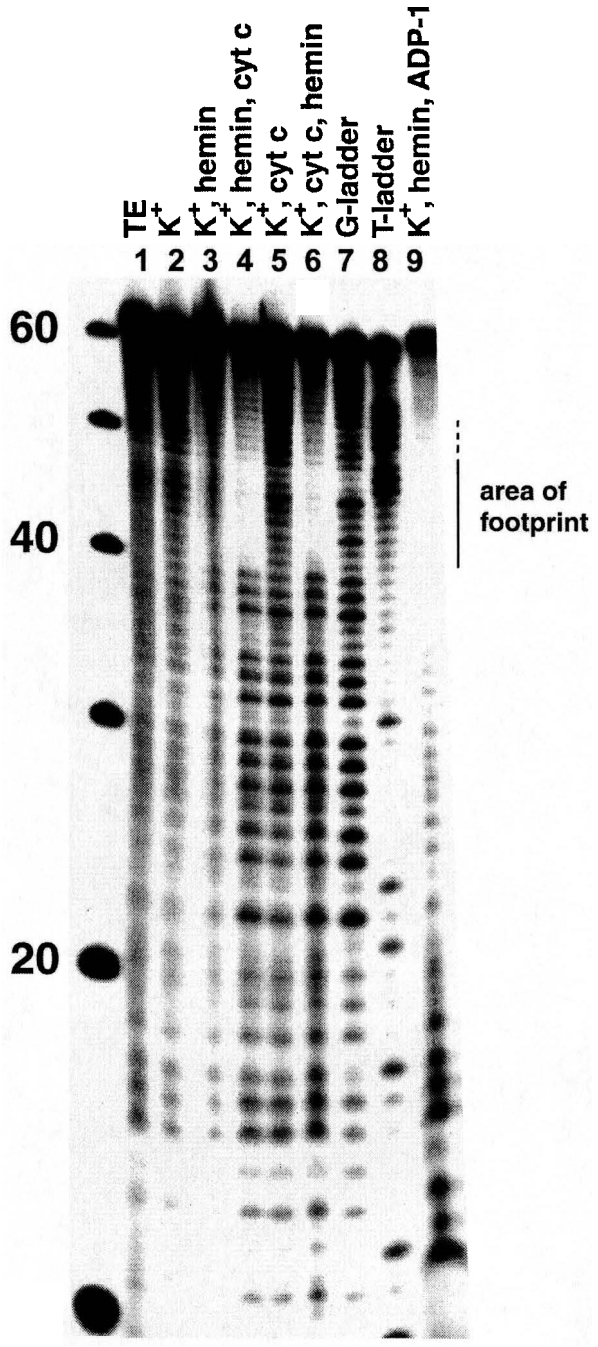
The chemical probing experiments utilizing dimethyl sulfate and diethylpyrocarbonate (above), which monitor the conformation of individual bases in the DNA and also the hydrogen-bonding or other contacts made by the DNA bases, did not yield any obvious signature for the binding of cytochrome *c* to CH6A (Figure 5-6 & Figure 5-7). We therefore attempted to 'footprint' this interaction using hydroxyl radicals generated *in situ* by the Fenton Reaction (Weidner *et al.*, 1989). Hydrogen

Figure 5-7 Chemical probing analysis of DEPC-modified CH6A.
Samples in lanes 1 and 7 were probed in TE buffer. Lanes 2 and 8: DNA folded in methylation buffer (containing 20 mM K^+). Lanes 3 and 9: 5 μM hemin added to DNA folded in methylation buffer. Lanes 4 and 10: 10 μM cytochrome *c* added to DNA-hemin in methylation buffer. Lanes 5 and 11: 10 μM cytochrome *c* added to DNA folded in methylation buffer in the absence of hemin.Lanes 6 and 12: 5 μM hemin added to DNA folded in buffer with 10 μM cytochrome *c*. Lane C is a diethylpyrocarbonate control where non-modified DNA is subject to piperidine treatment. Solid arrows indicate areas of protection in (a) and areas of hyper-reactivity in (b). Open arrows indicate bases showing little or no change. Bracketed regions show regions of following the addition of methylation buffer alone. Little change is seen when buffer alone is added to the DNA (lanes 1, 7 versus 2, 8). Protection of central guanines is seen when hemin is added (lane 3) and correlates to the appearance of hyper-reactive adenines 28 and 36 (lane 9).


CTRL
TE
K+
K+; hemin
K+; hemin, cyt c
K+; cyt c
K+; cyt c, hemin
7
8
9
10
11
12
A50
A36
A28
A20-
A22
A15 -
A17
A11
A9
A8
A7

abstraction from the 1' position of the deoxyribose sugars of DNA leads to a breakage of the DNA chain; and hydroxyl radicals are therefore useful as probes for the sugar-phosphate backbone of DNA, rather than for the bases. Figure 5-8 shows that the pattern of cleavage changed relatively slightly when the CH6A DNA was folded in potassium buffer, either in the absence (lane 2) or presence (lane 3) of hemin. When cytochrome *c* was added to the DNA-hemin complex (lane 4), however, a clear footprint was observed from residues ~40-55. This protected stretch lies just outside of the G-quadruplex forming region, but incorporates a portion of the double helical region (as predicted, shown in Figure 5-2, and also from the diethylpyrocarbonate -protection experiments, Figure 5-7). Once again, the order of addition of hemin and cytochrome *c* appeared not to be important (lanes 4 and 6)-- both orders of addition generated the footprint. However, cytochrome *c* in the absence of hemin (lane 5) did not generate the unequivocal footprint seen in the dual presence of hemin and cytochrome. As an added control to test for the specific locus of cytochrome binding upon the DNA, we used a highly basic peptide known to bind to RNA (Churcher *et al.*, 1993). ADP-1 is a 36 amino acid arginine-rich peptide that corresponds to the 37-72 amino acid region of the HIV Tat protein. ADP-1 might be expected to bind promiscuously, via mainly coulombic interactions, with the aptamer DNA. Lane 9 shows the footprint produced when incubating hemin-complexed CH6A with 12 μM ADP-1. The very large footprint produced by the small peptide is indicative of a high degree of binding, yet largely non-site-specific interaction. The stoichiometry of binding is most likely estimated to be much larger than 1:1.

Figure 5-8 Fenton reaction footprinting analysis of CH6A.
0.5 μM CH6A was folded in a selection buffer A (containing 140 mM monovalent salt). The reaction was initiated by the addition of 10 μM H_2O_2 and quenched with 5 mM thiourea. Samples were run in a 10% denaturing gel. 'G' and 'T' lanes (lanes 7 & 8) show the G and T sequencing ladders, respectively. Lane 1: DNA in TE buffer. Lane 2: DNA folded in selection buffer A (containing 20 mM K^+). Lane 3: 10 μM hemin added to folded DNA. Lane 4: 10 μM cytochrome *c* added to DNA folded with hemin. Lane 5: 10 μM cytochrome *c* added to DNA folded in the absence of hemin. Lane 6: the order of hemin and cyt *c* addition was reversed. In the dual presence of hemin and cyt *c* (lanes 4 and 6) a footprint could be seen, indicated by the bracketed region. Lane 9: control reaction with 12 μM ADP-1 peptide added to hemin-folded CH6A.


TE
K+
K+, hemin
K+, hemin, cyt c
K+, cyt c
K+, cyt c, hemin
G-ladder
T-ladder
K+, hemin, ADP-1
1 2 3 4 5 6 7 8 9
60
40
20
area of
footprint

5.4 Discussion

5.4.1 The Binding of Hemin to the DNA Aptamer

Earlier searches for single-stranded DNA (and RNA) sequences capable of folding to form porphyrin binding sites (Li *et al.*, 1996; Travascio *et al.*, 1998) had found guanine-rich sequences, that folded to form G-quadruplexes, to be important for such binding sites. The precise mode of interaction of hemin with such quadruplex sites has not been elucidated, but it has been proposed, from energetic considerations, that the preferred binding site of a porphyrin (or metalloporphyrin) to a quadruplex may be *via* stacking interactions upon the terminal guanine quartets of a quadruplex, rather than *via* intercalation between quartets (Travascio *et al.*, 1999; Travascio *et al.*, 1998; Travascio *et al.*, 2001). The observed hyperchromicity and sharpening of the porphyrin Soret absorption upon binding to a quadruplex has previously been attributed to the porphyrin being in a more hydrophobic environment (Li and Sen, 1998; Travascio *et al.*, 1998). The methylation and carboxyethylation probing (using dimethyl sulfate and diethylpyrocarbonate) data obtained with the CH6A aptamer are fully consistent with the conserved guanine-rich region (a) being the hemin-binding site and (b) forming a guanine-quadruplex folded structure. Although the precise mode of interaction of hemin with CH6A DNA was not yet clear; a key feature of *this* aptamer (as compared to those analyzed in previous work (Li *et al.*, 1996; Travascio *et al.*, 1999; Travascio *et al.*, 1998; Travascio *et al.*, 2001)) was that the formation of its folded quadruplex structure required the binding of hemin. In other words, hemin binding to CH6 or its deletion mutants appeared to operate *via* an induced fit mechanism, where ligand-binding itself contributed to the formation of the binding site.

