INCREASING THE CHEMICAL FUNCTIONALITY OF DNA ENZYMES

by

Curtis Hon Ming Lam

B.Sc. (Honours), University of British Columbia, 2002

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

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES
(Chemistry)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)

January 2011

Abstract

Deoxyribonucleic acid (DNA) has long been known as the storage of genetic material. DNA is chemically stable and predominantly found as a double helix of antiparallel complementary strands. Due to the development of *in vitro* selection techniques, artificial single-stranded DNA molecules have been discovered that can catalyze a range of reactions including the sequence-specific cleavage of ribonucleic acids (RNA). Such DNA enzymes (DNAzymes) are being studied for the *in vivo* cleavage of mRNA. Compared to protein enzymes, DNAzymes have far fewer functional groups to employ for catalysis. In order to increase the chemical functionality of DNAzymes, eight modified nucleoside triphosphates were synthesized: three dUTP's modified at the 5-position with either a phenol, phenylboronic acid or guanidinium and five dATP's modified at the 8-position with imidazoles that are attached with linkers of various composition, length and flexibility. After the synthesis of the modified nucleotides, the incorporation of seven of them into oligonucleotides by DNA polymerases was studied. Under primer extension conditions, the guanidinium- and phenol-modified dUTP's were found to be good substrates for several DNA polymerases including *Pfu* (exo–) and Vent (exo–). Under PCR conditions, both of the modified nucleotides gave rise to properly sized products as well. The modified dATP's, however, were found to be very poor substrates. Only two of them were incorporated by Sequenase V2.0. After establishing conditions under which the modified nucleotides could be incorporated, the phosphoramidite of the guanidinium-modified nucleotide was used for the solid phase synthesis of oligonucleotides, and the phenol-modified nucleotide was used in an *in vitro* selection. Oligonucleotides containing the guanidinium groups were

found to exhibit enhanced duplex stability. The guanidinium-modified nucleoside phosphoramidite was also used to synthesize variants of the divalent metal cation-dependent DNAzyme, 10-23. Variants containing guanidinium groups in the substrate binding regions were found to display reduced rates of RNA cleavage. Using the phenol-modified dUTP, an *in vitro* selection process gave rise to DNAzyme STA17, isolated from the eleventh generation, that could catalyze self-cleavage in the presence of divalent metal cations such as Mg^{+2}, Zn^{+2} and Mn^{+2}. This DNAzyme was found to be inhibited by Hg^{+2}.

Preface

The experiments in this dissertation were designed primarily by Dr. David M. Perrin and myself, and experiments were performed by myself unless otherwise indicated.

Chapter 2. This chapter was written by myself. The synthesis and enzymatic incorporation of the guanidine-modified 2′-deoxyuridine triphosphate was included in a manuscript: Hollenstein, M.; Hipolito, C. J.; Lam, C. H.; Perrin, D. M. *Nucleic Acids Res.* **2009**, *37*, 1638-1649. The manuscript was written by Dr. Marcel Hollenstein. Only the experiments from the manuscript that were performed by myself have been included in this chapter. Diol-functionalized aminomethyl Novagel resin was synthesized by Dr. Curtis Harwig.

Chapter 3. The synthesis and enzymatic incorporation of the imidazole-modified 2′-deoxyadenosine triphosphates have been published: Lam, C.; Hipolito, C.; Perrin, D. M. *Eur. J. Org. Chem.* **2008**, 4915-4923. This manuscript was written by myself. The initial development of the syntheses was aided by Christopher Hipolito, and all the experiments were performed by myself. 8-[2-(4-Imidazoylyl)-ethylamino]-2′-deoxyadenosine triphosphate was synthesized by Dr. David Dietrich.

Chapter 4. This work has been published. Lam, C.H.; Perrin D.M. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 5119–5122. This manuscript was written by myself, and all the experiments were performed by myself.

Chapter 5. The initial development of the *in vitro* selection was aided by Dr. Marcel Hollenstein. The cloning of individual selection sequences and the preparation of sequencing samples were performed by Christopher Hipolito.

Table of contents

List of tables

List of figures

List of schemes

List of abbreviations

A	adenosine
Ac	acetyl
Ac_2O	acetic anhydride
AIBN	2,2′-azo-bis-isobutyrylnitrile
Arg	L-arginine
ARS	Alizarin Red S
ATP	adenosine triphosphate
Boc	*tert*-butyloxycarbonyl
br	broad
Bst	Family A DNA polymerase from *Bacillus stearothermophilus*
Bz	benzoyl
C	cytidine
calcd.	calculated
CPP	cell-penetrating peptide
d	doublet
dA	2′-deoxyadenosine
dATP	2′-deoxyadenosine triphosphate
dC	2′-deoxycytidine
DCC	*N*,*N*′-dicyclohexylcarbodiimide
dCTP	2′-deoxycytidine triphosphate
dd	doublet of doublets
ddATP	dideoxyadenosine triphosphate
ddTTP	dideoxythymidine triphosphate
dG	2′-deoxyguanosine
dGTP	2′-deoxyguanosine triphosphate
DMF	dimethylformamide
DMSO	dimethylsulfoxide
DMTr	4,4′-dimethoxytrityl
DNA	deoxyribonucleic acid
DNAzyme	DNA enzyme
dNTP’s	2′-deoxynucleoside triphosphates
Dpo4	DNA polymerase 4

dT	2′-deoxythymidine
DTT	(2*S*),(3*S*)-dithiothreitol
dTTP	thymidine triphosphate
dU	2′-deoxyuridine
dUTP	2′-deoxyuridine triphosphate
EDC	1-ethyl-3-(3-dimethylaminopropyl) carbodiimide
EDTA	ethylenediaminetetraacetic acid
equiv.	equivalents
ESI	electrospray ionization
Et	ethyl
EtOAc	ethyl acetate
FAD	flavin adenine dinucleotide
$FADH_2$	flavin adenine dinucleotide (reduced form)
G	guanosine
HCV	hepatitis C virus
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His	L-histidine
HPLC	high performance liquid chromatography
HRMS	high resolution mass spectrometry
k_{cat}	catalytic rate constant (multiple turnover number)
K_m	Michaelis constant of a substrate
k_{obs}	observed rate constant
LNA	locked nucleic acid
Lys	L-lysine
M	molecular weight, used in the description of molecular ion adducts from mass spectrometry, i.e. $[M+H]^+$
m	multiplet
MALDI	matrix assisted laser desorption ionization
Me	methyl
MeCN	acetonitrile
mRNA	messenger ribonucleic acid
MS	mass spectrometry
N40	random 40 nucleobase DNA
NBS	*N*-bromosuccinimide
NHS	*N*-hydroxysuccinimide
NMR	nuclear magnetic resonance

nt	nucleotides
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
Pfu	Family B DNA polymerase from *Pyrococcus furiosus*
Ph	phenyl
pK_a	negative log of acid dissociation constant
PNA	peptide nucleic acid
PNK	polynucleotide kinase
ppm	parts per million
py	pyridine
q	quartet
rC	cytidine (as opposed to the 2′-deoxycytidine), used in the drawing of DNAzyme 2D structures
R_f	thin layer chromatography retention factor
RNA	ribonucleic acid
RNase	ribonuclease
s	singlet
SELEX	systematic evolution of ligands by exponential enrichment
siRNA	small interfering ribonucleic acid
T	thymidine
t	triplet
Taq	Family A DNA polymerase from *Thermus aquaticus*
TEN	tris, EDTA, NaCl buffer
Tf	trifluoromethylsulfonyl
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TLC	thin layer chromatography
T_m	melting temperature
TMS	trimethylsilyl
TOF	time of flight
T_r	retention time
Tris	tris(hydroxymethyl)aminomethane
Trt	trityl (triphenylmethyl)
U	uridine
UV	ultraviolet
Vent	Family B DNA polymerase derived from *Thermococcus*

	litoralis
Δ	heat
δ	chemical shift
λ_{max}	wavelength of maximum absorbance

Acknowledgements

I would like to thank my supervisor Dr. David M. Perrin for his support and guidance. He was critical to my growth and development as a researcher. I would also like to thank all the past and present members of the Perrin Lab. It was a pleasure to share a workspace with every one of them. In particular, Dr. Curtis Harwig, Dr. Marcel Hollenstein and Dr. Jonathan P. May were all excellent mentors who were never hesitant to give me advice or assistance. Fellow graduate students Dr. Ali Asadi, Mr. Antoine Blanc, Dr. David Dietrich, Mr. Christopher Hipolito, Ms. Ying Li, Mr. Jonathan Pellicelli and Dr. Richard Ting were all present during the ups and downs of my graduate research. Thanks must also be extended to my only student Ms. Sandy Wong.

The work in this dissertation would not have been possible without the help of many members of the UBC community. I would like to thank all the staff of the offices, shops and services of the Chemistry Department. Whether I needed something fixed, analyzed or ordered, I could always find help. I would also like to thank the laboratory of Dr. George Mackie for allowing me to use their phosphorimager when our machine was in need of repair.

This dissertation was proofread by Dr. David M. Perrin, Dr. Stephen G. Withers, Dr. Bibiana K. Y. Wong and Mr. Christopher Hipolito. Thank you for all your suggestions, corrections and feedback.

Finally, I would like to thank my family, especially my parents, and friends for their faith in me and continued support. They gave me motivation to explore my interests and finish a very long journey.

Chapter 1: General introduction

1.1 Nucleic acids

Nucleic acids are one of the four major families of biomolecules that are found in all forms of life. They are linear biopolymers that were once thought to only be responsible for the storage of genetic information in the form of deoxyribonucleic acids (DNA) and the transfer of genetic information in the form of ribonucleic acids (RNA). However, since the early 1980's, this view has changed significantly. Naturally occurring ribozymes (catalytic RNA) were found to facilitate a number of processes including RNA cleavage,[1] RNA self-cleavage,[2,3] RNA splicing[4] and peptide bond formation.[5,6] These discoveries along with the fact that RNA can serve as a genome have led to the "RNA world" hypothesis,[7,8] which postulates that life forms were once based on a primitive form of RNA that could both store genetic information like DNA and act as catalysts like protein enzymes. In such a world, RNA would have been required to catalyze many more reactions than it presently does. Both the "RNA world" hypothesis and the discovery of natural ribozymes intensified research towards artificial ribozymes and led to development of *in vitro* selection, a methodology commonly referred to as SELEX (systematic evolution of ligands by exponential enrichment).[9-11]

The use of *in vitro* selection was first reported by several groups in 1990.[9-11] It is an iterative combinatorial process that allows for the scanning of very large, random libraries of nucleic acids in relatively short amounts of time to give rise to individual molecules that can perform a desired task highly effectively. SELEX techniques have been used for the discovery of new ribozymes, DNAzymes (catalytic DNA) as well as

aptamers,[12-16] which are oligonucleotide sequences that are not catalytic but rather ligands that can bind molecular targets. Hence aptamers are functionally very similar to antibodies.

Selections have also led to the discovery of ribozymes that can catalyze a very wide range of reactions, many of which have not been observed in nature. Some of these reactions include Diels-Alder condensations,[17-19] Michael additions,[20] alcohol oxidation,[21] urea synthesis,[22] decarboxylative Claisen condensation[23] and aminoacylation.[24,25] *In vitro* selection has also been used on DNA. Although DNA in living systems is almost exclusively found in double-stranded form for information storage, single-stranded DNA isolated using *in vitro* selection has been shown to be able to perform many of the same functions that have been associated with RNA. DNAzymes can cleave RNA,[26-28] catalyze a Diels-Alder condensation,[29] cleave DNA,[30,31] ligate RNA,[32] metallate porphyrin[33] and form nucleopeptide linkages.[34]

In the search for catalysts, DNAzymes have several advantages over their ribozyme counterparts. First of all, the large scale synthesis of DNA is much less expensive than RNA. DNA solid phase synthesis is higher yielding and DNA phosphoramidite reagents are less expensive. Secondly, DNA is significantly more stable than RNA; under physiological conditions, DNA has a non-enzymatic phosphodiester cleavage half-life that is approximately 100 000 times greater than that of RNA.[35,36] Thirdly, the *in vitro* selection of a DNAzyme requires fewer steps than the selection of a ribozyme. This last issue will be discussed in more detail in section 1.6.

1.2 Nucleotide and nucleic acid structure

Nucleotides are the subunits of nucleic acids. Each nucleotide consists of a

nitrogenous base, a pentose sugar and 1–3 phosphate groups (Figure 1.1A). Linear polymers of these nucleotides (1 phosphate per nucleotide) attached together through a sugar-phosphate backbone constitute DNA and RNA. In DNA, the sugar constituent 2′-deoxyribose is attached to one of four different nitrogenous bases: two purines, adenine and guanine, and two pyrimidines, thymine and cytosine as shown in Figure 1.1B. In the case of RNA, ribose and uracil (also a pyrimidine) are substituted for 2′-deoxyribose and thymine, respectively. Nucleic acids can also be found as double-stranded material in the form of DNA/DNA or DNA/RNA duplexes. Nitrogenous bases from one strand can interact with the nitrogenous bases from another strand to give rise to a relatively stable double helix. Adenine can form two hydrogen bonds with thymine (or uracil), and guanine can form three hydrogen bonds with cytosine (Figure 1.1B). Opposite strands of nucleic acids are known as complementary sequences that hybridize in an antiparallel manner.

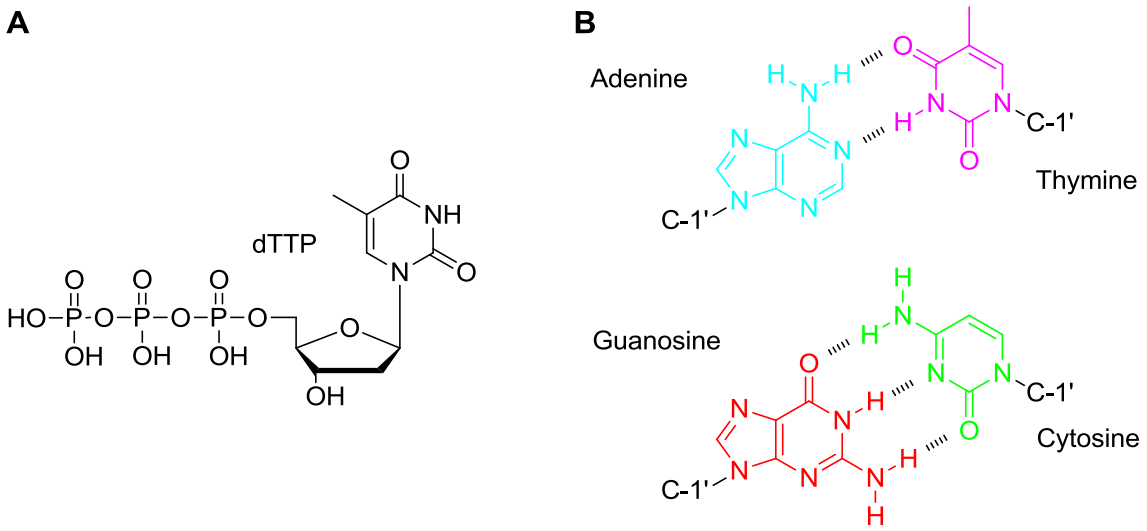


Figure 1.1 DNA structure. A, nucleotide dTTP; B, DNA base pairing.

1.3 Synthesis of DNA

1.3.1 Primer elongation

The synthesis of DNA can be achieved in one of two ways: primer elongation and solid phase synthesis. Primer elongations involve the use of DNA polymerases and 2′-deoxynucleoside triphosphates (dNTP's) (Figure 1.2). Two DNA oligonucleotides, a longer sequence that serves as the template, and a smaller primer sequence complementary to the 3′ end of the template, are annealed. DNA polymerases will then bind to this annealed complex and sequentially add nucleotides to the primer as directed by the template in the 5′ to 3′ direction to synthesize oligonucleotides that can be more than 10 000 bases in length. Each addition of a nucleotide gives rise to a molecule of pyrophosphate (PP_i) as a byproduct.

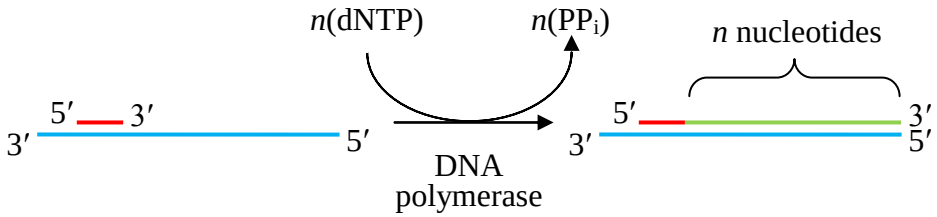


Figure 1.2 DNA synthesis by primer extension. Primer (——); template strand (——); incorporated portion of oligonucleotide product (——).

1.3.2 Solid phase synthesis

Solid phase synthesis involves the synthesis of DNA on a solid support in almost exclusively organic solvents (Figure 1.3).[37-39] The automated process is very efficient and is capable of producing oligonucleotides up to two hundred bases in length in amounts that are much higher than can be achieved with syntheses that commonly employ DNA

polymerases. The synthesis involves the sequential addition of nucleoside phosphoramidite molecules in the 3′ to 5′ direction. Briefly, a nucleoside attached onto a solid support at the 3′-OH is deprotected to expose the 5′-hydroxyl group. The phosphoramidite of the required DNA base is added to give rise to a product that is one unit longer. Following coupling, a capping procedure is performed. Unreacted hydroxyl groups are blocked by acetylation such that these "failed sequences" are no longer extended in subsequent rounds of synthesis. After the capping procedure, the newly formed phosphite triester is oxidized to a phosphate. The elongated product can then be specifically deprotected again and coupled with the next nucleoside phosphoramidite. When the synthesis is complete, base is used to cleave the oligonucleotide from the resin, remove the cyanoethyl groups from the phosphates as well as remove protection groups from the nitrogenous bases. Finally, the oligonucleotide can be purified by high performance liquid chromatography (HPLC) or polyacrylamide gel electrophoresis (PAGE). The solid phase synthesis of RNA uses a procedure that is very similar to what is shown in Figure 1.3 but is lower yielding due to lower coupling efficiencies.[40]

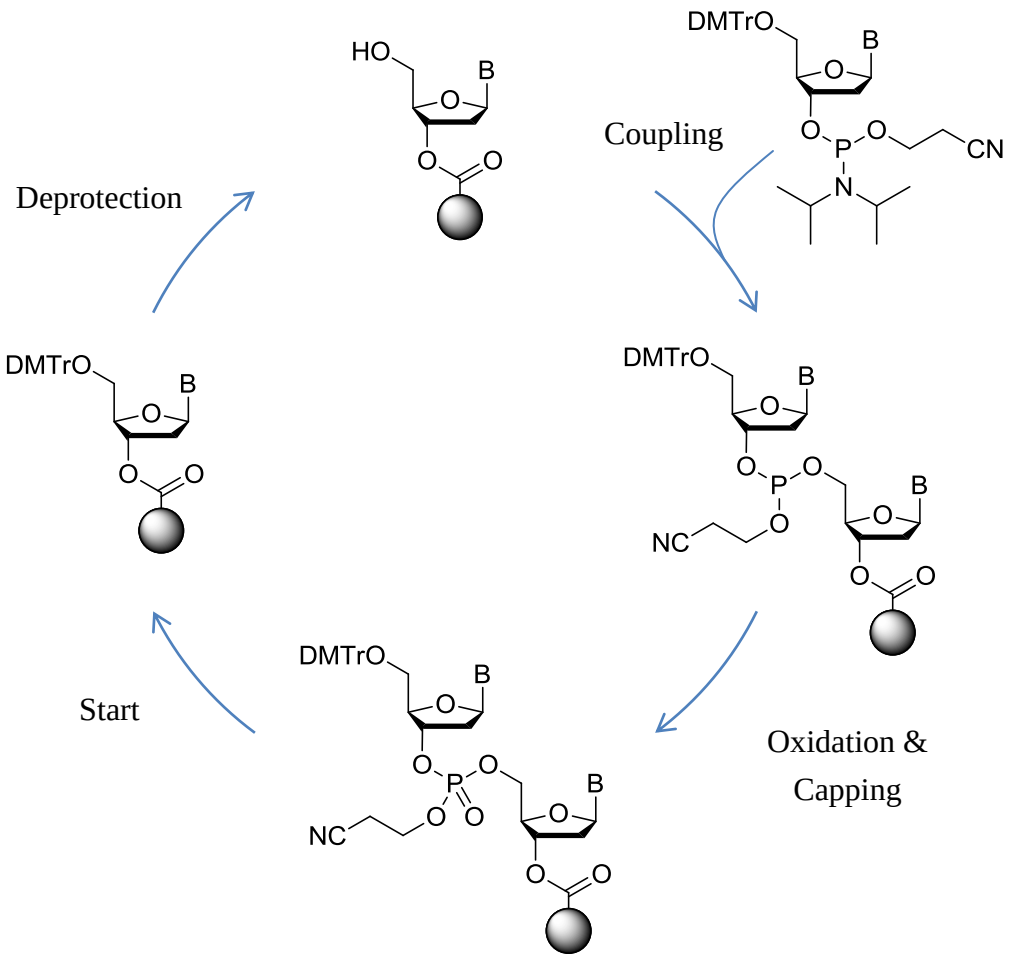


Figure 1.3 DNA solid phase synthesis. DMTr – 4,4′-dimethoxytrityl group.

1.4 Overview of SELEX

As mentioned earlier, *in vitro* selection is a technique that can isolate individually active molecules from very large random libraries of nucleic acids.[13,14,41] In Figure 1.4, a simple schematic of a typical selection is outlined. Starting from a nucleic acid pool of up to 10^{16} individually different nucleic acid sequences, a screening process is carried out in which the pool of oligonucleotides is challenged to perform a certain function. The experiment is designed such that the sequences of desired activity are separated from the inactive sequences. The isolated sequences are then amplified by the polymerase chain reaction (PCR)[42] to the size of the original nucleic acid pool. This subsequent pool of

sequences is one of enhanced desired activity and fewer individually different nucleic acid strands. The selection process is then repeated on the subsequent pool of sequences. In the initial rounds of selection, very few sequences from the pool will be active, often so few that activity cannot be detected. However, after several rounds of selection (usually 5–15 rounds), activity will be observed. The nucleic acid pool will have accumulated a significant population of sequences of desired activity and a relatively small number of individually different sequences. At this point, the pool of sequences can be cloned into bacteria such that each colony of bacteria will have plasmids that contain only one nucleic acid sequence. Finally, PCR amplification and sequencing of nucleic acids from individual colonies reveal the identities of the functional molecules.

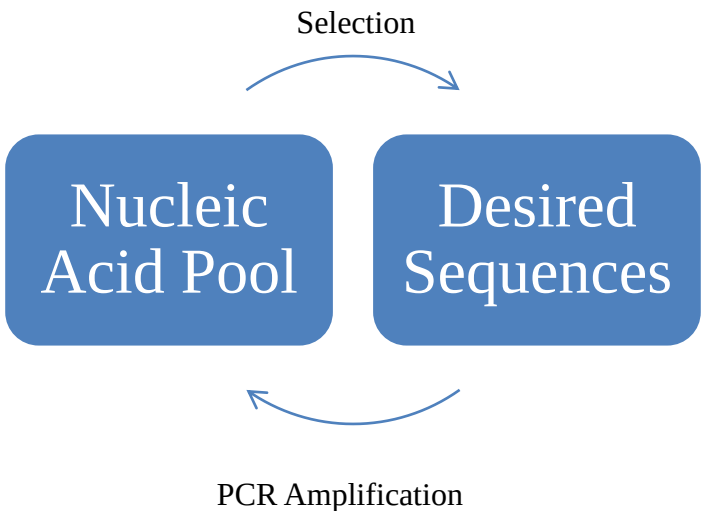


Figure 1.4 Selection schematic.

1.5 RNA-cleaving DNAzymes

Of the many reactions that can be catalyzed by DNAzymes as listed in section 1.1, DNAzymes that enhance the cleavage of ribonucleotide phosphodiester bonds have

probably been the most well studied.[26,28,43-45] Two examples of these DNAzymes are 10-23 and 8-17,[27,46] as shown in Figure 1.5. There is no general method for the naming of DNAzymes, but the names are often derived from the number of rounds of selections that are performed prior to the identification of the catalyst. Both DNAzymes 10-23 and 8-17 were derived from the same selection: 10-23 was the 23rd clone from round 10 and 8-17 was the 17th clone from round 8. These DNAzymes typically have a relatively small catalytic motif that is connected to two substrate binding regions, one attached to the 5′-hydroxyl and the other one attached to the 3′-hydroxyl of the catalytic motif. The binding regions hybridize with the desired RNA target. Upon binding, the catalytic motif facilitates the cleavage of the target to give rise to two shorter strands of RNA that subsequently dissociate from the substrate binding regions. Since only complementary RNA substrates will be bound, DNAzymes cleave RNA sequence-specifically. As well, the activity of a catalytic motif is normally not affected significantly if the sequence of the binding regions is altered.[27,41] Hence by manipulating the composition of the substrate binding domains, DNAzyme variants using the same catalytic motif can be synthesized to cleave almost any desired RNA sequence making DNAzymes very versatile endonucleases.[47-50]

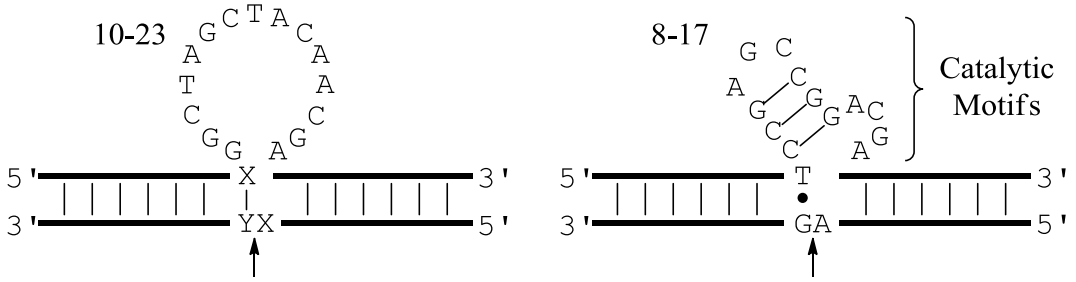


Figure 1.5 Secondary structures of the divalent metal-dependent RNA-cleaving DNAzymes 10-23 and 8-17. The DNAzyme (top strand) binds the RNA substrate (bottom strand) through Watson-Crick base pairing. The position of RNA cleavage is indicated by the arrow. Y = U or C; X = A or G.

1.5.1 Mechanism of RNA cleavage

The cleavage of RNA by small DNAzymes is considered to be very similar to the specific base catalyzed cleavage of RNA.[51,52] As shown in Figure 1.6, a deprotonated 2′-oxygen at the cleavage location nucleophilically attacks the phosphodiester in an in-line fashion to give what is believed to be a dianionic pentacoordinated phosphorane intermediate.[53-56] This very short-lived intermediate does not undergo pseudorotation[54,57] and may be kinetically indistinguishable from a transition state,[56] as was shown by a recent kinetic isotope effect study.[58] The 5′-oxygen from the nucleotide downstream from the cleavage location is then released to give rise to a 2′-3′ cyclic phosphate product. The 5′-oxygen leaving group is then protonated.

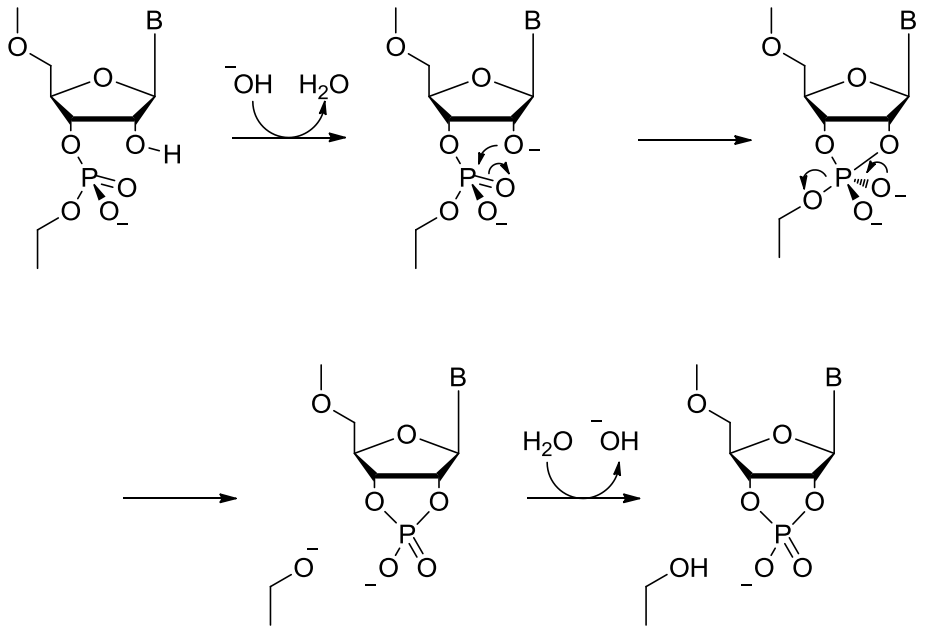


Figure 1.6 Mechanism of specific base-catalyzed RNA cleavage.

1.5.2 The role of divalent metal cations in DNAzyme catalysis

Divalent metal cations play very important roles in the function of DNAzymes. In order to fold up into compact three dimensional structures, single-stranded nucleic acids such as DNA require the presence of divalent metal cations to reduce repulsion between negatively charged phosphodiesters.[59,60] As well, divalent metal cations can accelerate the cleavage of RNA.[61-63]

Due to the limited number of chemical functionalities that can be found in nucleic acids, DNAzymes are usually selected in the presence of a divalent metal cation cofactor. Both 10-23 and 8-17 were selected in the presence of Mg^{+2} and require it for catalytic activity.[27] Compared to DNAzymes such as 10-23 and 8-17, unmodified DNAzymes that have excluded divalent metal cation cofactors such as Mg^{+2} have generally exhibited significantly inferior rate constants.[64-66] The divalent metal cations can serve several roles.[52] They can participate in general acid/base catalysis. A divalent metal-coordinated

hydroxide can abstract a proton from the 2′-hydroxyl group, and a divalent metal-coordinated water molecule can donate a proton to the 5′-oxygen. As well, the divalent metal cations can act as Lewis acids to stabilize negative charges on the 2′-oxygen, 5′-oxygen and the non-bridging phosphate oxygen atoms. [52,67]

1.5.3 Ribonuclease A cleavage of RNA

Ribonuclease A (RNase A)[68] is a non-sequence-specific protein endonuclease that can facilitate RNA cleavage efficiently in the absence of divalent metals. Unlike most DNAzymes (and ribozymes), RNase A does not require the aid of divalent metal cation cofactors for catalysis. Instead, it relies on two histidine residues to carry out acid/base catalysis (Figure 1.7). His-12 acts as the base to deprotonate the 2′-hydroxyl group to facilitate the in-line nucleophilic displacement of the nucleoside. His-119 acts as the acid to donate a proton to the 5′-oxygen to enhance displacement. There is also a nearby lysine residue that aids in the stabilization of negative charge on the transition state.[68] Two research labs have tried to obtain DNAzymes that act as RNase A mimics.[69,70] Their research will be discussed in more detail in section 1.8.3.

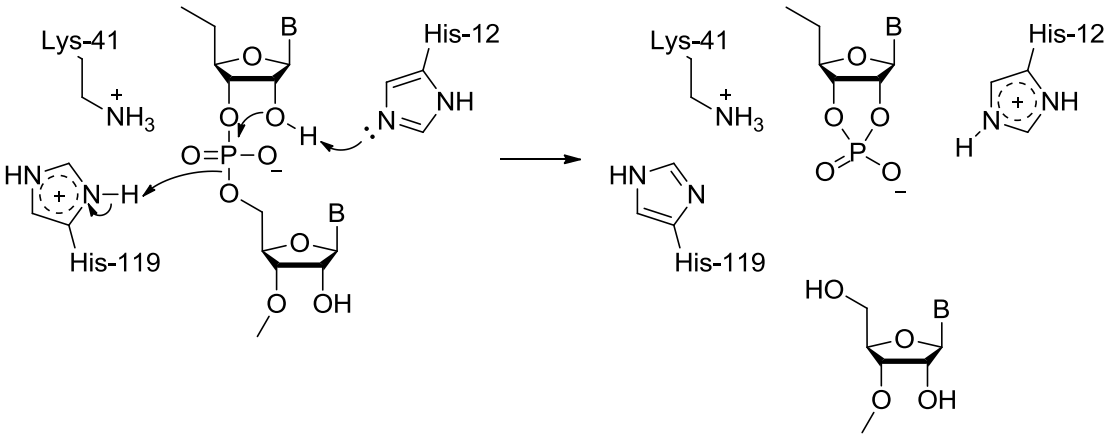


Figure 1.7 Mechanism of RNA cleavage by RNase A.

1.6 *In vitro* selection of a pair of RNA-cleaving DNAzymes

The *in vitro* selection process that was used to isolate the RNA-cleaving DNAzymes 8-17 and 10-23 is outlined in Figure 1.8.[27,41,71] A solid phase synthesized library of oligonucleotide sequences containing a random fifty nucleotide sequence and two primer binding regions (for PCR and primer extensions) on each end is used as a template for the primer extension of a biotinylated RNA-containing primer. Biotin binds extraordinarily strongly to streptavidin and allows the primer extension products to be immobilized on streptavidin affinity columns.[72] The bound double-stranded nucleic acids are treated with low concentrations of sodium hydroxide to remove the non-biotinylated template strands. Subsequently, the remaining single-stranded modified sequences were treated with a selection buffer (10 mM $MgCl_2$, 1 M NaCl, 50 mM Tris, pH 7.5) at 37 ℃. Under these conditions, the oligonucleotides adopt different three-dimensional structures. Those sequences that are capable of facilitating intramolecular RNA cleavage will liberate themselves from the streptavidin affinity columns.

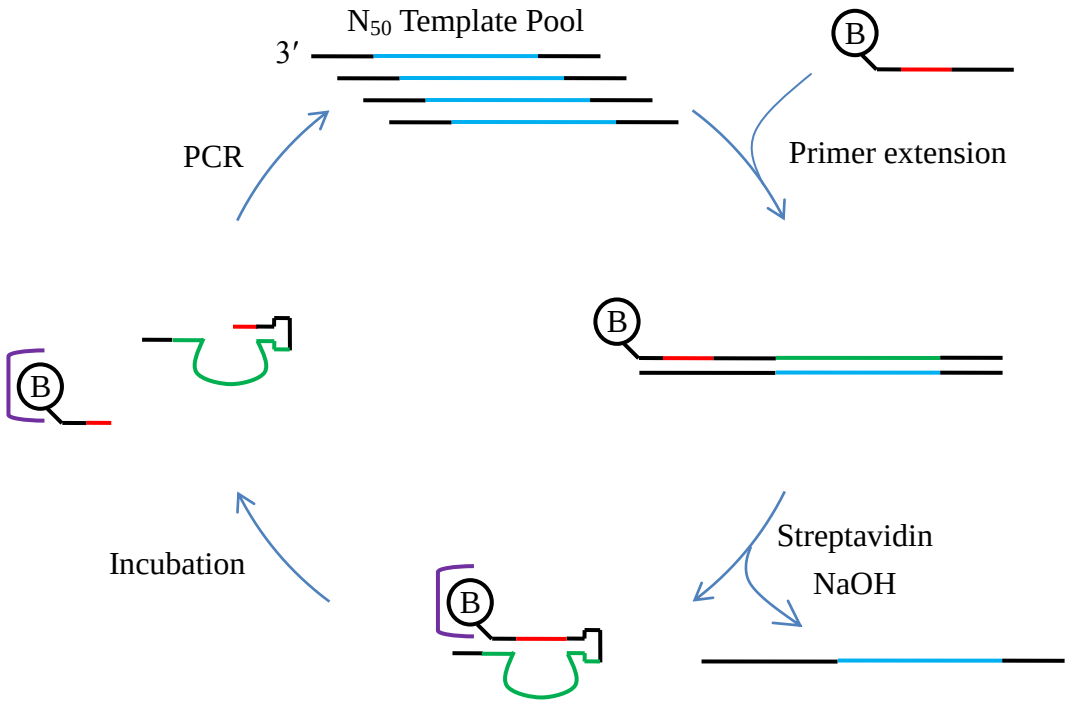


Figure 1.8 RNA-cleaving DNAzyme selection scheme. (——) DNA primer regions; (——) N_{50} template region; (——) RNA; (——) N_{50} random region; (B) biotin; (——) streptavidin magnetic particles.

Usually, selection conditions for early rounds are not very stringent, that is, the solution environment is very favourable for the desired activity. Divalent metal cation concentrations, which facilitate DNAzyme catalyzed RNA cleavage, are usually high and incubation time is typically long to ensure all potentially active sequences can catalyze self-cleavage and be carried over to the next round of selection. It should be noted that every sequence in a library (active or non-active) has a non-zero chance of catalyzing self-cleavage. Hence, in each and every round of a selection, there will always be sequences collected that are not active DNAzymes. Statistically speaking, if the existing pool of oligonucleotides had 1 active sequence for every 10^9 inactive sequences and the

active sequences accelerated the reaction by a factor of 10^6, then the material that is collected from a library of 10^9 sequences during that round of selection would be composed of 1 active sequence and 1 000 inactive sequences. Based on these numbers, the cleavage due to inactive species is very often more than the cleavage due to active species in the earlier generations. As the selection process advances into subsequent rounds, activity due to inactive sequences becomes much less significant. There will be many more active sequences and many copies of each of them as well. At this point, the stringency of the selection buffer can be gradually increased in order to isolate DNAzymes of the highest activity. Schlosser *et al.* performed and monitored two selections simultaneously: one with increasing stringency and one with constant stringency.[73] They found that the selection that employed increasing stringency gave rise to a more catalytically active sequence pool and subsequently superior DNAzymes.

The selection step, which can also be called the positive selection, is often preceded by a negative selection. A negative selection is a subtraction step that removes non-specific sequences, active sequences that also display other unwanted properties. The selection that gave rise to 8-17 and 10-23 did not employ a negative selection. Both DNAzymes were selected in the presence of Mg^{+2}, but have been shown to also be active in other divalent metal ions such as Ca^{+2} and Mn^{+2}.[46,74] In fact, these DNAzymes displayed superior catalysis in the presence of Mn^{+2} when compared with Mg^{+2}. Bruesehoff *et al.* performed two parallel selections for a divalent cobalt-dependent DNAzyme.[75] One selection was subjected to a solution containing competing cations as a negative selection such that sequences that were active in the presence of those cations were removed. The other selection was performed with no pretreatments. As a result, the selection that employed a negative selection step gave rise to nucleic acids that were most

catalytically active in Co^{+2} whereas the other selection produced DNAzymes that were more active in Zn^{+2} and Pb^{+2} than in Co^{+2}. The studies of Bruesehoff *et al.* demonstrate that DNAzymes can be effective sensors for divalent metal cations when negative selections are incorporated into the selection protocol.

Once the nucleic acid sequences have been collected from the eluted selection buffer, they are purified and then amplified using PCR. The PCR gives rise to two different but equally-sized products: encoding strands and complementary non-coding strands. In order to be able to separate the two products, a special primer must be used for the PCR. In the selection for 10-23 and 8-17, the non-encoding strand primer contained a 3′-terminal ribonucleotide.[27] Upon alkaline hydrolysis, the double stranded PCR product gave rise to a longer encoding strand and a shorter non-encoding strand. The encoding strand can then be purified and used as template for the next round of selection. Another method of removing the non-encoding strand is to have a 5′-phosphorylated primer for that strand. The phosphorylated, non-encoding strand can then be digested with exonucleases.[76]

As was alluded to earlier, a ribozyme selection requires additional procedures to be carried out. The reason for this is because PCR generally requires a DNA template and produces DNA products. Therefore, isolated RNA sequences must be reverse transcribed into complementary DNA prior to PCR, and the PCR products must be transcribed back into RNA for the next round of selection.

After several rounds of selection and subsequent cloning, individually active DNAzymes are isolated. At this point, additional rounds of selection can be performed on a library that is generated from the mutagenization of the most active clone sequence(s). The reselection is performed because the activity of the DNAzyme that is isolated is not

necessarily the highest possible. For a random region of 50 nucleotides, there are 10^{30} possible sequences. A selection can only sample a very small portion of these sequences. Instead of selecting from every possible sequence, only families of sequences are analyzed. By reselecting, DNAzymes of enhanced activity that were not present in the original pool can be obtained.[27]

DNAzymes are typically selected as self-cleaving species where the identity of each sequence is directly linked to the activity of each sequence. This linkage is necessary for the selective amplification and identification of catalysts. However, catalysts that perform intramolecular cleavage, otherwise known as *cis*-cleaving catalysts, are not true catalysts and have far fewer applications when compared to DNAzymes that are capable of multiple turnover. The term *trans*-cleavage is used to describe DNAzymes that can perform intermolecular cleavage. In most circumstances, *cis*-cleaving DNAzymes can be reconfigured into *trans*-cleaving DNAzymes. For 10-23 and 8-17, the *trans*-cleaving DNAzymes were obtained by chemically synthesizing oligonucleotides using solid phase synthesis according to the sequences of the clones that excluded everything outside of the substrate binding regions. These oligonucleotides can then be incubated with RNA substrates for intermolecular catalysis. The *cis*-cleaving and *trans*-cleaving DNAzyme 8-17 are shown in Figure 1.9. The first image is the DNAzyme that was obtained from the selection. In the second image, solid phase synthesis is used to synthesize a DNAzyme that is reduced to only the catalytic motif and the substrate-binding regions. The third image shows the substrate and catalytic motif requirements of the DNAzyme. The nucleotides in the substrate binding regions can be changed without significant loss to catalytic activity provided that base pairing is maintained.

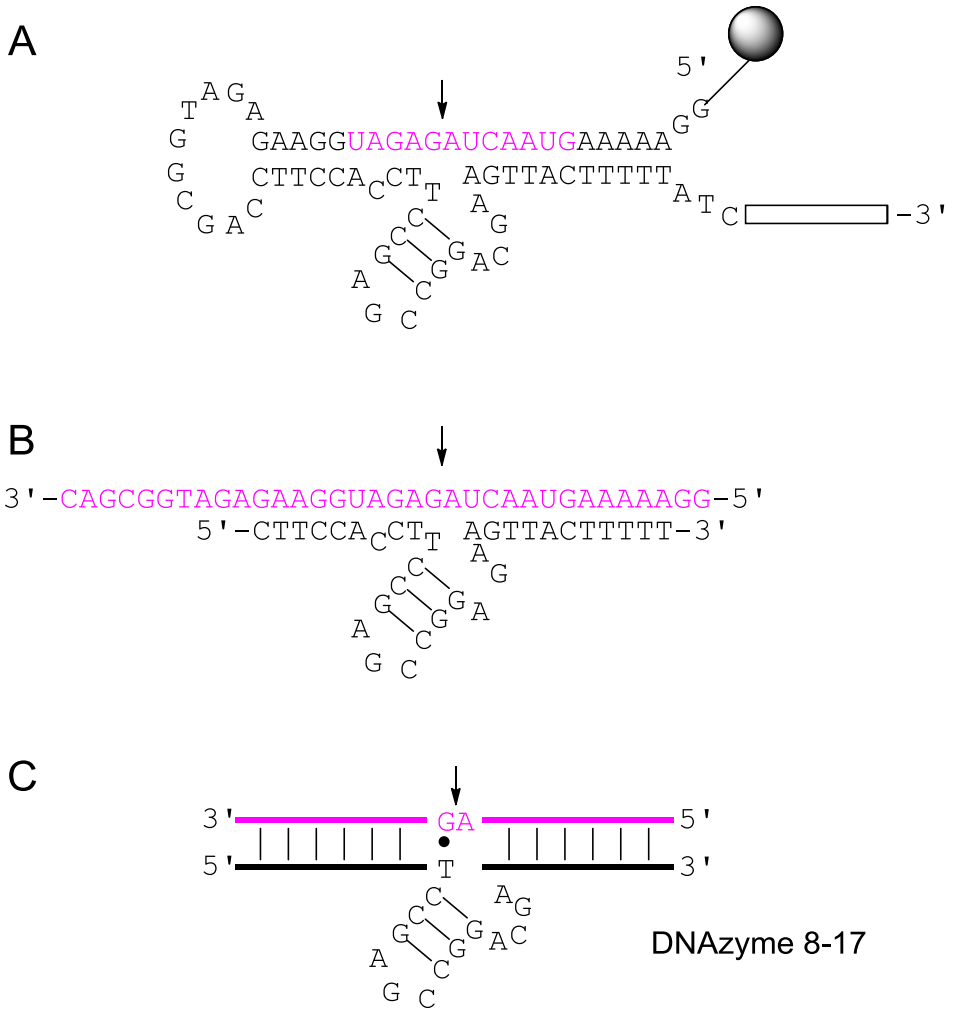


Figure 1.9 Structure of DNAzyme 8-17. A, *cis*-cleaving DNAzyme obtained from the selection; B, *trans*-cleaving DNAzyme obtained from removing nucleotides outside of the substrate binding regions, and RNA substrate based on the sequence of the primer that was used for the selection; C, *trans*-cleaving DNAzyme with essential nucleotides for the DNAzyme and RNA substrate shown. Ball, biotin; purple residues, RNA; arrow, cleavage location; rectangle, 31 nucleotides corresponding to the PCR primer site and the remainder of the N_{50} random region from the selection.

1.7 Divalent metal-independent DNAzymes

In physiological conditions, the concentrations of divalent metals are low,[77-79] much less than what is typically required for optimum ribophosphodiester bond cleavage by

divalent metal-dependent DNAzymes, and hence research groups have attempted to discover divalent metal-independent DNAzymes that would be more applicable to *in vivo* applications. Geyer and Sen performed a selection in the absence of divalent metal cations to evaluate divalent metal-independent DNAzyme catalysis.[64] Their efforts led to the discovery of the DNAzyme G3. G3 cleaved an RNA/DNA chimeric sequence at an observed rate constant of about 0.0015 min^{-1} in the presence of 250 mM NaCl and 1 mM EDTA at 25 ℃ and pH 7.0. Although this rate is ~10^2–10^4 fold less than what has been observed for divalent metal-dependent DNAzymes, this rate represents a rate enhancement of ~2 × 10^7 over the background reaction without the DNAzyme. Their results along with the results of two other groups,[65,66] who have also discovered similar divalent metal-independent DNAzymes, show that DNA can indeed perform ribophosphodiester bond cleavage without the aid of divalent metal cations. Another method of achieving divalent metal-independent DNAzyme catalysis is to use modified DNA.

1.8 Modified DNA

1.8.1 Nucleobase-modified nucleotides

To date, most nucleic acid catalysts have been far inferior to protein enzymes in terms of catalytic rate constants. RNase A performs non-sequence-specific RNA cleavage with kinetic perfection,[80] a term that is used to describe enzymes where the rate limiting step is the rate of diffusion rather than the rate of catalysis. RNase A cleaves RNA with an observed rate constant of 10^4–10^5 min^{-1} [68] and a catalytic efficiency (k_{cat}/K_m) that can exceed 10^9 $M^{-1}min^{-1}$. DNAzymes have been shown to exhibit similar k_{cat}/K_m values as well, but this is only because very favourable K_m values can be obtained with increased

substrate binding region lengths. Even under non-physiological conditions, DNAzymes have only been able to achieve a rate constant of ~10^1 min^{-1}.[27] One of the possible reasons for the discrepancy in rates of catalysis is that nucleic acid catalysts are chemically very limited compared to protein enzymes that employ amino acids containing twenty different side chains. In an effort to improve the catalytic rates and functional diversity of nucleic acids, many research groups have synthesized modified nucleoside triphosphates (dNTP's) that contain appended side chains for usage in *in vitro* selections.[81] In Figure 1.10, a small sampling of the modifications that have been attached is shown.[82-85] In addition to these four, other modifications that have been added to DNA include fluorescent groups,[86,87] photoreagents,[88,89] boronic acids[90,91], nearly every one of the amino acid side chains[84,92-96] and other groups as well.[97-99]

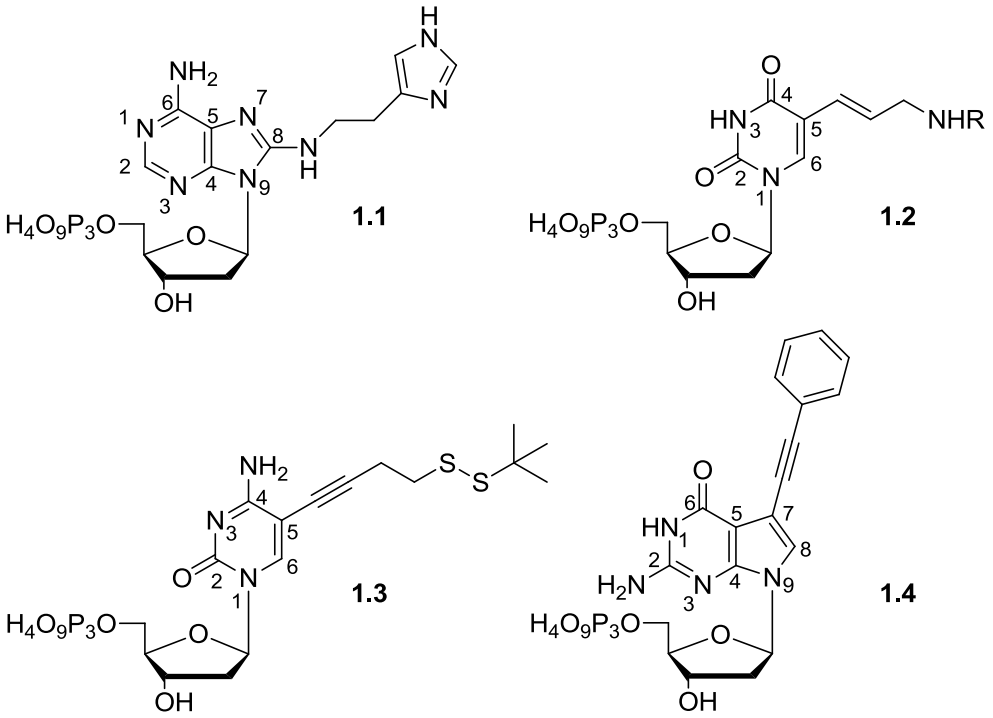


Figure 1.10 Modified dNTP's that are substrates for at least one DNA polymerase. "R" represents various modifications that are attached through an amide bond.

The generation of modified nucleotide-containing DNA libraries for selections begins with the synthesis of modified dNTP's. Typically, these monomers are used in place of their natural counterparts in primer extensions. A modified-dGTP residue would replace every dGTP, and a modified-dCTP would replace every dCTP. The resulting modified DNA pool is then incubated to isolate self-cleaving species, and the isolated material is amplified by PCR in the presence of natural dNTP's to generate a new template pool for the synthesis of modified sequences for the next round of selection. If the modified nucleotides are excellent substrates for the DNA polymerases that are used for PCR, the modified nucleotides can be directly used in the PCR to generate a modified DNA pool for the next round of selection. This PCR step would bypass the primer extension step.[84,92-94]

In order to be used for the synthesis of modified DNA in *in vitro* selections, modified nucleotides must satisfy two conditions.[85] First, the nucleoside triphosphate must be a substrate for DNA polymerases. If this requirement is not met, the modified nucleotide will be of very little use. Second, the sequences that contain modified residues must be able to act as suitable templates for PCR amplification. This second condition ensures that sequence information is retained during iterative rounds of selection. The modified DNA typically contains modifications that are attached to the non-base pairing face of the nucleobases. This preserves Watson-Crick base pairing[100] and hence facilitates both incorporation by DNA polymerases and correct transfer of sequence information. In order to preserve the base pairing of modified nucleotides, modifications have typically been attached to the 5-position of pyrimidines,[84,93,96,99,101] the 8-position of purines[69,84,95] and the 7-position of 7-deazapurines.[84,94,95]

Due to the difficulties of working with multiply charged organic molecules as well as the instability of nucleoside triphosphate species, the synthesis, isolation, characterization and use of nucleoside triphosphates can be rather difficult. In a typical synthesis, the modification is first attached to the nucleoside, and the addition of phosphates[102] is usually reserved for the last step. Purification of nucleoside triphosphates involves a combination of ion-exchange chromatography and/or HPLC. Because triphosphates are less stable in acidic environments, they are generally isolated as metal, ammonium or triethylammonium salts rather than the protonated acidic forms.

Once modified nucleoside triphosphates have been used in an *in vitro* selection and active *cis*-cleaving DNAzymes have been identified, modified nucleoside phosphoramidites must be synthesized and used for the solid phase synthesis of the *trans*-cleaving DNAzymes. Modified nucleoside phosphoramidites are fabricated such that the modifications are suitably protected and compatible with the general protocol of nucleic acid solid phase synthesis. (section 1.3.2)

1.8.2 Incorporation of modified nucleotides

Upon successful synthesis of modified nucleoside triphosphates, conditions must be established under which the modified residues can be incorporated into oligonucleotides. This process begins with finding appropriate DNA polymerases. There are several families of DNA polymerases based on sequence homology. The polymerases that have been used for the incorporation of modified residues in the literature [84,94,96,103] typically belong to one of three families: families A, B and Y. Polymerases of family A include Sequenase V2.0, *Thermus aquaticus* (*Taq*), *Bacillus stearothermophilus* (*Bst*) and Klenow. Polymerases of family B are usually enzymes of high fidelity. Three members that are

commonly used are polymerases isolated from *Pyrococcus furiosus* (*Pfu*), *Pyrococcus woesi* and *Thermococcus litoralis* (Vent). Y family polymerases are relatively small, low fidelity enzymes that are known for their ability to bypass lesions. One member of this family is DNA polymerase IV (Dpo4).[104] Isolated from *Sulfolobus solfataricus*, this polymerase is thermostable, but is also active at physiological temperatures. All of the above DNA polymerases were either originally isolated from different organisms (particularly prokaryotes) or were genetically engineered versions of the same DNA polymerases with special properties. One such engineered property that is very common is the elimination of a 3′–5′ proofreading exonuclease activity. Polymerases that exhibit this property have "(exo–)" attached to their names.

Aside from the DNA polymerase used, there are also several other variables that can affect the incorporation of nucleoside triphosphates. The pH, temperature, DMSO concentration, nucleotide concentration and enzyme concentration can all be varied. Changing the pH is an interesting approach as it will affect and in most cases lower the activity of the enzyme. At the same time, the pH change can alter the state of attached functional group modifications such that the incorporation of the modified nucleoside triphosphates may be enhanced. DMSO can be used to help destabilize secondary structures in DNA templates or primers. Incorporation experiments in the literature are sometimes performed in the presence of 5–10 % DMSO.[105,106]

1.8.3 Modified DNAzymes

The first modified nucleotide to be used in a DNAzyme selection was a 5-imidazole-functionalized-2′-deoxyuridine. This nucleotide consists of urocanic acid coupled with 5-aminoallyl-2′-deoxyuridine triphosphate through an amide bond. The

nucleotide was used in place of dTTP in a selection process to give rise to a catalytic sequence 16.2-11 that can catalyze the self-cleavage of RNA in the presence of micromolar concentrations of Zn^{+2} (Figure 1.11). They proposed that the imidazoles may be involved in binding a Zn^{+2} ion that participates in catalysis. The phosphoramidite of the modified dUTP was subsequently synthesized to obtain a catalyst capable of multiple turnover. The *trans*-cleaving DNAzyme had a maximum catalytic rate constant of 3.1 min^{-1} in the presence of 30 μM Zn^{+2} at pH 7.5 and 37 ℃, and 3 out of the 4 modified nucleotides found in the catalytic motif were found to be essential for activity.

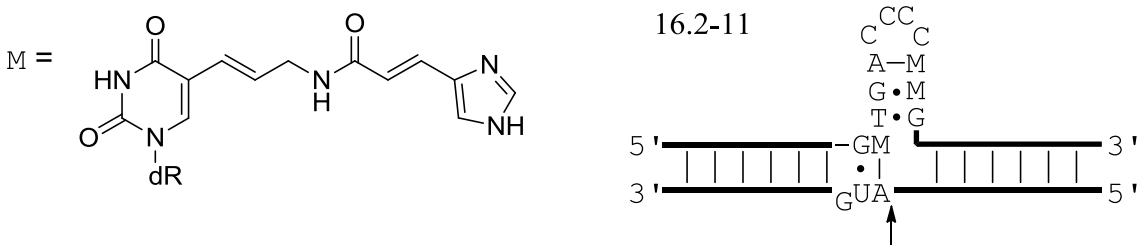


Figure 1.11 Modified RNA-cleaving DNAzyme 16.2-11 (*trans*-cleaving). DNAzyme (top strand) is shown with bound substrate (bottom strand).

A year later, the discovery of 9_{25}-11, a DNAzyme that used two different modified nucleotides, was published (Figure 1.12).[69] 8-[2-(4-Imidazolyl)ethylamino]-2′-deoxyadenosine was substituted for 2′-deoxyadenosine, and 5-aminoallyl-2′-deoxyuridine was substituted for thymidine. This selection set out to find a RNase A mimic and hence was performed in the absence of divalent metals. The imidazole would perform the function of histidine, and the allylamine would perform the function of lysine. The selection gave rise to a

self-cleaving sequence that was able to cleave a DNA/RNA chimeric substrate that contained one embedded cytidine with a rate constant of 0.037 min^{-1} at 24 ℃ and pH 7.5. This DNAzyme was subsequently converted to a catalyst capable of multiple turnover.[107,108] Mechanistic investigations on this *trans*-cleaving species have shown that the catalytic reaction occurs through general acid/base catalysis, and hence 9_{25}-11 is a true RNase A mimic.[109]

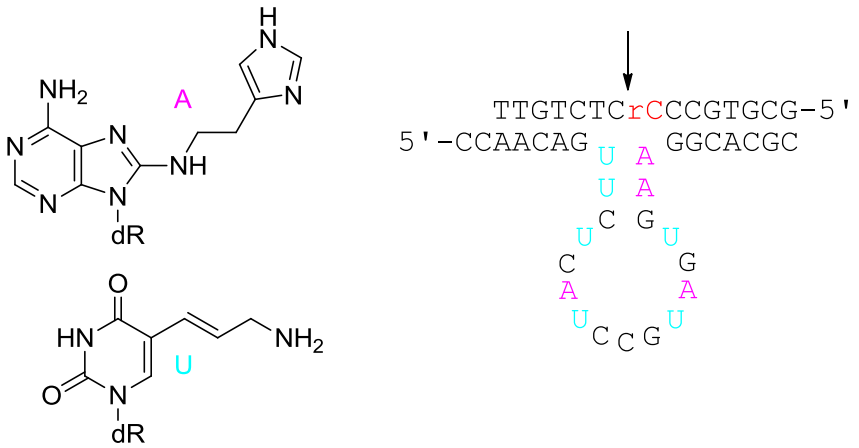


Figure 1.12 DNAzyme 9_{25}-11 (*trans*-cleaving) with bound substrate. The modified nucleotides 8-[2-(4-imidazolyl)ethylamino]-2′-deoxyadenosine (A) and 5-aminoallyl-2′-deoxyuridine (U) are shown in purple and blue, respectively. The location of cleavage following the cytidine (rC) residue is indicated by the arrow, and dR represents 2′-deoxyribose.

Sidorov *et al.* discovered a divalent metal-independent modified DNAzyme as well.[70] They also used nucleotides that contained imidazoles and amines to facilitate catalysis except the amine was attached to 7-deaza-2′-deoxyadenosine (**1.6**) and the imidazole was attached to 2′-deoxyuridine (**1.5**) (Figure 1.13). The catalytic sequence can cleave a twelve nucleotide RNA target sequence specifically at one of two possible locations with a rate constant that is somewhat higher than that of 9_{25}-11. The

modifications were shown to be necessary for optimal catalysis, but when natural dNTP's were used in place of either **1.5** or **1.6** for the synthesis of the DNAzyme, the observed rate of catalysis was reduced to 10 or 30 % of the optimal activity, respectively. It is not clear whether the modifications take part in catalysis or play a purely structural role, and this DNAzyme was not converted to a *trans*-cleaving catalyst.

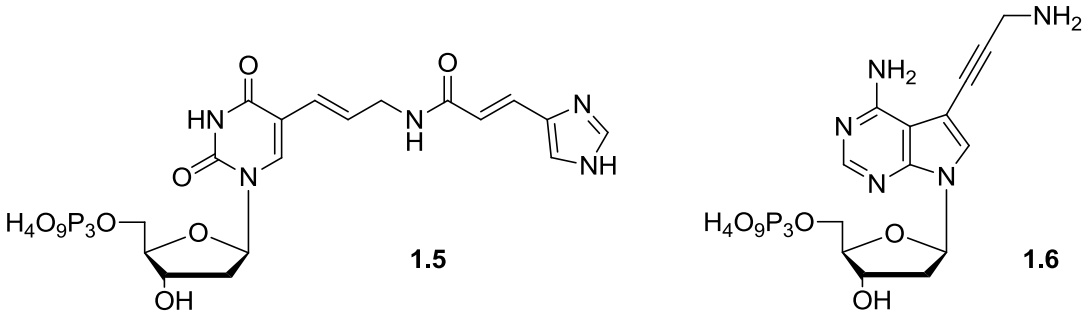


Figure 1.13 Two modified nucleoside triphosphates that have been used in the *in vitro* selection of a RNase A mimic DNAzyme.

The selections that gave rise to the DNAzyme discussed above and 9_{25}-11 both demonstrate that the addition of extra functionality enhances the catalytic ability of DNA. The observed cleavage rate constants in the absence of divalent metal cations were improved at least five-fold when compared to unmodified catalysts. The rate constants of 9_{25}-11 and the DNAzyme of Sidorov *et al.* compare very favourably against the rates of divalent metal-dependent DNAzymes that are incubated under physiological concentrations of divalent magnesium.[77-79,110]

1.9 DNAzyme cleavage of mRNA *in vivo*

The development of RNA-cleaving DNAzymes that are functional under

physiological conditions has paved the way for the potential use of DNAzymes *in vivo*.[111,112] Many studies have employed DNAzymes for gene expression control in cellular environments, and almost all of these endeavours have utilized 10-23 due to its simple cleavage site requirements. *In vivo* targets have included RNA associated with hepatitis C virus (HCV),[113] human immunodeficiency virus,[114,115] penicillin-binding protein[116] and vascular endothelial growth factor receptor 2.[117] All of these studies have shown that DNAzyme therapeutics have a lot of potential.

Although DNAzymes have been used against various targets *in vivo*, the practical use of RNA-cleaving DNAzymes for therapeutic applications has been limited by several challenges. The activity of divalent metal-dependent DNAzymes such as 10-23 is far from optimal under physiological conditions.[77-79] Also, DNAzymes are not very stable intracellularly and display poor membrane permeability.[71,118] As discussed earlier, the first issue is being addressed by the search for DNAzymes that can either perform catalysis rapidly under physiological conditions or function in the absence of divalent metal cations altogether. Recently, Young *et al.* studied the use of DNAzyme 10-23 for the regulation of gene expression in mammalian cells.[119] They found that the gene silencing that was observed was predominantly not due to the RNA-cleaving activity of the DNAzyme. Instead, the gene silencing was due to antisense mechanisms. Their results further highlight the need for DNAzymes of greater catalytic rates under physiological conditions. The other two issues, membrane permeability and cellular stability, apply to all forms of nucleic acid-based therapeutics including ribozymes and antisense oligonucleotides. Antisense therapy involves the use of oligonucleotides to base pair specifically with RNA targets to either permanently deactivate them or trigger cleavage of the resulting duplex by cellular RNase H.[120,121] Although DNAzymes are

much more stable than ribozymes, DNAzymes are still susceptible to nucleases in the cellular medium.[122]

In an attempt to address oligonucleotide cellular stability, the backbone of oligonucleotides in general has been modified extensively.[121] These modified scaffolds strive to preserve the general structure of oligonucleotides and have been introduced into the substrate binding regions of DNAzymes.[123-126] The modified oligonucleotide backbones that have been studied include phosphorothioate DNA,[127,128] LNA (locked nucleic acid),[129,130] 2′-O-methyl RNA[131] and PNA (peptide nucleic acid)[132] as shown in Figure 1.14. The use of LNA-modified oligonucleotides is particularly interesting in that LNA displays significantly increased affinity for nucleic acid targets. This makes LNA-derived therapeutics more capable of accessing cleavage sites that are obstructed by the three dimensional structure of the RNA target molecule. To improve the cellular uptake of oligonucleotides, complexing delivery vehicles have been used, synthetic conjugates have been directly attached and neutral/positive oligonucleotide backbones, including PNA, have been explored.[49,133]

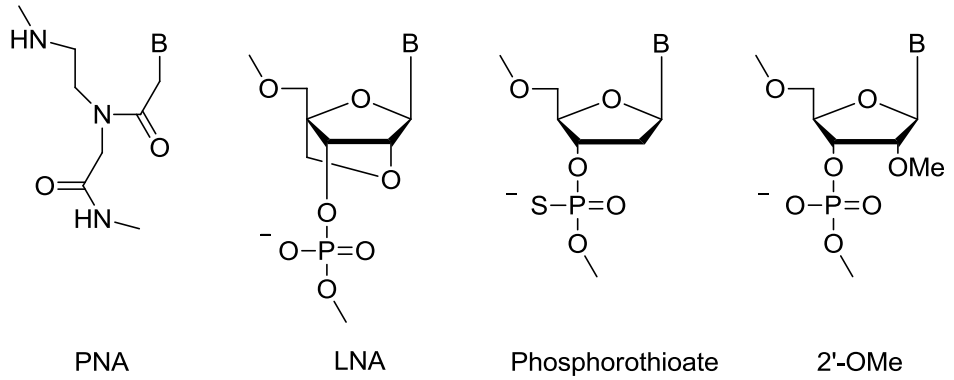


Figure 1.14 Modified oligonucleotides of enhanced cellular stability.

1.10 Research initiatives

This thesis is focused on the development of DNAzymes with expanded protein-like chemical functionality. As mentioned in section 1.8.1 and 1.8.2, many efforts have been directed towards the synthesis and enzymatic incorporation of modified 2′-deoxynucleoside triphosphates.[134-136] In particular, two modified dNTP's, 8-[(2-(4-imidazolyl)ethylamino)]-2′-deoxyadenosine and 5-aminoallyl-2′-deoxyuridine triphosphates, were successfully used for the *in vitro* selection of DNAzyme 9_{25}-11 (section 1.8.3).[69] Both of these two modified nucleotides have linkers between the nucleobase and the functional group that are relatively shorter than most other similarly modified nucleotides.[83,92-94,101] Because DNAzymes such as 9_{25}-11 are conformationally more flexible than protein enzymes, it is hypothesized that the use of modified dNTP's that have short linkers will reduce the entropy associated with the free rotation of appended functional groups thereby resulting in DNAzymes of enhanced activity. As such, several modified nucleotides with relatively short linker arms were synthesized. These nucleic acid monomers were then investigated for their ability to be incorporated into oligonucleotides by readily available DNA polymerases. Upon establishment of conditions under which the modified residues could be used, one of these molecules was used for the *in vitro* selection of a DNAzyme sensor and another was converted to a phosphoramidite. This phosphoramidite was used for the synthesis of DNAzymes.

Chapter 2 discusses the synthesis and enzymatic incorporation of two new nucleotides with modifications attached to a linker arm from the 5-position of 2′-deoxyuridine. They were subjected to primer extension reactions under a wide range of conditions. PCR was performed on both of them as well. Chapter 3 evaluates the importance of the linker that attaches functional groups to nucleotides. Modified

nucleoside triphosphates containing linkers of different lengths and composition between the 8-position of 2′-deoxyadenosine and an imidazole group were synthesized. These non-natural dATP's were then tested as substrates for DNA polymerases in the same manner as the modified dUTP's in the previous chapter. Following the synthesis and enzymatic incorporation of a guanidinium-containing 2′-deoxyuridine in Chapter 2 as well as the successful usage of this nucleotide in several DNAzyme selections,[137,138] the phosphoramidite of this nucleotide was synthesized. In Chapter 4, this phosphoramidite was used for the solid phase synthesis of oligonucleotides to study its effects on both duplex stability and the activity of the very thoroughly studied DNAzyme 10-23. Chapter 5 discusses an *in vitro* selection that was originally aimed at identifying sugar-sensing DNAzymes. However, only non-sugar-sensing divalent metal-dependent DNAzymes were isolated. Chapter 6 provides a synopsis of this dissertation as well as a discussion about future experiments that may be carried out.

Chapter 2: Functionalizing 2′-deoxyuridine triphosphates

2.1 Introduction

The synthesis of modified nucleoside triphosphates has mostly been accomplished via the functionalization of the 5-position of 2′-deoxyuridine.[82,95-97,101] Such modified nucleotides have generally been shown to be relatively good substrates for DNA polymerases.[103] Typically, an allyl or propargyl amino group is attached to the 5-position of 2′-deoxyuridine triphosphate (Figure 2.1), and the resulting nucleotide then serves as a precursor that can be coupled to different activated esters. This method allows for the rapid synthesis of modified nucleotides with a very wide range of functional groups. In this manner, Sakthivel and Barbas[83] reacted 5-aminoallyl-2′-deoxyuridine triphosphate (**2.2**) with *N*-hydroxysuccinimide esters that contained functionalities such as imidazole, pyridine and phenol groups to give nine different modified nucleoside triphosphates. Others have used 5-aminomethyl (**2.1**),[89] 5-aminopropargyl (**2.3**)[84] as well as 5-(2-(6-aminohexylamino)-2-oxoethyl)-2′-deoxyuridine triphosphates (**2.4**).[92,99] We chose to synthesize two modified 2′-deoxyuridine triphosphates in a similar manner as well.

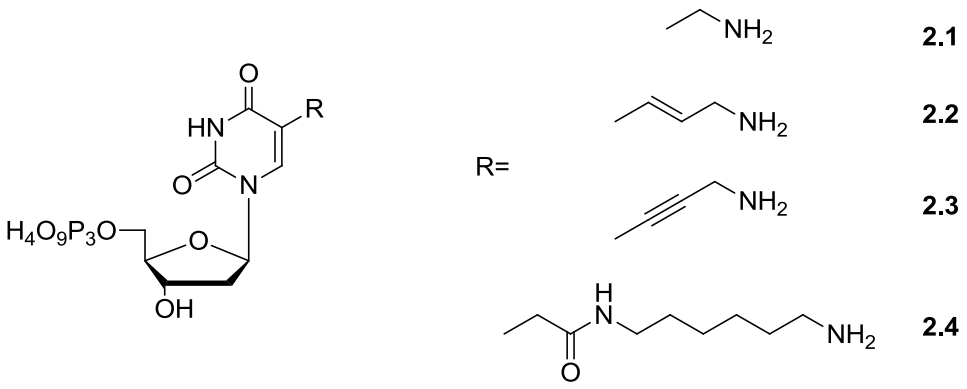


Figure 2.1 Amine-functionalized 2′-deoxyuridine triphosphates.

This chapter focuses on the attachment of two modifications onto 2′-deoxyuridine: a guanidine (or guanidinium ion under physiological conditions) group and a boronic acid group. The guanidinium group adds several dimensions to nucleic acids. The guanidinium group provides a cation that is charged over a very wide pH range allowing it to recognize, bind and orient anionic groups.[139,140] As such, the guanidinium groups of arginines are very often found in the ligand binding sites of enzymes that act on anionic substrates or cofactors.[139,141] Because the positive charge is distributed amongst the three nitrogen atoms such that multiple hydrogen bonds can be formed, the guanidinium group can also stabilize nucleic acid duplexes and triplexes.[142,143] Both these properties could play a very significant role in the three-dimensional folding of a DNAzyme. Finally, guanidinium groups can potentially mimic the function of arginine residues in cell-penetrating peptides (CPP's) to enhance the membrane permeability of nucleic acids.[144]

Three guanidinium-containing nucleoside triphosphates can be found in the literature (Figure 2.2). Jäger *et al.*[84] synthesized **2.5** and **2.6**, both originating from **2.3**. One has a triple bond in the linker arm (**2.5**) and the other only has single bonds in the

linker arm (**2.6**). Both triphosphates were found to be substrates for DNA polymerases and could be used for PCR. Another group of researchers attached *N*-(6-guanidinohexyl)carbamoylmethane to 2′-deoxyuridine triphosphate (**2.7**).[93] Triphosphate **2.7** has subsequently been used in the *in vitro* selection of a glutamate-binding aptamer.[145] All three of these triphosphates can be readily used for the synthesis of guanidine-modified DNA.

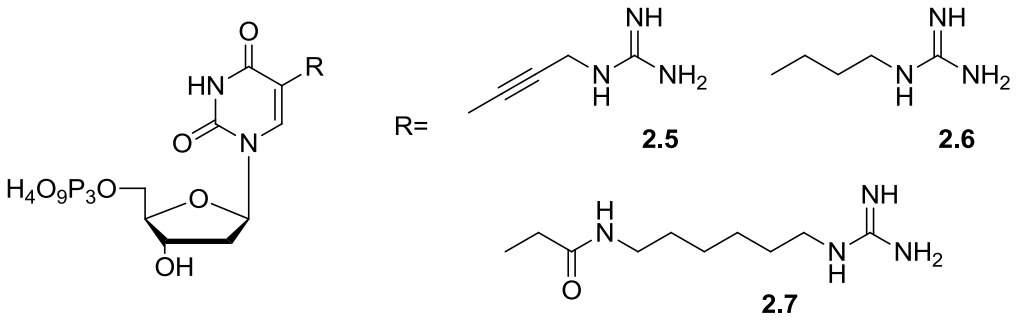


Figure 2.2 Guanidine-modified nucleotides synthesized by others.[84,93]

The recognition of sugars by chemically synthesized molecules has been rather difficult in aqueous environments because the structure of carbohydrates is dominated by hydroxyl groups, and hydroxyl groups are very similar to water.[146] One method of synthesizing receptors that can discriminate between water and sugar substrates is to use molecules functionalized with boronic acids.[147-152] Boronic acids are known to react with *cis*-1,2 and *cis*-1,3 diols to form covalently-bonded five- and six-membered cyclic esters, respectively and can reversibly form such covalent bonds with sugars in water.[147,153-155] Although unmodified nucleic acids have been used to target sugar-containing substrates,[156-158] boronic acid-containing nucleic acids should be able to target sugars with increased specificity and sensitivity.

Three boronic acid-modified nucleotides have been synthesized by others (Figure 2.3). Stolowitz *et al.* synthesized two derivatives of boronic acid-functionalized 2′-deoxyuridines (**2.8** and **2.9**) to be used in hybridization assays for the detection of target genes.[91] Also, Lin *et al.* published the synthesis of a boronic acid-containing nucleotide (**2.10**) that also used a relatively longer linker.[90] The same research group subsequently used **2.10** in the *in vitro* selection of an aptamer for the glycoprotein fibrinogen.[159]

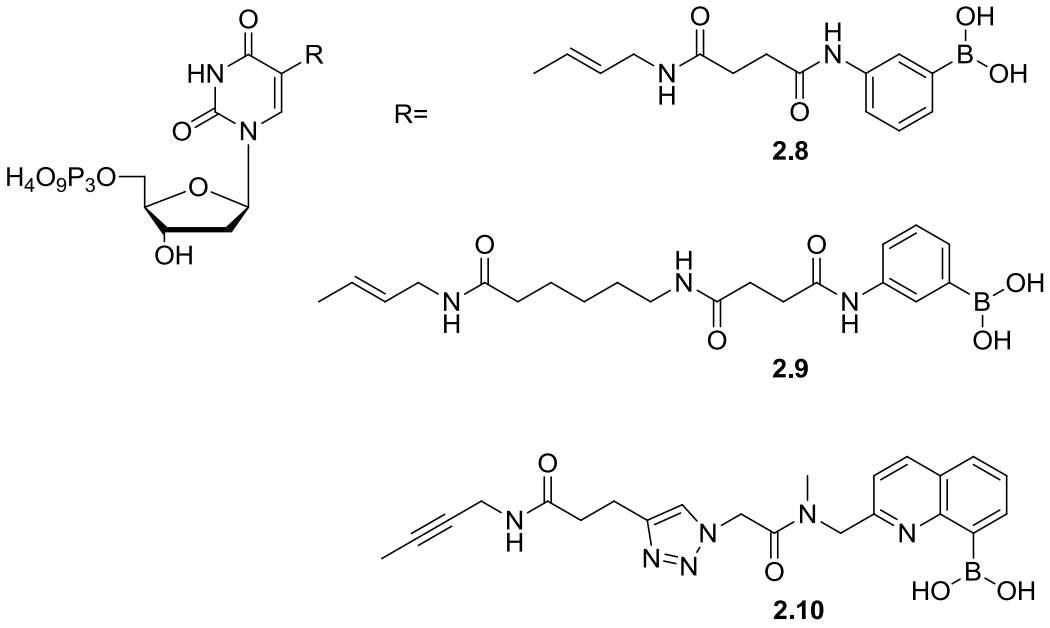


Figure 2.3 Boronic acid-modified nucleotides synthesized by others.[90,91]

In an effort to produce DNAzymes of greater catalytic activity and increase the ability of DNA to recognize carbohydrates, the synthesis of two modified nucleoside triphosphates, **2.11** and **2.12** (Figure 2.4), was undertaken. A guanidinium group was added to **2.2**, and a phenylboronic acid was attached to **2.1**. After the modified

nucleotides were synthesized, they were then used in primer extensions to find conditions under which the modified residues could be incorporated into oligonucleotides. Triphosphate **2.11** is very similar to **2.5** and **2.6** and will supplement what has already been established in terms of the DNA polymerase incorporation of 2′-deoxyuridine triphosphates that have closely attached guanidinium groups. Compared to **2.8**, **2.9** and **2.10**, triphosphate **2.12** contains an arylboronic acid that is situated much closer to the nucleobase. The relative shortness of the linker arm that connects the nucleotide with the modification will allow the evaluation of a modified 2′-deoxyuridine with a shorter linker in terms of polymerase incorporation and DNAzyme activity.

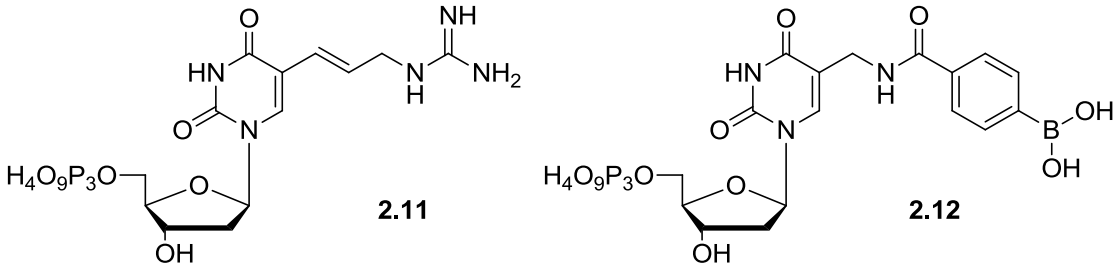
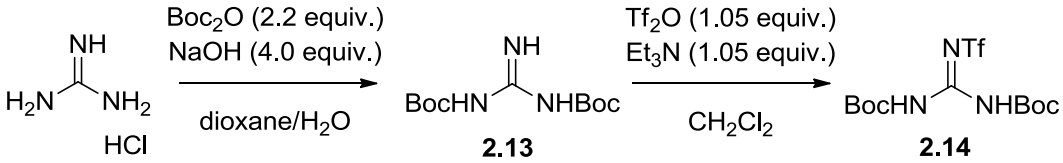


Figure 2.4 5-Position modified 2′-deoxyuridine triphosphates.

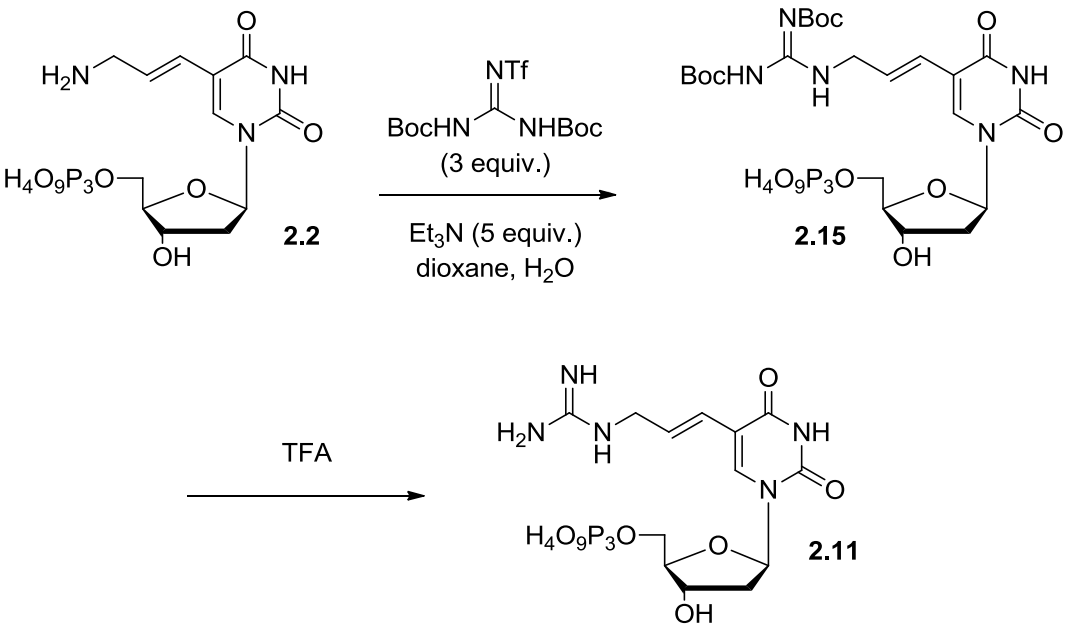
2.2 Synthesis of triphosphate 2.11

The preparation of the guanidinium-functionalized nucleoside triphosphate began with the synthesis of a guanidinylating agent according to the literature (Scheme 2.1). Guanidinium hydrochloride was deprotonated with sodium hydroxide in water and dioxane, and the resulting solution of guanidine was reacted with di-*tert*-butyl dicarbonate (Boc anhydride) to give the bis-protected **2.13**. Compound **2.13** was then mixed with triflic anhydride in the presence of triethylamine to give **2.14**.[160] Initially, the

mild triflating agent *N*-phenyl-bis(trifluoromethanesulfonimide) was used in place of triflic anhydride, but this was unsuccessful. The remaining steps for the synthesis of **2.11** are shown in Scheme 2.2. The guanidinylating agent **2.14** was added to the commercially available 5-aminoallyl-2′-deoxyuridine triphosphate (**2.2**) in the presence of triethylamine, dioxane and water to give the Boc-protected triphosphate **2.15**. A very brief deprotection treatment with neat trifluoroacetic acid (TFA) gave rise to the desired triphosphate **2.11**. This last step was a particularly low yielding step (23 %). Another research group using the same deprotection method for a related molecule observed a similar yield as well.[84] Triphosphate **2.11** was purified first by preparative TLC (thin layer chromatography) and then HPLC.



Scheme 2.1 Guanidinylating reagent synthesis.

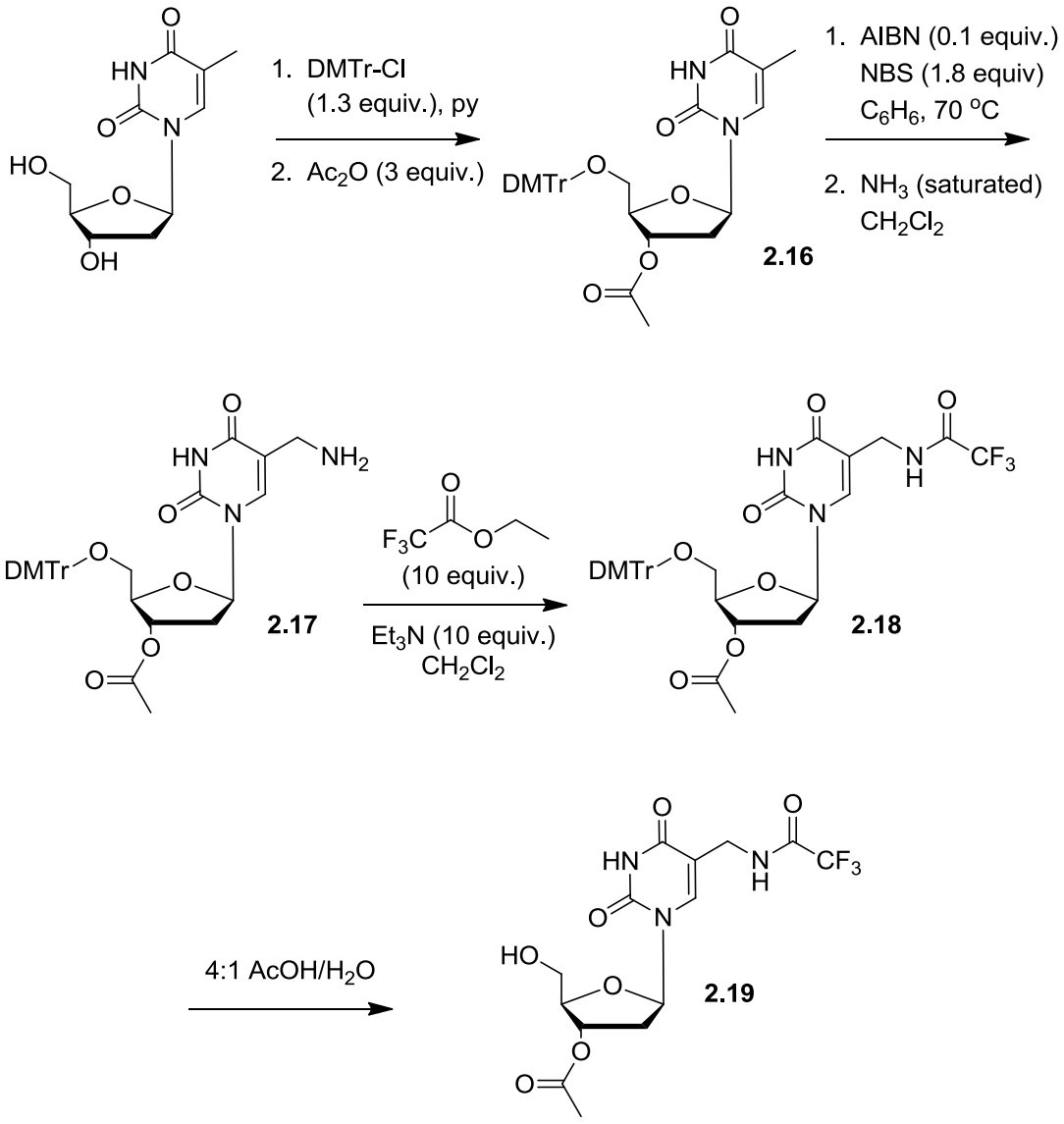


Scheme 2.2 Synthesis of **2.11**.

2.3 Synthesis of 2.12

In contrast to the synthesis of **2.11**, which began with a commercially available triphosphate, the synthesis of **2.12** began at the nucleoside level. Thymidine has been previously converted to **2.1** by Kolpashchikov *et al.*[89] We chose to use a different triphosphorylation method, a method that we were more familiar with and as a result required a somewhat alternate synthetic path to be taken. In our synthesis, all the precursors could be readily purified by silica flash chromatography.

2.3.1 Nucleoside preparation

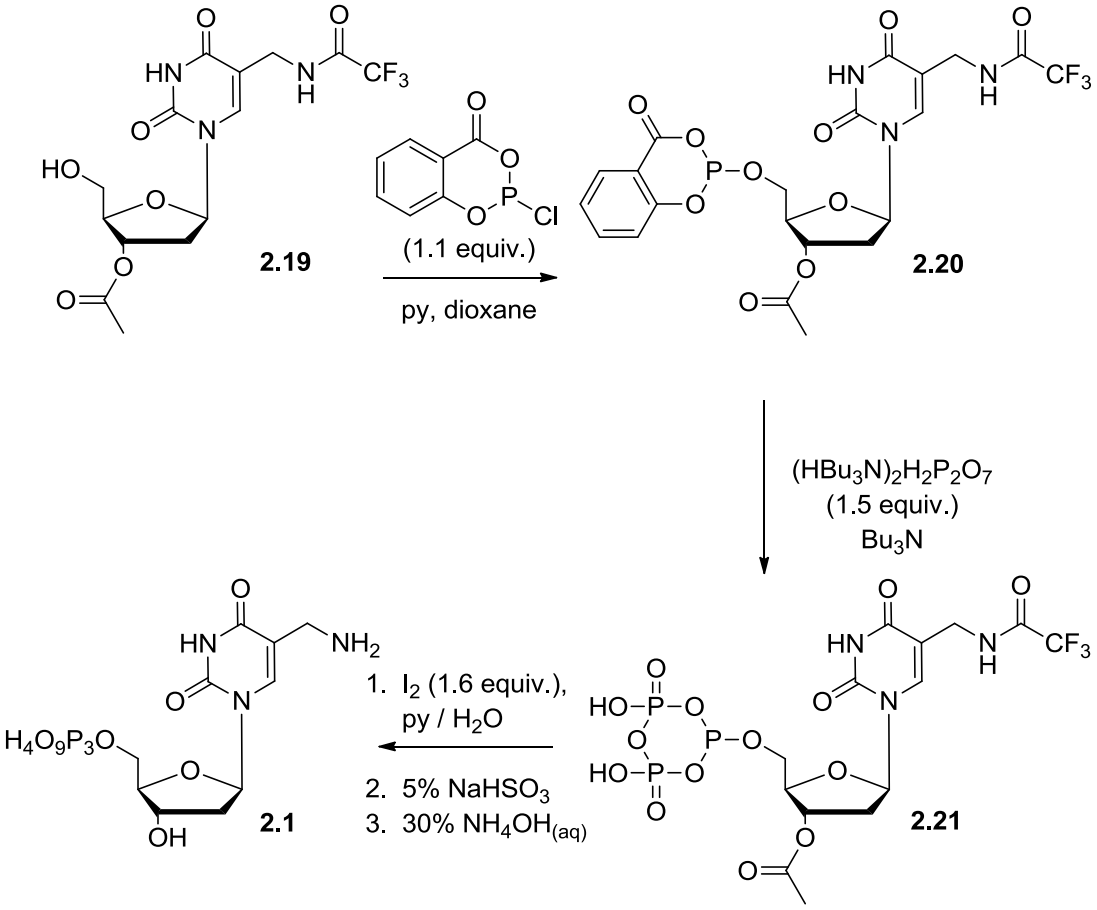


Scheme 2.3 Synthesis of nucleoside **2.19**.

The synthesis of 5-aminomethyl-2′-deoxyuridine triphosphate began with the preparation of the suitably protected modified nucleoside as shown in Scheme 2.3. Thymidine in pyridine (py) was treated first with 4,4′-dimethoxytrityl chloride (DMTr-Cl) followed by acetic anhydride (Ac_2O) to give the doubly protected **2.16**. This nucleoside

was then monobrominated in the presence of *N*-bromosuccinimide (NBS), 2,2′-azobis(2-methylpropionitrile) (AIBN) and benzene, conditions that have been previously used on 5′,3′-diacetylthymidine.[161] Due to the somewhat acidic nature of the bromination reaction, a reasonable amount of detritylation was observed and hence the yield of the reaction was not very high. The resulting crude product was dissolved in methylene chloride, and ammonia was bubbled into this solution. Under anhydrous conditions, the bromide was displaced, and the acetyl group left in place to give **2.17**. The free amine was then masked using ethyltrifluoroacetate in the presence of triethylamine and methylene chloride to form **2.18**. Finally, the DMTr group was removed in the presence of acetic acid/water (4:1) to give nucleoside **2.19**.[162]

2.3.2 Triphosphorylation



Scheme 2.4 Triphosphorylation of **2.19**.

The protected nucleoside **2.19** was converted to the triphosphate product **2.1** using a "one pot" five step synthesis developed by Ludwig and Eckstein (Scheme 2.4).[163] The nucleoside was first reacted with 2-chloro-4*H*-1,3,2-benzodioxaphosphorin-4-one. Bis-tributylammonium pyrophosphate was then added to the resulting phosphite (**2.20**) to displace salicylate. This is a double displacement that forms a dioxonucleosidyltriphosphite (**2.21**). The commercially available tributylammonium salt

of pyrophosphate[164] was used for its low water content and solubility in organic solvents such as DMF. NMR studies have confirmed that the trivalent phosphorus is indeed attached to the 5′-oxygen.[163] Following iodine oxidation and quenching with sodium bisulfite, aqueous ammonium hydroxide was added to open up the triphosphate ring and perform global deprotection. The first two steps of the triphosphorylation process were monitored by ^{31}P-NMR as illustrated in Figure 2.5. The phosphorinone and pyrophosphate reagents were found at 147.4 ppm and –9.3 ppm, respectively. Attachment of the phosphorinone to the 5′-hydroxyl group gave rise to a pair of singlet peaks at 126.5 and 125.5 ppm. The addition of tributylammonium pyrophosphate gave rise to an ABX spectrum: a triplet for the phosphite centre that is further shifted to 108.5 ppm and two pseudo quartets for the phosphates at –19.7 and –20.2 ppm. A minor peak was found in each and every spectrum between 10 and –5 ppm. The minor peak in Figure 2.5B corresponds to phosphoric acid, and the other minor peaks between 10 and –5 ppm in Figure 2.5 correspond to organophosphate species. The crude product was purified by preparative TLC.

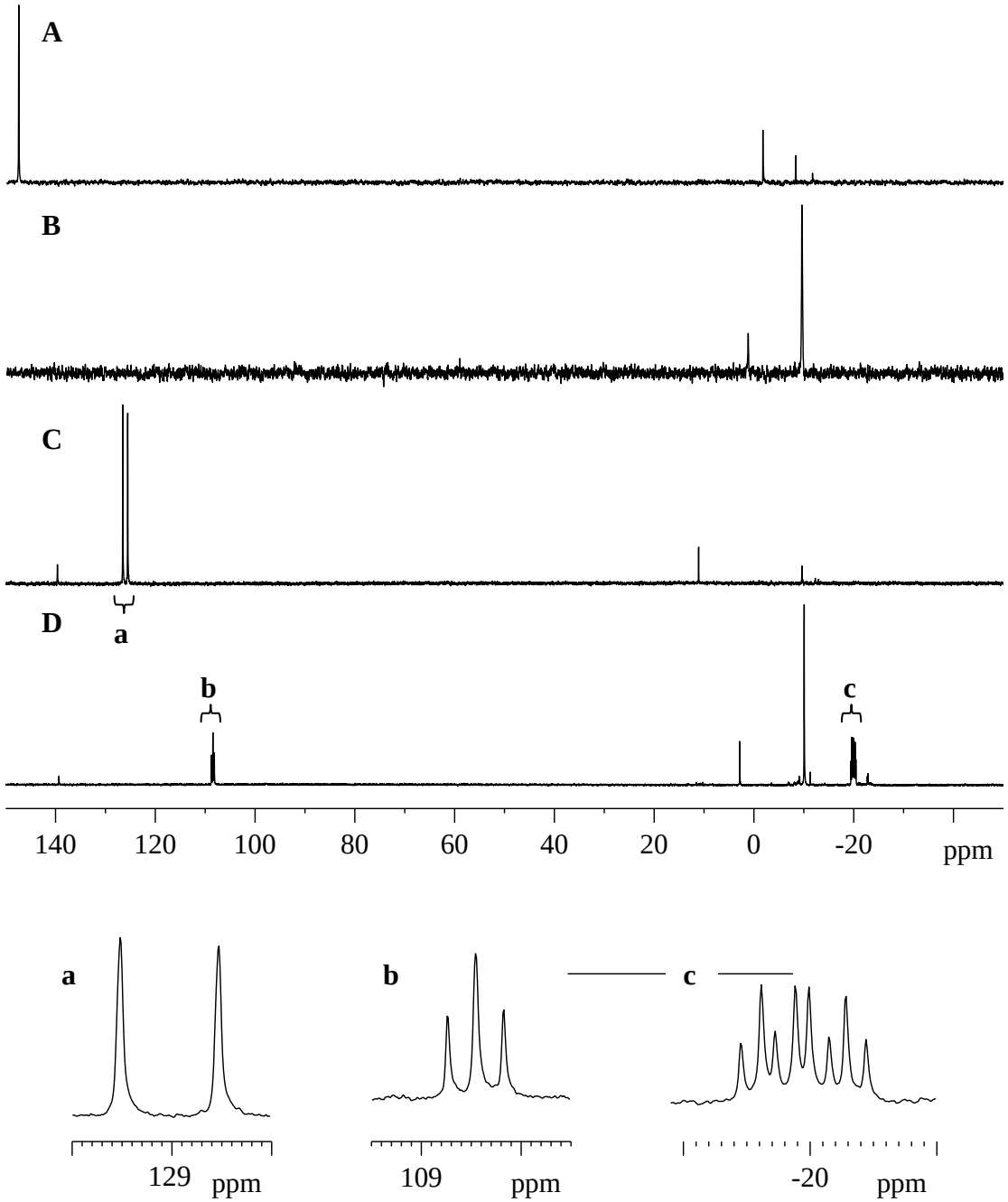
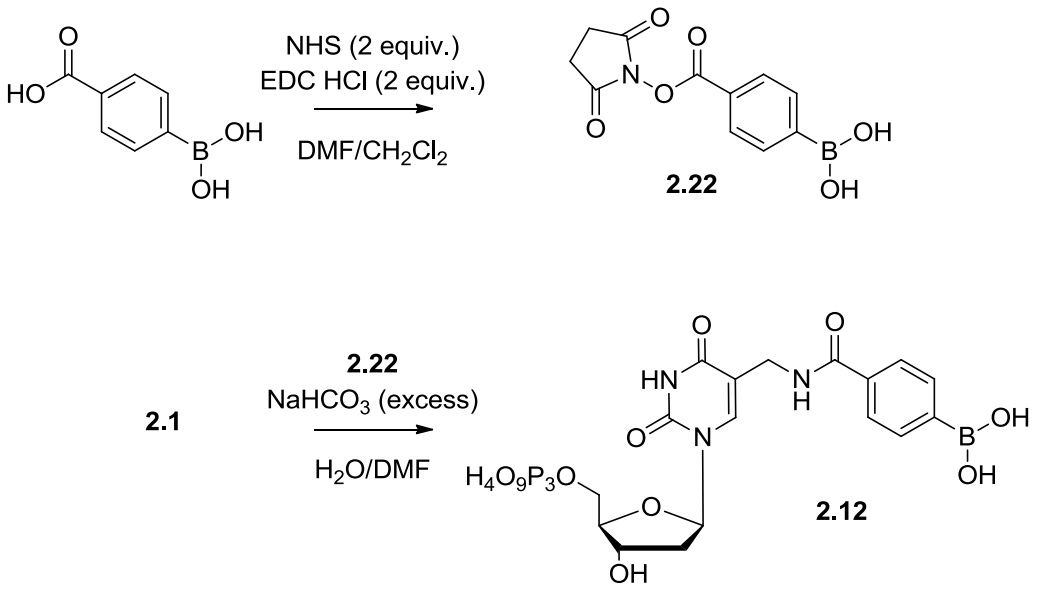


Figure 2.5 ^{31}P-NMR monitoring of the triphosphorylation progress of **2.19**. A. 2-Chloro-4*H*-1,3,2-benzodioxaphosphorin-4-one in dioxane. B. Tributylammonium pyrophosphate in DMF and dioxane. C. Reaction after the addition of phosphorinone. D. Reaction after the addition of tributylammonium pyrophosphate. Region "a" is a zoomed in view of C (128–126 ppm). Regions "b" & "c" are zoomed in views of D (109–108 ppm and –19 – –21 ppm, respectively). Calculated ABX parameters for D: δ_A = –19.7 ppm, δ_B = –20.2 ppm, J_{AB} = 25.9 Hz, J_{AX} = 43.7 Hz and J_{BX} = 47.3 Hz. NMR (122 MHz for A and B, 163 MHz for C and D) spectra were obtained at room temperature.

2.3.3 Conversion of 2.1 to 2.12

As shown in Scheme 2.5, compound **2.22**, as prepared by Masuda *et al.*,[165] was obtained by reacting 4-carboxyphenylboronic acid with *N*-hydroxysuccinimide (NHS) in the presence of water soluble carbodiimide (EDC). The NHS-ester was then added to **2.1** under slightly basic conditions to give what was initially thought to be triphosphate **2.12**. This product was a more mobile material on TLC than the starting material due to the conversion of the amine to an amide. The product was purified by preparative TLC and then HPLC.



Scheme 2.5 Synthesis of triphosphate **2.12**.

2.3.4 Preparation of 2.12 by an alternate route

The synthesis of **2.12** was initially attempted with a synthetic route that differed from what is illustrated in Schemes 2.4 and 2.5 in that the phenylboronic acid was attached to the nucleoside prior to triphosphorylation. However, the triphosphorylation procedure was not successful. During the synthesis of **2.10**, Lin *et al.*[90] also attempted

boronic acid functionalization prior to triphosphorylation using a different triphosphorylation method, one that did not involve an oxidation procedure, and were unsuccessful as well.

2.4 Characterization of 2.12

Numerous different techniques were used in an attempt to characterize the product that was isolated. Initially, ^{1}H-NMR (in D_2O) and MALDI-TOF MS were performed. No peaks corresponding to product or related compounds could be identified through MALDI, but in the ^{1}H-NMR spectrum, the proton shifts in the aromatic region correlated to the presence of a *para*-substituted phenyl group. Therefore, the NHS-ester coupling was thought to be successful. Later on, attempts were also made to try to verify that the boronic acid was not lost by some unknown process. Although others have used NHS-esters of unprotected phenyl boronic acids for the functionalization of amines,[165,166] it was felt that the verification of the presence of the boronic acid was still necessary. As boronic acids are known to interact with diols, two diol tests were performed on the product. The two diols that were used were Alizarin Red S (ARS)[167] and diol-functionalized aminomethyl Novagel resin that was prepared according to Li *et al.*[168] by Dr. Curtis Harwig (Figure 2.6). These two diols were chosen for their compatibility with aqueous conditions. ARS is a general optical reporter for boronic acids, and the diol-containing resin can be used for boronic acid affinity chromatography. Under buffered conditions (pH 7.4), a pink solution of ARS becomes orange in the presence of phenylboronic acids. For ARS, the triphosphate was used, and for the resin, both triphosphate and radio-labelled oligonucleotides were used. Unfortunately, neither of the two tests verified the presence of a boronic acid. Subsequently, the use of boron-11 NMR

was explored. Following the scaling up of the reactions shown in Schemes 2.4 and 2.5, ^{11}B-NMR revealed a peak at 18.8 ppm. Further verification of the presence of a boronic acid was obtained by the mass spectrometry analysis of the products obtained from the reaction of the NHS-ester **2.22** with benzylamine.

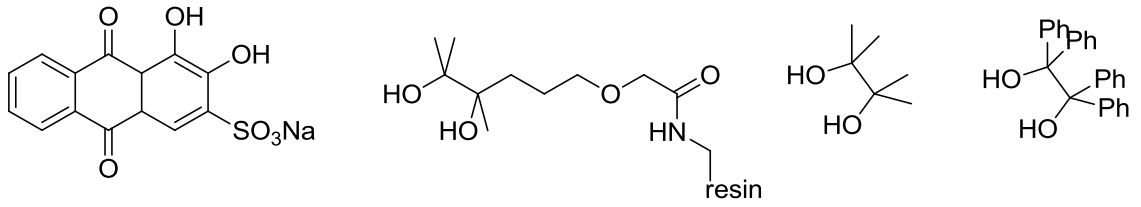
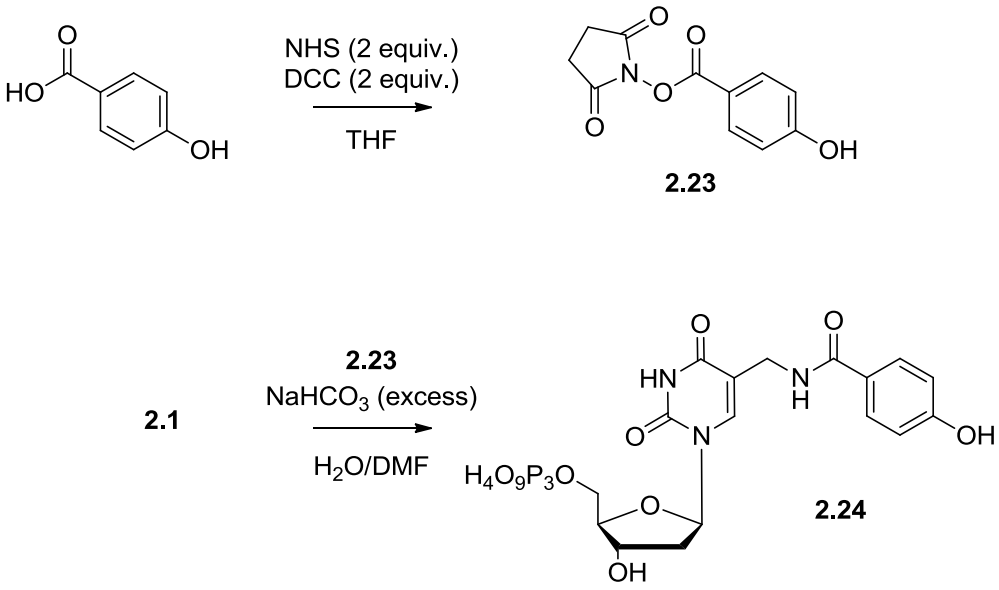


Figure 2.6 Boronic acid-binding diols. Left to right: ARS, diol-functionalized aminomethyl Novagel resin, pinacol, tetraphenylpinacol.

2.5 Synthesis of a phenol-modified 2′-deoxyuridine triphosphate (2.24)

Although ^{11}B-NMR revealed that our product sample contained some form of boron, the ^{11}B-NMR data was inconsistent with the lack of interaction with reporter diols, the same interactions that were to be exploited for the eventual targeting of sugars. The modified nucleoside triphosphate also did not appear to react with hydrogen peroxide, an oxidant that should oxidize the boronic acid to a hydroxyl group. Later on, an alternate and improved MALDI protocol did give rise to a peak corresponding to the mass of a phenol-modified nucleoside triphosphate where the boronic acid is replaced by a hydroxyl group. The presence of a hydroxyl group rather than a boronic acid suggested that triphosphate **2.12** may have undergone oxidative deboronylation. Arylboronic acids tend to be quite stable to oxidative deboronylation from air or aqueous oxygen. However, under very low concentrations during purification, the arylboronic acid was perhaps more susceptible to this destructive side reaction. To verify whether oxidative deboronylation

had indeed occurred, the synthesis of the phenol-containing 2′-deoxyuridine triphosphate (**2.24**) was performed. A phenol-modified nucleotide would give DNAzymes the side chain of tyrosine. Sakthivel *et al.*[83] have previously characterized a phenol-modified 2′-deoxyuridine. Triphosphate **2.24** differs from their nucleotide in that the linker arm is four carbon atoms shorter.



Scheme 2.6 Synthesis of triphosphate **2.24**.

The synthesis of **2.24** was very similar to the synthesis of **2.12** (Scheme 2.6). 4-Hydroxybenzoic acid was converted to the NHS ester **2.23** according to literature by reaction with NHS and *N*,*N*′-dicyclohexylcarbodiimide (DCC).[169] The resulting NHS ester was coupled to **2.1** to give the desired nucleoside triphosphate. Upon isolation of **2.24**, it was found that both **2.24** and what was thought to be **2.12** shared identical mass spectra, proton NMR, absorbance spectra, HPLC retention time and TLC mobility. Hence, the product that was isolated from our earlier synthesis of **2.12** was in fact **2.24**.

2.6 Probing the oxidative deboronylation of 2.12

When it was confirmed that quantitative oxidative deboronylation had occurred, the first thought was that this occurred during HPLC purification. If the irradiation from the photodiode array caused oxidative deboronylation, the resulting boric acid would still be isolated along with **2.24** to give a detectable signal with ^{11}B-NMR. However, the reaction requires oxygen, a gas that was probably almost absent due to the vacuum degasser of the HPLC system. If oxidative deboronylation occurred prior to HPLC purification, boron-containing species would be removed and no peak should be observed with ^{11}B-NMR. When the nucleotide couplings shown in Schemes 2.5 and 2.6 were performed in tandem, the reactions gave rise to products that co-migrated on TLC (dioxane/H_2O/NH_4OH 6:4:1). When the functionalization of **2.1** was attempted with the pinacol- and tetraphenylpinacol-protected versions of **2.22** (diol structures shown in Figure 2.6), the same product on TLC was observed. The pinacol-containing reaction was worked up and purified as well. The isolated material was found to exhibit the same properties as **2.24**. All these observations seemed to indicate that oxidative deboronylation was occurring during the NHS-ester coupling reaction. If this indeed happened, the signal that was observed with ^{11}B-NMR could not have originated from the synthesis (Scheme 2.5) and the ^{11}B-NMR signal was probably due to trace amounts of boric acid in the HPLC solvents.

During the writing of this dissertation, we decided to revisit this reaction to acquire information that would perhaps be helpful for the synthesis of **2.12** or other boronic acid-containing nucleotides. New batches of reagents and reactants were prepared, and reactions of benzylamine with either **2.22** or tetraphenylpinacol-protected **2.22** did not

show any signs of oxidative deboronylation. However, when the reaction illustrated in Scheme 2.5 was repeated, a product with a very different TLC mobility was found. Instead of having increased mobility on TLC when compared to **2.1**, the product exhibited decreased mobility. It is unclear as to what led to this unexpected result, but the reaction was repeated twice and gave rise to the same unexpected product rather than **2.24**. Previously, preparative-TLC purification was performed with an eluent that contained dioxane. Although the concentrations of peroxides in dioxane ($\leq$ 1 mg/L) should not be able to quantitatively oxidize the nucleotide, dioxane was avoided as a precaution. Triphosphate **2.12** was successfully identified using MALDI as the $[M–2H_2O]^-$ adduct of 608 *m/z* as shown in Figure 2.7. The isotopic peaks in the spectrum also matched what would be expected for the product, namely that boron exists as 19.9 % ^{10}B and 80.1 % ^{11}B. This product was also readily oxidized to **2.24** with aqueous hydrogen peroxide. It should be noted that there is a small peak at 616 *m/z* that corresponds to **2.24** in Figure 2.7. The identity of the peak at 639 *m/z* could not be deciphered. Enzymatic incorporation of actual **2.12** could be challenging if contaminating amounts of **2.24** are always present. If **2.24** is a better substrate than **2.12**, **2.24** could be preferentially incorporated even if it is only present at very low concentrations.

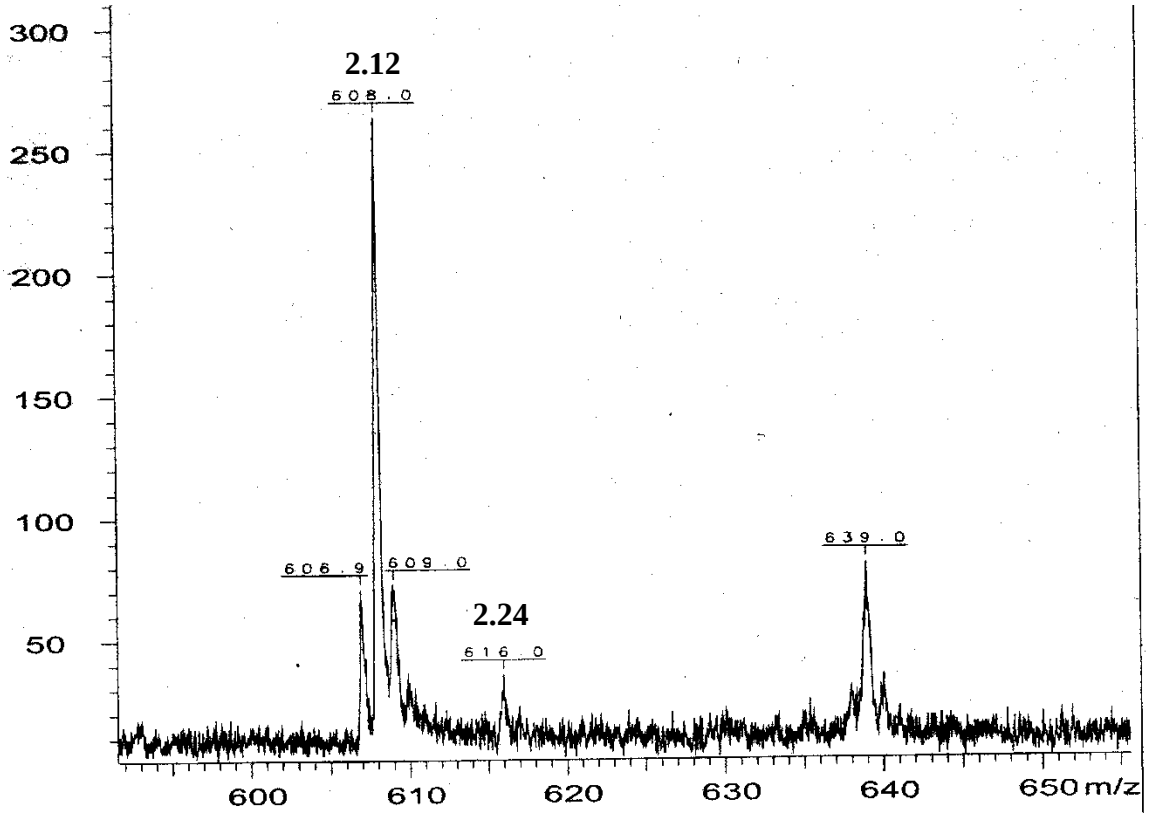


Figure 2.7 MALDI analysis of **2.12**. Analysis was performed in negative ion mode.

2.7 Enzymatic incorporation of 2.11

Following the successful synthesis of **2.11**, it was used in incorporation studies to establish conditions under which the modified nucleotide could be used. The DNA polymerases that were used were generally polymerases that lacked a 3′-5′ proofreading ability and have been used by other groups on modified residues. Primer extensions were performed with one 30 nucleotide primer **P2.1**, which was radiolabeled with a 5′-^{32}P-phosphate, and three different 55 nucleotide templates. The composition of the templates excluding the primer binding regions is shown in Figure 2.8. Template **T2.2** directs the incorporation of a total of seven modified residues and up to two successive modified residues. Template **T2.1** is somewhat more demanding. It requires five modified

residues to be incorporated successively and an additional five modified residues to be incorporated. Incorporations using templates that require consecutive repetitions of the same modified nucleotide are particularly difficult for DNA polymerases to perform.[84] These types of templates force the accumulation of steric bulk and charge into a localized area that can dramatically affect the ability of the polymerases to associate with the double helix. This can result in a reduced rate or a termination of incorporation. Template **T2.3** directs the incorporation of five T analogs and will be used for the incorporation of more than one different modified nucleotide.

The primer extensions were carried out in two different ways: 1) in the presence of only one dNTP and 2) under full extension conditions (all 4 dNTP's). In the presence of only the modified nucleotide of interest, the incorporation of the modified nucleotide can be directly compared to the incorporation of the unmodified counterpart. Both **T2.1** and **T2.2** will be used for such incorporations. Incorporations under full extension conditions will determine whether the modified nucleotide is suitable for the synthesis of modified DNA libraries for SELEX experiments.

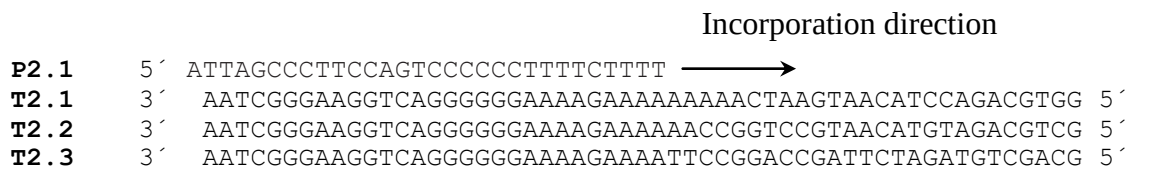


Figure 2.8 Primer and template sequences used for the enzymatic incorporation of **2.11** and **2.24**.

2.7.1 Incorporation of 2.11 with Sequenase V2.0 DNA polymerase

The incorporation of **2.11** was first attempted with Sequenase V2.0 DNA polymerase. The successful incorporation using **T2.2** template is shown in the denaturing polyacrylamide gel in Figure 2.9. Because product bands migrate from the top to the bottom, the unextended primer bands are found at the bottom of the gel and the fully extended products are found at the top of the gel. Lanes 1 and 2 correspond to unextended primer in the presence or absence of Sequenase V2.0. Sometimes, the exonuclease activities of DNA polymerases can result in the slow digestion of the DNA. Lane 2 verified that this did not occur under the conditions that were used. Lanes 3–5 were performed in the presence of just one nucleotide. The product band in lane 3 corresponds to the incorporation of one dideoxy residue, which terminates the incorporation. In lane 4, there are two bands. The more intense band corresponds to the incorporation of three dTTP's, and the less intense band corresponds to the incorporation of two dTTP's. The incorporation of an extra nucleotide was probably due to the absence of the other nucleotides and the lack of a proofreading activity. A similar pattern is observed in lane 5 as well. Lanes 6 and 7 were performed under full extension conditions. The bands in lane 6 and 7 correspond to fully extended unmodified and modified products, respectively.

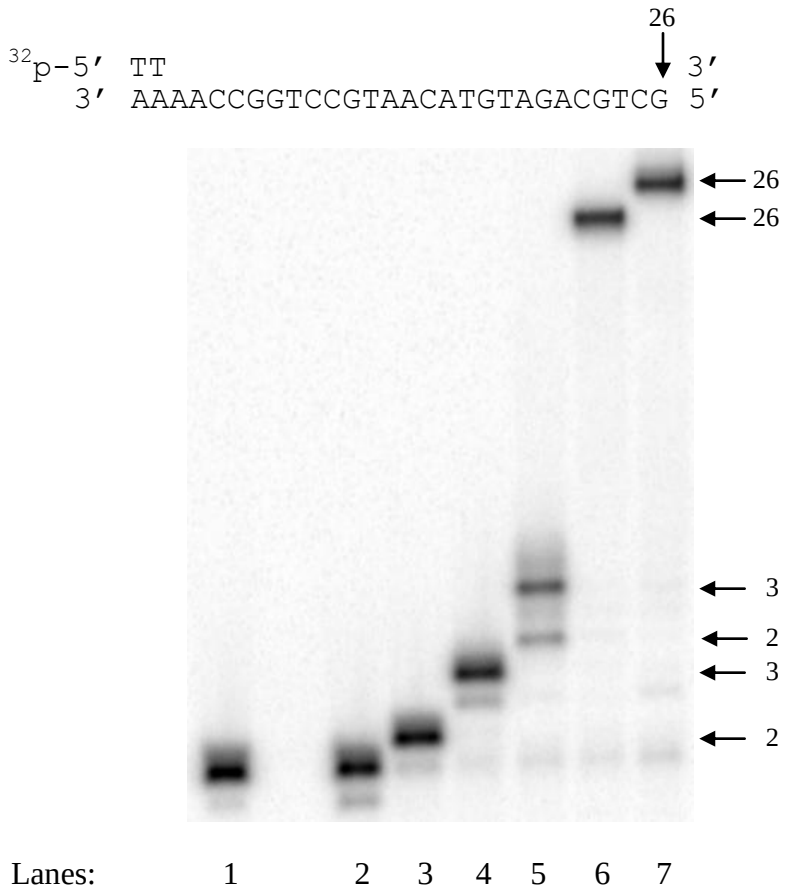


Figure 2.9 Sequenase V2.0 DNA polymerase incorporation of **2.11**. Assay performed at 37 ℃ using template **T2.2**. Lane 1, primer only; lane 2, primer and enzyme; lane 3, ddTTP; lane 4 dTTP; lane 5, **2.11**; lane 6, dGTP, dATP, dCTP and dTTP; lane 7, dGTP, dATP, dCTP and **2.11**. ddTTP was used at 20 μM and all other nucleotides were used at 50 μM. Numbers to the right of the gel image correspond to the number of nucleotides incorporated.

Incorporation of **2.11** gave rise to products that moved more slowly in denaturing PAGE when compared to the unmodified controls. Comparing lanes 4 and 5, the electrophoretic mobility retardation due to the incorporation of **2.11** was approximately equivalent to the retardation due to the incorporation of two dTTP's. Modifications typically reduce the electrophoretic mobility of oligonucleotides such that the presence of

product bands of lower electrophoretic mobility indicates that modified residues have been incorporated. The retardation is due to a combination of the molecular bulk/size of the modification and the addition of a cationic charge. As mentioned in section 2.1, Jäger *et al.* performed the synthesis and enzymatic incorporation of two guanidinium-modified nucleotides that contained either a propynyl (**2.5**) or a propyl linker (**2.6**).[84] They found that the incorporation of **2.5** produced product bands of lower resolution that were prone to smearing. These observations were not found in the incorporations of **2.6** or our incorporation of **2.11**. The incorporation shown in Figure 2.9 was also repeated with the more difficult template **T2.1** to also give fully extended products as well (results not shown).

After establishing that **2.11** could be efficiently incorporated by Sequenase V2.0, incorporations simultaneously involving several different modified residues were attempted. Three different modified residues were used: 8-histaminyl-dATP, 5-aminoallyl-dCTP and **2.11**. As shown in Figure 2.10, all three of them could simultaneously be incorporated by Sequenase V2.0 to give full length products. Lane 4 used only 8-histaminyl-dATP, lane 5 used both 8-histaminyl-dATP and 5-aminoallyl-dCTP and lane 6 used all three modified nucleotides for full extension assays. Compared to the unmodified control in lane 3, the incorporations shown in lanes 4 and 5 gave rise to similar amounts of full length product, and the incorporation shown in lane 6 gave rise to a lower yield of full length product. The incorporation of each of the three modified nucleotides led to further gel retardation in the resulting modified oligonucleotide products.

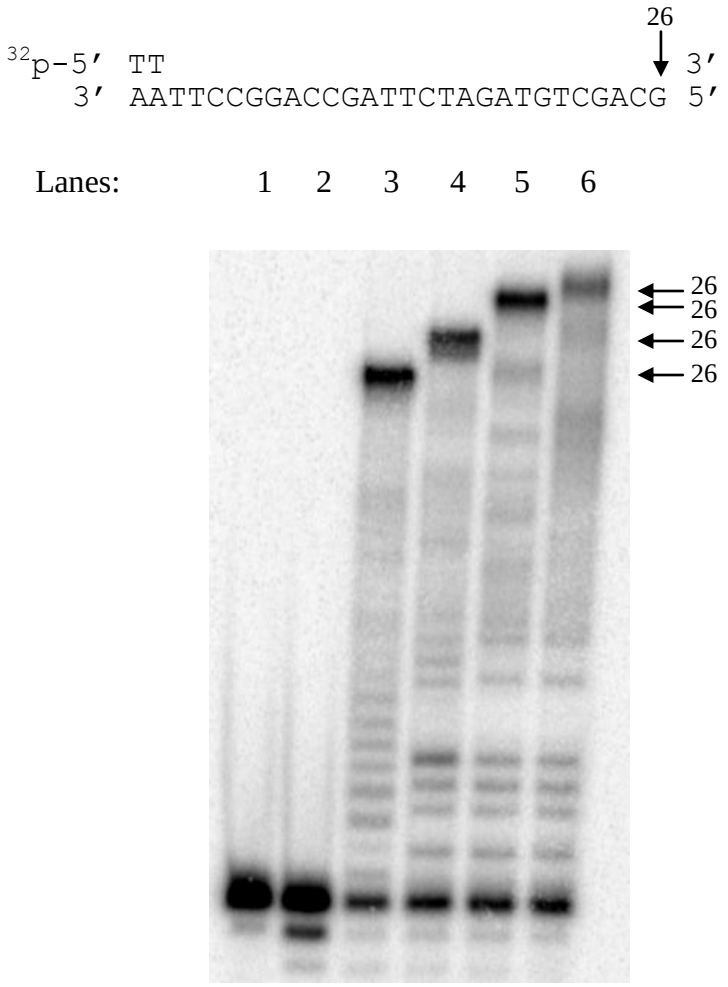


Figure 2.10 Sequenase V2.0 DNA polymerase incorporation of three different modified nucleotides. Assay performed at 37 ℃ using **T2.3**. Lane 1: primer only; lane 2: primer and enzyme; lane 3: dGTP, dATP, dCTP, dTTP; lane 4: dGTP, dCTP, dTTP, 8-histaminyl-dATP; lane 5: dGTP, dTTP, 8-histaminyl-dATP, 5-aminoallyl-dCTP; lane 6, dGTP, 8-histaminyl-dATP, 5-aminoallyl-dCTP, **2.11**. All nucleoside triphosphates were used at 50 μM. Numbers to the right of the gel image correspond to the number of nucleotides incorporated.

The incorporation of 8-histaminyl-dATP, 5-aminoallyl-dCTP and **2.11** in identical conditions allows constructs simultaneously containing free ammonium, imidazole and guanidinium groups to be enzymatically synthesized for *in vitro* selections. DNAzyme

9_{25}-11 was discovered using 8-histaminyl-dATP and **2.2**.[69] The three modified nucleotide system shown in lane 6 of Figure 2.10 is more sophisticated and could potentially give rise to more effective DNAzymes.

2.7.2 Incorporation of 2.11 with other DNA polymerases

To further evaluate the incorporation of triphosphate **2.11**, several other DNA polymerases were used in incorporation assays as well. In Figure 2.11, incorporation studies using Vent (exo–), *Pfu* (exo–), *Taq* and Klenow (exo–) DNA polymerases with **T2.2** are shown. It was found that three of these DNA polymerases could efficiently incorporate **2.11** to give full length products of somewhat retarded electrophoretic mobility. Incorporation using *Taq* gave rise to fully extended product as well as a much shorter band.

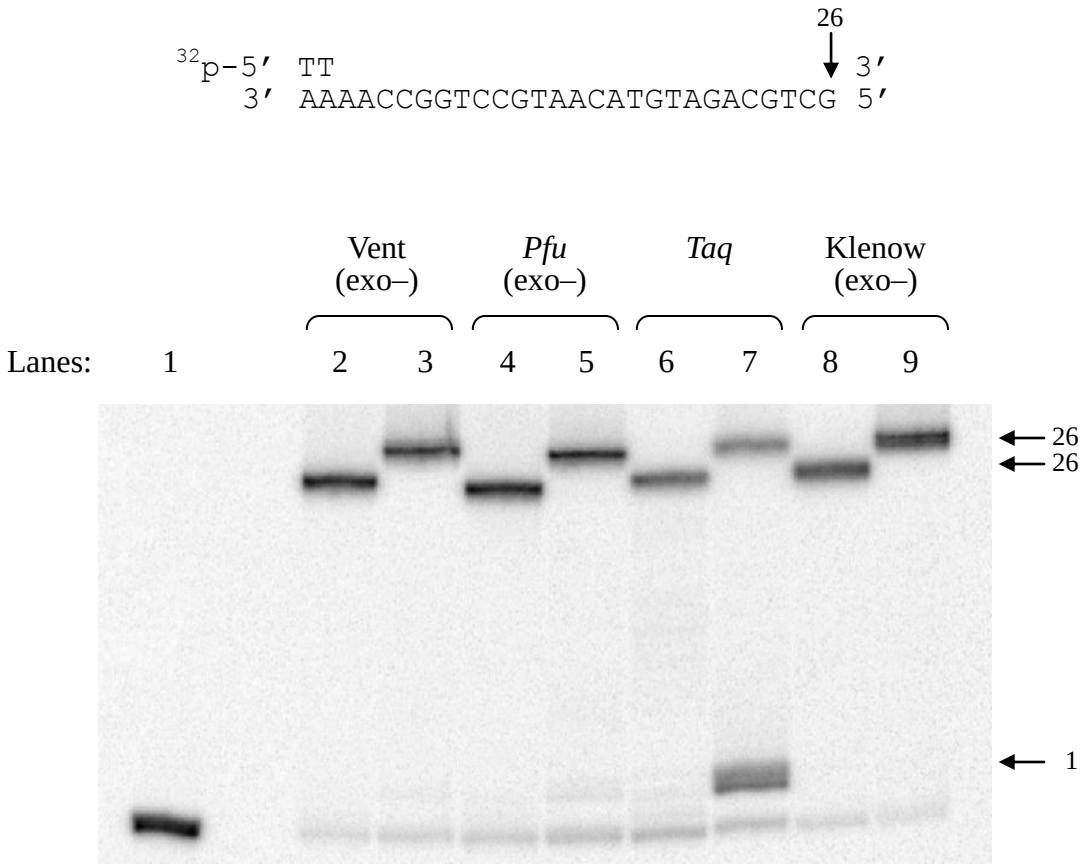


Figure 2.11 Incorporation with **2.11** under full extension conditions. Lane 1, primer only; lanes 2–8, incorporations with 4 different DNA polymerases: lanes 2 and 3, Vent (exo–); lanes 4 and 5, *Pfu* (exo–); lanes 6 and 7, *Taq*; lanes 8 and 9, Klenow (exo–). Reactions were performed for 30 min using nucleoside triphosphates at 50 μM and template **T2.2**. Lanes 2, 4, 6, 8 used dTTP, dATP, dCTP and dGTP. Lanes 3, 5, 7, 9 used **2.11**, dATP, dCTP, and dGTP. Lanes 1–7 were incubated at 72 ℃ and lanes 8 and 9 were incubated at 37 ℃. Numbers to the right of the gel image correspond to the number of nucleotides incorporated.

2.7.3 PCR with 2.11

Following the successful incorporation of **2.11** under primer extension conditions, PCR was attempted using unmodified templates. PCR experiments with modified nucleotides are much more difficult than the primer extensions of the same substrates.

The process produces material that has modifications on both strands rather than just one, and hence the synthesized modified DNA strands must act as templates in subsequent cycles for exponential amplification. If the modified DNA strands are poor templates for polymerization, very little product will be produced. Initially, 20 s extension times per cycle did not give rise to any modified product. When the elongation period was increased to 5 min, both the reaction with natural nucleotides and the reaction with **2.11** substituted for dTTP gave rise to properly sized products. Figure 2.12 shows the PCR of a 120 nucleotide template using Vent (exo–) on an agarose gel. Similar to the polyacrylamide gels, the largest and smallest oligonucleotides are found at the top and bottom areas of the gel, respectively. Compared to the lane 3 control, the modified product did not display any appreciable gel retardation on the agarose gel. Also, the PCR of **2.11** was not particularly efficient as the yield was only a third of the control. Based on the results in Figure 2.12, modified DNA synthesized using **2.11** is a satisfactory template for PCR, and hence, triphosphate **2.11** should have no problem being used for a selection that employs an unmodified amplification of what are initially modified templates.

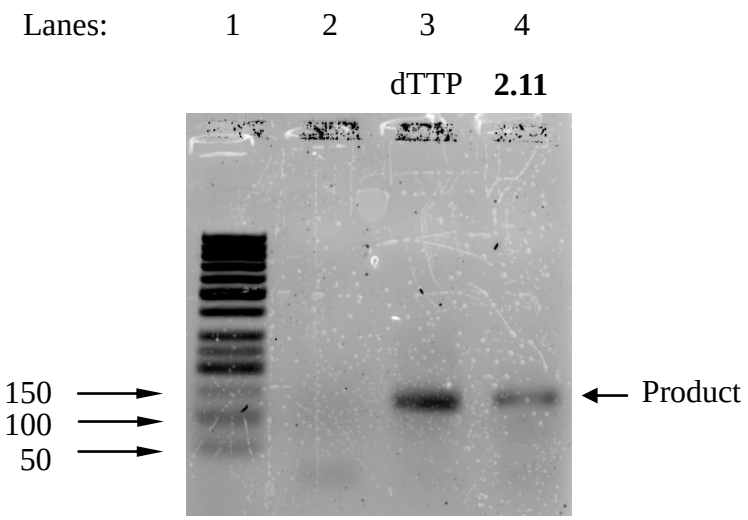


Figure 2.12 PCR with **2.11** and Vent (exo–). Total reaction mixture (10 μL) consisted of 10 ng plasmid template, 0.5 μM of each primer, 0.2 mM of each dNTP and 1 unit of enzyme. Reactions were performed at 1X thermopol buffer with a final Mg^{+2} concentration of 7 mM. Lane 1: DNA marker (50–1000 bp); lane 2: no template; lane 3: all 4 natural dNTP's; lane 4: dATP, dGTP, dCTP, **2.11**.

2.8 Incorporation of 2.24

After enzymatic uptake experiments were performed on **2.11**, similar studies were performed on **2.24** and its precursor **2.1**. The incorporation of **2.1** had not yet been studied, and the incorporation of it would give some insight into the incorporation of modified residues in general. Incorporation studies were performed in the same manner as the incorporation studies that were shown earlier in this chapter using the same primer and templates that are shown in Figure 2.8. Initially, primer extensions were performed to determine the maximum number of modified residues that could be incorporated in a row. Samples of template **T2.1** annealed to primer **P2.1** were polymerized by seven different DNA polymerases in the absence of the other 3 natural dNTP's.

The incorporation results using two DNA polymerases, Sequenase V2.0 and Klenow (exo–), are shown in Figure 2.13. Both triphosphates **2.1** and **2.24** were used as substrates for the 2 DNA polymerases at 4 different concentrations: 4, 12, 40 and 100 μM. As one would expect, the incorporation of either modified nucleotide resulted in gel bands of retarded mobility when compared to the unmodified controls. The retardation of gel bands obtained using **2.1** was due to a combination of the molecular bulk and positive charge of the modification, and the retardation of gel bands obtained using **2.24** was only due to the molecular bulk of the modification. Comparing lanes 5 and 10 in Figure 2.13B, the magnitude of retardation due to the two different modifications were approximately equal. Triphosphate **2.1** was a better substrate than **2.24** for both of the DNA polymerases used in Figure 2.13. At a concentration of at least 40 μM, 3 residues of **2.1** could be incorporated consecutively by Sequenase V2.0 and Klenow (exo–). In the case of **2.24**, Sequenase could only incorporate one modified residue and Klenow (exo–) could only incorporate 2 modified residues in a row.

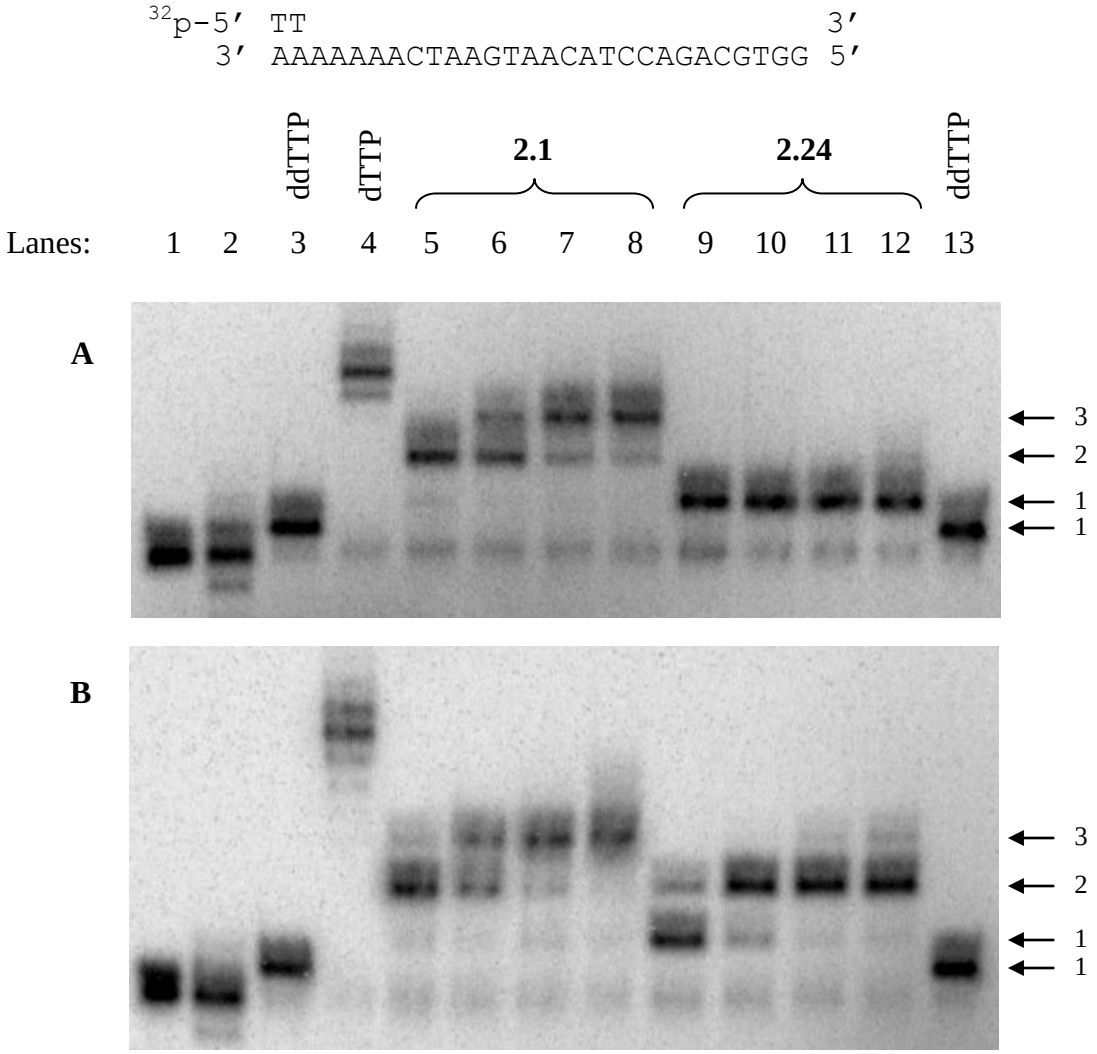


Figure 2.13 Incorporation of **2.1** and **2.24** by Sequenase V2.0 (A) and Klenow (exo–) (B). Assays were performed at 37 ℃ using template **T2.1**. Lane 1, primer only; lane 2, primer and enzyme; lane 3, ddTTP (10 μM); lane 4 dTTP (10 μM); lanes 5–8, **2.1** used at 4, 12, 40, 100 μM, respectively; lanes 9–12, **2.24** used at 4, 12, 40, 100 μM, respectively; lane 13, ddTTP (10 μM). Numbers on the right of the gel image correspond to the number of nucleotides incorporated.

Triphosphate **2.24** was also used as a substrate for Vent (exo–) as shown in Figure 2.14. The reactions shown were performed at very low concentrations of nucleotides in order to observe ladders of bands corresponding to the incorporation of 1–5 nucleotides.

As can be seen, **2.24** is a very good substrate for Vent (exo–). A nucleotide concentration of only 5.0 μM was sufficient for all 5 modified residues to be incorporated (lane 6). As was also observed in Figure 2.13, the incorporation of **2.24** gave rise to gel bands of retarded mobility. The product band corresponding to the incorporation of five unmodified residues matched the mobility of the product band corresponding to the incorporation of three residues of **2.24**.

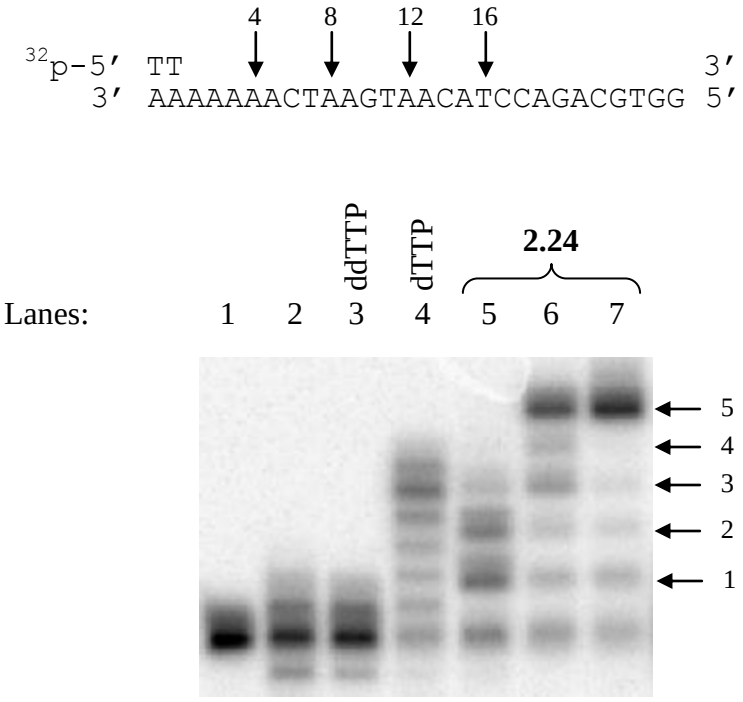


Figure 2.14 Incorporation of **2.24** by Vent (exo–). Reactions performed at 66 ℃ for 30 min using **T2.1** template. Lane 1: primer only; lane 2: primer and enzyme; lane 3: 5 μM ddTTP; lane 4: 2 μM dTTP; lanes 5, 6, 7: **2.24** 1, 5, 20 μM, respectively. Numbers on the right of the gel image correspond to the number of modified nucleotides incorporated.

The other polymerases that were used for the incorporation of **2.1** and **2.24** in the absence of the other 3 dNTP's (gel images not shown) and the results in Figures 2.13 and 2.14 are summarized in Table 2.1. The mesophilic DNA polymerases, Sequenase V2.0

and Klenow (exo–), were the most effective for incorporating **2.1**. Three consecutive residues of **2.1** could be incorporated. The use of thermophilic polymerases for the incorporation of **2.1** tended to produce a distribution of product bands. In the case of the incorporation of **2.24**, the opposite was true. The thermophilic DNA polymerases Vent (exo–), *Pfu* (exo–) and Dpo4 were all much more effective at incorporating **2.24** compared to the mesophilic DNA polymerases. All of those three DNA polymerases could incorporate the five consecutive modified residues of **2.24** that were required by template **T2.1**.

	Number of consecutive nucleotides incorporated	
DNA polymerase	**2.1**	**2.24**
Sequenase V2.0	3	1
Klenow (exo–)	3	2
Taq	0–2	0–1
Bst	1–2	1
Vent (exo–)	1–4	5
Pfu (exo–)	2–3	5
Dpo4	2	5

Table 2.1 Incorporation of **2.1** and **2.24** by DNA polymerases: maximum number of modified residues that can be incorporated consecutively in the absence of the other 3 dNTP's. Incorporation experiments were performed using template **T2.1** at 37 ℃ for Sequenase V2.0, Klenow (exo–), Dpo4 and at 68 ℃ for *Taq*, Vent (exo–), *Bst*, *Pfu* (exo–). Incorporations were incubated for 2 h, and the nucleotide substrates were used at a concentration of 100 μM. The incorporation of modified residues was limited to 5 by the template.

2.8.1 Incorporation of 2.24 under full extension conditions

The four DNA polymerases in Table 2.1 that incorporated **2.24** the best were used under full extension conditions. The results are shown in Figure 2.15. The template used for the incorporations, **T2.2**, only required up to two modified residues to be incorporated in a row. Both Vent (exo–) and *Pfu* (exo–) were able to produce full length products that were retarded in mobility compared to the unmodified control. However, the full extensions using Klenow (exo–) and Dpo4 gave rise to many truncation products and very little fully extended product. Based on the data in Table 2.1, it was a surprise that the incorporation reaction using Dpo4 was not very successful. In the presence of only **2.24**, five residues are incorporated (Table 2.1) whereas the inclusion of the other three natural nucleotides resulted in only 2–3 residues incorporated as illustrated in lane 4 of Figure 2.16.

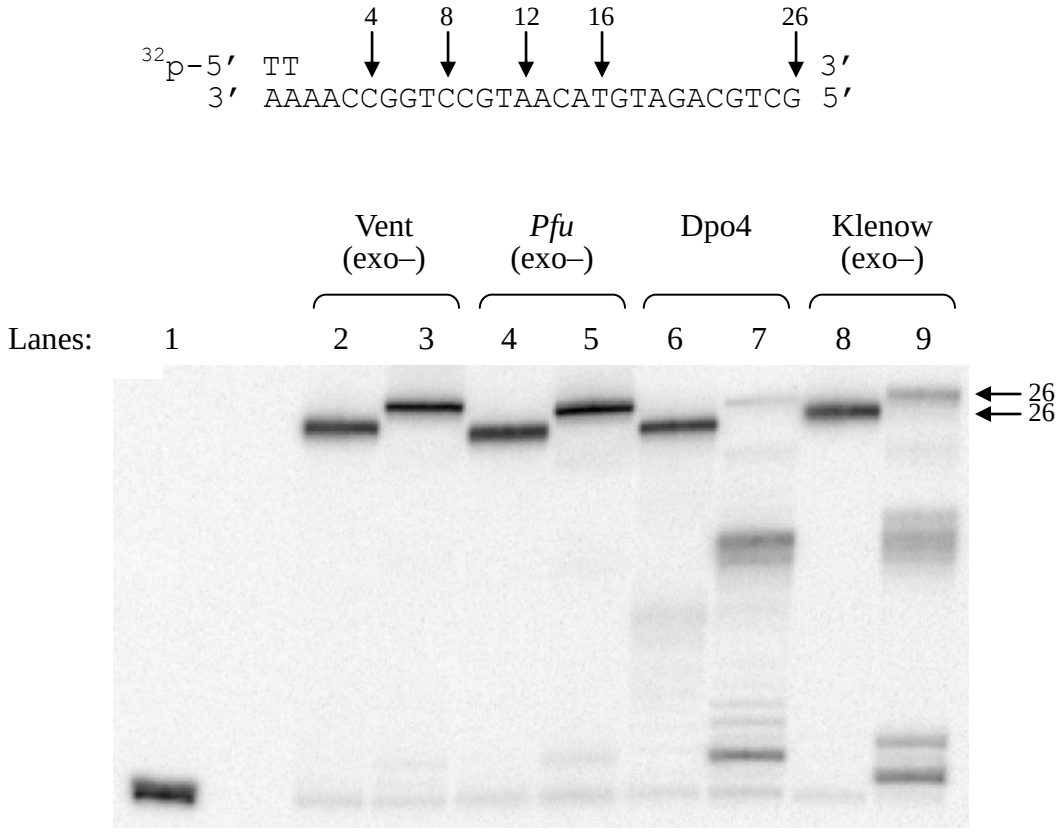


Figure 2.15 Incorporation under full extension conditions. Lane 1: primer only; lanes 2 and 3: Vent (exo–); lanes 4 and 5: *Pfu* (exo–); lanes 6 and 7: Dpo4; lanes 8 and 9: Klenow (exo–). Reactions were performed for 30 min using nucleoside triphosphates at 50 μM and template **T2.2**. Lanes 2, 4, 6, 8 used dTTP, dATP, dCTP and dGTP. Lanes 3, 5, 7, 9 used **2.24**, dATP, dCTP, and dGTP. Lanes 1–5 were incubated at 72 ℃ and lanes 6–9 were incubated at 37 ℃. Numbers to the right of the gel image correspond to the number of nucleotides incorporated.

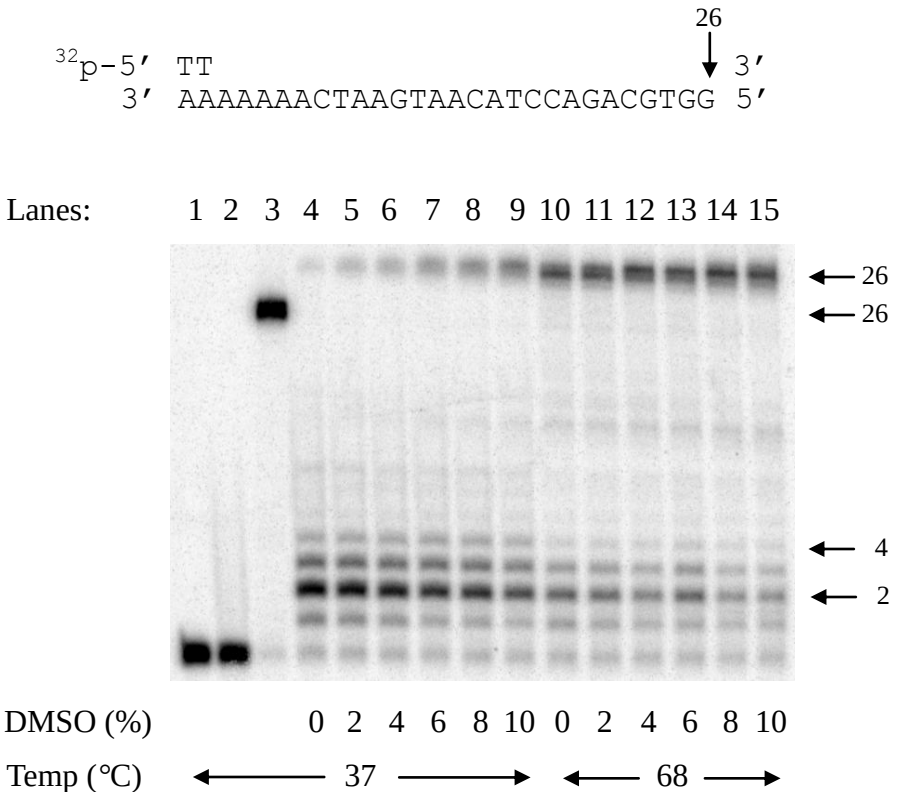


Figure 2.16 Dpo4 incorporation under various conditions. Full extension conditions employed 50 μM dNTP's, 6 ng/μL Dpo4, **T2.1** template and increasing amounts of DMSO (% by volume). Lane 1: primer; lane 2: primer and Dpo4; lane 3: dATP, dTTP, dGTP, dCTP; lanes 4–15: dATP, dGTP, dCTP, **2.24**. Lanes 1–9: reactions incubated at 37 °C for 2 h; lanes 10–15: reactions incubated at first 37 °C for 2 h and then at 68 °C for 30 min. Numbers to the right of the gel image correspond to the number of nucleotides incorporated.

To better understand the incorporation of **2.24** using Dpo4 under full extension conditions, optimization was performed to try to obtain a more substantial yield of fully extended product. When the polymerase amount was increased from 6 to 20 ng/μL or the $MgCl_2$ concentration was varied from 4 to 25 mM, no appreciable enhancement in incorporation was observed. On the other hand, the addition of DMSO and an increase in

incubation temperature both gave rise to improved yields of the expected product. As shown in Figure 2.16, the addition of up to 10 % DMSO and incubation at 68 ℃ rather than 37 ℃ together greatly enhanced the incorporation such that the band corresponding to fully extended product (lane 15) became the most intense band. However, the amount of fully extended modified DNA produced was still significantly less than that of the unmodified nucleotide incorporation in lane 3.

2.8.2 PCR with 2.24

The PCR of **2.24** was studied in the same manner as for **2.11** to produce the same double-stranded 120 base pair product. In Figure 2.17, the PCR results are shown for two different DNA polymerases. The first image (A) corresponds to the results obtained with *Pfu* (exo–). As can be seen, the correctly sized product was generated in both the unmodified and modified lanes. The modified lane gave about 35 % yield when compared to the control. The second image (B) corresponds to the PCR using Vent (exo–). This set of reactions also generated the desired product, and the yield was somewhat higher as well (55 % compared to unmodified control). However, two artifacts were observed on the agarose gel. The lane corresponding to the reaction where no template was used (lane 3) was found to produce a smear that was not visible in either of the two reactions that were performed in the presence of template. Also, the unmodified reaction generated, in addition to the desired product, small amounts of products that were in excess of 1 000 nucleotides in length. These long products were only found in the unmodified reaction and not the modified reaction. Both of the artifacts were found to be reproducible. When the extension time was decreased from 5 min to 20 s where no modified product was produced, the artifacts would not be present (data not shown). The

very large products were only observed with the 5 min reaction time because the long reaction time allowed the Vent (exo–) to replicate the whole plasmid template using natural dNTP's. During each cycle of the PCR, primers that were annealed onto the original template would produce product that was the same length as the template. As only the original circular plasmid template would give rise to the long products, each cycle of the PCR produced the same amount of very large product. In 20 cycles of amplification, the concentration of the plasmid would be multiplied by a factor of 21. Products that are not produced by exponential amplification are usually not visible. However, because the plasmid template was larger than the desired product by a factor of about 40, light bands of plasmid template product were observed. In the case of the reactions containing **2.24** and using Vent (exo–), the absence of the very large products was probably due to two reasons: 1) the incorporation of **2.24** is slow compared to dTTP and 2) the presence of incorporated residues of **2.24** slows down the incorporation of subsequent nucleotides.

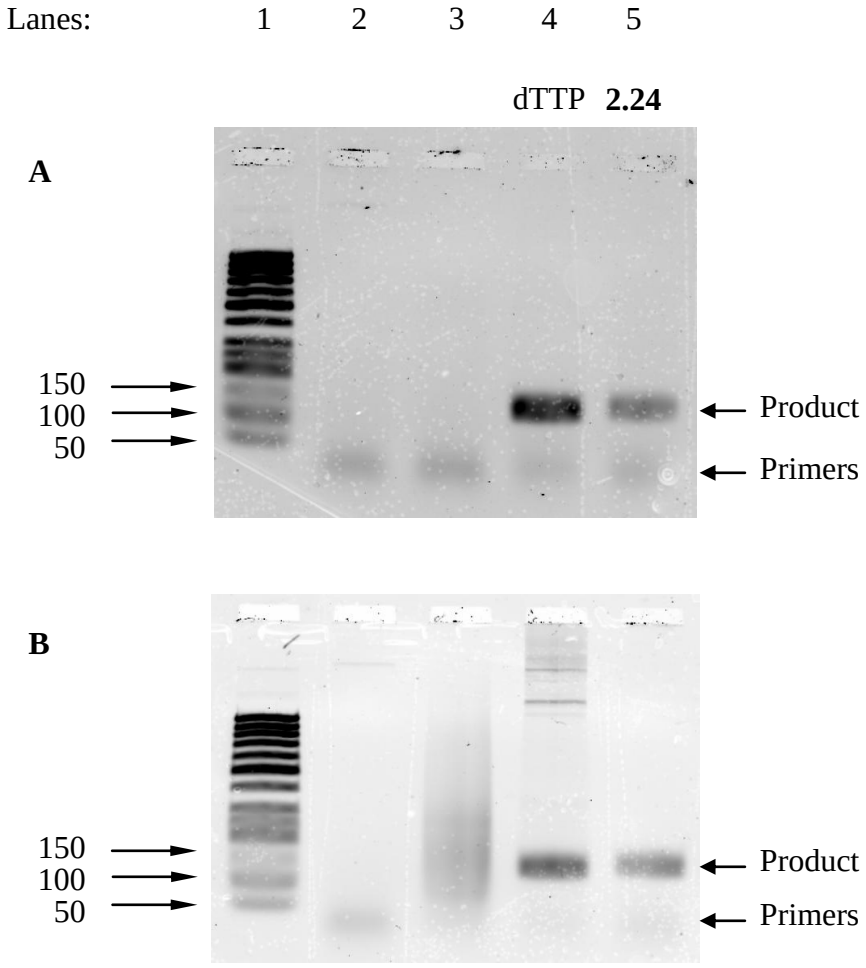


Figure 2.17 PCR with **2.24**. A. PCR with *Pfu* (exo–); B. PCR with Vent (exo–). Total reaction mixture (10 μL) consisted of 10 ng plasmid template, 0.5 μM of each primer, 0.2 mM of each dNTP and 1 unit of enzyme. Reactions were performed at 1X thermopol buffer with a final Mg^{+2} concentration of 7 mM. Lane 1: DNA marker (50–1000 bp); lane 2: no dNTP's; lane 3: no template; lane 4: all 4 natural dNTP's; lane 5: dATP, dGTP, dCTP, **2.24**.

2.9 Discussions and conclusions

Triphosphate **2.11** was shown to be an excellent substrate for DNA polymerases from both families A and B. The nucleotide was incorporated by every polymerase that was used including Sequenase V2.0, Klenow (exo–), *Pfu* (exo–), Vent (exo–) and *Taq* to

give rise to full length products. The incorporation of **2.11** using *Taq* was the only reaction where truncation products could be found. In lane 7 of Figure 2.11, the incorporation using *Taq* gave rise to two bands, a lighter band corresponding to desired fully extended material and a more intense but much shorter band. The identity of this shorter band can be deduced by comparison with the incorporations shown in Figure 2.9 that were also performed using template **T2.2** and **2.11**. In lane 5 of Figure 2.9, the gel shift of the product band (as referenced with unextended and fully extended material) corresponding to the incorporation of two modified nucleotides was approximately double the gel retardation that was observed for the shorter band in lane 7 of Figure 2.11. Based on the comparison of the two figures, the shorter band in lane 7 of Figure 2.11 should be due to the incorporation of one modified nucleotide. The template that was used in both Figures 2.9 and 2.11, **T2.2**, required two residues of **2.11** to be incorporated in a row in two different locations (Figure 2.18), one of which was at the very beginning of the incorporation. At the first location, *Taq* was able to incorporate the first modified nucleotide but not necessarily the second modified nucleotide. The second location that required two consecutive modified residues to be incorporated was not problematic as no other truncation products were observed. The different incorporation results at the two locations of **T2.2** that required two consecutive residues of **2.11** to be incorporated suggest that the incorporation of two consecutive residues of **2.11** by *Taq* is template-dependent.

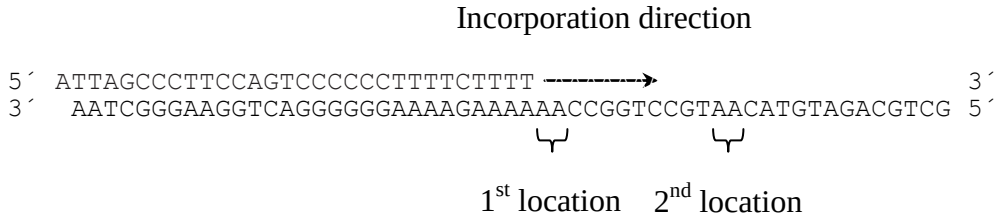


Figure 2.18 Primer extension using primer **P2.1** and template **T2.2**.

Triphosphate **2.24** was also found to be a substrate for DNA polymerases. Family B polymerases, Vent (exo–) and *Pfu* (exo–), could incorporate the modified residue efficiently to give full length products. Dpo4 and Klenow (exo–) were much less capable of using **2.24** as a substrate as the enzymes gave rise to little desired product. Both enzymes were found to have difficulties incorporating residues at the two locations of template **T2.2** that required two consecutive modified nucleotides to be incorporated (Figure 2.18). As mentioned in section 2.8.1, Dpo4 can effectively use **2.24** as a substrate only if none of the other three natural nucleotides is present. The presence of the other three natural nucleotides reduced the ability of Dpo4 to incorporate **2.24**. Of the conditions that were varied in an attempt to improve the use of **2.24** as a substrate for Dpo4 under full extension conditions, increasing the temperature from 37 ℃ to 65 ℃ was found to be somewhat effective. This improvement was probably due to the disruption of structures/complexes that are inhibitory to further incorporation. The incorporation improvement shown with the addition of DMSO in the reaction was probably acting in the same way as well.

The precursor triphosphate **2.1** was found to be an inferior substrate compared to **2.11** and **2.24**. None of the DNA polymerases tested could incorporate five consecutive

nucleotides of **2.1** (Table 2.1). Klenow (exo–) and Sequenase V2.0 were found to be the best polymerases for use with **2.1**, but only a maximum of three consecutive modified nucleotides could be incorporated. The closely related triphosphates **2.2** and **2.3,** which have cationic amines attached with a linker arm that is 3 atoms in length rather than 1, have been shown by others to be much better substrates for DNA polymerases than **2.1**.[84,101,103]

In this chapter, the synthesis of three modified 2′-deoxyuridine triphosphates and the enzymatic incorporation of two of them were discussed. The guanidinium-containing **2.11** was shown to be a very good substrate for DNA polymerases from both families A and B, and has subsequently been used in two *in vitro selections* for the metal-independent self-cleavage of DNA/RNA chimeric sequences. The first was discovered using a 20 nucleotide random region that gave rise to a DNAzyme, 9-86, with a self-cleavage rate constant of 0.17 min^{-1} (Figure 2.19).[138] The second was derived from a 40 nucleotide random region that gave rise to a DNAzyme with a self-cleavage rate constant of 0.50 min^{-1} (Figure 2.19).[137] This second DNAzyme, 10-66, was engineered into a *trans*-cleaving catalyst as well. The synthesis of the boronic acid-containing triphosphate **2.12** led to isolation of what turned out to be the phenol-containing **2.24**. Triphosphate **2.24** was found to be an excellent substrate for two DNA polymerases, both members of family B. Although **2.12** was eventually isolated as well, this nucleotide has not yet been used in incorporation assays.

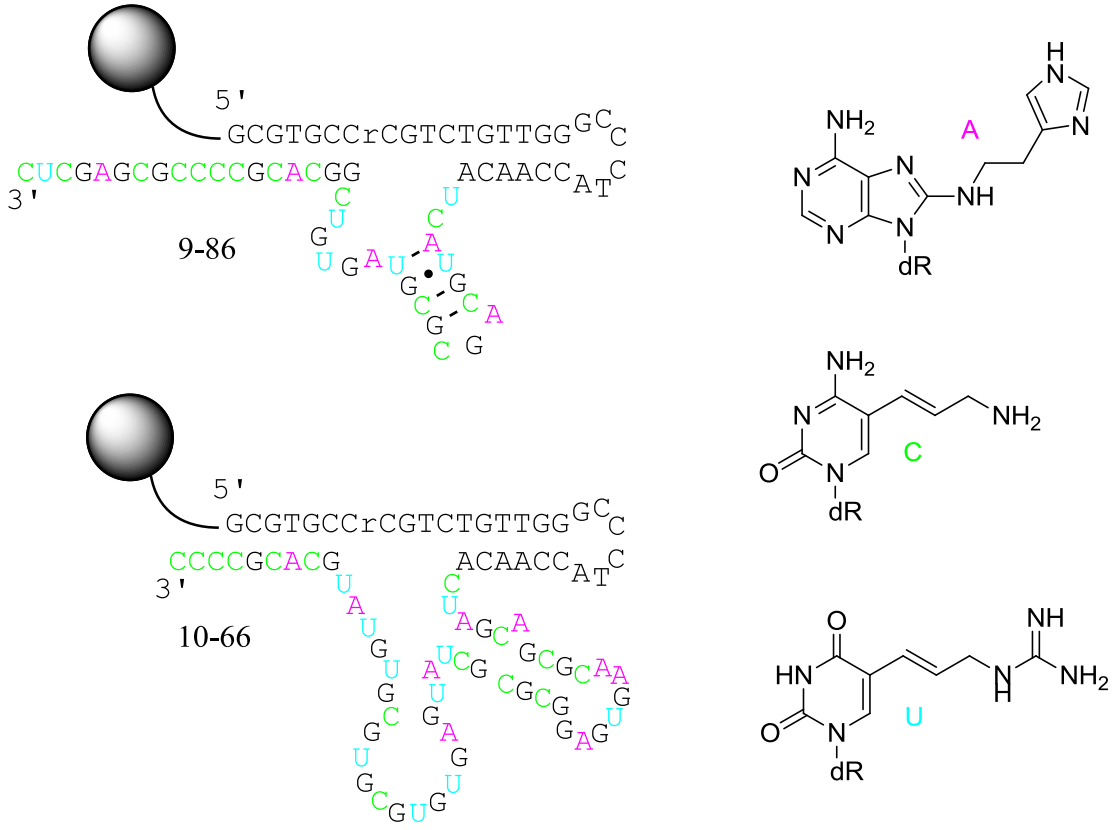


Figure 2.19 DNAzymes 9-86 and 10-66. The modified nucleotides 8-[2-(4-imidazolyl)ethyl]-2′-deoxyadenosine (A), 5-aminoallyl-2′-deoxycytidine (C) and 5-guanidinoallyl-2′-deoxyuridine (U) are shown in purple, green and blue, respectively. The location of cleavage following the cytidine (rC) residue is indicated by the arrow, dR represents 2′-deoxyribose and the ball represents biotin.

2.10 Experimental

All starting materials and solvents were purchased from either Sigma-Aldrich or Fisher Scientific. 5-Aminoallyl-2′-deoxyuridine triphosphate (**2.2**) in lithium salt form was purchased from TriLink Biotechnologies. Anhydrous solvents were obtained either by distillation or treatment with molecular sieves (4 Å). Unless otherwise indicated, all reactions were performed with flamed-dried glassware and nitrogen atmospheres. Thin layer chromatography (TLC) and preparative-TLC were performed on precoated glass-backed plates of silica gel 60 F_{254} from EMD Chemicals. Flash chromatography was carried out on Silicycle silica gel (230–400 mesh). Silica column dimensions are shown as the length multiplied by the diameter. NMR spectra were collected using either a Bruker AV-300, inverse AV-400 or direct AV-400 instrument. ^{1}H and ^{13}C NMR spectra were referenced to the signal of the solvent. ^{31}P NMR was externally referenced to phosphoric acid. ESI-MS were collected with a Bruker Esquire-LC, Micromass LCT or Waters LC/MS instrument in positive ion mode.