Analysis of the visible region of the DNA-hemin complex spectrum, shown in Figure 5-9, indicated a marked similarity with that of the previously characterized PS2M DNA-hemin complex (Li *et al.*, 1996; Travascio *et al.*, 1999; Travascio *et al.*, 1998; Travascio *et al.*, 2001). Absorption peaks were observed at 630 nm and 500nm, corresponding to the D and E bands respectively, along with a general decrease in the α band region. Travascio and Sen (Travascio *et al.*, 1998) had previously attributed these changes to the presence of ligand-to-metal charge transfer transitions, with the heme being a hexacoordinate high-spin species. Extensive EPR work had demonstrated that axial ligation of the iron center was mediated by a water molecule and, probably a guanine residue within the oligonucleotide (Travascio *et al.*, 2001; Witting *et al.*, 2001).

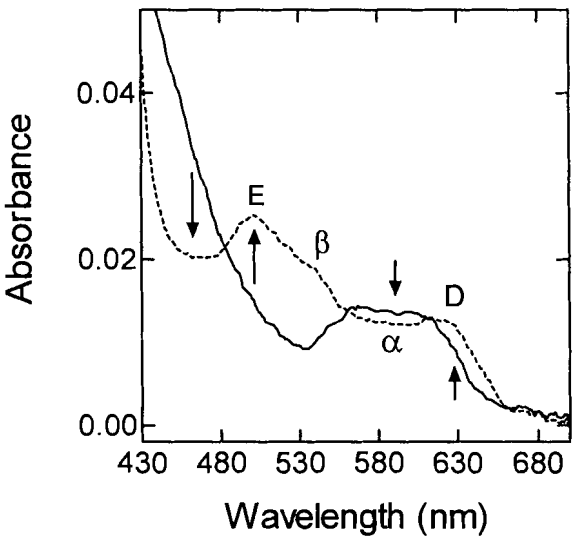


Figure 5-9 Spectra of the visible region of DNA-complexed hemin and free hemin. 1.5 μM hemin (solid lines) in buffer A (see Materials and Methods) and 20 μM CH6A DNA folded with 1.5 μM hemin in buffer A (dotted line).

Given the similarity of the spectra of the previously characterized PS2M-heme complexes to the DNA-heme complex reported here, as well as the guanine-rich natures of both PS2.M and CH6A, it is expected that the heme in the CH6A-hemin complex also exists as a high-spin hexacoordinate species. Further experiments will help to confirm this.

5.4.2 The Binding of Cytochrome *c* to the DNA Aptamer

The binding of cytochrome *c* to its normal biological redox partners has been studied using difference spectroscopy methods, under ionic conditions similar to the ones used in this study. The binding of cytochrome *c* to cytochrome *c* peroxidase was determined to have K_d of 2 mM at an ionic strength of 200 mM (Erman and Vitello, 1980). K_d values for the binding of cytochrome *c* to cytochrome *c* oxidase ranged from 0.3 – 3.0 μM at an ionic strength of 100 mM and 8 nM - 1 μM at a 50 mM ionic strength (Michel and Bosshard, 1984). The association of these proteins is highly sensitive to the ionic strength of the solution, with optimal binding (as might be expected for proteins interacting mainly electrostatically) at very low salt concentrations. The dissociation constants that we have measured for the binding of cytochrome *c* to our DNA:hemin complex are within the range of these values.

The dimethyl sulfate - and diethylpyrocarbonate - protection studies, which did not yield a clear footprint of the binding of cytochrome *c*, suggested that interaction of the protein with the DNA was not primarily through contacts to the 'major groove face' of individual adenine or guanine bases. Footprinting using hydroxyl radicals, however, indicated that the binding of cytochrome *c* rendered the sugar-phosphate backbone of a stretch of 12-15 nucleotides on CH6A significantly less accessible to solvent. According

to the folding scheme shown in Figure 5-2, and the diethylpyrocarbonate -protection results (Figure 5-7) elements of this sequence stretch should fold to form a duplex. However, the hydroxyl footprinting evidence did suggest that the cytochrome *c* interaction was mainly with one of the two strands (nucleotides ~40 to ~55 of CH6A, but not so well with nucleotides ~10 to ~25-- see Figure 5-7 & Figure 5-8) of the putative part-duplex flanking the hemin-binding quadruplex. To conceptualize how a single molecule of cytochrome *c* (with a diameter of ~40 Å at its longest axis) might produce such a 15 nucleotide footprint, the DNA could be partially wrapped around the protein, with the single stranded regions affording flexibility to kink the duplex elements. Figure 5-10 shows a model for this proposed interaction. Horse heart cytochrome *c* is known to contain 19 lysine and 2 arginine residues, whose positively charged side chains could interact electrostatically with the negatively charged phosphates of the DNA backbone. Such a purely electrostatic interaction, however, would be expected to be somewhat non-specific. In studies of sequence-specific DNA-binding proteins it has been found that direct and water-mediated hydrogen bonds between the protein side chains and the DNA bases and backbone are important for specific recognition (Pabo and Sauer, 1992; Seeman *et al.*, 1976). Studies on RNA interactions with proteins (Aboul-ela *et al.*, 1995; Aboul-ela *et al.*, 1996; Leulliot and Varani, 2001; Puglisi *et al.*, 1992) have indicated that along with significant conformational changes that occur within the RNA itself, there are also base-amino acid side chain contacts that afford specificity. It is also possible that in the CH6A-cytochrome interaction a degree of specificity may be afforded by direct (possibly hydrogen-bonding) interactions between the cytochrome side chains and the DNA phosphate groups as well as functionalities on the DNA bases that are not subject to

the dimethyl sulfate and diethylpyrocarbonate footprinting assays. The role of the G-quadruplex domain in positioning the cytochrome *c* on CH6A is hard to gauge at this point, although it is important to emphasize that G-quadruplexes have twice the negative charge density per unit length compared to double helices (Wellinger and Sen, 1997; Williamson, 1994); therefore, in terms of having a higher electrostatic potential per unit length (relative to a duplex), a G-quadruplex would be an excellent site for interaction by the positively charged cytochrome *c*. There are also a number of proteins that have been reported to bind G-quadruplex structures (Arimondo *et al.*, 2000; Fang and Cech, 1993; Jing *et al.*, 2000; Macaya *et al.*, 1993; Oliver *et al.*, 2000), including: thrombin (Fang and Cech, 1993), HIV1 integrase (Jing *et al.*, 2000) and fd gene 5 protein (Oliver *et al.*, 2000). However, in our studies the hydroxyl radical footprint of cytochrome *c* binding (Figure 5-8) does not indicate extensive backbone sugar-phosphate protection for the quadruplex nucleotides; moreover, interactions of the protein with the grooves of the quadruplex are not amenable to footprinting by dimethyl sulfate or diethylpyrocarbonate, since the dimethyl sulfate - and diethylpyrocarbonate -reactive guanine N7 functionalities are already resistant to modification owing to their participation in quartet formation.

5.4.3 Allostery in the Binding of the Two Ligands

Recently, a number of papers have reported allosteric behaviour in nucleic acids. Most of these involve the binding of a single ligand to a nucleic acid molecule, with a concomitant change in the structure and/or properties of the nucleic acid distal from the ligand-binding site. The allosteric variants of the hammerhead ribozyme developed by