MALDI-TOF mass spectra were collected using a Bruker Biflex instrument in negative ion mode. A solution of 2,4,6-trihydroxyphenone (47 mg, 0.25 mmol) in ethanol (0.5 mL) and a solution of ammonium $citrate_{(aq)}$ (2.4 %) were mixed in a four to one ratio for the matrix. Prior to sample loading, cation exchange beads (Bio-Rad AG50W-X8 resin converted to the ammonium form) were added to each nucleoside triphosphate sample (2 μl, 100–500 μM), and the samples were left at room temperature for at least 15 min. Matrix solution (1 μL) was first applied to the MALDI target and allowed to dry. Triphosphate solution (1 μL) was added and allowed to dry as well. Care was taken to make sure that at least a few of the cation exchange beads were loaded. Finally, additional matrix solution was added (0.5 μL).

HPLC purification was performed on an Agilent 1100 system using a Phenomenex Jupiter 10μ C4 300A column at a flow rate 1 mL/min. HPLC solvents were prepared with one of two buffer systems: 50 mM triethylammonium bicarbonate (pH 7.5) or 50 mM triethylammonium acetate (pH 7.0). The length of time for each injection run was 30 min. The purification solvent gradient was applied during the first 18 min of the injection run, and the other 12 min were used for the rinsing of the column in preparation for the next injection.

Time (min)	% MeCN/H_2O
0	0
10	1
18	25
19	50
24	50
25	0
30	0

Table 2.2 HPLC gradient system.

All incorporations were performed using autoclaved materials. Water was treated with diethyl pyrocarbonate (1 μL/10 mL) prior to autoclaving. Natural dNTP's and Klenow (exo–), were purchased from Fermentas Life Sciences. T4 polynucleotide kinase, *Bst* DNA polymerase large fragment, Vent (exo–) and *Taq* DNA polymerase were purchased from New England Biolabs. Sequenase version 2.0 T7 DNA polymerase, inorganic pyrophosphatase and single-stranded DNA binding protein (SSB) were purchased from GE Healthcare. DNA Polymerase IV was purchased from Trevigen.

^{32}P-γ-ATP was purchased from Perkin Elmer. Radioactivity and fluorescence was visualized using a Typhoon 9200 variable mode imager from GE Healthcare.

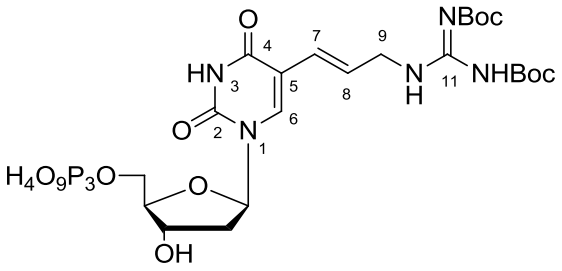
2.10.1 Synthesis

_N,N′_-Di-Boc-guanidine (2.13)

Synthesized according to Feichtinger *et al.*[160] R_f = 0.63 (3:2 EtOAc/hexanes). ^{1}H NMR (400 MHz, $CDCl_3$, 25 °C): δ = 1.49 (s, 18H). MS (ESI^+) 260.3 $[M+H]^+$.

_N,N′_-Di-Boc-_N′′_-trifluoromethanesulfonylguanidine (2.14)

Synthesized according to Feichtinger *et al.*[160] R_f = 0.71 (3:7 EtOAc/hexanes). ^{1}H NMR (300 MHz, DMSO, 25 °C): δ = 11.05 (s, br, 2H), 1.44 (s, 18H). MS (ESI^+) 414.2 $[M+Na]^+$.

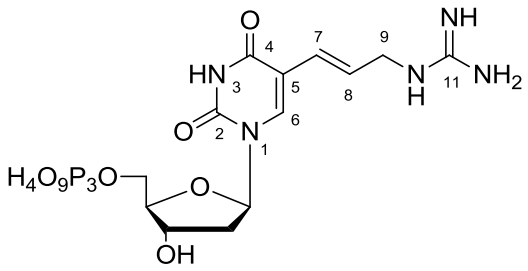


$C_{23}H_{38}N_5O_{18}P_3$
Exact Mass: 765.142
Mol. Wt.: 765.49

5-(3-_N,N′_-Diboc-guanidinoprop-1-enyl)-2′-deoxyuridine triphosphate (2.15)

The commercially available lithium salt of **2.2** (2 μmol) was suspended with AG 50W-X8 Biorad cation exchange beads in triethylammonium form (20 mg) in a buffered solution of triethylammonium bicarbonate (200 μL, 1 M). The solution of **2.2** was filtered and evaporated to give rise to the triethylammonium salt of **2.2**. The triethylammonium salt of **2.2** (0.5 μmol) was then suspended in a mixture of H_2O (1.7 μL) and dioxane (5.1 μL).

Compound **2.14** (0.6 mg, 1.5 μmol) was added to this solution, and the reaction was thoroughly agitated by a vortexer until everything was fully dissolved. After 5 min, triethylamine (0.35 μL, 2.5 μmol) was added to the reaction. The reaction was left at room temperature and agitated occasionally. After standing at room temperature overnight, the reaction was dried down and resuspended in H_2O. The extinction coefficient of **2.2** (7100 $M^{-1}cm^{-1}$ at 290 nm) was used to quantify the product. Unreacted triphosphate **2.2** (0.125 μmol, 25%) and triphosphate **2.15** (0.375 μmol, 75%) were isolated by preparative-TLC. The desired product **2.15** was found to be pure by TLC; R_f = 0.42 (6:4:1 dioxane/H_2O/NH_4OH). MS (MALDI$^-$): 764.6 $(M–1)^-$. λ_{max} = 235, 289 nm.

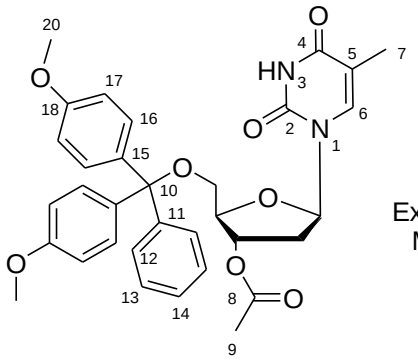


$C_{13}H_{22}N_5O_{14}P_3$
Exact Mass: 565.03
Mol. Wt.: 565.26

5-(3-Guanidinoprop-1-enyl)-2′-deoxyuridine triphosphate (2.11)

Trifluoroacetic acid (17 μL) was added to triphosphate **2.15** (0.375 μmol), and the resulting mixture was agitated by a vortexer for 2 min. Et_2O (170 μL) at –78 ℃ was added to the reaction to terminate it. The mixture was then centrifuged, and the supernatant was removed. The pellet was washed with 3 portions of Et_2O (85 μL) and dried under vacuum. TLC showed that the reaction was not quite complete. The crude product was dissolved in H_2O and purified by preparative-TLC. Triphosphate **2.11** was further purified on HPLC (T_r = 4.7 min) to give 85 nmol (23%) of product. HPLC was

performed using a 50 mM triethylammonium bicarbonate (pH 7.5) acetonitrile/H_2O system. The extinction coefficient of **2.2** (7100 $M^{-1}cm^{-1}$ at 290 nm) was used to quantify the product. R_f = 0.19 (6:4:1 dioxane/H_2O/NH_4OH). MS (MALDI$^-$): 564.0 $(M–1)^-$. λ_{max} = 239, 289 nm.

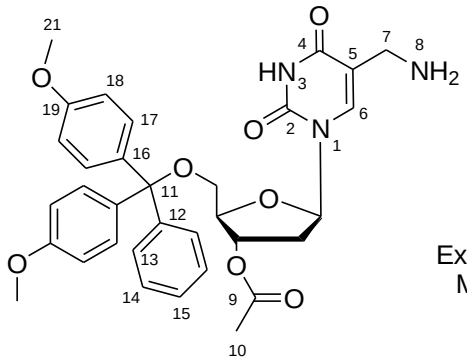


$C_{33}H_{34}N_2O_8$
Exact Mass: 586.2315
Mol. Wt.: 586.6317

5′-*O*-(4,4′-Dimethoxytrityl)-3′-*O*-acetylthymidine (2.16)

Thymidine (3.00 g, 12.4 mmol) was dried by co-evaporation from pyridine. The resulting solid was suspended in pyridine (12 mL) and added a solution of 4,4′-dimethoxytrityl chloride (5.46 g, 16.1 mmol) in pyridine (12 mL) at room temperature. After stirring the reaction at room temperature for 2 hr, acetic anhydride (3.50 mL, 37 mmol) was added. The reaction was stirred at room temperature overnight. The reaction was concentrated to a gum and resuspended in methylene chloride. This solution was washed with $NaHCO_{3(aq)}$ (5%) and dried with sodium sulfate (anhydrous). The solution was concentrated and purified by silica flash chromatography (14.5×5.0 cm column). The product was eluted with EtOAc/hexanes (1:1) to give 7.01 g (97%) of a white foam. R_f = 0.56 (1:9 MeOH/$CHCl_3$). ^{1}H NMR (400 MHz, CD_2Cl_2, 25 °C): δ = 8.43 (s, 1H, H3), 7.53 (s, 1H, H6), 7.41 (d, J = 8.0 Hz, 2H, H12), 7.34–7.23 (m, 7H, H13, H14, H16), 6.85 (d, J = 8.8 Hz, 4H, H17), 6.40–6.35 (m, 1H, H1′), 5.44–5.41 (m, 1H, H3′), 4.13–4.11 (m, 1H, H4′),

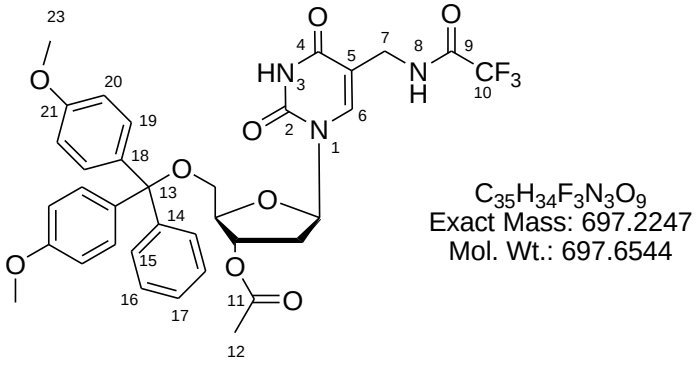
3.78 (s, 6H, H20), 3.46–3.37 (m, 2H, H5′), 2.48–2.36 (m, 2H, H2′), 2.06 (s, 3H, H9), 1.40 (s, 3H, H7). ^{13}C NMR (100 MHz, CD_2Cl_2, 25 ℃): δ = 170.7 (C8), 164.2 (C4), 159.2 (C18), 150.9 (C2), 144.9 (C11), 135.8, 135.7 (C15), 135.6 (C6), 128.4 (C12), (C13), 127.5 (C14), 113.6 (C17), 111.7 (C5), 87.4 (C10), 84.7 (C1′), 84.4 (C4′), 75.4 (C3′), 64.1 (C5′), 55.6 (C20), 38.2 (C2′), 21.1 (C9), 11.8 (C7). HRMS (ESI$^+$) calcd. for $C_{33}H_{35}N_2O_8Na$: 609.2213; found 609.2216.



5′-*O*-(4,4′-Dimethoxytrityl)-3′-*O*-acetyl-5-(aminomethyl)-2′-deoxyuridine (2.17)

N-Bromosuccinimide (2.17 g, 12.2 mmol) and 2,2′-azobis(2-methylpropionitrile) (115 mg, 0.7 mmol) were added to a solution of **2.16** (5.87 g, 10 mmol) in benzene (200 mL) at 70 ℃. After 30 min of stirring at 70 ℃, additional *N*-bromosuccinimide (1.08 g, 6.1 mmol) and 2,2′-azobis(2-methylpropionitrile) (58 mg, 0.35 mmol) were added to ensure full conversion by monitoring with TLC. After an additional 25 min of stirring at the same temperature, the solvent was removed. The resulting red solid was resuspended in methylene chloride (100 mL). Ammonia gas was bubbled into this solution at room temperature for 1 hr. The reaction was then concentrated *in vacuo*, redissolved in methylene chloride, washed with aq. $NaHCO_3$ (5 %), dried with Na_2SO_4 (anhydrous) and

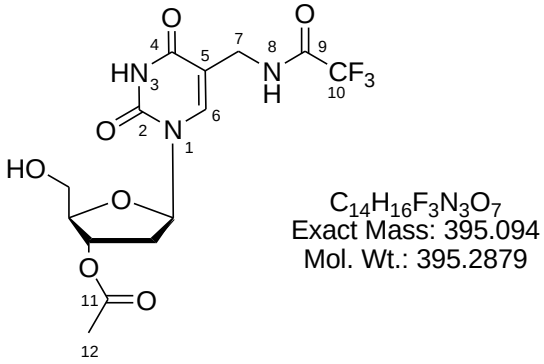
purified on flash chromatography (12.5×5.0 cm column). The product was eluted with MeOH/$CHCl_3$ (1:19) to give 2.17 g (36 %) of an orange solid. Unreacted **2.16** (0.57 g) was recovered from the column as well. R_f = 0.25 (1:9 MeOH/$CHCl_3$). ^{1}H NMR (400 MHz, CD_2Cl_2, 25 ℃): δ = 7.62 (s, 1H, H6), 7.42 (d, J = 7.6 Hz, 2H, H13), 7.34–7.23 (m, 7H, H14, H15, H17), 6.86 (d, J = 8.9 Hz, 4H, H18), 6.38 (dd, J = 8.4, 6.0 Hz, 1H, H1′), 5.43–5.40 (m, 1H, H3′), 4.15–4.12 (m, 1H, H4′), 3.78 (s, 6H, H21), 3.44–3.42 (m, 2H, H5′), 3.07–2.98 (m, 2H, H7), 2.50–2.37 (m, 2H, H2′), 2.06 (s, 3H, H10). ^{13}C NMR (100 MHz, CD_2Cl_2, 25 ℃): δ = 170.7 (C9), 163.6 (C4), 159.2 (C19), 150.7 (C2), 144.9 (C12), 136.2 (C6), 135.8, 135.7 (C16), 130.5 (C17), 128.5 (C14), 128.3 (C13), 127.5 (C15), 116.7 (C5), 113.6 (C18), 87.3 (C11), 84.9 (C1′), 84.4 (C4′), 75.3 (C3′), 64.1 (C5′), 55.6 (C21), 39.3 (C7), 38.1 (C2′), 21.1 (C10). HRMS (ESI$^+$) calcd. for $C_{33}H_{36}N_3O_8$: 602.2502; found 602.2515.



5′-*O*-(4,4′-Dimethoxytrityl)-3′-*O*-acetyl-5-[*N*-(trifluoroacetyl)aminomethyl]-2′-deoxy uridine (2.18)

Ethyl trifluoroacetate (2.14 mL, 18 mmol) was added to a solution of **2.17** (1.08 g, 1.8 mmol) in triethylamine (2.5 mL, 18 mmol) and methylene chloride (25 mL) at room

temperature. After 3.5 hr of stirring at room temperature, the reaction was concentrated *in vacuo* and purified on silica gel (12.5×2.5 cm column). The product was eluted with EtOAc/hexanes (2:3) to give 0.93 g (74 %) of a white foam. R_f = 0.5 (1:9 MeOH/$CHCl_3$). ^{1}H NMR (300 MHz, $CDCl_3$, 25 ℃): δ = 7.86 (s, 1H, H6), 7.39–7.21 (m, 9H, H15, H16, H17, H19), 6.84 (d, *J* = 9.0 Hz, 4H, H20), 6.31 (dd, *J* = 8.7, 5.7 Hz, 1H, H1′), 5.45–5.41 (m, 1H, H3′), 4.15–4.12 (m, 1H, H4′), 3.77 (s, 6H, H23), 3.61–3.56 (m, 2H, H5′), 3.50–3.38 (m, 2H, H7), 2.52 (ddd, *J* = 14.1, 5.7, 1.5 Hz, 1H, H2′), 2.38 (ddd, *J* = 14.1, 8.7, 6.0 Hz, 1H, H2′), 2.03 (s, 3H, H12). ^{13}C NMR (100 MHz, CD_2Cl_2, 25 ℃): δ = 170.6 (C11), 163.7 (C4), 159.2 (C21), 157.0 (C9), 150.4 (C2), 144.9 (C14), 139.6 (C6), 135.8, 135.6 (C18), 130.5, 130.4 (C19), 128.4 (C15), (C16), 127.5 (C17), 116 (C10), 113.8 (C20), 109.9 (C5), 87.5 (C13), 85.6 (C1′), 84.7 (C4′), 75.0 (C3′), 64.0 (C5′), 55.6 (C23), 38.6 (C2′), 37.3 (C7), 21.1 (C12). HRMS (ESI$^+$) calcd. for $C_{35}H_{34}F_3N_3O_9Na$: 720.2145; found 602.2515.



$C_{14}H_{16}F_3N_3O_7$
Exact Mass: 395.094
Mol. Wt.: 395.2879

3′-*O*-Acetyl-5-[*N*-(trifluoroacetyl)aminomethyl]-2′-deoxyuridine (2.19)

Compound **2.18** (21 mg, 30 μmol) was dissolved in 4:1 acetic acid/water (0.5 mL) and stirred at room temperature for 25 min. The reaction was concentrated *in vacuo* and

purified on silica gel (12.5×1.0 cm column). The product was eluted with EtOAc/hexanes (4:1) to give 11 mg (92 %) of a white solid. R_f = 0.31 (1:9 MeOH/$CHCl_3$). ^{1}H NMR (400 MHz, $CDCl_3$, 25 °C): δ = 9.24 (s, 1H, H3), 8.23 (s, 1H, H6), 7.76–7.67 (m, 1H, H8), 6.26 (dd, J = 7.8, 6.1 Hz, 1H, H1′), 5.39–5.34 (m, 1H, H3′), 4.23–4.13 (m, 3H, H7, H4′), 3.96 (dd, 1H, H5′), 3.89 (dd, 1H, H5′), 2.54–2.47 (m, 1H, H2′), 2.33–2.42 (m, 1H, H2′), 2.09 (s, 3H, H12). ^{13}C NMR (100 MHz, $CDCl_3$, 25 °C): δ = 170.6 (C11), 164.0 (C4), 158.1 (C9), 150.3 (C2), 141.2 (C6), 115.7 (C10), 109.3 (C5), 86.4 (C1′), 86.0 (C4′), 75.2 (C3′), 62.4 (H5′), 38.3 (H2′), 36.5 (C7), 20.9 (C12). HRMS (ESI^+) calcd. for $C_{14}H_{16}N_3O_7F_3Na$: 418.0838; found 418.0842.

5-Aminomethyl-2′-deoxyuridine triphosphate (2.1)

Nucleoside **2.19** (19.8 mg, 50 μmol) was dried by coevaporation with pyridine overnight in a NMR tube. The resulting gum was suspended in dioxane (345 μL) and pyridine (115 μL) and sealed with a septum. A solution of 2-chloro-4*H*-1,3,2-benzodioxaphosphorin-4-one (11.1 mg, 55 μmol) in dioxane (55 μL) is added. The NMR tube was mixed on a vortex mixer periodically over a period of 15 min. A well mixed solution of tri-*n*-butylammonium pyrophosphate (35.6 mg, 75 μmol) in tri-*n*-butylamine (50 μL) and DMF (150 μL) was then added. Following an additional 15 min of periodic mixing on a vortex mixer, I_2 (19.9 mg, 79 μmol) in H_2O (20 μL) and pyridine (980 μL) was added. The solution was mixed periodically over 20 min, and excess I_2 was quenched by the drop by drop addition of aqueous sodium bisulfite (5 %). The reaction was transferred into a 25 mL flask and evaporated to dryness. The resulting solid was then resuspended in H_2O (5 mL) and left to stand at room temperature for 30 min. Ammonium hydroxide (10 mL, 30 %) was added to the mixture. After stirring at

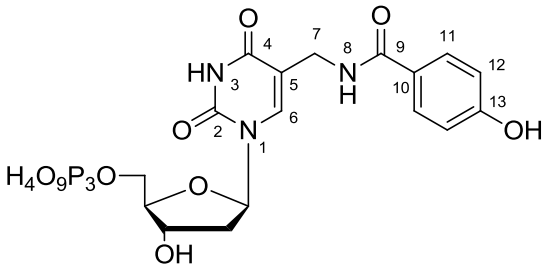
room temperature for an additional 3 h, the solution was dried down. The crude product was purified by preparative-TLC using an eluent of dioxane/H_2O/NH_4OH (6:4:1); R_f = 0.08. The synthesis and purification gave rise to 18.9 μmol (38 %) of material as quantified by the extinction coefficient of 8600 $cm^{-1}M^{-1}$ at 268 nm.[89] MS (MALDI$^-$): *m*/*z* = 495.6 $(M–1)^-$. λ_{max} = 265 nm.

4-Boronylbenzoyl-*N*-hydroxysuccinimide ester (2.22)

Synthesized according to Masuda *et al.*[165] Purification was performed with silica gel (20×1.5 cm column). The product was eluted with MeOH/$CHCl_3$ (1:9). 4-Carboxyphenylboronic acid (84 mg, 0.5 mmol) was converted to 100 mg of a white foam (76 %) R_f = 0.39 (1:9 MeOH/$CHCl_3$). ^{1}H NMR (300 MHz, CD_3OD, 25 °C): δ = 8.11 (d, *J* = 8.0 Hz, 2H), 7.90–7.78 (m, 2H), 2.92 (s, 4H).

4-Hydroxybenzoyl-*N*-hydroxysuccinimide ester (2.23)

Synthesized according to Guo *et al.*[169] R_f = 0.19 (3:7 EtOAc/hexanes). ^{1}H NMR (400 MHz, $CDCl_3$, 25 °C): δ = 10.84 (s, 1H), 7.94 (d, *J* = 8.8 Hz, 2H), 6.96 (d, *J* = 8.8 Hz, 2H), 2.86 (s, 4H).

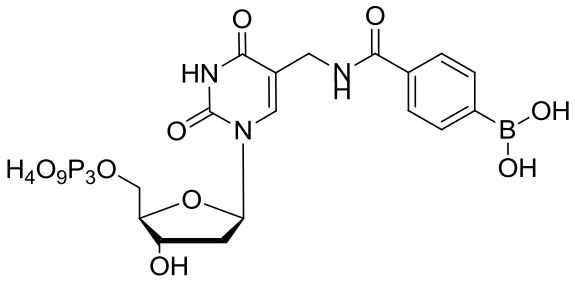


$C_{17}H_{22}N_3O_{16}P_3$
Exact Mass: 617.02
Mol. Wt.: 617.28

5-(4-Hydroxybenzoylaminomethyl)-2′-deoxyuridine triphosphate (2.24)

A solution of **2.23** (0.24 mg, 1.0 μmol) in DMF (5 μL) was added to triphosphate **2.1**

(0.25 μmol) in aqueous sodium bicarbonate (10 μL, 0.4 M). The reaction was left at room temperature with periodic agitation. After 3 h, the reaction was concentrated, resuspended with water and purified by preparative-TLC dioxane/H_2O/NH_4OH (6:4:1); R_f = 0.13. Triphosphate **2.24** was further purified on HPLC (retention time of 8.7 min) using a 50 mM triethylammonium acetate (pH 7.0) acetonitrile/H_2O system. Product (571 nmol, 57 % yield) was isolated as the tetrakis(triethylammonium) salt. ε_{258} = 17 900 $cm^{-1}M^{-1}$. ^1H NMR (400 MHz, D_2O, 25 °C, spectrum acquired with solvent suppression pulse program): δ = 7.92 (s, 1H), 7.81–7.75 (m, 2H), 7.06–6.99 (m, 2H), 6.35–6.29 (m, 1H), 4.75–4.60 (m, 1H), 4.42–4.12 (m, 4H), 2.47–2.39 (m, 2H) MS (MALDI$^-$): *m*/*z* = 615.9 $(M–1)^-$. λ_{max} = 258 nm.



$C_{17}H_{23}BN_3O_{17}P_3$
Exact Mass: 645.03
Mol. Wt.: 645.11

5-(4-Boronobenzoylaminomethyl)-2′-deoxyuridine triphosphate (2.12)

Triphosphate **2.1** (0.25 μmol) was suspended in aqueous sodium bicarbonate (10 μL, 0.4 M). Compound **2.22** (0.26 mg, 1.0 μmol) was dissolved in DMF (5 μL). The second solution was added to the first solution in 3 portions. At t = 0 min, 2.5 μL was initially added, and at two subsequent 30 min intervals, 1.25 μL was added. The reaction was left at room temperature for a total of 2 h with periodic agitation. Reaction was evaporated, resuspended in water and purified by TLC using a solvent system of 8:4:1 MeCN/H_2O/NH_4OH. R_f = 0.4 (6:4:1 MeCN/H_2O/NH_4OH). MS (MALDI$^-$): *m*/*z* = 608.0

(M–2H_2O adduct). λ_{max} = 240 nm. Half of the purified material dissolved in H_2O (5 μL), to which was added H_2O_2 (0.75 μL, 1 %). The reaction was left at room temperature for 2 h with occasional agitation with a vortexer. Material was fully converted to **2.24**. MS (MALDI⁻): *m*/*z* = 615.7. Further purification of **2.12** was attempted by HPLC (retention time of 8.3 min) using a 50 mM triethylammonium acetate (pH 7.0) acetonitrile/H_2O system. After the collected fractions were lyophilized, it was found that approximately (visualized by TLC) a third of the desired product **2.12** had been converted to **2.24**.

2.10.2 Polymerase assays

Solutions/buffers.

Loading solution: Formamide (27 mL), $EDTA_{(aq)}$ (3 mL, 0.5 M), xylene $cyanol_{(aq)}$ (300 μL, 0.05 %) and bromophenol $blue_{(aq)}$ (300 μL, 0.05 %).

Elution buffer: 1 % $LiClO_4$, 10 mM Tris pH 7.0 in water.

Labelling of P2.1 at the 5′ end: Primer **P2.1** (50 pmol), polynucleotide kinase (PNK) buffer (3 μL, 10X), ^{32}P-γ-ATP (25 μCi, 1 μL), PNK (1 μL) was diluted to a final volume of 30 μL. The reaction was incubated at 37 °C for 2 h and then terminated by heating at 65 °C for 20 min. Loading solution (30 μL) was added to the reaction, and the resulting mixture was purified by denaturing PAGE (10 %). Product was visualized by autoradiography and cut out. Gel material was broken up with a flame-sealed pipette tip and eluted with elution buffer at 65 °C. The liquids were evaporated and ethanol precipitated. Finally, the isolated solid was redissolved and loaded onto a G25 spin column.

DNA templates: (5′ to 3′, italicized residues correspond to the primer binding regions)

P2.1 ATTAGCCCTTCCAGTCCCCCCTTTTCTTTT.

T2.1

GGTGCAGACCTACAATGAATCAAAAA*AAAAGAAAAGGGGGGACTGGAAGGGCTAA*.

T2.2

GCTGCAGATGTACAATGCCTGGCCAA*AAAAGAAAAGGGGGGACTGGAAGGGCTAA*.

T2.3

GCAGCTGTAGATCTTAGCCAGGCCTT*AAAAGAAAAGGGGGGACTGGAAGGGCTAA*.

Primer extension protocol: Reactions were prepared in a final incorporation volume of 10 μL. Typically, a 5′-^{32}P-labelled primer (**P2.1**, ca. 1 pmol) was annealed with the appropriate template (ca. 1.5 pmol) in the presence of enzyme buffers. Single-stranded binding protein (SSB) (0.2 μL/reaction), pyrophosphatase (0.3 μL/reaction), and dithiothreitol (DTT) (0.5–1 μL/ reaction) were added to the annealed oligonucleotides as a cocktail. Experiments were incubated at 37 ℃ for Sequenase V2.0, Dpo4 and Klenow (exo–) or 68 ℃ for *Taq*, Vent (exo–), *Bst*, and *Pfu* (exo–). SSB and pyrophosphatase were omitted at elevated incubation temperatures. The enzymes were added last; 1–6 units of enzyme were used. In the case of Dpo4, 0.06 μg of enzyme was used. Following incubation, loading solution (30 μL) was added to each incorporation reaction and denaturing PAGE (13.3 %) was carried out.

PCR. Reactions (10 μL each) were carried out in 0.2 mM of each dNTP, 1 ng/μL of template (mutated *Taq* gene in an expression vector), 0.5 μM of each primer, 1 unit of enzyme, 1X thermopol buffer with a final $MgSO_4$ concentration of 7 mM. Reactions were performed on a Techne TC-312 thermocycler for 20 cycles of 94 ℃ for 20 s, 50 ℃ for 20 s and 72 ℃ for 5 min. Aliquots of the reactions were mixed with a solution consisting of 7 % sucrose, 0.001 % bromophenol blue and 0.001 % xylene cyanol and loaded onto 2 % agarose for analysis by gel electrophoresis. The gels were visualized by ethidium bromide staining. A 50 bp DNA ladder from Fermentas was used as a marker. The 2 primers used were 91A: 5′-TTCGGCGTCCCGCGGGAGGCCCTCCAGCCCCT-3′ and 92D: 5′-GTAAGGGATGGCTAGCTCCTGGGA-3′. The template used consisted of a mutated *Taq* gene in an expression vector from Christopher Hipolito. The region that was being amplified including the primer regions consisted of 120 bases: 5′-TTCGGCGTCCCGCGGGAGGCCCTCCAGCCCCTGATGCGCCGGGCGGCCAAG ACCATCAACTTCGGGGTCCTCTACGGCATGTCGGCCCACCGCCTCTCCCAGGA GCTAGCCATCCCTTAC-3′. Relative yields were quantified by fluorescence.

Chapter 3: Synthesis and enzymatic incorporation of modified derivatives of 8-histaminyl-dATP

3.1 Introduction

The imidazole of histidine plays a cherished role in biological catalysis. Due to the near neutral pK_a of its conjugate acid form, imidazoles are very often found in the active sites of enzymes where they facilitate general acid-base catalysis as well as chelate metals. As a result, several research teams have explored the use of imidazoles in the context of DNAzymes. The first attempt was the use of L-histidine as a cofactor for the cleavage of ribophosphodiester bonds.[170] The resulting *in vitro* selection-derived DNAzyme performed catalysis only in the presence of L-histidine or a very closely related analog. Subsequently, modified nucleotides containing appended imidazoles were synthesized and incorporated into DNA[69,83,92,96,99] for the *in vitro* selection of DNAzymes that could catalyze the self-cleavage of ribonucleotide phosphodiester bonds in a divalent metal-dependent[171] or independent manner.[69,70,137,138] As discussed in Chapter 1, one of the imidazole-containing modified nucleotides that has been successfully used for the *in vitro* selection of ribonucleotide phosphodiester bond cleavage in the absence of divalent metals was 8-histaminyl-dATP (**3.1**, Figure 3.1).[69]

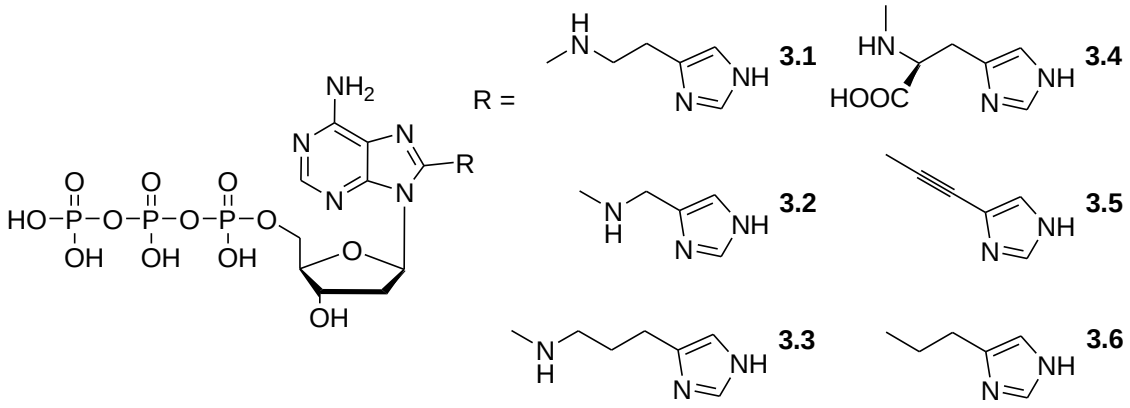


Figure 3.1 Imidazole-containing 8-modified 2′-deoxyadenosine triphosphates.

Thus far, the rate constants for the self-cleaving DNAzymes that have been isolated with **3.1** have been inferior to what have been observed for many divalent metal-dependent self-cleaving DNAzymes.[27,28] One hypothetical explanation for these observations is that the linker that connects the imidazole to 2′-deoxyadenosine is suboptimal. Due to the fact that single-stranded DNA is very flexible, a shorter or rigidified linker arm may allow potential DNAzymes to better align its substrate in a more ordered fashion. Also, the linker composition may be significant as well. Having a nitrogen atom at the 8-position of 2′-deoxyadenosine can contribute to Hoogsteen base pairing.[172,173] This interaction can play a significant role in the three dimensional structure of resulting DNAzymes.

A search of the literature will reveal that most modified nucleotides have linkers that are longer than what is found in **3.1**.[69,83,88,92,96,99] These nucleotides are generally good substrates for DNA polymerases under both primer extension and PCR conditions. If the linker is short, appended functional groups tend to interfere with the ability of the polymerase to bind to the growing strand such that incorporation of such modified

nucleotides becomes more difficult. Modified nucleotides that are poor substrates for DNA polymerases will result in modified nucleic acid libraries that are biased towards sequences with fewer modified nucleotides. Such modified nucleic acid libraries will exclude a significant portion of the possible sequences that are available for a selection. With regards to the above observations, it is important to note that there is no way of predicting what types of linkers are optimal for DNAzyme activity. Although modified nucleotides with short linkers tend to be poor substrates for DNA polymerases, modified nucleotides with short linkers cannot be ignored. For example, triphosphate **3.1** is not a particularly good substrate for DNA polymerases and cannot be used in PCR.[85] In spite of this, the modified residue was successfully used in several *in vitro* selections for ribonucleotide phosphodiester self-cleaving DNAzymes,[69,137,138] in which at least one imidazole has been shown to be directly involved in catalysis.[69,109] The successful usage of **3.1** in *in vitro* selections suggests that only reasonably efficient incorporation of modified nucleotides is necessary for DNAzymes of enhanced functionality.

In order to study the importance of the linker in 2′-deoxyadenosine 8-position modified nucleotides with the long term goal of determining whether shorter linkers provide superior catalysts, five analogs of **3.1** were synthesized containing linkers of different flexibilities, lengths and composition. Incorporation studies of these analogs were subsequently performed. The five linkages (Figure 3.1) that were synthesized for the first time and then studied in terms of polymerization are: aminomethyl (**3.2**), aminopropyl (**3.3**), (*S*)-1-amino-1-carboxyethyl (**3.4**), ethynyl (**3.5**), and ethyl (**3.6**). Triphosphates **3.2** and **3.3** will assess the effect of linker length on DNAzyme catalytic rates. Triphosphate **3.4** utilizes a linker of the same length as **3.1** but contains a carboxylate group. Therefore, the modification represents a histidinyl appendage. As the

carboxylate group can chelate metal cations or directly participate in catalysis, it was hoped that **3.4** would provide two important functionalities. The last two triphosphates, **3.5** and **3.6,** lack the amine in the linker and were chosen to evaluate the importance of hydrogen bonding at that position. Both **3.5** and **3.6** have linkers that are the same length as **3.2**. Compared to **3.2**, **3.5** is more rigid, and **3.6** is more flexible.

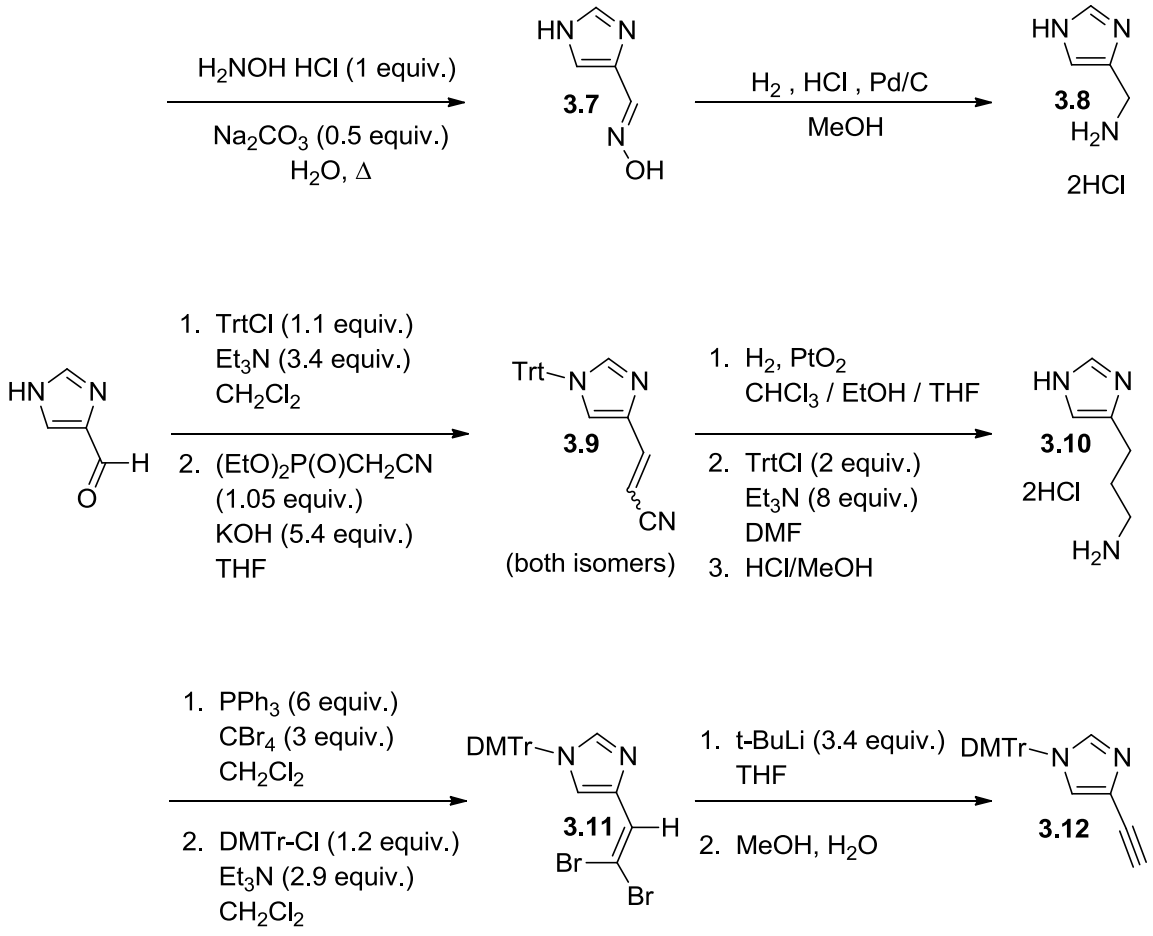
3.2 Synthesis

The synthesis of the modified dATP′s began with the synthesis of imidazoles with the different linkers. These linker-containing imidazoles were then attached to the 8-position of 2′-deoxyadenosine either by nucleophilic aromatic substitution or Sonagashira coupling,[174] and finally, the nucleosides were triphosphorylated.

3.2.1 Synthesis of imidazole precursors

Three imidazole precursors were synthesized from the relatively inexpensive 4-imidazolecarboxaldehyde as shown in Scheme 3.1. 4-Aminomethylimidazole dihydrochloride (**3.8**) was prepared according to literature. Sodium carbonate was added to a solution of the aldehyde and hydroxylamine hydrochloride, and the resulting mixture was refluxed to give 4-imidazolecarboxaldehyde oxime (**3.7**).[175] This oxime was then hydrogenated in methanolic hydrogen chloride.[176] The synthesis of 4-(3-aminopropyl)imidazole dihydrochloride (**3.10**) was also performed as published but with a modified purification protocol. 4-Imidazolecarboxaldehyde was tritylated, and the resulting product was then reacted with diethyl cyanomethylphosphonate in the presence of potassium hydroxide.[177] The resulting nitrile isomers (separable by silica flash chromatography) (**3.9**) were hydrogenated with platinum oxide and chloroform.[178] After

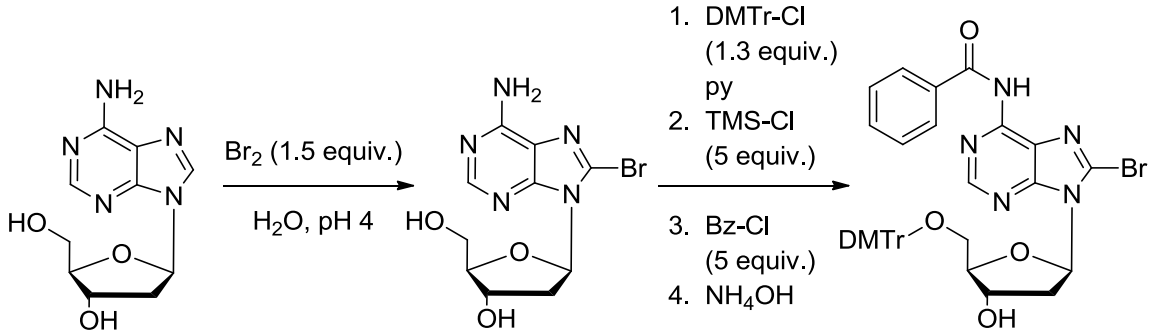
the free alkylamine of the crude product was tritylated, the resulting bistritylated species was deprotected in hydrochloric acid. The synthesis of [1-(4,4′-dimethoxytrityl)-1*H*-4-imidazolyl]ethyne (**3.12**) was adapted from the synthesis of the trityl protected version[179] that employed the Corey-Fuchs reaction.[180] A mixture of carbon tetrabromide and triphenylphosphine was reacted with 4-imidazolecarboxaldehyde to give a dibromoalkene (**3.11**). Following dimethoxytritylation, the alkene was converted to the alkyne in the presence of *tert*-butyl lithium.



Scheme 3.1 Imidazole precursor synthesis.

3.2.2 Synthesis of modified nucleosides

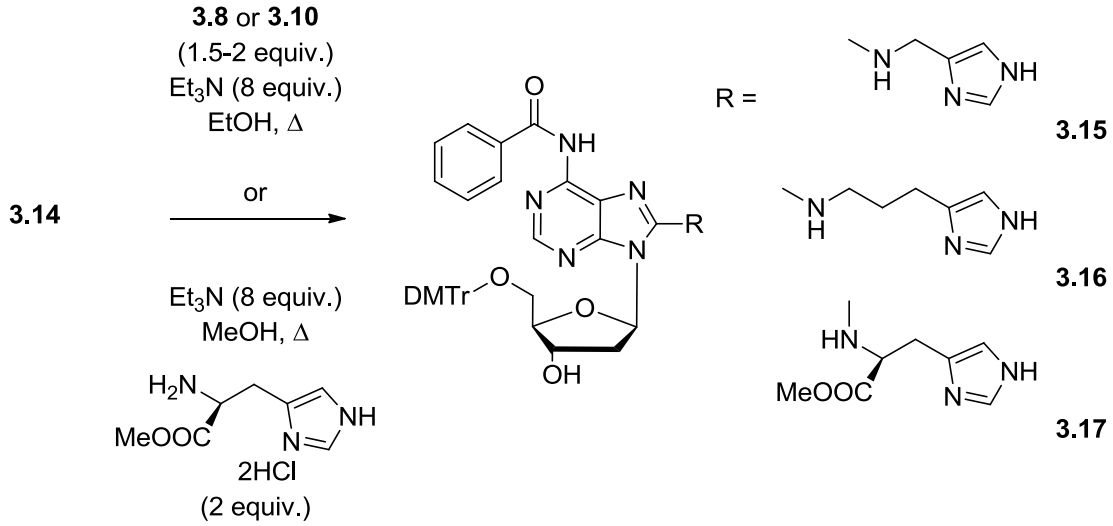
All of the modified nucleosides were derived from N^6-benzoyl-8-bromo-5′-*O*-(4,4′-dimethoxytrityl)-2′-deoxyadenosine (**3.14**). This fully protected nucleoside was prepared following established procedures (Scheme 3.2). 2′-Deoxyadenosine was first brominated in bromine water.[162] The 8-bromo-2′-deoxyadenosine (**3.13**) was then reacted sequentially with DMTr-Cl, trimethylsilyl (TMS) chloride, and finally benzoyl chloride (Bz-Cl). The DMTr-Cl protects the 5′-hydroxyl group, the TMS-Cl protects the 3′-hydroxyl group, and the benzoyl chloride (2 equivalents) protects the exocyclic amine on the nucleobase. Ammonium hydroxide deprotection removes the TMS group as well as one of the two benzoyl groups attached to the exocyclic amine to give the desired precursor.[181,182]



Scheme 3.2 Synthesis of the protected 2′-deoxyadenosine precursor.

The 2′-deoxyadenosines containing imidazolylalkylamines at the 8-position were synthesized by nucleophilic aromatic substitution (Scheme 3.3).[107,183] 4-Aminomethylimidazole dihydrochloride (**3.8**) and 4-(3-aminopropyl)imidazole

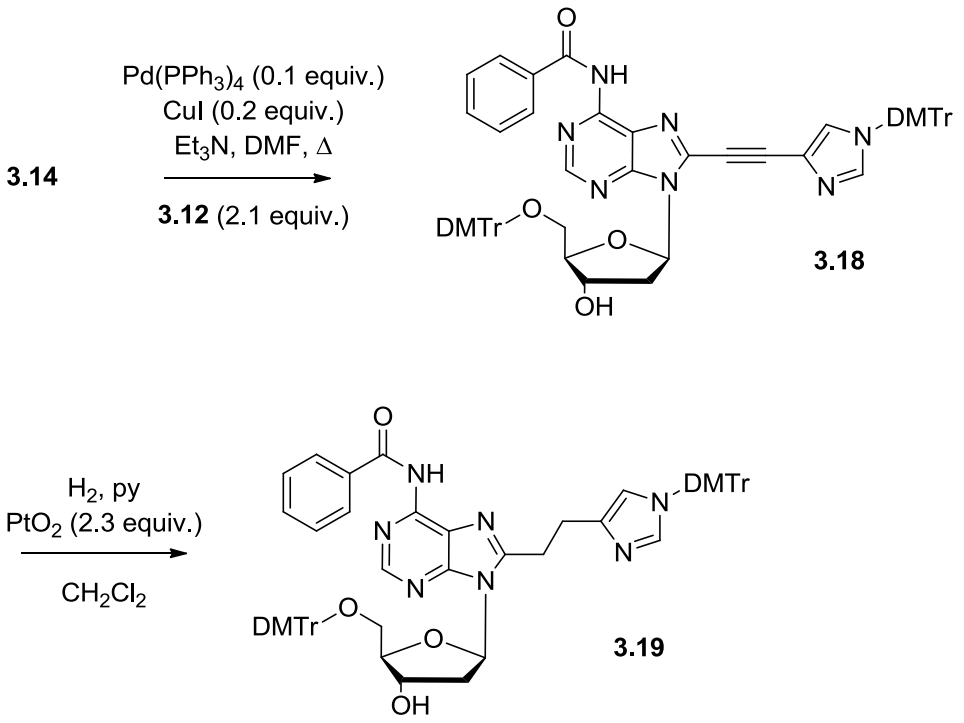
dihydrochloride (**3.10**) were heated with nucleoside **3.14** in the presence of triethylamine and ethanol to give **3.15** and **3.16**, respectively. Both of these nucleosides were readily prepared using this protocol. The displacement of the bromide with histidine methyl ester under the same conditions, however, was found to be quite slow. This was probably due to the extra steric bulk of the carboxylate methyl ester. The reaction was instead performed in methanol, to prevent transesterification, with prolonged heating to give the desired product **3.17** in low yield. Several byproducts were visible on TLC. The most significant byproduct was isolated and found to be debenzoylated **3.14**. Debenzoylation was due to the basic nature of the conditions used in Scheme 3.3 and prolonged heating. The debenzoylation of the product was observed as well, and hence it was often difficult to decide when to terminate the reaction.



Scheme 3.3 Nucleophilic aromatic substitution of **3.14**.

The synthesis of the 2′-deoxyadenosines with carbon atoms at the 8-position

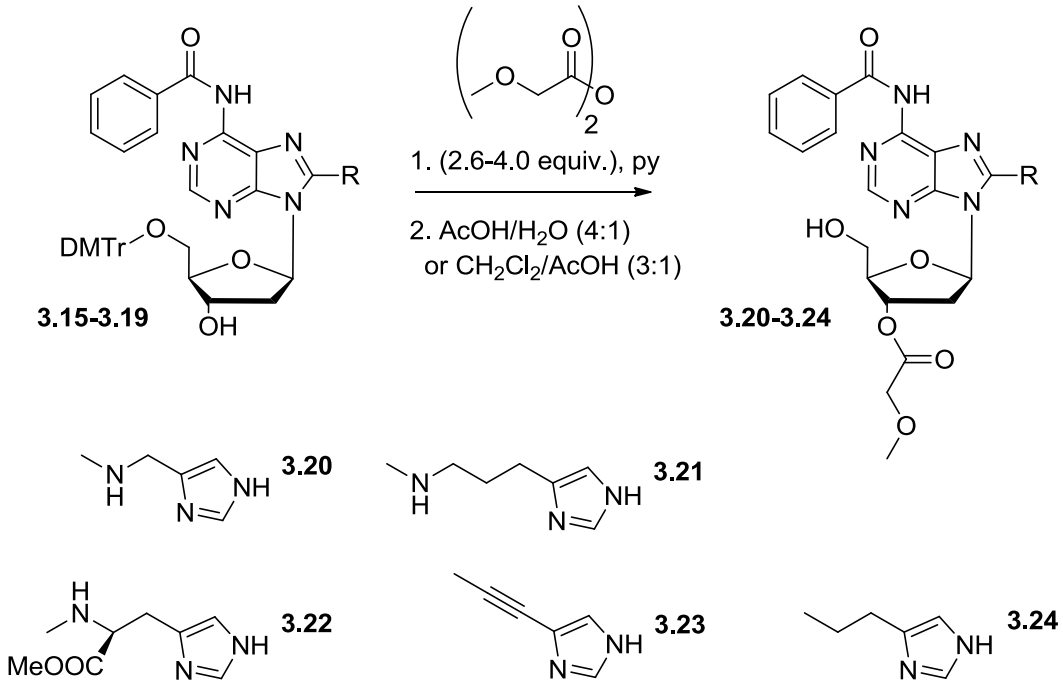
began with a Sonagashira coupling (Scheme 3.4).[184] Initial attempts to perform this reaction with 4-ethynylimidazole were unsuccessful. When the same conditions were used on the same imidazole prior to deprotection (**3.12**), a very good yield of the desired nucleoside (**3.18**) was obtained. Also, no additional deprotection step was necessary because the DMTr group on the imidazole was fully compatible with our synthesis. The ethyl linkage (**3.19**) was derived from the ethynyl linkage by catalytic hydrogenation of **3.18** in the presence of platinum oxide, pyridine and methylene chloride. The ratio of pyridine to methylene chloride was observed to be very important. A ratio of 1:60 was found to work very well. When this ratio was decreased, considerable deprotection of the DMTr group was observed, and when this ratio was increased, the reaction was slow. In an effort to obtain the *cis*-alkene linkage as well, **3.18** was separately hydrogenated with Lindlar's catalyst and palladium over barium sulfate. These conditions were unsuccessful and led to a combination of unreacted starting material and/or complete reduction to **3.19**.



Scheme 3.4 Synthesis of the ethynyl and ethyl linkage-containing 2′-deoxyadenosines.

3.2.3 Triphosphorylation

Prior to triphosphorylation, the 3′-hydroxyl groups had to be masked with methoxyacetyl groups, and the DMTr groups had to be removed (Scheme 3.5).[162] Reaction of the modified 2′-deoxyadenosines with methoxyacetic anhydride and pyridine followed by deprotection in acetic acid/water (4:1) gave the desired triphosphate precursors. This protocol was effective for all the modified nucleosides except for **3.18**, in which case a poor yield for the deprotection was observed. More mild conditions consisting of a mixture of methylene chloride/acetic acid (3:1) were found to give much higher yields.



Scheme 3.5 Triphosphate precursor preparation.

Following the removal of the dimethoxytrityl group, the modified nucleosides were converted to the triphosphates using the same multistep "one-pot" synthesis that was employed in Chapter 2.[163] Nucleosides **3.20**, **3.21**, **3.23** and **3.24** were all successfully converted to triphosphates **3.2**, **3.3**, **3.5** and **3.6** in this way, respectively. The isolation of **3.4** required an initial deprotection with lithium carbonate to convert the methyl ester to the free acid prior to ammonia deprotection, which otherwise would have given an amide. Purification of the products was performed first on preparative TLC and then on HPLC.

The triphosphates were isolated in low yields of 3-23 %. According to ^{31}P NMR that was used to monitor the multistep reactions, no irregularities could be observed. A possible reason for low yields would be incomplete removal of the benzoyl protecting group. Because the triphosphates were isolated in sufficient quantities for enzymatic

incorporation and combinatorial selection, the triphosphorylation reactions were only performed once or twice with little optimization.

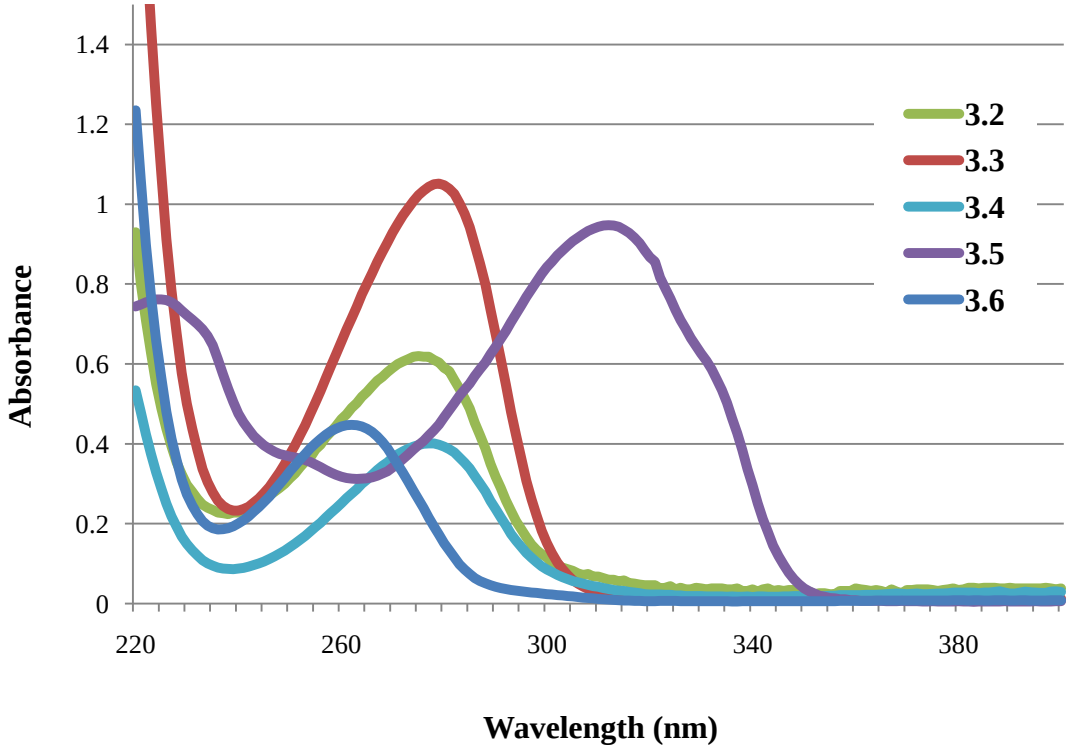


Figure 3.2 Ultraviolet absorbance of the modified dATP's. The spectra were not normalized for concentration.

The ultraviolet absorbance of the modified dATP's is shown in Figure 3.2. The attachment of the modifications to the 8-position of dATP shifted the absorbance maximum of 259 nm to longer wavelengths. The absorbance maximum was shifted to 262 for the alkyl linkage and a range of 276–279 for the 8-amino linkages. The spectrum of **3.5** differed dramatically from the others. Not only was the absorbance maximum shifted to 312 nm, but the shape of the curve was also not smooth like the other spectra.

These differences were probably due to the fact that the unsaturation of the linker arm resulted in a fully conjugated system between the imidazole and the adenine nucleobase.

3.3 Enzymatic incorporation of modified dATP's

Incorporation studies were performed for the new modified dATP's in Figure 3.1 as well as triphosphate **3.1**. The latter was included to use it as a control and to study its incorporation in more detail. Most of the incorporation studies did not include the histidine-containing 2′-deoxyadenosine triphosphate (**3.4**) as this modified residue was found to photodecompose during purification. Primer extensions were performed with a 30 nucleotide primer (**P2.1**) and four 55 nucleotide templates.

The four templates that were used are shown in Figure 3.3. Templates **T3.1** and **T3.3** require up to two modified residues to be incorporated successively and a total of eight and seven modified dATP's to be incorporated, respectively. Template **T3.1** is the relatively more challenging template because it requires one more modified nucleotide to be incorporated, and the locations of the required modified residues span a smaller region of the template. Both **T3.1** and **T3.3** were used to test the synthesis of fully extended products with the modified dATP's. Templates **T3.2** and **T3.4** require five and ten modified dATP's to be incorporated in a row, respectively. These two templates were mainly used in the absence of natural nucleotides dC, dG and dT to gauge the maximum number of modified dATP's that could be incorporated in a row.

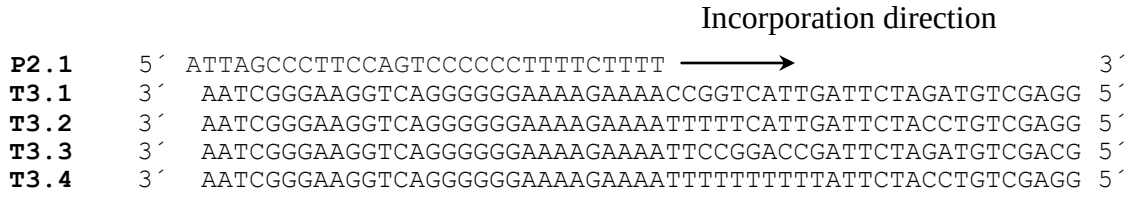


Figure 3.3 Primer and template sequences used for enzymatic incorporation in this chapter.

3.3.1 Incorporations with Sequenase V2.0 DNA polymerase

Since the enzyme Sequenase V2.0 DNA polymerase has been shown to incorporate **3.1**, this enzyme would very likely be able to incorporate the very similarly constructed modified dATP's that have been made. Figure 3.4 shows the results of incorporation using template **T3.1**. Lanes 1–6 correspond to control reactions. Based on the template **T3.1**, lanes 3–5 should have given rise to product bands that correspond to the incorporation of 4, 5 and 4 nucleotides, respectively. Instead, products of increased lengths, corresponding to the incorporation of mismatched nucleotides, were found in each of these three control lanes. These elongated products were probably due to a combination of the use of pyrophosphatase, which cleaves the inorganic pyrophosphate byproduct of nucleotide incorporation, and a non-proofreading DNAzyme. Product bands similar to what is shown in lanes 3–5 have been observed by Jäger *et al.* for modified nucleotide incorporations that were also performed in the presence of pyrophosphatase and a non-proofreading DNA polymerase.[84] Compared to the fully extended unmodified control (lane 6), both **3.2** and **3.1** were incorporated efficiently to give rise to full length products. For triphosphates **3.3** and **3.6,** full length products were produced in very low

amounts and many truncation products were observed. Triphosphate **3.5** was not a substrate of Sequenase V2.0 as lane 10 looks identical to lane 5 where no dATP was added. On the denaturing polyacrylamide gel, all the modified primer extension products migrated less than the unmodified extension product. Since imidazole should be neutral in charge in the TBE buffer environment of pH 8.3, the observed gel retardation was probably due to the additional bulk of the modifications. The two extensions involving the longer linker substrates **3.3** and **3.1** were slightly more retarded than the two extensions involving shorter linker substrates **3.2** and **3.6** (very, very faint band).

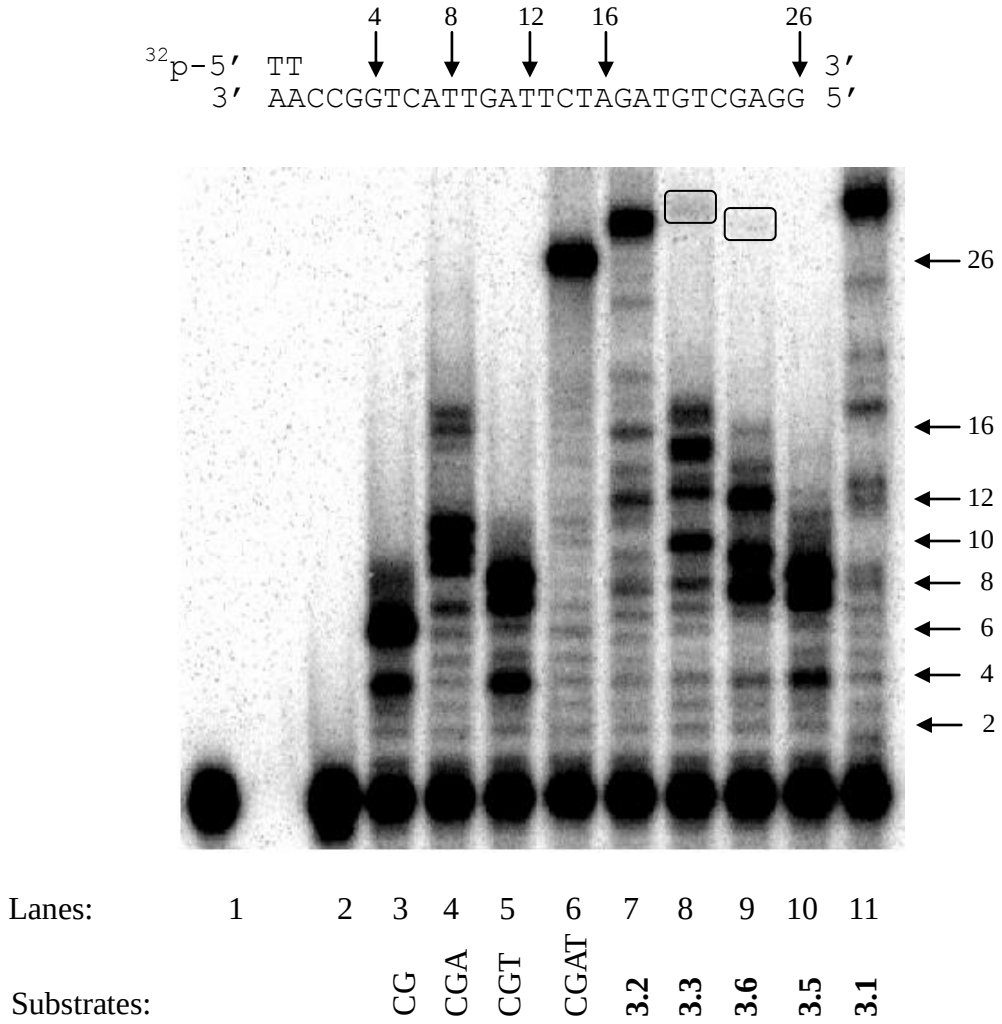


Figure 3.4 Sequenase V2.0 primer extensions of modified dATP`s at 37 ℃ using **T3.1** template. Lane 1, primer only; lane 2, primer and enzyme; lane 3, dCTP and dGTP; lane 4, dCTP, dGTP and dATP; lane 5, dCTP, dGTP and dTTP; lanes 6–11 all contain dCTP, dGTP and dTTP; lane 6, dATP; lane 7, **3.2**; lane 8, **3.3**; lane 9, **3.6**; lane 10, **3.5**; lane 11, **3.1**. The two boxes highlight the faint fully extended product bands for the incorporations of **3.3** and **3.6**. All nucleotides were used at 50 μM. Numbers on the side indicate the approximate number of incorporated unmodified nucleotides at that location of the gel lanes.

In an attempt to improve the incorporations with Sequenase V2.0, enzyme buffers of

pH 6.5 and 8.5 were used rather than the commercially supplied pH 7.5 buffer (results not shown). In the absence of dGTP, dTTP, and dCTP, incorporations with template **T3.2** at pH 8.5 showed that the incorporation of **3.3** was somewhat improved while the incorporation of the other dNTP's was unaffected. However, when the incorporations shown in Figure 3.4 were repeated at pH 8.5, the use of triphosphates **3.2**, **3.3**, **3.5** and **3.6** produced no fully extended product. Only the incorporation of **3.1** gave rise to a fully extended product, but the product band was very faint.

3.3.2 Incorporations with DNA Polymerase IV

Following the examination of Sequenase V2.0, incorporations employing other DNA polymerases that lacked a 3′-5′ exonuclease activity were investigated. The DNA polymerases that were used included Klenow (exo–), Vent (exo–), *Taq*, *Bst* large fragment, Phusion and *Sulfolobus solfataricus* DNA Polymerase IV (Dpo4). Among these DNA polymerases, only Dpo4 was able to incorporate the modified dATP's. The other polymerases either exhibited very little incorporation or none at all.

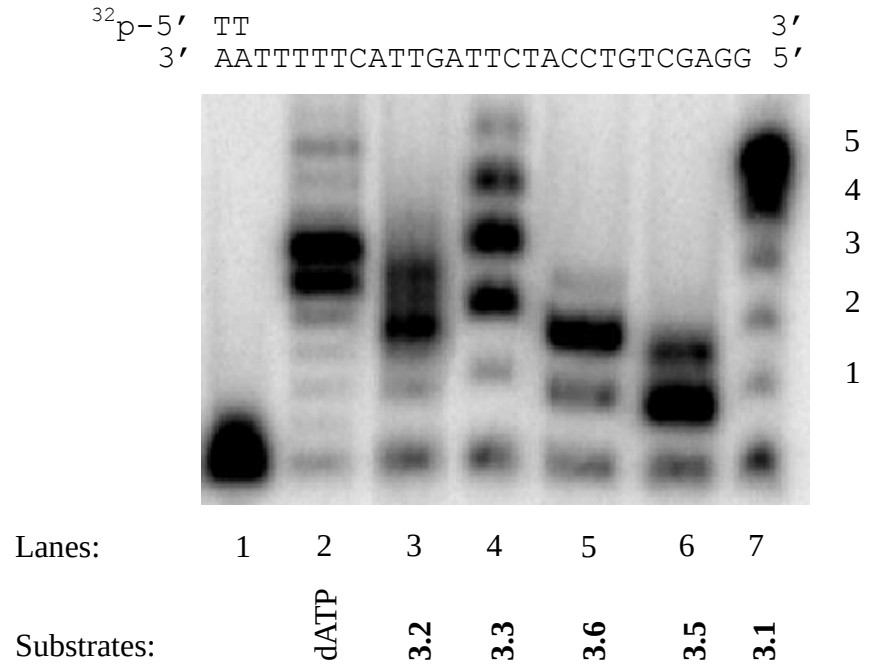


Figure 3.5 Dpo4 primer extensions of modified dATP`s at 37 ℃ using **T3.2** template. Lane 1, primer and enzyme; lane 2, dATP (5 μM); lane 3, **3.2**; lane 4, **3.3**; lane 5, **3.6**; lane 6, **3.5**; lane 7, **3.1**. Modified dATP's were used at 100 μM. Numbers to the right of the gel image indicate the number of modified dATP's incorporated.

Dpo4 incorporation of the modified residues in the absence of natural substrates is shown in Figure 3.5. Using **T3.2** template, all of the modified dATP's were to some degree incorporated. Up to five modified residues in a row were incorporated for both **3.3** and **3.1**, and up to two modified residues in a row were incorporated for the other three modified nucleotides. Initially, these results were very promising. Dpo4 seemed to be able to incorporate 8-modified dATP's much better than Sequenase V2.0. However, when the other three unmodified nucleotides were included in the experiment that is shown in Figure 3.5 (results not shown), the incorporation results were very different. When **3.3** and **3.1** were used under full extension conditions, not only were no full length products

produced, but the extensions also did not go any further than what was observed in Figure 3.5. In the case of **3.1**, Dpo4 was able to incorporate five modified residues consecutively, but could not incorporate unmodified substrates after that position. This was similar to what was observed for **2.24** with Dpo4 in Chapter 2.

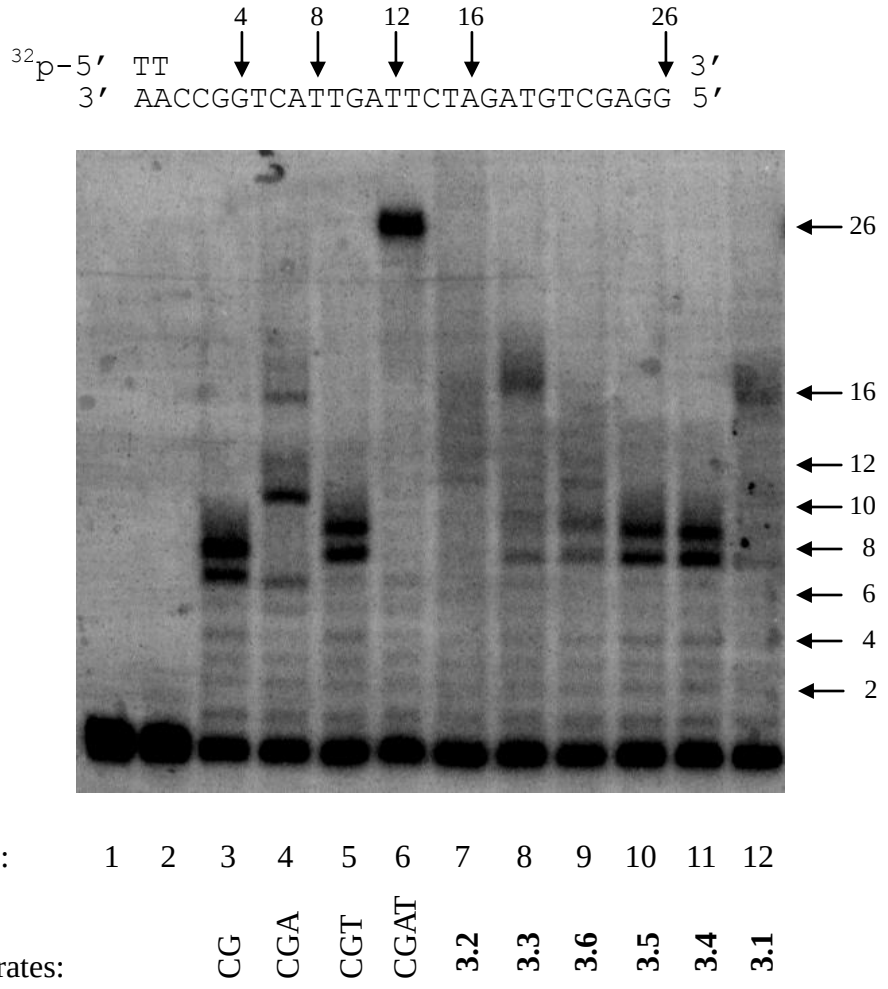


Figure 3.6 Dpo4 primer extension with **T3.1** template at 37 °C. Nucleotides were used at 50 μM. Lane 1, primer only; lane 2, primer and Dpo4; lane 3, dCTP and dGTP; lane 4, dCTP, dGTP and dATP; lane 5, dCTP, dGTP and dTTP; lanes 6-12 all contain dCTP, dGTP and dTTP; lane 6, dATP; lane 7, **3.2**; lane 8, **3.3**; lane 9, **3.6**; lane 10, **3.5**; lane 11, **3.4**; lane 12, **3.1**. Numbers to the right indicate the approximate number of incorporated unmodified nucleotides at that location of the gel lanes.

With these results in mind, full extensions of the modified-dATP's with Dpo4 and **T3.1** template were investigated (Figure 3.6). This was the same experiment as what is shown in Figure 3.4 with Dpo4 substituted for Sequenase V2.0. Since template **T3.1** required up to two modified residues to be incorporated successively rather than five (**T3.2**), it was thought that both **3.3** and **3.1** would give rise to full length incorporation products. Unfortunately, no bands of fully extended product could be found for any of the modified residues. The reactions involving **3.1**–**3.3** showed some extension of the primer whereas **3.4**–**3.6** showed little or no extension.

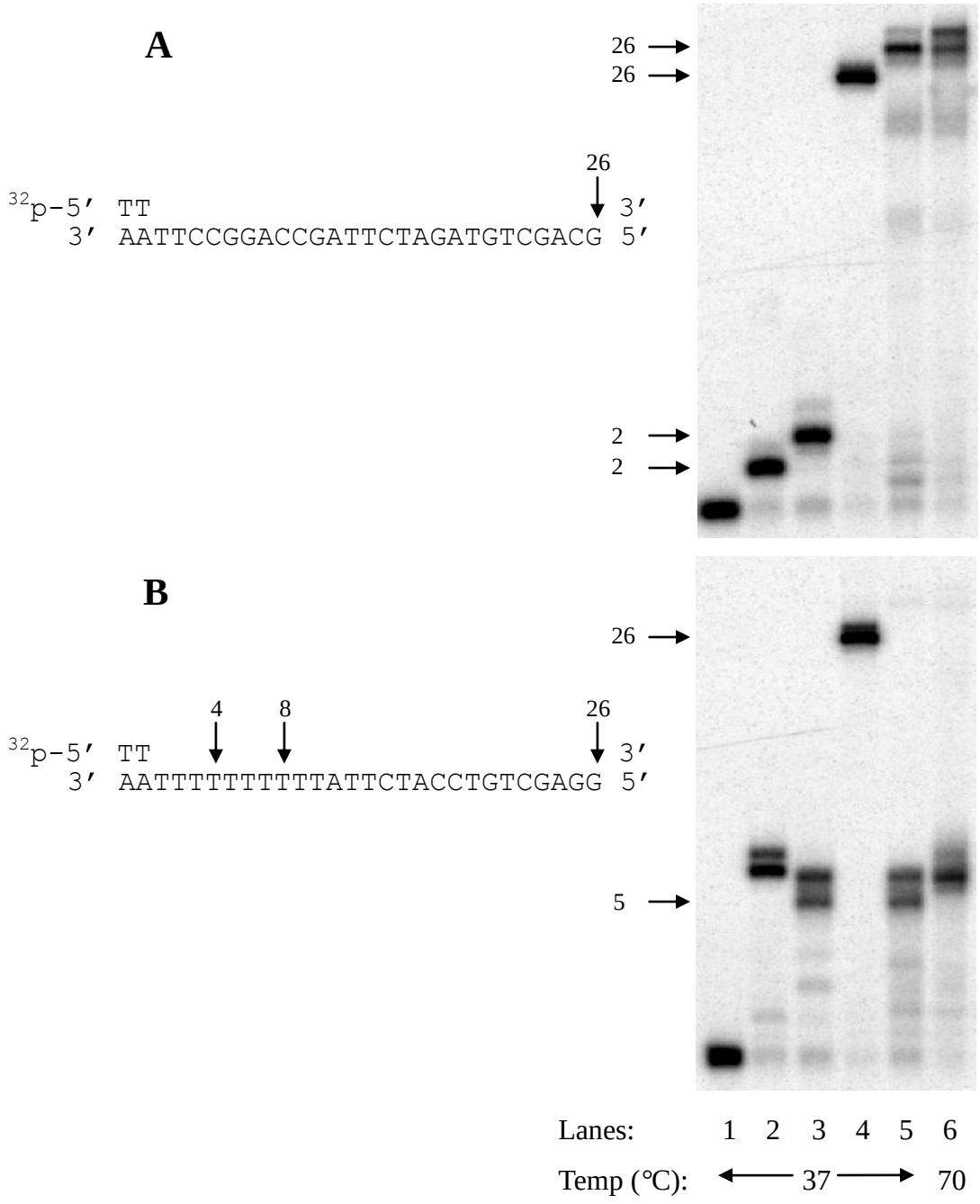


Figure 3.7 Dpo4 primer extension of **3.1** using **T3.3** (A) and **T3.4** (B) templates. Incorporations were incubated for 2 hr at 37 °C (lanes 1-5) or at 70 °C (lane 6). Natural nucleotides were used at 50 μM, and modified nucleotides were used at 100 μM. Lane 1, primer and enzyme; lane 2, dATP; lane 3, **3.1**; lane 4, dATP, dGTP, dTTP, dCTP; lanes 5, 6, **3.1**, dGTP, dTTP, dCTP. Numbers to the left of the gel images represent the number of incorporated nucleotides (modified/unmodified) for the corresponding product bands.

Since Dpo4 preferred **3.1** as a substrate over the other modified residues, two more templates were used with Dpo4 and **3.1** (Figure 3.7). In an effort to produce fully extended primer extension products, the less difficult template **T3.3** was used. As well, since it had been shown that Dpo4 could not incorporate a natural nucleotide after incorporating 5 residues of triphosphate **3.1** (results not shown), we wanted to establish whether Dpo4 could incorporate more than 5 consecutive residues of **3.1** as required by the template **T3.4**. In contrast to template **T3.1**, incorporation with **T3.3** (Figure 3.7A) at both 37 and 70 ℃ gave rise to full length extension products. In the case of the incorporation at 70 ℃, a slightly longer product was produced as well. This longer product may be the result of the incorporation of one extra nucleotide after the template, a process called blunt end addition. Dpo4 is one of several DNA polymerases that have been found to catalyze the non-templated addition of nucleotides onto the 3′-terminus.[185] Using **T3.4** under the same conditions (Figure 3.7B) gave rise to product bands corresponding to the incorporation of five or six modified dATP's consecutively.

3.3.3 Incorporation results analysis

Under the incorporation conditions that were attempted in sections 3.3.1 and 3.3.2, triphosphate **3.1** was found to be the best substrate. It could be efficiently incorporated by both Sequenase V2.0 and Dpo4. Primer extensions employing the two analogs of **3.1** that contained linkers that were shortened or elongated by one methylene unit were somewhat successful as well. The modified dNTP's that did not contain an amine in the linker arm were found to be more difficult to incorporate. Since the linkers of both **3.5** and **3.6** contain the same number of atoms as the linker in **3.2**, this incorporation difficulty was probably not due to linker length. A more conceivable explanation for this would be that

an amine at the 8-position could potentially form a hydrogen bond with the 5′-oxygen on the sugar.[186] In double-stranded DNA, nucleotides are arranged in the *anti* conformation in which the 8-position is on the same side as the sugar-phosphate backbone. The attachment of steric bulk at the 8-position may have hindered the ability of the modified residues to adopt an *anti* conformation.[187] The presence of a hydrogen bond between the sugar and the linker would have stabilized the *anti* conformation and hence improved the incorporation of the modified residues that contained an amine at the 8-position.

Although the 8-modified dATP's that were synthesized in this study were somewhat difficult to incorporate, they could still be used for *in vitro* selections. As mentioned earlier, a selection was performed using free histidine as a cofactor for the discovery of unmodified DNAzymes that could catalyze the cleavage of ribonucleotide phosphodiester bonds.[170] Since only a very small number of histidines will be bound by the DNAzyme, the imidazoles of the bound histidines are probably primarily responsible for catalysis rather than the folding of the DNAzyme. Therefore, a DNAzyme with a very small number of properly positioned catalytic imidazoles should be functional. In search of a Hg^{+2} sensing DNAzyme, Hollenstein *et al.* used Sequenase V2.0, 5-aminoallyl-dUTP (**2.2**) and **3.1** to synthesize a modified DNA library containing a 40 base random region.[188] An *in vitro* selection that was performed on this library gave rise to a DNAzyme, 10-13, that contained only 4 modified dA's in the random region. With this in mind, both **3.2** and **3.3** will serve as candidates for the *in vitro* selection of potentially more catalytically active DNAzymes.

3.4 Photodecomposition of triphosphate 3.4

3.4.1 Incorporation of triphosphate 3.4

During the incorporation of the modified dATP's that were synthesized, it was found that **3.4** appeared to be incorporated by both Sequenase V2.0 and Klenow (exo–). However, no change in mobility of product bands on denaturing PAGE was observed when compared to unmodified bands. In Figure 3.8, the incorporation using Sequenase V2.0 DNA polymerase of **3.4** at four different concentrations is shown. One can see that the band in lane 3 due to the incorporation of one dideoxyadenosine lines up with the bands corresponding to single nucleotide incorporation in lanes 5 and 6. This was in stark contrast to the other 4 modified dATP's in that their product bands were all retarded in mobility on denaturing PAGE. The dATP in question, 8-histidinyl-dATP (**3.4**), has a negative charge on its carboxylate group that should enhance the mobility of products on denaturing PAGE, but it seemed rather coincidental that this mobility enhancement would appear to almost exactly cancel out the gel retardation from the steric bulk of the attached modification to give a gel mobility that was identical to elongation products that used unmodified dATP/ddATP.