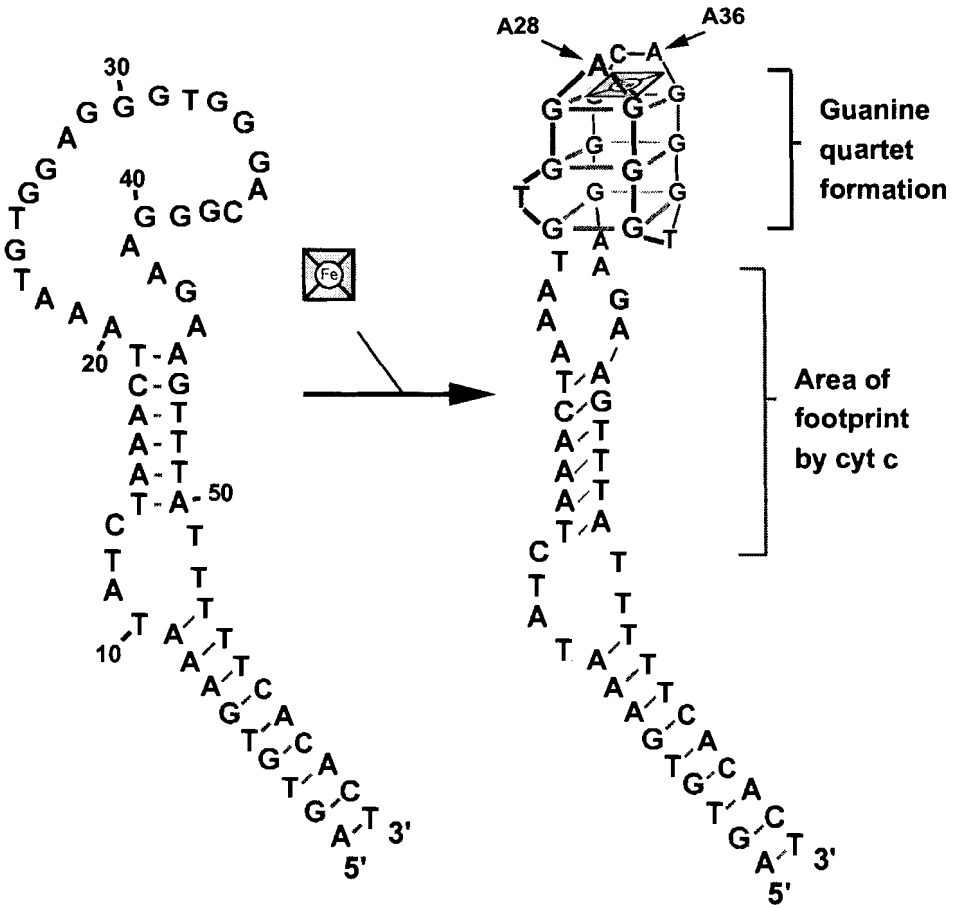


Figure 5-10 Proposed structural model for the binding of cytochrome and hemin to CH6A DNA.

Breaker and colleagues are the best examples of this (reviewed in (Soukup and Breaker, 2000)). Other examples include the selection of an allosteric RNA-cleaving ribozyme responsive to doxycycline (Piganeau *et al.*, 2001), and the selection of an allosteric RNA ligase responsive to an ATP cofactor (Robertson and Ellington, 1999).

Recently, however, Jose *et al.* (Jose *et al.*, 2001) have reported the construction of a variant of the hammerhead ribozyme that requires the binding of two effectors, flavin mononucleotide (FMN) and theophylline, for catalysis of phosphodiester bond cleavage. In their modular design the hammerhead motif as a catalytic entity was conjoined to two different and proximal binding sites for the two effectors. The hammerhead motif has also been used in developing dimeric ribozymes, termed 'maxizymes,' that can respond to RNA effector molecules both *in vitro* and *in vivo* (Warashina et al., 2000). In a different approach, two separate RNA selection pools for binding to Cibacron blue and cholic acid, respectively, were fused together (Wu and Curran, 1999). The resultant pool was then re-selected for allosteric behavior, such that the binding of one ligand negatively affected the binding of the other. In our system, as shown above, the binding of hemin to the DNA aptamer stimulates the binding of cytochrome *c* to the aptamer.

Figure 5-11 shows a thermodynamic cycle for the binding of the two ligands to the aptamer, i.e. the two possible pathways to the ternary complex, CH6A-hemin-cytochrome *c*. The two pathways describe a different order of binding of the two ligands. In this paper, we report values for three of the four equilibrium (association) constants, K_H, K_C, and K_{HC}, shown in Figure 5-11. I was unable to find a convenient or reliable method for measuring the fourth equilibrium constant, K_{CH}. In this chapter, I report that $K_{HC} > K_C$ (i.e. the hemin-DNA complex binds cytochrome *c* more strongly than does the

DNA itself). The thermodynamic cycle therefore mandates that $K_{CH} > K_H$ (in other words the cytochrome *c*-DNA complex should bind hemin more strongly than the uncomplexed DNA binds hemin). Future experiments would attempt to measure directly the equilibrium constant K_{CH}.

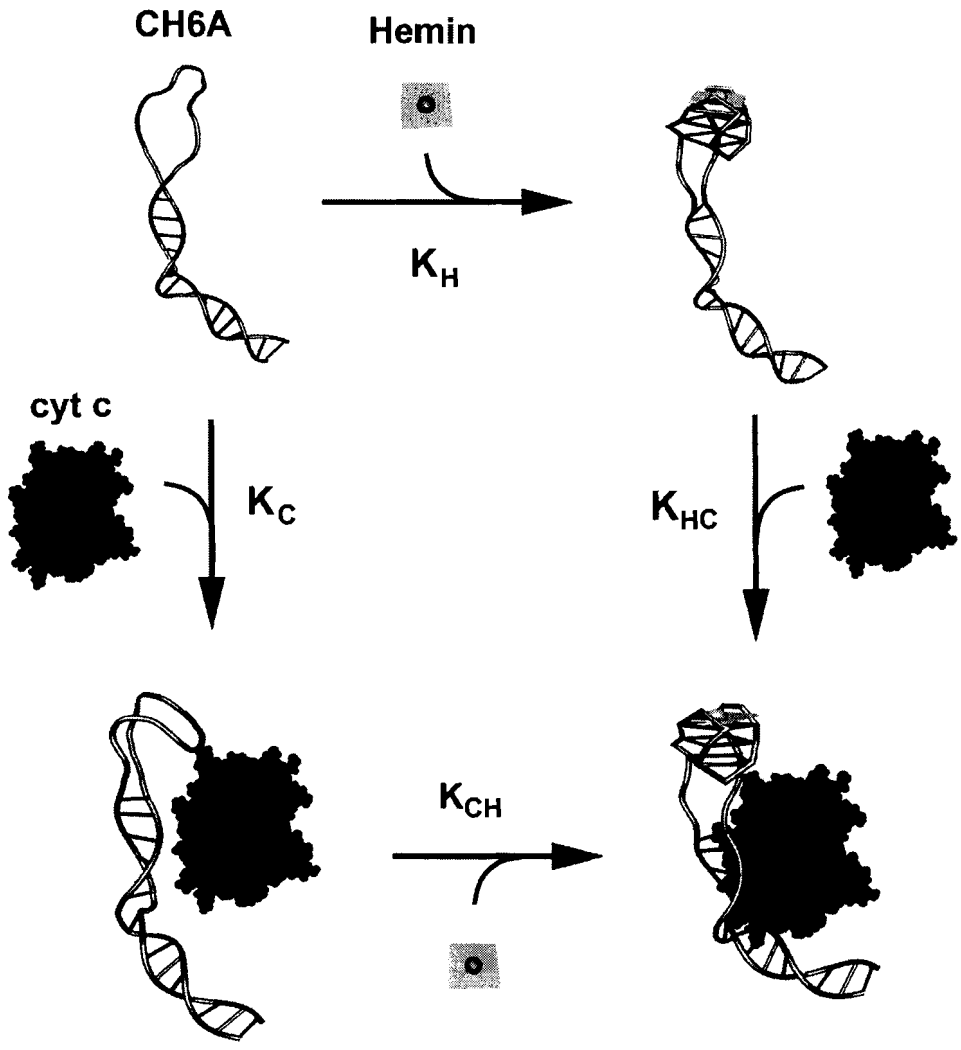


Figure 5-11 Thermodynamic pathways towards the formation of the ternary complex, CH6A-hemin-cytochrome *c*.

5.4.4 The Spatial Relationship Between Hemin and Cytochrome *c* Within the Ternary Complex.

The binding of CH6A DNA (in the absence *or* presence of complexed hemin) to cytochrome *c* was observed to perturb the electronic absorption spectrum of cytochrome *c*. This phenomenon has been observed in other studies, of cytochrome *c* binding to its protein redox partners cytochrome b_5 (Mauk *et al.*, 1982), cytochrome *c* oxidase (Michel and Bosshard, 1984) and cytochrome *c* peroxidase (Erman and Vitello, 1980), as well as to monoclonal antibodies specific for cytochrome *c* (Raman *et al.*, 1992). Positive difference spectra in the Soret peak are believed to be the results of a change in the local environment of the exposed heme crevice of cytochrome *c* (Raman *et al.*, 1992). The data presented in this paper suggest that a DNA-hemin complex is able to perturb the cytochrome *c* spectrum in a way comparable to the perturbation caused by the docking to cytochrome *c* of its various hemoprotein electron-transfer partners. This raises the possibility of electron transfer between the two heme centres in this ternary complex, separated presumably by both protein and DNA, and raises questions of what the most optimal path might be for such a transfer. Future experiments will be sought on the ternary complex, including spectroscopic studies and determining the distance between the two iron centers.