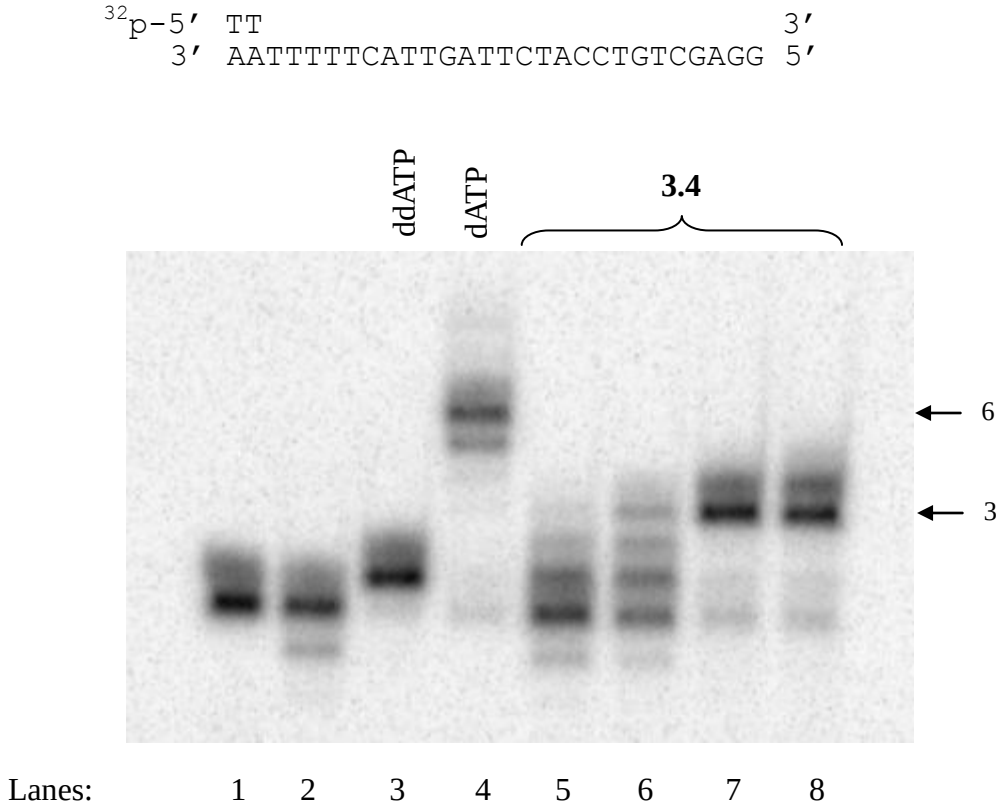


Figure 3.8 Sequenase V2.0 DNA polymerase incorporation of **3.4**. Primer extensions were performed at 37 ℃ and pH 8.5 with template **T3.2**. Lane 1: primer only; lane 2: primer and enzyme; lane 3: 10 μM ddATP; lane 4: 10 μM dATP; lanes 5-8: increasing concentrations of 8-histidinyl-dATP 4, 12, 40, 100 μM. Numbers to the right of the gel image correspond to the number of nucleotides incorporated.

Previously, it was established that 8-[2-(4-imidazolyl)ethylthio]-2′-deoxyadenosine triphosphate would undergo photodecomposition on exposure to incident UV light from a UV detector in the final step of purification.[189] During HPLC purification, small amounts of the modified nucleotide were converted to dATP. Since dATP is the preferred substrate for DNA polymerases, even at very low concentrations dATP would be preferentially incorporated over a modified substrate that is present at a much higher concentration. Based on these previous results, it was suspected that the bands in Figure 3.8 resulted

from the incorporation of a contaminating substrate and not the desired substrate. As a result, we began to question whether a similar decomposition was occurring to transform **3.4** into a different substance more closely related to dATP. This would explain the mobility shifts that were observed by denaturing PAGE. To test this, samples of **3.4** were deliberately irradiated with a hand held UV lamp (254 nm) for increasing amounts of time before the modified nucleotide was used for enzymatic incorporations. In Figure 3.9, incorporation of **3.4** was shown to improve with samples that were irradiated for increasing amounts of time. The results in Figure 3.9 suggest that the modified nucleoside triphosphate was susceptible to UV irradiation, and the resulting decomposition product was a substrate for DNA polymerases.

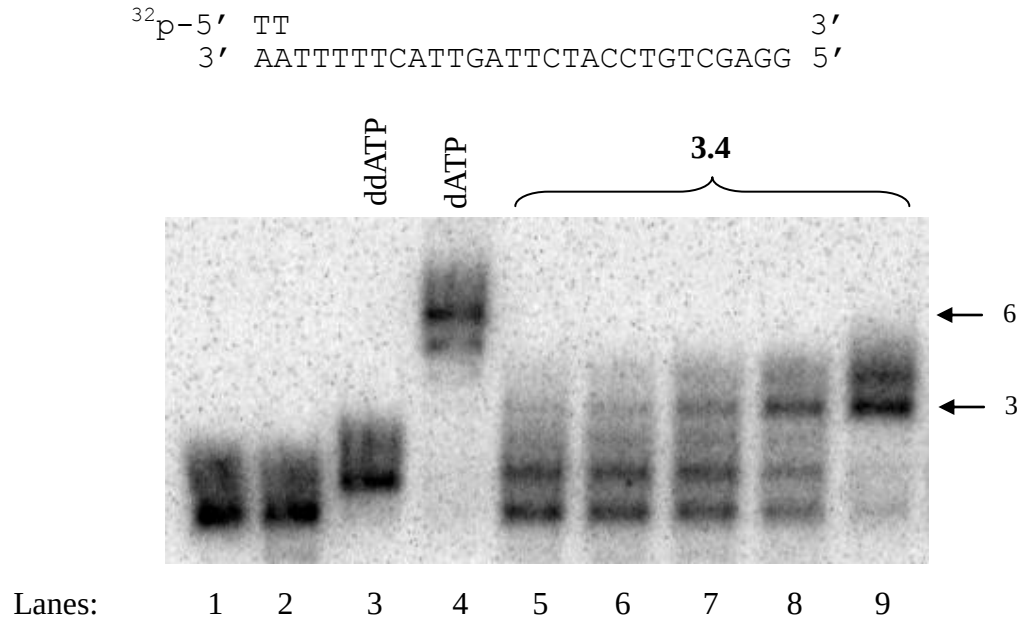


Figure 3.9 Sequenase V2.0 DNA polymerase incorporation of **3.4** that had been irradiated with UV light. Primer extensions were performed at 37 ℃ and pH 8.5 with template **T3.2**. Lane 1: primer only; lane 2: primer and enzyme; lane 3: 10 μM ddATP; lane 4: 10 μM dATP; lanes 5-9: 12 μM 8-histidinyl-dATP that had been irradiated with UV light for 0, 1, 5, 10 and 30 min. Numbers to the right of the gel image correspond to the number of nucleotides incorporated.

3.4.2 Identification of photodecomposition products

A sample of **3.4** was irradiated with a hand held UV lamp (254 nm) for prolonged amounts of time and monitored by UV. It was found that the absorbance maximum moved from 277 nm to 273 nm as illustrated in Figure 3.10. A mass spectrum of this sample revealed a complete absence of the mass corresponding to **3.4** and the appearance of a mass peak that matched that of 8-amino-dATP. A search through the literature found the very similar molecule 8-amino-dAMP, which has an absorbance maximum that is found at the slightly longer wavelength of 274.7 nm.[190] Due to these observations, it

seemed highly likely that one of the decomposition products was 8-amino-dATP. Two plausible mechanisms for the formation of 8-amino-dATP are shown in Scheme 3.6. Homolytic cleavage of the bond between the 8-amine and the α-carbon gives rise to a pair of radicals. The nitrogen radical can then perform a hydrogen abstraction to give 8-amino-dATP. The reaction could also proceed in a Norrish type II manner[191] to give a 1,4 diradical that then fragments to also give 8-amino-dATP. Both of these mechanisms would also give rise to 1 equivalent of urocanic acid. According to the literature, urocanic acid exists as an equilibrium of the *cis* and *trans* isomers. Both of these isomers have significant absorbances in the same region as 8-amino-dATP.[192] However, no additional absorbance intensity/peaks could be found in the UV spectrum of the photoirradiated sample (Figure 3.10) suggesting that if urocanic acid were produced, it may have been further photolyzed to a heretofore uncharacterized product that remains unknown in this context. Urocanic acid was also not found in a mass spectrum of the reaction. From these observations, it is uncertain whether the photodecomposition occurs through either of the two mechanisms. We therefore must perform additional experiments in order to understand the mechanism of the photodecomposition.

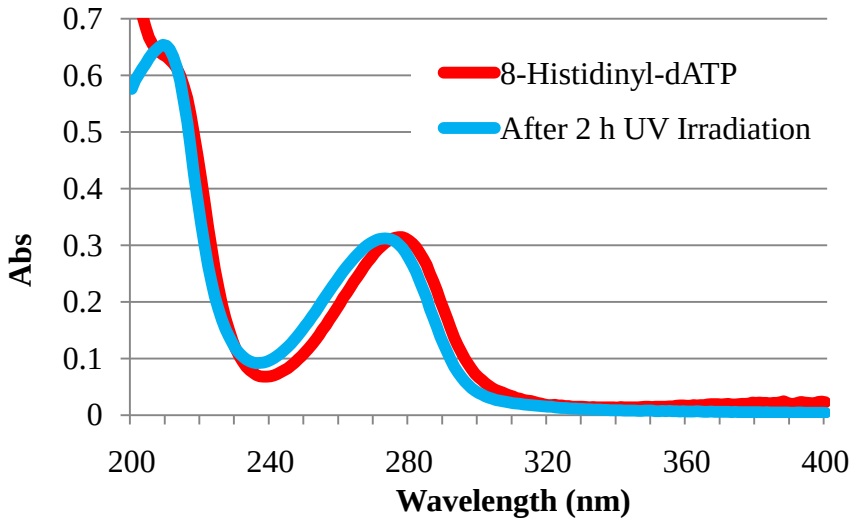
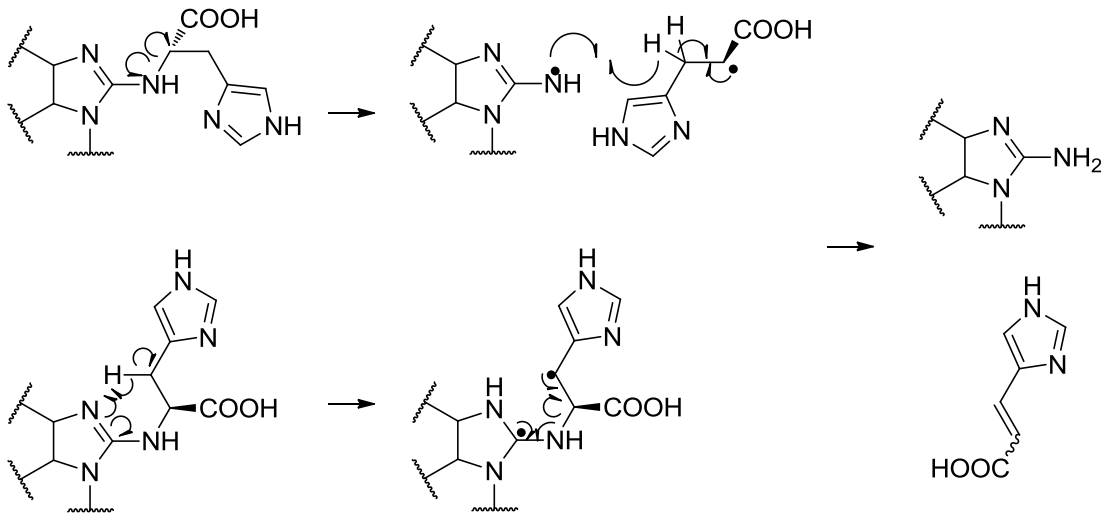


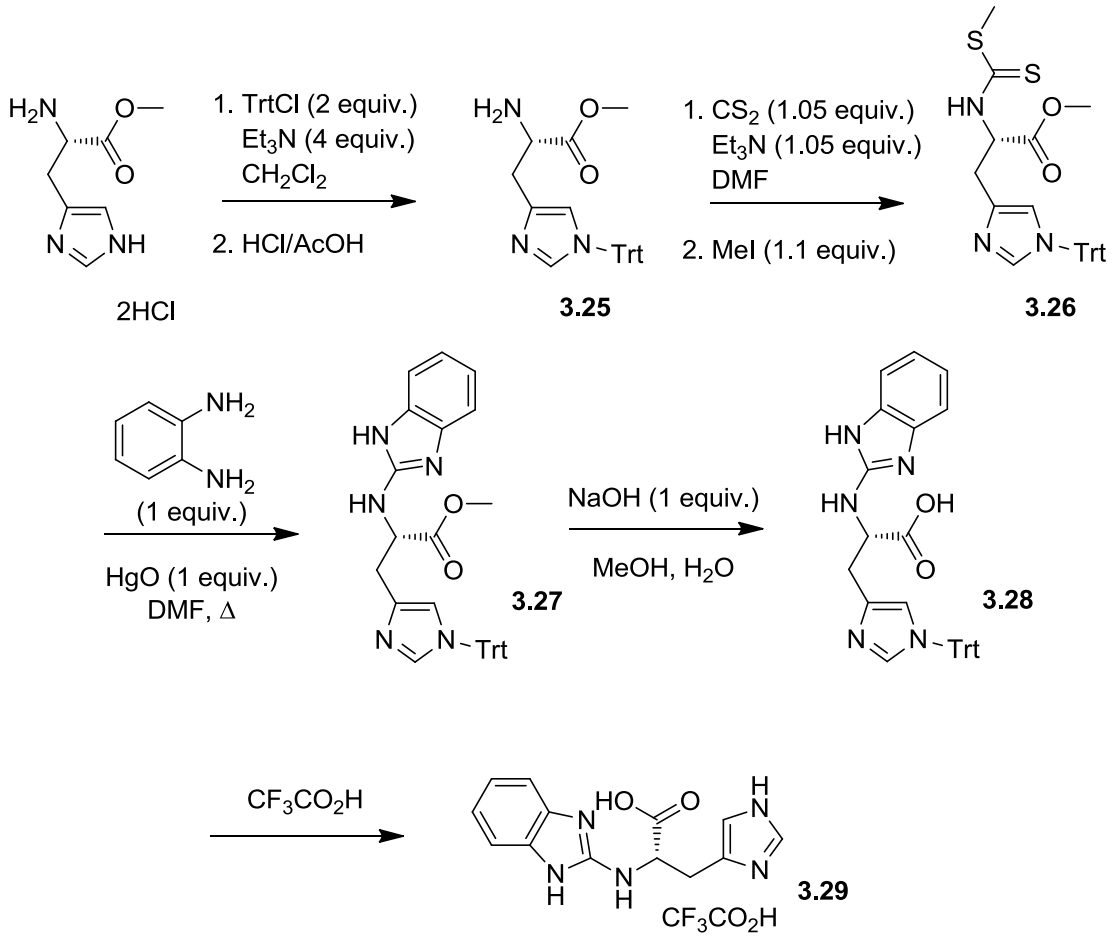
Figure 3.10 Photodecomposition of **3.4**. A 400 µL solution of 12.5 µM **3.4** was irradiated with a hand held UV lamp (254 nm) for 2 h at room temperature. The λ_{max} was shifted from 277 nm to 273 nm.



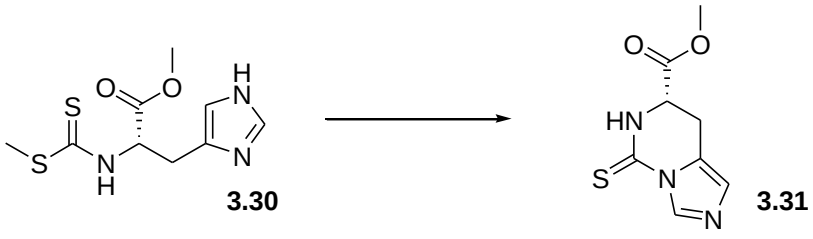
Scheme 3.6 Two plausible mechanisms for the photodecomposition of **3.4**.

3.4.3 Synthesis of 2-((*S*)-1-carboxy-2-(4-imidazolyl)ethylamino)benzimidazole

In order to study the photodecomposition in more detail and investigate whether this phenomenon can be generalized to small molecules of synthetic interest, we prepared a simple analog of **3.4** that could be more readily scaled up. Thus we substituted benzimidazole for dATP to ask whether this simple aromatic chromophore would be photoexcited for chemical transformation. The synthesis is outlined in Scheme 3.7. Histidine with a single trityl group on the imidazole was obtained by the method of Himes *et al*.[193] Histidine methyl ester dihydrochloride was tritylated on both the imidazole and the α-amine in the presence of triethylamine and methylene chloride. The resulting product was then selectively deprotected to give **3.25**. The imidazole-protected histidine was then reacted with carbon disulfide and triethylamine, and to the resulting product was added one equivalent of methyliodide to give the dithiocarbamate (**3.26**) in high yield. This reaction was initially attempted without a trityl group on the imidazole, which gave rise to the byproduct **3.31** in addition to the non-tritylated desired product (**3.30**) (Scheme 3.8). When **3.30** was used for the next step of synthesis, it was entirely converted to the intramolecular cyclization product. Compound **3.26** was heated with phenylenediamine and red mercuric (II) oxide in DMF to give benzimidazole **3.27**.[194] The methyl ester was removed with sodium hydroxide, and the trityl group was removed with trifluoroacetic acid. The final desired product **3.29** was recrystallized as the trifluoroacetate salt.



Scheme 3.7 Synthesis of benzimidazole **3.29**.



Scheme 3.8 Intramolecular cyclization of non-tritylated dithiocarbamate.

3.4.4 Photoirradiation of benzimidazole 3.29

The benzimidazole **3.29** was photoirradiated with a hand held UV lamp in a manner identical to that used for triphosphate **3.4**. Figure 3.11 shows the UV spectrum of **3.29** before and after two hours of irradiation from a hand held UV lamp (254 nm). Unlike the triphosphate, the absorbance intensity of **3.29** decreases over this time suggesting that the benzimidazole ring was decomposing. The absorbance maximum of the sample migrated from 276 nm to 274 nm. A more concentrated sample was irradiated in the same way and analyzed by TLC (EtOH/EtOAc/NH_4OH 6:3:1). The resulting plate showed significant streaking along with two distinguishable spots, the more pronounced of which corresponded to the starting material **3.29** and a smaller of which corresponded to 2-aminobenzimidazole. Mass spectrometry of this photodecomposition only showed peaks corresponding to the masses of **3.29** and 2-aminobenzimidazole.

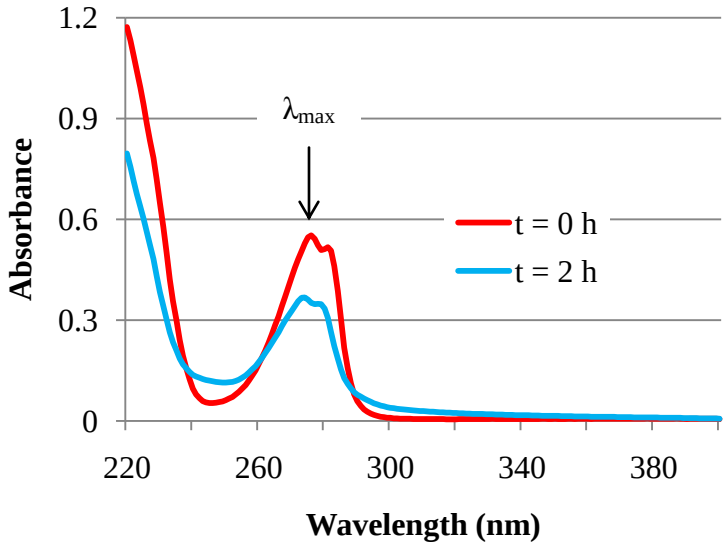


Figure 3.11 Photoirradiation of benzimidazole **3.29**. A 400 µL solution of 56 µM material was irradiated with a hand held UV lamp (254 nm) for 2 h at room temperature. The λ_{max} was shifted from 276 nm to 274 nm.

3.5 Conclusions

The synthesis and enzymatic incorporation of five 8-modified imidazole-containing dATP's were studied. For the incorporations, the modified residues with aminomethyl (**3.2**) and aminopropyl (**3.3**) linkages could be incorporated by Sequenase V2.0 to give full length products when the template did not require too many modified residues to be incorporated successively. Because DNAzymes should only require a small number of imidazoles for catalysis, at least two of the modified residues will be satisfactory substrates for *in vitro* selections in the search of improved catalysis. Unexpectedly, the addition of a carboxylate group to the aminoethyl linkage resulted in a modified 2′-deoxyadenosine triphosphate (**3.4**) that underwent photodecomposition. In an effort to evaluate this reaction, a benzimidazole analog of the triphosphate was synthesized. The photodecomposition of this analog was found to be more complicated and slower than that of the triphosphate. The mechanism of this photodecomposition is still unknown.

3.6 Experimental

The chemicals, enzymes, materials, methods and equipment used are almost identical to Chapter 2. Phusion DNA polymerase was obtained from New England Biolabs. 8-[2-(4-Imidazoylyl)-ethylamino]-2′-deoxyadenosine triphosphate was obtained from Dr. David Dietrich.

HPLC purification of the nucleoside triphosphates was performed on an Agilent 1100 system using a Phenomenex Jupiter 10μ C4 300A column. A flow rate of 1 mL/min and eluents containing 50 mM ammonium acetate (pH 7) were used. The gradient systems are shown in Table 3.1. Purification was performed from 0–18 min, and washing of the column was performed from 18–30 min.

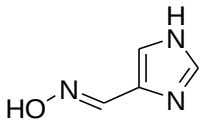
Time (min)	Program 3A % MeCN/H_2O	Program 3B % MeCN/H_2O
0	1	0
10	5	4
18	25	
19	50	
24	50	
25	0	
30	0	

Table 3.1 HPLC gradient systems.

MALDI-TOF mass spectra were collected using a Bruker Biflex instrument in positive ion mode. A solution of 3-hydroxypicolinic acid (30 mg) in acetonitrile (0.25 mL) and H_2O (0.25 mL) and a solution of ammonium citrate$_{(aq)}$ (2.4 %) were mixed in a four to one ratio for the matrix. Prior to sample loading, cation exchange beads (Bio-Rad AG50W-X8 resin converted to the ammonium form) were added to each nucleoside

triphosphate sample (2 µl, 100–500 µM), and the samples were left at room temperature for at least 15 min. Matrix solution (1 µl) was first applied to the MALDI target followed by nucleoside triphosphate sample (1 µl), and the solutions were mixed thoroughly. Care was taken to make sure that at least a few of the cation exchange beads were loaded onto the MALDI target plate as well.

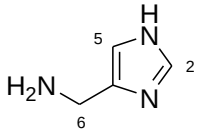
3.6.1 Synthesis



$C_4H_5N_3O$
Exact Mass: 111.0433
Mol. Wt.: 111.102

4-Imidazolecarboxaldehyde oxime (3.7)

Synthesized according to Hubball and Pyman.[175] R_f = 0.42 (MeOH/$CHCl_3$ 2:5). ^{1}H NMR (300 MHz, CH_3OH, 25 °C): δ = 7.72–7.70 (m, 1H), 7.37 (s, 1H). MS (ESI^+) 112.1 $[M+H]^+$.

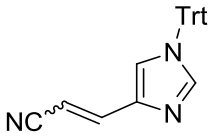


2HCl

$C_4H_9Cl_2N_3$
Exact Mass: 169.0174
Mol. Wt.: 170.0404

4-Aminomethylimidazole dihydrochloride (3.8)

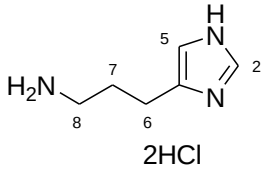
Synthesized according to Turner *et al.*[176] Methanolic HCl was obtained by adding acetyl chloride to methanol while cooling. ^{1}H NMR (300 MHz, CH_3OH, 25 °C): δ = 8.93 (s, 1H, H2), 7.65 (s, 1H, H5), 4.25 (s, 2H, H6). MS (EI^+) 97 $[M]^+$.



$C_{25}H_{19}N_3$
Exact Mass: 361.16
Mol. Wt.: 361.44

3-(*N*-Trityl-4-imidazolyl)acrylonitrile (3.9)

Preparation according to Griffith and Dipietro[177] gave rise to a mixture of *cis*/*trans* isomers. Only the trans isomer was isolated pure. R_f = 0.60 (EtOAc/$CHCl_3$ 1:9). ^{1}H NMR (300 MHz, $CHCl_3$, 25 °C): δ = 7.44 (s, 1H), 7.37–7.31 (m, 9H), 7.15 (d, J = 16.0 Hz, 1H), 7.12–7.06 (m, 6H), 6.99–6.98 (m, 1H), 6.03 (d, J = 16.0 Hz, 1H). MS (ESI$^+$) 384.4 $[M+Na]^+$.

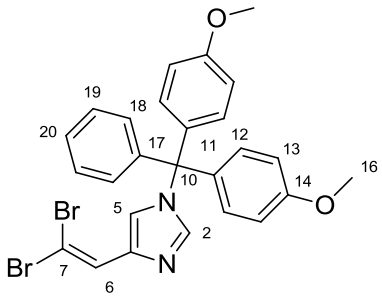


$C_6H_{13}Cl_2N_3$
Exact Mass: 197.0487
Mol. Wt.: 198.0935

4-(3-Amino)propylimidazole dihydrochloride (3.10)

3-(*N*-Trityl-4-imidazolyl)acrylonitrile (**3.9**) (1.81 g, 5.0 mmol, mixture of *cis*/*trans* isomers) was converted into **3.10** according to the procedure of Adger and Surtees.[178] However, in our hands, recrystallization of the product could not be performed as described. The crude product and trityl chloride (2.79 g, 10 mmol) were suspended in Et_3N (5.6 mL, 40 mmol) and DMF (50 mL). After stirring at room temperature for 12 h, the solvent was removed. The crude product was redissolved in CH_2Cl_2, washed with aq. $NaHCO_3$ (5 %), and dried with sodium sulfate (anhydrous). The bis-tritylated intermediate was purified on silica gel (17×3.4 cm column). A white solid (1.03 g, 34 %)

was eluted with EtOAc/hexanes (3:2); R_f = 0.5. This product was then suspended in aq. HCl (25 mL, 3 M), and acetone was added until all solid components dissolved. The contents were refluxed for 1 h. After removal of the acid, the product was recrystallized with EtOH/EtOAc to give 216 mg of white crystals (overall yield of 22 %).[1]H NMR (300 MHz, D_2O, 25 °C): δ = 8.44 (s, 1H, H2), 7.12 (s, 1H, H5), 2.91 (t, J = 7.8 Hz, 2H, H8), 2.70 (t, J = 7.6 Hz, 2H, H6), 1.91 (dt, J = 7.8, 7.6 Hz, 2H, H7). MS (ESI^+) 126.2 $[M+H]^+$.

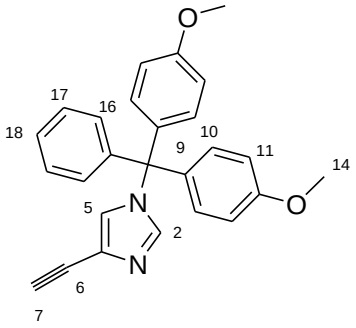


$C_{26}H_{22}Br_2N_2O_2$
Exact Mass: 552.0048
Mol. Wt.: 554.2731

4-(2,2-Dibromovinyl)-1-(4,4′-dimethoxytrityl)imidazole (3.11)

Carbon tetrabromide (10 g, 30.2 mmol) was suspended in CH_2Cl_2 (125 ml), and triphenylphosphine (15.75 g, 60.0 mmol) was added to this solution at –15 °C to give a red-orange mixture. The cold bath was then removed. After 15 min of stirring, 4-imidazolecarboxaldehyde (1.00 g, 10.4 mmol) was added at 0 °C. The reaction was stirred at room temperature for 2 hr and was then transferred to a separatory funnel. The product was extracted with $HCl_{(aq)}$ (100 ml, 5 M) and washed with CH_2Cl_2 (75 ml) 5 times. The extract was dried down and resuspended in $NaOH_{(aq)}$ (60 ml, 5M). The product was extracted with 3 portions of CH_2Cl_2 (60 ml) and dried with magnesium sulfate (anhydrous). The organic extracts were evaporated. The resulting solid was dissolved in CH_2Cl_2 (125 ml) and Et_3N (4 ml, 28.7 mmol). 4,4′-Dimethoxytrityl chloride

(4.33 g, 12.7 mmol) was added to the mixture. After 75 min of stirring at room temperature, the reaction was washed with 3 portions of $NaHCO_{3(aq)}$ (5 %) and dried over magnesium sulfate (anhydrous). The organic extract was concentrated and purified by flash chromatography (20×3.0 cm column). The product eluted in EtOAc/hexanes (1:4) to give 2.68 g (47 %) of a yellow foam. R_f = 0.57 (EtOAc/hexanes 3:2). ^{1}H NMR (300 MHz, CD_2Cl_2, 25 °C): δ = 7.46 (s, 2H, H5, H6), 7.38 (s, 1H, H2), 7.37–7.34 (m, 3H, H19, H20), 7.14–7.10 (m, 2H, H18), 7.06 (d, *J* = 9.0 Hz, 4H, H12), 6.86 (d, *J* = 9.0 Hz, 4H, H13), 3.81 (s, 6H, H16). ^{13}C NMR (100 MHz, CD_2Cl_2, 25 °C): δ = 159.7 (C14), 143.2 (C17), 138.7 (C6), 136.9 (C4), 134.8 (C11), 132.3 (C2), 131.3 (C12), 129.8 (C19), 128.4 (C18), 128.3 (C20), 122.2 (C5), 113.6 (C13), 86.0 (C7), 75.4 (C10), 55.7 (C16). MS (ESI^+) 575.0 $[M+Na]^+$.

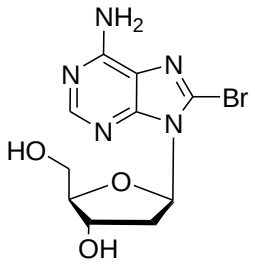


$C_{26}H_{22}N_2O_2$
Exact Mass: 394.1681
Mol. Wt.: 394.4651

2-[1-(4,4′-Dimethoxytrityl)-1*H*-imidazol-4yl]ethyne (3.12)

Compound **3.11** (2.68 g, 4.8 mmol) was suspended in THF (100 ml) and cooled to –78 °C. t-Butyl lithium (20 ml, 34 mmol) was added to this solution. The cold bath was then removed. After 5 min of stirring, MeOH (10 ml) was slowly added followed by the addition of H_2O (30 ml). The THF was evaporated. The resulting crude product was extracted with CH_2Cl_2 (60 ml), dried over sodium sulfate (anhydrous) and purified by

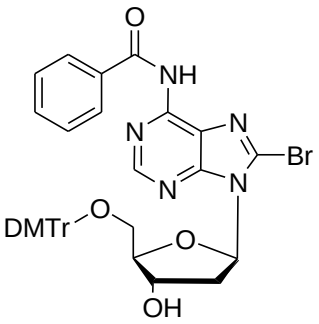
silica flash chromatography (23×2.0 cm column). An off-white solid, 1.53 g (81 %), was eluted with EtOAc/hexanes (2:3). R_f = 0.50 (EtOAc/hexanes 3:2). ^{1}H NMR (300 MHz, $CDCl_3$, 25 ℃): δ = 7.36 (s, 1H, H2), 7.36–7.24 (m, 3H, H17, H18), 7.20–7.15 (m, 3H, H5, H16), 7.00 (d, J = 9.0 Hz, 4H, H10), 6.82 (d, J = 9 Hz, 4H, H11), 3.79 (s, 6H, H14), 3.03 (s, 1H, H7). MS (ESI$^+$) 417.1 $[M+Na]^+$.



$C_{10}H_{12}BrN_5O_3$
Exact Mass: 329.0124
Mol. Wt.: 330.138

8-Bromo-2′-deoxyadenosine (3.13)

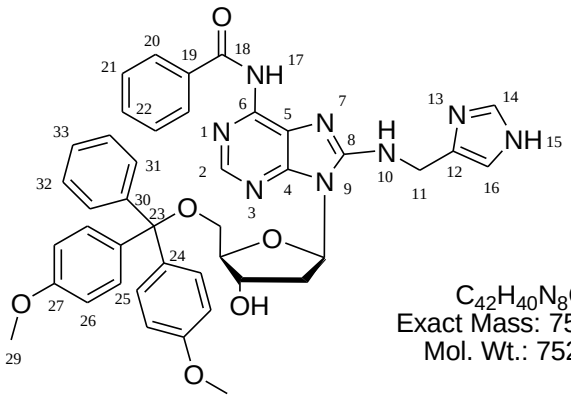
Synthesized according to Ikehara and Kaneko.[181] R_f = 0.22 (MeOH/$CHCl_3$ 1:9). ^{1}H NMR (300 MHz, DMSO, 25 ℃): δ = 8.09 (s, 1H), 7.50 (s, br, 2H), 6.32–6.24 (m, 1H), 5.37–5.30, (m, 1H), 5.30–5.21 (m, 1H), 4.51–4.43 (m, 1H), 3.91–3.83 (m, 1H), 3.68–3.59 (m, 1H), 3.52–3.41 (m, 1H), 3.29–3.17 (m, 1H), 2.23–2.13 (m, 1H). MS (ESI$^+$) 352.0 $[M+Na]^+$.



$C_{38}H_{34}BrN_5O_6$
Exact Mass: 735.1692
Mol. Wt.: 736.6105

***N*^{6}-Benzoyl-5′-*O*-(4,4′-dimethoxytrityl)-8-bromo-2′-deoxyadenosine (3.14)**

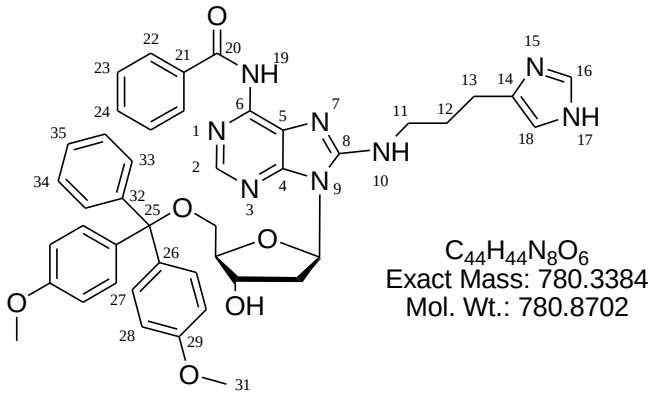
Synthesis of **3.14** was performed according to Singh *et al.*[182] except the order of the reactions was altered. The DMTr group was installed prior to benzoylation rather than after. R_f = 0.39 (MeOH/$CHCl_3$ 1:9). ^{1}H NMR (400 MHz, CD_3OD, 25 °C): δ = 8.48 (s, 1H), 8.09–8.04 (m, 2H), 7.70–7.64 (m, 1H), 7.61–7.55 (m, 2H), 7.36–7.30 (m, 2H), 7.23–7.14 (m, 7H), 6.78–6.68 (m, 4H), 6.59–6.53 (m, 1H), 5.51–5.48 (m, 1H), 4.20–4.14 (m, 1H), 3.78–3.65 (m, 8H), 2.47–2.37 (m, 1H). MS (ESI$^+$) 758.3 $[M+Na]^+$.



***N*^{6}-Benzoyl-5′-*O*-(4,4′-dimethoxytrityl)-8-((4-imidazolyl)methylamino)-2′-deoxyadenosine (3.15)**

4-(Aminomethyl)imidazole dihydrochloride (**3.8**) (460 mg, 2.72 mmol) was suspended in Et_3N (1.52 ml, 10.9 mmol) and EtOH (10 ml). This mixture was added to a solution of **3.14** (1.00 g, 1.36 mmol) in EtOH (10 ml). After 4.5 hrs of stirring at 65°C, the solvent was removed, and the reaction was redissolved in CH_2Cl_2 and washed with $NaHCO_{3(aq)}$ (5 %). The organic extracts were dried over Na_2SO_4 (anhydrous) and purified on silica gel (19×2.5 cm column). The desired product was eluted with MeOH/$CHCl_3$ (5:95) to give 650 mg (63 %) of a slightly off white solid. R_f = 0.31 (MeOH/$CHCl_3$ 1:9). ^{1}H NMR (400 MHz, CD_2Cl_2, 25 °C): δ = 8.38 (s, 1H, H2), 8.05 (d, J = 7.3 Hz, 2H, H20), 7.61 (s,

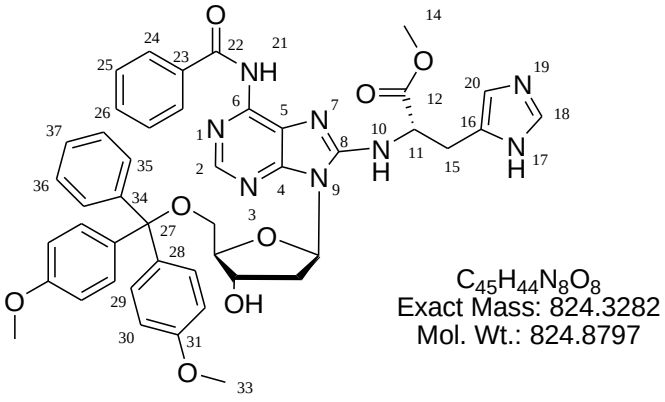
1H, H14), 7.61 (t, *J* = 7.4 Hz, 1H, H22), 7.53 (dd, *J* = 7.4, 7.3 Hz, 2H, H21), 7.40 (d, *J* = 6.7 Hz, 2H, H31), 7.35–7.25 (m, 7H, H33, H32, H25), 6.82 (d, *J* = 8.8 Hz, 4H, H26), 6.47 (dd, *J* = 8.9, 5.7 Hz, 1H, H1′), 6.39 (s, 1H, H16), 4.75–4.70 (m, 1H, H3′), 4.15–4.00 (m, 2H, H11, H4′), 3.77 (s, 6H, H29), 3.60–3.50 (m, 1H, H11), 3.50 (dd, *J* = 10.6, 2.5 Hz, 1H, H5′), 3.41 (dd, *J* = 10.6, 3.0 Hz, 1H, H5′), 2.88 (ddd, *J* = 13.5, 8.9, 6.3 Hz, 1H, H2′), 2.30 (ddd, *J* = 13.5, 5.7, 1.9 Hz, 1H, H2′). ^{13}C NMR (100 MHz, CD_2Cl_2, 25 °C): δ = 166.5 (C18), 159.5 (C27), 154.7 (C8), 154.4 (C4), 148.8 (C2), 144.7 (C30), 143.9 (C6), 135.9 (C14), 135.7 (C24), 134.1 (C12), 133.2 (C22), 130.9, 130.8 (C25), 129.3 (C21), 129.0 (C31), 128.6 (C32), 128.6 (C20), 127.9 (C33), 124.1 (C5), 113.8 (C26), 87.6 (C23), 86.8 (C4′), 84.1 (C1′), 71.9 (C3′), 63.5 (C5′), 55.8 (C29), 39.3 (C2′), 37.3 (C11). HRMS (ESI^+) calcd for $C_{42}H_{41}N_8O_6$ 753.3149, found 753.3160.



N^6-Benzoyl-5′-*O*-(4,4′-dimethoxytrityl)-8-(3-(4-imidazolyl)propylamino)-2′-deoxyadenosine (3.16)

Compound **3.16** was synthesized in the same way as **3.15**. 4-(3-Aminopropyl)imidazole dihydrochloride (**3.10**) (182 mg, 0.92 mmol) and **3.14** (450 mg, 0.61 mmol) were converted into 190 mg (40 %) of a flaky, slightly off-white solid. R_f = 0.21 (MeOH/$CHCl_3$ 1:9). ^{1}H NMR (400 MHz, CD_3OD, 25 °C): δ = 8.28 (s, 1H, H2), 8.04 (d,

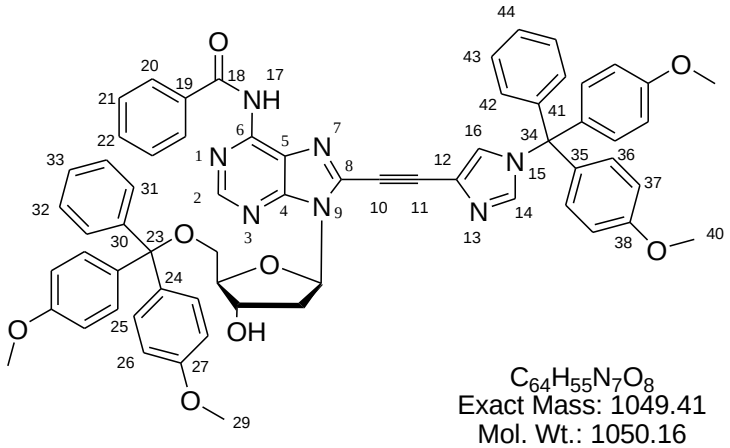
J = 7.3 Hz, 2H, H22), 7.62 (t, *J* = 7.3 Hz, 1H, H24), 7.51 (t, *J* = 7.3 Hz, 2H, H23), 7.49 (s, 1H, H16), 7.31 (d, *J* = 6.8 Hz, 2H, H33), 7.25-7.13 (m, 7H, H35, H34, H27), 6.78–6.68 (m, 5H, H28, H18), 6.26 (t, *J* = 6.4 Hz, 1H, H1′), 4.78 (ddd, *J* = 6.5, 5.0, 3.9 Hz, 1H, H3′), 4.15–4.05 (m, 1H, H4′), 3.70 (s, 3H, H31), 3.69 (s, 3H, H31), 3.52 (ddd, *J* = 13.3, 6.5, 6.4 Hz, 1H, H2′), 3.45–3.25 (m, 4H, H5′, H11), 2.58 (t, *J* = 7.3 Hz, 2H, H13), 2.32 (ddd, *J* = 13.3, 6.4, 5.0 Hz, 1H, H2′), 1.93 (q, *J* = 7.3 Hz, 2H, H12). ^{13}C NMR (100 MHz, CD_3OD, 25 °C): δ = 168.0 (C20), 160.1, 160.0 (C29), 156.1 (C8), 154.6 (C4), 149.0 (C2), 146.1 (C32), 144.4 (C6), 137.1, 136.9 (C26), 135.7 (C16), 135.4 (C14), 133.7 (C24), 131.3, 131.2 (C27), 129.7 (C23), 129.4 (C34), 129.2 (C22), 128.7 (C33), 127.8 (C35), 124.9 (C5), 118.0 (C18), 114.0 (C28), 87.6 (C4′), 87.5 (C1′), 85.2 (C25), 72.6 (C3′), 64.7 (C5′), 55.7 (C31), 43.4 (C11), 38.2 (C2′), 30.0 (C12), 24.7 (C13). HRMS (ESI$^+$) calcd for $C_{44}H_{44}N_8O_6$ 781.3462, found 781.3459.



N^6-Benzoyl-5′-*O*-(4,4′-dimethoxytrityl)-8-(1-methylcarboxy-2-(4-imidazolyl)ethylamino)-2′-deoxyadenosine (3.17)

L-Histidine methyl ester dihydrochloride (480 mg, 2 mmol) and Et_3N (1.1 ml, 7.9 mmol) were suspended in MeOH (4 ml). Nucleoside **3.14** (735 mg, 1 mmol) was added, and the

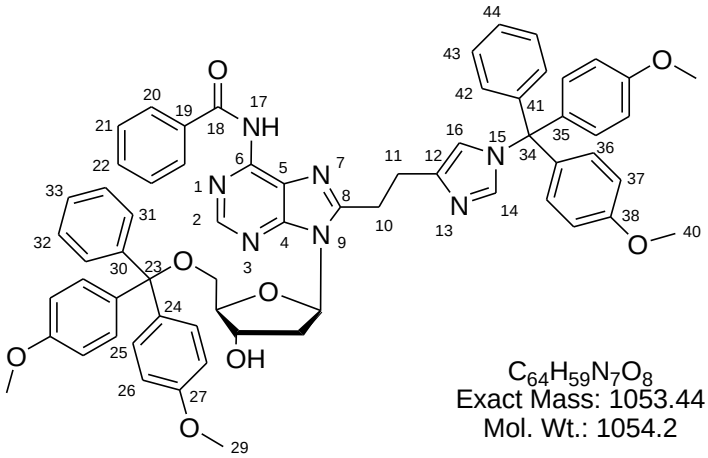
mixture was stirred at 50 °C for 2 days. Following the removal of solvent, the crude product was redissolved in CH_2Cl_2 and washed with $NaHCO_{3(aq)}$ (5 %). Sodium sulfate (anhydrous) was used to dry the organic solution. Purification of the product was done on silica gel (19.5×2.0 cm column). The product was eluted in MeOH/$CHCl_3$ (6:94) to give 191 mg (12 %) of a light yellow powder. R_f = 0.09 (MeOH/$CHCl_3$ 1:9). ^{1}H NMR (400 MHz, CD_2Cl_2, 25 °C): δ = 8.97 (s, 1H, H21), 8.31 (s, 1H, H2), 7.96 (d, *J* = 7.4 Hz, 2H, H24), 7.59 (t, *J* = 7.4 Hz, 1H, H26), 7.50 (t, *J* = 7.4 Hz, 2H, H25), 7.45–7.15 (m, 9H, H37, H36, H35, H29), 7.39 (s, 1H, H18), 6.85–6.70 (m, 5H, H30, H20), 6.23 (t, *J* = 6.6 Hz, 1H, H1′), 4.82 (dd, *J* = 6.2, 4.0 Hz, 1H, H11), 4.61 (m, 1H, H3′), 4.11 (m, 1H, H4′), 3.72 (2s, 6H, H33), 3.64 (s, 3H, H14), 3.4–3.1 (m, 5H, H15, H5′, H2′), 2.32 (ddd, *J* = 13.5, 6.6, 4.0 Hz, 1H, H2′). ^{13}C NMR (100 MHz, CD_2Cl_2, 25 °C): δ = 172.3 (C12), 165.7 (C22), 158.9 (C31), 153.6 (C8), 153.3 (C4), 148.6 (C2), 145.3 (C34), 144.2 (C6), 136.2 (C28), 135.5 (C18), 134.4 (C16), 132.9 (C26), 130.4, 130.3 (C29), 129.1 (C25), 128.4 (C36), 128.2 (C24), 128.1 (C35), 127.1 (C37), 123.2 (C5), 118.6 (C20), 113.3 (C30), 86.6 (C4′), 86.4 (C1′), 84.7 (C27), 72.4 (C3′), 64.2 (C5′), 56.1 (C11), 55.5 (C33), 52.8 (C14), 38.0 (C2′), 29.4 (C15). HRMS (ESI^+) calcd for $C_{45}H_{45}N_8O_8$ 825.3360, found 825.3367.



N^6-Benzoyl-5′-*O*-(4,4′-dimethoxytrityl)-8-(*N*-(4,4′-dimethoxytrityl)(4-imidazolyl)ethynyl)-2′-deoxyadenosine (3.18)

This reaction was performed in the complete absence of O_2 and H_2O. Catalysts were transferred in a glove box, and solvents were degassed by 3 freeze thaw cycles under vacuum. A catalyst suspension was prepared by adding cuprous iodide (92 mg, 0.48 mmol) and tetrakis(triphenylphosphine)palladium (280 mg, 0.24 mmol) into DMF (40 ml). This mixture was added to a solution of **3.14** (1.77 g, 2.4 mmol) and 2-[1-(4,4′-dimethoxytrityl)-1*H*-imidazol-4yl]ethyne (**3.12**) (2.00 g, 5.07 mmol) in DMF (80 ml) and Et_3N (20 ml). After the reaction was stirred at 45°C for 22 h, the solvent was removed. The crude product was redissolved in CH_2Cl_2, washed with $NaHCO_{3(aq)}$ (5 %), and dried over $NaSO_4$ (anhydrous). Purification by silica gel flash chromatography (15×4.0 cm column), with Et_3N/EtOAc (1:1000) as the eluent, afforded 1.98 g (78 %) of compound **3.18** as a yellow-green foam. Additional silica flash chromatograpy Excess starting material [**3.12**; 0.78 g (1.98 mmol)] was isolated on passing through an additional silica column using 0.1/60/40 % Et_3N/EtOAc/hexanes as an eluent. R_f = 0.44 (EtOAc). ^{1}H NMR (400 MHz, CD_2Cl_2, 25 °C): δ = 8.94 (s, 1H, H17), 8.53 (s, 1H, H2), 7.98 (d, J = 7.3 Hz, 2H, H20), 7.70–7.12 (m, 15H, H44, H43, H42, H33, H32, H31, H22, H21, H16,

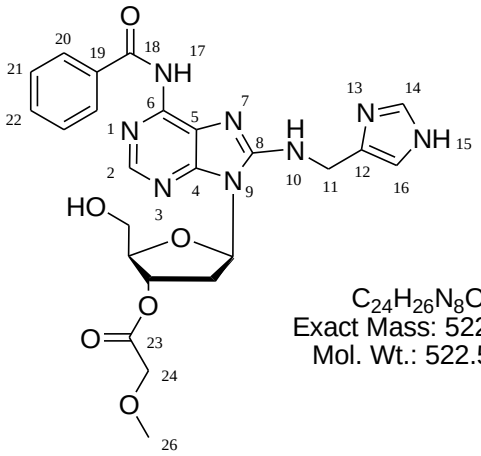
H14), 7.25 (d, *J* = 8.9 Hz, 4H, H25), 7.06 (d, *J* = 9.0 Hz, 4H, H36), 6.88 (d, *J* = 9.0 Hz, 4H, H37), 6.75 (d, *J* = 8.9 Hz, 2H, H26), 6.72 (d, *J* = 8.9 Hz, 2H, H26), 6.64 (dd, *J* = 7.3, 5.8 Hz, 1H, H1′), 5.02–4.96 (m, 1H, H3′), 4.14–4.08 (m, 1H, H4′), 3.82 (s, 6H, H40), 3.75 (s, 3H, H29), 3.74 (s, 3H, H29), 3.52–3.35 (m, 3H, H5′, H2′), 2.42 (ddd, *J* = 13.0, 7.3, 5.0 Hz, 1H, H2′). ^{13}C NMR (100 MHz, CD_2Cl_2, 25 °C): δ = 164.6 (C18), 159.7 (C38), 158.8 (C27), 152.9 (C2), 151.3 (C4), 149.4 (C6), 145.4 (C30), 142.7 (C41), 140.2 (C14), 137.8 (C8), 136.4, 136.3 (C24), 134.3 (C35), 133.0 (C19), 132.3 (C22), 131.3 (C36), 130.4, 130.2 (C25), 129.7 (C43), 129.2 (C21), 128.9 (C44), 128.5 (C42), 128.4 (C32), 128.1 (C20), 128.0 (C31), 126.9 (C33), 123.7 (C5), 120.9 (C16), 113.8 (C37), 113.2 (C26), 92.0 (C10), 86.7 (C23), 86.4 (C4′), 85.5 (C1′), 78.8 (C11), 75.9 (C34), 72.8 (C3′), 64.5 (C5′), 55.7 (C40), 55.5 (C29), 37.8 (C2′). HRMS (ESI$^+$) calcd for $C_{64}H_{55}N_7O_8$ 1072.4010, found 1072.4011.



N^6-Benzoyl-5′-*O*-(4,4′-dimethoxytrityl)-8-(*N*-(4,4′-dimethoxytrityl)(4-imidazolyl)ethyl)-2′-deoxyadenosine (3.19)

Compound **3.18** (500 mg, 0.48 mmol) and PtO_2 (250 mg, 1.10 mmol) was suspended in

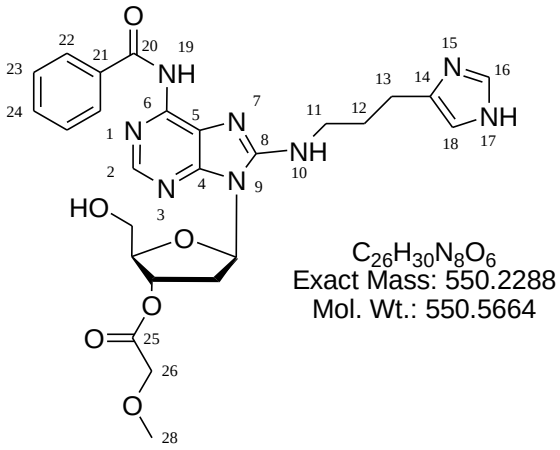
pyridine (2.5 ml) and CH_2Cl_2 (150 ml). This mixture was stirred at room temperature under H_2 (1 atm) for 4 h. Following the removal of solvent, the crude product was dissolved into MeOH. Filter paper was used to remove the catalyst. The solvent was then evaporated. The resulting material was redissolved in CH_2Cl_2, washed with $NaHCO_{3(aq)}$ (5 %), and dried over Na_2SO_4 (anhydrous). Purification by silica gel flash chromatography (12.5×3.0 cm column) using Et_3N/MeOH/$CHCl_3$ (0.1:2:98) gave 265 mg of **3.19** (53 %). R_f = 0.38 (MeOH/$CHCl_3$ 1:9). ^{1}H NMR (400 MHz, CD_2Cl_2, 25 °C): δ = 9.00 (s, 1H, H17), 8.46 (s, 1H, H2), 7.93 (d, *J* = 7.5 Hz, 2H, H20), 7.60 (t, *J* = 7.4 Hz, 1H, H22), 7.51 (dd, *J* = 7.5, 7.4 Hz, 2H, H21), 7.36 (d, *J* = 6.7 Hz, 2H, H31), 7.32 (s, 1H, H14), 7.30–7.12 (m, 10H, H44, H43, H33, H32, H25), 7.03 (d, *J* = 6.9 Hz, 2H, H42), 6.96 (d, *J* = 8.1 Hz, 4H, H36), 6.80–6.68 (m, 8H, H37, H26), 6.56 (s, 1H, H16), 6.39 (t, *J* = 6.8 Hz, 1H, H1′), 4.85–4.79 (m, 1H, H3′), 4.15–4.07 (m, 1H, H4′), 3.76 (s, 6H, H40), 3.74 (s, 3H, H29), 3.73 (s, 3H, H29), 3.57–3.49 (m, 1H, H2′), 3.39–3.27 (m, 4H, H10, H5′), 3.15–3.00 (m, 2H, H11), 2.23 (ddd, *J* = 13.4, 6.8, 4.0 Hz, 1H, H2′). ^{13}C NMR (100 MHz, CD_2Cl_2, 25 °C): δ = 164.4 (C18), 159.0 (C38), 158.4 (C27), 156.1 (C8), 152.3 (C4), 151.3 (C2), 148.2 (C6), 144.7 (C30), 142.9 (C41), 139.0 (C12), 138.4 (C14), 136.0 (C24), 135.9 (C24), 134.6 (C35), 134.1 (C22), 132.5 (C19), 130.9 (C36), 130.0 (C25), 129.9 (C25), 129.4 (C43), 128.8 (C20), 128.2 (C42), 127.9 (C32), (C44), 127.7 (C21), (C31), 126.7 (C33), 113.2 (C26), 113.0 (C37), 86.2 (C23), 86.1 (C4′), 84.4 (C1′), 74.4 (C34), 73.0 (C3′), 63.9 (C5′), 55.3 (C40), 55.2 (C29), 36.8 (C2′), 28.1 (C10), 26.6 (C11). HRMS (ESI^+) calcd for $C_{64}H_{59}N_7O_8$ 1054.4503, found 1054.4507.



N^6-Benzoyl-3′-*O*-methoxyacetyl-8-((4-imidazolyl)methylamino)-2′-deoxyadenosine (3.20)

Compound **3.15** (200 mg, 266 μmol) was dried by co-evaporation from pyridine under vacuum overnight. After dissolving the sample in pyridine (2.0 ml), 75 % methoxyacetic anhydride (234 μl, 980 μmol) was added. The reaction was stirred at room temperature for 1 hr. The reaction was concentrated to a gum and then resuspended in acetic acid/H_2O (2.0 ml, 80 %). After 2.0 hours at room temperature, the reaction was concentrated and purified on silica gel (17×1.5 cm column). The product was eluted in MeOH/$CHCl_3$ (7:93) to give 106 mg (76 %) of a light, pale yellow solid. R_f = 0.11 (MeOH/$CHCl_3$ 1:9). ^{1}H NMR (400 MHz, $CDCl_3$, 25 °C): δ = 8.40 (s, 1H, H2), 8.02 (d, J = 7.3 Hz, 2H, H20), 7.66 (s, 1H, H14), 7.58 (t, J = 7.3 Hz, 1H, H22), 7.48 (t, J = 7.3 Hz, 2H, H21), 6.99 (s, 1H, H16), 6.57 (dd, J = 10.1, 5.4 Hz, 1H, H1′), 5.50 (d, J = 6.4 Hz, 1H, H3′), 4.51 (d, J = 15.2 Hz, 1H, H11), 4.41 (d, J = 15.2 Hz, 1H, H11), 4.15–4.11 (m, 1H, H4′), 4.06 (s, 2H, H24), 3.96 (d, J = 11.6 Hz, 1H, H5′), 3.86 (dd, J = 11.6, 2.0 Hz, 1H, H5′), 3.43 (s, 3H, H26), 2.75 (ddd, J = 14.2, 10.1, 6.4 Hz, 1H, H2′), 2.26 (dd, J = 14.2, 5.4 Hz, 1H, H2′). ^{13}C NMR (100 MHz, $CDCl_3$, 25 °C): δ = 169.9 (C23), 166.1 (C18), 153.6 (C4), 153.0 (C8), 148.0 (C2), 143.1 (C6), 134.3 (C14), 133.4 (C19), 132.5 (C22), 132.0 (C12), 128.5

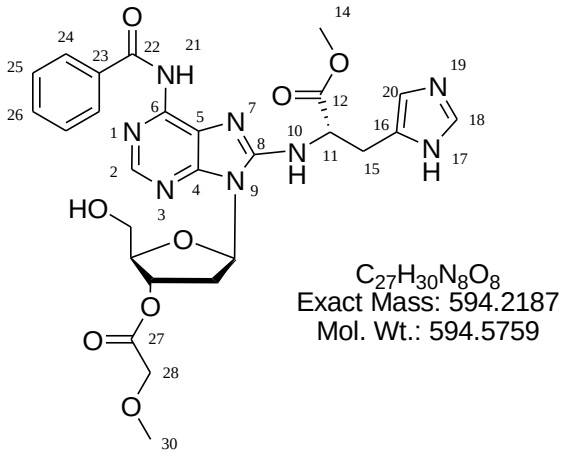
(C21), 127.8 (C20), 123.0 (C5), 119.4 (C16), 85.2 (C4′), 83.5 (C1′), 76.3 (C3′), 69.4 (C24), 61.4 (C5′), 59.2 (C26), 37.1 (C11), 35.3 (C2′). HRMS (ESI^+) calcd for $C_{24}H_{27}N_8O_6$ 523.2054, found 523.2043.



N^6-Benzoyl-3′-*O*-methoxyacetyl-8-(3-(4-imidazolyl)propylamino)-2′-deoxyadenosine (3.21)

Compound **3.21** was prepared in the same way as **3.20**. Compound **3.16** (150 mg, 192 μmol) was converted into 103.3 mg (98 %) of a light, pale-yellow solid. R_f = 0.48 (MeOH/$CHCl_3$ 1:4). ^{1}H NMR (400 MHz, $CDCl_3$, 25 °C): δ = 8.47 (s, 1H, H2), 8.02 (d, *J* = 7.2 Hz, 2H, H22), 7.66 (s, 1H, H16), 7.53 (t, *J* = 7.3 Hz, 1H, H24), 7.45 (dd, *J* = 7.3, 7.2 Hz, 2H, H23), 6.77 (s, 1H, H18), 6.63 (dd, *J* = 10.3, 5.3 Hz, 1H, H5′), 5.54 (d, *J* = 6.5 Hz, 1H, H3′), 4.20–4.16 (m, 1H, H4′), 4.07 (s, 2H, H26), 3.99 (d, *J* = 10.9 Hz, 1H, H5′), 3.90 (dd, *J* = 10.9, 1.7 Hz, 1H, H5′), 3.55–3.35 (m, 2H, H11), 3.44 (s, 3H, H28), 2.77 (ddd, *J* = 14.0, 10.3, 6.5 Hz, 1H, H2′), 2.70–2.60 (m, 2H, H13), 2.26 (dd, *J* = 10.3, 5.3 Hz, 1H, H2′), 2.12–2.00 (m, 1H, H12), 1.90–1.80 (m, 1H, H12). ^{13}C NMR (100 MHz, $CDCl_3$, 25 °C): δ = 169.9 (C25), 165.9 (C20), 153.0 (C8), 152.8 (C4), 148.1 (C2), 143.3 (C6), 133.9 (C21), 132.3 C24), 128.5 (C23), 127.8 (C22), 121.7 (C5), 114.7 (C18), 85.2 (C4′),

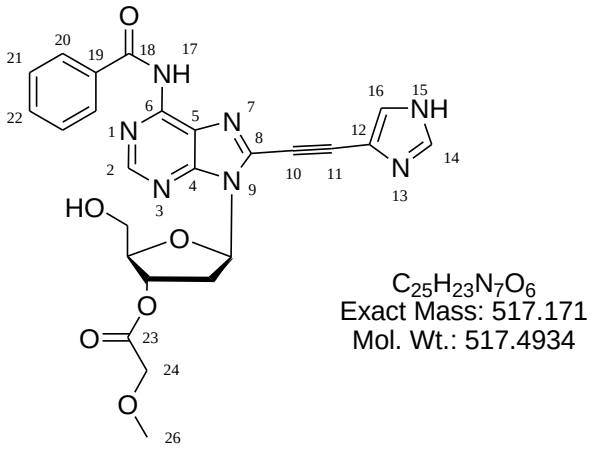
83.4 (C1′) 76.5 (C3′), 69.5 (C26), 61.3 (C5′), 59.3 (C28), 41.6 (C11), 35.2 (C2′), 28.4 (C12), 22.6 (C13). HRMS (ESI$^+$) calcd for $C_{26}H_{31}N_8O_6$ 551.2367, found 551.2362.



N^6-Benzoyl-3′-*O*-methoxyacetyl-8-(1-methylcarboxy-2-(4-imidazolyl)ethylamino)-2′-deoxyadenosine (3.22)

Nucleoside **3.17** (36 mg, 0.043 mmol) was dried by co-evaporation from pyridine under vacuum overnight. After dissolving the sample in pyridine (0.6 ml), 90 % methoxyacetic anhydride (30.8 µl, 174 µmol) was added. The reaction was stirred at room temperature for 100 min. The reaction was concentrated to a gum and redissolved in CH_2Cl_2. The organic solution was washed with $NaHCO_{3(aq)}$ (5 %) and dried with sodium sulfate (anhydrous). Following removal of the solvent, the crude product was resuspended in acetic acid/H_2O (0.6 ml, 80 %). After 2.25 hours at room temperature, the reaction was concentrated and purified on silica gel (11.5×1.5 cm column). The product was eluted in MeOH/$CHCl_3$ (8:92) to give 20.6 mg (80 %) of a white powder. R_f = 0.11 (MeOH/$CHCl_3$ 1:9). ^{1}H NMR (400 MHz, $CDCl_3$, 25 °C): δ = 8.74 (s, 1H, H21), 8.47 (s, 1H, H2), 7.88 (d, *J* = 7.6 Hz, 2H, H24), 7.54 (t, *J* = 7.4 Hz, 1H, H26), 7.50 (s, 1H, H18), 7.44 (dd, *J* = 7.6, 7.4 Hz, 2H, H25), 6.88 (s, 1H, H20), 6.60 (dd, *J* = 10.0, 5.4 Hz, 1H, H1′), 5.63 (d, *J* = 6.4

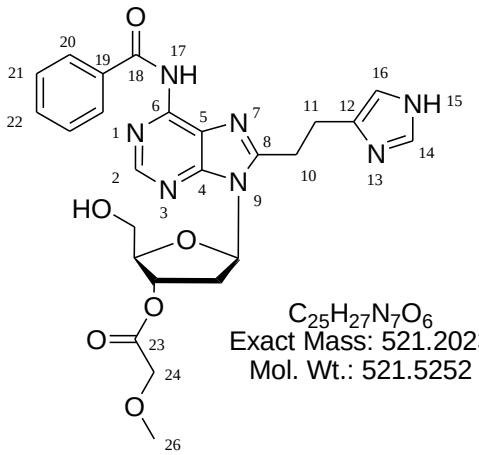
Hz, 1H, H3′), 4.85–4.75 (m, 1H, H11), 4.12 (s, 1H, H4′), 4.09 (s, 2H, H28), 4.02 (d, *J* = 12.1 Hz, 1H, H5′), 3.98 (d, *J* = 12.1 Hz, 1H, H5′), 3.70 (s, 3H, H14), 3.47 (s, 3H, H30), 3.27 (d, *J* = 13.3 Hz, 1H, H15), 3.15–3.00 (m, 2H, H15, H2′), 2.30 (dd, *J* = 14.2, 5.4 Hz, 1H, H2′). ^{13}C NMR (75 MHz, $CDCl_3$, 25 ℃): δ = 172.4 (C12), 169.9 (C27), 165.3 (C22), 152.7 (C4), 152.5 (C8), 148.4 (C2), 143.8 (C6), 133.9 (C26), 132.3 (C23), 128.6 (C25), 127.6 (C24), 122.0 (C5), 115.0 (C20), 85.4 (C1′), 83.2 (C4′), 76.0 (C3′), 69.7 (C28), 60.6 (C5′), 59.5 (C30), 57.6 (C11), 52.5 (C14), 35.3 (C2′), 29.7 (C15). HRMS (ESI$^+$) calcd for $C_{27}H_{31}N_8O_8$ 595.2264, found 595.2265.



N^6-Benzoyl-3′-*O*-methoxyacetyl-8-((4-imidazolyl)ethynyl)-2′-deoxyadenosine (3.23)

Compound **3.18** (250 mg, 0.240 mmol) was dried by co-evaporating with pyridine overnight. The nucleoside was dissolved in pyridine (2.5 ml), and methoxyacetic anhydride (150 μl, 625 μmol) was added to the solution. After 45 min of stirring at room temperature, the solvent was removed. The resulting gum was resuspended in CH_2Cl_2, washed with $NaHCO_{3(aq)}$ (5 %) and dried with Na_2SO_4 (anhydrous). Following the removal of the solvent, pyridine (1 ml) and H_2O (1 ml) were added to the crude product, and the mixture was left under vacuum until dry. The crude product was then dissolved in

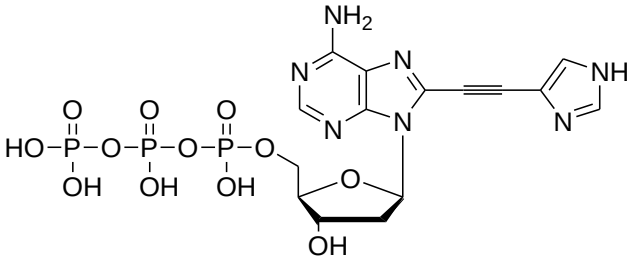
CH_2Cl_2 (5.4 ml), acetic acid (1.8 ml) and H_2O (0.2 ml). After 16 hours of stirring at room temperature, the solvent was removed. The reaction contents were again resuspended in CH_2Cl_2, washed with $NaHCO_{3(aq)}$ (5 %) and dried with Na_2SO_4 (anhydrous). Purification on silica gel chromatography (17.5×1.5 cm column) with MeOH/$CHCl_3$ (5:95) gave 70.5 mg (57 %) of a yellow solid. R_f = 0.09 (MeOH/$CHCl_3$ 1:9). ^{1}H NMR (300 MHz, $CDCl_3$, 25 °C): δ = 8.64 (s, 1H, H2), 7.96 (d, *J* = 7.4 Hz, 2H, H20), 7.59 (s, 1H, H14), 7.52 (t, *J* = 7.4 Hz, 1H, H22), 7.46 (s, 1H, H16), 7.42 (t, *J* = 7.4 Hz, 2H, H21), 6.60 (dd, *J* = 9.2, 5.6 Hz, 1H, H1′), 5.59 (d, *J* = 6.5 Hz, 1H, H3′), 4.23–4.19 (m, 1H, H4′), 4.05 (s, 2H, H24), 3.90 (d, *J* = 12.5 Hz, 1H, H5′), 3.83 (dd, *J* = 12.5, 1.6 Hz, 1H, H5′), 3.38 (s, 3H, H26), 3.14 (ddd, *J* = 14.1, 9.2, 6.5 Hz, 1H, H2′), 2.46 (dd, *J* = 14.1, 5.6 Hz, 1H, H2′). ^{13}C NMR (100 MHz, $CDCl_3$, 25 °C): δ = 169.9 (C23), 165.4 (C19), 152.3 (C2), 150.3 (C4), 149.8 (C6), 136.9 (C8), 133.2 (C19), 132.8 (C22), 128.6 (C21), 128.0 (C20), 123.4 (C5), 86.9 (C1′), (C4′), 76.2 (C3′), 69.5 (C24), 62.7 (C5′), 59.2 (C26), 36.8 (C2′). HRMS (ESI$^+$) calcd for $C_{25}H_{23}N_7O_6$ 540.1608, found 540.1611.



N^6-Benzoyl-3′-*O*-methoxyacetyl-8-((4-imidazolyl)ethyl)-2′-deoxyadenosine (3.24)

Compound **3.24** was prepared in the same way as **3.20**. Compound **3.19** (150 mg, 0.14

mmol) was converted into 45.4 mg (61 %) of a light orange solid that was eluted by silica flash chromatography (15.5×2.0 cm column) with MeOH/$CHCl_3$ (6:94) as eluent. R_f = 0.12 (MeOH/$CHCl_3$ 1:9). ^{1}H NMR (300 MHz, $CDCl_3$, 25 °C): δ = 8.64 (s, 1H, H2), 8.09 (d, *J* = 7.2 Hz, 2H, H20), 7.72 (s, 1H, H14), 7.65 (t, *J* = 7.3 Hz, 1H, H22), 7.56 (dd, *J* = 7.3, 7.2 Hz, 2H, H21), 6.89 (s, 1H, H16), 6.39 (dd, *J* = 8.6, 6.1 Hz, 1H, H1′), 5.63 (d, *J* = 6.4 Hz, 1H, H3′), 4.20–4.15 (m, 1H, H4′), 4.17 (s, 2H, H24), 3.89 (dd, *J* = 12.3, 3.1 Hz, 1H, H5′), 3.82 (dd, *J* = 12.3, 3.7 Hz, 1H, H5′), 3.46 (s, 3H, H26), 3.45–3.30 (m, 3H, H10, H2′), 3.19 (t, *J* = 6.7 Hz, 2H, H11), 2.37 (ddd, *J* = 14.0, 6.1, 1.4 Hz, 1H, H2′). ^{13}C NMR (75 MHz, $CDCl_3$, 25 °C): δ = 169.5 (C23), 165.9 (C18), 154.2 (C4), 152.7 (C8), 150.8 (C2), 148.7 (C6), 132.9 (C22), 128.6 (C21), 127.9 (C20), 123.0 (C5), 86.6 (C1′), 85.4 (C4′), 76.2 (C3′), 69.4 (C24), 62.5 (C5′), 59.2 (C26), 36.4 (C2′), 29.4 (C10), 28.1 (C11). HRMS (ESI$^+$) calcd for $C_{25}H_{27}N_7O_6$ 522.2101, found 522.2103.

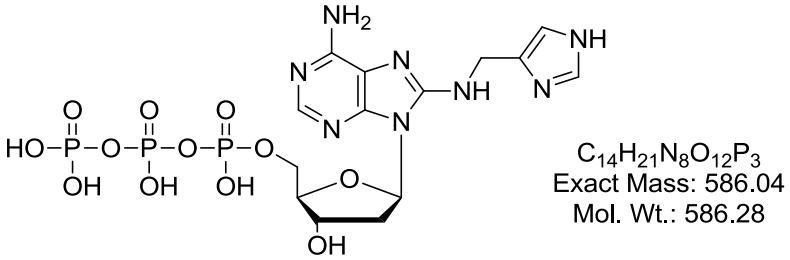


$C_{15}H_{18}N_7O_{12}P_3$
Exact Mass: 581.02
Mol. Wt.: 581.26

8-[2-(4-Imidazolyl)ethynyl]-2′-deoxyadenosine triphosphate (3.5)

Compound **3.23** (26 mg, 50 μmol) was dried by co-evaporation from pyridine under vacuum overnight in a 5 mm diameter NMR tube. The reaction was monitored by ^{31}P NMR spectroscopy. The starting material was suspended in dioxane (0.3 mL) and pyridine (0.1 mL). Following the addition of a solution of 2-chloro-4*H*-1,3,2-benzodioxaphosphorin-4-one (11.1 mg, 55 μmol) in dioxane (55 μL),

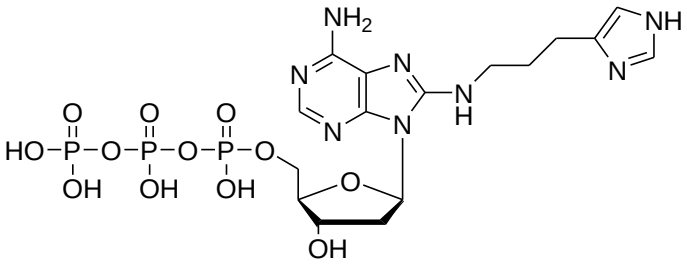
the reaction was agitated on a vortexer thoroughly and then periodically over 15 min. Bis(tri-*n*-butylammonium) pyrophosphate (35.6 mg, 75 μmol) in DMF (150 μL) and tri-*n*-butylamine (50 μL) were mixed thoroughly and added to the reaction. Following another 15 min of agitation on a vortexer, I_2 (20 mg, 79 μmol) in H_2O (20 μL) and pyridine (980 μL) was added. After 20 min of periodic agitation on a vortexer, the reaction contents were transferred to a 25 mL flask. Aq. $NaHSO_3$ (5 %) was added drop by drop until excess I_2 was quenched. Once the reaction had been evaporated to dryness, it was left to stand at room temperature in H_2O (5 mL) for 30 min. Then concentrated ammonium hydroxide (10 mL) was added. The flask was stoppered and while stirring was heated at 50 °C for 1 h. Following evaporation of the reaction, the contents were purified by preparative TLC using dioxane/H_2O/NH_4OH (6:4:1) as eluent; R_f = 0.14. The appropriate bands were eluted with H_2O and further purified by HPLC. Triphosphate **3.5** eluted at 8.0 min in Program 3A. Yield 23 %. UV: λ_{max} = 312 nm; λ_{min} = 263 nm. MS (MALDI$^+$): *m*/*z* = 582.1. ^{31}P NMR (121.5 MHz, D_2O, 25 °C): δ = –9.9 (m, 1 P, P_γ), –10.6 (m, 1 P, P_α), –22.3 (m, 1 P, P_β) ppm.



8-[(4-Imidazolyl)methylamino]-2′-deoxyadenosine triphosphate (3.2)

Compound **3.2** was synthesized in the same way as **3.5** and eluted at 5.4 min in Program 3A. Yield 8.6 %. R_f = 0.11 (dioxane/H_2O/NH_4OH 6:4:1). UV: λ_{max} = 276 nm; λ_{min} = 238

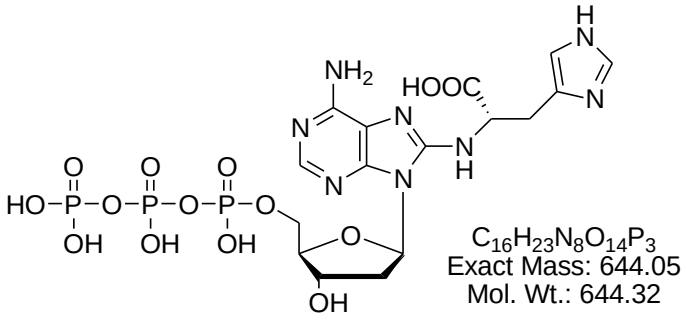
nm.MS (MALDI$^+$): *m*/*z* = 587.2. ^{31}P NMR (121.5 MHz, D_2O, 25 °C): δ = –8.3 (m, 1 P, P_γ), –10.6 (m, 1 P, P_α), –21.3 (m, 1 P, P_β) ppm.



$C_{16}H_{25}N_8O_{12}P_3$
Exact Mass: 614.08
Mol. Wt.: 614.34

8-[3-(4-Imidazolyl)propylamino]-2′-deoxyadenosine triphosphate (3.3)

Compound **3.3** was synthesized in the same way as **3.5** and eluted at 7.1 min in Program 3A. Yield 3.3 %. R_f = 0.22 (dioxane/H_2O/NH_4OH 6:4:1). UV: λ_{max} = 279 nm; λ_{min} = 239 nm. MS (MALDI$^+$): *m*/*z* = 615.1.

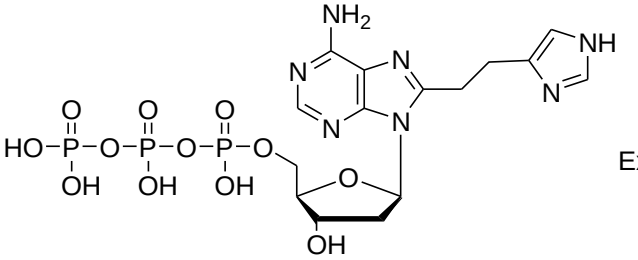


$C_{16}H_{23}N_8O_{14}P_3$
Exact Mass: 644.05
Mol. Wt.: 644.32

8-[(*S*)-1-Carboxy-2-(4-imidazolyl)ethylamino]-2′-deoxyadenosine triphosphate (3.4)

Compound **3.4** was synthesized in the same way as **3.5** except the deprotection was performed differently. After the reaction was left to stand in H_2O (5 mL) for 30 min, aq. Li_2CO_3 (0.2 M, 10 mL) was added. The reaction was stirred at 50 °C for 5.5 h and at room temperature overnight. The contents were once again dried. H_2O (2.5 mL) and concentrated aqueous ammonia (10 mL) were then used for deprotection. This was

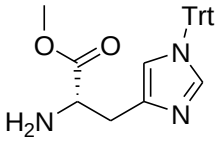
carried out at 50 °C for 3 h. Compound **3.4** eluted at 4.1 min in Program 3B. Yield 2.6 %. R_f = 0.13 (dioxane/H_2O/NH_4OH 6:4:1). UV: λ_{max} = 277 nm; λ_{min} = 239 nm. MS (MALDI$^+$): *m*/*z* = 644.9. ^{31}P NMR (121.5 MHz, D_2O, 25 °C): δ = –8.7 (m, 1 P, P_γ), –10.2 (m, 1 P, P_α), –21.6 (m, 1 P, P_β) ppm.



$C_{15}H_{22}N_7O_{12}P_3$
Exact Mass: 585.05
Mol. Wt.: 585.3

8-[2-(4-Imidazolyl)ethyl]-2′-deoxyadenosine triphosphate (3.6)

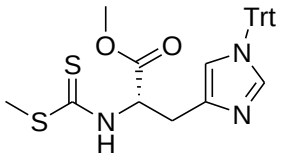
Compound **3.6** was synthesized in the same way as **3.5** and eluted at 5.6 min in Program 3A. Yield 8.3 %. R_f = 0.18 (dioxane/H_2O/NH_4OH 6:4:1). UV: λ_{max} = 262 nm; λ_{min} = 236 nm. MS (MALDI$^+$): *m*/*z* = 586.2.



$C_{26}H_{25}N_3O_2$
Exact Mass: 411.1947
Mol. Wt.: 411.4956

1-Trityl-L-histidine methyl ester (3.25)

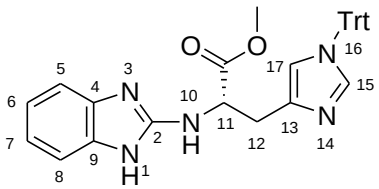
Monotritylated amino acid prepared according to Himes *et al.*[193] R_f = 0.18 (MeOH/$CHCl_3$ 1:9). ^{1}H NMR (400 MHz, $CDCl_3$, 25 °C): δ = 7.38–7.32 (m, 10H), 7.18–7.11 (m, 6H), 6.58 (s, 1H), 3.76 (dd, *J* = 6.8, 4.8 Hz, 1H), 3.60 (s, 3H), 2.92 (dd, *J* = 14.4, 4.8 Hz, 1H), 2.81 (dd, *J* = 14.4, 6.8 Hz, 1H), 1.96 (s, br, 2H). MS (ESI$^+$) 434.2 $[M+Na]^+$.



$C_{28}H_{27}N_3O_2S_2$
Exact Mass: 501.1545
Mol. Wt.: 501.6629

***N*-[(Methylthio)thioxomethyl]-1-trityl-L-histidine (3.26)**

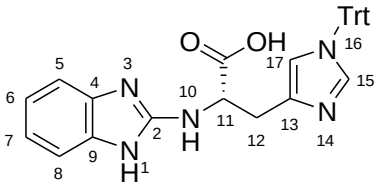
Compound **3.25** (823 mg, 2.0 mmol) was dissolved in DMF (10 mL) to give a yellow solution. Upon the addition of Et_3N (290 µL, 2.1 mmol) and then carbon disulfide (127 µL, 2.1 mmol), the solution became orange. The reaction was stirred at room temperature for 100 min. Methyl iodide (137 µL, 2.2 mmol) was added, and the mixture was heated to 65 °C. The heating of this brown solution was maintained for 70 min. At this time, the reaction was evaporated to dryness, redissolved in CH_2Cl_2, washed with $NaHCO_{3(aq)}$ (5 %) and dried over sodium sulfate (anhydrous). The product was purified by flash chromatography (14×3.0 cm column) and eluted in MeOH/$CHCl_3$ (1:49) to give 888 mg (89 %) of a yellow-orange oil. This oil solidified after being stored at room temperature for about a week. R_f = 0.71 (MeOH/$CHCl_3$ 1:9). ^{1}H NMR (400 MHz, CD_2Cl_2, 25 °C): δ = 9.44 (d, J = 6.2 Hz, 1H, NH), 7.41 (s, 1H, Imid-H2), 7.40–7.29 (m, 9H, PheH), 7.17–7.08 (m, 6H, PheH), 6.59 (s, 1H, Imid-H5), 5.40–5.31 (m, 1H, H_α), 3.59 (s, 3H, OCH_3), 3.08 (d, J = 4.5 Hz, 2H, H_β), 2.61 (s, 3H, SCH_3). HRMS (ESI$^+$) calcd for $C_{28}H_{27}N_3O_2NaS_2$ 524.1442, found 524.1446.



$C_{33}H_{29}N_5O_2$
Exact Mass: 527.2321
Mol. Wt.: 527.6157

2-((*S*)-1-Methylcarboxy-2-(1-trityl-(4-imidazolyl))ethylamino)benzimidazole (3.27)

A solution of **3.26** (1.03 g, 2.0 mmol) and mercuric oxide (red) (0.43 g, 2 mmol) in DMF (21 mL) was slowly added phenylenediamine (0.22 g, 2 mmol). The mixture was heated to 50 ℃. Shortly after heating, a greenish yellow precipitate began to appear. After 3 hr, the reaction was cooled to room temperature. The mixture was filtered and purified on silica flash chromatography (15×2.0 cm column). The product eluted in EtOAc/hexanes (9:1) to give 674 mg (64 %) of a yellow foam. TLC of this material revealed small amounts of an impurity. The product was again purified on silica gel (14×2.0 cm column) using the same solvent system to give 602 mg (57 %) of a yellow foam. R_f = 0.47 (MeOH/$CHCl_3$ 1:9). ^{1}H NMR (400 MHz, CD_2Cl_2, 25 ℃): δ = 7.44 (s, 1H, H15), 7.38–7.02 (m, 17H, H5, H8, Phe-H), 7.02–6.94 (m, 2H, H6, H7), 6.62 (s, 1H, H17), 4.89–4.81 (m, 1H, H11), 3.61 (s, 3H, CH_3), 3.14 (dd, J = 14.7, 4.0 Hz, 1H, H12), 3.04 (dd, J = 14.7, 6.2 Hz, 1H, H12). HRMS (ESI$^+$) calcd for $C_{33}H_{29}N_5O_2Na$ 550.2216, found 550.2213.