5.4.5 Summary

DNA aptamers were selected for their ability to bind simultaneously to the protein cytochrome *c* and to the metalloporphyrin hemin. Such aptamers each contained a conserved guanine-rich core, analogous to sequences shown previously to form a hemin-binding site when folded. The detailed study of CH6A indicated that in the presence of

hemin the guanine-rich core of the aptamer folded to form a guanine quadruplex. Both hemin and potassium ions were required for this folding. The binding of fully oxidized cytochrome *c* to this DNA-hemin complex resulted in an absorbance difference spectrum in the Soret region, which could be used as an indicator of binding behavior. It was found that cytochrome *c* bound more tightly to the folded CH6A DNA-hemin complex than to the folded CH6A DNA alone. Footprinting experiments showed the binding site of the cytochrome *c* to be a partial duplex element of the aptamer, immediately flanking its guanine-rich hemin-binding site. The order of addition of hemin and cytochrome *c* appeared not to affect either the formation rate or the structure of the final ternary complex. The ternary complex represents the docking of a nucleic acid-heme complex to cytochrome *c* (a protein-heme complex). Future experiments will focus on investigating the spatial arrangement between cytochrome *c* and CH6A and ultimately, electron-transfer between the two iron centers through intervening protein and DNA. The quenching of fluorescent porphyrin molecules bound to the DNA could be used to determine the distances between the two porphyrins. Ultimately, it would be of interest to observe the optimal path of electron transfer between the two heme centres.

CHAPTER 6

CONCLUSION

6.1 Summary of Results

The goal of this thesis was to explore both the photocatalytic potential of nucleic acids and the ability of nucleic acids to bind to redox-active proteins. The discovery of ribozymes and the development of *in vitro* selection techniques have allowed researchers to investigate the catalytic potential of nucleic acids, as well as their ability to bind to ligands. In the first part of my thesis, I described the selection and characterisation of a novel DNAzyme that is able to catalyse a photochemical reaction. *In vitro* selection was used to obtain deoxyribozymes that had the remarkable ability to use light to repair thymine dimers without the aid of a photosensitizer. Subsequent characterisation of the sensitizer-independent deoxyribozyme determined that thymine dimer repair correlated to the folding of an intramolecular guanine quadruplex. Deletion constructs of the enzyme, which truncated sequences from either end, showed that the 3' binding arm of the enzyme could be removed while maintaining catalytic activity. Cross-linking studies located the thymine dimer close to the guanine residues of the guanine quadruplex. UV1C could not effectively repair RNA-containing substrates containing uracil dimers. The efficiency of repair was greatly affected by type of the sugar moiety present on the pyrimidine dimer, showing highest preference to deoxyribose sugars. This indicated that UV1C was highly specific for the thymidine dimer substrate. In this sense, UV1C

exhibited properties of a "classical" enzyme, one that has high specificity for its substrate. Based on thermodynamic calculations it was determined that electron donation from a photoexcited guanine was thermodynamically feasible. Substitution of guanine residues located within the quartet with inosine resulted in a reduced activity at most positions. Also, the absence of detectable oxidative damage to guanines may have indicated a fast rate of back electron transfer from the thymine dimer following repair.

In a separate project, *in vitro* selection was used to generate a DNA aptamer that simultaneously bound to the electron transfer protein cytochrome *c*, and to the iron-containing porphyrin, hemin. Chemical probing studies indicated the presence of a guanine quadruplex in the structure with complementary stems. Spectroscopic experiments indicated that the DNA aptamer affects the local environment of the heme in cytochrome *c* in a manner similar to that to its natural protein-binding partner. The binding site of cytochrome *c* onto the DNA was revealed by hydroxyl radical footprinting.

6.2 Implications to the RNA World

Both of the projects presented in my thesis bear relevance to the RNA World Hypothesis. The finding that DNA could photorepair thymine dimers in the absence of extraneous sensitizing cofactors, suggests that photochemical reactions might well have constituted a significant part of the chemical/catalytic repertoire of a putative RNA World. Interestingly, many electron transfer proteins contain nucleotide-like cofactors, such as FADH as found in photolyase. A primordial role for the formation and breakage of intermolecular thymine dimers has been hypothesized by Lewis & Hanawalt (Lewis and Hanawalt, 1982). The atmosphere of the early Earth was believed to have been

subjected to high levels of ultraviolet radiation, owing to the absence of a stratospheric ozone layer. This would have posed a serious threat to a largely nucleic acid population of an RNA World. A string of nucleotides that is able to repair itself from UV damage could have implications for the survival of early life forms.

A theoretical gap exists between the proposed first signs of life of a putative RNA World (Gilbert, 1986) and the first cellular life possibly resembling modern archaebacteria. Many of these bacteria were able to survive in the extreme and anaerobic conditions of the early Earth environment, use sulfur and nitrogen as food sources, and survive in extreme heat, high salt, or high pH. Extant archaebacteria that live under these conditions have evolved proteins that bind metalloporphyrins as well as other redox cofactors to carry out their reduction/ oxidation reactions (Schafer *et al.*, 1999; Thauer, 1998). In a possible transition from the RNA world to a protein-nucleic acid world, some period may have existed where nucleic acids, with the help of cofactors, may have participated in the redox processes that would eventually be superseded by proteins. The ability of a DNA molecule to bind hemin and then recognize an electron transfer protein offers a prototype for investigating the ability of nucleic acids to participate in electron transfer reactions.

BIBLIOGRAPHY

Aboul-ela, F., Karn, J. and Varani, G. (1995) The structure of the human immunodeficiency virus type-1 TAR RNA reveals principles of RNA recognition by Tat protein. *J. Mol. Biol.*, **253**, 313-332.

Aboul-ela, F., Karn, J. and Varani, G. (1996) Structure of HIV-1 TAR RNA in the absence of ligands reveals a novel conformation of the trinucleotide bulge. *Nucleic Acids Res.*, **24**, 3974-3981.

Arimondo, P.B., Riou, J.F., Mergny, J.L., Tazi, J., Sun, J.S., Garestier, T. and Helene, C. (2000) Interaction of human DNA topoisomerase I with G-quartet structures. *Nucleic Acids Res.*, **28**, 4832-4838.

Ban, N., Nissen, P., Hansen, J., Moore, P.B. and Steitz, T.A. (2000) The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. *Science*, **289**, 905-920.

Barbier, B., Charlier, M. and Maurizot, J.C. (1984) Photochemical cross-linking of lac repressor to nonoperator 5-bromouracil-substituted DNA. *Biochemistry*, **23**, 2933-2939.

Bartel, D.P. and Szostak, J.W. (1993) Isolation of new ribozymes from a large pool of random sequences. *Science*, **261**, 1411-1418.

Bartel, D.P. and Unrau, P.J. (1999) Constructing an RNA world. *Trends Cell Biol.*, **9**, M9-M13.

Behrens, C. and Carell, T. (2003) Excess electron transfer in flavin-capped, thymine dimer-containing DNA hairpins. *Chem. Comm.*, 1632-1633.

Benner, S.A. and Ellington, A.D. (1991) RNA world. *Science*, **252**, 1232.

Benner, S.A., Ellington, A.D. and Tauer, A. (1989) Modern metabolism as a palimpsest of the RNA world. *Proc. Natl. Acad. Sci. U.S.A.*, **86**, 7054-7058.

Bhattacharyya, A. and Lilley, D.M. (1989) The contrasting structures of mismatched DNA sequences containing looped-out bases (bulges) and multiple mismatches (bubbles). *Nucleic Acids Res.*, **17**, 6821-6840.

Bock, L.C., Griffin, L.C., Latham, J.A., Vermaas, E.H. and Toole, J.J. (1992) Selection of single-stranded DNA molecules that bind and inhibit human thrombin. *Nature*, **355**, 564-566.

Brautigan, D.L., Ferguson-Miller, S. and Margoliash, E. (1978) Mitochondrial cytochrome c: preparation and activity of native and chemically modified cytochromes *c*. *Methods Enzymol.*, **53**, 128-164.

Breaker, R.R. (1997) *In Vitro* Selection of Catalytic Polynucleotides. *Chem. Rev.*, **97**, 371-390.

Breaker, R.R. and Joyce, G.F. (1994) A DNA enzyme that cleaves RNA. *Chem. Biol.*, **1**, 223-229.

Breaker, R.R. and Joyce, G.F. (1995) A DNA enzyme with Mg(2+)-dependent RNA phosphoesterase activity. *Chem. Biol.*, **2**, 655-660.

Burrows, C.J. and Muller, J.G. (1998) Oxidative Nucleobase Modifications Leading to Strand Scission. *Chem. Rev.*, **98**, 1109-1152.

Cai, J., Yang, J. and Jones, D.P. (1998) Mitochondrial control of apoptosis: the role of cytochrome *c*. *Biochim. Biophys. Acta.*, **1366**, 139-149.

Cairns, M.J., Saravolac, E.G. and Sun, L.Q. (2002) Catalytic DNA: a novel tool for gene suppression. *Curr. Drug Targets*, **3**, 269-279.

Calvert, J.G. and Pitts, J.N. (1966) *Photochemistry*. Wiley, New York.

Carell, T., Burgdorf, L.T., Kundu, L.M. and Cichon, M. (2001) The mechanism of action of DNA photolyases. *Curr. Opin. Chem. Biol.*, **5**, 491-498.

Carmi, N., Balkhi, S.R. and Breaker, R.R. (1998) Cleaving DNA with DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **95**, 2233-2237.

Cech, T.R. and Bass, B.L. (1986) Biological catalysis by RNA. *Annu. Rev. Biochem.*, **55**, 599-629.

Cech, T.R., Herschlag, D., Piccirilli, J.A. and Pyle, A.M. (1992) RNA catalysis by a group I ribozyme. Developing a model for transition state stabilization. *J. Biol. Chem.*, **267**, 17479-17482.

Chapman, K.B. and Szostak, J.W. (1995) Isolation of a ribozyme with 5'-5' ligase activity. *Chem. Biol.*, **2**, 325-333.

Chinnapen, D.J. and Sen, D. (2004) A deoxyribozyme that harnesses light to repair thymine dimers in DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 65-69.

Churcher, M.J., Lamont, C., Hamy, F., Dingwall, C., Green, S.M., Lowe, A.D., Butler, J.G., Gait, M.J. and Karn, J. (1993) High affinity binding of TAR RNA by the human immunodeficiency virus type-1 tat protein requires base-pairs in the RNA stem and amino acid residues flanking the basic region. *J. Mol. Biol.*, **230**, 90-110.

Ciesiolka, J. and Yarus, M. (1996) Small RNA-divalent domains. *RNA*, **2**, 785-793.

Cobaleda, C. and Sanchez-Garcia, I. (2001) RNase P: from biological function to biotechnological applications. *Trends Biotechnol.*, **19**, 406-411.

Cochran, A.G., Sugasawara, R. and Schultz, P.G. (1988) Photosensitized cleavage of a thymine dimer by an antibody. *J. Am. Chem. Soc.*, **110**, 7888-7890.

Conn, M.M., Prudent, J.R. and Schultz, P.G. (1996) Porphyrin metalation catalyzed by a small RNA molecule. *J. Am. Chem. Soc.*, **118**, 7012-7013.

Costa, R.M., Chigancas, V., Galhardo Rda, S., Carvalho, H. and Menck, C.F. (2003) The eukaryotic nucleotide excision repair pathway. *Biochimie*, **85**, 1083-1099.

Cuenoud, B. and Szostak, J.W. (1995) A DNA metalloenzyme with DNA ligase activity. *Nature*, **375**, 611-614.

Dandliker, P.J., Holmlin, R.E. and Barton, J.K. (1997) Oxidative thymine dimer repair in the DNA helix. *Science*, **275**, 1465-1468.

Dandliker, P.J., Nunez, M.E. and Barton, J.K. (1998) Oxidative charge transfer To repair thymine dimers and damage guanine bases in DNA assemblies containing tethered metallointercalators. *Biochemistry*, **37**, 6491-6502.

Davis, K.A., Abrams, B., Lin, Y. and Jayasena, S.D. (1996) Use of a high affinity DNA ligand in flow cytometry. *Nucleic Acids Res.*, **24**, 702-706.

Dieckmann, T., Butcher, S.E., Sassanfar, M., Szostak, J.W. and Feigon, J. (1997) Mutant ATP-binding RNA aptamers reveal the structural basis for ligand binding. *J. Mol. Biol.*, **273**, 467-478.

Dieckmann, T., Suzuki, E., Nakamura, G.K. and Feigon, J. (1996) Solution structure of an ATP-binding RNA aptamer reveals a novel fold. *RNA*, **2**, 628-640.

Dotse, A.K., Boone, E.K. and Schuster, G.B. (2000) Remote cis-syn-thymine [2 + 2] dimers are not repaired by radical cations migrating in duplex DNA. *J. Am. Chem. Soc.*, **122**, 6825-6833.

Ekland, E.H. and Bartel, D.P. (1996) RNA-catalysed RNA polymerization using nucleoside triphosphates. *Nature*, **383**, 192.

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

Erman, J.E. and Vitello, L.B. (1980) The binding of cytochrome c peroxidase and ferricytochrome c. A spectrophotometric determination of the equilibrium association constant as a function of ionic strength. *J. Biol. Chem.*, **255**, 6224-6227.

Falk, J.E. and Smith, K.M. (1975) *Porphyrins and metalloporphyrins : a new edition based on the original volume by J. E. Falk.* Elsevier Scientific Pub. Co., Amsterdam ; New York.

Fang, G. and Cech, T.R. (1993) Characterization of a G-quartet formation reaction promoted by the beta-subunit of the Oxytricha telomere-binding protein. *Biochemistry*, **32**, 11646-11657.

Fedorova, O., Su, L.J. and Pyle, A.M. (2002) Group II introns: highly specific endonucleases with modular structures and diverse catalytic functions. *Methods*, **28**, 323-335.

Feldman, A.R. and Sen, D. (2001) A new and efficient DNA enzyme for the sequence-specific cleavage of RNA. *J. Mol. Biol.*, **313**, 283-294.

Fortini, P., Pascucci, B., Parlanti, E., D'Errico, M., Simonelli, V. and Dogliotti, E. (2003) The base excision repair: mechanisms and its relevance for cancer susceptibility. *Biochimie*, **85**, 1053-1071.

Geyer, C.R. and Sen, D. (1997) Evidence for the metal-cofactor independence of an RNA phosphodiester-cleaving DNA enzyme. *Chem. Biol.*, **4**, 579-593.

Geyer, C.R. and Sen, D. (2000) Use of intrinsic binding energy for catalysis by a cofactor-independent DNA enzyme. *J. Mol. Biol.*, **299**, 1387-1398.

Giese, B. (2002a) Electron transfer in DNA. *Curr. Opin. Chem. Biol.*, **6**, 612-618.

Giese, B. (2002b) Long-distance electron transfer through DNA. *Annu. Rev. Biochem.*, **71**, 51-70.

Giese, B., Wessely, S., Sporman, M., Lindemann, L., Meggars, E. and Michel-Beyerle, M.E. (1999) On the mechanism of long-range electron transfer through DNA. *Angew. Chem. Int. Ed.*, **38**, 996-998.

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

Gold, L., Brown, D., He, Y., Shtatland, T., Singer, B.S. and Wu, Y. (1997) From oligonucleotide shapes to genomic SELEX: novel biological regulatory loops. *Proc. Natl. Acad. Sci. U.S.A.*, **94**, 59-64.

Golden, M.C., Collins, B.D., Willis, M.C. and Koch, T.H. (2000) Diagnostic potential of PhotoSELEX-evolved ssDNA aptamers. *J. Biotechnol.*, **81**, 167-178.

Gorodkin, J., Heyer, L.J. and Stormo, G.D. (1997a) Finding common sequence and structure motifs in a set of RNA sequences. *Proc. Int. Conf. Intell. Syst. Mol. Biol.*, **5**, 120-123.

Gorodkin, J., Heyer, L.J. and Stormo, G.D. (1997b) Finding the most significant common sequence and structure motifs in a set of RNA sequences. *Nucleic Acids Res.*, **25**, 3724-3732.

Green, D.R. and Reed, J.C. (1998) Mitochondria and apoptosis. *Science*, **281**, 1309-1312.

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.

Hardin, C.C., Watson, T., Corregan, M. and Bailey, C. (1992) Cation-dependent transition between the quadruplex and Watson-Crick hairpin forms of d(CGCG3GCG). *Biochemistry*, **31**, 833-841.

Helene, C. and Charlier, M. (1977) Photosensitized splitting of pyrimidine dimers by indole derivatives and by tryptophan-containing oligopeptides and proteins. *Photochem. Photobiol.*, **25**, 429-434.

Henderson, E., Hardin, C.C., Walk, S.K., Tinoco, I., Jr. and Blackburn, E.H. (1987) Telomeric DNA oligonucleotides form novel intramolecular structures containing guanine-guanine base pairs. *Cell*, **51**, 899-908.

Higgins, D.G., Thompson, J.D. and Gibson, T.J. (1996) Using CLUSTAL for multiple sequence alignments. *Methods Enzymol.*, **266**, 383-402.

Hitomi, K., Kim, S.T., Iwai, S., Harima, N., Otoshi, E., Ikenaga, M. and Todo, T. (1997) Binding and catalytic properties of Xenopus (6-4) photolyase. *J. Biol. Chem.*, **272**, 32591-32598.