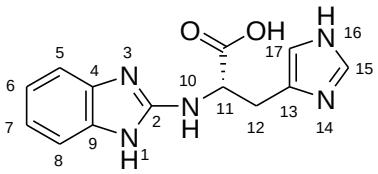


$C_{32}H_{27}N_5O_2$
Exact Mass: 513.2165
Mol. Wt.: 513.5891

2-((*S*)-1-Carboxy-2-(1-trityl-(4-imidazolyl))ethylamino)benzimidazole (3.28)

To a solution of **3.27** (106 mg, 0.20 mmol) in MeOH (4 mL) was added $NaOH_{(aq)}$ (0.20 mL, 1.0 M) and H_2O (1.8 mL). After this mixture was stirred at room temperature for 3 hr, it was diluted with H_2O and extracted with CH_2Cl_2. The organic extracts were dried over Na_2SO_4 (anhydrous). Purification on silica flash chromatography (19×1.0 cm column)

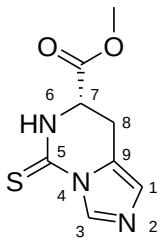
isolated 66 mg (64 %) of a white powder that eluted in MeOH/$CHCl_3$ (3:47). R_f = 0.07 (MeOH/$CHCl_3$ 1:9). ^{1}H NMR (400 MHz, CD_2Cl_2, 25 °C): δ = 10.15–10.00 (s, br, 1H, OH), 7.57 (s, 1H, H15), 7.39–7.22 (m, 11H, H5, H8, Phe-H), 7.19–7.05 (m, 8H, H6, H7, Phe-H), 6.77 (s, 1H, H17), 4.48–4.40 (m, 1H, H11), 3.55–3.46 (m, 1H, H12), 2.96–2.85 (m, 1H, H12). HRMS (ESI$^+$) calcd for $C_{32}H_{28}N_5O_2$ 514.2243, found 514.2239.



$C_{13}H_{13}N_5O_2$
Exact Mass: 271.1069
Mol. Wt.: 271.2746

2-((*S*)-1-Carboxy-2-(4-imidazolyl)ethylamino)benzimidazole (3.29)

Compound **3.28** (387 mg, 0.75 mmol) was dissolved in trifluoroacetic acid (20 mL) and stirred at room temperature for 40 min. The reaction was concentrated and resuspended in a mixture of H_2O and CH_2Cl_2. The liquids were neutralized with ammonium hydroxide (2 %). Concentration of the aqueous layer gave 0.729 g of a white solid. This white solid was suspended in H_2O (5.5 mL) and centrifuged. The resulting white pellet (260 mg) was recrystallized in a mixture of EtOH and H_2O to give 96 mg (33 %) of very tiny white needles. NMR revealed that the product was isolated as the trifluoroacetamide salt. ^{1}H NMR (400 MHz, DMSO, 25 °C): δ = 8.31 (s, 1H, H15), 7.30–7.21 (m, 2H, H5, H8), 7.17 (s, 1H, H17), 7.07–7.01 (m, 2H, H6, H7), 4.67 (dd, *J* = 7.9, 4.9 Hz, 1H, H11), 3.24 (dd, *J* = 15.1, 7.9 Hz, 1H, H12). ^{13}C NMR (300 MHz, DMSO, 25 °C): δ = 172.2 (CO), 152.2 (C1), 134.3 (C15), 133.7 (C3, C8), 131.5 (C13), 121.4 (C5, C6), 116.4 (C17), 111.6 (C4, C7), 55.2 (C11), 27.8 (C12). HRMS (ESI$^+$) calcd for $C_{13}H_{14}N_5O_2$ 272.1137, found 272.1139. λ_{max} = 276 nm (ε = 9800 $cm^{-1}M^{-1}$).



$C_8H_9N_3O_2S$
Exact Mass: 211.0415
Mol. Wt.: 211.241

(*S*)-5,6,7,8-Tetrahydro-5-thioimidazo[1,5-c]pyrimidine-7-carboxylic acid methyl ester (3.31)

^{1}H NMR (300 MHz, CD_2Cl_2, 25 °C): δ = 8.44 (s, 1H, H3), 7.82 (s, br, 1H, H6), 6.89 (s, 1H, H1), 4.31 (dd, *J* = 10.8, 5.1 Hz, 1H, H7), 3.85 (s, 3H, CH_3), 3.43 (dd, *J* = 15.7, 5.1 Hz, 1H, H8), 3.17–3.04 (m, 1H, H8′). HRMS (ESI^+) calcd for $C_8H_{10}N_3O_2S$ 212.0494, found 212.0492.

3.6.2 Incorporations

Primer extensions were performed in the same manner as Chapter 2.

DNA templates: (5′ to 3′, italicized residues correspond to the primer binding regions) **P2.1** ATTAGCCCTTCCAGTCCCCCCTTTTCTTTT. **T3.1** GGAGCTGTAGATCTTAGTTACTGGCC*AAAAGAAAAGGGGGGACTGGAAGGGCTAA*. **T3.2** GGAGCTGTCCATCTTAGTTACTTTTT*AAAAGAAAAGGGGGGACTGGAAGGGCTAA*. **T3.3** GCAGCTGTAGATCTTAGCCAGGCCTT*AAAAGAAAAGGGGGGACTGGAAGGGCTAA*. **T3.4** GGAGCTGTCCATCACATTTTTTTTTT*AAAAGAAAAGGGGGGACTGGAAGGGCTAA*.

Chapter 4: The incorporation of guanidinium-modified oligonucleotides into duplex structures and the substrate binding regions of a DNAzyme

4.1 Introduction

A great deal of research has been devoted towards oligonucleotides due to their promise as therapeutic agents: aptamers are being designed to target many different proteins; small interfering RNA (siRNA),[195,196] antisense,[120,121] ribozyme and DNAzyme strategies are being used to regulate gene expression; and triplex forming oligonucleotides are being applied to the inhibition of mRNA transcription.[13,15,197-199] In the pursuit of oligonucleotides as therapeutic agents, many efforts have been devoted towards the synthesis of modified oligonucleotides of increased cellular stability, enhanced target affinity, superior cellular uptake, and in certain situations, improved catalytic activity. One particular modification of oligonucleotides that can potentially satisfy all of the above objectives is the guanidinium group. Guanidinium groups have been added to the sugar,[143] nucleobase[142,200] and phosphate[201] of oligonucleotides. The addition of guanidinium cations to oligonucleotides can increase cellular uptake by reducing their anionic nature[200] and is thought to mimic the function of arginines in many arginine-rich CPP's.[144] *In vitro* selections have exploited the use of guanidinium modifications as well. Oligonucleotides with guanidinium-modified nucleobases have given rise to Mg^{+2}-independent DNAzymes of high activity[137,138] and a glutamate-binding aptamer.[145]

As discussed in Chapter 1, DNAzymes such as 8-17 and 10-23 are presently being explored for their ability to degrade any given sequence of mRNA *in vivo*.[111,112] In an

effort to enhance the target affinity and *in vivo* stability of DNAzymes, research has been focused on the chemical modification of the substrate binding regions, which generally does not affect catalysis. Vester *et al.* have substituted locked nucleic acids (LNA), residues that have been shown to increase the thermostability of helical structures, into the substrate binding regions of DNAzyme 10-23.[123] They discovered that the LNA-containing DNAzymes were more catalytically active than their unmodified counterparts. LNA was subsequently used for the synthesis of 8-17 as well.[125] Other groups have explored the use of phosphorothioates, 2′-OMe sugars[124] as well as appended intercalators[126] with DNAzyme 10-23.

In order to address the effect of guanidinium groups in terms of target affinity and DNAzyme 10-23 catalysis, we have 1) synthesized and characterized a modified 2′-deoxyuridine phosphoramidite (**4.1**, Figure 4.1) that contains a guanidinium group that is attached at the 5-position, 2) studied its effect on the stability of oligonucleotide duplex structures and 3) carried out kinetic assays on RNA-cleaving DNAzyme 10-23[27] variants that were fabricated with the newly made phosphoramidite and contained substrate binding regions that targeted hepatitis C virus (HCV) RNA.[113,202] The target molecule **4.1** is the phosphoramidite of **2.11**, a modified nucleoside triphosphate that was successfully used in two separate *in vitro* selection experiments that gave rise to self-cleaving DNAzymes in the complete absence of Mg^{+2}.[137,138] The modification is attached to the non Watson-Crick face of the nucleobase such that the base pairing and sugar-phosphate backbone are not affected. These properties facilitate the incorporation of modified nucleoside triphosphates by DNA polymerases for the synthesis of modified nucleic acid libraries. Earlier research on oligonucleotides containing guanidinium modifications have shown that duplex stability is enhanced.[142,143] Our modified oligonucleotides should

behave similarly and hence are hypothesized to enhance catalysis by facilitating enhanced substrate affinity, especially in low Mg^{+2} (physiological ~0.5 mM)[77-79] and low substrate environments where the unmodified 10-23 is not particularly active.[46]

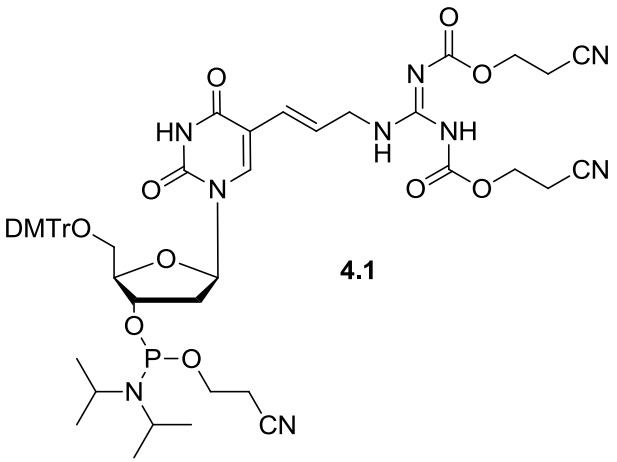
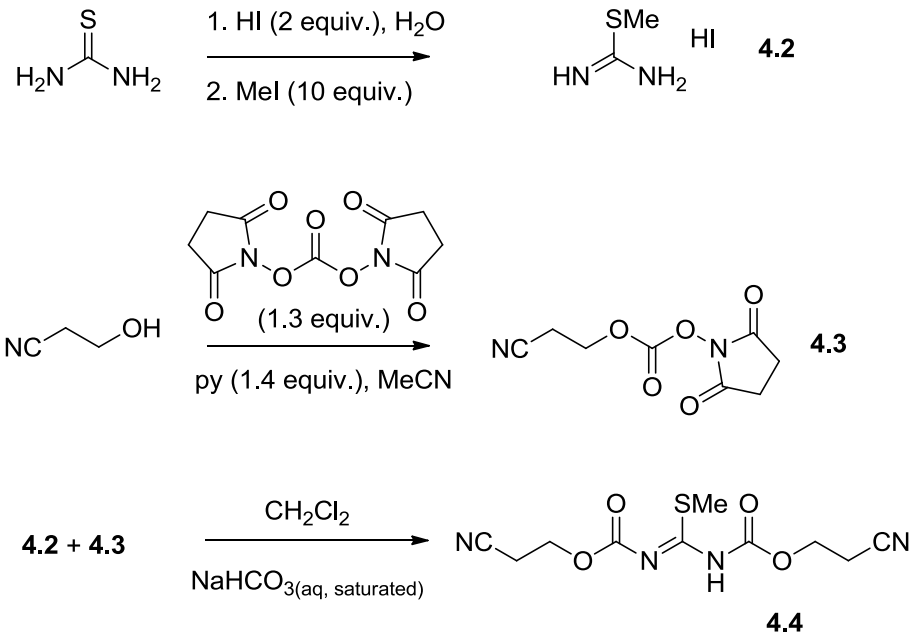


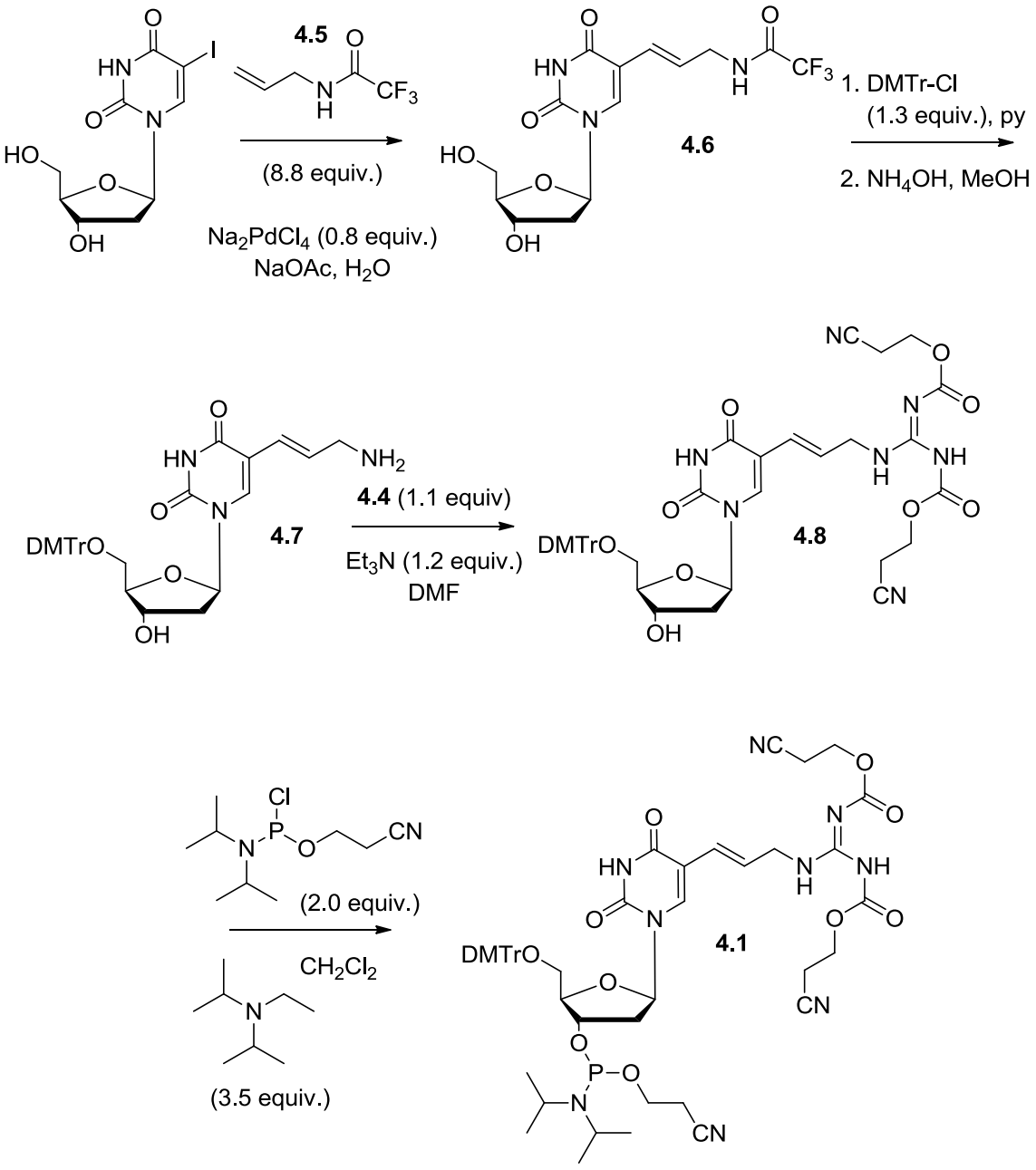
Figure 4.1 Guanidine-modified 2′-deoxyuridine phosphoramidite.

4.2 Synthesis

The synthesis of the desired phosphoramidite began with the guanidinylating agent, which was synthesized according to the literature as shown in Scheme 4.1. Thiourea was protonated with hydroiodic acid and then methylated with iodomethane to give *S*-methylisothiourea (**4.2**).[203] 2-Cyanoethanol was reacted with *N*,*N*′-disuccinimidyl carbonate[204] in the presence of pyridine to give the NHS-ester **4.3**.[205] Two equivalents of this product were reacted with **4.2** under basic conditions to form **4.4**.[143]



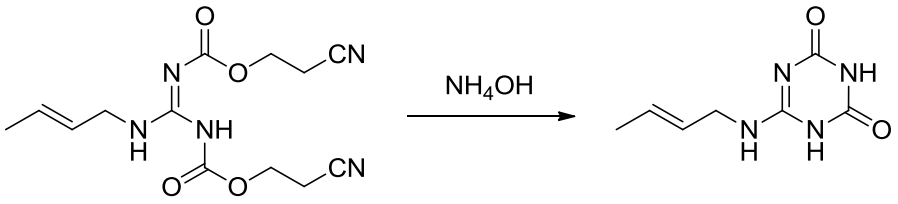
Scheme 4.1 Guanidinylating agent synthesis.



Scheme 4.2 Synthesis of phosphoramidite **4.1**.

The rest of the phosphoramidite synthesis is highlighted in Scheme **4.2**. The preparation of **4.7** was done according to the literature. In a buffered environment (pH = 5.2), commercially available 5-iodo-2′-deoxyuridine was coupled with 3-trifluoroacetylamino-1-propene (**4.5**, prepared according to Wlassoff *et al.*[206]) in the

presence of nearly stoichiometric amounts of sodium tetrachloropalladate (II) to give **4.6**.[83] This nucleoside was protected with a 4,4′-dimethoxytrityl group,[207] and the trifluoroacetyl group was removed with ammonium hydroxide.[171] Compound **4.7** was then reacted with **4.4** to give the guanidinylated nucleoside **4.8**. The final step of the synthesis involved phosphitylation in the presence of diisopropylethylamine and CH_2Cl_2. Using standard solid phase techniques, **4.1** was successfully incorporated into oligonucleotides. The deprotection of the modified DNA product required a 24 h treatment with piperidine to first remove the 2-cyanoethanol groups prior to routine deprotection with ammonium hydroxide. This extra step was necessary to avoid the possibility of the formation of a triazinyl byproduct (Scheme 4.3).[208] MALDI-TOF MS data for the modified oligonucleotides used for melting temperature studies are shown in Table 4.1.



Scheme 4.3 Triazine byproduct formation.

	Calculated	Found
O4.2	5131	5132
O4.3	5214	5214

Table 4.1 MALDI-TOF MS data for modified melting temperature oligonucleotides (*m/z*). Data was collected in negative ion mode.

4.3 Duplex stability

The helical stability studies were performed with the oligonucleotides shown in Table 4.2, and the results are shown in Tables 4.3 and 4.4. The modified DNA oligonucleotides **O4.2** and **O4.3** were annealed to unmodified complementary RNA (**O4.4**) or DNA (**O4.5**). The double helices were then slowly heated up and cooled down while the absorbance at 260 nm was being recorded. As expected, the modifications increased duplex stability. Melting temperatures for both DNA/RNA and DNA/DNA helices were increased by ~2.3 and 2.5 ℃ per modified nucleotide, respectively. These results were in agreement with studies that were done on closely related oligonucleotides.[142]

O4.1	d-TTC TTT TTC TTC TCT TT
O4.2	d-TTC TTT TTC TTC TCT MT
O4.3	d-TTC TTM TTC TTC TCT MT
O4.4	r-CUC AAA GAG AAG AAA AAG AAC U
O4.5	d-CTC AAA GAG AAG AAA AAG AAC T
O4.6	d-CTC A*C*A GAG AAG AAA AAG AAC T
O4.7	d-CTC A*G*A GAG AAG AAA AAG AAC T

Table 4.2 Oligonucleotides used for melting temperature studies (5′–3′). Guanidinium-containing 2′-deoxyuridines are indicated with M.

Oligonucleotide duplex				T_m	Description
O4.1	5′	TTCTTTTTCTTCTCTTT	3′	49.3	DNA/RNA
O4.4	3′	UCAAGAAAAAGAAGAGAAACUC	5′		
O4.1	5′	TTCTTTTTCTTCTCTTT	3′	40.8	DNA/DNA
O4.5	3′	TCAAGAAAAAGAAGAGAAACTC	5′		
O4.1	5′	TTCTTTTTCTTCTCTTT	3′	36.3	DNA/DNA
O4.6	3′	TCAAGAAAAAGAAGAGA*C*ACTC	5′		1 mismatch
O4.1	5′	TTCTTTTTCTTCTCTTT	3′	37.0	DNA/DNA
O4.7	3′	TCAAGAAAAAGAAGAGA*G*ACTC	5′		1 mismatch
O4.2	5′	TTCTTTTTCTTCTCT**M**T	3′	51.5	DNA/RNA
O4.4	3′	UCAAGAAAAAGAAGAGAAACUC	5′		
O4.2	5′	TTCTTTTTCTTCTCT**M**T	3′	43.3	DNA/DNA
O4.5	3′	TCAAGAAAAAGAAGAGAAACTC	5′		
O4.2	5′	TTCTTTTTCTTCTCT**M**T	3′	37.7	DNA/DNA
O4.6	3′	TCAAGAAAAAGAAGAGA*C*ACTC	5′		1 mismatch
O4.2	5′	TTCTTTTTCTTCTCT**M**T	3′	38.4	DNA/DNA
O4.7	3′	TCAAGAAAAAGAAGAGA*G*ACTC	5′		1 mismatch
O4.3	5′	TTCTT**M**TTCTTCTCT**M**T	3′	53.9	DNA/RNA
O4.4	3′	UCAAGAAAAAGAAGAGAAACUC	5′		
O4.3	5′	TTCTT**M**TTCTTCTCT**M**T	3′	45.8	DNA/DNA
O4.5	3′	TCAAGAAAAAGAAGAGAAACTC	5′		

Table 4.3 Oligonucleotide duplex melting temperatures. Data measured in 100 mM NaCl, 10 mM sodium phosphate (pH = 7.0) and 1 μM of each strand. Absorbance at 260 nm measured. Numerical data from 4 separate experiments of forward and reverse melting temperatures were averaged. Standard deviation: 0.4 ℃ (based on the observed melting temperature of the duplex **O4.1/O4.5**). **M** – guanidine-modified 2′-deoxyuridine.

	O4.4 RNA		**O4.5** DNA		**O4.6** DNA 1 mismatch		**O4.7** DNA 1 mismatch	
	T_m (℃)	ΔT_m	T_m (℃)	ΔT_m	T_m (℃)	ΔT_m	T_m (℃)	ΔT_m
O4.1	49.3		40.8		36.3		37.0	
O4.2	51.5	+2.2	43.3	+2.5	37.7	+1.4	38.4	+1.4
O4.3	53.9	+4.6	45.8	+5.0	NE		NE	

Table 4.4 Summary of duplex melting temperatures from Table 4.3. Data measured in 100 mM NaCl, 10 mM sodium phosphate (pH = 7.0) and 1 μM of each strand. Absorbance at 260 nm measured. Numerical data from 4 separate experiments of forward and reverse melting temperatures were averaged. Standard deviation: 0.4 ℃ (based on the observed melting temperature of the duplex **O4.1/O4.5**). NE - T_m not evaluated.

In order to test the base pairing specificity of our guanidinium-modified dU, a mismatch base pair was introduced into the DNA/DNA helices. Oligonucleotide **O4.2** was annealed to two additional DNA oligonucleotides **O4.6** and **O4.7**, which contain a C and G, respectively, in place of the A that base pairs with the guanidinium-modified dU. As expected, the melting temperatures of the resulting mismatched duplexes were significantly reduced. Compared to mismatched duplexes with unmodified residues, it was found that the mismatched duplexes with modified residues were destabilized more substantially. These results suggest that the guanidinium-modified 2′-deoxyuridine not only retains proper Watson-Crick base pairing interactions, but also exhibits increased base pairing specificity with dA, at least for the specific location of the oligonucleotides that were tested.

4.4 Kinetic studies of DNAzyme 10-23 variants

Kinetic studies were performed using 10-23 DNAzyme variants that contained 2, 4, or 6 guanidinium residues as shown in Figure 4.2. The design of these DNAzyme

variants was based on the unmodified DNAzyme **O4.9**. Previously, this DNAzyme was used for the *in vitro* cleavage at position 699 of HCV RNA spanning the 5′-untranslated region-core genomic position 1–976, and a similar but more active DNAzyme was used for the *in vivo* cleavage of the same HCV RNA as well.[113] With these studies in mind, DNAzymes of the same length containing modified residues in either the 5′-substrate binding region, 3′-substrate binding region or both substrate binding regions were synthesized.

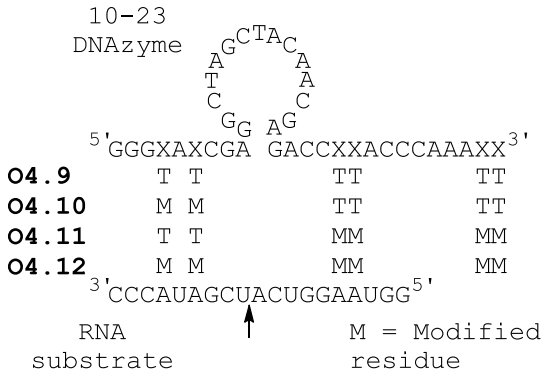


Figure 4.2 Sequences of the 10-23 DNAzyme variants and the RNA substrate (**O4.8**). The site of cleavage is noted with an arrow.

To study the effect of the addition of modifications in the substrate binding regions, multiple turnover experiments were performed (Figure 4.3). Under physiological conditions (e.g. 0.5 mM Mg^{+2}, 150 mM NaCl, pH 7.4, 37 ℃), experiments were performed under two concentration pairs of 19 ribonucleotide substrate (**O4.8**) and DNAzyme: 5 μM and 0.25 μM or 50 nM and 5 nM, respectively. The higher substrate concentration was chosen because it should be greater than the K_M of the DNAzymes used.[27,46] The lower substrate concentration reflects conditions where substrate binding

may limit catalysis such that modified catalysts with lower dissociation rates could improve catalysis. As can be seen in Table 4.5, all of the modified DNAzymes exhibited reduced catalytic rates. The presence of guanidinium groups in the 3′ substrate binding regions was particularly poorly tolerated as the catalytic rates of both DNAzymes **O4.11** and **O4.12** were significantly inferior to that of the unmodified control. The addition of guanidinium groups to the 5′ substrate binding region resulted in only a slight reduction in catalysis.

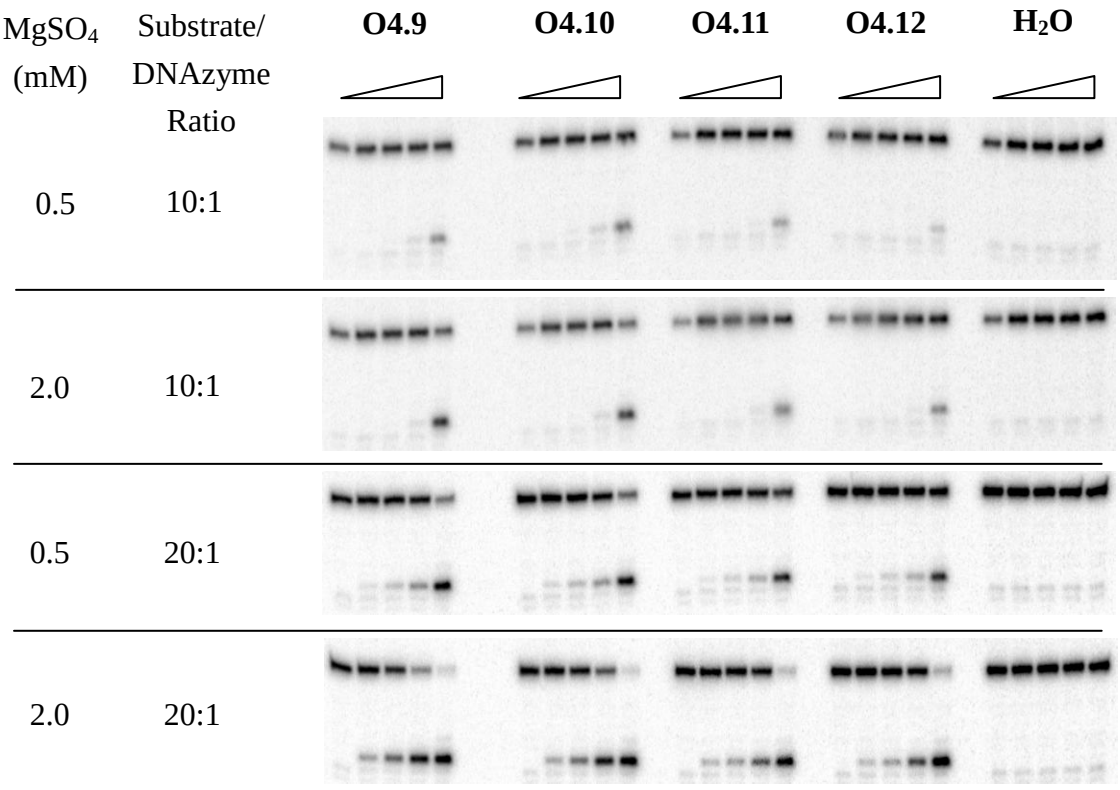


Figure 4.3 Denaturing PAGE (20 %) of DNAzyme 10-23 cleavage reactions. Four DNAzymes and a H_2O control were incubated with RNA substrate (**O4.8**) under two concentrations of $MgSO_4$, 0.5 mM and 2 mM. Final concentrations of the catalyst and substrate were either 5 and 50 nM or 0.25 μM and 5 μM, respectively. Kinetic reactions were performed in 150 mM NaCl, 50 mM Tris-HCl (pH 7.5), 37 ℃. Aliquots were removed and quenched at 0, 0.5, 1.5, 4.5, 25 h with increasing time indicated by the triangle above each set of lanes.

Mg^{+2} (mM)	0.5		2.0	
Substrate/DNAzyme ratio	10:1	20:1	10:1	20:1
DNAzyme **O4.9**	$(1.5 \pm 0.1) \times 10^{-3}$	$(11 \pm 1) \times 10^{-3}$	$(5 \pm 1) \times 10^{-3}$	$(5.2 \pm 0.9) \times 10^{-2}$
DNAzyme **O4.10**	$(1.5 \pm 0.3) \times 10^{-3}$	$(9 \pm 2) \times 10^{-3}$	$(4 \pm 2) \times 10^{-3}$	$(3.9 \pm 0.9) \times 10^{-2}$
DNAzyme **O4.11**	$(0.7 \pm 0.1) \times 10^{-3}$	$(8 \pm 4) \times 10^{-3}$	$(2 \pm 1) \times 10^{-3}$	$(2.1 \pm 0.1) \times 10^{-2}$
DNAzyme **O4.12**	$*(0.3 \pm 0.1) \times 10^{-3}$	$(3.6 \pm 0.8) \times 10^{-3}$	$(1.1 \pm 0.3) \times 10^{-3}$	$(1.6 \pm 0.3) \times 10^{-2}$

Table 4.5 DNAzyme rate constants (min^{-1}) from multiple turnover experiments. Reaction conditions: 150 mM NaCl, 50 mM Tris-HCl (pH 7.5), 37 ℃. Ratio of substrate to product evaluated with densitometry. Rate constants were obtained by first dividing the ratio of cleaved product to original substrate at certain time points of the reaction by time and then multiplying by the substrate to catalyst ratio. The asterisk denotes a value that was obtained as an average of 2 experiments. All other values were averages of 3 experiments performed on different days. The errors shown are the standard deviations of the rate constants obtained.

In addition to what is shown in Figure 4.3, four other experiments were performed (gel images included in the appendix). Two of these experiments involved the use of substrates of different lengths: a shorter, 17-nucleotide RNA substrate and a longer, 25 nucleotide RNA substrate that was designed to hybridize to the entire substrate binding region of the DNAzymes. Compared to the 19-nucleotide substrate (**O4.8**), the shorter substrate was cleaved much less effectively by unmodified or modified catalysts because the substrate was not long enough to form catalyst-substrate complexes of sufficient stability. The longer substrates were also cleaved less effectively. This was due to product release inhibition as was observed by others.[46] The third experiment analyzed catalysis under the very low substrate concentration of 25 nM. Under multiple turnover reactions, where the Mg^{+2} concentration was varied from a physiologically relevant value of 0.5 mM[77-79,110] to a non-physiological value of 10.0 mM, none of the modified catalysts was

superior to the unmodified control.

Finally, since none of our modified catalysts displayed superior kinetic activity compared to the unmodified counterpart, melting temperature measurements of the dissociation of the 10-23 variants with the 19 nucleotide RNA substrate (**O4.8**) were also performed. These determinations were performed in either a Mg^{+2} concentration of 0 or 0.5 mM. In the absence of Mg^{+2}, no cleavage of the substrate is observed over the course of the experiment; however, T_m values under these conditions may not correctly reflect the active catalyst as the divalent metal cation plays a role in the proper folding of such catalytic motifs. At a Mg^{+2} concentration of 0.5 mM, substrate cleavage is negligible over the course of the T_m assay. The T_m results under the two different conditions were quite similar and are summarized in Table 4.6. DNAzyme **O4.10** was somewhat unusual as the **O4.10**/substrate duplexes were slightly less stable than the unmodified control; however, the more heavily modified **O4.11** and **O4.12** gave rise to duplexes that were significantly more stable. These T_m results reinforce the fact that our modified residue enhances DNAzyme target affinity.

	Without Mg^{+2}	With Mg^{+2} (0.5 mM)
O4.9	44.8	44.9
O4.10	43.6	44.4
O4.11	48.6	49.4
O4.12	50.7	51.3

Table 4.6 Melting temperatures for DNAzyme/substrate (19 ribonucleotide, **O4.8**) duplexes. Data measured in 100 mM NaCl, 10 mM sodium phosphate (pH = 7.0) and 1 μM of each strand. Absorbance at 260 nm measured. T_m values in the absence of Mg^{+2} represent the forward melting temperatures of 1 experiment. T_m values in the presence of Mg^{+2} represent the average of 1 pair of forward and reverse experiments. Standard deviation: 0.7 ℃ (based on melting temperature values from **O4.9** with Mg^{+2}.

4.5 Discussion and conclusions

Even though the modified catalysts exhibited enhanced target affinity, they did not show any catalytic improvement. A possible explanation for this is that the modification alleviates the need for Mg^{+2} for substrate binding but does not reduce the need for Mg^{+2} in terms of catalysis. As is summarized in Table 4.5, all of the guanidinium-modified catalysts were kinetically less active under the conditions tested. The rate constants that were obtained were reduced by roughly a factor of two with increasing modification. These minor changes in rates suggest the modified catalysts exhibit generally uniform binding between the transition state and the ground state ($\Delta\Delta G^{\ddagger}$ remains largely unchanged, <1 kcal/mol). In fact, the modified substrate binding in the ground state may be slightly preferred over the modified substrate binding in the transition state. This would account for the decrease in catalytic rates.

To summarize, a guanidinium-containing 2′-deoxyuridine phosphoramidite was synthesized. It was used for the solid phase synthesis of DNA to evaluate helical stability and DNAzyme catalysis. Both DNA/RNA and DNA/DNA duplexes were found to be stabilized by the modifications. Although none of the modified DNAzymes studied were able to enhance catalysis, they were all found to be catalytically active. Hence, 10-23 DNAzymes have been successfully modified with residues that can potentially enhance membrane permeability. A perhaps more promising path in the search of DNAzymes of high catalytic ability and cellular uptake would be to select modified catalysts *de novo*[69,70,109,209] in the absence of Mg^{+2}, particularly DNAzymes that employ guanidinium cations, imidazoles and cationic amines.[137,138] Finally, the phosphoramidite **4.1** will also be used in the future to evaluate the multiple turnover cleavage of *in vitro* selected DNAzymes that have been obtained with the triphosphate version of this nucleotide.

4.6 Experimental

General. Reactants and reagents were obtained from either Fisher Scientific or Sigma-Aldrich. Solvents were dried either by using molecular sieves (4 °) or distillation. Thin layer chromatography was carried out on silica gel 60 F_{254} glass backed plates from EMD Chemicals, and flash chromatography was carried out on silica (230-400 mesh) from Silicycle. Silica column dimensions are shown as the length multiplied by the diameter. NMR experiments were done using either a Bruker AV-300 or AV-400 spectrometer. ^{1}H and ^{13}C NMR spectra were referenced to the signal of the solvent. Phosphoric acid was used as an external reference for ^{31}P NMR. ESI-MS was performed on a Waters/Micromass LCT, and MALDI-TOF MS was performed on a Bruker Biflex instrument. MALDI was performed in the same way as Chapter 2, and calibrations were performed using the most abundant isotopic ion. Modified oligonucleotides were synthesized by the University Core DNA Services at the University of Calgary, and unmodified oligonucleotides were synthesized by Integrated DNA Technologies. The modified oligonucleotides were synthesized with a 6-carboxyfluorescein group at the 5′ end for future use in *in vivo* fluorescence experiments. UV spectrometry was performed on a Beckman Coulter DU800 spectrophotometer.

4.6.1 Synthesis

_S_-Methylisourea Hydroiodide (4.2)[203]

^{1}H NMR (400 MHz, $CDCl_3$, 25 °C): δ = 8.91 (s, br, 4H), 2.55 (s, 3H).

_N_-(2-Cyanoethoxycarbonyloxy)succinimide (4.3)[205]

^{1}H NMR (300 MHz, $CDCl_3$, 25 °C): δ = 4.54 (t, J = 4.5 Hz, 2H), 2.83 (s, 6H). MS (ESI^+)

235.2 $[M+Na]^+$.

_N,N′_-Bis-(2-cyanoethoxycarbonyloxy)-2-methyl-2-thiopseudourea (4.4)[208]

R_f = 0.66 (1:9 MeOH/$CHCl_3$). ^1H NMR (300 MHz, $CDCl_3$, 25 °C): δ = 4.41–4.33 (m, 4H), 2.78 (t, J = 6.3 Hz, 4H), 2.43 (s, 3H). MS (ESI^+) 285.2 $[M+H]^+$.

3-Trifluoroacetylamino-1-propene (4.5)[206]

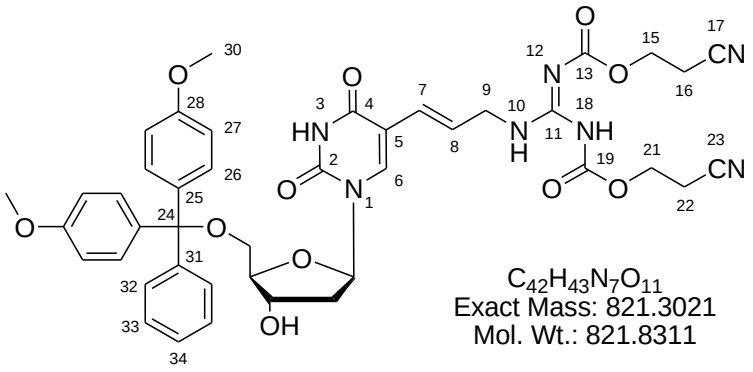
^1H NMR (300 MHz, $CDCl_3$, 25 °C): δ = 6.34 (s, br, 1H), 5.95–5.75 (m, 1H), 5.30–5.20 (m, 2H), 4.00–3.90 (m, 2H). MS (ESI^+) 176.1 $[M+Na]^+$.

5-(3-Trifluoroacetylamino-1-propenyl)-2′-deoxyuridine (4.6)[83]

R_f = 0.19 (1:9 MeOH/$CHCl_3$). ^1H NMR (400 MHz, CD_3OD, 25 °C): δ = 8.10 (s, 1H), 6.47–6.39 (m, 1H), 6.27–6.18 (m, 2H), 4.37–4.30 (m, 1H), 3.92–3.88 (m, 2H), 3.88–3.84 (m, 1H), 3.78–3.65 (m, 2H), 2.27–2.13 (m, 2H). MS (ESI^+) 402.1 $[M+Na]^+$.

5′-_O_-(4,4′-Dimethoxytrityl)-5-(3-amino-1-propenyl)-2′-deoxyuridine (4.7)[171]

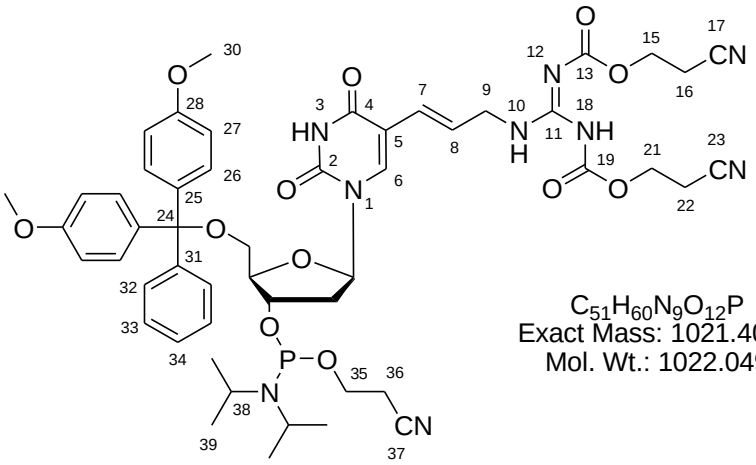
R_f = 0.05 (1:9 MeOH/$CHCl_3$). ^1H NMR (300 MHz, $CDCl_3$, 25 °C): δ = 7.65 (s, 1H), 7.43 (d, J = 8.0 Hz, 2H), 7.34–7.20 (m, 7H), 6.84 (d, J = 8.9 Hz, 4H), 6.39–6.28 (m, 2H), 5.67 (d, J = 5.92 Hz, 1H), 4.58–4.52 (m, 1H), 4.06–4.01 (m, 1H), 3.77 (s, 6H), 3.44–3.31 (m, 2H), 3.04–2.99 (m, 2H), 2.44–2.25 (m, 2H). MS (ESI^+) 608.3 $[M+Na]^+$.



5′-*O*-(4,4′-Dimethoxytrityl)-5-(3-(*N*,*N*′-bis-(2-cyanoethoxycarbonyloxy)guanidinyl)-1-propenyl)-2′-deoxyuridine (4.8)

Nucleoside **4.7** (381 mg, 0.65 mmol) was added to a solution of the guanidinylating agent **4.4** (203 mg, 0.72 mmol) in DMF (3 mL). This mixture was quickly added Et_3N (109 µL, 0.78 mmol) and stirred at room temperature. After 2.5 hr, the reaction was evaporated to dryness, resuspended in CH_2Cl_2, washed with $NaHCO_{3(aq)}$ (5%) and dried over sodium sulfate (anhydrous). Purification was performed on silica flash chromatography. The desired product eluted in MeOH/$CHCl_3$ (1:24) to give 476 (89 %) mg of a white foam. R_f = 0.36 (1:9 MeOH/$CHCl_3$). ^{1}H NMR (400 MHz, CD_2Cl_2, 25 °C): δ = 11.68 (s, 1H, H18), 8.84 (s, br, 1H, H3), 8.12 (m, 1H, H10), 7.70 (s, 1H, H5), 7.42 (d, *J* = 7.4 Hz, 2H, H32), 7.35–7.21 (m, 7H, H26, H33, H34), 6.85 (d, *J* = 8.3 Hz, 4H, H27), 6.32–6.00 (m, 2H, H8, H1′), 5.57 (d, *J* = 15.8 Hz, 1H, H7), 4.60–4.55 (m, 1H, H3′), 4.34 (t, *J* = 6.3 Hz, 2H, H21), 4.25 (t, *J* = 6.4 Hz, 2H, H15), 4.07–4.03 (m, 1H, H4′), 3.83–3.71 (m, 2H, H9), 3.78 (s, 6H, H30), 3.42–3.34 (m, 2H, H5′), 2.76 (t, *J* = 6.3 Hz, 2H, H22), 2.71 (t, *J* = 6.4 Hz, 2H, H16), 2.46–2.28 (m, 2H, H2′). ^{13}C NMR (100 MHz, CD_2Cl_2, 25 °C): δ =163.4 (C13), 162.0 (C4), 159.2 (C28), 155.9 (C11), 153.3 (C19), 149.7 (C2), 144.9 (C31), 137.8 (C6), 135.9, 135.7 (C25), 130.4 (C26), 128.4 (C32,C33), 127.4 (C34), 126.3 (C8), 124.5 (C7), 117.7 (C17), 116.8 (C23), 113.6 (C27), 111.7 (C5), 87.2 (C24), 86.6 (C4′), 85.3 (C1′),

72.4 (C3′), 63.9 (C5′), 61.1 (C21), 60.3 (C15), 55.6 (C30), 44.0 (C9), 41.4 (C2′), 18.5 (C16), 18.3 (C22). HRMS (ESI$^+$) calcd for $C_{42}H_{44}N_7O_{11}$ 822.3099, found 822.3084.



5′-*O*-(4,4′-Dimethoxytrityl)-3′-*O*-(2-cyanoethyl-*N*,*N*′-diisopropylphosphoramidyl)-5-(3-(*N*,*N*′-bis-(2-cyanoethoxycarbonyloxy)guanidinyl)-1-propenyl)-2′-deoxyuridine (4.1)

Compound **4.8** (410 mg, 0.50 mmol) was dried by co-evaporation from pyridine, and residual pyridine was removed by co-evaporation from CH_2Cl_2. The resulting foam was suspended in CH_2Cl_2 (12.5 mL). Diisopropylethylamine (160 µL, 0.92 mmol) and then 2-cyanoethyl *N*,*N*′-diisopropylchlorophosphoramidite (110 µL, 0.49) was added at room temperature. The reaction was about half complete after 1 hr of stirring at room temperature and remained unchanged overnight. Additional diisopropylethylamine (160 µL) and 2-cyanoethyl *N*,*N*′-diisopropylchlorophosphoramidite (110 µL) were added, and the reaction was terminated by adding MeOH (100 µL). The reaction mixture was diluted with CH_2Cl_2, washed with saturated $NaHCO_3$ and dried over Na_2SO_4 (anhydrous). Purification on silica flash chromatography (16×2.0 cm column) gave 343 mg (67 %) of a

white foam that eluted in EtOAc/hexanes (4:1). R_f = 0.48 (1:9 MeOH/$CHCl_3$). ^{1}H NMR (400 MHz, CD_2Cl_2, 25 ℃): δ = 11.68 (s, 1H, H18), 8.81 (s, br, 1H, H3), 8.09 (m, 1H, H10), 7.74 (s, 1H, H5), 7.46–7.40 (m, 2H, H32), 7.35–7.22 (m, 7H, H26, H33, H34), 6.89–6.81 (m, 4H, H27), 6.40–6.31 (m, 2H, H8, H1′), 5.53 (d, *J* = 15.8 Hz, 1H, H7), 4.71–4.64 (m, 1H, H3′), 4.32 (t, *J* = 6.3 Hz, 2H, H21), 4.25 (t, *J* = 6.4 Hz, 2H, H15), 4.23–4.13 (m, 1H, H4′), 3.85–3.52 (m, 6H, H9, H35, H38), 3.79 (s, 6H, H30), 3.50–3.30 (m, 2H, H5′), 2.76 (t, *J* = 6.3 Hz, 2H, H22), 2.72 (t, *J* = 6.4 Hz, 2H, H16), 2.63 (t, *J* = 6.4 Hz, 1H, H36), 2.60–2.45 (m, 1H, H2′), 2.46 (t, *J* = 6.3 Hz, 1H, H36), 2.40–2.32 (m, 1H, H2′) 1.20–1.13 (m, 9H, H39), 1.10–1.05 (m, 3H, H39). ^{13}C NMR (100 MHz, CD_2Cl_2, 25 ℃): δ = 163.3, 162.0, 159.2, 155.9, 153.3, 149.7, 149.6, 144.9, 137.8, 135.8, 135.7, 130.5, 128.5, 128.4, 127.5, 127.4, 126.2, 124.5, 118.2, 118.0, 117.7, 116.7, 113.6, 111.7, 87.1, 86.1, 85.8, 85.4, 85.3, 74.0, 73.8, 73.6, 73.4, 63.5, 63.4, 61.1, 60.3, 58.8, 58.7, 58.6, 58.5, 55.6, 44.0, 43.7, 43.6, 43.5, 40.5, 40.4, 24.7, 24.6, 20.8, 20.7, 20.6, 18.5, 18.3. ^{31}P NMR (121 MHz, CD_2Cl_2, 25 ℃): δ = 149.5, 149.4. HRMS (ESI^+) calcd for $C_{51}H_{60}N_9O_{12}Na$ 1044.3997, found 1044.3990.

4.6.2 Guanidinium-modified oligonucleotides

Oligonucleotide concentration determination. The molar absorptivity of the unmodified oligonucleotide (382700 $M^{-1}cm^{-1}$) and the molar absorptivity of the 6-carboxyfluorescein group (20960 $M^{-1}cm^{-1}$) was used to determine the concentration of the modified oligonucleotides at 260 nm.

Melting temperature. Oligonucleotides at 1 μM were prepared in 100 mM NaCl and 10 mM sodium phosphate (pH = 7.0) and annealed. Samples were monitored for absorbance

at 260 nm with data points at about every 1 ℃ from 20–85 ℃. For the DNAzyme/substrate studies, T_m's were obtained for solutions in the absence of Mg^{2+}. $MgSO_4$ was then added to these same solutions to a final concentration of 0.5 mM and reannealed prior to another set of T_m measurements. These studies were measured from 15–60 ℃ to keep the cleavage of the substrate to a minimum.

Kinetics. Catalyst and substrate solutions were prepared in 150 mM NaCl, 50 mM Tris-HCl (pH 7.5), 0.5–10 mM $MgSO_4$ and reactions were initiated by adding equal volumes of substrate solution to catalyst solution. Trace amounts of the substrate were labeled at the 5′ end with ^{32}P-phosphate (radiolabelled substrate prepared in the same manner as the radiolabelled primer **P2.1** from Chapter 2). Final concentrations of the catalyst and substrate were either 5 and 50 nM or 0.25 μM and 5 μM, respectively. Reactions were incubated at 37 ℃ in the presence of mineral oil. Aliquots were removed and quenched with a 3x volume of loading buffer [a solution of formamide (27 mL), $EDTA_{(aq)}$ (3 mL, 0.5 M), xylene $cyanol_{(aq)}$ (300 μL, 0.05 %) and $bromophenol\ blue_{(aq)}$ (300 μL, 0.05 %)]. Samples were loaded onto denaturing PAGE (20 %) and analyzed with a phosphorimager from GE Healthcare.

Chapter 5: Towards the *in vitro* selection of a DNAzyme sensor

5.1 Introduction

Sensors are devices that produce signals upon exposure to real-world conditions. They give us the ability to acquire information from our environment that extends beyond our five senses, and hence sensors play an integral role in our quality of life. They are found in automobiles, electronics, safety devices, medical equipment and much, much more. One area of intense research involves the search for sensors that can detect and monitor certain molecules or ions, particularly those that can be found inside a human body. These types of sensors can aid in the detection of toxic substances and facilitate the monitoring, diagnosis, and treatment of patients.

Every sensor has two essential components: target identification and signal generation. The target can range from small ions to large biopolymers and even whole cells, and signal transduction converts the recognition event into something that is more easily measured or quantified such as an optical or electrochemical signal. One commonly used sensor is the one that is used for the monitoring of glucose in blood for patients of diabetes.[210,211] This sensor uses glucose oxidase as the recognition component to oxidize glucose and reduce the cofactor flavin adenine dinucleotide (FAD). The resulting $FADH_2$ is oxidized by the signal generating electrode to give rise to a measurable current. The glucose sensor is highly specific, sensitive, fast and relatively low cost, all properties that are necessary for an ideal sensor. One group of molecules that has promising sensor properties is that of sensors derived from nucleic acids.[212,213] Nucleic acids are fully capable of binding targets with high affinity and specificity in the form of aptamers,[13-16] and this binding action can be coupled to a structural change that

leads to the generation of a signal. There are two types of nucleic acid sensors: aptamers and nucleic acid enzymes. This chapter will be focused on DNAzymes[13,41,214] that act as sensors. Sensors of this type will catalyze a reaction only in the presence or absence of a certain analyte.

Sensors derived from nucleic acids can be found in nature in the form of riboswitches.[215-218] Riboswitches are domains that are usually found in the untranslated regions of mRNA's where they can directly bind to metabolite targets for the regulation of gene expression. The binding process typically triggers structural changes that affect transcription termination or translation initiation. While most riboswitches are aptamer-based, two riboswitches have been shown to be ribozymes.[219,220] One of these ribozyme riboswitches, the glucosamine-6-phosphate riboswitch,[220-222] regulates the translation of glmS mRNA that is responsible for the glmS enzyme. The protein enzyme catalyzes the production of glucosamine-6-phosphate, a molecule that is required for the biosynthesis of cell walls for Gram-positive bacteria. In the presence of glucosamine-6-phosphate, the ribozyme is activated and performs cleavage in the non-coding region of the glmS mRNA that leads to the eventual degradation of the mRNA. This results in the lower production of glucosamine-6-phosphate.

Compared to natural nucleic acid enzyme sensors, the number of artificial nucleic acid enzyme sensors is much more plentiful. Although DNAzymes would be the preferred polymers for such applications due to their relative stability, ease of synthesis and low cost compared to their RNA counterparts, numerous examples of both ribozyme and DNAzyme sensors can be found in the literature. Some of the sensors are derived from unmodified nucleic acids whereas a few have also made use of modifications.[212] Nucleic acid catalysts come in two types. In the first group of these nucleic acid sensors,

the target molecule is directly involved with catalysis in the active site either as a cofactor or an inhibitor. Catalysts of this type have been found to be sensors for divalent zinc,[171,223] lead,[27,224-227] uranate,[228] and copper.[229] The second group of these nucleic acid sensors are allosteric in nature. The distal binding of the target molecule to the active site triggers the proper folding of the catalytic motif and results in catalysis. These sensors can recognize a very wide range of molecules. In addition to metal cations,[188,230] allosteric nucleic acid enzymes have also been used for the detection of ATP,[231-233] flavin mononucleotide[234,235] and theophylline.[235]

The aforementioned allosteric catalysts are predominantly sensors that were rationally designed whereby the aptamer portion, catalytic portion or both portions have previously been derived by independent *in vitro* selection experiments. The attachment of the two domains involves a trial and error process that results in allosteric catalysts that are usually less catalytically active than the original catalyst and have less affinity for the allosteric factor than the original aptamer. In order to obtain superior allosteric catalysts, a selection, where one or both of the domains are replaced by a random oligomeric region, can be carried out. Through selections, stem II of the hammerhead ribozyme has been transformed into an allosteric binding region for targets that include ADP, caffeine, aspartame and metal cations.[236-239] *In vitro* selections have also been applied to allosteric ribozyme ligases as well.[240,241] In this chapter, the search for a sugar-recognizing DNAzyme will be discussed. Starting from a library of sequences with no pre-existing catalytic or binding motifs, an *in vitro* selection was carried out to find catalysts that would perform ribonucleotide phosphodiester self-cleavage in the presence of a target sugar. To facilitate the detection of target sugars, the library was synthesized with what was originally thought to be 5-(4-boronobenzoylaminomethyl)-2′-deoxyuridine

triphosphate (**2.12**). Of course, as discussed in Chapter 2, we subsequently showed that this triphosphate was in fact 5-(4-hydroxybenzoylaminomethyl)-2′-deoxyuridine triphosphate (**2.24**), but the selection was still carried out to completion. The selection process and the characterization of the isolated DNAzymes will be discussed.

5.2 *In vitro* selection

The selection of a sugar-sensing DNAzyme was performed using the same streptavidin-biotin methodology that has been used by others, as shown in Figure 5.1; a biotinylated primer containing a single cytidine residue was polymerized with a template containing 40 degenerate positions in the presence of the modified nucleotide **2.24** to synthesize a modified oligonucleotide library of greater than 10^{13} sequences. The library was constructed such that the random sequence domain, which must be responsible for both sugar recognition and catalysis, is brought into close proximity to the target self-cleavage site. The sequences were initially incubated in an environment of 50 mM HEPES pH 7.4, 200 mM NaCl, 0.5 mM $MgCl_2$ and 0.1 mM $ZnSO_4$ for the negative selection and an additional 1 mM each of four different carbohydrates for the positive selection.

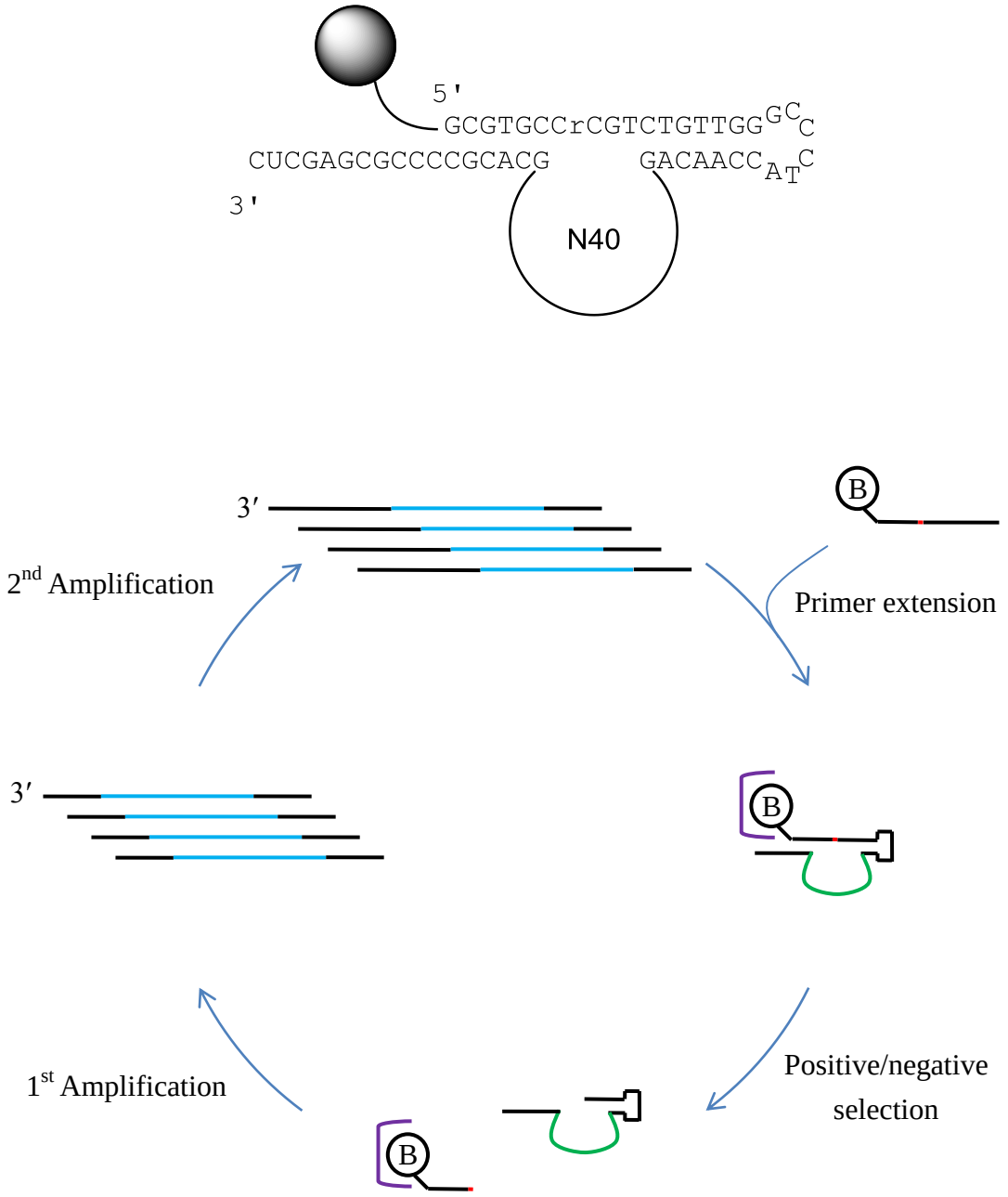


Figure 5.1 Sugar-sensing DNAzyme selection. Upper: DNA pool sequences for the selection. U – modified residue; N40 – random 40 modified-nucleobase region; ball – biotin; rC – cytidine. Lower: selection scheme. (——) DNA primer regions; (——) N_{40} template region; (——) RNA; (——) N_{40} random region; (B) biotin; (——) streptavidin magnetic particles.

During iterative rounds of *in vitro* selection, a library that initially contains very few sequences that exhibit the desired activity can easily be contaminated with unwanted sequences, particularly during the earlier rounds of the selection. This contamination would increase the number of rounds of selection required and may even prevent the selection from being successful. The use of modified nucleotides further complicates the selection because generally, modified nucleotides are poorer substrates for polymerase-mediated incorporation, and modified templates are more difficult to amplify. Due to all this, several denaturing PAGE purifications were performed to minimize contamination from unwanted sequences. After the cleaved sequences were isolated from the positive selection, the modified oligonucleotides were purified by denaturing PAGE to isolate only the sequences that were of the appropriate size. This step was predominantly responsible for removing uncleaved species that passed through the selection procedure. The purified cleaved oligonucleotides were then amplified with two PCR's, and the products of both PCR reactions were purified by denaturing PAGE as well. The first PCR used one set of primers to produce double-stranded product that was the same length as the cleavage material. In the second PCR, one of the primers was substituted to give a larger product that fully hybridizes with the selection primer.

Using three denaturing PAGE purifications and two PCR's for every round of selection, two selections were attempted on the same oligonucleotide library. The first selection involved the search for sugar-sensing DNAzymes, and the second selection was a less demanding one that searched for a self-cleaving DNAzyme. It was not known whether our modified DNA library would be capable of catalyzing self-cleavage, and hence, the second selection would act as a control for the first. If no activity could be detected from the second selection, then it would not be a surprise to see no activity in the

first selection as well. In fact however, the opposite was observed. Activity began appearing in the sixth round of the first selection. No activity was observed in the second selection, and it was discontinued after the ninth round.

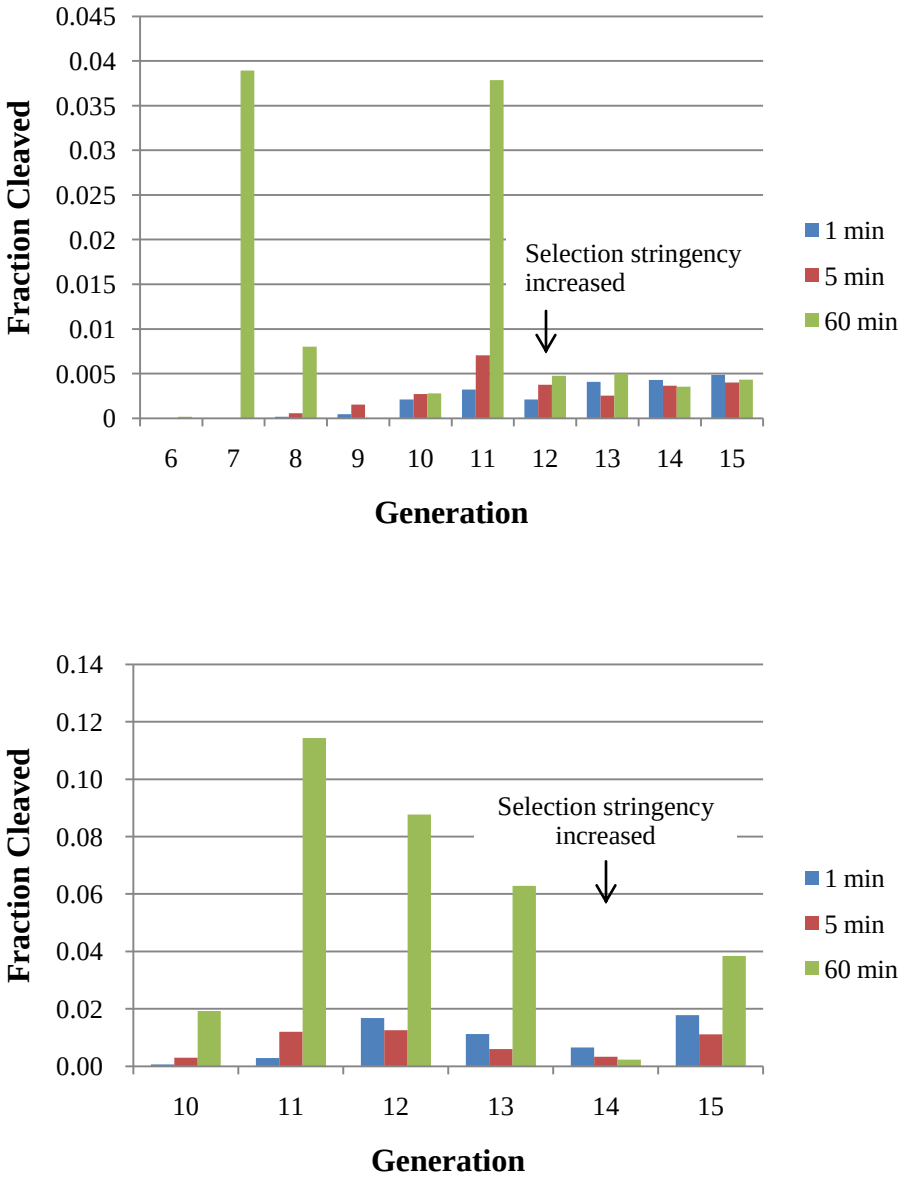


Figure 5.2 Progress of the sugar sensor selection. Top: after round 9, the selection was divided into two. Generations 10–15 correspond to the stachyose selection. Bottom: ATP selection. Cleaved species were collected only after 1 h of incubation for rounds 6 and 7, and the collected material after 1 h of incubation for generation 9 was not analyzed by denaturing PAGE.

The progress of the first selection is illustrated in Figure 5.2. In the initial rounds, the positive selection was carried out for 1 h. As substantial activity began to appear, the

selection products were collected in three different solutions of selection buffer. After 1 min of selection time, the selection buffer was removed, and then fresh selection buffer was added. The removal of selection buffer was repeated twice after a cumulative time of 5 and 60 min. Each of these three selection samples were separately used as templates for 1st amplification. Since the cleaved material from the earlier time intervals should consist of faster cleaving DNAzymes, more 1st amplification material from the earlier time intervals was used as template for the 2nd amplification in an attempt to favour DNAzymes of superior catalytic rates.

The positive selection was initially carried out in a "soup" of four *cis*-diol-containing molecules: stachyose, sialic acid, D-mannoheptose and ATP (Figure 5.3). It was thought that the presence of several potential target allosteric factors would enhance the probability of finding a sugar-sensing DNAzyme. Upon discovery of a DNAzyme, kinetic studies would be performed on the catalyst to determine its target specificity. After the ninth round, it was realized that the modified 2′-deoxyuridine that was being used did not contain a boronic acid as described in Chapter 2. However, since the selection appeared to display sugar-dependent activity, the selection was continued. Prior to the tenth round of selection, kinetic experiments were performed with DNAzymes that were synthesized using sequences from the ninth generation. It was found that the DNAzyme population showed enhanced self-cleavage in the presence of ATP and stachyose. As a result, the sugar-sensing DNAzyme selection was split into two selections after the ninth round, one incubated in the presence of stachyose and the other one, in the presence of ATP. The concentrations of the sugars began at 1 mM. For ATP, the concentration remained constant throughout, and for stachyose, the concentration was decreased to 0.2 mM during the 12th round of selection. The negative selections were

carried out for 1–6 h.

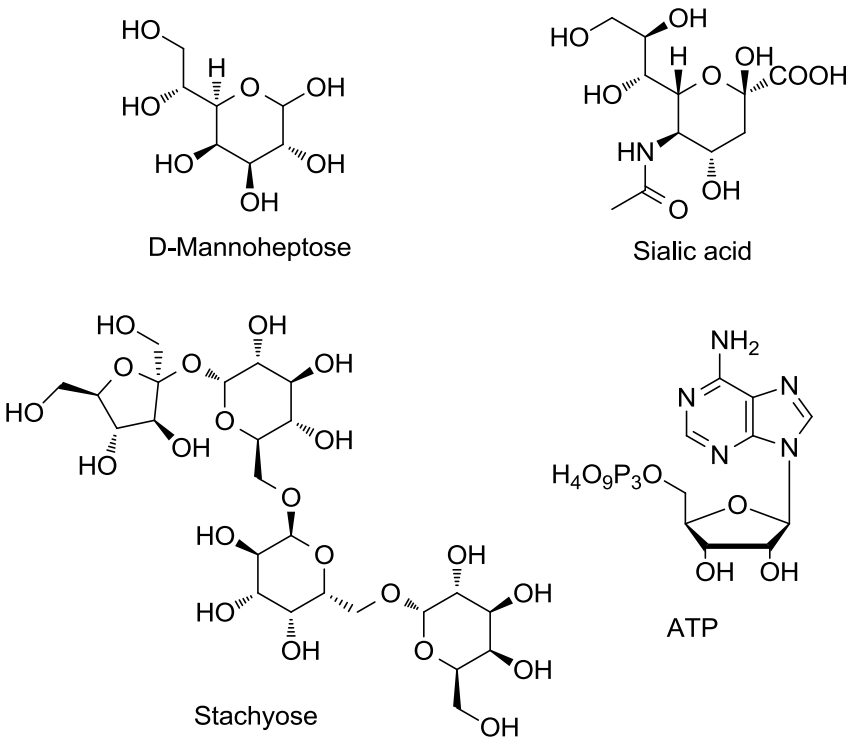


Figure 5.3 *cis*-Diol-containing molecules used in the positive selection of the sugar-sensing selection.

The denaturing polyacrylamide gel from round 8 of the sugar sensor and non-sugar sensor selections is shown in Figure 5.4 as a representative gel for the selection process. Lanes 1 and 2 are the control lanes for uncleaved and cleaved material, respectively. In the absence of an identifiable activity band for the selections, as was the case for lane 3 for this round, the lane 2 control would be used to judge where the cleaved products would be. Lane 3 corresponds to the non-sugar-sensing selection, and lanes 4–6 correspond to the sugar-sensing selection. There are several items in Figure 5.4 that need to be mentioned. Although triphosphate **2.24** was a very good substrate for Vent (exo–), truncation products were still found in the selection as can be seen in lanes 1 and 2. These

shorter sequences will predominantly remain attached to the streptavidin magnetic particles. Also, uncleaved and cleaved species are very well separated due to a size difference of 28 nucleotides. The most noticeable aspect of this gel is that lanes 3–6 all have detectable amounts of uncleaved material. This is a function of incomplete separation of cleaved from uncleaved material and is attributed to the imperfect retention by the streptavidin magnetic particles. Without gel purification, these uncleaved sequences would contaminate the selection.

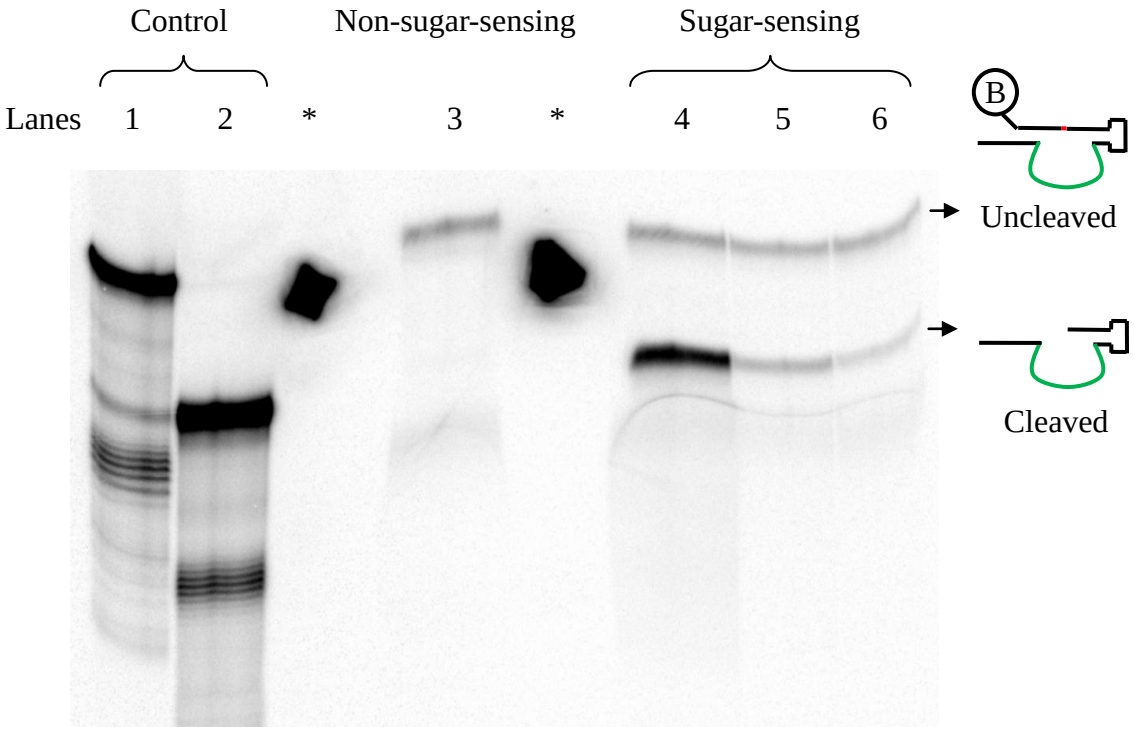


Figure 5.4 Round 8 of both the sugar sensor and non-sugar sensor selections on denaturing PAGE (7 %). Lanes 1 and 2, control lanes corresponding to uncleaved material and base-treated material, respectively; lanes marked with (*) correspond to the two polygons that were used as radioactive tabs for imaging and locating of desired product bands for purification; lane 3, positive non-sugar-sensing selection (1 h incubation); lanes 4–6, positive sugar-sensing selection 5–60, 2–5, 0–1 min (incubation time interval), respectively. Oligonucleotides: (——) DNA primer regions; (——) RNA; (——) N_{40} random region; (B) biotin.

5.3 Cloning and sequence analysis

After 15 rounds of selection, cloning was performed on the material obtained from generation 11 of the stachyose selection and generation 15 of the ATP selection by Christopher Hipolito. The reason generation 11 of the stachyose selection was cloned was because the selection did not seem to progress from rounds 12–15. The sequences obtained are shown in Table 5.1, and an analysis of the sequences is shown in Table 5.2. In addition to the sequence information corresponding to the 40 nucleotide random region, an extra nucleotide on the 5′ end is shown as well. In the original N_{40} template, the extra nucleotide was a dG. However, because this nucleotide was not part of the primer binding region, mutation due to misincorporation at this location could occur. Each of the 31 sequences obtained from the cloning process were individually different. The average length of the random region was about 40 nucleotides, and the average number of modified residues was 8.8 (standard deviation of 2.1), a value that is not significantly different from the statistically random number of 10. Often, over multiple rounds of selection, the number of modified residues found in a library can be lowered due to poor incorporation of more heavily modified sequences or poor amplification by PCR. Neither of these two situations should have played a significant role as the incorporation and amplification of triphosphate **2.24** was shown to be at least satisfactory in Chapter 2.

Initial kinetic analysis of the clones (4 time points per clone) revealed that a large proportion of them, 17 out of 31 sequences, were catalytically active. However, although the selection seemed to indicate that a sugar-sensing population of oligonucleotides had been obtained, none of the isolated clones displayed any expected allosteric activity. Each and every clone was incubated in the absence of sugars or in the presence of either ATP

or stachyose, and each clone displayed the same self-cleavage activity under the three different environments. Each clone was separately tested with both ATP and stachyose because both selections originated from the same ninth generation pool. The lack of allosteric enhancement was even more surprising considering that the selection performed in the absence of sugars did not yield activity. The non-allosteric activity was found to be very dependent on the concentration of the divalent metal cations present. The clones were first tested in an environment consisting of 50 mM HEPES pH 7.4, 200 mM NaCl, 10 mM $MgCl_2$ and 1 mM $ZnSO_4$, and those that were found to be active were tested at the lower concentrations of 1 mM $MgCl_2$ and 0.1 mM $ZnSO_4$ as well.

Clone	Sequence (5′–3′)
ATP2	GAAACTAGTTGACGGAGGCACTGCATGGTGAGTGCTGTTC
ATP3	GATACCAGCCCGTACGCAGGCACGTTACCCTACAAGCAAAA
ATP13	GAAACTAGTTGACGACGACACGCCAACGCCCTGTCGTGTATA
ATP14	GAAACTAGTTGACGGAGGTACGGTCATGAGGCGGTGTGGTCTA
ATP16	GATAGTAGTTCACGGAGGCACTACTAAGTATGGTGTGCAAA
ATP18	GAAGCTAGTTGACGCAGGCACGGAGTGGGTGAGCGTGGTTA
ATP20	GAAACTAGTTGACGCAGGTACCCCTACACAGTGTAAACATA
ATP22	GATACTAGTTGACGAGGGCATGCCCTCCCTCCCGTAGTGTA
ATP37	GAAACTAGTTGACGGGGGCATAGGTCAGTATGCTTTGTTT
ATP47	GAAACAAGCTGCACGAAGGCACGCACGGTCCGTATCGGCCTA
ATP101	AACTGCCATATTACCTACAATGGCGGAGTGCGCCCTATGATC
ATP102	GATAGCAGCGTCTACGAAGGCACCATGGTCCTAGCATGTTA
ATP105	GAAGCAAGCTACACGAAGGCACGCACACTGGTGCCTGGTTA
ATP108	GAAACTAGTTGACGCAGACACTGCTAAGTATCGTGAGCAAG
ATP110	GAACTGAGCCCAGACGCCGGCTAAAAGTCGAAATGGTGTTA
ATP114	GATATCAGCCGATACGAAGGCAAACAATATGCCTTTCATTG
ATP115	GAAACTAGTTGACGCAGGCATTATGCAGTATCTCATAGAAA
STA1	GAAACGAGCTCGAACGAAGGCACGATGGCCAATCACCTATTG
STA4	GAAATTAGTTGACGCAGGCATTGCACTGTGAATGGTGTTA
STA17	GAAACTAGTTGACGAAGGCACTGCTGACTATCGAGTGCAAA
STA22	GATACTAGTTGACGCAGGCACGATTGAGTATCCCCTTCACAA
STA46	GATACTAGTTGACGCAGGCACTAATGACCATCCCGTGCATA
STA48	GATACGAGCCGTACGAAGGCACGATGCGCAATCGCCTATTG
STA53	GCACATGAGTGTGAGACACGCAGTGGTCCGGATTGGGCAA
STA54	GAACTGAGCTCAGACGCAGGCTAAACGCTGAAATAGCGCTA
STA107	GATACTAGTTGACGAAGACACTAATGACTATCTAGTGCATA
STA109	GATACTAGTTGATGATGGCACTCCCTAGCAGTCTACTACT
STA112	GCATCCTGCTCGCTGTGGAGGCCACCAAATCGCTATGGAAT
STA113	GAAACTAGTTGACGGAGGCACTTCTGAGGATCGTGTCCACA
STA120	GATACTAGTTGACGAAGGCACTATCGCGTTTTACGTGCAAA
STA124	GAAACTAGTTGACGCAGGCACTTCTGAGCATCTCCTTCACAA

Table 5.1 Clone sequences isolated from the sugar-sensing selection. Only the sequence of the random region and one extra residue on the 5′ end is shown. ATP clones correspond to sequences obtained from the ATP selection and the stachyose (STA) clones correspond to sequences obtained from the stachyose selection.

Clone	Random region length	# of modified residues	Relative Activity	
			A	B
ATP2	39	10	+++	+
ATP3	40	5	NA	NE
ATP13	41	8	+	NE
ATP14	42	10	+++	+
ATP16	40	10	++	NA
ATP18	40	8	+++	+
ATP20	40	8	NA	NE
ATP22	40	10	NA	NE
ATP37	39	13	NA	NE
ATP47	41	5	+	NE
ATP101	41	10	NA	NA
ATP102	40	9	+	NA
ATP105	40	6	NA	NA
ATP108	40	8	NA	NA
ATP110	40	7	NA	NA
ATP114	40	10	NA	NA
ATP115	40	10	+	NA
STA1	41	6	+++	NA
STA4	39	12	NA	NA
STA17	40	8	++++	++
STA22	41	10	+++	+
STA46	40	9	+++	+
STA48	40	7	++	NA
STA53	39	7	NA	NE
STA54	40	6	++	NA
STA107	40	11	NA	NA
STA109	39	12	NA	NA
STA112	40	9	NA	NA
STA113	40	9	+++	+
STA120	40	11	+++	+
STA124	41	10	+++	+

Table 5.2 Properties of the clones isolated from the sugar-sensing selection. Rates of self-cleavage were performed under 2 different conditions: A) $MgSO_4$ (10 mM), $ZnSO_4$ (1 mM), HEPES (50 mM) pH 7.4, NaCl (200 mM); B) $MgSO_4$ (1 mM), $ZnSO_4$ (0.1 mM), HEPES (50 mM) pH 7.4, NaCl (200 mM). The reactions were analyzed after 15, 60, 180

min and overnight treatment. Relative activities: (++++) ≥ 50 % complete after 15 min; (+++) ≥ 50 % complete after 60 min; (++) ≥ 50 % complete after 180 min; (+) ≥ 50 % complete after overnight reaction; (NA) < 50 % complete after overnight reaction; (NE) – not evaluated.

Table 5.3 shows a summary of the catalytic activities of the non-allosteric DNAzymes. The stachyose selection clones gave rise to a greater proportion of active clones as well as DNAzymes of higher activity. This result is not what was expected as the ATP selection displayed more activity during selection, and the ATP selection was also carried out for four additional rounds. As can be seen in Table 5.4, the majority of the DNAzymes share a common motif where 17 out of the first 21 residues in the random region are identical. This domain may be a common catalytic motif.

Relative activity	++++	+++	++	+
ATP clones	0	3	1	4
STA clones	1	6	2	0

Table 5.3 Summary of relative activities of the isolated clones from the sugar-sensing selection. Relative activities were obtained in the presence of $MgSO_4$ (10 mM), $ZnSO_4$ (1 mM), HEPES (50 mM) pH 7.4, NaCl (200 mM). Relative activities: (++++) ≥ 50 % complete after 15 min; (+++) ≥ 50 % complete after 60 min; (++) ≥ 50 % complete after 180 min; (+) ≥ 50 % complete after overnight reaction.