Hix, S., Kadiiska, M.B., Mason, R.P. and Augusto, O. (2000) In vivo metabolism of tert-butyl hydroperoxide to methyl radicals. EPR spin-trapping and DNA methylation studies. *Chem. Res. Toxicol.*, **13**, 1056-1064.

Huizenga, D.E. and Szostak, J.W. (1995) A DNA aptamer that binds adenosine and ATP. *Biochemistry*, **34**, 656-665.

Husain, I., Griffith, J. and Sancar, A. (1988) Thymine dimers bend DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **85**, 2558-2562.

Hush, N.S. and Cheung, A.S. (1975) Ionization potentials and donor properties of nucleic acid bases and related compounds. *Chem. Phys. Lett.*, **34**, 11-13.

Illangasekare, M., Sanchez, G., Nickles, T. and Yarus, M. (1995) Aminoacyl-RNA synthesis catalyzed by an RNA. *Science*, **267**, 643-647.

Impey, H.L., Applegate, T.L., Haughton, M.A., Fuery, C.J., King, J.E. and Todd, A.V. (2000) Factors that influence deoxyribozyme cleavage during polymerase chain reaction. *Anal. Biochem.*, **286**, 300-303.

Ito, T. and Rokita, S.E. (2003) Excess electron transfer from an internally conjugated aromatic amine to 5-bromo-2'-deoxyuridine in DNA. *J. Am. Chem. Soc.*, **125**, 11480-11481.

Ito, T. and Rokita, S.E. (2004) Criteria for efficient transport of excess electrons in DNA. *Angew. Chem. Int. Ed.*, **43**, 1839-1842.

Jacobsen, J.R., Cochran, A.G., Stephans, J.C., King, D.S. and Schultz, P.G. (1995) Mechanistic Studies of Antibody-Catalyzed Pyrimidine Dimer Photocleavage. *J. Am. Chem. Soc.*, **117**, 5453-5461.

Jaschke, A. (2001) Artificial ribozymes and deoxyribozymes. *Curr. Opin. Struct. Biol.*, **11**, 321-326.

Jaschke, A. and Seelig, B. (2000) Evolution of DNA and RNA as catalysts for chemical reactions. *Curr. Opin. Chem. Biol.*, **4**, 257-262.

Jencks, W.P. (1987) *Catalysis in Chemistry and Enzymology*. Dover, New York.

Jiang, F., Kumar, R.A., Jones, R.A. and Patel, D.J. (1996) Structural basis of RNA folding and recognition in an AMP-RNA aptamer complex. *Nature*, **382**, 183-186.

Jing, N., Marchand, C., Liu, J., Mitra, R., Hogan, M.E. and Pommier, Y. (2000) Mechanism of inhibition of HIV-1 integrase by G-tetrad-forming oligonucleotides in Vitro. *J. Biol. Chem.*, **275**, 21460-21467.

Job, P. (1928) *Ann. Chim.*, **9**, 113-203.

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**, 1319-1325.

Jortner, J., Bixon, M., Langenbacher, T. and Michel-Beyerle, M.E. (1998) Charge transfer and transport in DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **95**, 12759-12765.

Jose, A.M., Soukup, G.A. and Breaker, R.R. (2001) Cooperative binding of effectors by an allosteric ribozyme. *Nucleic Acids Res.*, **29**, 1631-1637.

Joyce, G.F. (1989) Amplification, mutation and selection of catalytic RNA. *Gene*, **82**, 83-87.

Joyce, G.F. (2004) Directed evolution of nucleic acid enzymes. *Annu. Rev. Biochem.*, **73**, 791-836.

Kavarnos, G.J. (1993) *Fundamentals of photoinduced electron transfer*. VCH Publishers, New York, NY.

Keniry, M.A. (2000) Quadruplex structures in nucleic acids. *Biopolymers*, **56**, 123-146.

Khachigian, L.M. (2004) DNAzymes as molecular agents that manipulate Egr-1 gene expression. *Biochem. Pharmacol.*, **68**, 1023-1025.

Kim, S.T., Heelis, P.F., Okamura, T., Hirata, Y., Mataga, N. and Sancar, A. (1991) Determination of rates and yields of interchromophore (folate-flavin) energy transfer and intermolecular (flavin-DNA) electron transfer in Escherichia coli photolyase by time-resolved fluorescence and absorption spectroscopy. *Biochemistry*, **30**, 11262-11270.

Kim, S.T., Li, Y.F. and Sancar, A. (1992) The third chromophore of DNA photolyase: Trp-277 of Escherichia coli DNA photolyase repairs thymine dimers by direct electron transfer. *Proc. Natl. Acad. Sci. U.S.A.*, **89**, 900-904.

Kim, S.T. and Sancar, A. (1991) Effect of base, pentose, and phosphodiester backbone structures on binding and repair of pyrimidine dimers by Escherichia coli DNA photolyase. *Biochemistry*, **30**, 8623-8630.

Kim, S.T., Sancar, A., Essenmacher, C. and Babcock, G.T. (1993) Time-resolved EPR studies with DNA photolyase: excited-state FADH0 abstracts an electron from Trp-306 to generate FADH-, the catalytically active form of the cofactor. *Proc. Natl. Acad. Sci. U.S.A.*, **90**, 8023-8027.

Komori, H., Masui, R., Kuramitsu, S., Yokoyama, S., Shibata, T., Inoue, Y. and Miki, K. (2001) Crystal structure of thermostable DNA photolyase: pyrimidine-dimer recognition mechanism. *Proc. Natl. Acad. Sci. U.S.A.*, **98**, 13560-13565.

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.

Leulliot, N. and Varani, G. (2001) Current topics in RNA-protein recognition: control of specificity and biological function through induced fit and conformational capture. *Biochemistry*, **40**, 7947-7956.

Lewis, R.J. and Hanawalt, P.C. (1982) Ligation of oligonucleotides by pyrimidine dimers--a missing 'link' in the origin of life? *Nature*, **298**, 393-396.

Li, Y. and Breaker, R.R. (1999) Phosphorylating DNA with DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **96**, 2746-2751.

Li, Y., Geyer, C.R. and Sen, D. (1996) Recognition of anionic porphyrins by DNA aptamers. *Biochemistry*, **35**, 6911-6922.

Li, Y., Liu, Y. and Breaker, R.R. (2000) Capping DNA with DNA. *Biochemistry*, **39**, 3106-3114.

Li, Y. and Sen, D. (1996) A catalytic DNA for porphyrin metallation. *Nat. Struct. Biol.*, **3**, 743-747.

Li, Y. and Sen, D. (1998) The modus operandi of a DNA enzyme: enhancement of substrate basicity. *Chem. Biol.*, **5**, 1-12.

Lin, C.H. and Patel, D.J. (1997) Structural basis of DNA folding and recognition in an AMP-DNA aptamer complex: distinct architectures but common recognition motifs for DNA and RNA aptamers complexed to AMP. *Chem. Biol.*, **4**, 817-832.

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

Lorsch, J.R. and Szostak, J.W. (1994) In vitro evolution of new ribozymes with polynucleotide kinase activity. *Nature*, **371**, 31-36.

Macaya, R.F., Schultze, P., Smith, F.W., Roe, J.A. and Feigon, J. (1993) Thrombin-binding DNA aptamer forms a unimolecular quadruplex structure in solution. *Proc. Natl. Acad. Sci. U.S.A.*, **90**, 3745-3749.

Marcus, R.A. and Sutin, N. (1985) Electron transfers in chemistry and biology. *Biochim. Biophys. Acta.*, **811**, 265-322.

Mathews, II, Padmanabhan, K.P., Tulinksy, A. and Sadler, J.E. (1994) Structure of a nonadecapeptide of the fifth EGF domain of thrombomodulin complexed with thrombin. *Biochemistry*, **33**, 13547-13552.

Matsunaga, T., Hieda, K. and Nikaido, O. (1991) Wavelength dependent formation of thymine dimers and (6-4) photoproducts in DNA by monochromatic ultraviolet light ranging from 150 to 365 nm. *Photochem. Photobiol.*, **54**, 403-410.

Mauk, A.G., Mauk, M.R., Moore, G.R. and Northrup, S.H. (1995) Experimental and theoretical analysis of the interaction between cytochrome *c* and cytochrome *b5*. *J. Bioenerg. Biomembr.*, **27**, 311-330.

Mauk, M.R., Reid, L.S. and Mauk, A.G. (1982) Spectrophotometric analysis of the interaction between cytochrome *b5* and cytochrome *c*. *Biochemistry*, **21**, 1843-1846.

Maxam, A.M. and Gilbert, W. (1977) A new method for sequencing DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **74**, 560-564.

McKay, D.B. and Wedekind, J.E. (1999) Small Ribozymes. In Gesteland, R.F., Cech, T.R. and Atkins, J.F. (eds.), *The RNA World*. Cold Spring Harbor Laboratory Press, New York, pp. 265-286.

Mees, A., Klar, T., Gnau, P., Hennecke, U., Eker, A.P., Carell, T. and Essen, L.O. (2004) Crystal structure of a photolyase bound to a CPD-like DNA lesion after in situ repair. *Science*, **306**, 1789-1793.

Mergny, J.L., Phan, A.T. and Lacroix, L. (1998) Following G-quartet formation by UV-spectroscopy. *FEBS Lett.*, **435**, 74-78.

Michel, B. and Bosshard, H.R. (1984) Spectroscopic analysis of the interaction between cytochrome *c* and cytochrome *c* oxidase. *J. Biol. Chem.*, **259**, 10085-10091.

Mignotte, B. and Vayssiere, J.L. (1998) Mitochondria and apoptosis. *Eur. J. Biochem.*, **252**, 1-15.

Miller, S.L. (1953) A production of amino acids under possible primitive earth conditions. *Science*, **117**, 528-529.

Miller, S.L. (1974) The atmosphere of the primitive earth and the prebiotic synthesis of amino acids. *Orig. Life*, **5**, 139-151.

Miller, S.L. and Urey, H.C. (1959) Organic compound synthesis on the primitive Earth. *Science*, **130**, 245-251.

Moore, P.B. (1999) Structural motifs in RNA. *Annu. Rev. Biochem.*, **68**, 287-300.

Morris, K.N., Jensen, K.B., Julin, C.M., Weil, M. and Gold, L. (1998) High affinity ligands from in vitro selection: complex targets. *Proc. Natl. Acad. Sci. U.S.A.*, **95**, 2902-2907.

Muller, J.G., Duarte, V., Hickerson, R.P. and Burrows, C.J. (1998) Gel electrophoretic detection of 7,8-dihydro-8-oxoguanine and 7, 8-dihydro-8-oxoadenine via oxidation by Ir (IV). *Nucleic Acids Res.*, **26**, 2247-2249.

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

Nissen, P., Ippolito, J.A., Ban, N., Moore, P.B. and Steitz, T.A. (2001) RNA tertiary interactions in the large ribosomal subunit: the A-minor motif. *Proc. Natl. Acad. Sci. U.S.A.*, **98**, 4899-4903.

Ogata, R. and Gilbert, W. (1977) Contacts between the lac repressor and the thymines in the lac operator. *Proc. Natl. Acad. Sci. U.S.A.*, **74**, 4973-4976.

Oliver, A.W., Bogdarina, I., Schroeder, E., Taylor, I.A. and Kneale, G.G. (2000) Preferential binding of fd gene 5 protein to tetraplex nucleic acid structures. *J. Mol. Biol.*, **301**, 575-584.

Pabo, C.O. and Sauer, R.T. (1992) Transcription factors: structural families and principles of DNA recognition. *Annu. Rev. Biochem.*, **61**, 1053-1095.

Pan, T. and Uhlenbeck, O.C. (1992) A small metalloribozyme with a two-step mechanism. *Nature*, **358**, 560-563.

Payne, G., Heelis, P.F., Rohrs, B.R. and Sancar, A. (1987) The active form of Escherichia coli DNA photolyase contains a fully reduced flavin and not a flavin radical, both in vivo and in vitro. *Biochemistry*, **26**, 7121-7127.

Pecourt, J.M., Peon, J. and Kohler, B. (2001) DNA excited-state dynamics: ultrafast internal conversion and vibrational cooling in a series of nucleosides. *J. Am. Chem. Soc.*, **123**, 10370-10378.

Pettigrew, G.W. and Moore, G.R. (1987) *Cytochromes c : biological aspects*. Springer-Verlag, Berlin ; New York.

Piganeau, N., Jenne, A., Thuillier, V. and Famulok, M. (2001) An Allosteric Ribozyme Regulated by Doxycyline. *Angew. Chem. Int. Ed.*, **40**, 3503.

Preuss, M., Schmidt, W.G., Seino, K., Furthmuller, J. and Bechstedt, F. (2004) Ground- and excited-state properties of DNA base molecules from plane-wave calculations using ultrasoft pseudopotentials. *J. Comput. Chem.*, **25**, 112-122.

Prudent, J.R., Uno, T. and Schultz, P.G. (1994) Expanding the scope of RNA catalysis. *Science*, **264**, 1924-1927.

Puglisi, J.D., Tan, R., Calnan, B.J., Frankel, A.D. and Williamson, J.R. (1992) Conformation of the TAR RNA-arginine complex by NMR spectroscopy. *Science*, **257**, 76-80.

Raman, C.S., Jemmerson, R., Nall, B.T. and Allen, M.J. (1992) Diffusion-limited rates for monoclonal antibody binding to cytochrome c. *Biochemistry*, **31**, 10370-10379.

Robertson, M.P. and Ellington, A.D. (1999) *In vitro* selection of an allosteric ribozyme that transduces analytes to amplicons. *Nat. Biotechnol.*, **17**, 62-66.

Rosen, G.M. (1999) *Free radicals : biology and detection by spin trapping*. Oxford University Press, New York.

Saenger, W. (1984) *Principles of nucleic acid structure*. Springer-Verlag, New York.

Sancar, A. (2003) Structure and function of DNA photolyase and cryptochrome blue-light photoreceptors. *Chem. Rev.*, **103**, 2203-2237.

SantaLucia, J., Jr. (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl. Acad. Sci. U.S.A.*, **95**, 1460-1465.

Santoro, S.W. and Joyce, G.F. (1997) A general purpose RNA-cleaving DNA enzyme. *Proc. Natl. Acad. Sci. U.S.A.*, **94**, 4262-4266.

Sassanfar, M. and Szostak, J.W. (1993) An RNA motif that binds ATP. *Nature*, **364**, 550-553.

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.

Scaiano, J.C. (1989) *CRC handbook of organic photochemistry*. CRC Press, Boca Raton, Fla.

Scannell, M.P., Fenick, D.J., Yeh, S.R. and Falvey, D.E. (1997) Model Studies of DNA Photorepair: Repuction Potentials of Thymine and Cytosine Cyclobutane Dimers Measured by Fluorescence Quenching. *J. Am. Chem. Soc.*, **119**, 1971-1977.

Schafer, G., Engelhard, M. and Muller, V. (1999) Bioenergetics of the Archaea. *Microbiol. Mol. Biol. Rev.*, **63**, 570-620.

Schultze, P., Macaya, R.F. and Feigon, J. (1994) Three-dimensional solution structure of the thrombin-binding DNA aptamer d(GGTTGGTGTGGTTGG). *J. Mol. Biol.*, **235**, 1532-1547.

Seeman, N.C., Rosenberg, J.M. and Rich, A. (1976) Sequence-specific recognition of double helical nucleic acids by proteins. *Proc. Natl. Acad. Sci. U.S.A.*, **73**, 804-808.

Seidel, C.A., Schulz, A. and Sauer, M.H. (1996) Nucleobase-specific quenching of fluorescent dyes. *J. Phys. Chem.*, **100**, 5541-5553.

Sen, D. and Gilbert, W. (1988) Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature*, **334**, 364-366.

Sen, D. and Gilbert, W. (1990) A sodium-potassium switch in the formation of four-stranded G4-DNA. *Nature*, **344**, 410-414.

Sheppard, T.L., Ordoukhanian, P. and Joyce, G.F. (2000) A DNA enzyme with N-glycosylase activity. *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 7802-7807.

Shih, I.H. and Been, M.D. (2002) Catalytic strategies of the hepatitis delta virus ribozymes. *Annu. Rev. Biochem.*, **71**, 887-917.

Sigurdsson, S.T., Thomson, J.B. and Eckstein, F. (1998) Small Ribozymes. In Simons, R.W. and Grunberg-Manago, M. (eds.), *RNA Structure and Function*. Cold Spring Harbor Laboratory Press, New York, pp. 339-376.

Smirnov, I. and Shafer, R.H. (2000) Lead is unusually effective in sequence-specific folding of DNA. *J. Mol. Biol.*, **296**, 1-5.

Smith, F.W., Schultze, P. and Feigon, J. (1995) Solution structures of unimolecular quadruplexes formed by oligonucleotides containing Oxytricha telomere repeats. *Structure*, **3**, 997-1008.

Soukup, G.A. and Breaker, R.R. (2000) Allosteric nucleic acid catalysts. *Curr. Opin. Struct. Biol.*, **10**, 318-325.

Sponer, J., Leszczynski, J. and Hobza, P. (1996) Structures and energies of hydrogen-bonded DNA base pairs: a nonempirical study with inclusion of electron correlation. *J. Phys. Chem.*, **100**, 1965-1974.