STA17	GAAACTAGTTGACGAAGGCACTGCTGACTATCGAGTGCAAA-- 41
STA120	GATACTAGTTGACGAAGGCACTATCGCGTTTTACGTGCAAA-- 41
STA113	GAAACTAGTTGACGGAGGCACTTCTGAGGATCGTGTCCACA-- 41
STA124	GAAACTAGTTGACGCAGGCACTTCTGAGCATCTCCTTCACAA- 42
STA22	GATACTAGTTGACGCAGGCACGATTGAGTATCCCCTTCACAA- 42
STA46	GATACTAGTTGACGCAGGCACTAATGACCATCCCGTGCATA-- 41
ATP18	GAAGCTAGTTGACGCAGGCACGG-AGTGGGTGAGCGTGGTTA- 41
ATP2	GAAACTAGTTGACGGAGGCACTG-CATGG-TGAGTGCTGTTC- 40
ATP14	GAAACTAGTTGACGGAGGTACGGTCATGAGGCGGTGTGGTCTA 43

Table 5.4 Sequence alignment analysis of the DNAzymes (5′–3′). Clustal W 2.0[242] was used to analyze the sequences of 9 of the 10 most active self-cleaving species shown in Tables 5.1 and 5.2. The number on the right is the length of the sequence analyzed, and the sequence region shown corresponds to the N_{40} random region and 1 residue at the 5′ end.

5.4 Analysis of clone STA17.

The catalytically most active self-cleaving DNAzyme obtained from the selection was clone STA17. The random region of this DNAzyme contained 40 nucleotides of which 8 were modified nucleotides. The predicted 2D structure using mfold[243,244] is shown in Figure 5.5. Two stem-loops are observed in the random region. Since the common motif that is shared by the most active DNAzymes corresponds to the smaller loop, it is likely that the smaller loop is responsible for catalysis and the other loop plays more of a structural role. It should be noted that this predicted structure does not take into account the modifications, which could have a significant effect on the 2D structure.

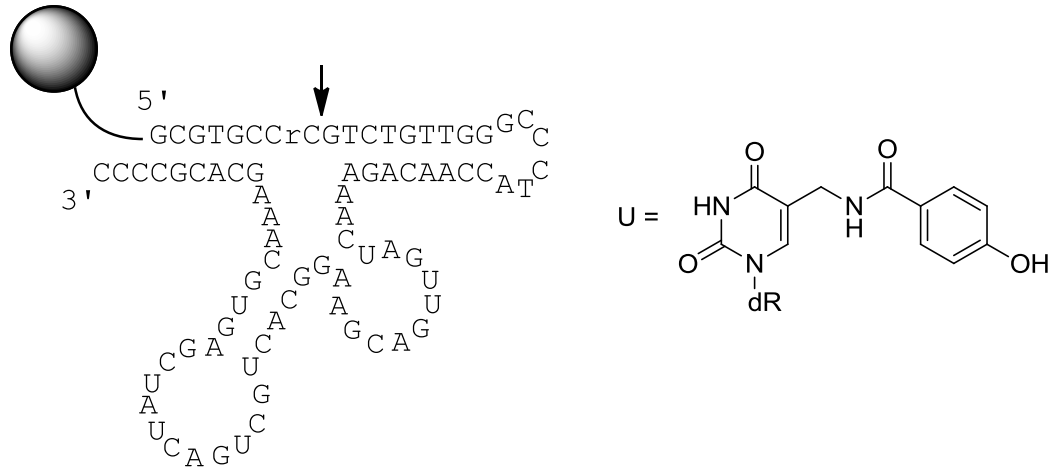


Figure 5.5 Predicted 2D structure of the self-cleaving STA17. Cytdine is represented by rC, self-cleavage location is indicated by the arrow, modified residues are denoted with a "U" and biotin is represented with the ball. Structure was obtained using mfold[243,244] with hybridization of the cytidine not allowed.

The kinetic analysis of STA17 is shown in Figure 5.6. In the gel, uncleaved DNAzyme (top band) catalyzes time-dependent self-cleavage to provide the cleaved product (lower band). After an hour, greater than 90 % of the cleavage reaction was complete in the presence of 50 mM HEPES pH 7.4, 200 mM NaCl, 1 mM $ZnSO_4$ and 10 mM $MgSO_4$ at room temperature. The bands were quantified and fit to a monophasic single-exponential equation to give an average rate constant of $k_{obs} = 0.20 \pm 0.02$ min^{-1}. It was found that this catalyst was active in the presence of either divalent zinc or magnesium with divalent zinc giving rise to enhanced rates of cleavage. In the absence of divalent metal cations, self-cleavage was not detected. When the catalyst was synthesized with dTTP rather than **2.24**, the resulting sequences in the presence of both divalent zinc and magnesium also did not reveal any self-cleavage. The modifications may directly participate in catalysis or just be involved in the proper folding of the catalyst.

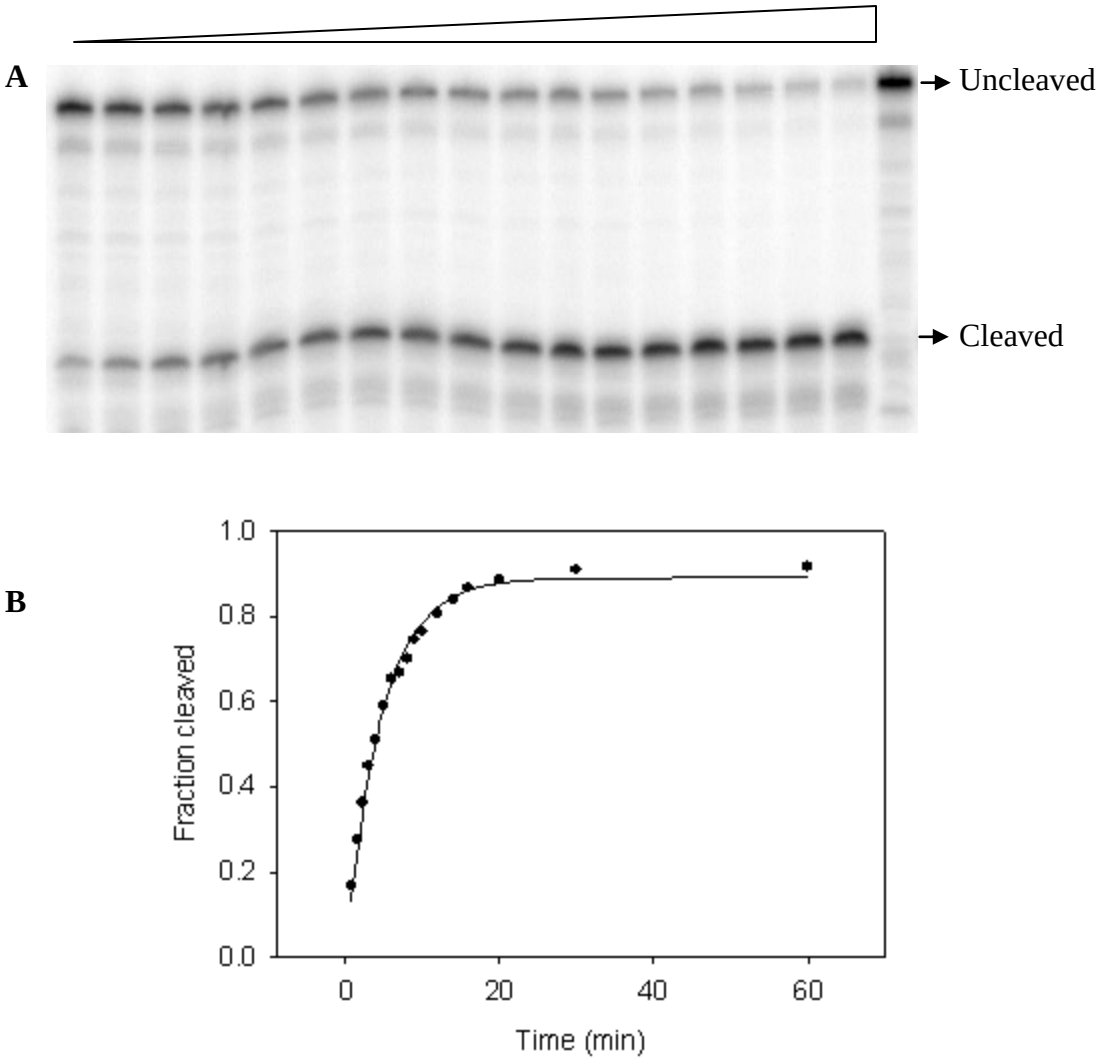


Figure 5.6 Self-cleavage kinetics for STA17. Experiment performed in 50 mM HEPES pH 7.4, 200 mM NaCl, 1 mM $ZnSO_4$, 10 mM $MgSO_4$ at room temperature. A) Denaturing PAGE (7 %) showing the relative amounts of cleaved and uncleaved material over a period of 60 min. Lanes correspond to aliquots quenched at 0.75, 1.5, 2.25, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, 20, 30, 60, 0 min. The triangle indicates increasing time. B) Graphical analysis of the reaction: $k_{obs} = 0.20 \pm 0.02$ min^{-1} ($R^2 > 0.99$).

The kinetic behaviour of DNAzyme STA17 was further analyzed through rescue experiments. As shown in Table 5.5, alkaline earth metal, transition metal and lanthanide cations (each at 1 mM) were substituted for Mg^{+2} and Zn^{+2}. Of the cations that were tested, only Ca^{+2} and Mn^{+2} ions enabled appreciable self-cleavage activity. Catalysis in the presence of Mn^{+2} was further enhanced. More extensive kinetic experiments performed using Mn^{+2} gave rise to an observed rate constant of 0.35 ± 0.03 min^{-1}. The DNAzyme STA17 performed self-cleavage more effectively in the presence of 1 mM Mn^{+2} than in the simultaneous presence of both 10 mM Mg^{+2} and 1 mM Zn^{+2}.

In an attempt to establish whether our phenol modified DNA could offer binding sites for metal cations, inhibition experiments were carried out with STA17. It was thought that if the DNAzyme was inhibited by one or a few metal cations, the DNAzyme could potentially be used as a sensor for those metal cations. As such, the same metal cations used for the rescue studies were also studied as inhibitors of STA17. Two organophosphate pesticides, fenitrothion and diazinon (Figure 5.7), were used as well. Organophosphate pesticides are used for the inhibition of acetylcholinesterase, an enzyme that degrades the neurotransmitter acetylcholine, by phosphorylation of a serine residue in the active site.[245] It was thought that the phenol modifications on STA17 could potentially be phosphorylated as well to give rise to an inhibited DNAzyme. The divalent and trivalent metal cations were used at 100 μM, and the pesticides were used at saturated concentrations (close to 100 μM). As shown in Table 5.5, inhibition of the self-cleavage reaction was observed with only Eu^{+3} and Hg^{+2}. Divalent mercury was particularly effective at inhibiting STA17. To investigate this further, the inhibition of STA17 was repeated with the lower Hg^{+2} concentrations of 10 and 30 μM. The denaturing PAGE analysis of these experiments is shown in Figure 5.8. Divalent mercury can effectively

inhibit STA17 at concentrations of 30 and 100 μM. At a Hg^{+2} concentration of 10 μM, the inhibition effect is much reduced.

	Rescue	Inhibition
Diazinon	NE	None
Fenitrothion	NE	None
Ca^{+2}	+	None
Ba^{+2}	none	None
Mn^{+2}	+++	None
Fe^{+2}	unsuccessful	None
Fe^{+3}	none	None
Ni^{+2}	none	None
Cu^{+2}	none	None
Hg^{+2}	none	+++
Eu^{+3}	none	+
La^{+3}	none	None

Table 5.5 STA17 kinetic experiments under rescue and inhibitory conditions. Rescue experiments were performed in the absence of divalent magnesium and zinc cations: divalent or trivalent metal ion (1 mM), HEPES (50 mM) pH 7.4, NaCl (200 mM). Inhibition experiments were performed using inhibitor (100 μM for metal ion inhibitors or saturated diazinon/fenitrothion), $MgSO_4$ (10 mM), $ZnSO_4$ (1 mM), HEPES (50 mM) pH 7.4, NaCl (200 mM). The use of Fe^{+2} for STA17 rescue gave rise to bands in the gel that did not migrate. NE – was not evaluated.

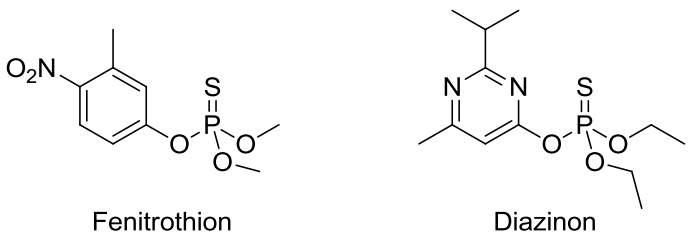


Figure 5.7 Organophosphate insecticides fenitrothion and diazinon.

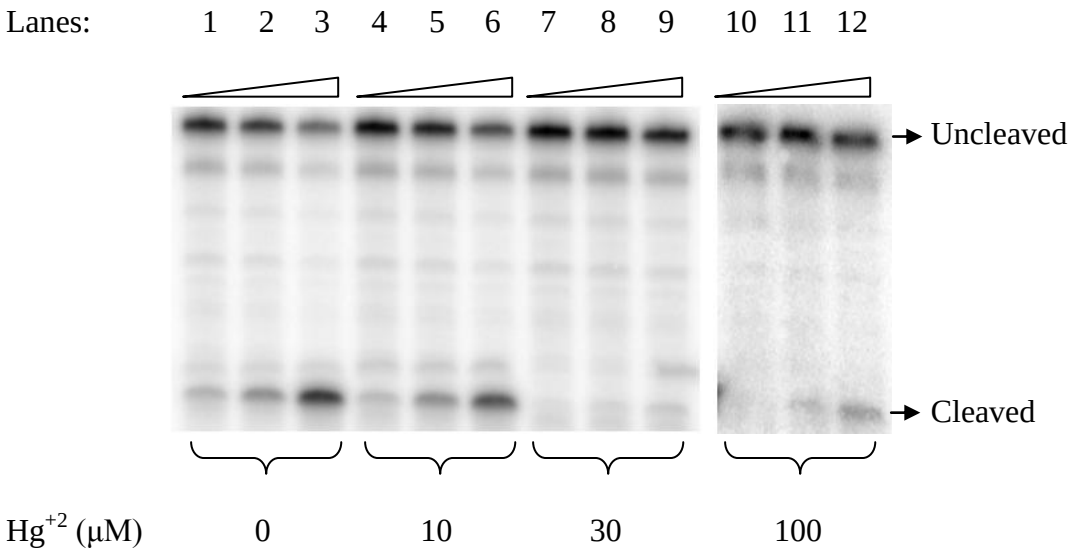


Figure 5.8 Inhibition of STA17 by Hg^{+2}. Lanes 1–3, control; lanes 4–6, 10 μM Hg^{+2}; lanes 7–9, 30 μM Hg^{+2}; lanes 10–12, 100 μM Hg^{+2}. Three time points per condition: 0.75, 3, 10 min. Reactions performed in $MgSO_4$ (10 mM), $ZnSO_4$ (1 mM), HEPES (50 mM) pH 7.4, NaCl (200 mM). Denaturing PAGE was performed on 7 % gels.

One last study that was performed on STA17 was the effect of pH on catalysis. The self-cleavage reaction was performed at 6 additional pH values between pH 5.9 and 8.9. As shown in Figure 5.9, STA17 was found to be most active at pH 7.4, the pH at which it was selected. On a log of the rate constant versus the pH curve, a straight line with a slope of 0.86 was found between pH 5.9 and 7.4 where the rate increases logarithmically. This log-linear relationship suggests that the rate-determining step of the reaction involves the deprotonation of a single group. At pH 7.9, the rate decreases somewhat, and above that, the rate of catalysis was rapidly lowered. This pH-dependent behaviour is similar to that of several other divalent metal cation-dependent ribophosphodiester-cleaving DNAzymes[28,46] including the modified DNAzyme

16.2-11[171] (section 1.8.3). This rapid decrease in activity could have been due to the deprotonation of a catalytically essential group such as a pK_a perturbed phenol group (4-hydroxybenzamide has a pK_a of 8.56).[246]

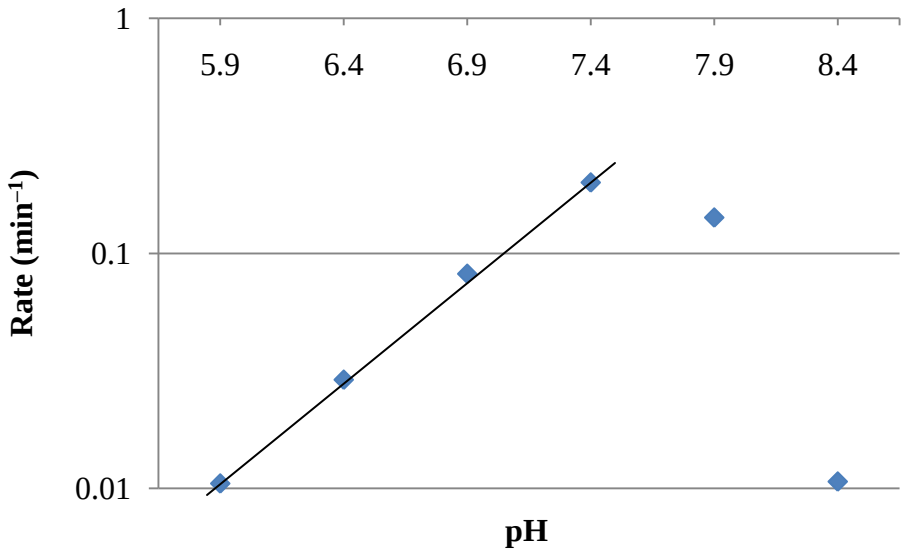


Figure 5.9 pH rate profile of STA17. Reactions performed in 50 mM buffer, 200 mM NaCl, 1 mM $ZnSO_4$, 10 mM $MgCl_2$. Fit to the formula of logy = mx + b, a slope of 0.86 was obtained ($R^2 > 0.99$).

5.5 Analysis of the selection.

Although the selection ultimately did not employ boronic acids, the selection process still gave rise to a population that appeared to display sugar-dependent activity. Self-cleavage was first detected in generation 6, and the selection process continued as the conditions of the selection gradually became more stringent. The number of sequences that were isolated during the positive selections was somewhat low (never more than 13 % cleavage after 1 h), but this was probably due to a very slow cleaving

population. Given more time, a much higher proportion of the selection would have cleaved. Later analysis of prolonged selection times revealed that this was indeed true (results not shown). Individual molecules from the sequence pools were subsequently cloned. It was thought that the absence of the boronic acids would be significant, but the behaviour of the selection led us to believe that the cloning process would have given us DNAzymes that would at least be somewhat responsive to the addition of carbohydrates. Hence, it was unexpected that none of the 17 active DNAzymes was more active in the presence of either ATP or stachyose, especially since the other selection performed in the absence of sugars did not give rise to activity. These results suggested that there were perhaps some problems with the selection process and that even if boronic acid-containing libraries had been used, the selection would still have been unsuccessful.

5.5.1 Negative selection

As discussed in section 1.6, negative selections have been effectively used for the *in vitro* selection of DNAzyme sensors. In our selection procedure, the negative selection was implemented to remove sequences that could catalyze self-cleavage in the absence of sugars thereby preventing these sequences from contaminating the positive selection. However, because all of the isolated DNAzymes displayed non-sugar-specific cleavage activity, the negative selection clearly was not effective. During the multiple rounds of selection, the amount of material that was eluted during the negative selection was always very little compared to the positive selection. There were no indications that a large number of non-sugar-specific sequences were present. In an attempt to understand what happened, the negative selection was carried out for longer periods of time. When the negative selection was extended from 1 h to 3 h, much more activity was observed in the

negative selection, but still somewhat less than the activity from the 1 h positive selection. However, the activity observed with an overnight negative selection was much more significant than the 1 h positive selection. This phenomenon was found to be present in both early and later generations. Figure 5.10 shows the denaturing PAGE of self-cleavage material that was synthesized using generation 7 product, the second round in which activity could clearly be observed. The figure illustrates what has just been described. Instead of decreasing over time, the activity of the negative selection seemed to be accelerating. The activity after 16.5 h of negative selection was more than 5.5 times the negative selection activity after 3 h. As well, the amount of activity in the positive selection in both negative selection scenarios was not very different. Provided that the sugars do not inhibit the self-cleavage reaction, the activity observed in both positive selections should predominantly be due to non-sugar-specific sequences.

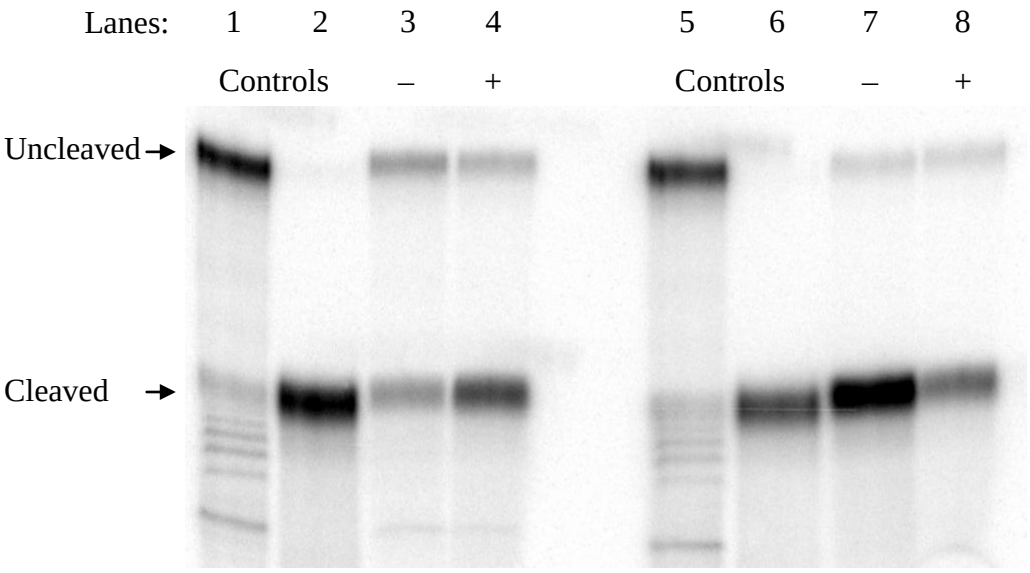


Figure 5.10 Negative selection analysis of sugar-sensing DNAzyme selection generation 7 by denaturing PAGE (7 %). Two identical selection samples were incubated in 50 mM HEPES pH 7.4, 200 mM NaCl, 0.1 mM $ZnSO_4$, 1 mM $MgCl_2$. The positive selection was carried out with 1 mM each of stachyose, sialic acid, D-mannoheptose and ATP. Lanes 1–2, controls; lane 3, 3 h negative selection; lane 4, 1 h positive selection; lane 5–6, controls; lane 7, 16.5 h negative selection; lane 8, 1 h positive selection.

5.5.2 Cleaved sequences remain on streptavidin magnetic particles.

Further analysis of the self-cleavage of STA17 under selection conditions revealed that cleaved sequences can stay bound onto the streptavidin magnetic particles rather than be collected with the selection buffer. This would explain why the negative selection seemed to be accelerating over time. During the first hour of the selection (negative selection), a large majority of the sequences that underwent self-cleavage would not be isolated because they would remain associated with the magnetic particles via non-specific interactions. If the negative selection is performed for longer periods of time, the streptavidin magnetic particles would probably slowly become saturated with cleaved

sequences such that no additional cleaved sequences could be retained. For each round, the selection was only carried out for 1 h. Any sequences that would have catalyzed self-cleavage were not very likely to be collected for exponential amplification. For the sugar-sensing selection, any sugar-specific DNAzymes that cleaved during the positive selection would be in equilibrium with the cleaved sequences that had already become associated with the streptavidin magnetic particles during the negative selection. Assuming that the number of cleaved yet bound oligonucleotides from the negative selection vastly outnumbered the oligonucleotides that underwent self-cleavage during the positive selection, the positive selection would isolate predominantly non-sugar-sensing sequences that would be amplified and carried to the next round of selection.

In an attempt to address the issue of cleaved sequences remaining attached to the magnetic particles via non-specific interactions, the cleavage of STA17 on streptavidin magnetic particles was analyzed in a variety of conditions. As streptavidin is very resistant to denaturants,[247] a 2 M solution of urea was used to try to remove cleaved material. The use of salmon sperm DNA for pretreating, as a washing step or directly in the selection buffer was explored as well. It was found that the use of 0.1 mg/mL salmon sperm DNA in the selection buffer was the most effective method for minimizing the number of cleaved sequences that remained attached to the streptavidin magnetic particles through non-specific interactions. Figure 5.11 illustrates this. A sample of DNAzyme STA17 was divided into 3 portions. One was used as an uncleaved control, and the other two were allowed to cleave in the presence or absence of salmon sperm DNA. The cleavage reactions were incubated with 50 mM HEPES pH 7.4, 200 mM NaCl, 1 mM $ZnSO_4$, 10 mM $MgCl_2$. After 30 min, the selection buffer from each reaction was

removed, and fresh buffer was added. The selection buffers were again removed after an additional 2 hours. For each reaction, the material from the strepavidin particles and the eluted buffers were loaded onto denaturing PAGE. As shown in Figure 5.11, the untreated reaction left a significant amount of material on the immobilized support (38 %, lane 2) whereas the presence of salmon sperm DNA greatly reduced the amount of material that was left behind (14 %, lane 5). This technique would be very useful for a new selection. In our selection, the technique would be ineffective as the population has already been dominated by non-sugar-sensing sequences.

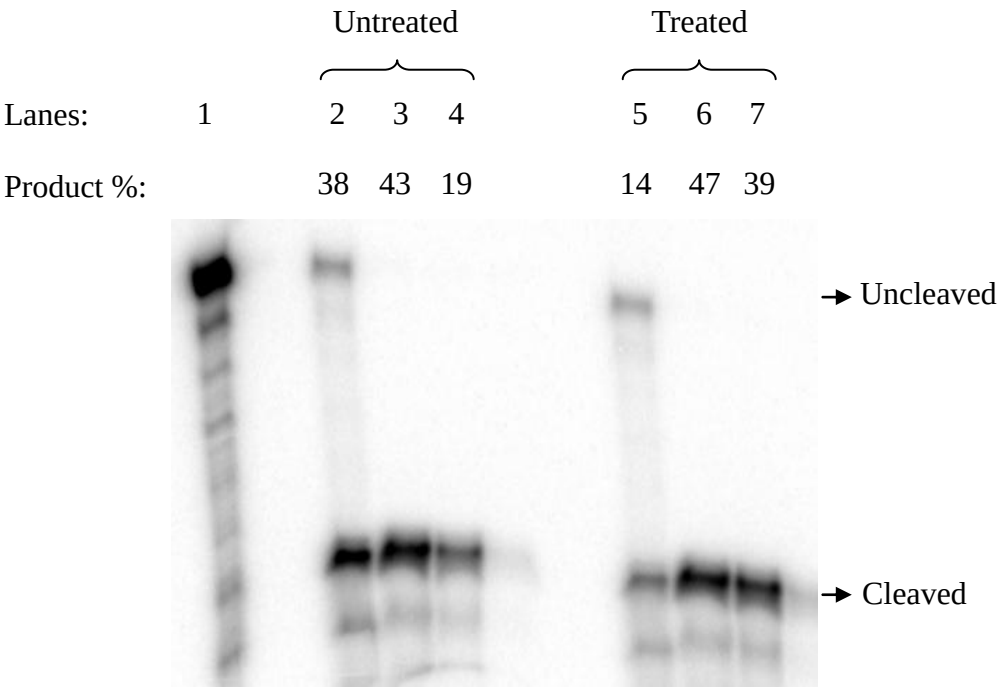


Figure 5.11 Use of salmon sperm DNA in the cleavage of STA17. STA17 divided into three samples: one control (lane 1), one carried out as usual (lanes 2–4) and one using incubation buffer that contained salmon sperm DNA (0.1 mg/mL) (lanes 5–7). Lane 1, control; lanes 2 and 5, streptavidin magnetic particles; lanes 3 and 6, 30 min elution; lanes 4 and 7, 180 min elution. Reaction conditions: 50 mM HEPES pH 7.4, 200 mM NaCl, 1 mM $ZnSO_4$, 10 mM $MgCl_2$. Denaturing PAGE was performed on 7 % gels.

Looking back at the progress of the selection in Figure 5.2, it can be concluded that the sugars played an important role. Of course, the carbohydrates had no effect on the self-cleavage of the library. However, it did have an influence on how many cleaved species were retained on the solid support. In the tenth round of selection, a common starting pool of sequences gave rise to much more activity in the presence of ATP than in the presence of stachyose. As well, when the concentration of stachyose was decreased from 1 mM to 0.2 mM in round 12 of the stachyose selection, a significant decrease in activity was observed. These observations show that the presence of either ATP or stachyose promoted the isolation of cleaved DNAzymes from the magnetic particles with ATP having a greater effect than stachyose. It is unclear how this process occurs, but ATP, being the monomer of RNA, may compete with cleaved sequences for the association with the solid support that is independent of biotin/streptavidin interactions. Looking back at Figure 5.2 again as well as Figure 5.4, it can be concluded that this process was not very fast since the addition of sugars to the DNAzymes did not immediately lead to significant amounts of material being isolated right away in the first minute of positive selection. The discovery of non-sugar-sensing DNAzymes such as STA17 was promoted by the ability of the sugars to facilitate the isolation of cleaved species. This was probably one of the primary reasons why self-cleavage activity was present in the sugar-sensing selection and absent in the non-sugar-sensing selection.

5.5.3 Relative activity of the isolated DNAzymes.

As was pointed out earlier (Table 5.3), compared to the ATP selection, the stachyose selection gave rise to a higher number of DNAzymes that also exhibited superior

catalytic activity. This result was probably due to the influence of the long negative selections that were performed on the ATP selection in the final two rounds of selection. During the longer negative selection, the more active species would be cleaved and subsequently removed from the selection. Although the cloned material from the stachyose selection went through four fewer rounds of selection, it was never treated with prolonged negative selections and hence gave rise to superior DNAzymes. The cloning of generation 11 of the ATP selection would likely have given rise to DNAzymes that displayed catalytic activity comparable to those from the stachyose selection. Based on the above reasoning, catalysts of greater activity than STA17 could have been isolated in the absence of negative selections altogether.

5.6 Discussion and conclusions.

In vitro selection enables the evaluation and isolation of a small number of nucleic acid sequences that perform a certain desired task out of a library of up to 10^{16} individually different sequences. The technique is very similar to a screening process, and in theory, if each round of selection that was performed managed to remove 90 % of the remaining unwanted sequences, after 16 rounds of selection, 10^{16} sequences would be reduced to only 1 (10^{0}) sequence. However, in reality this does not happen. *In vitro* selection is a long and laborious process where many different factors can affect the final outcome. The use of modified nucleotides further complicates the process. The synthesis of the modified library is more difficult than the synthesis of an unmodified library. As a result, many sequences may not be fully extended. As well, modified sequences that are isolated may not amplify properly during a 30 cycle PCR. It is known that the high PCR amplification of natural oligonucleotide templates can favour certain sequences over

others, potentially suppressing the amplification of desired active species.[248,249] This would be even more significant for modified oligonucleotides where modified nucleotides are poorly read through by polymerases. Finally, a selection must be carried out and cloned before it is known whether the selection was successful.

During our selection, it turned out that our selection method was capable of isolating DNAzymes, but was unable to separate sugar-specific species from non-sugar-specific species. The suspected root of the problem was the affinity of cleaved DNAzymes for the streptavidin-coated support. This was an unexpected development since the use of biotin and streptavidin/avidin has been exploited in a large number of selections including selections that have been used for modified residues as well. Perhaps, the phenol modification increased the affinity of DNA for the streptavidin magnetic particles. It should be noted that the ratio of sugar-specific to sugar-non-specific sequences in the original library was probably very low. This is because non-specific sequences without a sugar binding domain will tend to be shorter in length, and shorter sequences are present in greater numbers in a random library. If we hypothesize that in order to accommodate allosteric binding, the specific DNAzymes on average required a functional motif that was 5 bases longer than what was required for the non-specific DNAzymes, then the ratio of specific/non-specific sequences in the original pool would be $1:4^5$. Based on these theoretical numbers, even if cleaved sequences did not remain attached to the streptavidin-coated support, the selection would have still been very difficult to successfully perform. At a minimum though, we would have been in a much better position to monitor the selection progress and control the selection pressure. We would not have been led to believe that sugar-specific DNAzyme activity in the selection was present, and the selection would probably not have been cloned.

It is hard to predict how the selection would have progressed if indeed the boronic acid modification were present. Boronic acid groups would certainly have given potential DNAzymes a much better ability to bind (enhanced sensitivity) as well as differentiate target carbohydrates (enhanced specificity). The boronic acid-modified DNAzymes may have had less affinity for the streptavidin magnetic particles. That certainly would have changed the dynamics of the selection. However, if the boronic acid-modified DNA has affinity for the solid support in the same manner as the phenol-modified DNA and could catalyze divalent metal-dependent self-cleavage, then the negative selection difficulties that plagued the selection that used phenol-containing **2.24** would also affect a selection that uses the phenylboronic acid-containing **2.12**.

In conclusion, the *in vitro* selection that was originally aimed at isolating a sugar-specific allosteric DNAzyme was unsuccessful and ultimately led to the discovery of STA17. Using phenol-modified nucleotides, this divalent metal-dependent DNAzyme was found to perform self-cleavage with a rate constant of 0.35 ± 0.03 min^{-1} in the presence of Mn^{+2} (1 mM) and a rate constant of 0.20 ± 0.02 min^{-1} in the presence of Mg^{+2} (10 mM) and Zn^{+2} (1 mM). The self-cleaving sequence was found to be inhibited in the presence of divalent mercury at a concentration of 30 μM and could potentially be converted to a mercury-sensing DNAzyme.

5.7 Experimental

General. Reagents were obtained from Fisher Scientific, Bio-Rad Laboratories or Sigma-Aldrich. Vent (exo–) DNA polymerase, *Taq* DNA polymerase and λ-exonuclease were obtained from New England Biolabs. Sheared salmon sperm DNA was from Eppendorf. Steptavidin magnetic particles were purchased from Roche. pGEM-T-Easy Vector Systems kit was purchased from Promega. α-^{32}P-dGTP was obtained from Perkin Elmer. Metal salts used: $FeSO_4 \cdot 7H_2O$, $Fe(NO_3)_3 \cdot 9H_2O$, $BaCl_3$, $EuCl_3 \cdot xH_2O$, $CuSO_4 \cdot 5H_2O$ $ZnSO_4 \cdot 7H_2O$, $MnCl_2$, $NiCl_2 \cdot 6H_2O$ $CaCl_2$, $La(OAc)_3$, $MgCl_2 \cdot 6H_2O$, $MgSO_4 \cdot 7H_2O$, $Hg(OAc)_2$.

5.7.1 *In vitro* selection

Buffers/solutions.

TEN (tris, EDTA, NaCl): 40 mM Tris pH 7.5, 1 mM EDTA, 150 mM NaCl.

Loading solution: Formamide (27 mL), $EDTA_{(aq)}$ (3 mL, 0.5 M), xylene $cyanol_{(aq)}$ (300 μL, 0.05 %) and bromophenol $blue_{(aq)}$ (300 μL, 0.05 %).

Elution buffer: 1 % $LiClO_4$, 10 mM Tris pH 7.0 in water.

Selection buffer 1: 50 mM HEPES pH 7.4, 200 mM NaCl, 0.1 mM $ZnSO_4$, 0.5 mM $MgCl_2$.

Selection buffer 2: 50 mM HEPES pH 7.4, 200 mM NaCl, 0.1 mM $ZnSO_4$, 1 mM $MgCl_2$.

Selection buffer 3: 50 mM HEPES pH 7.4, 200 mM NaCl, 1 mM $ZnSO_4$, 10 mM $MgCl_2$.

First amplification cocktail (5X): MH4 (7.5 nmol), MH5 (9.0 nmol), dNTP's (345 nmol each), $MgSO_4$ (1.15 μmol), thermopol (115 μL, 10X) and water for a final volume of 230 μL.

Second amplification cocktail (5X): MH5 (20 nmol), MH14 (15 nmol), dNTP's (690

nmol each), $MgSO_4$ (3.45 μmol), thermopol (230 μL, 10X) and water for a final volume of 460 μL.

Kinetics quenching buffer: loading solution/100 mM biotin in DMF (99:1).

Oligonucleotides. (5′–3′)

Selection primer: Biotin-T_{20}GCGTGCC**rC**GTCTGTTGGGCCCTACCAACA

Library template:

GAGCTCGCGGGGCGTGCN_{40}CTGTTGGTAGGGCCCAACAGACG

PCR primers: MH4 phosphate-CGTCTGTTGGGCCCTACCA

MH5 GAGCTCGCGGGGCGTGC

MH14

phosphate-ACGACACAGAGCGTGCCCGTCTGTTGGGCCCTACCA

STA17 template:

GGGGCGTGCTTTGCACTCGATAGTCAGCAGTGCCTTCGTCAACTAGTTTCTGT

TGGTAGGGCCCAACAGACGGGCACGCTCTGTGTC

Denaturing PAGE. Polyacrylamide gels were made with acrylamide/bis-acrylamide 29:1 ratio. The gels were polymerized in the presence of TBE (1X) and used TBE (1X) as the electrolytic solution. Purifying gels were allowed to polymerize for at least 6 h prior to use. Oligonucleotides were analyzed and/or purified using one of two different sized gels: a regular sized sequencing gel 42 × 33 × 0.04 cm (L × W × D) or a minigel 170 × 165 × 1 mm. All oligonucleotide samples were heated for ~5 min at 95 ℃ and cooled to ~0 ℃ prior to loading. Gels were visualized by either autoradiography or UV shadowing. Radioactivity was detected with a Typhoon 9200 phosphorimaging scanner from GE

healthcare.

Gel elution. Gel bands were cut out and then broken up with a flame-sealed pipette tip. Elution buffer was added to the resulting paste like material, and the suspension was agitated with a vortexer and heated on a heat block at 65 ℃ for 15 min. The mixture was centrifuged, and the supernatant was collected. This elution process was repeated once for the large gels and twice for the minigels. Elutions were pooled and evaporated on a vacuum centrifuge either at room temperature for modified oligonucleotides or at 45 ℃ for natural oligonucleotides.

Spin columns. Columns prepared by loading G25 or G10 Sephadex into 1 mL pipette tips that have been plugged with silanized glass wool. Short spin columns and normal sized spin columns were ~2 and 5 cm, respectively.

PCR. PCR was performed for 30 cycles: 15 s at 54 ℃, 40 s at 75 ℃ and 15 s at 95 ℃ on a Techne TC-312.

Elongation. Template (15 μL, 30 pmol), primer (3.0 μL, 30 pmol) and thermopol buffer (4.0 μL, 10X) were combined and heated to 100 °C. The solution was slowly cooled to ~35 ℃ over a period of ~45 min. A solution of nucleotides consisting of dATP (2.0 μL, 2.0 nmol), **2.24** (2.0 μL, 2.0 nmol), dCTP (1.0 μL, 1.0 nmol), dGTP (1.0 μL, 1.0 nmol), α-^{32}P-dGTP (2.0 μL), DTT solution (0.4 μL, 40 nmol) and water (8.1 μL) was added to the annealing solution. The primer extension was initiated by the addition of Vent (exo–) (1.5 μL, 3 U) and covered with mineral oil (20 μL). The reaction was carried out for 1 h

at 72 ℃ and then quenched by the addition of $EDTA_{(aq)}$ (2.0 μL, 0.5 M). This solution was either used for the next step shortly after or stored in the freezer for use on a subsequent day.

Final concentrations for elongation:

Primer	0.75	μM
Template	0.75	μM
dCTP/dGTP	25	μM
dATP/**2.24**	50	μM
DTT	1	mM

Selection. A solution of streptavidin magnetic particles (50 μL) was magnetized and decanted. The particles were rinsed and magnetized 3 times with TEN buffer (100 μL each). The elongation reaction was then added to the particles and left at room temperature for 30 min with occasional agitation with a vortexer. The solution was magnetized and decanted. The particles were then rinsed, magnetized and decanted with a series of solutions: 2 solutions of TEN buffer (100 μL each), 5 × 100 μL solutions containing NaOH (0.1 M) and EDTA (1.0 mM), a solution of HEPES buffer (200 μL, 50 mM, pH 7.4) and water (100 μL). Negative selection buffer (100 μL) was added to the magnetic particles. The particles were left at room temperature with occasional agitation with a vortexer. After 1 h, the mixture was magnetized and decanted. This decanted liquid was collected as the negative selection. The 1 h incubation step was repeated with positive selection buffer (100 μL).

Selection work up. To the decanted selection solution was added 1st amplification cocktail (2 μL) and 1 % $LiClO_4$ in acetone (1 mL), and the resulting solution was agitated

with a vortexer, centrifuged for 15 min and decanted. Ethanol (0.5 mL) was then added to the pellet. This suspension was agitated with a vortexer, centrifuged for 15 min and decanted as well. Residual ethanol was removed by vacuum. Loading solution (20 μL) was added to the precipitated selection material in preparation for purification by denaturing PAGE.

Control lane preparation for the purification was prepared from the remaining magnetic particles from the selection. The particles were suspended in water (10 μL). This sample was divided into two by removing 2 × 5 μL samples for positive and negative controls. To the negative control was added loading solution (198 μL) and 100 mM biotin in DMF (2 μL). The positive control was first treated with 1 M NaOH (2 μL) for 20 min at 65 °C prior to the addition of loading solution and biotin solution.

The control and selection samples were heated at 95 °C for 5 min and then cooled rapidly to 0 °C. The samples (20 μL each) were loaded onto denaturing PAGE (7 %). Selection bands that lined up with the positive control were cut out, broken up and eluted. After the elutions were evaporated to ~100 μL, to each of the mixtures were added ethanol (1 mL) and 1st amplification cocktail (2 μL). The samples were agitated by a vortexer, centrifuged for 15 min, decanted and dried by vacuum. Each sample was resuspended in water (30 μL) and loaded onto a short G25 spin column.

First PCR amplification. A control sample was made by mixing N40 template (1 μL, 2 pmol) with water (30 μL). To each selection sample was added 1st amplification cocktail (8 μL), α-^{32}P-dGTP (0.75 μL) and Vent (exo–) (1 μL, 2U). Each sample was transferred to two PCR vials and thermocycled using a PCR machine. Each pair of samples was recombined. To each recombined sample was added phenol/chloroform/isoamyl alcohol (40 μL). Samples were agitated by a vortexer and centrifuged for 2 min. Ethanol (400 μL)

was added to each sample, and the mixtures were agitated on a vortexer, centrifuged for 15 min, decanted and dried on a heat block at 65 ℃. To each sample control, non-sugar selection, sugar selection were subsequently added water (35 μL), λ-exonuclease buffer (4 μL, 10X) and λ-exonuclease (1 μL, 5U). After incubation at room temperature for at least 90 min, 3 % $LiClO_4$ in acetone (300 μL) was added to each sample. The mixtures were agitated with a vortexer, centrifuged for 15 min and decanted. Ethanol (400 μL) was added to each sample, and the mixtures were agitated with a vortexer, centrifuged for 15 min, decanted and dried on a heat block at 65 ℃. To each sample was added water (15 μL), loading solution (15 μL) and loaded onto denaturing PAGE (10 %). Resulting radioactive gel bands were cut out, broken up and eluted. Each elution was reduced to ~100 μL. 3 % $LiClO_4$ in acetone (300 μL) was added to each sample, and the suspensions were agitated with a vortexer, centrifuged for 15 min and decanted. Ethanol (400 μL) was then added to each sample, and the suspensions were agitated with a vortexer, centrifuged for 15 min, decanted and dried on a heat block at 65 ℃.

Second PCR amplification. Water (25 μL) was added to each sample, and the mixtures were centrifuged for 2 min. To a 6 μL amount of supernatant from the first amplification was added 2^{nd} amplification cocktail (40 μL), water (150 μL) and Vent (exo–) (9 μL, 18 U). A control reaction was performed using N40 template (1 μL, 2 pmol), 2^{nd} amplification cocktail (20 μL), water (75 μL) and Vent (exo–) (4.5 μL, 9 U). The selection was divided into 4 PCR tubes, and the control was divided into 2 PCR tubes. All the sample tubes were run on the PCR machine. The PCR tube samples were then recombined into 100 μL samples. Phenol/chloroform/isoamyl alcohol (100 μL) was added to each sample, and the mixtures were agitated with a vortexer and centrifuged for 2 min. Each top layer was isolated, to which ethanol (1 mL) was added. The mixtures

were agitated with a vortexer, centrifuged for 15 min, decanted and dried on a heat block at 65 °C. Water (90 μL for control and 45 μL for selections) was added, each selection pair was recombined. To each of the three sample tubes was added λ-exonuclease buffer (10 μL, 10X) and λ-exonuclease (1 μL, 5U). Digestions were left standing at room temperature overnight. Samples were again precipitated first with phenol/chloroform and then ethanol. To each sample/control was added loading solution (40 μL) and NaOH (1 μL, 1M). Samples were purified on a minigel and product bands were identified by UV shadowing. Elutions were reduced to ~100 μL by using a vacuum centrifuge and precipitated with 10 volumes of ethanol. The resulting solid sample was resuspended in water (50 μL) and purified on a G25 spin column. The sample was diluted to 400 μL for UV quantification.

Cloning. Second PCR amplification product was used as template for PCR with *Taq* DNA polymerase to produce double stranded material with overhanging 3′-dA's. TA-cloning was then performed on this product using the pGEM-T-Easy Vector Systems kit. Using electroporation, the resulting plasmids were transformed into *E. coli* DH10B cells, and the cells were streaked onto lysogeny broth agar containing ampicillin (100 mg/L). Individual white colonies were picked and used to inoculate 1 mL samples of Terrific Broth-containing Plasmid Miniprep Kit (from Invitrogen). The mixture was restriction digested in the presence of EcoRI to screen for appropriately sized inserts. Single insert plasmids were submitted to the UBC Nucleic Acid Protein Service Unit for sequencing.

Selection conditions.

General: all selections performed at room temperature.

Non-sugar-sensing DNAzyme. Incubation carried out for 1 h in the presence of selection

buffer 1 (G1–G3) or selection buffer 2 (G4–G9).

Sugar-sensing DNAzyme. Negative selection performed for 1 h in selection buffer 1 (G1–G3) or selection buffer 2 (G4–G9). The positive selection was performed with negative selection buffer with 1 mM each of stachyose, sialic acid, D-mannoheptose and ATP. G8–G9: 3 separate positive selection solutions were used (0–1, 1–5, 5–60 min).

Stachyose selection. Negative selection carried out for 1 h (G10–G11) with selection buffer 1, 2 h (G12–G13) with selection buffer 2 or 2 h (G14–15) with selection buffer 3. Three separate positive selection solutions were used (0–1, 1–5, 5–60 min). Positive selection performed for 1 h in selection buffer 2 with 1 mM stachyose (G10–G11), selection buffer 2 with 0.2 mM stachyose (G12–G13) or selection buffer 3 with 0.2 mM stachyose (G14–G15).

ATP selection. Negative selection performed in selection buffer 2 for 1 h (G10–G11), selection buffer 2 for 1 h followed by selection buffer 2 with 1 mM dATP for 1 h (G12–G13) or selection buffer 3 for 1 h followed by selection buffer 3 with 1 mM dATP for 5 h (G14–G15). Three separate positive selection solutions were used (0–1, 1–5, 5–60 min). Positive selection carried out for 1 h in selection buffer 2 (G10–G13) with 1 mM ATP or selection buffer 3 (G14–G15) with 1 mM ATP.

5.7.2 Kinetic analysis

The catalyst was prepared in the same way as the modified library in the selection as described in section 5.7.1. Template (75 μL, 150 pmol), primer (15.0 μL, 150 pmol) and thermopol buffer (20.0 μL, 10X) were combined and heated to 100 °C. The solution was slowly cooled to ~35 ℃ over a period of ~45 min. A solution of nucleotides consisting of dATP (10.0 μL, 10.0 nmol), **2.24** (10.0 μL, 10.0 nmol), dCTP (5.0 μL, 5.0 nmol), dGTP

(5.0 μL, 5.0 nmol), α-^{32}P-dGTP (10.0 μL), DTT solution (2.0 μL, 200 nmol) and water (40.5 μL) was added to the annealing solution. The primer extension was initiated by the addition of Vent (exo–) (7.5 μL, 15 U) and covered with mineral oil (50 μL). The reaction was carried out for 1 h at 72 °C and then quenched by the addition of EDTA (10.0 μL, 5.0 μmol). A 36 μL sample of the primer extension was removed, and the rest was stored in the freezer. Streptavidin magnetic particle solution (45 μL) was washed 3 times with TEN buffer (100 μL), and the 36 μL primer extension sample was added to the particles. The mixture was left at room temperature for 30 min with periodic agitation on a vortexer. The particles were then rinsed, magnetized and decanted with a series of solutions: 2 solutions of TEN buffer (100 μL each), 5 × 100 μL solutions containing NaOH (0.1 M) and EDTA (1.0 mM) and a solution of HEPES buffer (200 μL, 50 mM, pH 7.4) and NaCl (200 mM). Water (264 μL) was added to the particles, and the suspension was divided into 3 portions for 3 identical self-cleavage reactions. For each reaction, 2 × 4 μL aliquots were removed for positive and negative controls. The water was decanted, and incubation buffer (80 μL) was added. Aliquots (4 μL each) were removed and added directly to quenching buffer (12 μL each) at 17 time points: 0.75, 1.5, 2.25, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, 20, 30, 60 min. Solutions were analyzed by denaturing PAGE (7 %).

Inhibition experiments. Primer extension product solution (4 μL / reaction) was incubated with streptavidin magnetic particles and treated in the usual way. After the decantation of HEPES buffer, water (30 μL / reaction) was added. Aliquots (4 μL each) were removed for positive/negative controls. Following the removal of water, a 22 μL solution of inhibitor (100 μM for divalent and trivalent metal cations or saturated diazinon/fenitrothion), $MgSO_4$ (10 mM), $ZnSO_4$ (1mM), HEPES (50 mM) pH 7.4, NaCl (200 mM) was added. Aliquots (4 μL each) were removed at times: 0.75, 3, 10 min and

added directly to quenching buffer (12 μL each).

Rescue experiments. These were predominantly carried out in the same way as the inhibition experiments. Following the removal of water, a 22 μL solution of divalent metal ion (1 mM), HEPES (50 mM) pH 7.4, NaCl (200 mM) was added. Aliquots were removed after: 5, 60 min and overnight reaction.

pH rate profile. The single-stranded STA17 DNAzyme on streptavidin magnetic particles was incubated in the presence of 50 mM buffer, 200 mM NaCl, 1 mM $ZnSO_4$, 10 mM $MgCl_2$. Experiments were performed at 6 different pH's: 5.9, 6.4, 6.9, cacodylic acid; 7.9, HEPES; 8.4, 8.9, TRIS. Aliquots were removed and quenched at 0.75, 2, 5, 12, 30, 75, 180 min. The rate of the reaction at pH 8.9 was very slow, so slow that a numerical value could not be obtained from the data. The kinetic experiments at each pH were performed once.

Chapter 6: Summary and future work

6.1 Summary of research

This thesis has addressed several aspects regarding the use of modified nucleotides with enhanced chemical functionality in DNAzymes. Chapters 2 and 3 discussed the synthesis of eight different modified nucleoside triphosphates with modifications tethered from either the 5-position of 2′-deoxyuridine or the 8-position of 2′-deoxyadenosine. Seven of these modified nucleotides were used in primer extension experiments with DNA polymerases for the synthesis of modified oligonucleotides. The modifications consisted of guanidinium groups (arginine-like), phenols (tyrosine-like), phenylboronic acids and imidazoles (histidine-like), all groups that have been appended to DNA either previously or while work for this thesis was being performed. Effort was made to keep the linkers that attached the modifications to the nucleobase relatively short. The use of these modified dNTP's in the *in vitro* selection of DNAzymes will allow the evaluation of whether shorter linker arms will eventually give rise to DNAzymes of superior catalytic rates. The incorporation studies of **2.11** and **2.24** reinforced what is now becoming obvious from work in our lab and in others: 5-position-modified dUTP's are generally good substrates for DNA polymerases. Both of the 2′-deoxyuridine derivatives that were tested were found to be incorporated efficiently by polymerases, particularly those from the B family, under primer extension and PCR conditions. The modified nucleotides studied in Chapter 3, however, were much more difficult to incorporate. In the literature, modifications have been attached to the 7-position of 7-deaza-2′-deoxyadenosine. In particular, Capek *et al.*[95] synthesized modified 2′-deoxyadenosine triphosphates with identical modifications attached at either the 8-position or the 7-deaza-position. Only the

7-modified-7-deaza-2′-deoxyadenosines were successfully incorporated. Their results and the results of others[84,94] have shown that future modifications on dA should be attached to the 7-position of 7-deaza-2′-deoxyadenosine. Chapter 4 illustrated the effect of incorporating modifications of **2.11** into oligonucleotides in the context of oligonucleotide duplexes and DNAzyme 10-23 substrate binding regions. Both situations resulted in more stable structures. Although no catalytic enhancement in the modified 10-23's was observed, the catalysts may display superior membrane permeability and are presently being tested by our collaborators. Chapter 5 described the steps that were taken to isolate a modified DNAzyme, STA17, starting from a library synthesized using **2.24**. The rates of the self-cleaving sequence STA17 were found to be highest in the presence of Mn^{+2}, and the DNAzyme was inhibited by Hg^{+2}.

6.2 Future directions

It has been 10 years since the first DNAzyme with extra chemical functionality was published,[171] but much has yet to be accomplished. From a therapeutic perspective, the most attractive characteristic of DNAzymes is their potential for sequence-specific RNA cleavage.[47-50] However, to this date there has not been an unmodified or modified DNAzyme discovered that can perform rapid cleavage of RNA in the absence of divalent metal cations with multiple turnover. With this goal in mind, more modified DNAzyme selections designed to cleave RNA sequences (rather than DNA/RNA chimeric sequences) in the absence of divalent metal cations with the help of imidazoles and amines will need to be performed.

Since research for this dissertation was first being performed, much progress has been made in terms of the synthesis and enzymatic incorporation of modified dNTP's.

There is now a very wide range of modifications that have been appended to dNTP's for use in *in vitro* selections. All the modified dNTP's in this dissertation that were incorporated with DNA polymerases will be used for selections. For instance, the modified dATP's in Chapter 3 that displayed polymerase incorporation will be used to evaluate the importance of the linker arm that attaches the functional group to the nucleotide. Hipolito *et al.* has recently performed an *in vitro* selection for ribonucleotide phosphodiester bond cleavage in the absence of divalent metal cations using **3.2**.[250]

The search for sugar-sensing DNAzymes should be continued. This will require another selection to be performed that will have to take into account some of the difficulties that were dealt with in Chapter 5. This selection could perhaps be performed with three modifications: imidazoles, amines and boronic acids. Such a system would give rise to DNAzymes that will be able to 1) catalyze the destruction of RNA in the absence of divalent metal cations and 2) recognize carbohydrates effectively with the boronic acids. If triphosphate **2.12** is found to be a substrate for Sequenase V2.0, then the nucleoside triphosphate can be used with **3.1** and the commercially available 5-aminoallyl-2′-deoxycytidine triphosphate for the selection. Otherwise, another imidazole-modified dNTP will need to be substituted for **3.1** since **3.1** is only effectively incorporated by Sequenase V2.0.

In terms of the use of functionally enhanced nucleotides for *in vitro* selection, both the incorporation of modified nucleotides and the use of modified oligonucleotides as templates for PCR can be improved. To improve incorporations, two DNA polymerases can be used simultaneously. In such a binary system, one polymerase could compensate for the deficiencies of another polymerase to give rise to modified DNA that cannot be obtained by either DNA polymerase on its own. While exploring the incorporation of

nucleotides with a system that included a non-natural base pair, Tae *et al.* achieved full extension of their template in the simultaneous presence of both DNA polymerases Klenow and Pol β.[251] In order to improve the PCR amplification of modified DNAzymes that are collected during the selection process, the modified sequences could first be reverse transcribed to an unmodified sequence prior to PCR amplification. Such a protocol would allow more heavily modified sequences to be amplified equally as well as less heavily modified sequences. To explore this, we performed an experiment using DNAzyme STA17 as shown in Figure 6.1 (details can be found in the appendix). After the cleaved sequences of STA17 were isolated, they were annealed with a primer containing a single ribonucleotide. Primer extension with natural dNTP's resulted in the synthesis of two complementary strands of the same length. After base digestion of the duplex, the two strands were separable by denaturing PAGE. The unmodified strand would then be used as the template for 1st amplification. Of course, since triphosphate **2.24** can be used in PCR, these steps were not necessary for our selection and were not used. STA17 was only used to verify that such a protocol could be carried out. This protocol adds extra steps to the *in vitro* selection process, but the extra steps may pay dividends for selections that have very heavily modified sequences such as the selections that were performed by Hollenstein *et al.*[137,188]

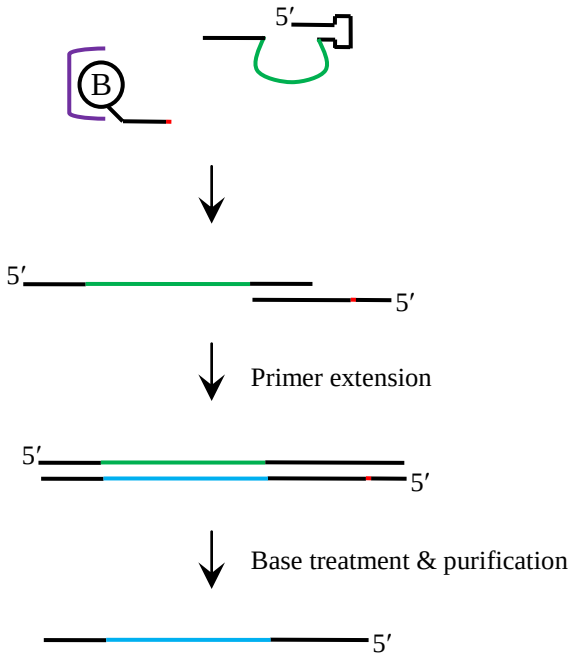


Figure 6.1 Transcription of modified DNA to unmodified DNA. (——) DNA primer regions; (——) RNA; (——) modified N_{40} random region; (——) unmodified N_{40} template region; (B) biotin; (——) streptavidin magnetic particles.

Finally, instead of trying to synthesize modified dNTP's that can be incorporated by DNA polymerases, the DNA polymerases can be mutated to incorporate modified dNTP's.[252] Two methods that are presently being used for the evolution of DNA polymerases are phage display[253,254] and compartmentalized self-replication (CSR).[255-257] These methods can address both the incorporation of modified dNTP's and the use of modified DNA as templates for PCR. The evolution of *Taq* by CSR is presently being pursued in our lab.

References

(1) Guerriertakada, C.; Gardiner, K.; Marsh, T.; Pace, N.; Altman, S. *Cell* **1983**, *35*, 849-857.

(2) Forster, A. C.; Symons, R. H. *Cell* **1987**, *50*, 9-16.

(3) Prody, G. A.; Bakos, J. T.; Buzayan, J. M.; Schneider, I. R.; Bruening, G. *Science* **1986**, *231*, 1577-1580.

(4) Kruger, K.; Grabowski, P. J.; Zaug, A. J.; Sands, J.; Gottschling, D. E.; Cech, T. R. *Cell* **1982**, *31*, 147-157.

(5) Nissen, P.; Hansen, J.; Ban, N.; Moore, P. B.; Steitz, T. A. *Science* **2000**, *289*, 920-930.

(6) Noller, H. F.; Hoffarth, V.; Zimniak, L. *Science* **1992**, *256*, 1416-1419.

(7) Gilbert, W. *Nature* **1986**, *319*, 618-618.

(8) Orgel, L. E. *Trends Biochem. Sci.* **1998**, *23*, 491-495.

(9) Tuerk, C.; Gold, L. *Science* **1990**, *249*, 505-510.

(10) Ellington, A. D.; Szostak, J. W. *Nature* **1990**, *346*, 818-822.

(11) Robertson, D. L.; Joyce, G. F. *Nature* **1990**, *344*, 467-468.

(12) Jayasena, S. D. *Clin. Chem.* **1999**, *45*, 1628-1650.

(13) Wilson, D. S.; Szostak, J. W. *Annu. Rev. Biochem.* **1999**, *68*, 611-647.

(14) Marshall, K. A.; Ellington, A. D. *Methods Enzymol.* **2000**, *318*, 193-214.

(15) Nimjee, S. M.; Rusconi, C. P.; Sullenger, B. A. *Annu. Rev. Med.* **2005**, *56*, 555-583.

(16) Bunka, D. H. J.; Stockley, P. G. *Nat. Rev. Microbiol.* **2006**, *4*, 588-596.

(17) Seelig, B.; Keiper, S.; Stuhlmann, F.; Jaschke, A. *Angew. Chem. Int. Ed.* **2000**,

39, 4576-4579.

(18) Tarasow, T. M.; Tarasow, S. L.; Eaton, B. E. *Nature* **1997**, *389*, 54-57.

(19) Agresti, J. J.; Kelly, B. T.; Jaschke, A.; Griffiths, A. D. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 16170-16175.

(20) Sengle, G.; Eisenfuhr, A.; Arora, P. S.; Nowick, J. S.; Famulok, M. *Chem. Biol.* **2001**, *8*, 459-473.

(21) Tsukiji, S.; Pattnaik, S. B.; Suga, H. *Nat. Struct. Biol.* **2003**, *10*, 713-717.

(22) Nieuwlandt, D.; West, M.; Cheng, X. Q.; Kirshenheuter, G.; Eaton, B. E. *Chembiochem* **2003**, *4*, 651-654.

(23) Ryu, Y.; Kim, K. J.; Roessner, C. A.; Scott, A. I. *Chem. Commun.* **2006**, 1439-1441.

(24) Illangasekare, M.; Sanchez, G.; Nickles, T.; Yarus, M. *Science* **1995**, *267*, 643-647.

(25) Lee, N.; Bessho, Y.; Wei, K.; Szostak, J. W.; Suga, H. *Nat. Struct. Biol.* **2000**, *7*, 28-33.

(26) Breaker, R. R.; Joyce, G. F. *Chem. Biol.* **1994**, *1*, 223-229.

(27) Santoro, S. W.; Joyce, G. F. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 4262-4266.

(28) Li, J.; Zheng, W.; Kwon, A. H.; Lu, Y. *Nucleic Acids Res.* **2000**, *28*, 481-488.

(29) Chandra, M.; Silverman, S. K. *J. Am. Chem. Soc.* **2008**, *130*, 2936-2937.

(30) Chandra, M.; Sachdeva, A.; Silverman, S. K. *Nat. Chem. Biol.* **2009**, *5*, 718-720.

(31) Carmi, N.; Shultz, L. A.; Breaker, R. R. *Chem. Biol.* **1996**, *3*, 1039-1046.

(32) Purtha, W. R.; Coppins, R. L.; Smalley, M. K.; Silverman, S. K. *J. Am. Chem. Soc.* **2005**, *127*, 13124-13125.

(33) Li, Y. F.; Sen, D. *Nat. Struct. Biol.* **1996**, *3*, 743-747.

(34) Pradeepkumar, P. I.; Hobartner, C.; Baum, D. A.; Silverman, S. K. *Angew. Chem. Int. Ed.* **2008**, *47*, 1753-1757.

(35) Li, Y. F.; Breaker, R. R. *J. Am. Chem. Soc.* **1999**, *121*, 5364-5372.

(36) Radzicka, A.; Wolfenden, R. *Science* **1995**, *267*, 90-93.

(37) Beaucage, S. L.; Caruthers, M. H. *Tetrahedron Lett.* **1981**, *22*, 1859-1862.

(38) Sinha, N. D.; Biernat, J.; Mcmanus, J.; Koster, H. *Nucleic Acids Res.* **1984**, *12*, 4539-4557.

(39) Beaucage, S. L.; Iyer, R. P. *Tetrahedron* **1992**, *48*, 2223-2311.

(40) Ogilvie, K. K.; Beaucage, S. L.; Schifman, A. L.; Theriault, N. Y.; Sadana, K. L. *Can. J. Chem.* **1978**, *56*, 2768-2780.

(41) Joyce, G. F. *Annu. Rev. Biochem.* **2004**, *73*, 791-836.

(42) Saiki, R. K.; Gelfand, D. H.; Stoffel, S.; Scharf, S. J.; Higuchi, R.; Horn, G. T.; Mullis, K. B.; Erlich, H. A. *Science* **1988**, *239*, 487-491.

(43) Breaker, R. R.; Joyce, G. F. *Chem. Biol.* **1995**, *2*, 655-660.

(44) Faulhammer, D.; Famulok, M. *Angew. Chem. Int. Ed.* **1996**, *35*, 2837-2841.

(45) Cruz, R. P. G.; Withers, J. B.; Li, Y. F. *Chem. Biol.* **2004**, *11*, 57-67.

(46) Santoro, S. W.; Joyce, G. F. *Biochemistry* **1998**, *37*, 13330-13342.

(47) Dass, C. R.; Choong, P. F. M.; Khachigian, L. M. *Mol. Cancer Ther.* **2008**, *7*, 243-251.

(48) Achenbach, J. C.; Chiuman, W.; Cruz, R. P. G.; Li, Y. *Curr. Pharm. Biotech.* **2004**, *5*, 321-336.

(49) Patil, S. D.; Rhodes, D. G.; Burgess, D. J. *AAPS J.* **2005**, *7*, E61-E77.

(50) Eckstein, F. *Exp. Opin. Biol. Ther.* **2007**, *7*, 1021-1034.

(51) Zhou, D. M.; Taira, K. *Chem. Rev.* **1998**, *98*, 991-1026.

(52) Kuimelis, R. G.; McLaughlin, L. W. *Chem. Rev.* **1998**, *98*, 1027-1044.

(53) Dennis, E. A.; Westheim.Fh *J. Am. Chem. Soc.* **1966**, *88*, 3432-3433.

(54) Westheim.Fh *Acc. Chem. Res.* **1968**, *1*, 70-78.

(55) Oivanen, M.; Kuusela, S.; Lonnberg, H. *Chem. Rev.* **1998**, *98*, 961-990.

(56) Perreault, D. M.; Anslyn, E. V. *Angew. Chem. Int. Ed.* **1997**, *36*, 432-450.

(57) Emilsson, G. M.; Nakamura, S.; Roth, A.; Breaker, R. R. *RNA* **2003**, *9*, 907-918.

(58) Harris, M. E.; Dai, Q.; Gu, H.; Kellerman, D. L.; Piccirilli, J. A.; Anderson, V. E. *J. Am. Chem. Soc.* **2010**, *132*, 11613-11621.

(59) Draper, D. E. *RNA* **2004**, *10*, 335-343.

(60) Anderson, C. F.; Record, M. T. *Annu. Rev. Phys. Chem.* **1995**, *46*, 657-700.

(61) Ikenaga, H.; Inoue, Y. *Biochemistry* **1974**, *13*, 577-582.

(62) Eichhorn, G. L.; Tarien, E.; Butzow, J. J. *Biochemistry* **1971**, *10*, 2014-2019.

(63) Breslow, R.; Huang, D. L. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 4080-4083.

(64) Geyer, C. R.; Sen, D. *Chem. Biol.* **1997**, *4*, 579-593.

(65) Faulhammer, D.; Famulok, M. *J. Mol. Biol.* **1997**, *269*, 188-202.

(66) Carrigan, M. A.; Ricardo, A.; Ang, D. N.; Benner, S. A. *Biochemistry* **2004**, *43*, 11446-11459.

(67) Takagi, Y.; Ikeda, Y.; Taira, K. *Top. Curr. Chem.* **2004**, *232*, 213-251.

(68) Raines, R. T. *Chem. Rev.* **1998**, *98*, 1045-1065.

(69) Perrin, D. M.; Garestier, T.; Helene, C. *J. Am. Chem. Soc.* **2001**, *123*, 1556-1563.

(70) Sidorov, A. V.; Grasby, J. A.; Williams, D. M. *Nucleic Acids Res.* **2004**, *32*, 1591-1601.

(71) Silverman, S. K. *Nucleic Acids Res.* **2005**, *33*, 6151-6163.

(72) Chaiet, L.; Wolf, F. J. *Arch. Biochem. Biophys.* **1964**, *106*, 1-5.

(73) Schlosser, K.; Lam, J. C. F.; Li, Y. F. *Nucleic Acids Res.* **2009**, *37*, 3545-3557.

(74) Bonaccio, M.; Credali, A.; Peracchi, A. *Nucleic Acids Res.* **2004**, *32*, 916-925.

(75) Bruesehoff, P. J.; Li, J.; Angustine, A. J.; Lu, Y. *Comb. Chem. High Throughput Screening* **2002**, *5*, 327-335.

(76) Higuchi, R. G.; Ochman, H. *Nucleic Acids Res.* **1989**, *17*, 5865-5865.

(77) Mulquiney, P. J.; Kuchel, P. W. *NMR Biomed.* **1997**, *10*, 129-137.

(78) Murphy, E.; Freudenrich, C. C.; Levy, L. A.; London, R. E.; Lieberman, M. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 2981-2984.

(79) Murphy, E.; Steenbergen, C.; Levy, L. A.; Raju, B.; London, R. E. *J. Biol. Chem.* **1989**, *264*, 5622-5627.

(80) Park, C.; Raines, R. T. *Biochemistry* **2003**, *42*, 3509-3518.

(81) Keefe, A. D.; Cload, S. T. *Curr. Opin. Chem. Biol.* **2008**, *12*, 448-456.

(82) Roychowdhury, A.; Illangkoon, H.; Hendrickson, C. L.; Benner, S. A. *Org. Lett.* **2004**, *6*, 489-492.

(83) Sakthivel, K.; Barbas, C. F. *Angew. Chem. Int. Ed.* **1998**, *37*, 2872-2875.

(84) Jager, S.; Rasched, G.; Kornreich-Leshem, H.; Engeser, M.; Thum, O.; Famulok, M. *J. Am. Chem. Soc.* **2005**, *127*, 15071-15082.

(85) Perrin, D. M.; Garestier, T.; Helene, C. *Nucleosides Nucleotides Nucleic Acids* **1999**, *18*, 377-391.

(86) Giller, G.; Tasara, T.; Angerer, B.; Muhlegger, K.; Amacker, M.; Winter, H. *Nucleic Acids Res.* **2003**, *31*, 2630-2635.

(87) Sarfati, S. R.; Pochet, S.; Guerreiro, C.; Namane, A.; Huynhdinh, T.; Igolen, J. *Tetrahedron* **1987**, *43*, 3491-3497.

(88) Wlassoff, W. A.; Dobrikov, M. I.; Safronov, I. V.; Dudko, R. Y.; Bogachev, V. S.; Kandaurova, V. V.; Shishkin, G. V.; Dymshits, G. M.; Lavrik, O. I. *Bioconjugate Chem.* **1995**, *6*, 352-360.

(89) Kolpashchikov, D. M.; Ivanova, T. M.; Boghachev, V. S.; Nasheuer, H. P.; Weisshart, K.; Favre, A.; Pestryakov, P. E.; Lavrik, O. I. *Bioconjugate Chem.* **2000**, *11*, 445-451.

(90) Lin, N.; Yan, J.; Huang, Z.; Altier, C.; Li, M. Y.; Carrasco, N.; Suyemoto, M.; Johnston, L.; Wang, S. M.; Wang, Q.; Fang, H.; Caton-Williams, J.; Wang, B. H. *Nucleic Acids Res.* **2007**, *35*, 1222-1229.

(91) Stolowitz, M. L.; Kaiser, R. J.; Patent, U., Ed. 1998.

(92) Kuwahara, M.; Hanawa, K.; Ohsawa, K.; Kitagata, R.; Ozaki, H.; Sawai, H. *Bioorg. Med. Chem.* **2006**, *14*, 2518-2526.

(93) Kuwahara, M.; Nagashima, J.; Hasegawa, M.; Tamura, T.; Kitagata, R.; Hanawa, K.; Hososhima, S.; Kasamatsu, T.; Ozaki, H.; Sawai, H. *Nucleic Acids Res.* **2006**, *34*, 5383-5394.

(94) Gourlain, T.; Sidorov, A.; Mignet, N.; Thorpe, S. J.; Lee, S. E.; Grasby, J. A.; Williams, D. M. *Nucleic Acids Res.* **2001**, *29*, 1898-1905.

(95) Capek, P.; Cahova, H.; Pohl, R.; Hocek, M.; Gloeckner, C.; Marx, A. *Chem. Eur. J.* **2007**, *13*, 6196-6203.

(96) Vaught, J. D.; Bock, C.; Carter, J.; Fitzwater, T.; Otis, M.; Schneider, D.; Rolando, J.; Waugh, S.; Wilcox, S. K.; Eaton, B. E. *J. Am. Chem. Soc.* **2010**, *132*, 4141-4151.

(97) Borsenberger, V.; Kukwikila, M.; Howorka, S. *Org. Biomol. Chem.* **2009**, *7*, 3826-3835.

(98) Gierlich, J.; Burley, G. A.; Gramlich, P. M. E.; Hammond, D. M.; Carell, T. *Org. Lett.* **2006**, *8*, 3639-3642.

(99) Ohbayashi, T.; Kuwahara, M.; Hasegawa, M.; Kasamatsu, T.; Tamura, T.; Sawai, H. *Org. Biomol. Chem.* **2005**, *3*, 2463-2468.

(100) Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 737-738.

(101) Lee, S. E.; Sidorov, A.; Gourlain, T.; Mignet, N.; Thorpe, S. J.; Brazier, J. A.; Dickman, M. J.; Hornby, D. P.; Grasby, J. A.; Williams, D. M. *Nucleic Acids Res.* **2001**, *29*, 1565-1573.

(102) Burgess, K.; Cook, D. *Chem. Rev.* **2000**, *100*, 2047-2059.

(103) Sawai, H.; Nagashima, J.; Kuwahara, M.; Kitagata, R.; Tamura, T.; Matsui, I. *Chem. Biodivers.* **2007**, *4*, 1979-1995.

(104) Boudsocq, F.; Iwai, S.; Hanaoka, F.; Woodgate, R. *Nucleic Acids Res.* **2001**, *29*, 4607-4616.

(105) Winship, P. R. *Nucleic Acids Res.* **1989**, *17*, 1266-1266.

(106) Jager, S.; Famulok, M. *Angew. Chem. Int. Ed.* **2004**, *43*, 3337-3340.

(107) Lermer, L.; Hobbs, J.; Perrin, D. M. *Nucleosides Nucleotides Nucleic Acids* **2002**, *21*, 651-664.

(108) Lermer, L.; Roupioz, Y.; Ting, R.; Perrin, D. M. *J. Am. Chem. Soc.* **2002**, *124*, 9960-9961.

(109) Thomas, J. M.; Yoon, J. K.; Perrin, D. M. *J. Am. Chem. Soc.* **2009**, *131*, 5648-5658.

(110) Suzuki, Y.; Komatsu, H.; Ikeda, T.; Saito, N.; Araki, S.; Citterio, D.; Hisamoto, H.; Kitamura, Y.; Kubota, T.; Nakagawa, J.; Oka, K.; Suzuki, K. *Anal. Chem.* **2002**, *74*, 1423-1428.

(111) Baum, D. A.; Silverman, S. K. *Cell. Mol. Life Sci.* **2008**, *65*, 2156-2174.

(112) Cairns, M. J.; Saravolac, E. G.; Sun, L. Q. *Curr. Drug Targets* **2002**, *3*, 269-279.

(113) Trepanier, J.; Tanner, J. E.; Momparler, R. L.; Le, O. N. L.; Alvarez, F.; Alfieri, C. *J. Viral Hepat.* **2006**, *13*, 131-138.

(114) Basu, S.; Sriram, B.; Goila, R.; Banerjea, A. C. *Antiviral Res.* **2000**, *46*,

125-134.

(115) Jakobsen, M. R.; Haasnoot, J.; Wengel, J.; Berkhout, B.; Kjems, J. *Retrovirology* **2007**, *4*, Article 29.

(116) Hou, Z.; Meng, J. R.; Niu, C.; Wang, H. F.; Liu, J.; Hu, B. Q.; Jia, M.; Luo, X. X. *Clin. Exp. Pharmacol. Physiol.* **2007**, *34*, 1160-1164.

(117) Zhang, L.; Gasper, W. J.; Stass, S. A.; Ioffe, O. B.; Davis, M. A.; Mixson, A. J. *Cancer Res.* **2002**, *62*, 5463-5469.

(118) Silverman, S. K. *Acc. Chem. Res.* **2009**, *42*, 1521-1531.

(119) Young, D. D.; Lively, M. O.; Deiters, A. *J. Am. Chem. Soc.* **2010**, *132*, 6183-6193.

(120) Uhlmann, E.; Peyman, A. *Chem. Rev.* **1990**, *90*, 543-584.

(121) Kurreck, J. *Eur. J. Biochem.* **2003**, *270*, 1628-1644.

(122) Akhtar, S.; Kole, R.; Juliano, R. L. *Life Sci.* **1991**, *49*, 1793-1801.

(123) Vester, B.; Lundberg, L. B.; Sorensen, M. D.; Babu, B. R.; Douthwaite, S.; Wengel, J. *J. Am. Chem. Soc.* **2002**, *124*, 13682-13683.

(124) Schubert, S.; Gul, D. C.; Grunert, H. P.; Zeichhardt, H.; Erdmann, V. A.; Kurreck, J. *Nucleic Acids Res.* **2003**, *31*, 5982-5992.

(125) Donini, S.; Clerici, M.; Wengel, J.; Vester, B.; Peracchi, A. *J. Biol. Chem.* **2007**, *282*, 35510-35518.

(126) Asanuma, H.; Hayashi, H.; Zhao, J.; Liang, X. G.; Yamazawa, A.; Kuramochi, T.; Matsunaga, D.; Aiba, Y.; Kashida, H.; Komiyama, M. *Chem. Commun.* **2006**, 5062-5064.

(127) Matsukura, M.; Shinozuka, K.; Zon, G.; Mitsuya, H.; Reitz, M.; Cohen, J. S.; Broder, S. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 7706-7710.

(128) Eckstein, F. *Antisense Nucleic Acid Drug Dev.* **2000**, *10*, 117-121.

(129) Singh, S. K.; Nielsen, P.; Koshkin, A. A.; Wengel, J. *Chem. Commun.*

1998, 455-456.

(130) Obika, S.; Nanbu, D.; Hari, Y.; Andoh, J.; Morio, K.; Doi, T.; Imanishi, T. *Tetrahedron Lett.* **1998**, *39*, 5401-5404.

(131) Lamond, A. I.; Sproat, B. S. *FEBS Lett.* **1993**, *325*, 123-127.

(132) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497-1500.

(133) Debart, F.; Abes, S.; Deglane, G.; Moulton, H. M.; Clair, P.; Gait, M. J.; Vasseur, J. J.; Lebleu, B. *Curr. Top. Med. Chem.* **2007**, *7*, 727-737.

(134) Hocek, M.; Fojta, M. *Org. Biomol. Chem.* **2008**, *6*, 2233-2241.

(135) Brudno, Y.; Liu, D. R. *Chem. Biol.* **2009**, *16*, 265-276.

(136) Weisbrod, S. H.; Marx, A. *Chem. Commun.* **2008**, 5675-5685.

(137) Hollenstein, M.; Hipolito, C. J.; Lam, C. H.; Perrin, D. M. *Chembiochem* **2009**, *10*, 1988-1992.

(138) Hollenstein, M.; Hipolito, C. J.; Lam, C. H.; Perrin, D. M. *Nucleic Acids Res.* **2009**, *37*, 1638-1649.

(139) Riordan, J. F.; Mcelvany, K. D.; Borders, C. L. *Science* **1977**, *195*, 884-886.

(140) Schug, K. A.; Lindner, W. *Chem. Rev.* **2005**, *105*, 67-113.

(141) Borders, C. L.; Riordan, J. F. *Biochemistry* **1975**, *14*, 4699-4704.

(142) Roig, V.; Asseline, U. *J. Am. Chem. Soc.* **2003**, *125*, 4416-4417.

(143) Prakash, T. P.; Puschl, A.; Lesnik, E.; Mohan, V.; Tereshko, V.; Egli, M.; Manoharan, M. *Org. Lett.* **2004**, *6*, 1971-1974.

(144) Wender, P. A.; Mitchell, D. J.; Pattabiraman, K.; Pelkey, E. T.; Steinman, L.; Rothbard, J. B. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 13003-13008.

(145) Ohsawa, K.; Kasamatsu, T.; Nagashima, J. I.; Hanawa, K.; Kuwahara, M.;

Ozaki, H.; Sawai, H. *Anal. Sci.* **2008**, *24*, 167-172.

(146) Davis, A. P.; Wareham, R. S. *Angew. Chem. Int. Ed.* **1999**, *38*, 2978-2996.

(147) Wang, W.; Gao, X. M.; Wang, B. H. *Curr. Org. Chem.* **2002**, *6*, 1285-1317.

(148) James, T. D.; Sandanayake, K. R. A. S.; Shinkai, S. *Angew. Chem. Int. Ed.* **1996**, *35*, 1910-1922.

(149) Yoon, J.; Czarnik, A. W. *J. Am. Chem. Soc.* **1992**, *114*, 5874-5875.

(150) Jin, S.; Wang, J. F.; Li, M. Y.; Wang, B. H. *Chem. Eur. J.* **2008**, *14*, 2795-2804.

(151) Edwards, N. Y.; Sager, T. W.; McDevitt, J. T.; Anslyn, E. V. *J. Am. Chem. Soc.* **2007**, *129*, 13575-13583.

(152) Dowlut, M.; Hall, D. G. *J. Am. Chem. Soc.* **2006**, *128*, 4226-4227.

(153) Springsteen, G.; Wang, B. H. *Tetrahedron* **2002**, *58*, 5291-5300.

(154) Yan, J.; Springsteen, G.; Deeter, S.; Wang, B. H. *Tetrahedron* **2004**, *60*, 11205-11209.

(155) Soundararajan, S.; Badawi, M.; Kohlrust, C. M.; Hageman, J. H. *Anal. Biochem.* **1989**, *178*, 125-134.

(156) Kawakami, J.; Kawase, Y.; Sugimoto, N. *Anal. Chim. Acta* **1998**, *365*, 95-100.

(157) Jeong, S.; Eom, T. Y.; Kim, S. J.; Lee, S. W.; Yu, J. *Biochem. Biophys. Res. Commun.* **2001**, *281*, 237-243.

(158) Yang, Q.; Goldstein, I. J.; Mei, H. Y.; Engelke, D. R. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 5462-5467.

(159) Li, M. Y.; Lin, N.; Huang, Z.; Du, L. P.; Altier, C.; Fang, H.; Wang, B. H. *J. Am. Chem. Soc.* **2008**, *130*, 12636-+.

(160) Feichtinger, K.; Zapf, C.; Sings, H. L.; Goodman, M. *J. Org. Chem.* **1998**, *63*, 3804-3805.

(161) Shiau, G. T.; Schinazi, R. F.; Chen, M. S.; Prusoff, W. H. *J. Med. Chem.* **1980**, *23*, 127-133.

(162) Potter, B. V. L.; Eckstein, F.; Uznanski, B. *Nucleic Acids Res.* **1983**, *11*, 7087-7103.

(163) Ludwig, J.; Eckstein, F. *J. Org. Chem.* **1989**, *54*, 631-635.

(164) Moffatt, J. G.; Khorana, H. G. *J. Am. Chem. Soc.* **1961**, *83*, 649-658.

(165) Masuda, T.; Nagasaki, T.; Tamagaki, S. *Supramol. Chem.* **2000**, *11*, 301-314.

(166) Hu, L. B.; Maurer, K.; Moeller, K. D. *Org. Lett.* **2009**, *11*, 1273-1276.

(167) Springsteen, G.; Wang, B. H. *Chem. Commun.* **2001**, 1608-1609.

(168) Li, W.; Burgess, K. *Tetrahedron Lett.* **1999**, *40*, 6527-6530.

(169) Guo, Z. F.; Sun, Y. R.; Zheng, S. L.; Guo, Z. H. *Biochemistry* **2009**, *48*, 1712-1722.

(170) Roth, A.; Breaker, R. R. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 6027-6031.

(171) Santoro, S. W.; Joyce, G. F.; Sakthivel, K.; Gramatikova, S.; Barbas, C. F. *J. Am. Chem. Soc.* **2000**, *122*, 2433-2439.

(172) Cubero, E.; Avino, A.; de la Torre, B. G.; Frieden, M.; Eritja, R.; Luque, F. J.; Gonzalez, C.; Orozco, M. *J. Am. Chem. Soc.* **2002**, *124*, 3133-3142.

(173) Avino, A.; Cubero, E.; Gonzalez, C.; Eritja, R.; Orozco, M. *J. Am. Chem. Soc.* **2003**, *125*, 16127-16138.

(174) Sonogashira, K.; Tohda, Y.; Hagihara, N. *Tetrahedron Lett.* **1975**, 4467-4470.

(175) Hubball, W.; Pyman, F. L. *J. Chem. Soc.* **1928**, 21-32.

(176) Turner, R. A.; Huebner, C. F.; Scholz, C. R. *J. Am. Chem. Soc.* **1949**, *71*, 2801-2803.

(177) Griffith, R. K.; Dipietro, R. A. *Synth. Commun.* **1986**, *16*, 1761-1770.

(178) Adger, B. M.; Surtees, J. *Synth. Commun.* **1987**, *17*, 223-227.

(179) Phillips, J. G.; Tedford, C. E.; Chaturvedi, N. C.; Ali, S. M.; Patent, U. S., Ed. 2000.

(180) Corey, E. J.; Fuchs, P. L. *Tetrahedron Lett.* **1972**, 3769-3772.

(181) Ikehara, M.; Kaneko, M. *Tetrahedron* **1970**, *26*, 4251-4259.

(182) Singh, D.; Kumar, V.; Ganesh, K. N. *Nucleic Acids Res.* **1990**, *18*, 3339-3345.

(183) Prakash, T. P.; Kumar, R. K.; Ganesh, K. N. *Tetrahedron* **1993**, *49*, 4035-4050.

(184) Tierney, M. T.; Grinstaff, M. W. *Org. Lett.* **2000**, *2*, 3413-3416.

(185) Fiala, K. A.; Brown, J. A.; Ling, H.; Kshetry, A. K.; Zhang, J.; Taylor, J. S.; Yang, W.; Suo, Z. C. *J. Mol. Biol.* **2007**, *365*, 590-602.

(186) Jordan, F.; Niv, H. *Biochim. Biophys. Acta* **1977**, *476*, 265-271.

(187) Lee, H. J.; Diavatis, T.; Tennakoon, S.; Yu, P. L.; Gao, X. L. *Biochim. Biophys. Acta, Gene Struct. Expression* **2007**, *1769*, 20-28.

(188) Hollenstein, M.; Hipolito, C.; Lam, C.; Dietrich, D.; Perrin, D. M. *Angew. Chem. Int. Ed.* **2008**, *47*, 4346-4350.

(189) Ting, R.; Lermer, L.; Perrin, D. M. *J. Am. Chem. Soc.* **2004**, *126*, 12720-12721.

(190) Zhang, B.; Wagner, G. K.; Weber, K.; Garnham, C.; Morgan, A. J.; Galione, A.; Guse, A. H.; Potter, B. V. L. *J. Med. Chem.* **2008**, *51*, 1623-1636.

(191) Wagner, P. J. *Acc. Chem. Res.* **1971**, *4*, 168-177.

(192) Brookman, J.; Chacon, J. N.; Sinclair, R. S. *Photochem. Photobiol. Sci.* **2002**, *1*, 327-332.

(193) Himes, R. A.; Park, G. Y.; Barry, A. N.; Blackburn, N. J.; Karlin, K. D. *J. Am. Chem. Soc.* **2007**, *129*, 5352-5253.

(194) Garin, J.; Melendez, E.; Merchan, F. L.; Merino, P.; Orduna, J.; Tejero, T. *J. Heterocycl. Chem.* **1990**, *27*, 321-326.

(195) Hannon, G. J. *Nature* **2002**, *418*, 244-251.

(196) Dykxhoorn, D. M.; Novina, C. D.; Sharp, P. A. *Nat. Rev. Mol. Cell Biol.* **2003**, *4*, 457-467.

(197) Alvarez-Salas, L. M. *Curr. Top. Med. Chem.* **2008**, *8*, 1379-1404.

(198) Bartel, D. P. *Cell* **2004**, *116*, 281-297.

(199) Opalinska, J. B.; Gewirtz, A. M. *Nat. Rev. Drug Discovery* **2002**, *1*, 503-514.

(200) Ohmichi, T.; Kuwahara, M.; Sasaki, N.; Hasegawa, M.; Nishikata, T.; Sawai, H.; Sugimoto, N. *Angew. Chem. Int. Ed.* **2005**, *44*, 6682-6685.

(201) Deglane, G.; Abes, S.; Michel, T.; Prevot, P.; Vives, E.; Debart, F.; Barvik, I.; Lebleu, B.; Vasseur, J. J. *Chembiochem* **2006**, *7*, 684-692.

(202) Oketani, M.; Asahina, Y.; Wu, C. H.; Wu, G. Y. *J. Hepatol.* **1999**, *31*, 628-634.

(203) Zepik, H. H.; Benner, S. A. *J. Org. Chem.* **1999**, *64*, 8080-8083.

(204) Ogura, H.; Kobayashi, T.; Shimizu, K.; Kawabe, K.; Takeda, K. *Tetrahedron Lett.* **1979**, 4745-4746.

(205) Manoharan, M.; Prakash, T. P.; Barber-Peoc'h, I.; Bhat, B.; Vasquez, G.; Ross, B. S.; Cook, P. D. *J. Org. Chem.* **1999**, *64*, 6468-6472.

(206) Wlassoff, W. A.; Albright, C. D.; Sivashinski, M. S.; Lvanova, A.; Appelbaum, J. G.; Salganik, R. I. *J. Pharm. Pharmacol.* **2007**, *59*, 1549-1553.

(207) Cook, A. F.; Vuocolo, E.; Brakel, C. L. *Nucleic Acids Res.* **1988**, *16*, 4077-4095.

(208) Prakash, T. P.; Pusehl, A.; Manoharan, M. *Nucleosides Nucleotides Nucleic Acids* **2007**, *26*, 149-159.

(209) Ting, R.; Thomas, J. M.; Lermer, L.; Perrin, D. M. *Nucleic Acids Res.* **2004**, *32*, 6660-6672.

(210) Wilson, R.; Turner, A. P. F. *Biosens. Bioelectron.* **1992**, *7*, 165-185.

(211) Oliver, N. S.; Toumazou, C.; Cass, A. E. G.; Johnston, D. G. *Diabetic Med.* **2009**, *26*, 197-210.

(212) Liu, J. W.; Cao, Z. H.; Lu, Y. *Chem. Rev.* **2009**, *109*, 1948-1998.

(213) Willner, I.; Shlyahovsky, B.; Zayats, M.; Willner, B. *Chem. Soc. Rev.* **2008**, *37*, 1153-1165.

(214) Franzen, S. *Curr. Opin. Mol. Ther.* **2010**, *12*, 223-232.

(215) Mandal, M.; Breaker, R. R. *Nat. Rev. Mol. Cell Biol.* **2004**, *5*, 451-463.

(216) Winkler, W. C.; Breaker, R. R. *Annu. Rev. Microbiol.* **2005**, *59*, 487-517.

(217) Vitreschak, A. G.; Rodionov, D. A.; Mironov, A. A.; Gelfand, M. S. *Trends in Genet.* **2004**, *20*, 44-50.

(218) Nudler, E.; Mironov, A. S. *Trends Biochem. Sci.* **2004**, *29*, 11-17.

(219) Lee, E. R.; Baker, J. L.; Weinberg, Z.; Sudarsan, N.; Breaker, R. R. *Science* **2010**, *329*, 845-848.

(220) Winkler, W. C.; Nahvi, A.; Roth, A.; Collins, J. A.; Breaker, R. R. *Nature* **2004**, *428*, 281-286.

(221) Klein, D. J.; Ferre-D'Amare, A. R. *Science* **2006**, *313*, 1752-1756.

(222) Cochrane, J. C.; Lipchock, S. V.; Strobel, S. A. *Chem. Biol.* **2007**, *14*, 97-105.

(223) Rajendran, M.; Ellington, A. D. *Anal. Bioanal.Chem.* **2008**, *390*, 1067-1075.

(224) Brown, A. K.; Li, J.; Pavot, C. M. B.; Lu, Y. *Biochemistry* **2003**, *42*, 7152-7161.

(225) Li, J.; Lu, Y. *J. Am. Chem. Soc.* **2000**, *122*, 10466-10467.

(226) Pan, T.; Uhlenbeck, O. C. *Nature* **1992**, *358*, 560-563.

(227) Pan, T.; Uhlenbeck, O. C. *Biochemistry* **1992**, *31*, 3887-3895.

(228) Liu, J. W.; Brown, A. K.; Meng, X. L.; Cropek, D. M.; Istok, J. D.; Watson, D. B.; Lu, Y. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 2056-2061.

(229) Cuenoud, B.; Szostak, J. W. *Nature* **1995**, *375*, 611-614.

(230) Liu, J.; Lu, Y. *Angew. Chem. Int. Ed.* **2007**, *46*, 7587-7590.

(231) Tang, J.; Breaker, R. R. *Chem. Biol.* **1997**, *4*, 453-459.

(232) Mei, S. H. J.; Liu, Z. J.; Brennan, J. D.; Li, Y. F. *J. Am. Chem. Soc.* **2003**, *125*, 412-420.

(233) Wang, D. Y.; Lai, B. H. Y.; Sen, D. *J. Mol. Biol.* **2002**, *318*, 33-43.

(234) Araki, M.; Okuno, Y.; Hara, Y.; Sugiura, Y. *Nucleic Acids Res.* **1998**, *26*, 3379-3384.

(235) Soukup, G. A.; Breaker, R. R. *Structure* **1999**, *7*, 783-791.

(236) Breaker, R. R. *Curr. Opin. Biotechnol.* **2002**, *13*, 31-39.

(237) Srinivasan, J.; Cload, S. T.; Hamaguchi, N.; Kurz, J.; Keene, S.; Kurz, M.; Boomer, R. M.; Blanchard, J.; Epstein, D.; Wilson, C.; Diener, J. L. *Chem. Biol.* **2004**, *11*, 499-508.

(238) Zivarts, M.; Liu, Y.; Breaker, R. R. *Nucleic Acids Res.* **2005**, *33*, 622-631.

(239) Ferguson, A.; Boomer, R. M.; Kurz, M.; Keene, S. C.; Diener, J. L.; Keefe, A. D.; Wilson, C.; Cload, S. T. *Nucleic Acids Res.* **2004**, *32*, 1756-1766.

(240) Robertson, M. P.; Ellington, A. D. *Nat. Biotechnol.* **1999**, *17*, 62-66.

(241) Robertson, M. P.; Ellington, A. D. *Nucleic Acids Res.* **2000**, *28*, 1751-1759.

(242) Larkin, M. A.; Blackshields, G.; Brown, N. P.; Chenna, R.; McGettigan, P. A.; McWilliam, H.; Valentin, F.; Wallace, I. M.; Wilm, A.; Lopez, R.; Thompson, J. D.; Gibson, T. J.; Higgins, D. G. *Bioinformatics* **2007**, *23*, 2947-2948.

(243) Mathews, D. H.; Sabina, J.; Zuker, M.; Turner, D. H. *J. Mol. Biol.* **1999**, *288*, 911-940.

(244) Zuker, M. *Nucleic Acids Res.* **2003**, *31*, 3406-3415.

(245) Aldridge, W. N. *Biochem. J.* **1950**, *46*, 451-460.

(246) Albert, A.; Serjeant, E. P. *The Determination of Ionization Constants*; Chapman and Hall: New York, 1984.

(247) Gonzalez, M.; Argarana, C. E.; Fidelio, G. D. *Biomol. Eng.* **1999**, *16*, 67-72.

(248) Polz, M. F.; Cavanaugh, C. M. *Appl. Environ. Microbiol.* **1998**, *64*, 3724-3730.

(249) Kanagawa, T. *J. Biosci. Bioeng.* **2003**, *96*, 317-323.

(250) Hipolito, C. J.; Hollenstein, M.; Lam, C. H.; Perrin, D. M. *Org. Biomol. Chem.* **2010**, (accepted).

(251) Tae, E. J. L.; Wu, Y. Q.; Xia, G.; Schultz, P. G.; Romesberg, F. E. *J. Am. Chem. Soc.* **2001**, *123*, 7439-7440.

(252) Loakes, D.; Holliger, P. *Chem. Commun.* **2009**, 4619-4631.

(253) Jestin, J. L.; Kristensen, P.; Winter, G. *Angew. Chem. Int. Ed.* **1999**, *38*, 1124-1127.

(254) Xia, G.; Chen, L. J.; Sera, T.; Fa, M.; Schultz, P. G.; Romesberg, F. E. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 6597-6602.

(255) Ghadessy, F. J.; Ong, J. L.; Holliger, P. *Proc. Natl. Acad. Sci. U.S.A.* **2001**,

98, 4552-4557.

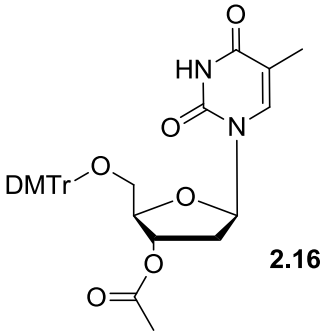
(256) Ghadessy, F. J.; Ramsay, N.; Boudsocq, F.; Loakes, D.; Brown, A.; Iwai, S.; Vaisman, A.; Woodgate, R.; Holliger, P. *Nat. Biotechnol.* **2004**, *22*, 755-759.

(257) Ong, J. L.; Loakes, D.; Jaroslawski, S.; Too, K.; Holliger, P. *J. Mol. Biol.* **2006**, *361*, 537-550.

Appendix

Chapter 2

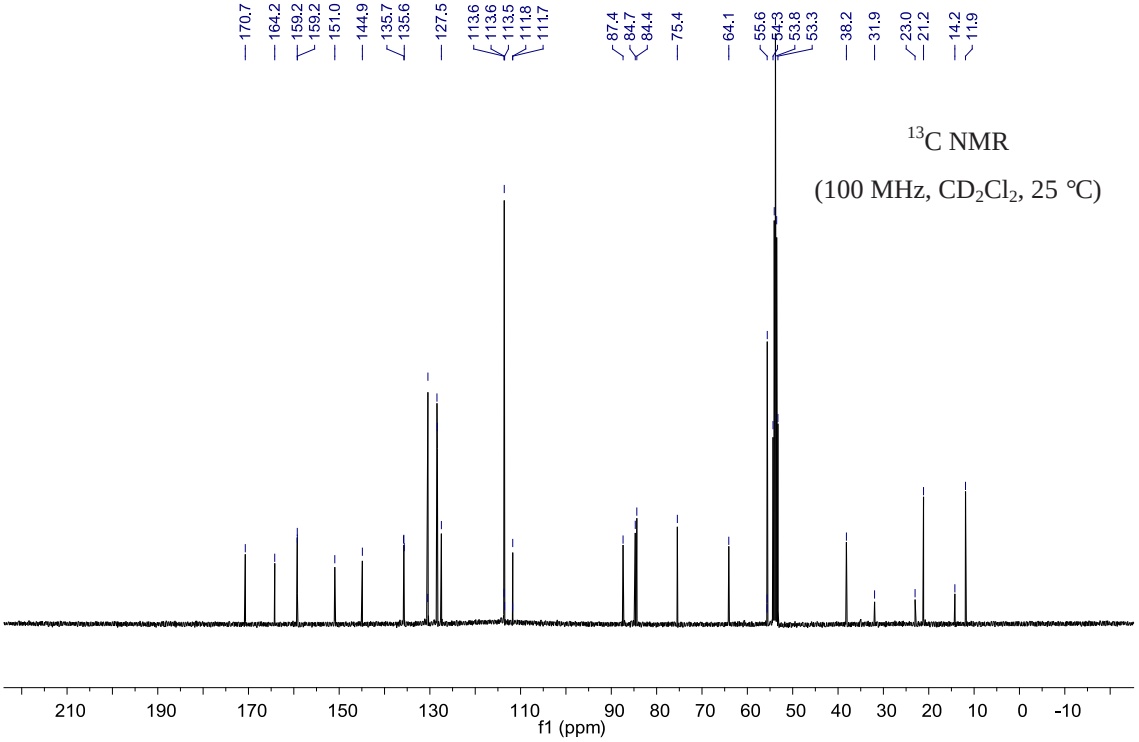
NMR spectra



2.16

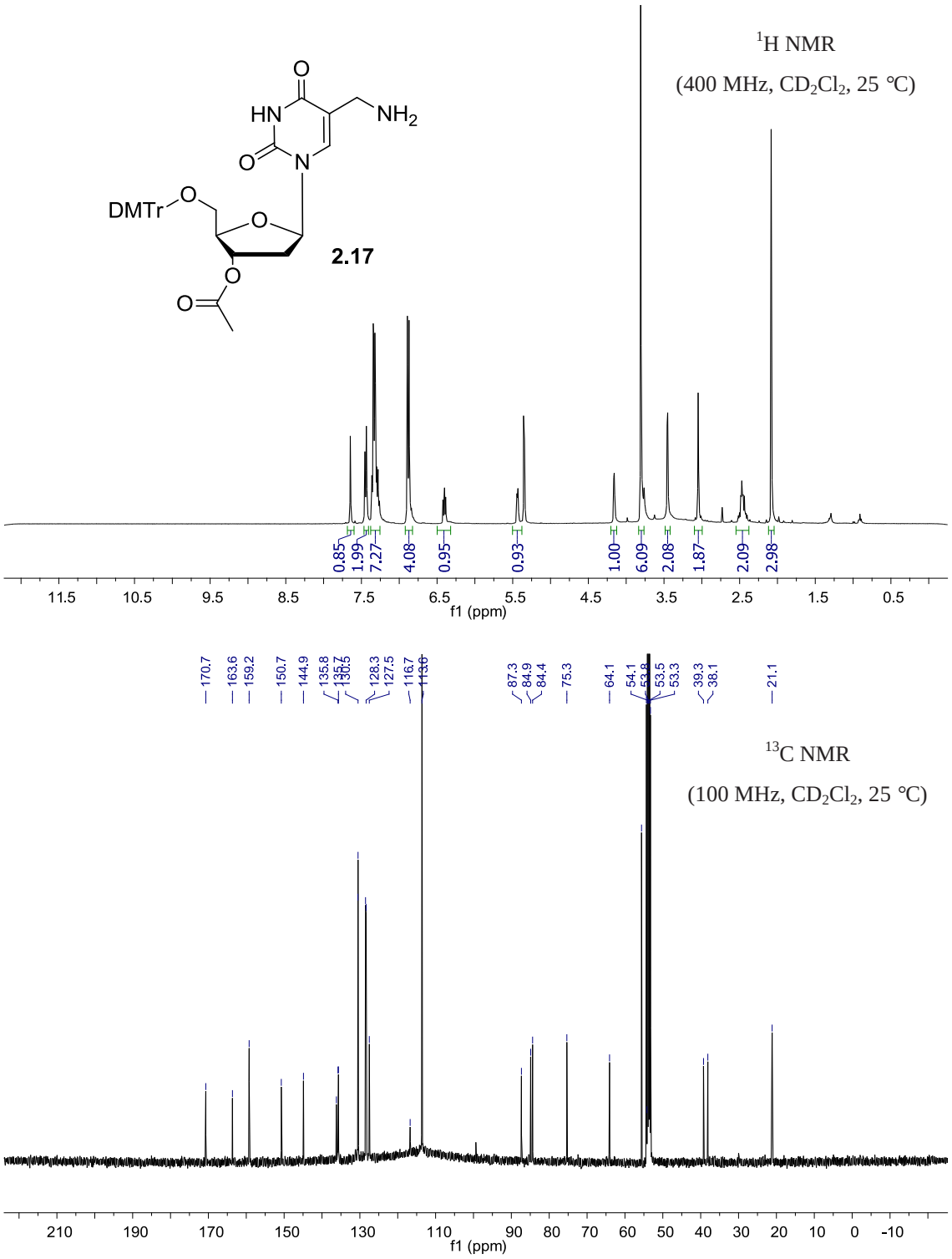
^{1}H NMR

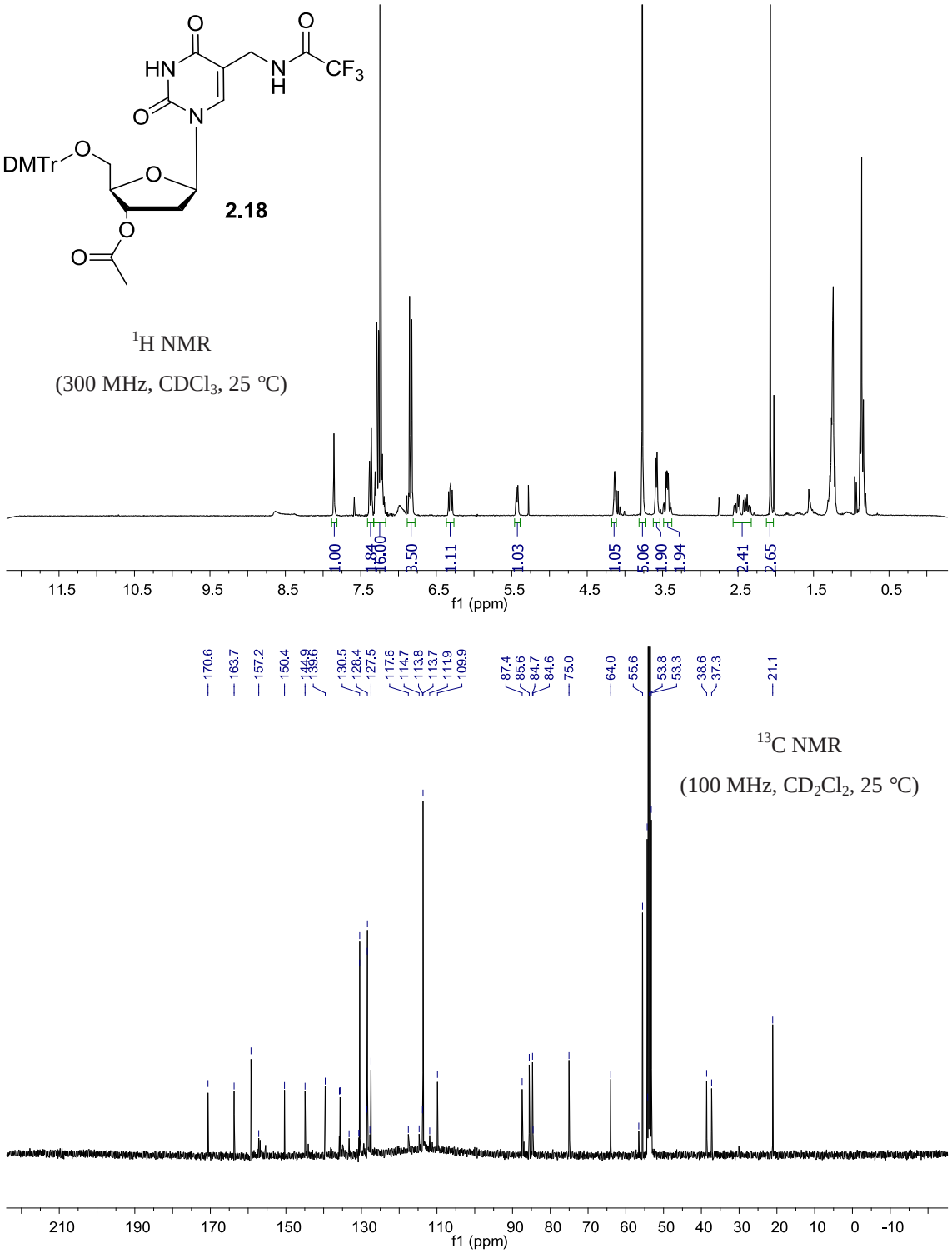
(400 MHz, CD_2Cl_2, 25 °C)

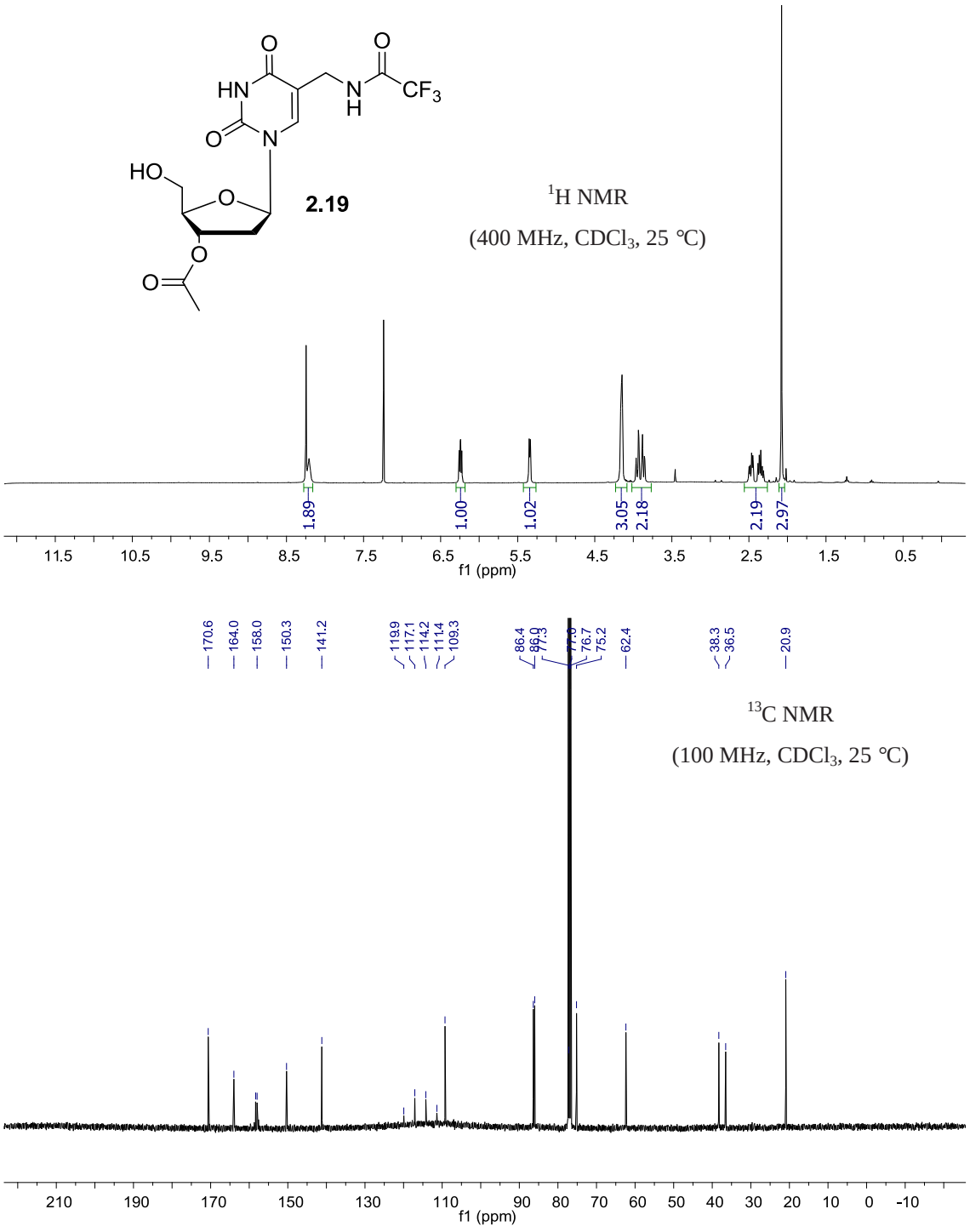


^{13}C NMR

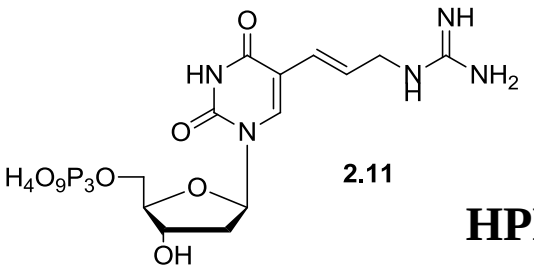
(100 MHz, CD_2Cl_2, 25 °C)


2.17
^{1}H NMR
(400 MHz, CD_2Cl_2, 25 °C)
0.85
1.99
7.27
4.08
0.95
0.93
1.00
6.09
2.08
1.87
2.09
2.98
f1 (ppm)
^{13}C NMR
(100 MHz, CD_2Cl_2, 25 °C)
170.7
163.6
159.2
150.7
144.9
135.8
128.3
127.5
116.7
87.3
84.9
84.4
75.3
64.1
54.1
53.5
53.3
39.3
38.1
21.1
f1 (ppm)

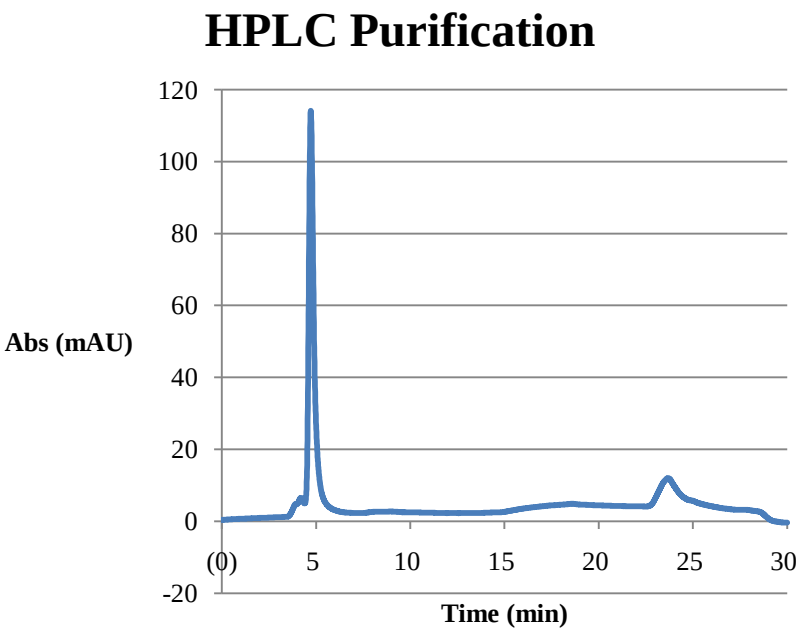

DMTr
2.18
^{1}H NMR
(300 MHz, $CDCl_3$, 25 ℃)
1.00
1.84
16.00
3.50
1.11
1.03
1.05
5.06
1.90
1.94
2.41
2.65
f1 (ppm)
^{13}C NMR
(100 MHz, CD_2Cl_2, 25 ℃)
170.6
163.7
157.2
150.4
144.9
139.6
130.5
128.4
127.5
117.6
114.7
113.8
113.7
111.9
109.9
87.4
85.6
84.7
84.6
75.0
64.0
55.6
53.8
53.3
38.6
37.3
21.1
f1 (ppm)

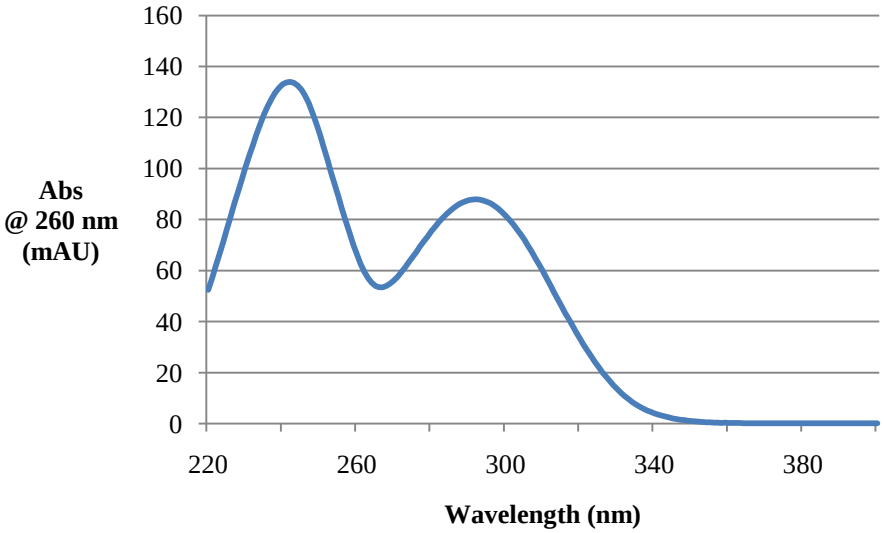

2.19
^{1}H NMR
(400 MHz, $CDCl_3$, 25 ℃)
1.89
1.00
1.02
3.05
2.18
2.19
2.97
11.5
10.5
9.5
8.5
7.5
6.5
5.5
4.5
3.5
2.5
1.5
0.5
f1 (ppm)
170.6
164.0
158.0
150.3
141.2
119.9
117.1
114.2
111.4
109.3
86.4
86.0
77.3
76.7
75.2
62.4
38.3
36.5
20.9
^{13}C NMR
(100 MHz, $CDCl_3$, 25 ℃)
210
190
170
150
130
110
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)

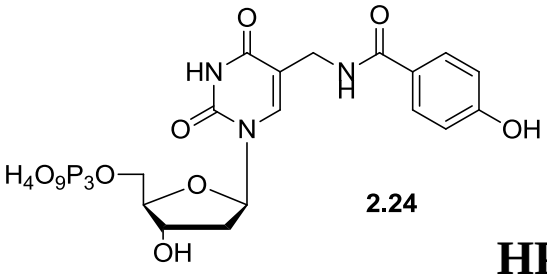
HPLC purification of modified 2′-deoxyuridine triphosphates



2.11

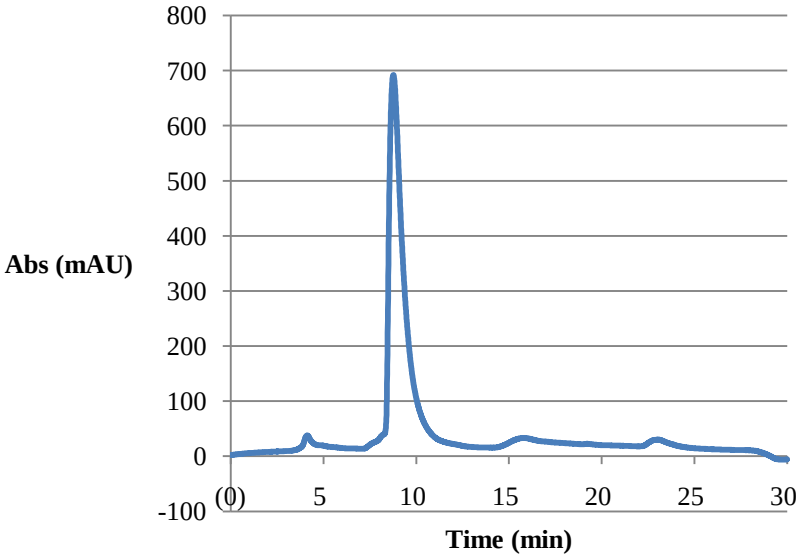


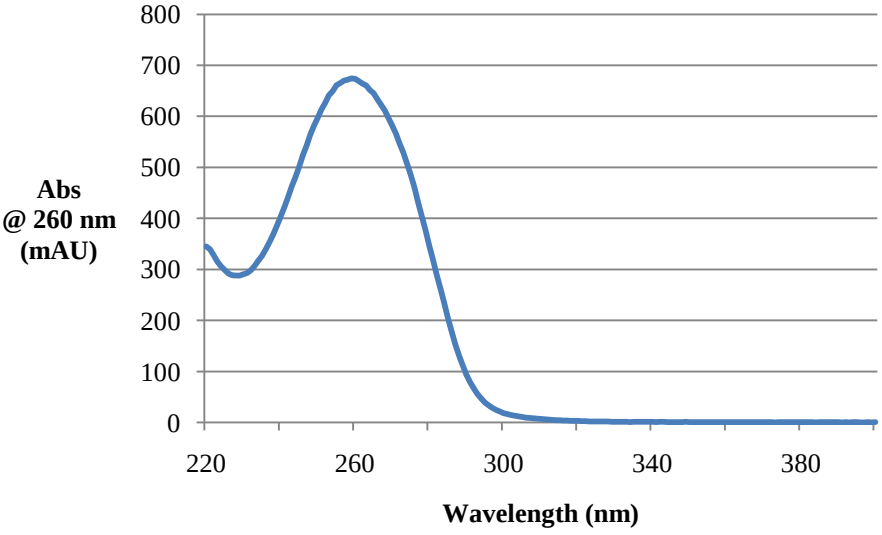


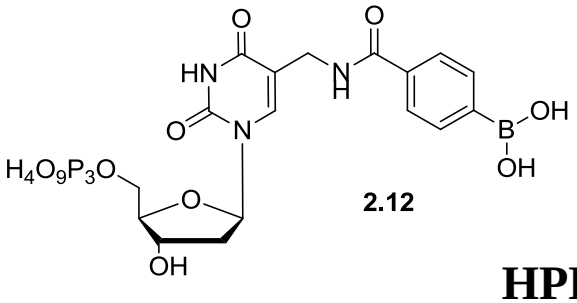


2.24

HPLC Purification

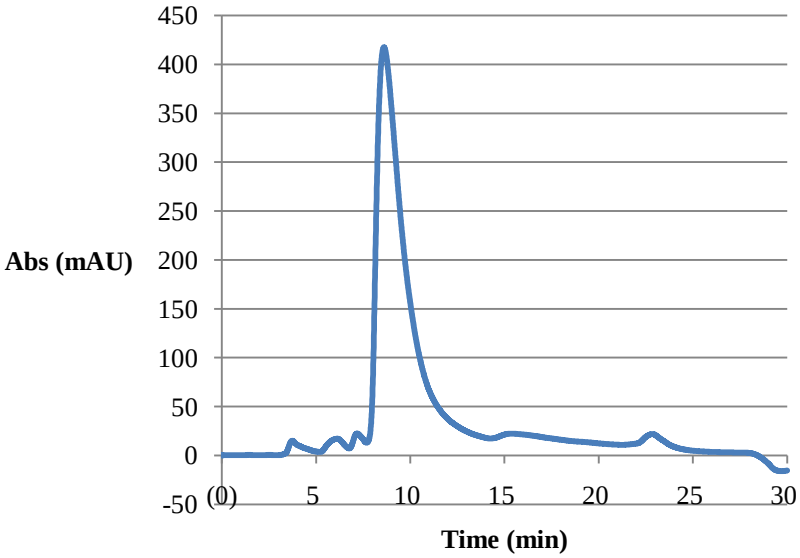


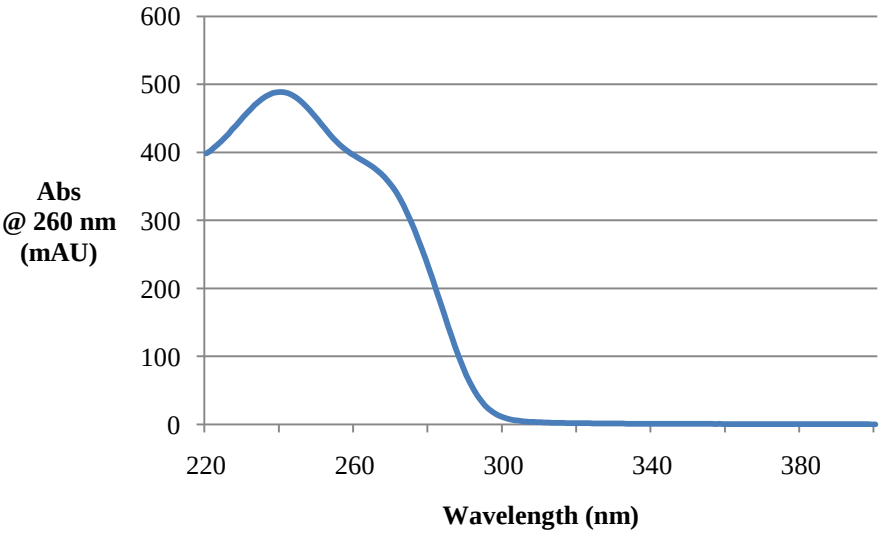




2.12

HPLC Purification

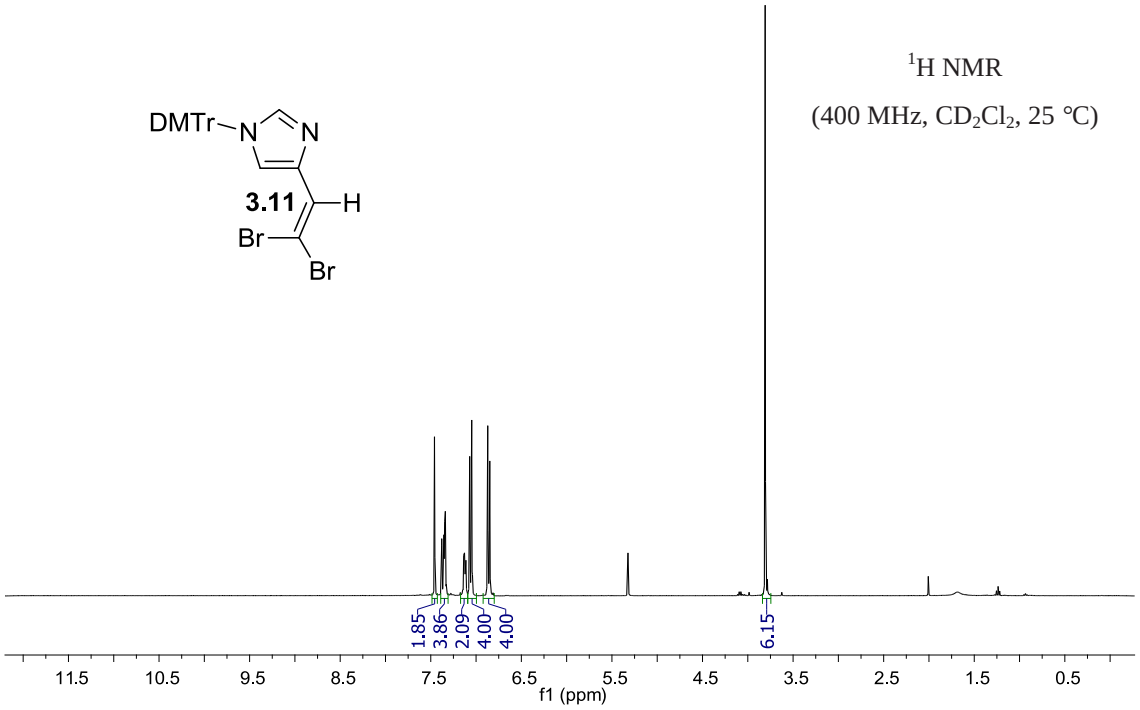


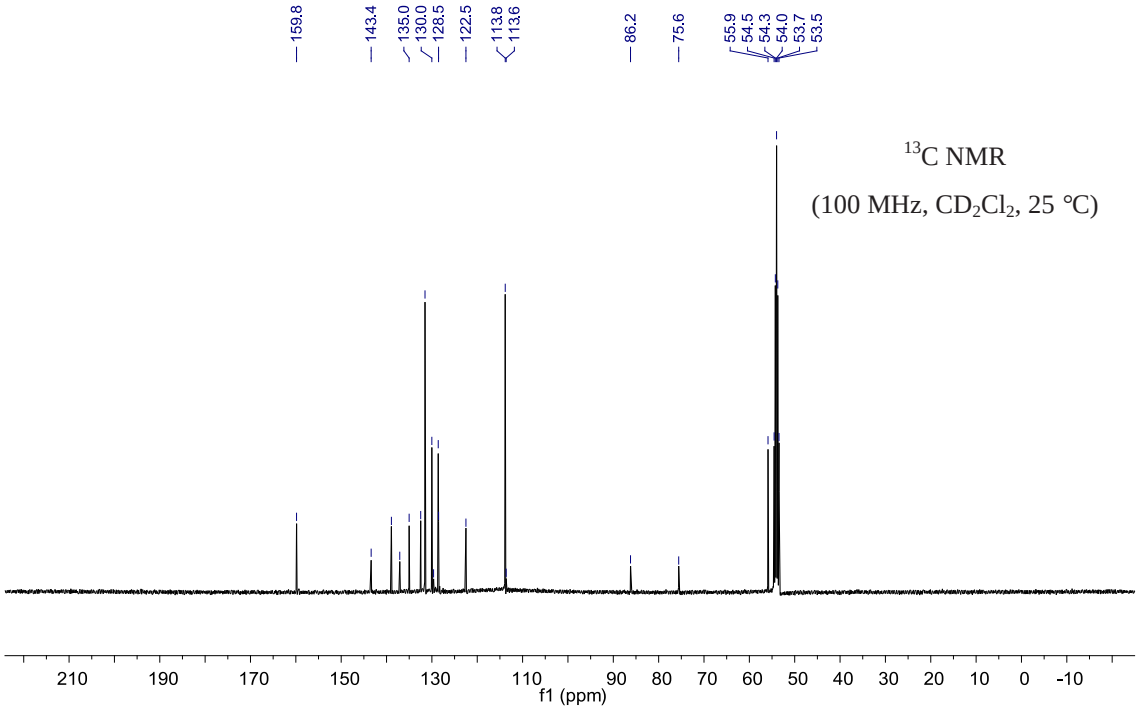


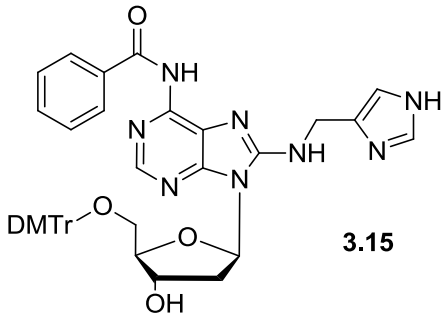
Chapter 3

NMR spectra

The NMR spectra of the modified nucleosides were very often done in the presence of two deuterated solvents due to problems of solubility. Using either deuterated chloroform or methylene chloride as solvents, it was found that the addition of one or two drops of deuterated methanol increased the solubility of the nucleosides significantly. Because of this, several spectra have peaks due to two deuterated solvents rather than just one. This method was particularly useful for ^{13}C NMR.


DMTr
N
N
3.11
H
Br
Br
1H NMR
(400 MHz, CD2Cl2, 25 °C)
1.85
3.86
2.09
4.00
4.00
6.15
11.5
10.5
9.5
8.5
7.5
6.5
5.5
4.5
3.5
2.5
1.5
0.5
f1 (ppm)


159.8
143.4
135.0
130.0
128.5
122.5
113.8
113.6
86.2
75.6
55.9
54.5
54.3
54.0
53.7
53.5
13C NMR
(100 MHz, CD2Cl2, 25 °C)
210
190
170
150
130
110
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)



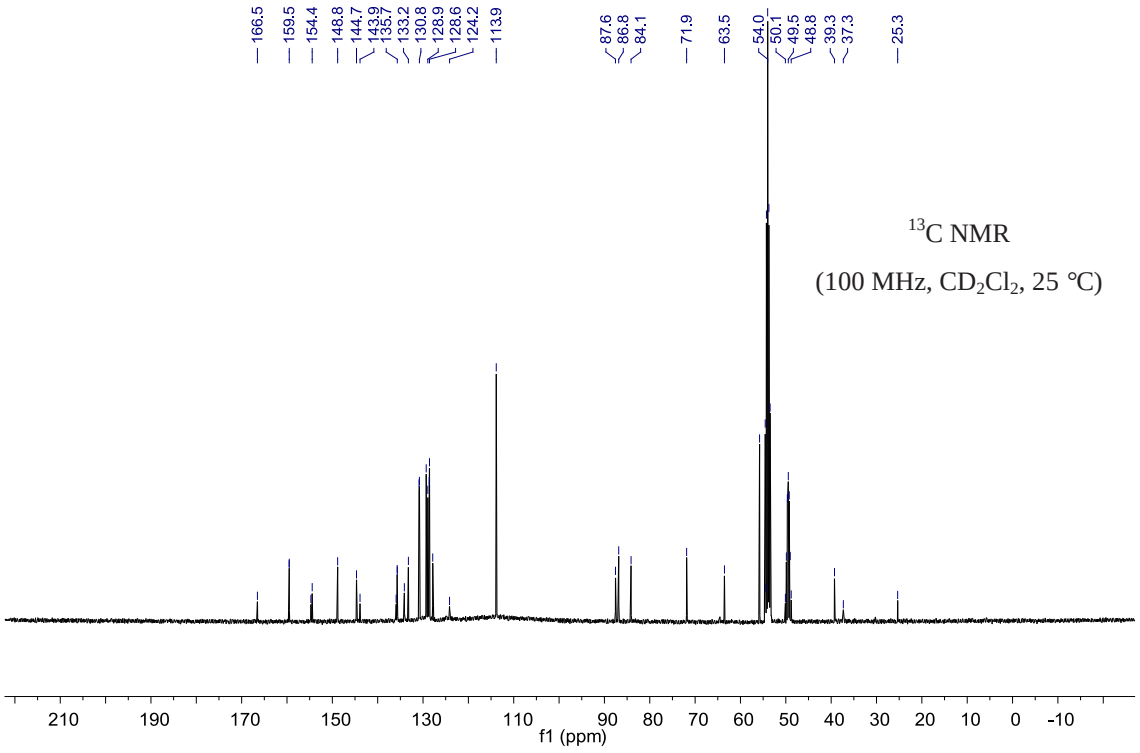
3.15

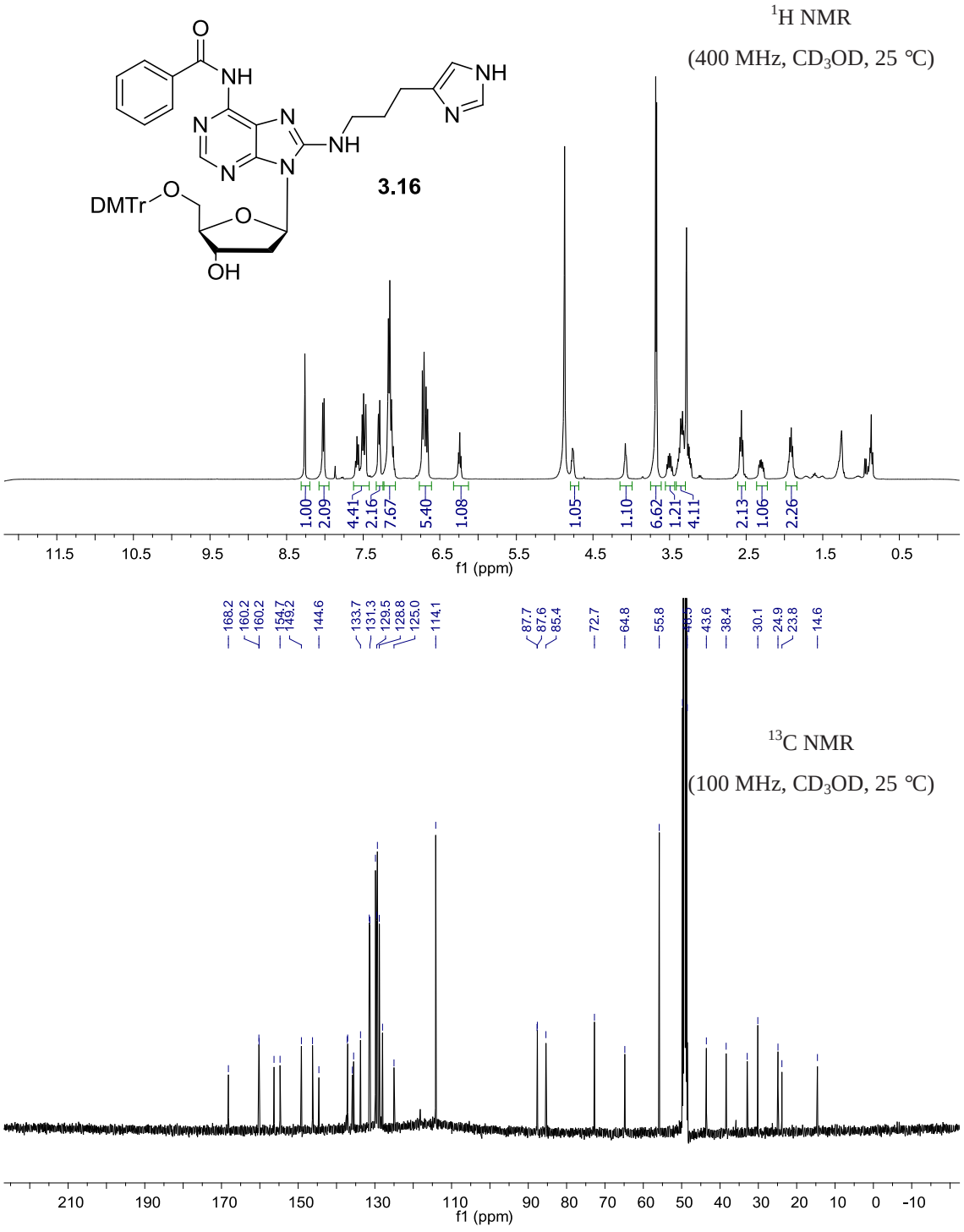
^{1}H NMR

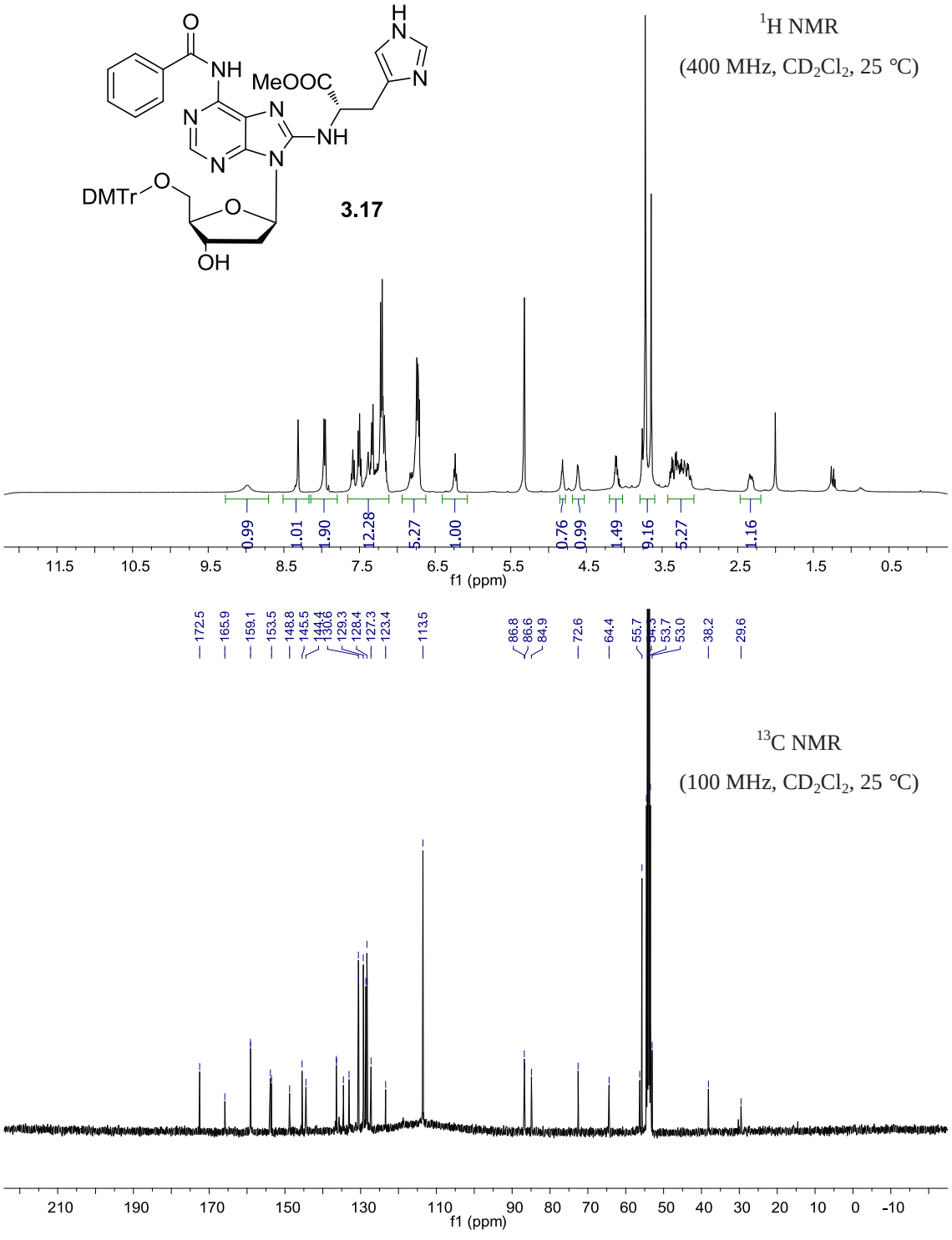
(400 MHz, CD_2Cl_2, 25 °C)

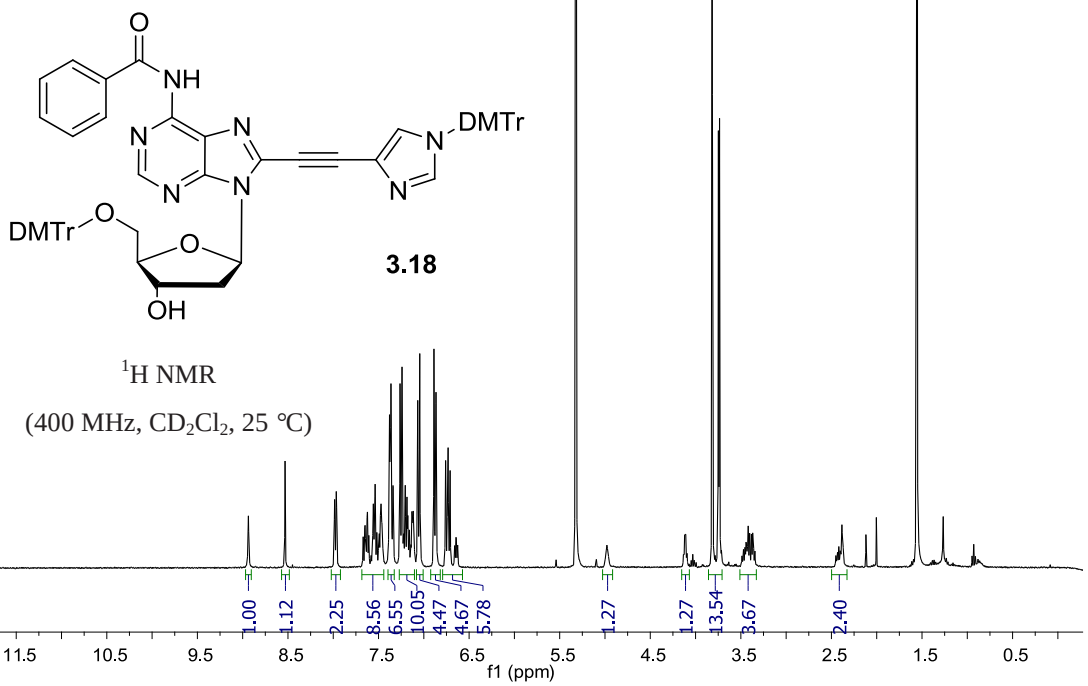
^{13}C NMR

(100 MHz, CD_2Cl_2, 25 °C)



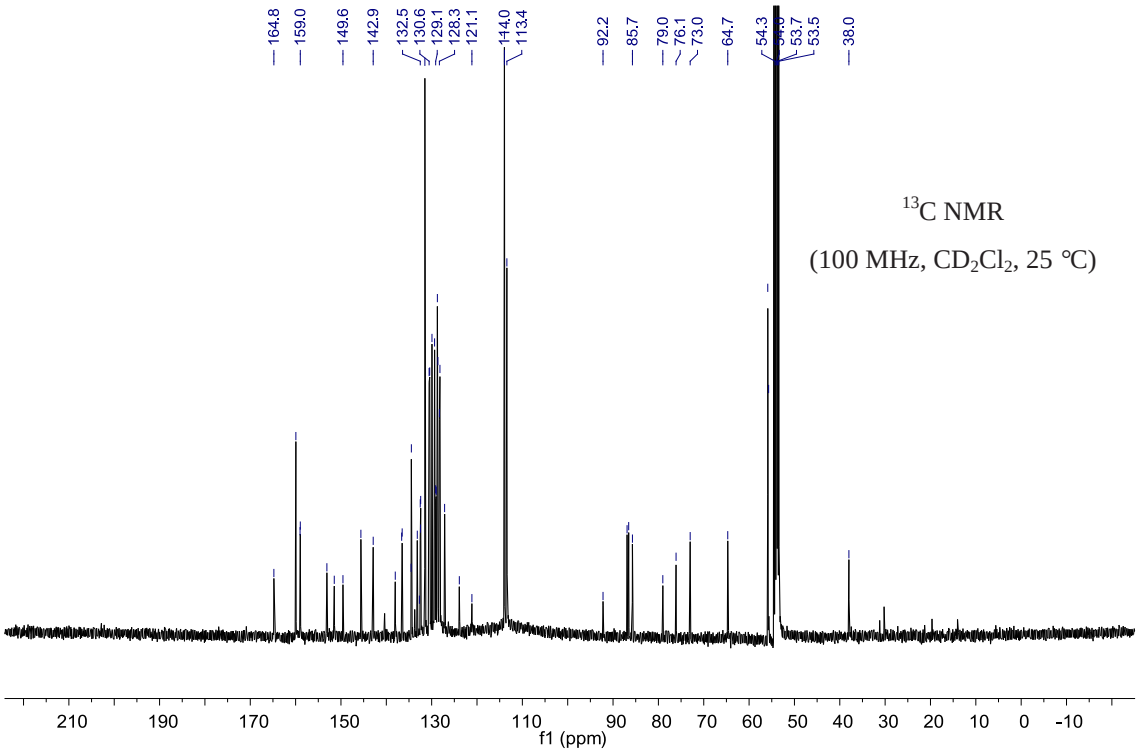

3.16
DMTr
OH
1H NMR
(400 MHz, CD3OD, 25 ℃)
1.00
2.09
4.41
2.16
7.67
5.40
1.08
1.05
1.10
6.62
1.21
4.11
2.13
1.06
2.26
f1 (ppm)
168.2
160.2
160.2
154.7
149.2
144.6
133.7
131.3
129.5
128.8
125.0
114.1
87.7
87.6
85.4
72.7
64.8
55.8
43.6
38.4
30.1
24.9
23.8
14.6
13C NMR
(100 MHz, CD3OD, 25 ℃)
f1 (ppm)


3.17
^{1}H NMR
(400 MHz, CD_2Cl_2, 25 °C)
0.99
1.01
1.90
12.28
5.27
1.00
0.76
0.99
1.49
9.16
5.27
1.16
f1 (ppm)
^{13}C NMR
(100 MHz, CD_2Cl_2, 25 °C)
f1 (ppm)



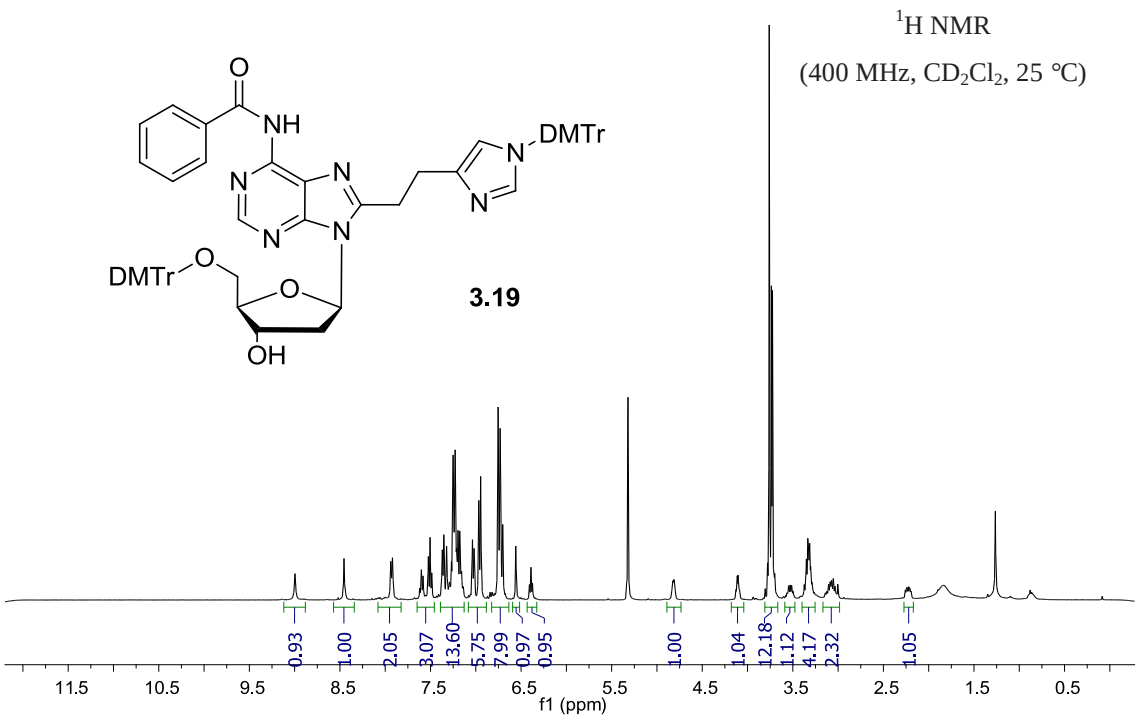
^{1}H NMR

(400 MHz, CD_2Cl_2, 25 ℃)



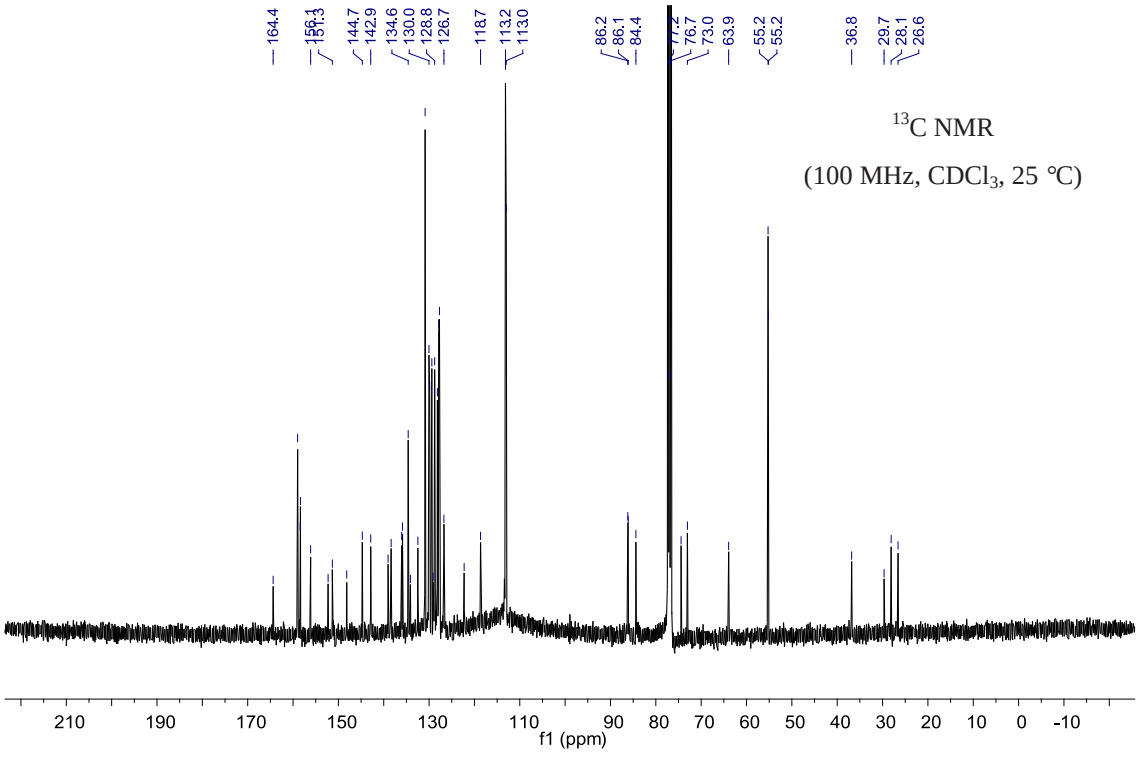
^{13}C NMR

(100 MHz, CD_2Cl_2, 25 ℃)



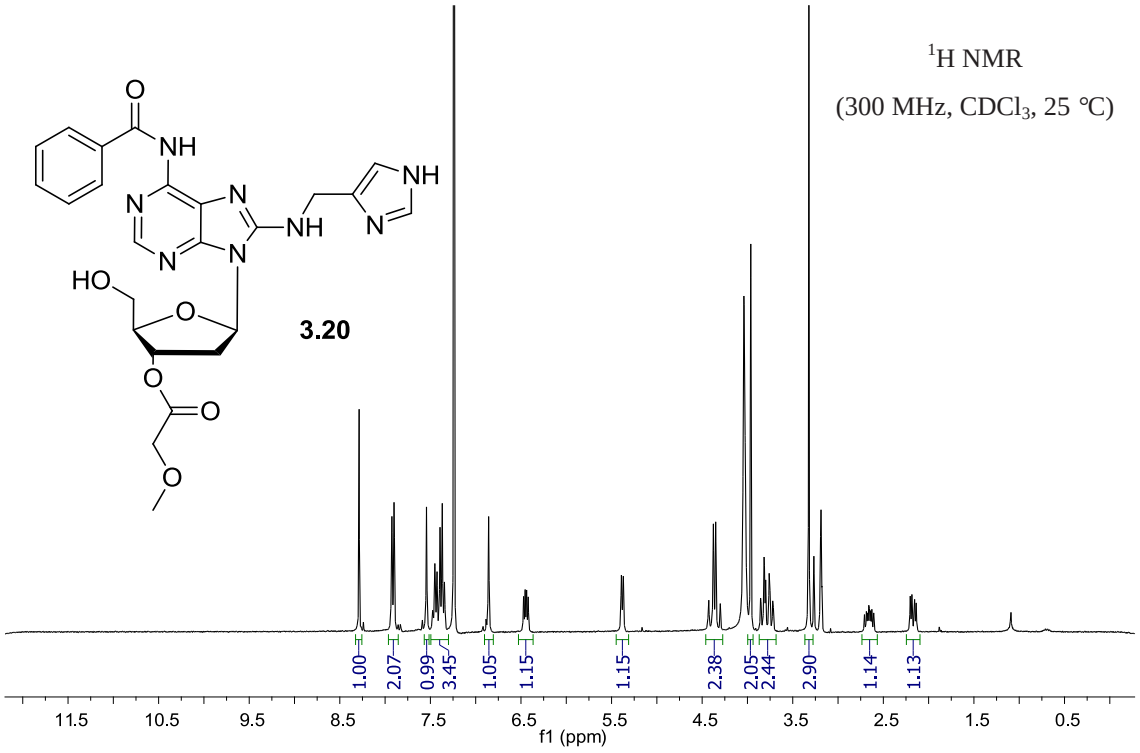
^{1}H NMR

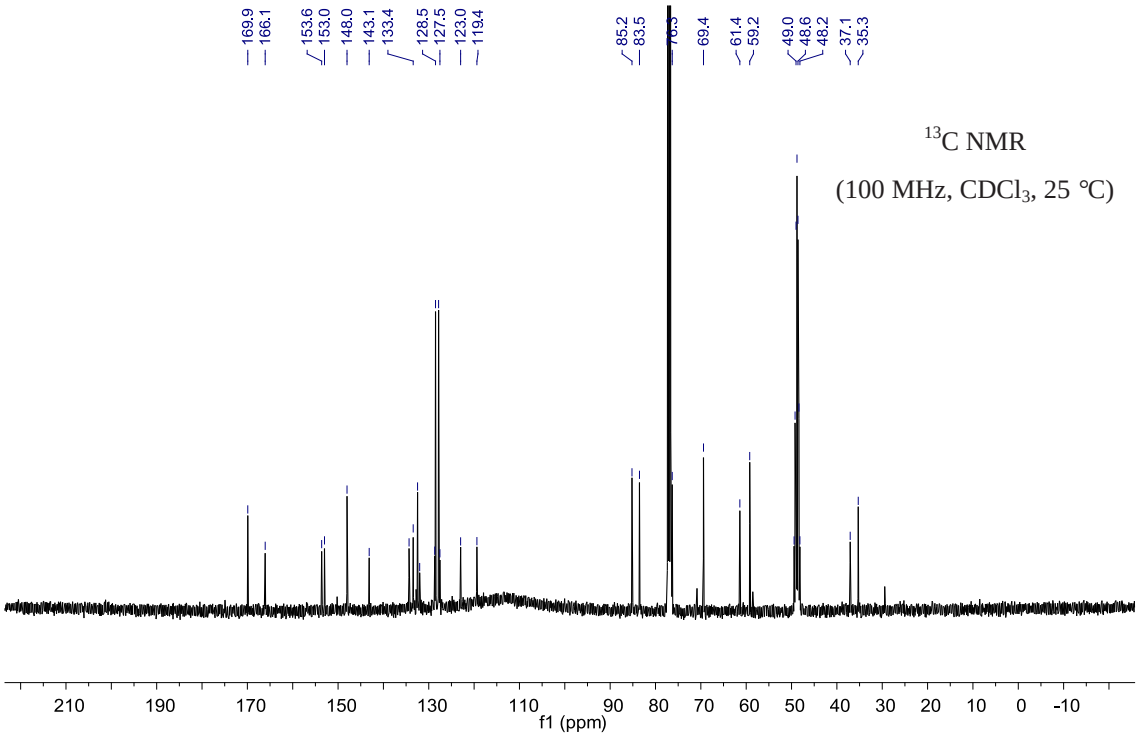
(400 MHz, CD_2Cl_2, 25 ℃)

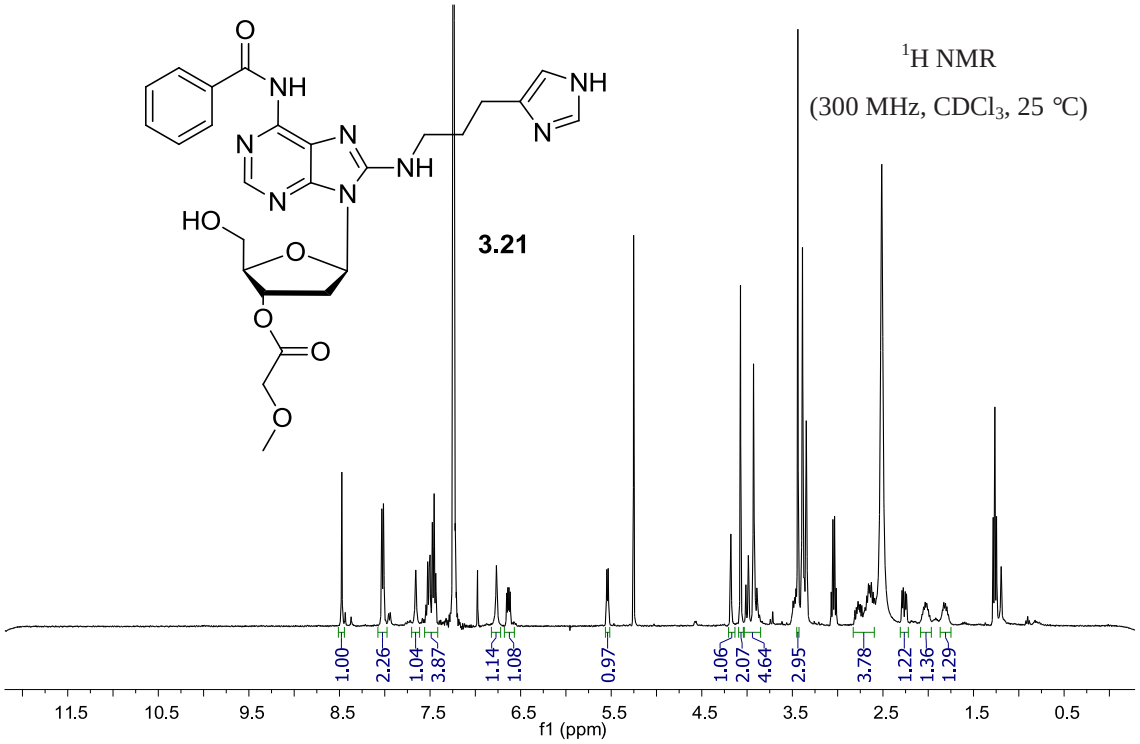


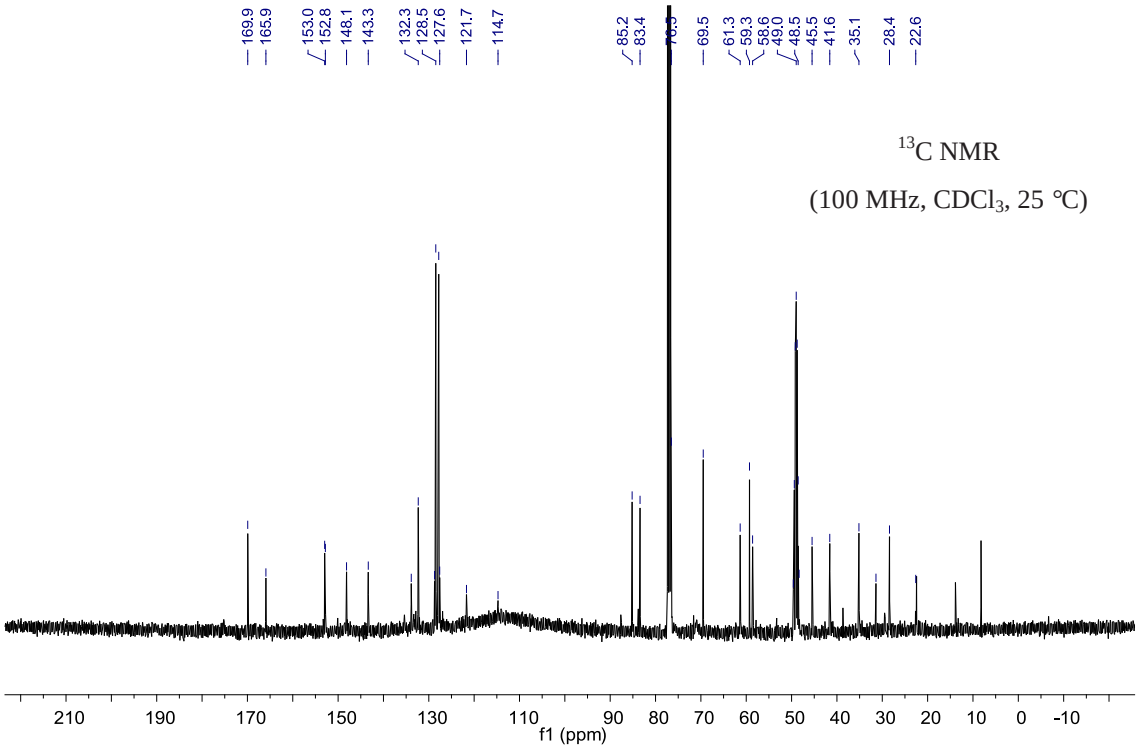
^{13}C NMR

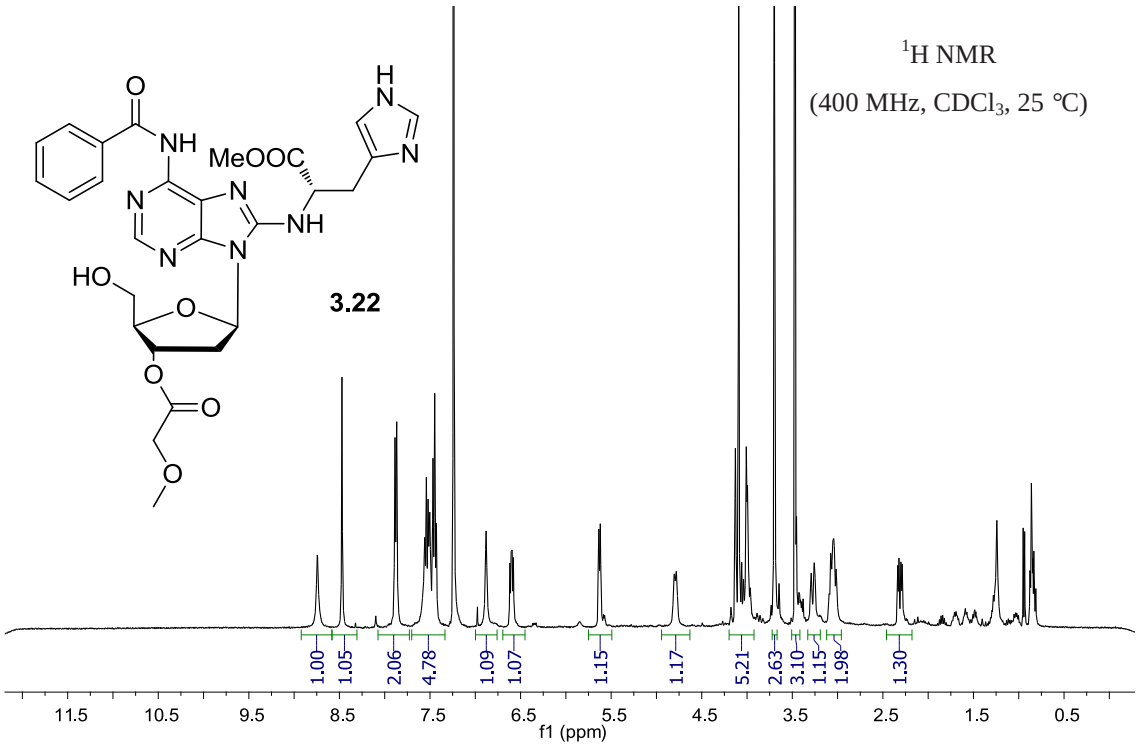
(100 MHz, $CDCl_3$, 25 ℃)