Steenken, S. and Jovanovic, S.V. (1997) How Easily Oxidizable Is DNA? One-Electron Reduction Potentials of Adenosine and Guanosine Radicals in Aqueous Solution. *J. Am. Chem. Soc.*, **119**, 617-618.

Stemp, E.D. and Barton, J.K. (2000) The flash-quench technique in protein-DNA electron transfer: reduction of the guanine radical by ferrocytochrome *c*. *Inorg. Chem.*, **39**, 3868-3874.

Stojanovic, M.N., Mitchell, T.E. and Stefanovic, D. (2002) Deoxyribozyme-based logic gates. *J. Am. Chem. Soc.*, **124**, 3555-3561.

Stonehuerner, J., Williams, J.B. and Millett, F. (1979) Interaction between cytochrome c and cytochrome b5. *Biochemistry*, **18**, 5422-5427.

Takagi, Y., Warashina, M., Stec, W.J., Yoshinari, K. and Taira, K. (2001) Recent advances in the elucidation of the mechanisms of action of ribozymes. *Nucleic Acids Res.*, **29**, 1815-1834.

Tamada, T., Kitadokoro, K., Higuchi, Y., Inaka, K., Yasui, A., de Ruiter, P.E., Eker, A.P. and Miki, K. (1997) Crystal structure of DNA photolyase from Anacystis nidulans. *Nat Struct Biol*, **4**, 887-891.

Thauer, R.K. (1998) Biochemistry of methanogenesis: a tribute to Marjory Stephenson. 1998 Marjory Stephenson Prize Lecture. *Microbiology*, **144**, 2377-2406.

Tian, Y., Adya, N., Wagner, S., Giam, C.Z., Green, M.R. and Ellington, A.D. (1995) Dissecting protein:protein interactions between transcription factors with an RNA aptamer. *RNA*, **1**, 317-326.

Travascio, P., Bennet, A.J., Wang, D.Y. and Sen, D. (1999) A ribozyme and a catalytic DNA with peroxidase activity: active sites versus cofactor-binding sites. *Chem. Biol.*, **6**, 779-787.

Travascio, P., Li, Y. and Sen, D. (1998) DNA-enhanced peroxidase activity of a DNA-aptamer-hemin complex. *Chem. Biol.*, **5**, 505-517.

Travascio, P., Witting, P.K., Mauk, A.G. and Sen, D. (2001) The peroxidase activity of a hemin--DNA oligonucleotide complex: free radical damage to specific guanine bases of the DNA. *J. Am. Chem. Soc.*, **123**, 1337-1348.

Tuerk, C. and Gold, L. (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science*, **249**, 505-510.

Tuschl, T., Sharp, P.A. and Bartel, D.P. (1998) Selection in vitro of novel ribozymes from a partially randomized U2 and U6 snRNA library. *EMBO J.*, **17**, 2637-2650.

Tuschl, T., Thomson, J.B. and Eckstein, F. (1995) RNA cleavage by small catalytic RNAs. *Curr. Opin. Struct. Biol.*, **5**, 296-302.

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

Unrau, P.J. and Bartel, D.P. (2003) An oxocarbenium-ion intermediate of a ribozyme reaction indicated by kinetic isotope effects. *Proc. Natl. Acad. Sci. U.S.A.*, **100**, 15393-15397.

Vande Berg, B.J. and Sancar, G.B. (1998) Evidence for dinucleotide flipping by DNA photolyase. *J. Biol. Chem.*, **273**, 20276-20284.

Vink, A.A. and Roza, L. (2001) Biological consequences of cyclobutane pyrimidine dimers. *J. Photochem. Photobiol. B*, **65**, 101-104.

Wagenknecht, H.A. (2003) Reductive electron transfer and transport of excess electrons in DNA. *Angew. Chem. Int. Ed. Engl.*, **42**, 2454-2460.

Walter, N.G. and Burke, J.M. (1998) The hairpin ribozyme: structure, assembly and catalysis. *Curr. Opin. Chem. Biol.*, **2**, 24-30.

Wan, C., Fiebig, T., Schiemann, O., Barton, J.K. and Zewail, A.H. (2000) Femtosecond direct observation of charge transfer between bases in DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 14052-14055.

Wang, K.Y., Krawczyk, S.H., Bischofberger, N., Swaminathan, S. and Bolton, P.H. (1993a) The tertiary structure of a DNA aptamer which binds to and inhibits thrombin determines activity. *Biochemistry*, **32**, 11285-11292.

Wang, K.Y., McCurdy, S., Shea, R.G., Swaminathan, S. and Bolton, P.H. (1993b) A DNA aptamer which binds to and inhibits thrombin exhibits a new structural motif for DNA. *Biochemistry*, **32**, 1899-1904.

Wang, S.Y. (1976) *Photochemistry and photobiology of nucleic acids*. Academic Press, New York.

Wang, W., Billen, L.P. and Li, Y. (2002) Sequence diversity, metal specificity, and catalytic proficiency of metal-dependent phosphorylating DNA enzymes. *Chem. Biol.*, **9**, 507-517.

Wang, Y., Hamasaki, K. and Rando, R.R. (1997) Specificity of aminoglycoside binding to RNA constructs derived from the 16S rRNA decoding region and the HIV-RRE activator region. *Biochemistry*, **36**, 768-779.

Warashina, M., Kuwabara, T. and Taira, K. (2000) Working at the cutting edge: the creation of allosteric ribozymes. *Structure Fold. Des.*, **8**, R207-212.

Wecker, M., Smith, D. and Gold, L. (1996) In vitro selection of a novel catalytic RNA: characterization of a sulfur alkylation reaction and interaction with a small peptide. *RNA*, **2**, 982-994.

Weidner, M.F., Millard, J.T. and Hopkins, P.B. (1989) Determination at single nucleotide resolution of the sequence specificity of DNA interstrand crosslinking agents in DNA fragments. *J. Am. Chem. Soc.*, **111**, 9270-9272.

Wellinger, R.E. and Sen, D. (1997) The DNA Structures at the Ends of Eukaryotic Chromosomes. *Eur. J. Cancer*, **33**, 735-749.

Wellinger, R.E. and Thoma, F. (1996) Taq DNA polymerase blockage at pyrimidine dimers. *Nucleic Acids Res.*, **24**, 1578-1579.

Wick, K.L. and Matthews, K.S. (1991) Interactions between lac repressor protein and site-specific bromodeoxyuridine-substituted operator DNA. Ultraviolet footprinting and protein-DNA cross-link formation. *J. Biol. Chem.*, **266**, 6106-6112.

Williams, K.P., Ciafre, S. and Tocchini-Valentini, G.P. (1995) Selection of novel Mg(2+)-dependent self-cleaving ribozymes. *EMBO J.*, **14**, 4551-4557.

Williamson, J.R. (1994) G-quartet structures in telomeric DNA. *Annu Rev Biophys Biomol Struct*, **23**, 703-730.

Williamson, J.R., Raghuraman, M.K. and Cech, T.R. (1989) Monovalent cation-induced structure of telomeric DNA: the G-quartet model. *Cell*, **59**, 871-880.

Willis, M.C., Hicke, B.J., Uhlenbeck, O.C., Cech, T.R. and Koch, T.H. (1993) Photocrosslinking of 5-iodouracil-substituted RNA and DNA to proteins. *Science*, **262**, 1255-1257.

Wilson, C. and Szostak, J.W. (1995) In vitro evolution of a self-alkylating ribozyme. *Nature*, **374**, 777-782.

Wilson, D.F., Vinogradov, S., Lo, L.W. and Huang, L. (1996) Oxygen dependent quenching of phosphorescence: a status report. *Adv. Exp. Med. Biol.*, **388**, 101-107.

Wilson, D.S. and Szostak, J.W. (1999) In vitro selection of functional nucleic acids. *Annu. Rev. Biochem.*, **68**, 611-647.

Witting, P.K., Travascio, P., Sen, D. and Mauk, A.G. (2001) A DNA oligonucleotide-hemin complex cleaves t-butyl hydroperoxide through a homolytic mechanism. *Inorg. Chem.*, **40**, 5017-5023.

Wu, L. and Curran, J.F. (1999) An allosteric synthetic DNA. *Nucleic Acids Res.*, **27**, 1512-1516.

Xu, Y. and Sugiyama, H. (2004) Highly efficient photochemical 2'-deoxyribonolactone formation at the diagonal loop of a 5-iodouracil-containing antiparallel G-quartet. *J. Am. Chem. Soc.*, **126**, 6274-6279.

Yeh, S.R. and Falvey, D.E. (1992) Model Studies of DNA Photorepair: Energetic Requirements for the Radical Anion Mechanism Determined by Fluorescence Quenching. *J. Am. Chem. Soc.*, **117**, 7709.

Zhang, B. and Cech, T.R. (1997) Peptide bond formation by in vitro selected ribozymes. *Nature*, **390**, 96-100.