3.20
^{1}H NMR
(300 MHz, $CDCl_3$, 25 ℃)
1.00
2.07
0.99
3.45
1.05
1.15
1.15
2.38
2.05
2.44
2.90
1.14
1.13
11.5
10.5
9.5
8.5
7.5
6.5
5.5
4.5
3.5
2.5
1.5
0.5
f1 (ppm)


169.9
166.1
153.6
153.0
148.0
143.1
133.4
128.5
127.5
123.0
119.4
85.2
83.5
76.3
69.4
61.4
59.2
49.0
48.6
48.2
37.1
35.3
^{13}C NMR
(100 MHz, $CDCl_3$, 25 ℃)
210
190
170
150
130
110
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)

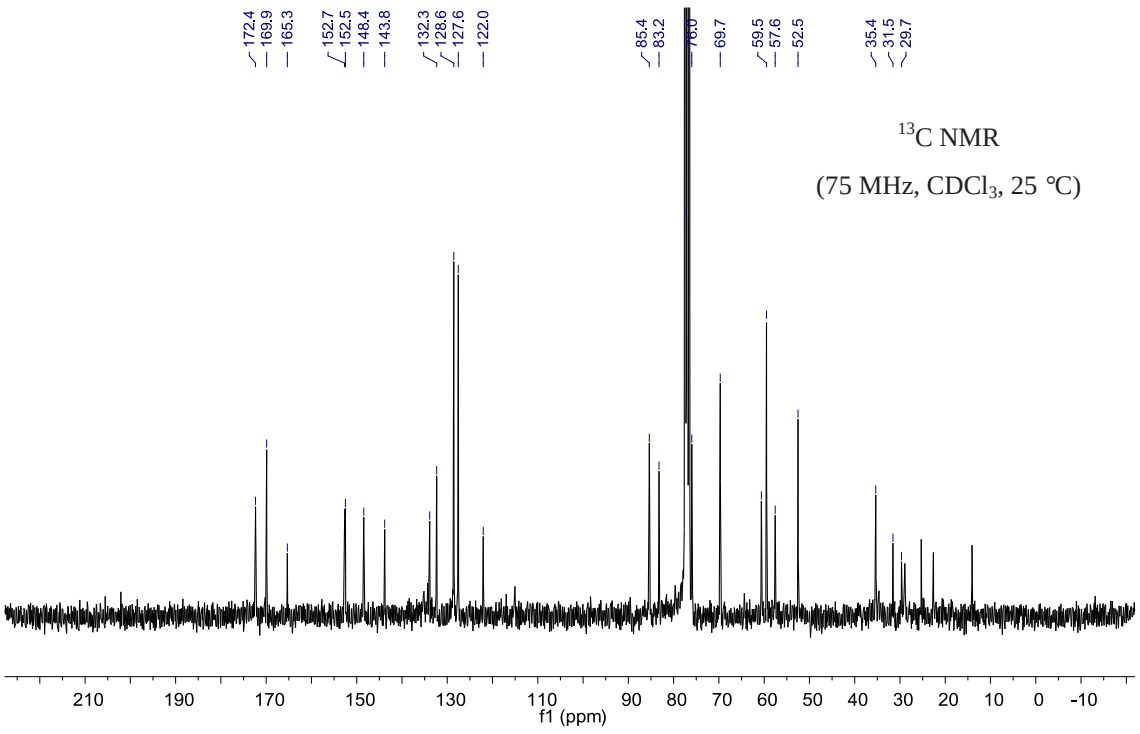

3.21
1H NMR
(300 MHz, CDCl3, 25 °C)
1.00
2.26
1.04
3.87
1.14
1.08
0.97
1.06
2.07
4.64
2.95
3.78
1.22
1.36
1.29
11.5
10.5
9.5
8.5
7.5
6.5
5.5
4.5
3.5
2.5
1.5
0.5
f1 (ppm)


169.9
165.9
153.0
152.8
148.1
143.3
132.3
128.5
127.6
121.7
114.7
85.2
83.4
76.8
69.5
61.3
59.3
58.6
49.0
48.5
45.5
41.6
35.1
28.4
22.6
13C NMR
(100 MHz, CDCl3, 25 °C)
210
190
170
150
130
110
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)



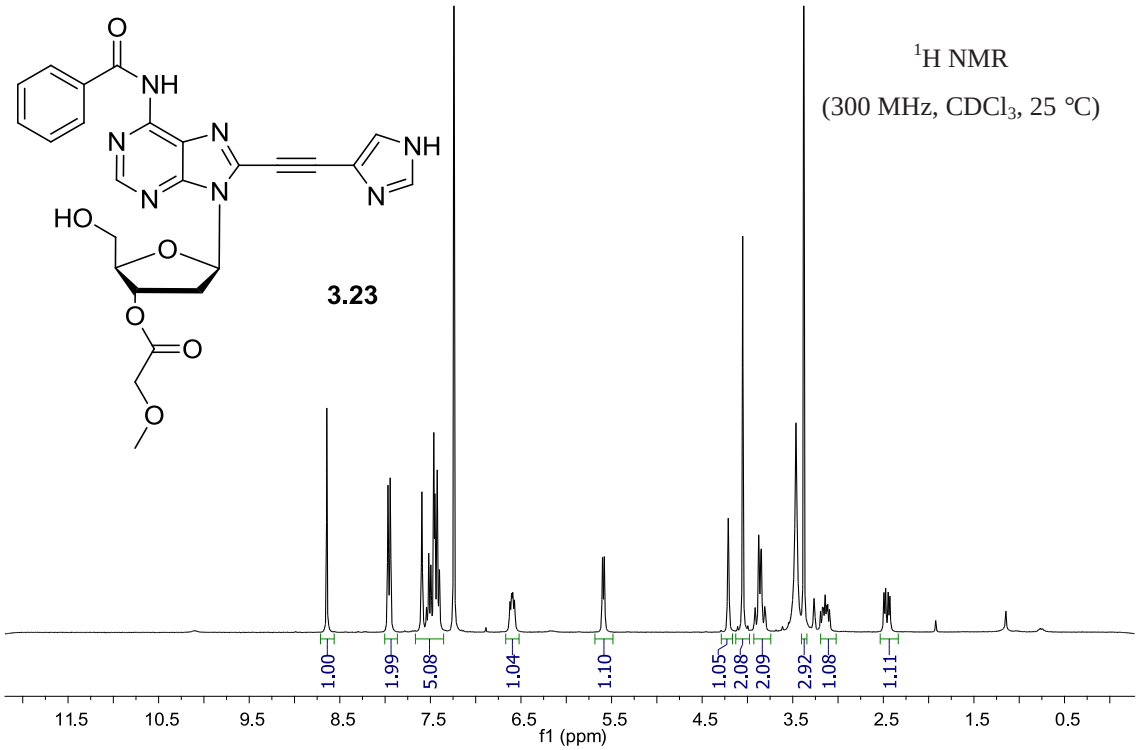
^{1}H NMR

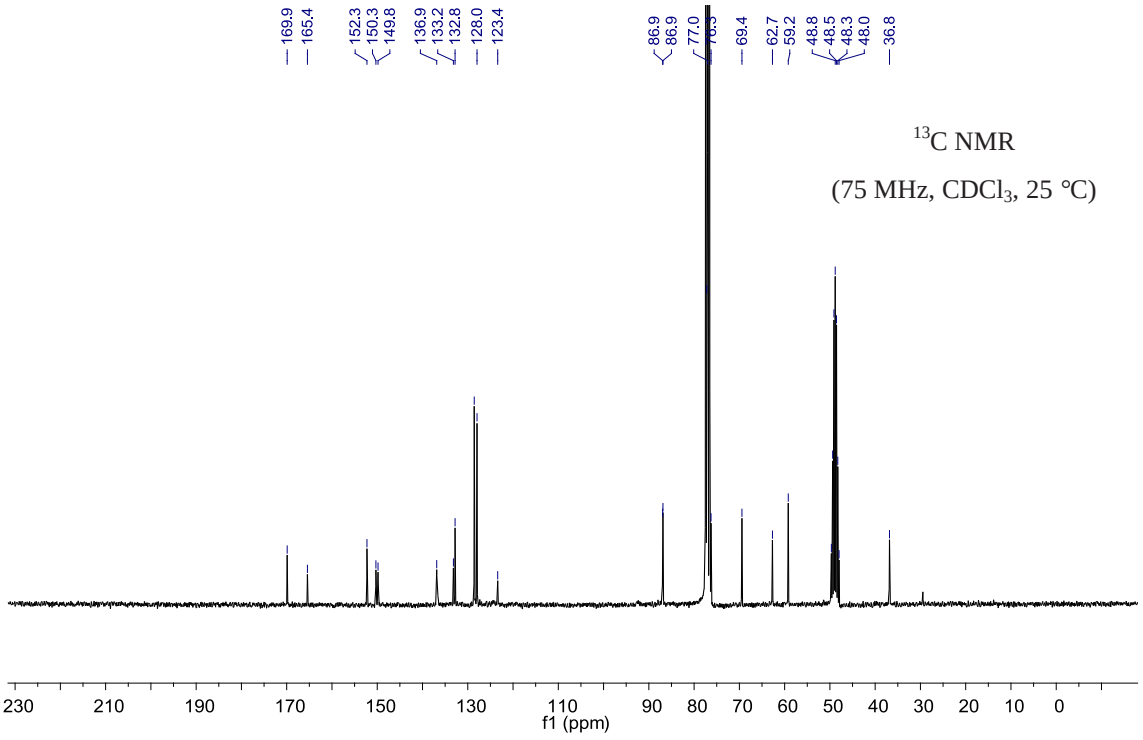
(400 MHz, $CDCl_3$, 25 ℃)

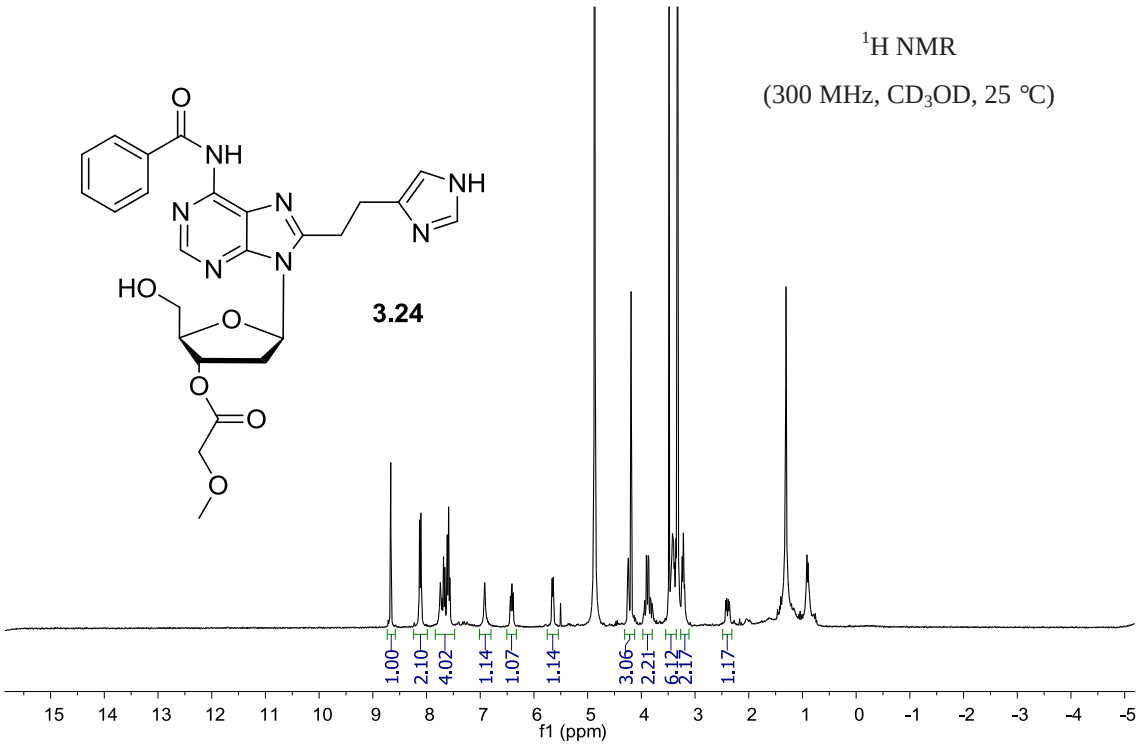


^{13}C NMR

(75 MHz, $CDCl_3$, 25 ℃)

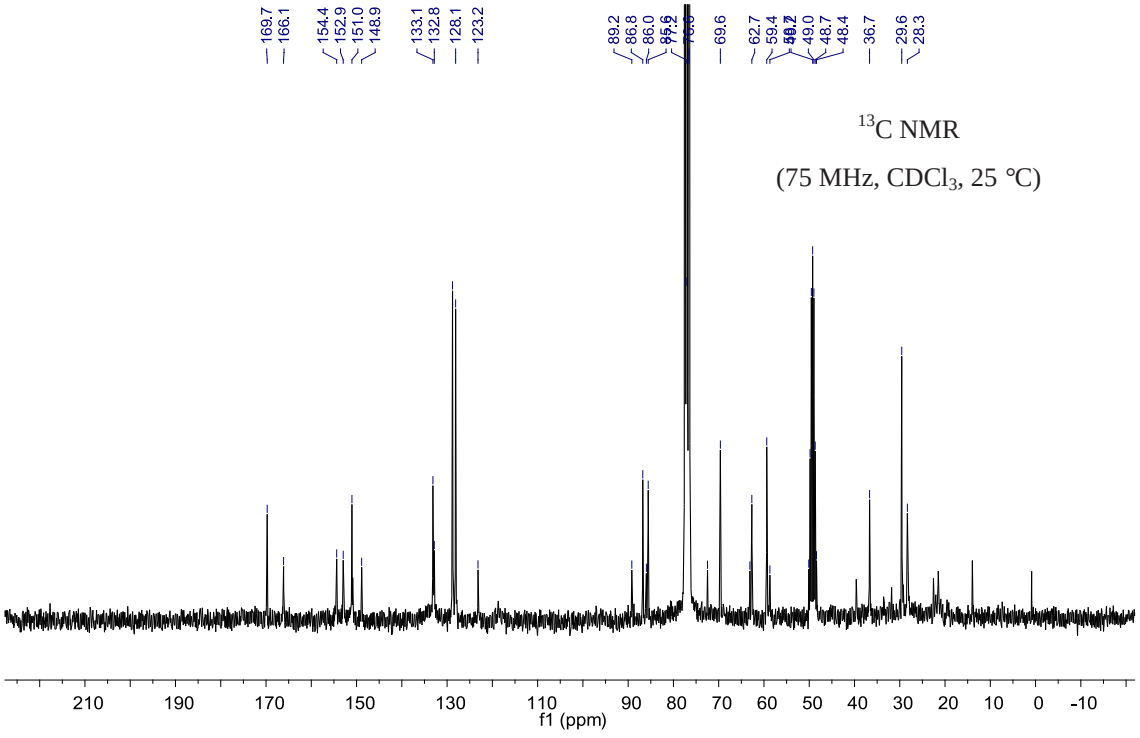

3.23
^{1}H NMR
(300 MHz, $CDCl_3$, 25 °C)
1.00
1.99
5.08
1.04
1.10
1.05
2.08
2.09
2.92
1.08
1.11
f1 (ppm)


169.9
165.4
152.3
150.3
149.8
136.9
133.2
132.8
128.0
123.4
86.9
86.9
77.0
76.3
69.4
62.7
59.2
48.8
48.5
48.3
48.0
36.8
^{13}C NMR
(75 MHz, $CDCl_3$, 25 °C)
f1 (ppm)



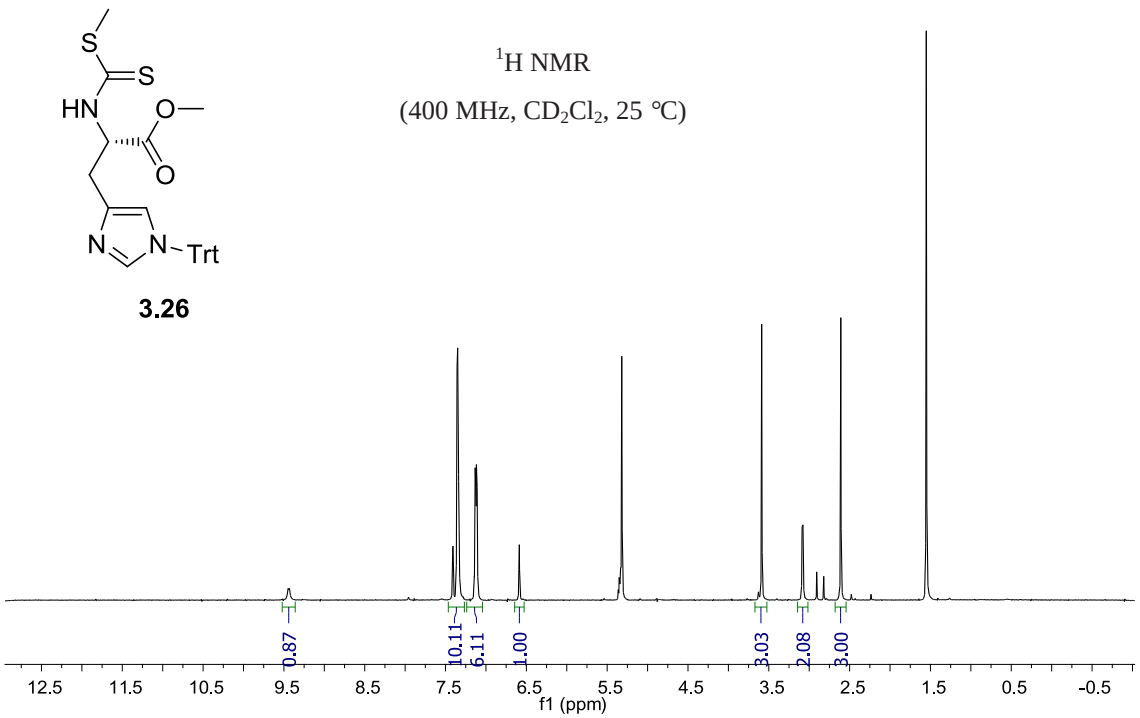
^{1}H NMR

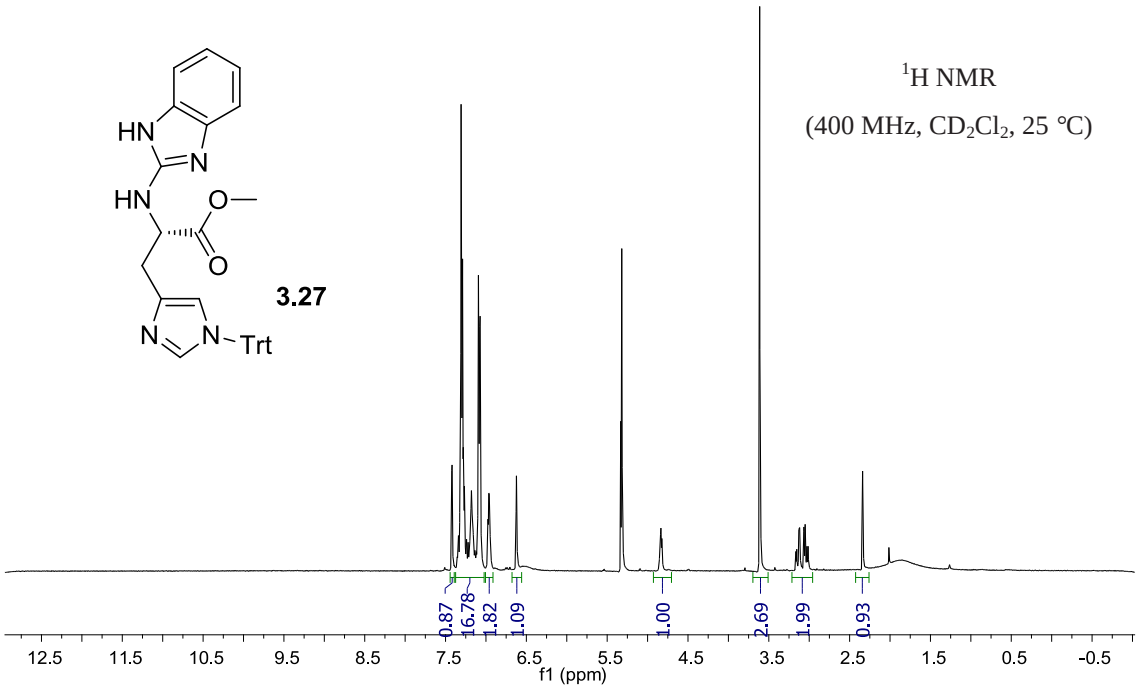
(300 MHz, CD_3OD, 25 ℃)

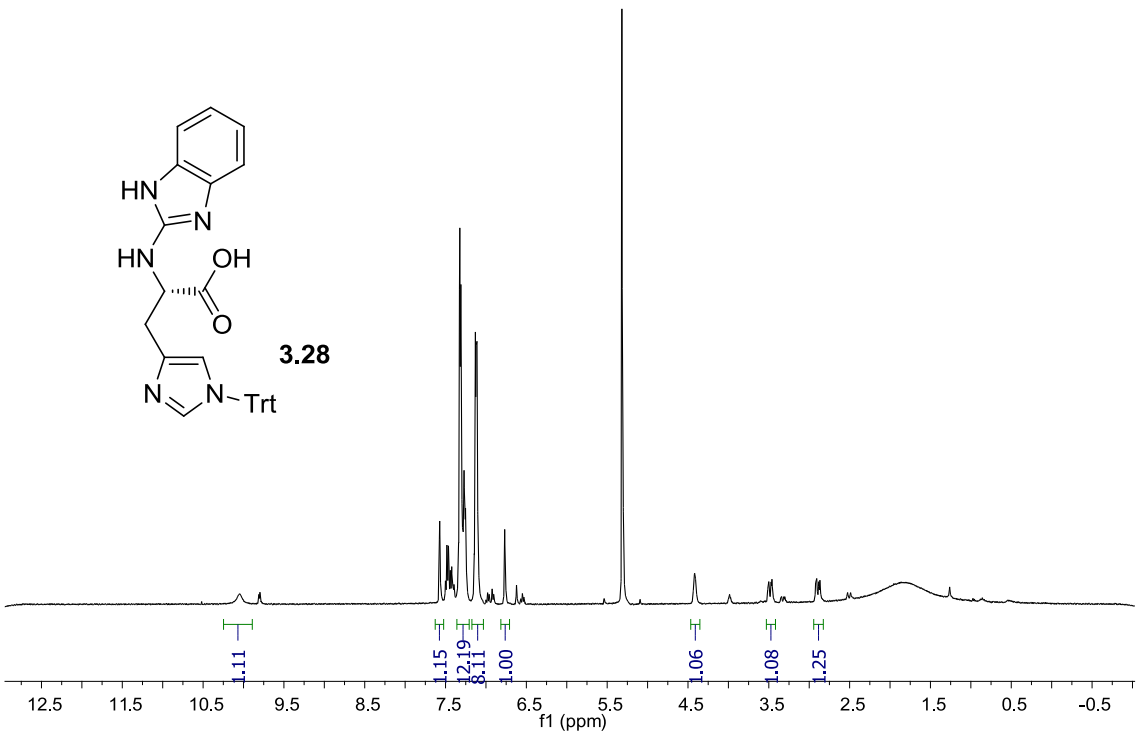


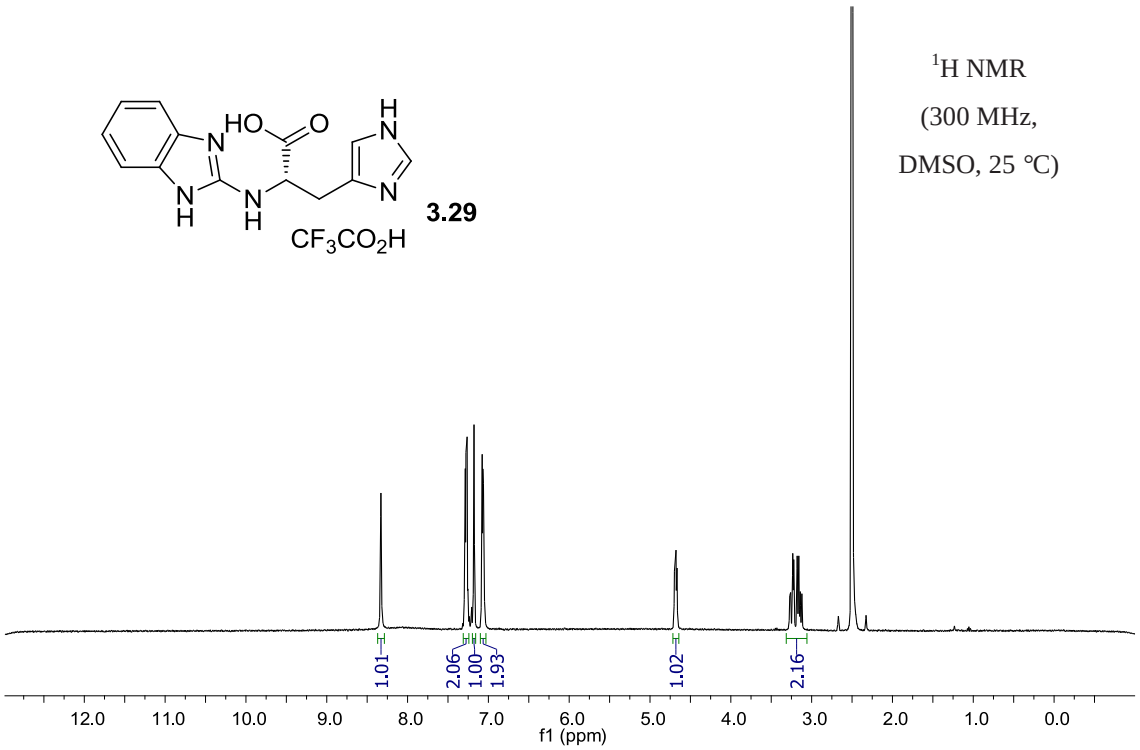
^{13}C NMR

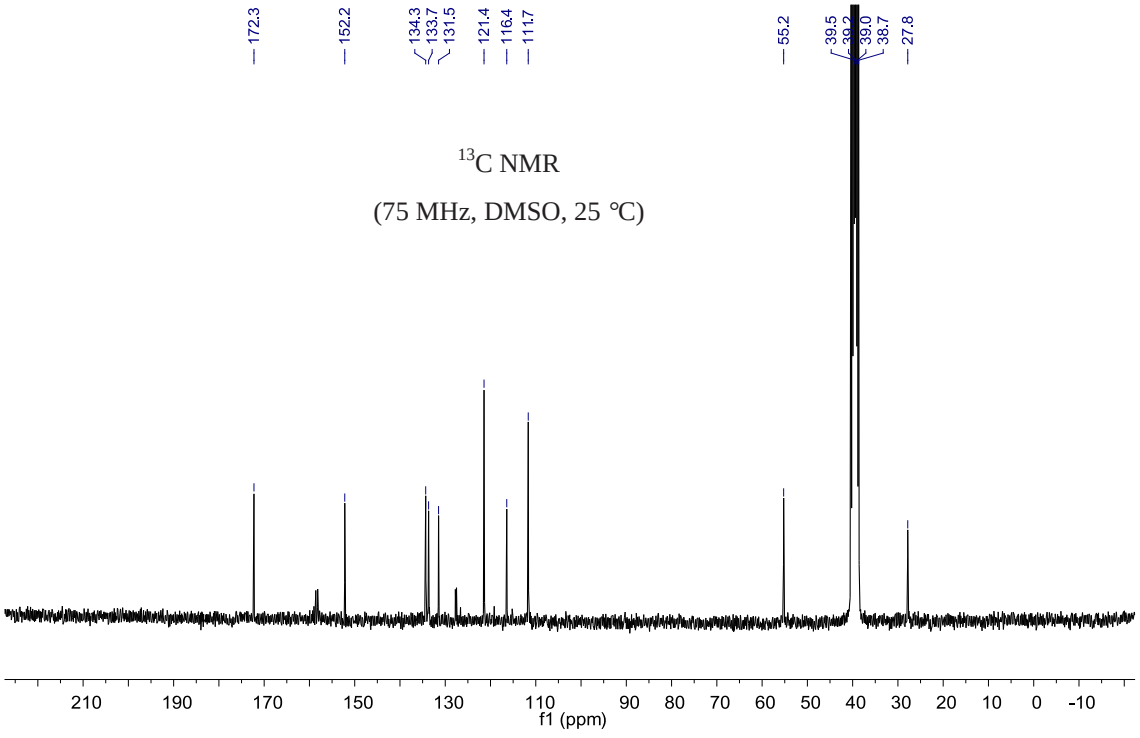
(75 MHz, $CDCl_3$, 25 ℃)

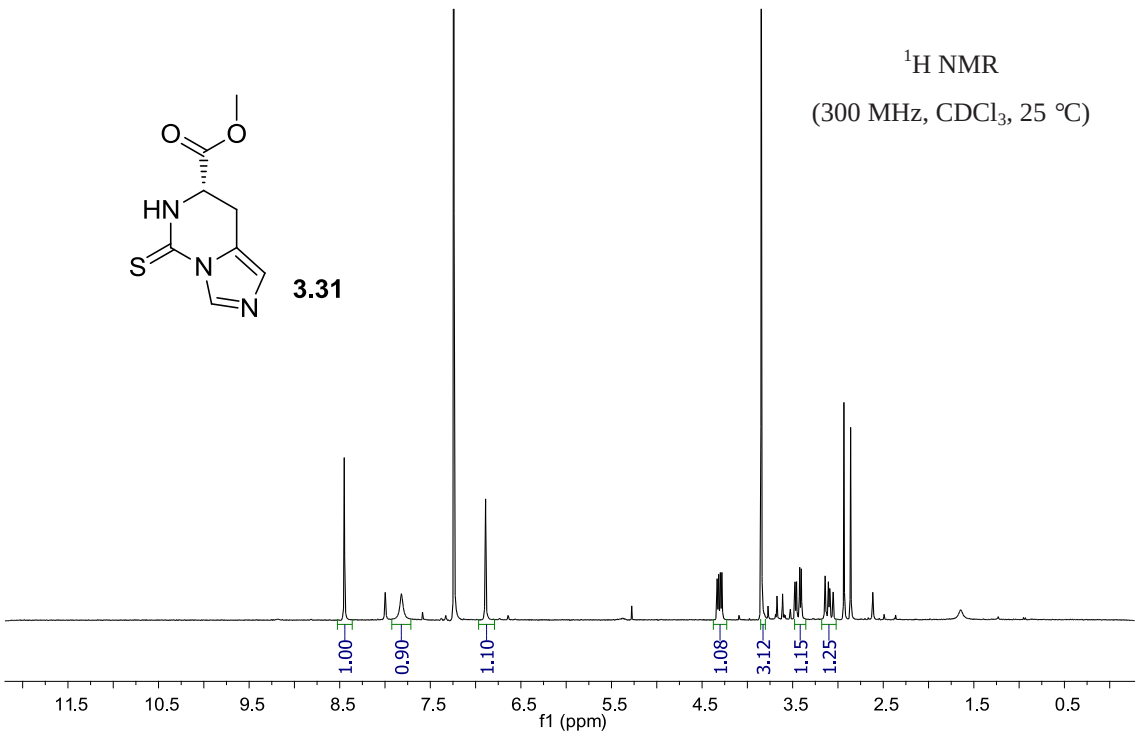

S
S
HN
O
O
N
N
Trt
3.26
^{1}H NMR
(400 MHz, CD_2Cl_2, 25 ℃)
0.87
10.11
6.11
1.00
3.03
2.08
3.00
12.5
11.5
10.5
9.5
8.5
7.5
6.5
5.5
4.5
3.5
2.5
1.5
0.5
-0.5
f1 (ppm)


HN
N
HN
O
O
N
N
Trt
3.27
^{1}H NMR
(400 MHz, CD_2Cl_2, 25 °C)
0.87
16.78
1.82
1.09
1.00
2.69
1.99
0.93
12.5
11.5
10.5
9.5
8.5
7.5
6.5
5.5
4.5
3.5
2.5
1.5
0.5
-0.5
f1 (ppm)

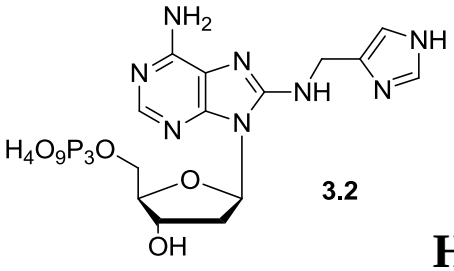

HN
N
HN
OH
O
N
N
Trt
3.28
1.11
1.15
2.19
3.11
1.00
1.06
1.08
1.25
12.5
11.5
10.5
9.5
8.5
7.5
6.5
5.5
4.5
3.5
2.5
1.5
0.5
-0.5
f1 (ppm)


HO
O
N
H
N
N
H
N
H
N
CF3CO2H
3.29
1H NMR
(300 MHz,
DMSO, 25 ℃)
1.01
2.06
1.00
1.93
1.02
2.16
12.0
11.0
10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
f1 (ppm)

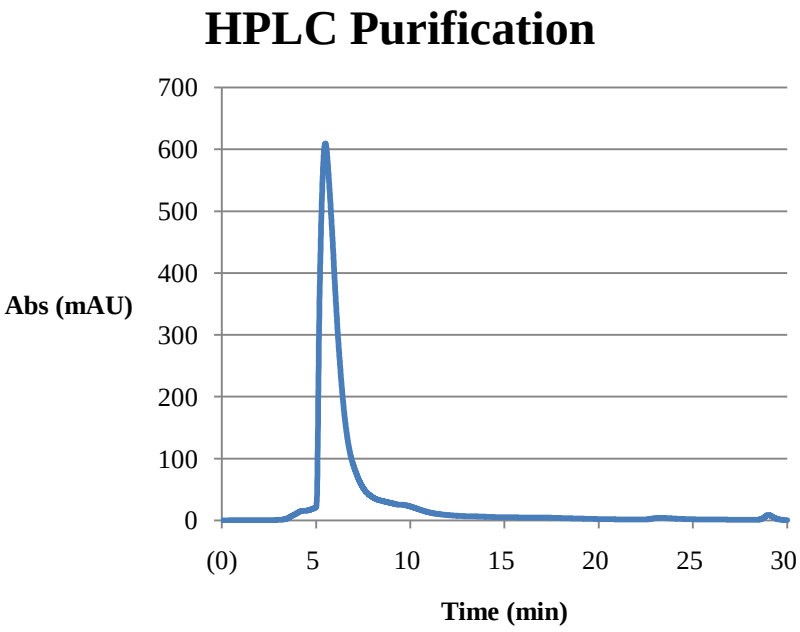

172.3
152.2
134.3
133.7
131.5
121.4
116.4
111.7
55.2
39.5
39.2
39.0
38.7
27.8
13C NMR
(75 MHz, DMSO, 25 ℃)
210
190
170
150
130
110
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)

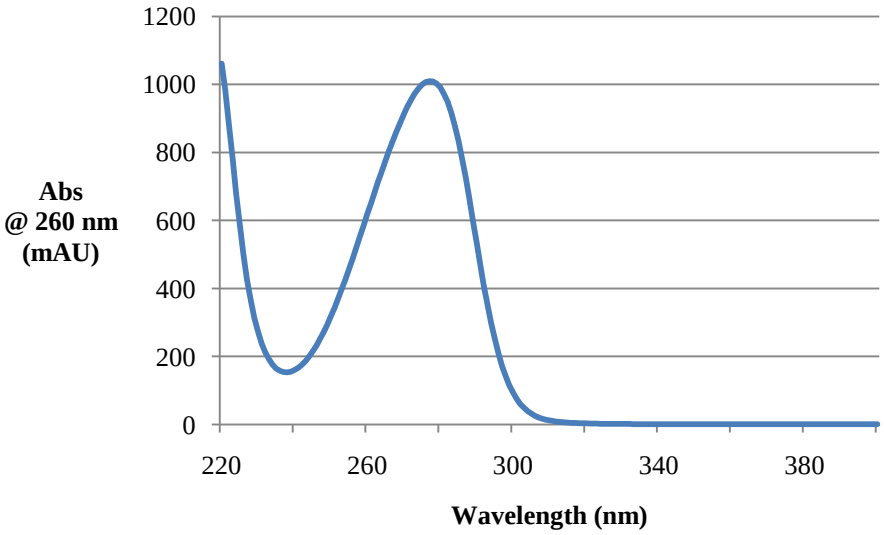

O
O
HN
S
N
N
3.31
1H NMR
(300 MHz, CDCl3, 25 ℃)
1.00
0.90
1.10
1.08
3.12
1.15
1.25
11.5
10.5
9.5
8.5
7.5
6.5
5.5
4.5
3.5
2.5
1.5
0.5
f1 (ppm)

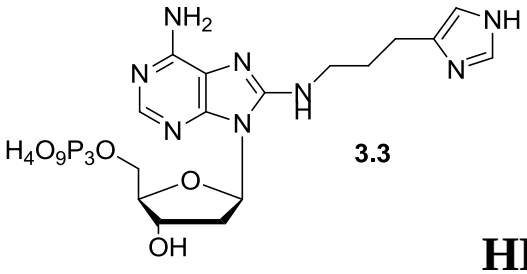
HPLC purification of modified 2′-deoxyadenosine triphosphates



3.2

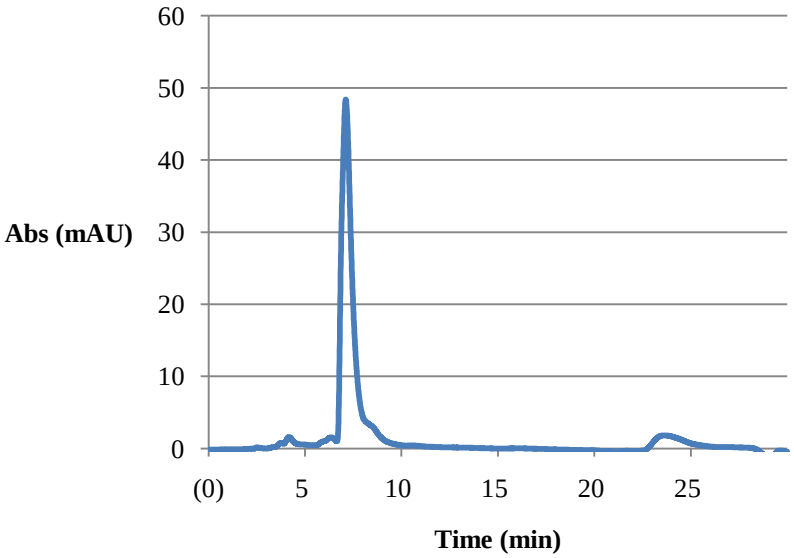


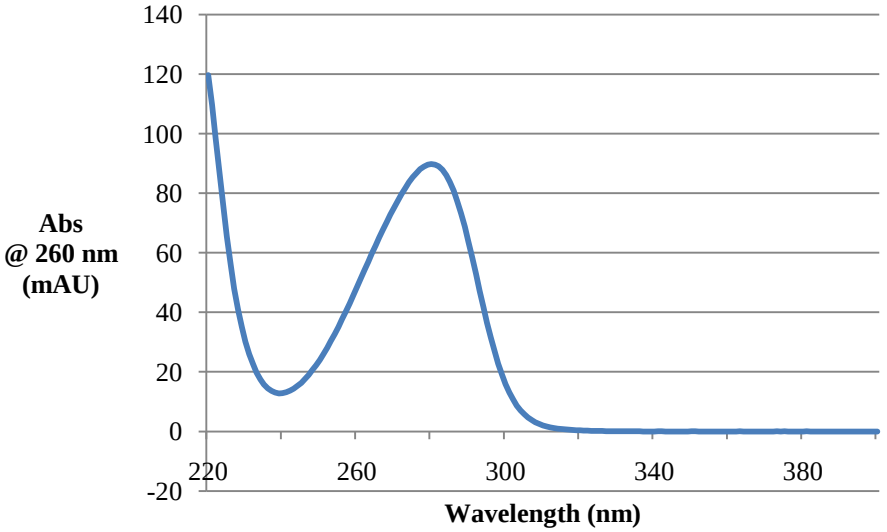


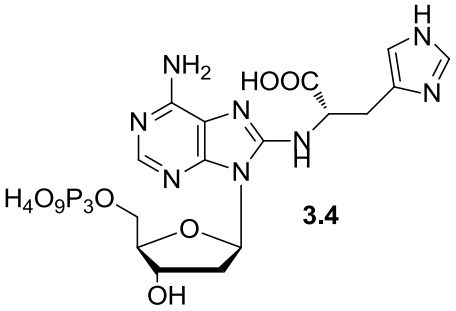


3.3

HPLC Purification

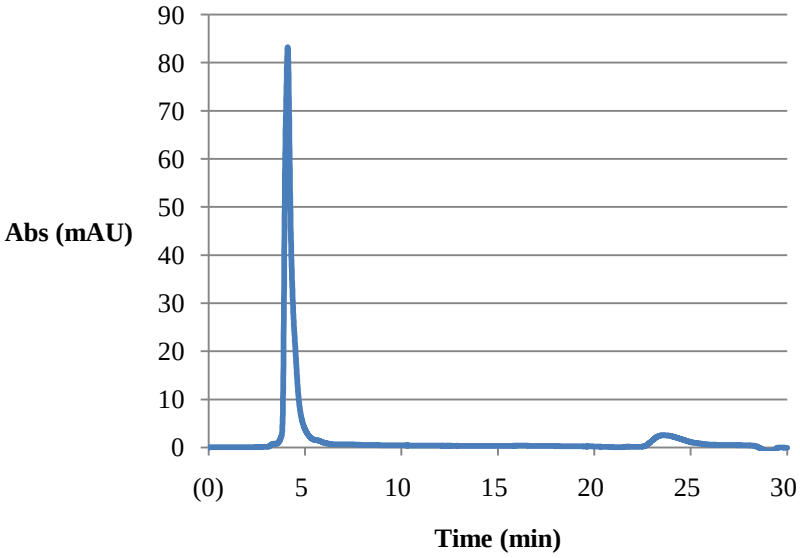


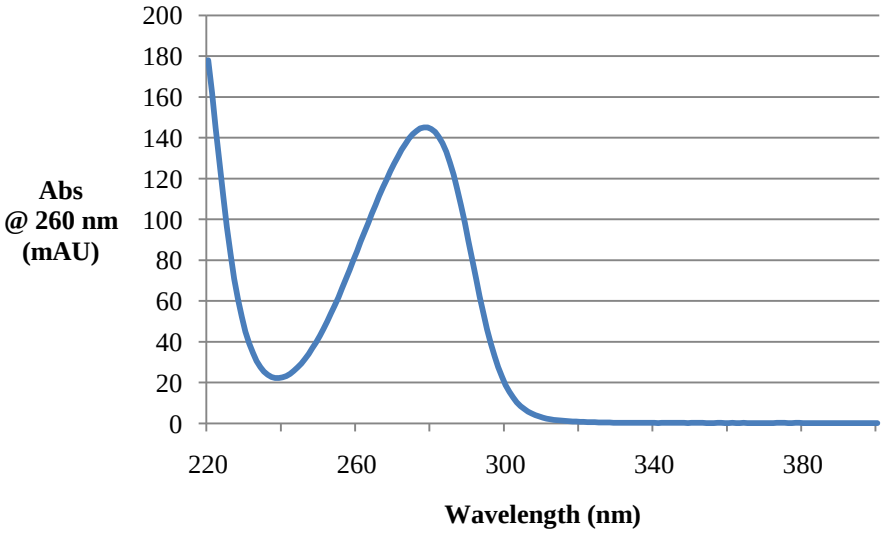


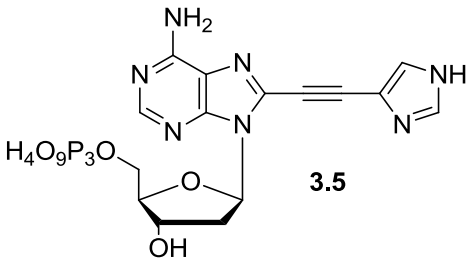


3.4

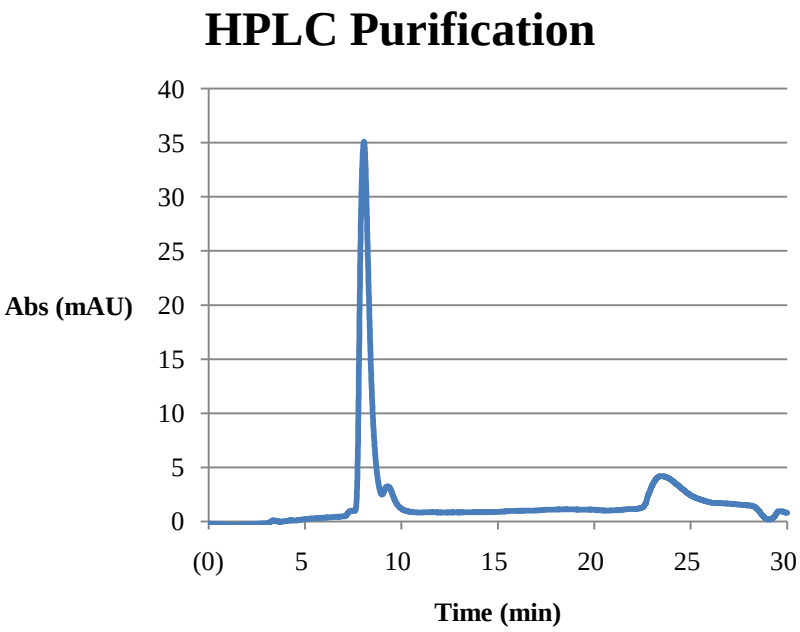
HPLC Purification

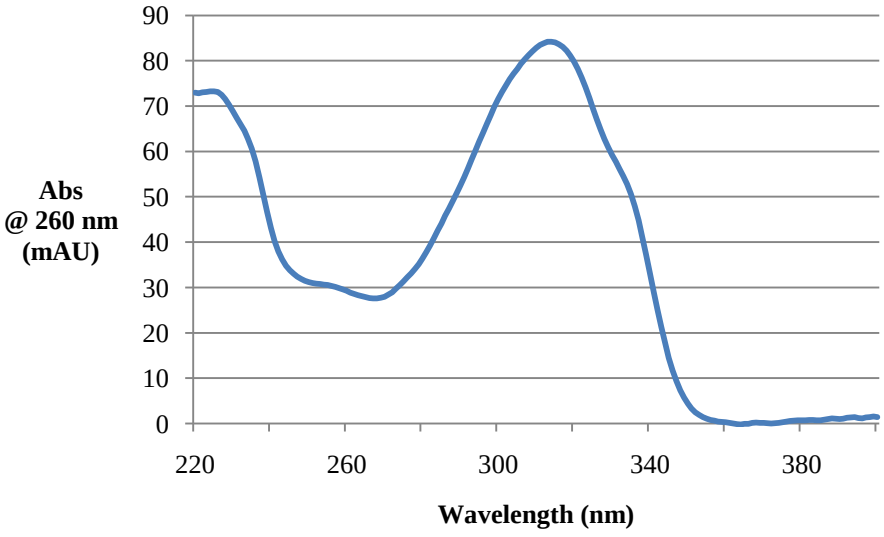


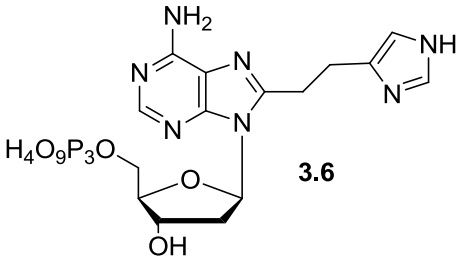



NH2
N
N
N
N
NH
N
H4O9P3O
O
OH

3.5

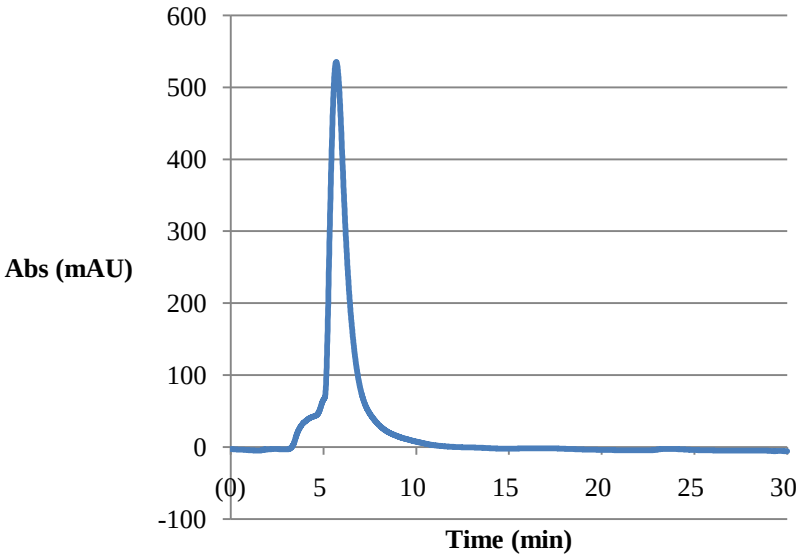

HPLC Purification
Abs (mAU)
40
35
30
25
20
15
10
5
0
(0)
5
10
15
20
25
30
Time (min)

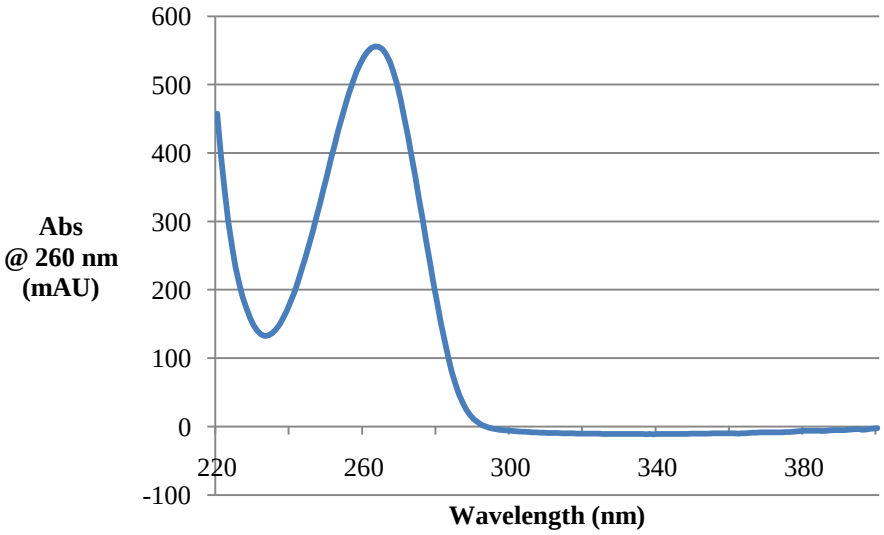

Abs
@ 260 nm
(mAU)
90
80
70
60
50
40
30
20
10
0
220
260
300
340
380
Wavelength (nm)



3.6

HPLC Purification





Chapter 4

Additional data for guanidinium-modified 10-23

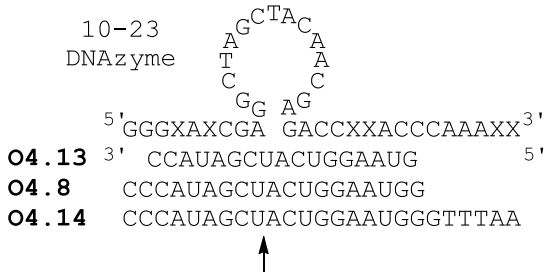


Figure A1. RNA substrates used for catalysis. X is either T or a guanidinium-modified residue. The location of cleavage is indicated by the arrow.

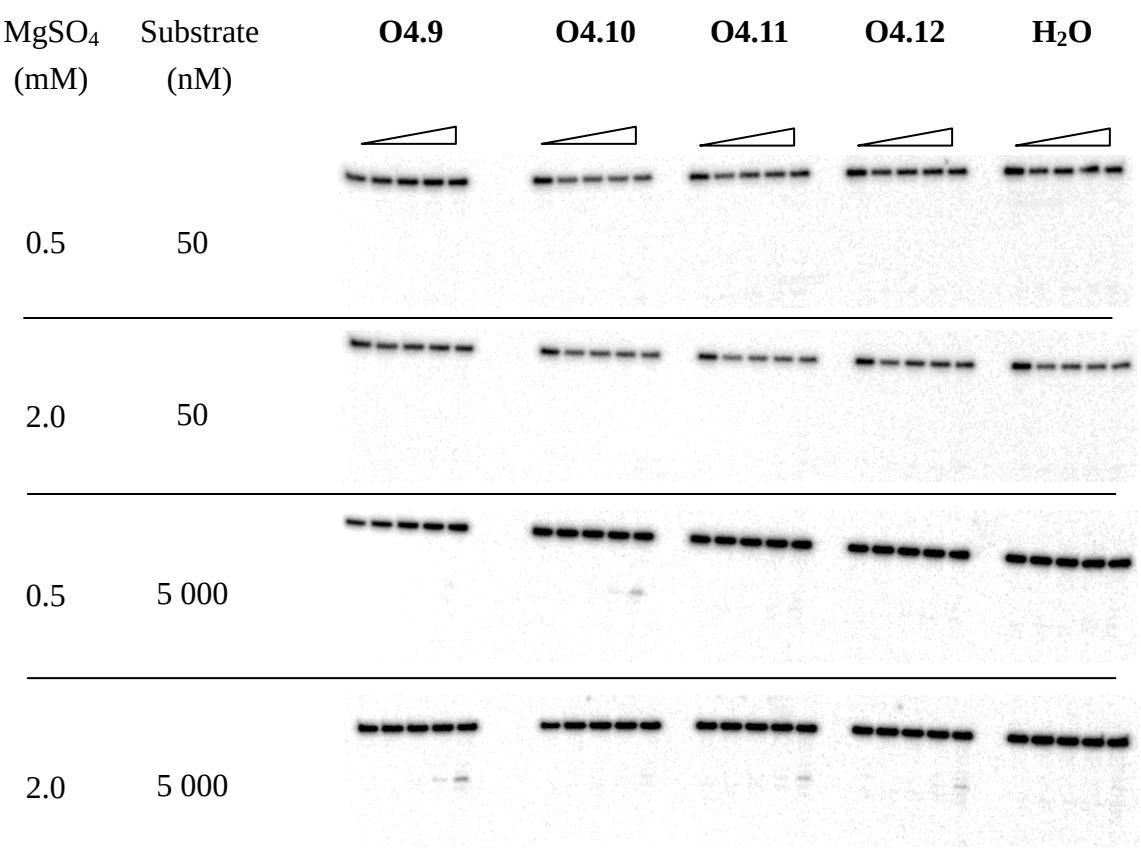


Figure A2. Denaturing PAGE (20 %) of 10-23 cleavage reactions using substrate **O4.13**. Four DNAzymes and a H_2O control were incubated with RNA substrate under different concentrations of $MgSO_4$. Final concentrations of the catalyst and substrate were either 5 and 50 nM or 0.25 μM and 5 μM, respectively. Kinetic reactions were performed in 150 mM NaCl, 50 mM Tris-HCl (pH 7.5), 37 ℃. Aliquots were removed and quenched at 0, 0.5, 1.5, 4:35, 29:25 h. The 2 sets of 5 lanes corresponding to **O4.10** catalysis at 5 000 nM substrate were in fact performed at 2.0 and 0.5 mM Mg^{+2} rather than 0.5 and 2.0 mM Mg^{+2}, respectively.

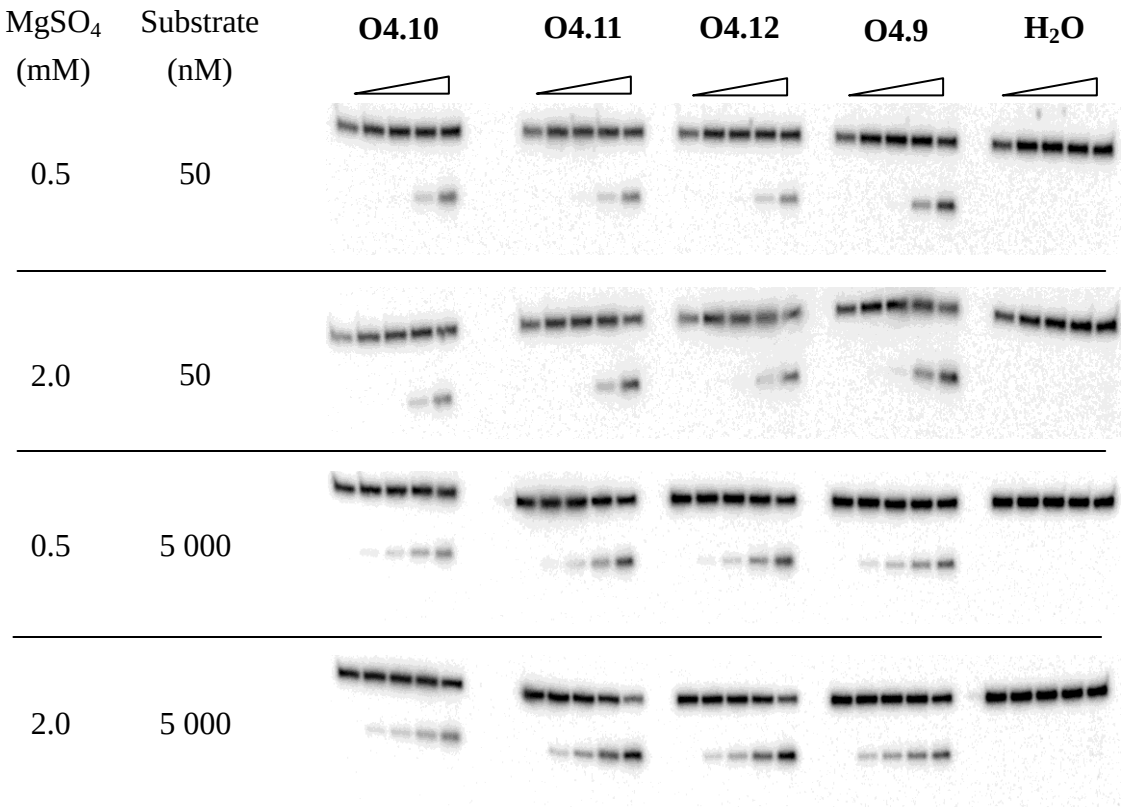


Figure A3. Denaturing PAGE (20 %) of 10-23 cleavage reactions using substrate **O4.14**. Four DNAzymes and a H_2O control were incubated with different concentrations of RNA substrate and $MgSO_4$. Final concentrations of the catalyst and substrate were either 5 and 50 nM or 0.25 μM and 5 μM, respectively. Kinetic reactions were performed in 150 mM NaCl, 50 mM Tris-HCl (pH 7.5), 37 ℃. Aliquots were removed and quenched at 0, 0.5, 1.5, 6, 21 h.

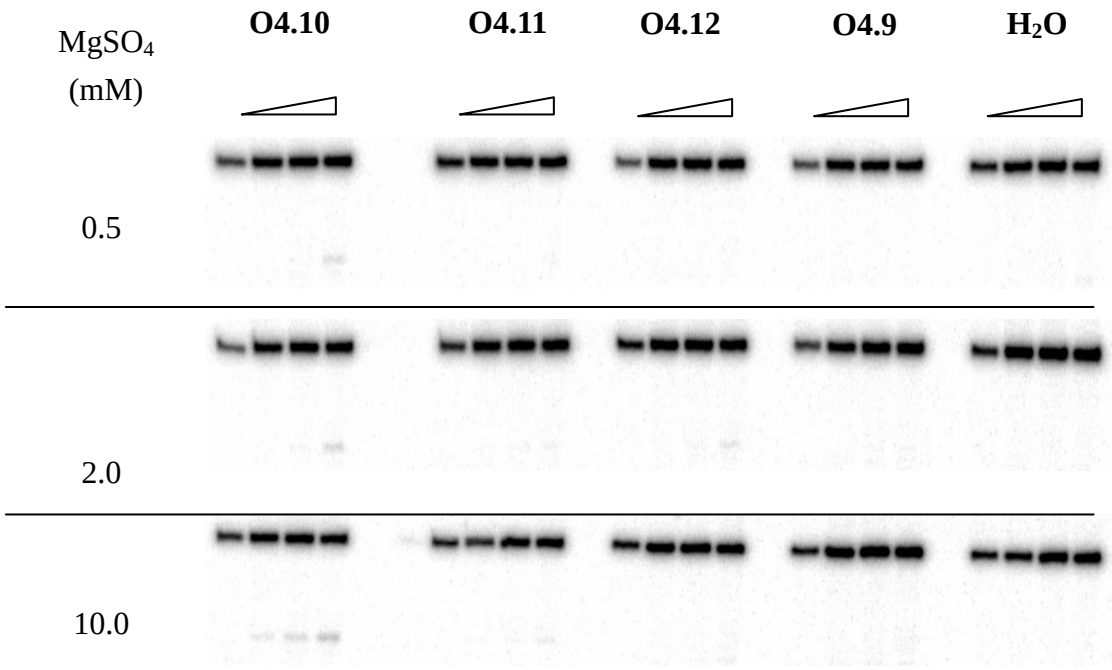
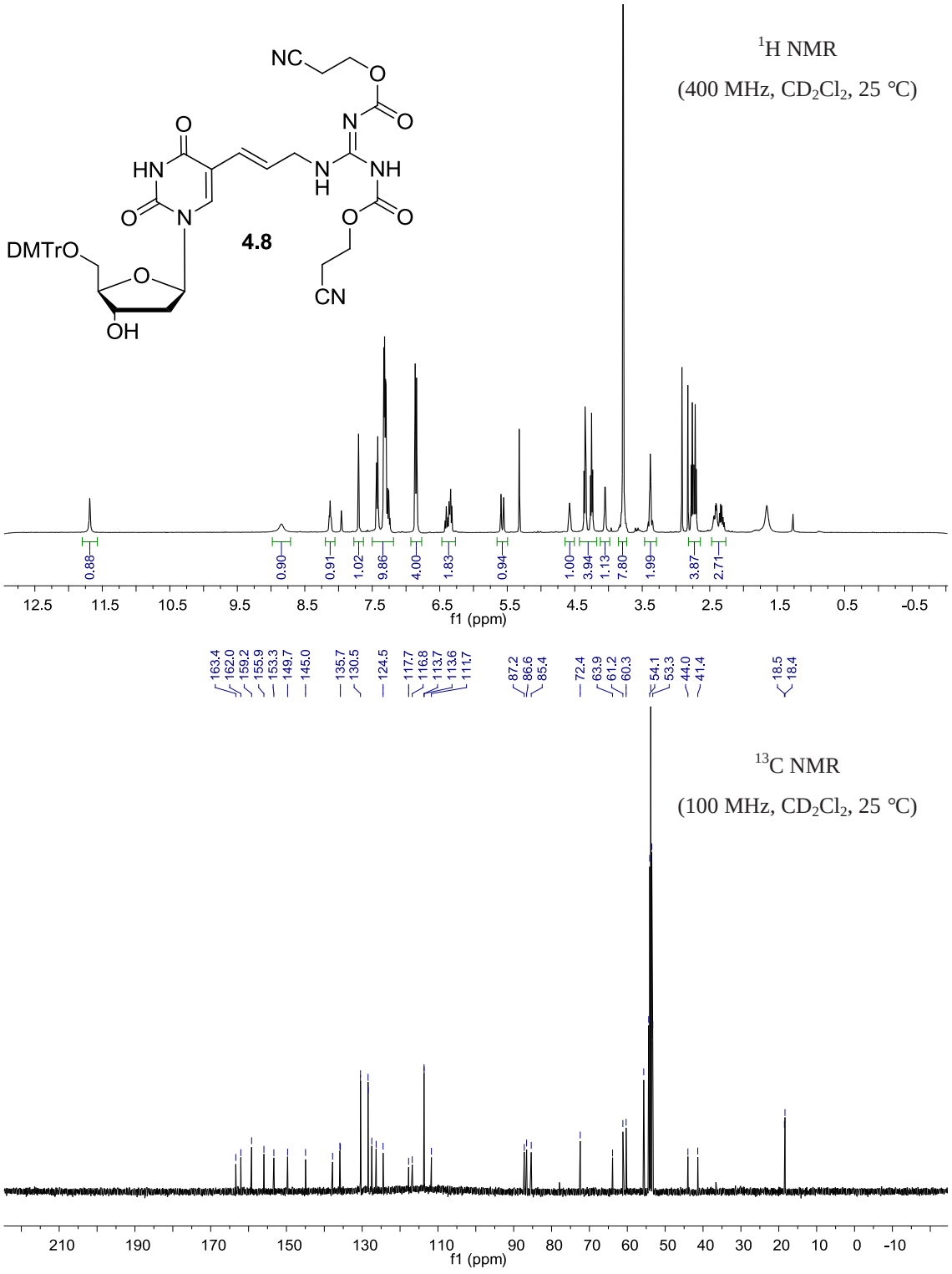
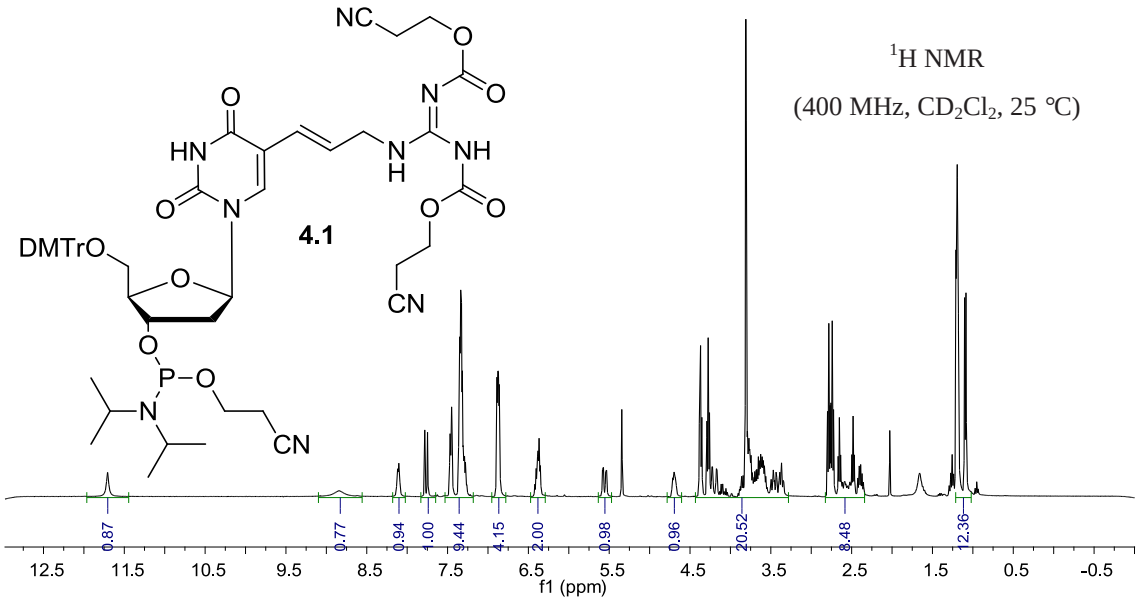
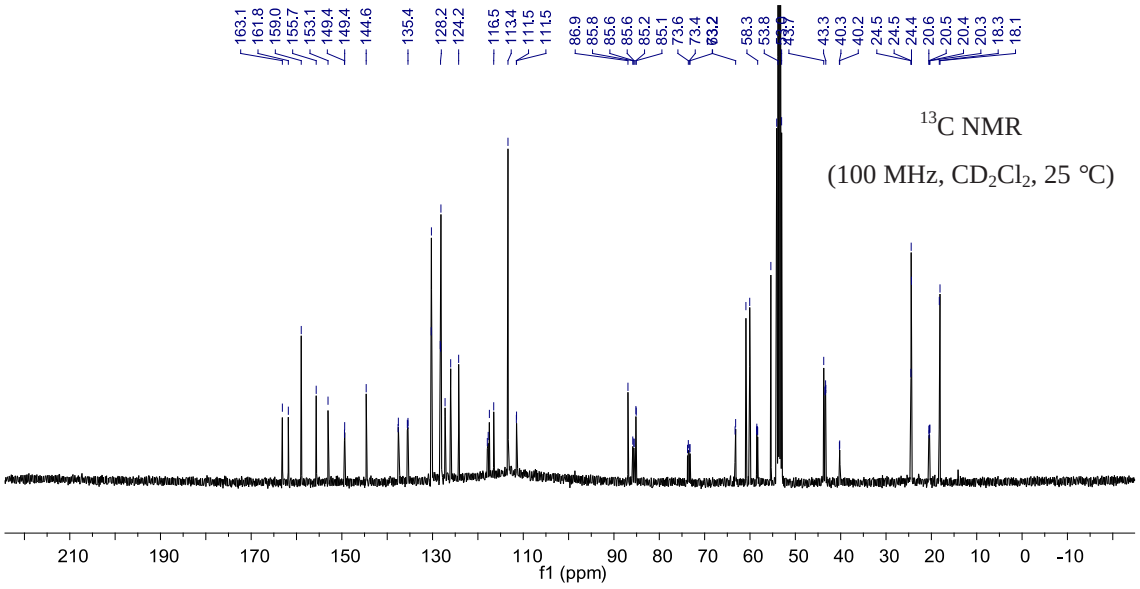


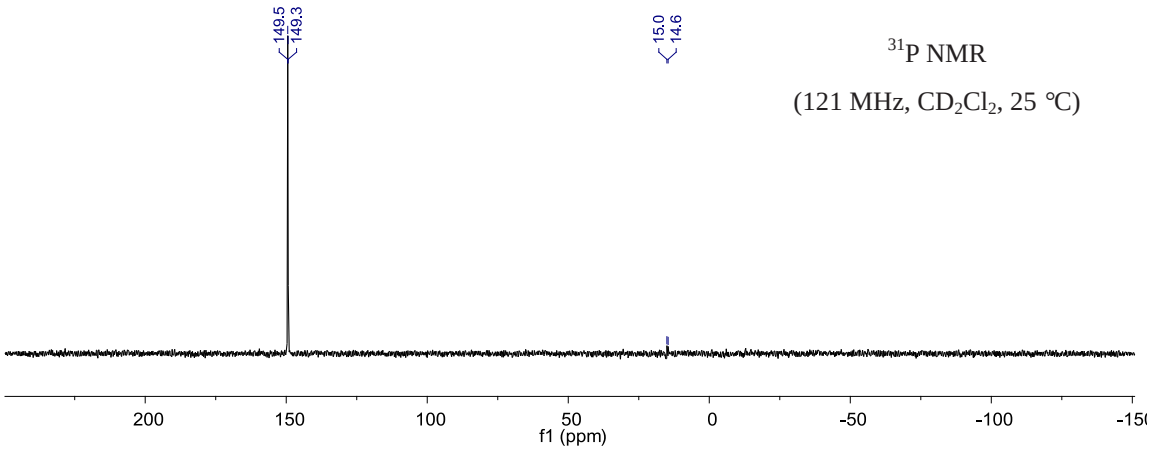
Figure A4. Denaturing PAGE (20 %) of 10-23 cleavage reactions using substrate **O4.8**. Four DNAzymes and a H_2O control were incubated with RNA substrate under different concentrations of $MgSO_4$. Final concentrations of the catalyst and substrate were 5 and 25 nM. Kinetic reactions were performed in 150 mM NaCl, 50 mM Tris-HCl (pH 7.5), 37 ℃. Aliquots were removed and quenched at 0, 20, 60, 180 min.

NMR spectra


NC
O
N
O
HN
N
H
NH
O
N
O
O
O
4.8
DMTrO
O
CN
OH
^{1}H NMR
(400 MHz, CD_2Cl_2, 25 °C)
0.88
0.90
0.91
1.02
9.86
4.00
1.83
0.94
1.00
3.94
1.13
7.80
1.99
3.87
2.71
12.5
11.5
10.5
9.5
8.5
7.5
6.5
5.5
4.5
3.5
2.5
1.5
0.5
-0.5
f1 (ppm)
163.4
162.0
159.2
155.9
153.3
149.7
145.0
135.7
130.5
124.5
117.7
116.8
113.7
113.6
111.7
87.2
86.6
85.4
72.4
63.9
61.2
60.3
54.1
53.3
44.0
41.4
18.5
18.4
^{13}C NMR
(100 MHz, CD_2Cl_2, 25 °C)
210
190
170
150
130
110
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)


NC
O
N
O
HN
O
N
H
NH
O
N
O
O
O
4.1
DMTrO
O
O
CN
O
P
O
N
CN
0.87
0.77
0.94
1.00
9.44
4.15
2.00
0.98
0.96
20.52
8.48
12.36
12.5
11.5
10.5
9.5
8.5
7.5
6.5
5.5
4.5
3.5
2.5
1.5
0.5
-0.5
f1 (ppm)
^{1}H NMR
(400 MHz, CD_2Cl_2, 25 ℃)


163.1
161.8
159.0
155.7
153.1
149.4
149.4
144.6
135.4
128.2
124.2
116.5
113.4
111.5
111.5
86.9
85.8
85.6
85.6
85.2
85.1
73.6
73.4
63.2
58.3
53.8
43.7
43.3
40.3
40.2
24.5
24.5
24.4
20.6
20.5
20.4
20.3
18.3
18.1
210
190
170
150
130
110
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)
^{13}C NMR
(100 MHz, CD_2Cl_2, 25 ℃)


149.5
149.3
15.0
14.6
200
150
100
50
0
-50
-100
f1 (ppm)
^{31}P NMR
(121 MHz, CD_2Cl_2, 25 ℃)

Chapter 6

Conversion of DNAzyme STA17 cleaved product into complementary unmodified product

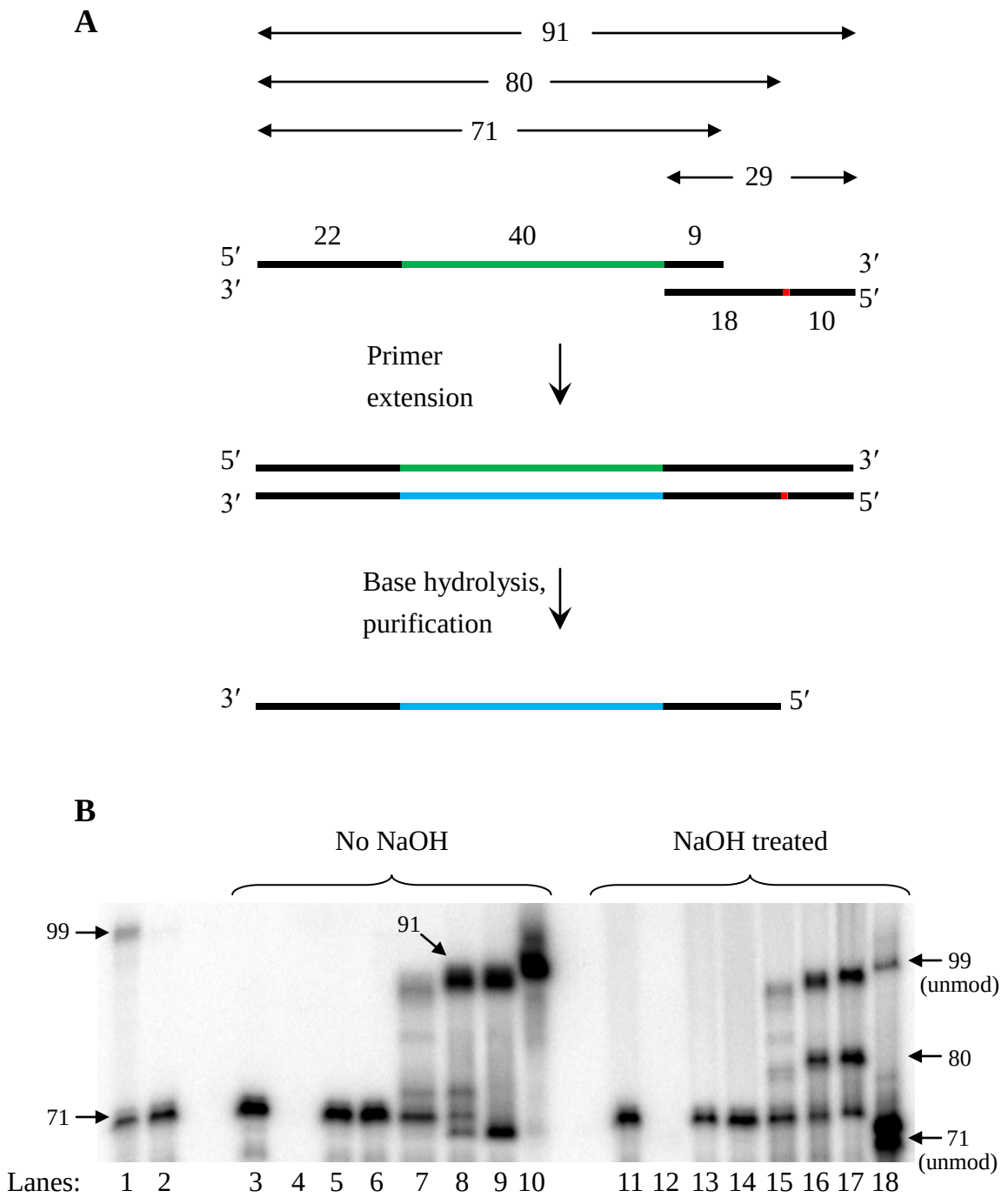


Figure A5 Conversion of STA17 cleaved product into the complementary unmodified

product. STA17 product [71 nucleotides (nt)] is used as a template for a primer extension with Vent (exo–) to give one modified strand and a complementary unmodified strand, both 91 nt. The two strands are treated with NaOH to produce an unmodified strand of decreased length (80 nt) A: schematic of the experiment and the size of all the oligonucleotides used and produced. (——) DNA primer regions; (——) RNA; (——) modified 40 nucleotide region; (——) unmodified 40 nucleotide region. B: denaturing PAGE (7 %) of the experiment. Numerical labels correspond to the number of nucleotides for each of the gel products. The biotinylated primer that was used for the original selection of STA17 was 28 nt in length. Lanes 1 and 2, STA17 showing uncleaved and cleaved DNAzyme as size markers; lanes 3 and 11, no polymerase; lanes 4 and 12, no template; lanes 5 and 13, no primer; lanes 6 and 14, 0 min extension time; lanes 7 and 15, 0.5 min extension time; lanes 8 and 16, 5 min extension time; lanes 9 and 17, 60 min extension time; lanes 10 and 18, uncleaved and cleaved unmodified STA17 as size markers. Lanes 1, 3–10 were not treated with NaOH; lanes 2, 11–18 were treated with NaOH.

Experimental. For general materials and methods, please refer to Chapter 5.

DNAzyme STA17 (30 pmol) was prepared as described in Chapter 5. The DNAzyme was incubated for 3.5 h in the presence of 100 μL selection buffer (50 mM HEPES pH 7.4, 200 mM NaCl, 1 mM $ZnSO_4$ and 10 mM $MgCl_2$). The liquid was desalted with a G25 spin column and evaporated to dryness. This material was resuspended in water (22 μL). Unmodified STA17 (15 pmol) was prepared in the same manner as the modified STA17 for use as a size control.

Annealing mixture: water (10 μL), template (cleaved STA17, 20 μL), primer (5′-GATCTAGCGArATGAGCTCGCGGGGCGTGC-3′, 2.5 μL, 10 μM) and thermopol (5 μL, 10X). Nucleotide solution: dCTP (1 μL, 1 mM), dGTP (1 μL, 1 mM), dATP (2 μL, 1 mM), dTTP (2 μL, 1 mM), DTT (0.4 μL, 100 mM) and α-^{32}P-dGTP (2 μL).

	Annealing mixture	Water	Template	Primer	Thermopol
No enzyme	3.75				
No template		3.0		0.25	0.5
No primer		1.25	2.0		0.5
t = 0 min	3.75				
t = 0.5 min	3.75				
t = 5 min	3.75				
t = 60 min	3.75				

Table A1 Primer extension sample preparation.

The 7 solutions shown in the above table were each added nucleotide solution (1.05 μL), Vent (exo–) (0.2 μL) and mineral oil (15 μL). All of the samples were incubated at 72 ℃. The 3 control reactions were incubated for 60 min, and the other 4 reactions were incubated for 0, 0.5, 5, and 60 min, respectively. Each reaction was quenched with $EDTA_{(aq)}$ (0.25 μL, 0.5 M). Water (35 μL) was added to each reaction. Each reaction was precipitated with 1 % $LiClO_4$ in acetone (400 μL) and rinsed with ethanol (200 μL). Each sample was then resuspended in water (5 μL) and divided into two portions. To one portion of each pair was added NaOH (0.25 μL, 10 M), and resulting mixtures were heated at 95 °C for 5 min. NaOH treated samples were precipitated with ethanol (200 μL), and the non-NaOH treated samples were evaporated. To each of the 14 samples were added loading solution (5 μL). Modified and unmodified DNAzyme controls (lanes 1, 2, 10, 18 of Figure A5) were prepared in the same manner as Chapter 5. All 18 samples were heated at 95 ℃ for 5 min and rapidly cooled prior to loading onto denaturing PAGE (7 %).